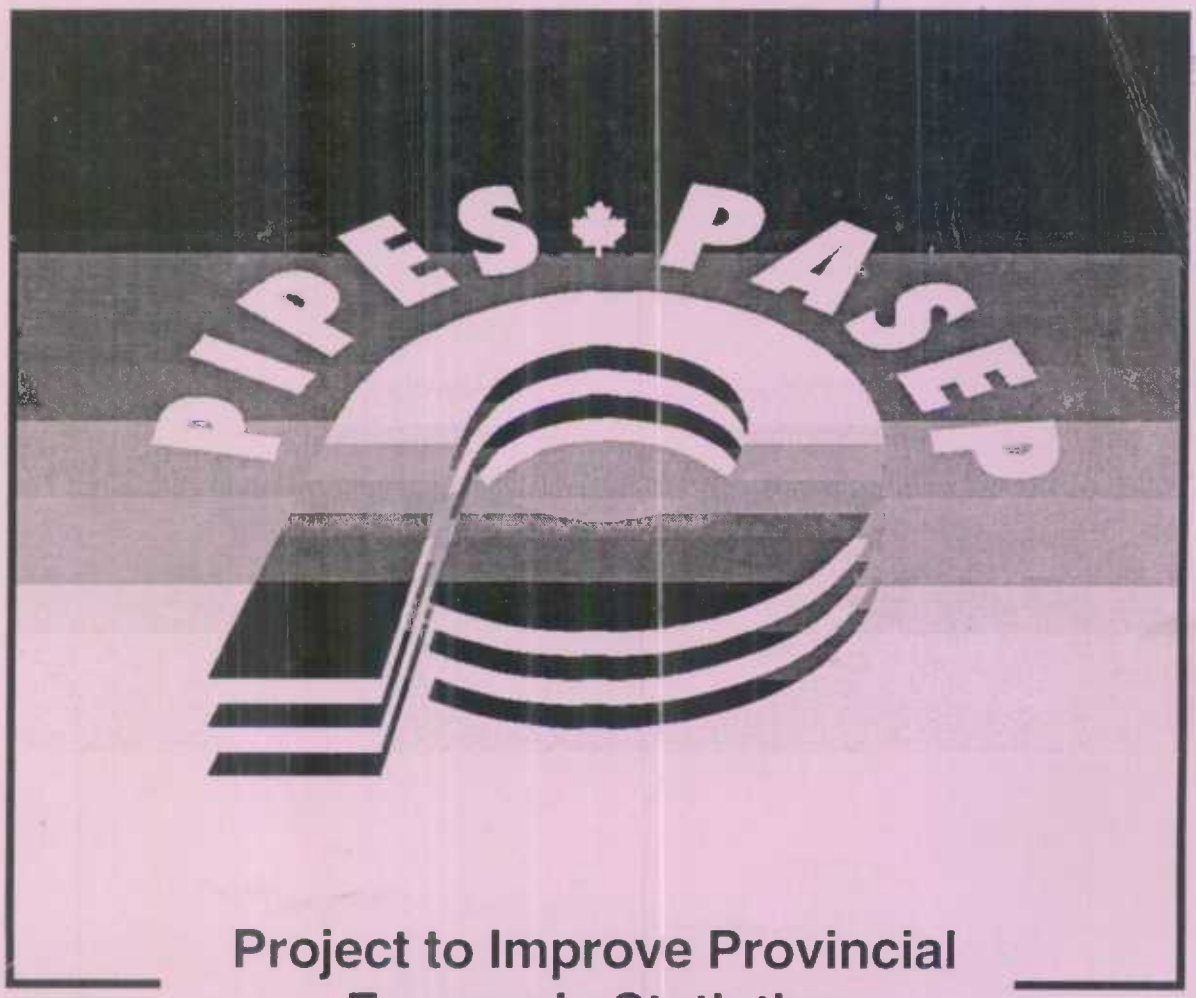
C.3

Catalogue no. 68N0003XPE

Unified Enterprise Statistics Program

STATISTICS CANADA STATISTIQUE CANADA
MAY MAI 26 1997



Project to Improve Provincial Economic Statistics

Statistics Canada Statistique Canada

Canada

For more information, please contact:

George Beelen, Director
Enterprise Statistics Division
Statistics Canada
Jean Talon Building
13th Floor, Section C-8
Ottawa, Ontario K1A 0T6

Telephone: (613) 951-6559
Fax: (613) 951-6196

This document is an internal working paper of Statistics Canada. It is written in somewhat technical language and is not intended for a general audience.

May 5, 1997

Table of Contents

EXECUTIVE SUMMARY 1

PIPES OBJECTIVES 1
A MORE INTEGRATED BUSINESS STATISTICS PROGRAM 2
OVERVIEW OF THE BASIC PLAN 3
IMPLEMENTATION TIMING 4

INTRODUCTION 7

STATISTICS FOR THE HST REVENUE ALLOCATION FORMULA 7
AN IMPLEMENTATION STRATEGY BASED ON TIGHTER INTEGRATION 7
PURPOSE OF THIS PAPER 9
STRUCTURE OF THIS PAPER 9

CHAPTER 1

PIPES REQUIREMENTS AND THEIR IMPLICATIONS FOR ECONOMIC STATISTICS 11

THE HARMONIZED SALES TAX AGREEMENT 11
SPECIFICATIONS OF THE HST/GST 11
FINAL DEMAND FOR GOODS AND SERVICES 15
INVESTMENT IN HOUSING 17
PRODUCTION/SUPPLY AND INTERMEDIATE USE OF COMMODITIES 18
PRODUCERS OF EXEMPT COMMODITIES IN THE BUSINESS SECTOR 19
PUBLIC SECTOR BODIES 20
GENERAL REQUIREMENTS OF THE NEW SPA 20

CHAPTER 2

OVERVIEW OF THE UNIFIED ENTERPRISE STATISTICS PROGRAM 21

FUNDAMENTAL PRINCIPLES OF THE UNIFIED ENTERPRISE STATISTICS PROGRAM 21
MAJOR ELEMENTS OF THE UNIFIED ENTERPRISE STATISTICS PROGRAM 22

CHAPTER 3

A CENSUS OF LARGE AND COMPLEX ENTERPRISES 27

INTRODUCTION 27
MAJOR ELEMENTS OF THE PROPOSED CCE 27
COMPLEX ENTERPRISES 28
CONTENT 30
RESPONSE BURDEN 31
PROFILING 32
PROCESSING 33

CHAPTER 4

A SURVEY OF ESTABLISHMENTS OF COMPLEX ENTERPRISES 35

INTRODUCTION 35
MAJOR ELEMENTS OF THE SURVEY OF ESTABLISHMENTS OF COMPLEX ENTERPRISES 35
COMPLEX ENTERPRISES 36
GOALS 38
RELATIONSHIP TO SURVEY OF SIMPLE ENTERPRISES 38
RELATIONSHIP TO SUB-ANNUAL SURVEYS 38
TIMELINESS 38
SAMPLING STRATEGY 39
COLLECTION STRATEGY 39
RESPONDENT RELATIONS 40
PROCESSING 40
ANALYSIS 41
OUTPUT 41
IMPLEMENTATION STRATEGY 42

CHAPTER 5

A SURVEY OF SIMPLE ENTERPRISES 43

SURVEY STRATEGY: AN OVERVIEW 43
SIMPLE ENTERPRISES 44
RESPONSE BURDEN AND THE SIMPLE ENTERPRISE 47

CHAPTER 6

SAMPLE DESIGN AND STRATIFICATION ISSUES 53

INTRODUCTION 53
OVERVIEW OF THE INTEGRATED SAMPLING AND ESTIMATION APPROACH 53
ADVANTAGES OF AN INTEGRATED APPROACH 54
QUALITY OF THE BUSINESS REGISTER 56
SAMPLING OF COMPLEX ENTERPRISES 56
SAMPLING OF SIMPLE ENTERPRISES 57
USE OF ADMINISTRATIVE DATA IN ESTIMATION 59
DATA QUALITY 60
QUALITY ASSURANCE 60
DATA QUALITY MEASUREMENT STUDIES 61

CHAPTER 7

CHARACTERISTICS OF THE IMPROVED BUSINESS REGISTER 63

THE REQUIREMENT FOR A SINGLE BUSINESS REGISTER 63
POPULATION COVERAGE 63
BUSINESS REGISTER BOUNDARIES 64
PRIMARY ATTRIBUTES OF THE BUSINESS REGISTER 64
IMPORTANCE OF TIMELY UPDATING FRAME INFORMATION (SUCH AS GBI) 65
RECONCILIATION OF THE BR WITH THE IOFD AND AGRICULTURE DIVISION FRAMES 65
LEVELS OF STATISTICAL UNITS 66
LINKING OPERATIONAL ENTITIES TO LEGAL ENTITIES 66

INTERFACES WITH SURVEYS 67
TIME TRAVEL 67
SYSTEMS 67
TIMETABLE 68
SUMMARY PROFILING INFORMATION FROM THE QUARTERLY SURVEY OF ENTERPRISES 68
CONCERNS 68

CHAPTER 8

THE INTEGRATED QUESTIONNAIRE 71

INTEGRATED QUESTIONNAIRE - THE CONCEPTUAL MODEL 71
DEVELOPMENT PLAN – OUTLINE OF ACTIVITIES 73
SPECIAL SITUATIONS 78

CHAPTER 9

CLASSIFICATION SYSTEMS 81

THE NORTH AMERICAN INDUSTRY CLASSIFICATION SYSTEM (NAICS) 81
COMMODITY CLASSIFICATION 84
OTHER CLASSIFICATIONS 88

CHAPTER 10

THE USE OF ADMINISTRATIVE DATA 89

INTRODUCTION 89
BACKGROUND 89
WHAT HAS CHANGED? 90
ASSUMPTIONS 92
PROPOSED USES OF TAX DATA 93
RISKS ASSOCIATED WITH EACH TAX SOURCE 95

CHAPTER 11

COMPUTER SYSTEMS AND DATABASE DESIGN REQUIREMENTS 97

INTRODUCTION 97
GENERAL PRINCIPLES SHAPING THE UES SYSTEMS ENVIRONMENT 99
INFRASTRUCTURE DATABASES 99
SOFTWARE INTEGRATION TOOLS 102
NEXT STEPS – THE SHORT-TERM STRATEGY 104

CHAPTER 12

IMPLEMENTATION ISSUES AND TIMING 105

THE PLANNING HORIZON IN PERSPECTIVE 105
SOME KEY ISSUES 107

CHAPTER 13

CONCLUSION 109

ORGANIZATIONAL IMPACT 109
ADVANTAGES OF THE ORGANIZATION CHANGES 110
LONG TERM IMPACT OF PIPES AND UESP 110

ANNEX 1

GLOSSARY OF TERMS 113

ANNEX 2

DATA REQUIREMENTS AND PRIORITIES FOR PIPES FROM THE SYSTEM OF NATIONAL ACCOUNTS 119

BACKGROUND 119
CRITERIA FOR DATA REQUIREMENTS 119
DATA COLLECTION PROBLEMS 121
SPECIFICATION OF REQUIREMENTS AND PRIORITIES 122

ANNEX 3

ADDITIONAL INFORMATION ABOUT ADMINISTRATIVE DATA SOURCES 133

BUSINESS TAX RETURNS WITH GENERAL INDEX OF FINANCIAL INFORMATION (GIFI) 133
GOODS AND SERVICES TAX (GST) FILES 135
PAYROLL DEDUCTION 136
CUSTOMS DOCUMENTS - IMPORT AND EXPORT SHIPMENTS 138

ANNEX 4

BUSINESS SURVEY AT STATISTICS CANADA 141

Executive Summary

PIPES objectives

Statistics Canada is embarking on a new project, the objective of which is to substantially upgrade the reliability and detail of the provincial economic statistics it produces.

Improvements to provincial economic statistics are needed in view of the federal-provincial agreement to harmonize the Goods and Services Tax with provincial sales taxes in Newfoundland, New Brunswick and Nova Scotia, announced on October 23, 1996. The single harmonized sales tax took effect on April 1, 1997. Statistics Canada will be the primary source of the data to be used in the formula for allocating the pooled revenues among the participating governments. More reliable annual provincial data than currently available are required for this purpose, and Statistics Canada must begin immediately to implement the proposed improvements so the new data can begin coming on stream in 1997 and following years, when they are needed.

The primary objective of PIPES is to measure final sales of goods and services accurately, on an annual basis by province, in sufficient detail and in a timely manner. The required information will be produced by a combination of much improved business surveys and an enhanced family expenditure survey plus information that can be derived from the judicious use of administrative records. The plan envisages the simultaneous assessment of each separate piece of information against all available data in the framework of a very detailed set of accounting entries — the provincial input-output tables. This will result in stable and accurate estimation of the value of final sales of goods and services, indeed in the best such estimation possible with existing methodologies.

PIPES will affect a wide range of statistical programs throughout the Bureau. Business surveys, in particular, will be very substantially affected. This paper presents a plan for implementing the business surveys part of the project.

One broad goal is to enhance the business statistics program so it can produce annual provincial economic statistics that are more reliable for all provinces and territories, and more equally reliable across provinces and territories, compared to the present situation. Another objective is to increase the amount of commodity and industry detail available in the statistics. Both goals suggest there will have to be a large increase in the survey response burden on business. This is indeed the case, although in the three Atlantic provinces that are harmonizing their sales taxes with the GST there will be an associated reduction in sales tax compliance cost that will more than offset the rise in survey response burden. Regardless, the implied increase in survey response burden associated with this undertaking is a major concern that will be kept foremost in mind as the project is implemented. This concern is one of the factors that led to the development of the following plan for a major restructuring of Statistics Canada's business data collection program.

Improvements in data quality are targeted in four areas: consistency, coherence, breadth, and depth. Consistency refers to constancy in the use of concepts, standards, statistical units, methodology, and terminology. Coherence refers to the internal inter-relatedness of data within an enterprise, currently obtained from many different surveys in most cases.

Breadth relates to the comprehensiveness of coverage in terms of filling gaps in coverage of industrial sectors and ensuring comprehensive, non-duplicative coverage of all business activity of complex enterprises. Finally, depth refers to the size of samples and number of data elements measured for industrial sectors. The UESP plan is designed to address PIPES requirements for a higher standard of quality in all four of these aspects.

A more integrated business statistics program

The PIPES implementation plan involves a fundamental reshaping of the underlying model for the collection and compilation of business statistics in the Bureau. A new model is required to improve consistency, coherence, breadth, and depth as noted in the previous paragraph, to permit a more effective management and control of the survey response burden that must be imposed on Canadian businesses and to facilitate detailed microdata analysis at the enterprise level.

The existing model for the business statistics program involves the application of over 700 distinct questionnaires within a framework of some 200 distinct business surveys. Each survey has its own sample design and selection process, its own processing and database systems, its own survey management and analysis staff and in some cases, its own separately managed survey frame.

So as not to overstate matters, it must be recognized that these existing survey vehicles do already have a very substantial degree of coherence and integration. Survey concepts are generally tied to the System of National Accounts framework. The Business Register Division provides a central frame from which numerous survey samples are drawn, as well as a centrally-managed program to provide estimates from tax data in place of survey-derived estimates, thus contributing both to integration and response burden reduction. A central pool of professional methodologists and systems experts , in the Methodology and Informatics Branches, supplies human resources for most of the surveys. Many, though not all of the surveys use the common software toolbox that has been developed over the past decade, consisting of DC2, GES, X11ARIMA and a number of other programs for data capture, imputation, estimation, aggregation, seasonal adjustment, analysis and tabulation. Most also use the common pool of human resources in the Operations and Integration Division and the Regional Operations Branch for data collection. Survey metadata, including questionnaires, are stored in a common database, the Statistical Data Documentation System (SDDS), and survey outputs are stored centrally in CANSIM. The Statistical Information System (SIS) provides a facility for combining and analyzing survey microdata from multiple sources, and the Company Contact System and Survey Impact Database (SID) offer partial means of tracking survey response burden for individual enterprises. Through these and other mechanisms, the business statistics program has already realized a high degree of integration.

Nevertheless, the program remains less well integrated than it could be. This paper outlines a restructuring plan to build a more coherent, more consistent, more comprehensive and more cost-effective system for gathering and compiling economic information, which when implemented will enable Statistics Canada to meet its new commitments with respect to the provision of high-quality annual provincial economic statistics.

Overview of the basic plan

The main elements in the proposed plan are as follows:

- Data collection from businesses will be treated in an enterprise-centric fashion. The data originating from a particular enterprise, whether collected from the headquarters, from one of its establishments or from one of its legal entities, will be kept together and treated as a single interrelated set of information. This will enable a reduction of duplication in questionnaires and an increase in the use of administrative data as an alternative information source, permit more varied types of aggregations and analyses, facilitate the enforcement of confidentiality restrictions and the prevention of residual disclosure, and assist in managing and controlling the response burden on individual enterprises.

- Instead of collecting data by means of multiple distinct business surveys, there will be an integrated corporate survey of all statistical enterprises. This integrated survey will have three parts:

    - An enterprise-level census of the approximately 8,000 complex enterprises, each having establishments in more than one industry and/or in more than one province;

    - A "take-all" survey of the approximately 60,000 establishments associated with the complex enterprises; and

    - A sample survey of the approximately 2,000,000 simple enterprises, each having establishments in only one industry and province.

- The integrated surveys of complex and simple enterprises will be carefully stratified to realize (a) approximately equal reliability across provinces and territories and (b) highest relative data quality for those industries with greatest priority. The priorities will depend on a variety of historical and current factors, notably the needs of existing data users (including cost recovery clients) and the HST revenue allocation formula requirements of PIPES.

- Instead of many non-integrated survey questionnaires, there will be a single, consolidated enterprise questionnaire, consisting of a main module directed to the enterprise and a number of additional schedules directed at the component establishments of the enterprise. The consolidated questionnaire will be managed within a relational database system so as to permit customized questionnaires to be generated for individual enterprises. It will be administered to respondents in electronic rather than paper form where this is acceptable to the enterprise.

- The North American Industrial Classification System (NAICS) will be fully implemented in the new Unified Enterprise Statistics Program.

- Revenue Canada's "General Index of Financial Information" (GIFI) will be used instead of survey vehicles to obtain balance sheet and income statement statistics for the simple enterprises. Other administrative data from the GST and payroll deduction systems will also be incorporated over time. Using these administrative sources instead of more lengthy surveys will permit us to produce reliable data with lower respondent burden than would otherwise be the case.

- All data collected from Canadian businesses will be warehoused in a standardized and interconnected set of relational microcomputer databases, where they can be easily accessed by any properly authorized data compiler or analyst in Statistics Canada. In other words, the raw data themselves, collected from Canadian businesses, will become a truly corporate resource. This will bring increased flexibility in developing new cross-cutting statistical products and in responding to requests for custom tabulations.

- All business survey samples, without exception, will be drawn from the central business register.

- The existing company contact database and monitoring system will be significantly enhanced to keep track, over time, of the total response burden put upon each Canadian enterprise. Policies and procedures will be put in place to ensure that all business survey contacts will be recorded in this database.

Implementation timing

The new Unified Enterprise Statistics Program is a highly ambitious undertaking. Its achievement promises to bring enormous benefits in the form of more reliable, detailed, timely, consistent and comprehensive statistical information, at minimal unit cost in terms of dollars and response burden.

Implementing this challenging vision will take time. At this stage the following three-step scenario is envisaged:

1. In early 1998 we will implement a new annual survey of the approximately 8,000 complex enterprises. This new survey will be closely integrated with the existing IOFD quarterly survey but augmented to include: (1) sufficient income statement detail to calculate enterprise value added, (2) enterprise capital expenditures and planned expenditures, (3) enterprise employment and (4) profiling information. The IOFD quarterly survey will be enhanced to be totally compatible with this new survey vehicle. In early 1998, for the 1997 reference year, we will also begin to implement (1) a new unified "take-all" annual survey of establishments attributable to the complex enterprises, and (2) a new unified annual survey of simple enterprises. In the initial year, these two integrated surveys will be implemented for a few industries where no business survey program now exists, and will collect fiscal year information to be calendarized, after the fact, by Statistics Canada. The specific industries to be covered in the 1998 surveys will include the construction industry as well as selected services and transportation industries, and fish aquaculture, — industries we will call the "1998 industries".

2. In 1999, for the 1998 reference year, we will expand the new unified annual survey of establishments attributable to the complex enterprises to cover as many as possible of the remaining industries. This means that agriculture, fishing and trapping, logging and forestry, mines and quarries, manufacturing, transportation, communication, finance, insurance and real estate, and business and personal services will be brought into the survey, as well as cross-cutting surveys covering broad economic sectors such as the employment and investment surveys.

3. In the year 2000, the process will be extended by expanding the unified enterprise survey program to include selected sub-annual business surveys along with the annual ones.

To realize these goals, the following supporting objectives will also have to be achieved:

- An integrated business register, combining the information in the IOFD frame with that in the BR, must be ready by the fall of 1997. Preliminary NAICS codes must be incorporated into the business register at the same time. This integrated business register must be expanded to reflect the information in the farm register as well, by the fall of 1998.

- The company contact management system must be sufficiently enhanced by early 1998, to track response burden by enterprise. By early 1998, the system must have complete coverage of the "1998 industries". By early 1999 it must have comprehensive coverage of all industries, including non-BTS surveys (e.g. Balance of Payments; Workplace and Employee Survey) as well as BTS surveys. In later years, the coverage will be extended to include sub-annual as well as annual surveys, to ensure that Statistics Canada's surveys of business enterprises can be managed in a consistent, integrated and coherent manner. The system must be efficient from an informatics point of view, and it must be supported by well-articulated corporate policies requiring all survey contacts to be registered in the system.

- Clean tax databases must be developed and made easily available on-line for the T1-GIFI, T2-GIFI, GST and PD7 sources. If GIFI implementation lags, then contingency plans must be developed to capture a sample of tax data in the interim.

- A new consolidated questionnaire must be conceived that (1) integrates all existing business questionnaires, (2) avoids duplication, (3) enforces consistent concepts and (4) limits the amount of information requested by recognizing what is already available from administrative sources.

- New "client-server" methodologies and computer systems must be developed and put in place that will facilitate:

  - Storing and accessing microdata from the unified surveys and the tax databanks;

  - Capturing, editing, imputing, estimating and aggregating the survey data that are collected;

  - Generating customized, integrated questionnaires for complex enterprises;

- Interfacing easily with the BR ; and
- Implementing the enhanced company contact management system.

- A "key provider manager" program must be put into place, involving the hiring and training of several senior company contact officers and the commencement of discussions with approximately 50 selected enterprises.
- A variety of field tests will be needed in 1997.

These plans require an immediate mobilization of Bureau resources and a collaborative team effort by employees. Rapid and revolutionary change of this type will inevitably be unsettling and we must be prepared to address honestly and openly the legitimate concerns of staff. Information must be well dispersed and discussed. At the same time, we must keep our eyes fixed on the objectives and the many benefits that will result from their realization.

Introduction

Statistics for the HST revenue allocation formula

On October 23, 1996 the Governments of Canada, New Brunswick, Nova Scotia and Newfoundland announced formal agreement on an arrangement to harmonize sales taxes. Effective April 1, 1997 there is a single harmonized sales tax rate of 15% in the three provinces, instead of three separate sales taxes and a distinct federal Goods and Services Tax. The revenues from the combined tax is collected by Revenue Canada and divided among the four governments by means of a specified revenue allocation formula. The formula is based on aggregate statistics compiled by Statistics Canada. This harmonized tax arrangement is expected to be simpler and less burdensome for consumers and businesses, and cheaper for government to administer.

During the intergovernmental discussions leading up to the harmonization agreement, it was considered important that the statistics used in the revenue allocation formula be highly reliable. The Government of Canada undertook to substantially improve the quality of the available provincial economic statistics and Statistics Canada drafted a plan to accomplish this objective. The plan subsequently put forward by the Bureau and accepted by the four governments entailed the following principal elements:

- Annual provincial input-output tables and greatly improved annual provincial economic accounts estimates;
- More frequent family expenditure surveys with a larger sample size, regular surveys of residential repairs and renovations, and a regular survey of travel expenditures;
- An expanded and improved business register database, recording the names of all Canadian businesses of significant size, their addresses, their industry classification and other key information necessary for survey-taking purposes; and
- A much improved business survey system, yielding more accurate, complete and detailed business data of roughly equal reliability for all twelve provinces and territories.

Statistics Canada was asked to start immediately on the project and to try to realize as many of the statistical improvements as possible for reference year 1997, the first year in which the new revenue allocation formula applies.

An implementation strategy based on tighter integration

This paper considers the full range of issues relating to the last, and to some extent the second-to-last of the four project elements just mentioned. Planning for the others is being done separately, although they are closely related.

Two alternative approaches were considered for implementing the project. The first, which was ultimately rejected, would have involved a broad scaling-up of the existing business statistics program to achieve the desired data quality improvements. Survey sample sizes would have increased, questionnaires would have lengthened and new surveys would have been initiated to fill data gaps, but the system itself and the underlying model upon which it is based would have remained largely unaltered. On the face of it, this approach might seem the surest way to achieve the project goals with minimal risk and delay. However, the scaling-up strategy is problematic for the following reasons:

1) The existing survey system provides no assurance of complete business sector coverage. Information is collected from individual establishments and any given business may have several of these. There is significant risk, in the current system, that some of these entities may be overlooked. The underlying model can be improved by giving it a more top-down focus, measuring revenues and expenditures at the enterprise level first and then determining their distribution across establishments.

2) The data quality improvements required for PIPES relate to improvement in at least four aspects: consistency, coherence, breadth, and depth. Consistency refers to more consistent use of concepts, standards, statistical units, methodology, and terminology. Coherence refers to the coherence of data within an enterprise. Breadth relates to the comprehensiveness of coverage in terms of filling gaps in measurement of industrial sectors and ensuring comprehensive non-duplicative coverage of all business activity of complex enterprises. Finally, depth refers to the size of samples and number of data elements measured for industrial sectors. It would have been impossible to address quality improvement in all these aspects of data quality using the scaling-up strategy.

3) Statistics Canada presently uses more than 700 questionnaires for its business survey-taking and there is duplication, inconsistency and a general lack of integration among them. Allowing this situation to continue and indeed, adding to the problem, would imply a higher-than-necessary burden on respondents. The project offers an opportunity to rationalize, integrate and simplify survey questionnaires, thereby making it easier for businesses to provide the required information.

4) No effective means is presently available to monitor and control the total response burden imposed on each individual business enterprise. Yet increasingly, businesses are seeking to negotiate that total burden one-on-one with Statistics Canada. The need for a database and monitoring system for this purpose will increase as the aggregate response burden rises and it would be difficult if not impossible to implement an adequate control mechanism without at the same time developing a more tightly integrated business survey system.

5) As the statistical system faces increasing demands for more detailed and regionally disaggregated information, the magnitude of the challenge involved in maintaining the confidentiality of data from individual businesses increases. The risk of residual disclosure, in particular, is much greater in a non-integrated data collection system. The maintenance of data confidentiality can be better ensured through the application of uniform methods.

6) Finally, there is a growing need for detailed micro-information at the enterprise level. The present system collects a wide range of detailed facts, primarily from the establishment, business location or legal entity, and stores the information in a constellation of non-standardized databases dispersed through the Bureau. The lack of standardization and general dispersal of information often makes it prohibitively expensive to conduct analyses across survey boundaries and this situation must be rectified.

So the scaling-up approach was rejected and the Bureau opted instead for a general restructuring of the business statistics program that will bring greater integration, consistency and coverage, will improve operating efficiency and will allow better management and control of the aggregate response burden.

Purpose of this paper

The first discussion version of this paper was distributed widely in Statistics Canada on December 20, 1996. PIPES implementation team members met widely across the Bureau with many divisions throughout January and into February 1997. This revised version of the paper has benefited enormously from the comments received at these meetings. The feedback received was detailed and thorough, with many suggestions both for how the long term vision might be clarified and how the implementation plans might be improved. Many divisions provided a considered response to all elements of the plan. These comments are reflected in this revised version of the paper.

This paper is intended to serve as a baseline document describing the long term UESP goals. As implementation of this plan proceeds, the process in reaching the long term vision outlined in this paper is expected to evolve in response to practical circumstances and further work. This paper should serve as a benchmark guiding document outlining the long term goals and characteristics of the business survey program and the rationale for change.

Structure of this paper

This paper has the following organizational structure.

It begins with a discussion of the PIPES objectives deriving from the use of data from the System of National Accounts in the HST allocation formula. The chapter discusses the implications of these objectives for the business statistics program.

The second chapter provides an overview of the broad structure of the Unified Enterprise Statistics Program (UESP). The goal of the chapter is to let the reader see the big picture before delving into the details.

Chapters 3, 4 and 5 consider the major component parts of the integrated business survey program. The three parts are (1) an enterprise-level census of enterprises with complex organizational structures, (2) a census of the establishments associated with complex enterprises and (3) a stratified survey of the simple enterprises.

The next chapter considers a range of sample design and stratification issues associated with the new survey framework.

Chapter 7 discusses key characteristics of the business register, which will be crucial to the success of the project.

Chapter 8 is devoted to the integrated questionnaire. Central to the plan is the idea that all existing business survey questionnaires will be brought into a single master collection vehicle, administered by means of a questionnaire database system that can generate customized versions of the integrated questionnaire appropriate for each individual business enterprise. This chapter explores the nature of the task involved in designing the integrated questionnaire.

Systems of classification are the subject of Chapter 9. The North American Industrial Classification System (NAICS) and the Central Product Classification System (CPCS) both play a key role in the plan.

The role of administrative data in the overall strategy is considered in Chapter 10. Administrative data from Revenue Canada are vital, both in permitting a much lower survey response burden and in generating highly accurate statistics at affordable cost.

Chapter 11 looks at a variety of issues related to computer systems. The plan envisages an integrated, distributed database containing survey and administrative micro-data that are easily accessible by authorized employees for program analysis and output purposes. New computer systems for questionnaire administration and manipulation, survey processing, response burden management, quality assessment and other purposes must be developed.

Chapter 12 looks at a variety of implementation and transition issues and proposes a sequence of steps that will allow greatly improved provincial statistics to be produced for reference year 1997.

The paper concludes with a chapter summarizing some of the key elements of the plan and the challenges ahead.

In addition, a number of annexes are included at the end of the paper. These provide a glossary of terms plus background information about Statistics Canada's existing survey questionnaires, information about Revenue Canada's principal administrative data source and the data requirements and priorities for PIPES from the System of National Accounts.

Chapter 1

PIPES Requirements and their Implications for Economic Statistics

With the signing of the agreement to harmonize the sales taxes of Newfoundland, Nova Scotia and New Brunswick with the federal GST new demands are placed on the Provincial Economic Accounts for greater consistency, accuracy, depth and detail. The following chapter outlines the general specifications of the Harmonized Sales Tax agreement as it relates to the Statistics Canada data used to determine provincial revenue entitlements. The attributes of a new System of Provincial Accounts (SPA) sufficiently complete and accurate to be used for revenue allocation are outlined. The remainder of this paper explains the proposal for a unified enterprise statistics program designed to provide the data necessary for the new SPA.

The Harmonized Sales Tax agreement

The HST rate is comprised of 7 percentage points for the federal component and 8 percentage points for the provincial component and will have the same basic operating rules as the GST applied in the other provinces and territories. Because the HST is in effect a tax on final sales but collected at the stages of production and distribution, the allocation of the federal and provincial portions under a strictly administrative approach would be unduly burdensome on businesses, requiring articulation of origin and destination of all transactions. This approach would also imply an enormous compilation and audit cost for the governments involved. Instead, the revenues from the combined tax are collected by Revenue Canada and divided among the four governments by means of a specified revenue allocation formula. Provincial Economic Accounts aggregates already published by Statistics Canada are used to calculate roughly 80% of the tax base in the revenue allocation formula while administrative records are used for the remainder.

Specifications of the HST/GST

The GST is a final consumption tax that is collected at the various stages of production and distribution of commodities to enhance compliance. Under the GST a given commodity is either:

1. taxable at 7% , which applies to most goods and services,

2. zero rated, which means taxable at 0% with full input credits on purchases; this applies to basic groceries, prescription drugs, any good or service exported, certain transportation services and purchases by aboriginal peoples, or

3. tax exempt, which means not taxable but the producers cannot claim input credits except on exported sales and public sector bodies can claim partial rebates on purchases. Examples of these commodities are residential rents, health care services, educational services, child care services, legal aid, financial services, ferry, road and bridge tolls, union dues and services provided by public sector bodies.

The HST has the same characteristics as the GST but the rate is 15%. To achieve a level playing field inside and outside the harmonized zone, additional rules apply regarding supplies bought outside the harmonized zone.

The revenue entitlement for each of the four governments is based on a general revenue allocation formula. The provincial share of the tax pool collected by Revenue Canada is determined by a provincial tax rate times a provincial tax base. The bulk of non-rebatable HST/GST is included in final expenditures made by households either on consumer goods and services or investment in new and existing housing due to the fact that businesses receive input tax credits. Some HST remains in the business sector because they produce or use exempt commodities in their production process. Public sector bodies and non-profit institutions also pay HST/GST on purchases and receive partial rebates. For the purposes of the formula, the base has been partitioned as follows:

1. Consumer expenditure
2. Housing
3. Business sector
4. Public sector bodies.

Each tax base can be calculated, in principal, by multiplying detailed final expenditure levels (excluding tax) times a taxable proportion for each commodity. This is in fact what is done for the base on consumer expenditures which uses provincial economic accounts data, and the non-financial business sector which uses commodity detail from the inter-provincial input-output tables. The housing base is determined by dividing the provincial economic accounts estimate of HST net of rebates on housing expenditures for a province by an average federal tax rate. The HST tax base for financial institutions and public sector bodies is determined from administrative databases at Revenue Canada. With regard to the four tax bases, Statistics Canada is required to provide data accounting for 100% of tax bases 1 and 2, for about 50% of tax base 3, and Revenue Canada provides the data for the remainder of tax base 3 and all of tax base 4. The following is a simplification of the revenue allocation formula for the purpose of exposition. The official formula can be found in Annex A of the Tax Harmonization Agreement.

$$\text{Provincial Base} = \underline{\textit{Consumer Base}} * \delta + \textit{Housing Base} + \textit{Business Base} * \delta$$

$$+ \text{Finance Base} + \text{Public Sector Base}$$

Where δ is a calibration factor reflecting the relationship of tax collections to tax estimates which may differ over time due to collection lags, statistical errors, or non-compliance. The bases in italics are databases provided by Statistics Canada and represent about 90 percent of the total base. There is no calibration factor for the housing base because it is implicit in the mechanics of how it is calculated.

$$\text{Provincial Entitlement} = \text{Provincial tax rate} * \text{Provincial Base}$$

This revenue allocation formula is applied using the most recent Provincial Economic Accounts (PEA) data available and used to allocate the monthly collections. The agreement specifies that interim adjustments to the payment entitlements are made periodically during the entitlement year (two or three times) and once a year thereafter

using the most recent PEA adjusted for population change. The agreement stipulates that the entitlement for any given year becomes "final" no more than 180 days following the publication of the "final" Provincial Input-output Accounts (PIOA) for that given period. In a side letter to the tax harmonization agreement, Department of Finance has stated that they expect the "final" PIOA to become available about three years after the reference period (in the year 2000 for reference year 1997) and no later than four years after the reference period.

The Current System of Provincial Economic Accounts

The System of ***National*** Accounts at Statistics Canada is comprised of Annual National Input-Output Tables available with a two and one-half year lag. These accounts provide the benchmarks for the production accounts of the Quarterly National Economic and Financial Accounts and the monthly constant price Gross Domestic Product by Industry accounts. The monthly and quarterly accounts are produced with a 55 to 60 day lag on the reference period.

The annual input-output estimates articulate the production, use and disposition of commodities by industries right through to final demand by consumers, governments, domestic businesses and foreign markets. All components of the national accounting system are fully integrated by benchmarking to the appropriate component of the balanced input-output system. In addition, the structural information obtained from the balanced tables forms a basis for projection for the more current monthly industry GDP accounts or the quarterly Income and Expenditure Accounts. For example, the annual ratios of imported commodities to final demand for those same commodities are used to project current estimates of final demand from the monthly import statistics, or the ratios of value added to gross production are used to calibrate the monthly GDP by industry accounts where only projectors of gross production are available.

The national system is a highly developed one where the estimates of GDP by the three methods, sum of value added by industry, sum of final expenditures plus inventory change, and sum of factor incomes (primary expenses), produce reasonably close estimates. In level, the results are rarely greater than one percent different from each other, even for the preliminary data, and the growth rates rarely diverge by more than 0.3 percentage points. This is one of the most robust set of national accounts in the world, largely due to the high quality national statistical system on which they are based.

The System of ***Provincial*** Accounts currently in place falls short of the national program in many ways. The conceptual system of allocation of national estimates by region is not as well developed as its national counterpart and no international standards exist. Significant differences still exist between the conceptual treatment of some items in the PIOA and the PEA. A feature of the basic accounting standards on which the National accounts are based is that the measures of savings and net lending of the various sectors of the economy consolidate to form the factor incomes and final expenditures for the production account. The mechanics of this same consolidation are problematic in the case of regional accounts. The basis for regional allocation of any particular variable which is best suited for the measurement of where production took place may not be appropriate for measures of savings and net lending/borrowing. The allocation of head office expenses and revenues is a case in point. The treatment of the FST which was a sales tax levied at the production point, and the predecessor of the GST/HST, is another example as allocating federal government revenues and expenditures is a version of the head office problem.

The system is incomplete. The PIOA exist for only two time periods, 1984 and 1990, based on availability of data on origin and destination of commodity flows. The Income and Expenditure parts of the system have been under development since the early seventies. Since data on inter provincial trade flows were unavailable for most time periods, GDP was estimated using the sum of factor incomes approach, final domestic demand and inventory change were estimated independently, and the net trade balances were calculated as a residual. This is a substantial shortfall from the national system where three independent estimates of GDP are available. More recently, estimates of inter-provincial trade flows and provincial international trade flows were developed for the period 1984 to 1989. Together with the structural information from the PIOA for 1990 and 1984, estimates for inter-provincial and international trade flows by province were estimated for the entire period of 1981 to the most current year first published in the fall of 1994. Estimates of provincial value added by industry at constant prices (PGDPI) were also developed annually from 1984 to the most recent data year and were published for the first time in the fall of 1994. (These had previously been published for only selected industries -- mostly goods producing ones.) Therefore three independent estimates of complete GDP by province have only existed for two years.

Full statistical integration of all measures of the provincial accounts has not been achieved since there are only two years of fully balanced provincial Input-output tables and conceptual differences still exist between the PIOA and PEA.

Finally, the variation of the estimates by the three measures of GDP is not nearly as consistent in level as for the national program and the growth rates for GDP can vary substantially. This is particularly problematic through the troughs of a business cycle as in 1991 when it is difficult to determine the cyclical characteristics of the provincial economies. A comparison of the two published constant price growth rates is shown on table 1.1 below.

Table 1.1 Provincial Gross Domestic Product

Year to Year percent change published as of May 1996[1]

	1990		*1991*		*1992*		*1993*	
	PEA	*PGDPI*	*PEA*	*PGDPI*	*PEA*	*PGDPI*	*PEA*	*PGDPI*
Newfoundland	-1.2	-1.0	-0.8	-2.7	0.4	-1.8	1.6	1.5
Prince Edward	0.9	-0.4	1.2	-0.5	2.3	2.3	1.3	1.3
Nova Scotia	-0.4	0.1	0.1	-0.5	1.9	0.9	1.4	1.4
New Brunswick	-1.6	-0.6	1.1	-0.3	3.8	0.9	4.1	2.1
Quebec	0.2	0.2	-2.0	-2.8	0.3	0.1	2.5	2.8
Ontario	-2.4	-2.6	-3.2	-2.8	1.0	0.3	1.2	1.6
Manitoba	2.9	1.4	-4.7	-4.0	1.2	1.4	0.9	0.3
Saskatchewan	9.1	6.5	1.4	-0.5	-5.0	-4.0	3.1	4.0
Alberta	2.4	3.1	0.1	0.4	0.3	1.2	5.7	6.0
British Columbia	1.9	1.9	0.4	0.9	2.4	3.4	2.4	3.4
Yukon	4.4	6.9	1.5	-0.7	7.9	12.4	-21.3	-18.0
North West Territories	4.9	-4.3	0.7	-0.2	-6.7	-1.8	4.3	0.8
Canada	*-0.2*	*-0.2*	*-1.8*	*-1.8*	*0.8*	*0.6*	*2.2*	*2.6*

[1]PEA is GDP at market prices and PGDPI is GDP at factor cost. Generally the growth rates of the two are similar.

Despite these shortcomings, the structure and methodology to produce the detail needed for the revenue allocation formula already exist. However, to achieve sufficiently robust estimates to be used for tax allocation purposes, these accounts must be of consistent quality across provinces and through time. To achieve that goal, an annual System of Provincial Accounts (SPA) which mirrors the national counter part in complexity and detail is needed, backed up by a statistical feeder system which can provide more consistent quality source data (both across provinces and through time). This implies annual Provincial Input-output Accounts fully integrated with the other parts of the system. It is through the balancing process of tracking commodities from their source of supply through to the ultimate end use that the accuracy of the estimates is tested and the projection techniques of the preliminary estimates are refined. This process forms the audit of the expenditure bases used in the allocation formula. The structure of the PIOA is also required to provide estimates of the "taxable proportions" used in the allocation formula.

The Basic Requirements of the new SPA

The requirements of the new SPA can be assessed in terms of the data needed to fill in the demand and supply equations of a fully integrated set of accounts. An upgrade in quality of survey results by province through larger samples, enhanced editing and up-to-date survey infrastructure is required. The initiative also means increased commodity detail at the final sales stages of distribution and annual estimates of origin and destination of shipments and wholesale sales. Gaps in the statistical system must also be addressed. The following sections articulate the most important requirements beginning with components of final demand because of the direct impact on the allocation formula.

Final Demand for Goods and Services

Personal expenditure

Personal expenditure is the largest component of expenditure based GDP, at over $475 billion a year, and provides the largest component of GST. Improvements in direct measures of personal expenditures are an important element of this proposed data improvement initiative.

Currently, provincial estimates of personal expenditure are based largely on surveys which measure sales by the suppliers of goods and services bought by households, complemented by with the occasional Family Expenditure Survey (FAMEX), a demand side measure. Three shortcomings exist with this approach:

1. These suppliers also sell commodities to businesses and governments for which a differing tax treatment exists. Sales broken out by class of customer are not always collected.

2. Sales are not collected by commodity in most cases, but commodity detail is imputed from past retail commodity surveys and related indicators.

3. The demand side measure is too infrequent to be used directly and is often only used as an allocator or as supplementary information.

In addition, almost all of the surveys used to measure personal expenditure were designed to produce reliable national estimates, and do not measure consistent quality among provinces.

The improved system should:

1. Include an annual FAMEX so that an annual demand side measure could be fully integrated into the commodity balancing process. It could be re-designed to be a provincial survey with a commodity structure more readily mapped to the Input-Output commodity classification and J-Series used in the PEA and in the allocation formula.

2. Improve other existing demand side surveys such as the Canadian Travel Survey.

3. Improve and augment supply-side data with a provincial and commodity content approach. These include:

    - Retail Trade / Retail Commodities;
    - Direct Selling;
    - New Motor Vehicle Sales / Motor Vehicle Dealer industry;
    - Restaurants, Caterers and Taverns industry;
    - Traveler Accommodation industry;
    - Amusement and Recreational Services industry;
    - Accounting and Legal Services industry;
    - Telecommunications industry.

The reconciliation of demand and supply is essential to the validation process. For example:

- Supply of personal services is essential to the measure of household consumption. These estimates can be compared to the Family Expenditure Survey results. Such a comparison could also be useful in calibrating the GST estimates based on the PEA with the GST actually collected. Consumption takes place on both "declared" and "undeclared" activities. However the GST is only collected from declared activity.

- The Retail Trade / Retail Commodity Survey should provide detailed information by province, by class of customer. For commodities like motor vehicle parts and repairs, gasoline or the personal/business use of motor vehicles, the data will have to be confronted with the FAMEX results.

- The Restaurants, Caterers and Taverns Survey information should be supplemented by the Family Expenditure Survey in the instances for which the class of customer could not be obtained or is unreliable.

Investment in Housing

Residential Construction is the second largest component in terms of the provincial allocation of non-recoverable GST and a large and volatile component of expenditure based GDP at around $40 billion a year. It includes three components:

- new housing,
- alterations and improvements,
- transfer costs.

Each one of these poses measurement challenges in terms of GST.

New Housing

The required input into the revenue allocation formula is GST/HST on the sales of new property.

In the current system of national/provincial accounts, investment in new residential construction is the value of the work put-in-place, not the value of sales of property. Data produced by Canada Mortgage and Housing Corporation on "absorption" of newly constructed housing units have been used in the past in the national system to impute the GST. These data do not cover centres under 50,000 in population, which could be a substantial limitation for use of this data to monitor sales issues inside the harmonized zone. Given recent budget cutting and restructuring at CMHC it may be necessary to find an appropriate STC survey vehicle to collect provincial value of the sales of property, by type of structure and by sales unit value. A supply side survey of construction could gather the information as well as data for estimates of work put-in-place and the input and output structure of the construction industry. It should improve current estimates of total expenditure based GDP as well as the validation process taking place through the commodity balancing of the provincial input-output system. This data source should also include the value of the land by type of structure allowing for a construction cost valuation (excluding land) and other cost information such as developer fees, architect and engineering fees and the inventory of completed but unsold houses.

Alterations and Improvements

Alterations and improvements on residential construction by homeowners and by landlords is a large and growing component of the housing base. The two most important data sources are:

- the Homeowner Repairs and Renovations Survey
- the Landlord Repairs and Renovations Survey

The Landlord Repairs and Renovations Survey was implemented for the first time for data year 1995. This one-year survey is sponsored by CMHC. The Homeowner Repairs and Renovations Survey had been sponsored by CMHC for the last few years. In both cases, there is no guarantee of continued funding. The continuance of these two surveys, with provincial enhancements, is important to PIPES.

Transfer Costs

The GST related to transfer costs represents one third of the value of the GST on residential housing. Currently, transfer costs are estimated by moving a benchmark last updated for statistical year 1971, using a derived number of unit sales based on data published by the Canadian Real Estate Board. These data are not statistically representative of all housing sales as they only represents a simple unit count of houses sold through the Multiple Listing Service.

A new Real Estate Agents Survey is needed to obtain revenue from real estate commissions net of discounts, by province, separating the residential component from the commercial component. This survey should also provide the gross sales value by type of structure so that average real-estate-commission rates by type of structure can be calculated.

There is also a requirement for more information on other transfer costs such as the cost of property assessment, building inspections, etc. This could be done through the redesigned Family Expenditure Survey and/or by directly surveying the relevant specialists.

Production/Supply and Intermediate Use of Commodities

Data Gaps

There are a number of data gaps that need to be filled to produce annual provincial input-output tables. For the following industries, little or no data exist for revenues, intermediate expenses and GDP components:

- Business management consultants;
- Purchased service inputs by the forestry, mining and manufacturing industries;
- Industries such as services incidental to mining and to transport;
- Courier services;
- Taxis, ferries and urban transit
- Waste management;
- Revenue details by commodity for water transport to improve margin rates and balance commodity supply and demand.

- Output and intermediate input structure of the construction industries by province. Input data by broad structure types (residential non-residential building, etc.) are needed. The data must come from producers of construction such as general contractors, special trade contractors, developers and own-account producers. The collection vehicle for the contractor part, in the past, was the Survey of Construction. This survey should be reviewed and expanded to include real estate and land developers. Appreciable values of own-account construction are produced in the highway and engineering area. In addition the Survey of Construction should include own-account producers.

- Data on inter-provincial trade flows should be obtained on an annual basis for manufacturing, wholesale and domestic service industries. Information on international services collected by Balance of Payments Division should reinstate the provincial dimension.

Producers of Exempt Commodities in the Business Sector

Producers of exempt commodities in the business sector, including financial institutions, suppliers of residential rent and health practitioners pay GST on their inputs but cannot claim any input tax credit. They are the "final consumer" of GST-taxable commodities. For these industries, the data validation exercise is essential because knowledge of the taxable and non-taxable commodity input purchases by the producers of exempt commodities is not sufficient for a final measurement of the GST/HST. The complication is that exempt producers may claim input tax credits when they export their outputs. Therefore, the destination of the output is essential in assessing the percentage of zero-rated inputs by industry, based on international exports or inter-provincial exports to the non-harmonized zone.

Specifically, to facilitate the balancing exercise in order to provide accurate data for the portion of business-sector inputs representing final consumption of those commodities, it is necessary to:

- Enumerate the value of services provided by head offices to operating entities in order to determine the proportion of activity effectively taking place in each province and to allocate the consumption of taxable inputs based on activity. It has been the national accounting practice of countries to impute output and distribute the expenses of the head office. There are some industries where the definition of the provincial establishment and allocating head office expenses are difficult e.g. finance, air transport. Some resources should be devoted to profile and identify head office, division, etc. to build the provincial detail.

- Improve commodity content and geographical allocation for the finance and insurance industries.

- Improve accuracy of the provincial allocation of investment in machinery and equipment through a revised sampling strategy for the Capital Expenditures Survey.

- Measure the taxable repair expenditures of landlords through the Landlord Repairs and Renovations Survey.

Public Sector Bodies

Although the allocation formula will use the administrative collections and rebates data to derive the GST paid by public sector bodies, the PEA must have accurate measures of the transactions made by these bodies in order to calibrate the entire system. Better expenditure and revenue data could be obtained through more systematic and intensive use of administrative data pertaining to municipalities, hospitals, school boards, universities and urban transit commissions and an improved local government survey.

The GST paid by the public sector bodies, obtained from Revenue Canada, combined with expenditure data are used to generate effective tax rates that is compared with Statistics Canada's tax blueprint and allow for detailed validation of the numbers. This improvement remains an essential feature of the complete commodity balancing and is also important for GST calibration.

General requirements of the new SPA

In summary, to provide provincial accounts sufficiently robust for tax allocation purposes, coverage of all industries must be insured. This means closing existing gaps in the areas of business and personal services and construction. In addition, to allocate activity properly to the provinces and territories it is essential that all establishments of enterprises are well covered, especially those complex enterprises which operate across provincial boundaries. The data used must be consistent and currently relevant based on one up-to-date, complete survey frame. The industrial and commodity classifications should be applied uniformly. Sufficient detail must be collected in terms of commodity content, class of customer and origin and destination of sales/shipments articulated by province and territory. And last but not least, the data must be reliable in terms of minimal revisions and consistent quality across provinces. Federal and provincial fiscal planning is dependent on the integrity of this set of data.

Chapter 2

Overview of the Unified Enterprise Statistics Program

The use of Statistics Canada data in the HST allocation formula creates the need for an across-the-board improvement in the quality of economic statistics needed to compile provincial economic accounts. Data quality has many dimensions. The UESP is intended to improve data quality in a comprehensive sense -- consistency, coherence, breadth and depth.

The UESP will increase data consistency by extending the use of common concepts, terminology and standards across the economic statistics program. An integrated survey methodology will be used with consistently defined statistical units across economic sectors.

The coherence of data will be enhanced as the UESP focus will be enterprise-centric, interrelating data about parts of enterprises into a coherent framework. Coherence analysis tools and the Key Provider Manager approach for very large enterprises will address internal data inconsistencies for complex enterprises.

The UESP will augment the breadth of the economic statistics program. The business economy will be more comprehensively measured as gaps in the coverage of industrial sectors are filled. The enterprise-centric focus of the UESP will ensure more comprehensive, non-duplicative coverage of the business activity of complex enterprises.

Finally the economic statistics program will be deepened in terms of sample sizes and data elements measured. Sample sizes will be expanded to more equally measure economic activity by province. Data elements such as commodity production and inter-provincial trade will be included across many sectors of the economy. As discussed in the introduction to this paper, two options for implementing the project were considered. The option of a broad scaling up of the existing business surveys was rejected for the reasons outlined in the introduction. Instead the Unified Enterprise Statistics Program (UESP) involving a comprehensive restructuring of the business survey program was adopted.

This paper outlines the master plan for the UESP which will change the business survey approach at Statistics Canada based on an integrated enterprise-centric approach to surveying businesses. This chapter provides an overview of the UESP and the next three chapters elaborate each of the three parts of the unified survey approach.

Fundamental principles of the Unified Enterprise Statistics Program

The proposal comprises four fundamental elements:

1) The unit of statistical observation and primary contact shall be the enterprise – in contrast to the establishment, the legal entity and the various other units currently in use. All data collected will be related explicitly and directly to the business enterprise with which they are associated. In general, data collected for enterprises that work out of more than one establishment, and where the

different establishments cross industrial or provincial boundaries, will include all the relevant details on product and gross value of production for each of those establishments.

2) Administrative data will be used in preference to survey vehicles. Corporation and personal income tax returns, GST compliance forms, payroll deduction statements, customs forms and other administrative sources will be used to the full; maximum use will be made of the Revenue Canada Business Number and generally, direct burden will only be increased if there is no satisfactory administrative alternative.

3) There will be one and only one survey frame – the Business Register. Independent business frames, and business frames which are only indirectly linked to the business register, will be phased out as part of PIPES implementation. The Business Register will be augmented to include businesses identified through the introduction of the Business Number.

4) A central contact measurement system will be built which measures response burden by enterprise in a central database. Without such a tool, Statistics Canada could not manage properly its relationship with each business enterprise.

Major elements of the Unified Enterprise Statistics Program

Simple and complex enterprises

There are about 2 million business enterprises in Canada. Approximately one million enterprises are incorporated and of these, some 18,000 enterprises have more than one location. About 11,500 of these enterprises have more than one statistical establishment. Some 3,500 of these "multi-establishment" enterprises have establishments in only one province and industry and are therefore considered "simple". Approximately 8,000 enterprises are "complex" in the sense of operating in more than one province or more than one industry or both. Diagram 2.1 summarizes the approximate numbers of enterprises and establishments associated with various parts of the survey strategy. Detailed tables about complex and simple enterprises by province, industry division of the SIC and economic impact are incorporated into the body of this report. The implementation of NAICS might have the effect of slightly increasing the number of "complex" enterprises because NAICS defines more detailed industries especially in service industries and therefore, there is a greater chance that multi-establishment enterprises will cross industry boundaries under NAICS.

A key challenge of the UESP is to collect reliable, complete and consistent provincial information on sales, expenses, product, customer and location of establishments associated with these 8,000 multi-province, multi-industry enterprises without imposing on them an intolerable response burden.

Goals

The main goals of the new Unified Enterprise Statistics Program are:

- To achieve complete coverage in the measurement of value-added, addressing not only missing sectors of the economy but also parts of enterprises that may currently be missed with the distributed survey approach.

- To link explicitly all data related to an enterprise, in order to promote a better understanding of the relationship of establishments to the enterprise as a whole.

- To build a single, integrated Business Register to serve as the survey frame for all business surveys and to improve the annual updating of enterprise structure information by regularly updating the profiles of complex enterprises.

- To improve the consistency of concepts and questionnaire content across surveys and to eliminate duplicate requests for essentially the same information from the same business. Existing surveys use a variety of concepts and terminology, and a rationalization of questionnaires to eliminate duplication and to standardize data elements will help to limit response burden.

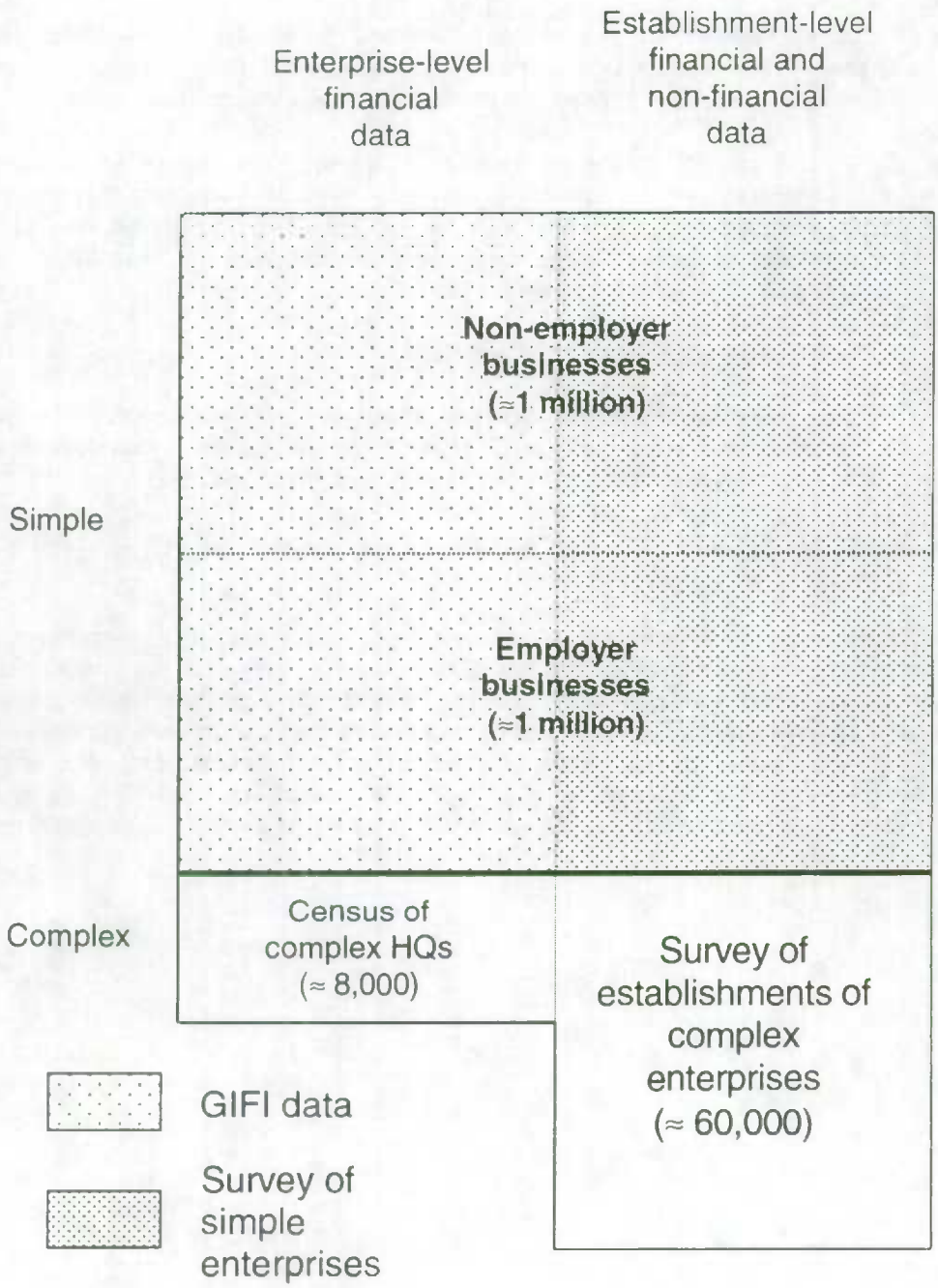
- To maximize the use of administrative data.

- To actively manage the response burden on individual businesses. Key components of this strategy are the creation of a central company contact database (at the enterprise level), the introduction of a "key provider management" program for the very largest enterprises, and the use of burden criteria in selecting enterprises in sampling. Burden on smaller enterprises will also be managed by use of short-form questionnaires where appropriate.

Establishments will be the standard tabulation unit within the UESP. However, the establishment may be replaced by other statistical or operating units in some instances. The establishment is the appropriate level for the collection of production characteristics. Use of the legal entity as a collection unit may result in reduced response burden when GIFI data become available. Regulated industries may require targeting of other non-standard units. It is a matter of weighing the advantages of consistently using one common statistical unit against those of meeting particular program requirements in a locally optimal way.

Diagram 2.1

The Unified Enterprise Statistics Program



Components of the survey program

Part I of the survey program will be a Census of Complex Enterprises, collecting information at the enterprise level about the consolidated balance sheet, the consolidated income statement (sufficient to calculate enterprise value added), investment, investment plans, employment and the enterprise legal and operating structures. Because of the multi-province and/or multi-industry aspect of these complex enterprises, tax data cannot provide the necessary provincial and industrial detail, so it is necessary to survey them. Complex enterprises will be asked to confirm, or update if necessary, a diagrammatic representation of its legal and operating structures.

Part II will be a survey (either a census or a sample) of establishments associated with complex enterprises. The survey will gather data on the income statement (sufficient to calculate value added), commodities produced and sold, class of customer, origin and destination, and other industry-specific information. To the extent possible, reporting relationships will be established at the enterprise level, although it is expected that some respondents will prefer that we deal directly with contacts at the establishment level.

Part III will consist of a sample of simple enterprises, i.e. those where the enterprise operates in only one province and industry. For these enterprises, information on the income statement and balance sheet will be obtained from tax data (GIFI) while other characteristics will be obtained by survey. As with the Part II survey, the other characteristics will vary by industry. Once complete GIFI data are available, the sample for the Part III survey may also serve to validate and calibrate the GIFI data, and to build regression models to be used for model-based estimation of characteristics not available from GIFI but correlated with them, similar to the approach now used with SEPH.

Sample design

The results from the Part II and III surveys will be combined to produce a wide range of establishment-level data. While there may be some content that is common across industries, much of the content will still be very sector-specific. The major issue from a sample design perspective, therefore, is how to design an efficient program of annual production surveys whose target populations are divided between complex enterprises and simple enterprises. The design needs to satisfy the objective of linking the production data to the enterprise data, the objective of improved consistency of concepts and content across annual surveys, the objective of managing response burden at the enterprise level, and the multiple sector-specific objectives of the Part III survey.

Key provider managers

One very important element in the proposed framework is the creation of a "Key Provider Manager" (KPM) program (an idea borrowed from the Australian Bureau of Statistics). This program will develop specially tailored working relationships with a progressively increasing number of the most important and/or problematic complex enterprises. All of Statistics Canada's data collection activities concerning these enterprises will be managed and controlled by the key provider manager. For example, questionnaires sent to these enterprises for the Enterprise Survey will be channeled through the KPM, as will any more highly specialized questionnaires to collect particular information.

Chapter

3

A Census of Large and Complex Enterprises

Introduction

The Canadian economy is dominated by about 8,000 complex enterprises which account for about 60% of total gross business income. It is essential to the Unified Enterprise Statistics Program that these enterprises be well understood and accurately measured not only at the national level but in terms of their activities in each province and territory. In estimating value-added for the total economy, the current business survey program has not succeeded in assuring complete and unduplicated coverage of all industries nor indeed of all the units in those complex enterprises in the industries which are surveyed. There are data gaps in the measurement of service activities taking place in their head offices and in the measurement of other strategic variables. For example, while good national data are available for corporate profits of complex enterprises, no information is collected on the associated provincial distribution.

Complex enterprises, whose component production units are classified to more than one industry and/or operate in several provinces, are by their very nature difficult to measure. As explained in Chapter 10, administrative sources will satisfy some of the data needs for these enterprises, but not all. A census of complex enterprises is crucial to the program, providing benchmark totals of operating revenues, operating profits, operating expenses, wages and salaries, and opening and closing inventories, for both the enterprise as a whole and for each of the production units comprising that enterprise. Tax data are inadequate for these complex enterprises because they provide no information about the provincial dimension or the industrial dimension of their activities. This census information, coupled with the same key economic variables from simple enterprises, will provide the desired economy-wide measurement of value-added. It will also provide the means to allocate enterprise characteristics to the production units in the enterprise and will facilitate the linkage of production unit data to the enterprise.

To implement this part of the unified enterprise program strategy, we need (1) an improved business register (see Chapter 7 for additional details), (2) a new questionnaire directed at the consolidated enterprise level and (3) various supporting systems (see Chapter 11). The rest of this chapter will elaborate on the strategy for the Census of Complex Enterprises (CCE).

Major elements of the proposed CCE

- The questionnaire for the Census of Complex Enterprises will consist of a set of questions addressed to the consolidated enterprise, a deconsolidation "spreadsheet" schedule requesting value-added data elements for each of the production units comprising that enterprise, a diagrammatic representation (portrait) of the legal structure, and a portrait of the operating structure. To this standard package will be

added additional schedules addressing other enterprise level characteristics (for Balance of Payments for example). Since the legal or operating structures can change at any time, the treatment of any changes discovered during the survey will have to be worked out as an imputation or edit follow-up.

- Industry-specific variables and those variables known only at the establishment level of the organization will be collected via schedules "attached" to the Part I questionnaire.

- Depending on respondent preferences, data for all schedules may be collected centrally from the enterprise headquarters or questionnaires may be delivered to any level in the enterprise. The timing of the mail-out of the various schedules may or may not coincide with that of the census questionnaire. Based on past survey experience, it is expected that in most cases schedules will be sent to the establishment level of the organization rather than collected entirely from the enterprise headquarters.

- The first CCE Part I survey will be conducted in early 1998, for the reference year 1997, for all complex enterprises. The deconsolidation spreadsheet and industry-specific schedules will be sent only to enterprises operating primarily in the "1998 industries", based on NAICS coding.

- A "Key Provider Manager" (KPM) program will be implemented to negotiate and coordinate the collection of data from selected large and complex enterprises. A KPM portfolio of enterprises, building up from about 50 in 1997 to as many as 300 in the year 2000 is foreseen.

- Reporting arrangements for the first census of complex enterprises will be established by a combination of personal contacts by the KPMs and possibly a pre-census reporting arrangements survey questionnaire, for the 1998 industries. It is anticipated that in most instances the enterprise contacts currently used by the IOFD quarterly survey will also be the contact for the Part I survey.

Complex enterprises

Complex enterprises are primarily defined as those which have production units in more than one industry and/or more than one province. A size criterion will also be applied to include large and simple enterprises and to exclude small but complex enterprises. More analysis is required to define precisely the universe of complex enterprises.

Table 3.1 provides a breakdown of most of the complex enterprises to be targeted by the Census of Complex Enterprises, by the 1980 SIC-E Division of the dominant establishment in the enterprise (as defined by revenue). See Table 4.1 for other enterprise characteristics.

Table 3.1 Complex Enterprises by SIC Industry Division

	SIC-80 Industry Division	Enterprises	Establishments
	Total	*8,119*	*56,614*
A	Agriculture	45	145
B	Fishing and Trapping	3	14
C	Logging and Forestry	26	97
D	Mining and Oil and Gas	174	994
E	Manufacturing	1,588	10,746
F	Construction	414	1,307
G	Transportation & Storage	368	2,327
H	Communications & Util.	284	1,595
I	Wholesale	1,885	8,945
J	Retail	948	15,710
K	Finance & Insurance	241	1,962
L	Real Estate & Insurance Op	198	1,247
M	Business Services	599	2,632
N	Government Services	149	971
O	Educational Services	52	195
P	Health Services	297	1,919
Q	Accommodation & Food	369	3,150
R	Other Services	479	2,658

Source: Business Register Division Tabulations – September 1996

Table 3.2 shows all multiple-establishment enterprises by the number of different industry divisions represented by establishments in the enterprise. This table includes all enterprises with more than one statistical establishment, classified to the industry of the dominant establishment in the enterprise. Each cell for the industry division row presents the counts of enterprises in that division with establishments in only that division, in that plus 1 other, 2 others, etc. For example, in agriculture there are 70 multi-establishment enterprises, with 41 of those with establishments only in agriculture, 28 with establishments in agriculture and one other industry division, and so on.

Note that no enterprise has establishments in more than 8 divisions and that 8,749 of the 11,605 (75%) have establishments in only 1 Division. 95% have establishments in the same division only or in just one other. This shows that very few enterprises are highly complex, at least in terms of their industrial composition as currently profiled.

Table 3.2 Multi Establishment Enterprises

	Number of Divisions →	Same	1 more	2 more	3 more	4 more	5 more	6 more	7 more	Total
	Total	*8,749*	*2,320*	*317*	*127*	*63*	*19*	*7*	*3*	*11,605*
A	Agriculture	41	28	1						70
B	Fishing & Trapping	6	1	2						9
C	Logging & Forestry	7	22	2						31
D	Mining and Oil & Gas	165	67	13	4	3				252
E	Manufacturing	1,085	541	91	28	19	9	2	2	1,777
F	Construction	320	177	21	3					521
G	Transportation & Storage	301	91	25	5	1	1	1		425
H	Communications & Util.	92	159	13	39	6	2	2		313
I	Wholesale	2,203	287	44	7	3	2	1		2,547
J	Retail	1,695	307	2,913	3	2	2	1	1	2,051
K	Finance & Insurance	232	42	8	2					284
L	Real Estate & Insurance Op	247	70	10	1					328
M	Business Services	658	77	13	3	1				752
N	Government Services	161	23	3		1				188
O	Educational Services	52	33	3		2				90
P	Health Services	332	77	7	15	15	1			447
Q	Accommodation & Food	592	171	12	1		2			778
R	Other Services	560	147	20	6	9				742
	-									

Source: Business Register Division Tabulations – September 1996

Content

The content of the enterprise survey was addressed by the Enterprise Task Force (headed by Brian Nemes). This task force proposed a questionnaire on consolidated enterprise characteristics supplemented with a spreadsheet for deconsolidated characteristics related to the value-added data elements. The questionnaire was to be accompanied by a portrait of the legal and operating structures of the enterprise on the Business Register. The respondent was to be asked to update the structures if necessary, and to complete the spreadsheet for establishments in the updated structure. Appropriate methods will be needed to ensure adequate information on structure changes to enable updates to be applied to the BR.

One issue of fundamental importance to the census questionnaire design, indeed for the whole survey strategy, is whether respondents are willing and able to provide the deconsolidation requested. Several versions of the deconsolidation spreadsheet have been proposed and these need to be tested with respondents. The data elements on the proposed deconsolidation spreadsheet appears to have much common with what is needed for updating the enterprise profile. Therefore, UESP will investigate whether these two components of the CCE can be combined.

As part of Phase 3 of the re-design of the Survey of Employment, Payrolls and Hours, approximately 1,000 large companies will be surveyed in 1997 to obtain a "deconsolidation" of the labour variables (among other data needs). The experience gained from that survey will be very useful in assessing the reaction of respondents to the proposed census deconsolidation.

A second issue for the CCE relates to the blending of the content of the IOFD financial surveys with that proposed for the census questionnaire. The IOFD annual and its companion quarterly financial survey provide the bases for a database of enterprise financial statistics that has a long history for a valuable constituency of users. A modified annual supplement with new questions on expense components and the number of employees and severance pay will be launched in early 1997 to add to the utility of that database. Different definitions for similar questions on the existing IOFD questionnaire and the future CCE questionnaire could confuse respondents and users so it will be necessary to harmonize these questions. Changes in the administrative data sources that also form part of the IOFD database will also have an impact on the concepts and definitions used in the IOFD surveys. (See Chapter 10 and Annex 3 for a detailed description of the administrative data sources and the issues they raise). The issue of CCE content requires further attention and testing.

Other data needs relating to the enterprise as a whole will be addressed by additional "schedules" attached to this questionnaire. For example, balance of payments data requirements, foreign ownership, capital expenditure investment intentions, and research and development activities will be considered. See Chapter 8 on the Integrated Questionnaire for more details.

As more experience is gained with the deconsolidation of enterprise-level characteristics, questions on intra-enterprise transfer pricing may be added to the questionnaire.

Response burden

The response burden implied by the CCE must be evaluated. Statistics Canada has made a concerted effort over many years to reduce the response burden of its business surveys, resulting in a current measure of 735,000 respondent hours in total for all businesses, and 532,000 (72% of total burden) for large businesses (not strictly comparable to "complex" businesses as defined here, but indicative). The deconsolidation question is new and therefore would be added to the response burden for those enterprises now surveyed. The census, of course adds to the response burden by including industries not now surveyed and expanding the coverage of those industries that are now only sampled. For example, the census would cover the construction industry, regulated industries such as telecommunications, and enterprises in education, health and culture that have in the past been addressed outside the Business and Trade Statistics Field. It could include businesses that have traditionally been surveyed by other federal and provincial departments, although changes of that nature will have to be negotiated. As well, the deconsolidation question would pick up production units that have been missed by existing survey programs.

In the longer term, administrative data may prove viable for some of the balance sheet and income statement data elements, especially for enterprises filing only one tax return. However, at least one census will be needed to evaluate the feasibility of this and to validate profiling and other information. For some enterprises, it may prove feasible to aggregate multiple GIFI reports to produce enterprise estimates. For others, whose profiles are quite static, the deconsolidation spreadsheet may not be needed every year.

A longer term prospect for reducing response burden may be found in a recent Canadian Institute of Chartered Accountants (CICA) exposure draft to establish standards for "Reporting Disaggregated Information about a Business Enterprise". The proposal "would require business enterprises to report certain information about the components of the enterprise (operating segments)", information paralleling that in the deconsolidation spreadsheet. If this becomes the established standard, then the information would be more readily available from companies or may even be supplied to RCT in conjunction with the GIFI information.

Profiling

A key component to the CCE is an accurate, timely profile to accompany the questionnaire. The Business Register Division maintains an annual cycle of legal and production entity profile updates for the largest 250 enterprises. Legal entity profiles exist in IOFD for all multi-legal-entity enterprises, updated using information from the Quarterly Survey of Financial Statistics and CALURA. Clearly these sets of profiles must be consolidated in the process of combining the IOFD and Business Register frames (see Chapter 5).

Respondents could react adversely to the census questionnaire if the profiles were badly out of date, so every effort will be made to update the profiles in advance of the census itself. However, most profiles will probably be updated by respondents at the time of the survey, which implies that the accompanying instructions must be well developed and tested.

More user-friendly tools are required to produce and maintain the profiles. A graphical interface to modify legal and operating structures will be developed. Improved feedback mechanisms are needed to incorporate the structural change signals generated by the Part II and Part III components of the UES program, as well as those generated by the sub-annual surveys. See Chapter 7 for more discussion on this issue.

A new method for identifying enterprises moving into scope for the census needs to be developed since the traditional signals from administrative data sources may not capture an increase in organizational complexity, although the CALURA staff have other methods that will help in identifying such changes. Also, the Business Register Division has been investigating the multi-province indicator found on the Payroll Deduction Account administrative source as a possible signal source.

As well, the current design of the Business Register and the processes to promote the "simple" (NIP units) to "complex" (IP units) need to be modified to reflect the definitions to be used in the UESP. The IP/NIP distinction will no longer be relevant under the UESP.

Processing

The census of complex enterprises is a new component of the program. As such, systems are required for each of the processing functions: questionnaire mail-out, collection, editing, imputation and estimation. The enhancement of existing systems in IOFD to meet these requirements will be investigated. More complex editing requirements are foreseen, linking data from the deconsolidation to that at the enterprise level, or to establishment information from Part II. More complex imputation requirements will use administrative data. The implementation of such strategies is contingent on the availability of links through the Business Register. See Chapter 11 for more detail on the database design and processing requirements.

The new GIFI data and GST data will have to be analyzed as they become available from RCT. The potential to substitute these administrative data elements for those collected by the census will be exploited to the extent possible.

Chapter

4

A Survey of Establishments of Complex Enterprises

Introduction

The Unified Enterprise Statistics Program represents a departure from current business survey practices by targeting the enterprise for the collection of value added benchmarks. It represents a coherent economy-wide approach to data collection, use of administrative data, survey processing and estimation. This presents us with a unique opportunity to introduce parallel improvements to the business surveys production statistics program. The purpose of this chapter is to outline an integrated (across industry sectors and within the UESP) approach to surveying the establishments of complex enterprises.

Major elements of the survey of establishments of complex enterprises

- The Unified Enterprise Statistics Program will collect all required annual data (enterprise level and establishment level) for complex enterprises.

- A "Key Provider Manager" program will be implemented to negotiate and coordinate the collection of data from selected complex enterprises.

- Industry specific variables and those variables known only at the establishment level of the organization will be collected via schedules "attached" to the Part I questionnaire.

- The schedules will have the same basic form for all industries with variations for industry specific data (such as commodity and origin and destination data). A rationalization of schedules to eliminate duplication and to standardize data elements will help to limit response burden. Wherever possible, administrative data will replace data collection from businesses.

- Part II is a process to collect production and other characteristics data in conjunction with subject matter divisions. Subject matter divisions will be responsible for integrating the data into their programs and performing analytical and dissemination functions.

- Where the enterprise so chooses, data for establishments will be collected centrally from the enterprise headquarters for the chosen unit of observation. Otherwise, we must be prepared to deliver questionnaires to any level within the enterprise.

- Part I will collect the full detail of value added. Where these data do not provide enough industrial and geographic detail to allocate value added to the province/industry level, these estimates will provide a control total for establishment level estimates. Part II will collect additional value added elements to the extent available by industry and province/territory.

- The stratification and sampling strategy adopted for Part I will also apply to Part II. Part I and Part II will together be a census of complex enterprises. This approach will yield a richer data base for integrating financial, value added and production data.

- The quality and level of detail of all production statistics should be maintained or improved in the new survey strategy.

Complex enterprises

Complex enterprises are those, which have establishments in more than one industry, in more than one province or those which have establishments linked to more than one legal entity. Administrative or enterprise level data do not provide the industrial or geographic detail required to replace current business surveys for the establishments of these enterprises. An idea of the number and characteristics of these enterprises is contained in the following tables. Note that the table presents four different breakdowns of the population of multi-establishment enterprises.

Table 4.1 Characteristics of Multi-Establishment Enterprises

Multi-Establishment Enterprises

	Enterprises	Establishments
Total	*11,605*	*67,491*
Single SIC-One Province	**3,486**	**10,877**
Multi SIC-One Province	**2,914**	**10,427**
Single SIC-Multi Province	**4,044**	**18,035**
Multi SIC-Multi Province	**1,161**	**28,152**
Single SIC	**7,530**	**28,912**
Multi SIC	**4,075**	**38,579**
Single Province	**6,400**	**21,304**
Multi Province	**5,205**	**46,187**
Multi SIC or Multi Province	**8,119**	**56,614**
Single SIC & Single Province	3,486	10,877

Source: Business Register Division Tabulations – September 1996

The 3,486 enterprises, which have establishments in only one province and in one industry, may require surveying only for other characteristics (see next chapter for details). Administrative (or CCE) data will provide most value added type information for the industry/province aggregation of these establishments. The 8,119 enterprises, which have establishments in more than one industry or in more than one province, are the complex enterprises, which will be covered by this survey. There are about 1,000 enterprises linked to more than one legal entity. The overlap between the enterprises meeting this complexity condition and the 8,119 enterprises above is thought to be very high.

Complex enterprises are concentrated in the Manufacturing, Wholesale and Retail sectors (54% of enterprises and 63% of establishments). Due to the size and contribution to estimates of these large enterprises, they are presently well represented in business surveys. Of the 4,447 enterprises classified to logging & forestry, manufacturing, wholesale and retail, 4,163 (94%), are covered by the respective current annual programs. The annual manufacturing survey program covers all 1,613 manufacturing, logging & forestry enterprises, the annual retail and wholesale trade survey includes 937 of the 948 complex retail enterprises and 1,613 of the 1,885 complex wholesale enterprises in its sample.

Table 4.2 All Enterprises on BR

	Enterprises			*Establishments*		
Industry	*Prov > 1 or SIC > 1*	*Prov = 1 and SIC = 1*	*Total*	*Prov >1 or SIC > 1*	*Prov = 1 and SIC = 1*	*Total*
Agriculture	45	52,403	52,448	145	52,435	52,580
Fishing, Trapping	3	5,776	5,779	14	5,791	5,805
Logging, Forestry	26	10,737	10,763	97	10,753	10,850
Mines, Quarries	174	5,404	5,578	994	5,538	6,532
Manufacturing	1,588	56,600	58,188	10,746	56,860	67,606
Construction	414	107,877	108,291	1,307	108,008	109,315
Transportation, Storage	368	36,580	36,948	2,327	36,677	39,004
Communication, Other	284	5,257	5,541	1,595	5,341	6,936
Wholesale	1,885	58,700	60,585	8,945	59,876	68,821
Retail	948	138,702	139,650	15,710	141,230	156,940
Finance, Insurance	241	30,212	30,453	1,962	30,399	32,361
Real Estate, Ins.	198	35,518	35,716	1,247	35,809	37,056
Business Services	599	96,572	97,171	2,632	96,822	99,054
Government services	149	6,016	6,165	971	6,117	7,088
Education Services	52	8,946	8,998	195	9,060	9,255
Health Services	297	72,411	72,708	1,919	72,866	74,785
Accommodation, Food	369	68,611	68,980	3,150	69,488	72,638
Other Services	479	109,282	109,761	2,658	109,925	112,583
All Industries	*8,119*	*905,604*	*913,723*	*56,614*	*912,995*	*969,609*

Source: Business Register Division Tabulations – September 1996

Establishments will be the standard tabulation unit within the UESP. However, the establishment may be replaced by other statistical or operating units in some instances. The establishment is the appropriate level for the collection of production characteristics. Use of the legal entity as a collection unit may result in reduced response burden when GIFI data become available. Regulated industries may require targeting of other non-standard units. It is a matter of weighing the advantages of consistently using one common statistical unit against those of meeting particular program requirements in a locally optimal way.

Goals

The Part I survey will attempt to collect a provincial and industrial allocation of value added. If the Part I survey is successful in this respect, Part II will be relieved of collecting some principal statistics for establishments. Eliminating data being collected in the "spreadsheet questionnaire" in the CCE could reduce the size of the schedules. If the spreadsheet questionnaire is not feasible, the Part II survey will collect principal statistics, commodity distributions, origin and destination information, production characteristics, and other industry specific data. These data would be used to allocate value added by province and industry. An objective of the survey is to improve the quality of estimates and make quality more equal across provinces. All data elements covered by existing business surveys will be included in the new survey with allowance for exceptions where justified.

There are many (about 90) annual business surveys with many more questionnaires, which now collect information from entities of complex enterprises. A variety of levels of the enterprise receive these questionnaires which may miss some segments of an enterprise's operation, may duplicate collection of some information and may use differing concepts and terminology. Integration of the data from these various surveys has not occurred to a satisfactory extent, nor have sampling or collection strategies been adequately coordinated.

The new census of complex enterprises offers an opportunity to coordinate the treatment of establishments for these complex enterprises in terms of sampling, the entity targeted, the questionnaire and data collection. The new survey will feature a coordinated sampling and collection strategy across all industries.

Relationship to survey of simple enterprises

Part I and Part III will produce an economy-wide estimate of value added. Part II and Part III will form a program to collect a wide range of establishment data by industry and province. All data collection for large and small enterprises will be integrated in a single three part collection vehicle in a coherent approach to business surveys.

Relationship to sub-annual surveys

Some areas have statistical programs with integrated annual and sub-annual surveys (e.g. P13 and Q5 surveys). The changes proposed to the annual survey programs will require a re-thinking of this survey approach and will require modifications to sub-annual surveys.

Timeliness

Annual data should be available within the year following the reference year. Timeliness has generally not been adequate for external users of production estimates and work has been underway across the BTS Field to improve timeliness over the past several years. The new strategy will strive to improve upon these gains. The timing of GIFI data is expected to improve relative to current tax data timeliness.

Sampling strategy

To improve the quality, completeness, and richness of the production statistics database, a census of the establishments of complex enterprises will be conducted. A distribution of the numbers and significance of output of the statistical units (company, establishment and location) in these enterprises by industry and province is required to determine if this approach will impose significant limitations on any individual industry statistical program. It is believed that this approach will not significantly increase response burden other than in industries not currently covered by STC surveys. The proposed approach has the disadvantage of increasing total response burden on these complex enterprises but has the advantage of producing a complete composite statistical profile of these enterprises. This is a feature that Statistics Canada has never been able to attain and which will permit the integration of all relevant data at the micro level.

An alternative approach would be to sample establishments within enterprises according to industry specific stratification thresholds. The sample would likely be very large because of the significance of these enterprises to most economic aggregates. This alternative reduces response burden on the enterprise but does not yield a complete statistical profile of these enterprises. This would complicate data collection and integration and might not result in a significant response burden reduction.

Whether the Survey of Establishments of Complex Enterprises should cover all or just a sample of establishments of complex enterprises will be decided after further analysis. However, the project goal is higher data quality and the large enterprises are the most important. They are better able to cope with the response burden than are small enterprises. For these reasons and on grounds of simplicity, the census approach is preferred at this stage.

Collection strategy

Data should be collected using standardized schedules for a common reporting period within each enterprise. The reporting period should be the fiscal year. These schedules will contain industry specific questions where necessary but will also be standardized and unduplicated with CCE and GIFI data. By maintaining contact information by establishment for these enterprises, the industry specific schedules can be customized to include only questions relevant to the specific reporting unit. Key Provider Managers will negotiate reporting arrangements for selected complex enterprises.

Work is required to develop standard variable definitions and concepts across industries. All current questionnaires will be reviewed to ensure that data elements are consistent and relevant. This will yield response burden reductions from the current situation in some cases. In others it will result in a significant increase in response burden, particularly for enterprises with establishments in industries not currently covered by annual production surveys.

Electronic questionnaires with the facility to import data from existing electronic files will be a reporting option for both Part I and Part II data. This feature has proved to be very popular with respondents in surveys with onerous reporting requirements. A collection system that goes a step further would be software that extracts the information we need from existing enterprise information systems. References to "questionnaire" in this paper should be taken to include paper and electronic versions.

We will develop a menu of collection options to satisfy a wide range of respondent capabilities and preferences. For the foreseeable future the menu will have to include mail-out, mail-back questionnaires; fax-out, fax-back questionnaires; and telephone collection as well as whatever electronic options appeal to respondents.

By maintaining contact and coverage information for each establishment, the schedules for the Part II survey could be delivered to any contact in the enterprise according to their preferences. The BR currently has the facility to keep track of contact information and questionnaire type for collection entities. Collection entities will be created to respect reporting arrangements specified by the enterprises. Some enterprises may wish to have all questionnaires (Part I and Part II) delivered to a single contact. This contact could complete them all or coordinate their completion by others at various points in the enterprise. Other enterprises may wish to have the schedules delivered to many different contacts.

In principle, with the new integrated questionnaire model, it should be possible eventually to re-arrange the distribution of questions across schedules, so that each contact is asked for only that data which that contact can provide.

A record keeping and reporting arrangements survey should be conducted to determine what the contact points are to be for the various schedules.

The Part II survey collection should be tested at the same time that the Part I collection is tested. Combining the two will be a more realistic test for both surveys.

Respondent relations

Respondent relations are a key element of any collection strategy. One aspect of this is feedback of survey results to respondents. Such feedback could involve extracts of survey data tailored to specific companies (details on their geographic region or industry). We have found that the best respondents are those that are also users, or at least can see tangible survey results. They appreciate the need for supplying the data and for its accuracy. Feedback could make data users out of more respondents. This is also an argument for producing data that are useful and pertinent for private sector users. Natural Resources Canada has been experimenting with providing respondents with data fact sheets and reference to Internet sites for detailed data as a means of motivating respondents. We should also build into the UESP a process of consulting with data suppliers and data users before we change either our data collection or our data outputs. Strategies for dealing with total non-responses of major respondents also need to be developed.

Processing

Data collection, capture, and at least the preliminary edit of all data will be the responsibility of the Enterprise Statistics Division. Processing will use common methods and systems. Generalized and/or existing processing systems will be reviewed to determine if they contain the required functionality and ease of use for this survey. Existing generalized software will be used, although some additional functionality is required and

improved software integration is essential. The complete data set of the survey will be much more comprehensive and complex than existing surveys, perhaps requiring multiple stages for edit and imputation with linkages for key variables only. For example, total sales may be imputed using other information pertaining to the enterprise, the commodity distribution of those sales may require information from other businesses for imputation, etc.

The quality of industry estimates produced by the subject matter areas will depend to a large extent on the processing of data from the large enterprises. Edit and verification of the data must be as comprehensive as possible to ensure the accuracy of these data.

BR feedback will originate from the profile questions in Part I, but may also arise from Part II survey activity. A BR feedback mechanism will be required, that provides timely update of the BR and does not conflict with feedback from sub-annual surveys. Survey operations staff should have access to the processing status of survey feedback. This will help to ensure that coverage and contacts of surveyed entities are up-to-date.

Analysis

Higher quality estimates will be achieved through improved and more consistent coverage of establishments and through more complete and intensive analysis by subject matter divisions. The complete database for each enterprise will permit a more thorough analysis.

Output

The output of this survey will be a data base containing all information collected (survey schedule or GIFI), calculated and/or imputed for establishments of complex enterprises. The database will also incorporate all data from the Part I and Part III surveys. The full set of data from all three components of the survey strategy as well as administrative data will be accessible to all subject matter divisions and analysts. Data from the databases will be integrated into the statistical programs of these divisions and integrated with output of any independent surveys. Analysis and dissemination will remain the responsibility of these areas. Having access to all available data for the enterprise and its establishments will give analysts more complete information and allow for more in depth analysis than is currently possible. In addition to the new data elements collected due to PIPES, new outputs will be developed on topics such as the performance of multi-industry versus single sector enterprises.

Chapter 11 contains a diagram of the proposed structure of the survey output databases.

Estimates derived from the survey will be for the calendar year. Since not all data will be collected on this basis, there is a requirement for calendarization techniques.

Implementation strategy

It is not possible to implement a survey of all establishments of complex enterprises by the spring of 1998, which covers all data requirements and all enterprises. A phased implementation will cover all data requirements of a subset of the complex enterprises. The subset will be those industries described in the overview as the 1998 industries. In industrial sectors where full implementation is not possible for the 1997 reference year, interim data improvements will be realized by scaling up existing survey samples. Linkage of NAICS-based estimates from the UESP to previous establishment estimates based on the 1980 SIC-E or SIC-C will be required for analytical and dissemination purposes. Standards Division will produce concordance tables to assist in the linkages.

Consultation with major national and provincial industry associations and regulatory bodies must begin in 1997 to inform them of our intentions and allow for their input to definitions, coverage and procedures.

The primary goal of the UESP is to supply to the Input-Output Division improved business survey data by province for 1997 on a NAICS basis so that Input-Output Division can estimate 1997 provincial Input-Output tables in the year 2000. The Survey of Establishments of Complex Enterprises will be vital to the Input-Output requirements due to the importance of these enterprises in most sectors of the economy.

Since 1997 is the first reference year of the new program and the first year for NAICS, not all elements of the UESP are feasible for 1997. The transition strategy consists of:

- for most industries, continuation of existing annual surveys with conversion to NAICS;
- for some industries such as retail, existing surveys will have expanded samples and more data elements;
- for a small number of previously unsurveyed industries (the "1998 industries"), a trial UESP will be done. A census of all complex enterprises (Part I) will be done along with Parts II and III of the UESP for only the 1998 industries.

Chapter

5

A Survey of Simple Enterprises

The Survey of Simple Enterprises (SSE) is the third of three survey vehicles comprising the Unified Enterprise Statistics Program. In many respects, it can be conceived as a blend of the other two survey vehicles — the Census of Complex Enterprises and the Survey of Establishments of Complex Enterprises — because simple enterprises contain both the management and administrative record-keeping and information capabilities of the former and the production and operational reporting capacity of the latter.

Survey strategy: an overview

Data for simple enterprises will be derived as much as possible from administrative sources; this reliance will be much greater than for complex enterprises because, by definition, simple enterprises operate in only province and industry and therefore administrative sources can be used directly, whether or not the simple enterprise is a single legal entity or multiple legal entity enterprise. Any consolidation problems for multiple legal entity enterprises are expected to be manageable.

Administrative data availability will be greatly enhanced by the implementation of the GIFI program by Revenue Canada (RC). The GIFI's provision of balance sheet and income statement information will permit Statistics Canada to greatly reduce the detail asked of simple enterprises. Businesses will be asked for only the most basic of financial data as "control variables" (i.e. as a means for evaluating the GIFI files as transmitted) or to 'benchmark' more detailed data reported by the simple enterprise (e.g. a respondent might be asked to report total revenue if subsequent questions ask for a breakdown of total revenue by distribution channel).

RC's implementation of the single Business Number will also represent an enhancement of the administrative data available to our program, insofar as it will facilitate the identification of businesses and the linkages between various taxation accounts.

The timing of this use of administrative data rests on RC's schedule for implementing the GIFI program and the Business Number. However, if the GIFI is not in place in time to meet our needs, additional data elements will not be added to the Survey of Small Enterprises. Significant resources would be required to temporarily modify questionnaires, methods and systems in response to a relatively short-term situation due to any delay in GIFI implementation. Instead, the tax estimates program will be modified to meet the needs of PIPES in the interim; this work must be undertaken reasonably quickly to allow the implementation of that portion of PIPES which proceeds for reference year 1997.

In addition to very limited detail on financial performance, simple enterprises will be asked to confirm profiling information (to determine whether they are still simple enterprises), and to report data elements currently collected and new data elements required by PIPES.

The PIPES requirements are outlined earlier in this paper; on the whole, these data elements (e.g. class of customer; destination, or origin and destination, of shipments; purchased inputs, etc.) will have an "across-the-board" applicability. PIPES-related content for industry-specific schedules will be determined on an industry-by-industry basis primarily by subject-matter divisions.

PIPES requirements must be integrated harmoniously with the "sector-specific" questions presently collected by the individual subject-matter divisions in support of internal and external data user needs (including cost-recovery survey questions). The remainder of this chapter will define and identify simple enterprises, outline the bureau's past experience with these enterprises regarding response burden, discuss the scope of the new systems requirements and examine administrative data availability.

Simple enterprises

Simple enterprises are defined as those that do not cross provincial or industrial activity boundaries. They therefore lack the "complexity" (at least in terms of organizational structure) of those enterprises which were the subject of the preceding two chapters.

Individually, there are many simple enterprises which have a significant impact on a given industry. Simple enterprises may be quite large in terms of revenue and employment. They may also be industry leaders in such aspects as innovation or sub-industrial specialization There are also simple enterprises with more than one establishment – that is, there are multi-establishment enterprises which respect provincial and industry classification boundaries, thereby remaining "simple" as defined in this paper. Issues surrounding these two types of relatively 'complex' small enterprises will be discussed later in this chapter.

In aggregate, simple enterprises represent a significant proportion of the total population of enterprises in each industry and in each province. The following tables show the number of simple enterprises on the Business Register in September 1996 in each SIC Industry Division and by province. Note that these counts of simple enterprises ignore non-employer businesses which are not currently on the BR.

Table 5.1 Number of Simple Enterprises by Industry Division, Canada, 1996

Industry Division	*Frequency*
Agriculture	52,403
Fishing and Trapping	5,776
Logging and Forestry	10,737
Mines and Quarries	5,404
Manufactures	56,600
Construction	107,877
Transportation and Storage	36,580
Communications	5,252
Wholesale Trade	58,700
Retail Trade	138,702
Finance and Insurance	30,217
Real Estate	35,518
Business Services	96,572
Government Services	6,016
Educational Services	8,946
Health Services	72,411
Accommodation, Food and Beverage	68,611
Other Services	109,282
Total, all Industry Divisions	*905,606*[1]

Table 5.2 Number of Simple Enterprises by Province, Canada, 1996

Province	*Frequency*
Newfoundland	16,644
Prince Edward Island	6,373
Nova Scotia	27,893
New Brunswick	24,927
Québec	226,434
Ontario	289,563
Manitoba	31,838
Saskatchewan	37,278
Alberta	102,246
British Columbia	139,185
Yukon Territory	1,458
Northwest Territories	1,767
Total, Canada	*905,606*

Source: Business Register Division Tabulations – September 1996

[1] Includes two enterprises not classified by industry division.

Simple enterprises that are not so simple

There are two types of simple enterprise that warrant particular attention: those that are especially large (in terms of revenue or some other indicator), and those that are somewhat complex (insofar as they contain more than one establishment).

Table 5.3 Number of Simple Enterprises with more than one establishment, by Industry Division, Canada, 1996

Industry Division	*Frequency*
Agriculture	25
Fishing and Trapping	6
Logging and Forestry	5
Mines and Quarries	78
Manufactures	189
Construction	107
Transportation and Storage	57
Communications	24
Wholesale Trade	662
Retail Trade	1,103
Finance and Insurance	48
Real Estate	130
Business Services	153
Government Services	39
Educational Services	386
Health Services	150
Accommodation, Food & Beverage	409
Other Services	268
Total, all Industry Divisions	*3,486*

Table 5.4 Number of Simple Enterprises with more than one establishment, by Province, Canada, 1996

Province	*Frequency*
Newfoundland	84
Prince Edward Island	32
Nova Scotia	95
New Brunswick	74
Québec	686
Ontario	1,154
Manitoba	103
Saskatchewan	189
Alberta	494
British Columbia	550
Yukon Territory	12
Northwest Territories	13
Total, Canada	*3,486*

Source: Business Register Division Tabulations – September 1996

One option is that simple enterprises with more than one establishment could be surveyed as a census, similar to that proposed for complex enterprises. Since multi-establishment simple enterprises have more potential to become complex, a census of these enterprises may be justified to help maintain accurate profiles. Large single establishment simple enterprises (those chosen in a "take-all" stratum) should also be periodically profiled.

Administrative sources may also be useful in identifying simple enterprises that become complex. The full implementation of the GIFI program by RC will play an important role in the early detection of any enterprise that has significantly increased its size and/or organizational complexity. Significant year-to-year changes in an enterprise's data from this detailed administrative source could serve to trigger a requirement to obtain profiling information from the enterprise in question. Discussions on the development of post-PIPES BR frame maintenance functions have included the delineation of a process to "graduate" an enterprise from simple to complex. The operational impacts (collection, imputation, estimation) of enterprises newly discovered to be complex are potentially great and will need careful consideration.

An important issue to be resolved is whether data need to be collected about each establishment of multi-establishment simple enterprises or only for the enterprise as a whole. Response burden will be lower if data are collected only for the enterprise. However there are issues to be resolved such as consolidation problems, reporting capabilities, and treatment of newly complex enterprises.

Response burden and the simple enterprise

As noted, some simple enterprises are quite large; the statistical system already imposes significant burden on such enterprises, often including them in "take-all" strata.

On the other hand, there are many simple enterprises at the other end of the scale that are rarely or never surveyed by Statistics Canada. Some fulfill the stereotypical definition of "mom-and-pop" businesses, such as those found in retail trade and several personal services sectors (e.g. barber and beauty salons). There are other industries containing a large number of relatively small (and simple) enterprises, including residential construction, residential real estate operators and professional groups like architects, consultants, accountants and bookkeepers and these do not currently carry a substantial response burden.

Data on small business are of analytical interest for many data users (internal and external) in isolation and in comparison with larger firms. It is often held that smaller enterprises are significantly different from larger firms; an important source of job growth, more innovative, etc.

The aggregate impact of these small enterprises on detailed data is relatively small (as suggested by the table at the beginning of this chapter). This is particularly true in larger provinces, given that the country's largest firms in virtually all industries are clustered in the largest population centres. Smaller firms are relatively more important in smaller provincial economies (suggesting a provincial skew in sampling rates) – an important factor for the UESP design.

The current situation

Currently, there are a variety of thresholds in place throughout the Bureau below which firms are excluded from survey activity. In the Services Division's regular business survey program, firms with annual revenues of less than $250,000 are excluded from survey samples. In IOFD, firms in some industries are excluded from reporting under CALURA if their annual revenues are less than $100,000. Similarly, capital expenditures surveys exclude firms on the basis of revenue thresholds.

Small simple businesses have less sophisticated record-keeping practices than larger firms. This is because their need to monitor detailed economic performance, in general, is not as great, given their relative operational and structural simplicity. They do not realize the same economies of scale in record-keeping -- and completing forms – that large enterprises can. There is no response burden measure readily available for simple enterprises since this is new to PIPES. We can, though, discuss the burden placed on the small business sector, assuming that this group roughly matches the simple enterprise grouping covered here. It has been estimated that small businesses (defined as those with fewer than 20 employees) spent roughly 203 thousand hours completing Statistics Canada survey forms in 1995. This represented a 10% decrease from the previous year.

To what degree will this response burden increase? The answer will depend on the demands made on these firms by the design of this survey, which in turn will depend on the information we require of them, our data quality targets, our capacity to devise new, simpler methods of obtaining required data and the availability of alternative sources.

Response burden implications of the new approach

A significant aspect of the new survey program will be the move away from thresholds and other practices which entirely exclude small enterprises from reporting. This reflects in part a recognition of the aggregate significance of small enterprises in certain provinces and industries; it also reflects the degree to which the new survey program is an attempt at a far more comprehensive understanding of the broader economy.

The Business Register is, at present, comprised of all firms which are employers (i.e. all those with Payroll Deduction Accounts); ultimately, non-employers will be added to the Business Register following the implementation of the single Business Number. By introducing non-employers which file under other tax accounts, (primarily the Goods and Services Tax) it is projected that about a million firms will be added. This will roughly double the number of records presently on the BR. Given that all businesses with an annual revenue of at least $30 thousand are required to maintain a GST remittance account, our surveying threshold could drop at least to that level across all industries.

The highest priority for additions of non-employers should be active BN accounts. This will add about 150,000 GST accounts remitting over the $30,000 threshold. Then, all other BN accounts regardless of remittance activity could be added, about 850,000 that are under the $30,000 threshold.

Potentially no firm on the Business Register larger than some very small size will be excluded from eligibility for sampling in the new survey program. To do so would introduce a downward bias into the estimates, and although it may be small at the Canada level, it will differ by province, and with a disproportionately large impact on smaller provinces. This implies some degree of new response burden on a sector of small business which previously had no burden — particularly the small non-employers who are yet to be added to the Business Register.

This is not to say that stratification strategies cannot be developed to reduce the burden on the smallest businesses. Short forms might be used for some respondents; that is, only a limited amount of content would be asked of very small enterprises, eliminating those elements deemed, to place an unrealistic burden on the record-keeping practices and resources of such enterprises. As well, the survey design could include a "take-none" stratum for very small enterprises with estimates for this stratum based entirely on administrative data and modelling for data elements not on administrative sources.

The strategy for estimating data used in the SEPH redesign will be applied more broadly. This strategy is to obtain detailed data elements from a relatively small sample and project these data elements over the entire population using administrative data and regression (or other modeling) methods. The GIFI source will be a stronger administrative source than ever before available. Modeling from small samples will be a strong factor in keeping response burden on small enterprises under control.

Survey activity in new sectors

The new survey program objective of better provincial data goes beyond the need to modify our present survey activity; it adds the requirement to develop programs for industries not presently being covered. Examples of affected industries include the construction sector and certain services incidental to various primary industries. These additional surveys will represent an addition to response burden.

Implementing the North American Industrial Classification System in reference year 1997 will include the development of programs for some industries separately and distinctly classified for the first time. The establishment of these programs will represent a small addition to response burden insofar as affected respondents will be asked a greater number of questions which will be more specific and focused on the firm's characteristics than questions asked in previous, more general surveys.

More comprehensive administrative data sources

The UESP will maximize Statistics Canada's use of administrative sources to maintain survey frames on the Business Register and to generate estimates of economic performance with less of a requirement for surveys. Appropriate sample stratification will help to maximize the efficiency of the survey samples and to minimize response burden on simple enterprises.

An integral part of the long-term overall strategy is that, for simple enterprises, balance sheet and income statement data from the GIFI will be sufficient and therefore not included on surveys intended for that target population. Survey activity for simple enterprises will focus on two requirements: first, data for understanding and calibrating data obtained through taxation files (like GIFI) and second, data required that are not available through administrative sources. The former refers to the collection of very basic

data for use in evaluating the quality of GIFI data; the latter includes those industry-specific elements presently collected by the regular survey program, including cost-recovery projects. While Revenue Canada is currently instituting the GIFI, not all firms will be filing under this procedure in time for our immediate needs. The Tax Estimates Program will be modified as necessary to meet the needs of PIPES in the interim.

Sector-specific data collection

For most subject-matter programs, the basic value added and employment content satisfies only the most fundamental collection requirements, such as of internal users (the System of National Accounts).

In addition to a wide variety of "sector-specific" standard data elements across business surveys, increasingly, Statistics Canada does cost-recovery work for external clients as a way to enhance and improve the information generated by regular survey programs. The content added to surveys has added greatly to the bureau's reputation.

It is essential to the overall strategy that this "sector-specific" content be reasonably incorporated as part of business survey strategy. Subject-matter staff should be involved in integrating their questions with the content of the new survey vehicles. One option might be to design the Survey of Small Enterprises so that there is a separate schedule or set of schedules for each industry. This would permit subject-matter divisions to customize questionnaire content to an extent while the more general objectives of the program are also met.

Some sector-specific questions are compatible with the objective of asking survey questions in order to validate taxation data. For example, a small enterprise classified as a software developer could be asked to provide "total sales of goods and services". This could then be checked against the amount they reported to Revenue Canada on the GIFI. The software developer could be asked to provide the percentage distribution of this amount by type of good or service commodity (i.e. custom programming under contract, consumer applications, business applications specific to certain industries, etc.), just as they are now asked in the Annual Survey of Software Development and Computer Services. Qualitative questions (for example "which marketing channels does this travel agency use?") could be appended to the schedule – again, much as we do now.

Some sector-specific schedules could become overwhelmingly long and difficult, if we are not careful. This redesign will offer the opportunity to modify or eliminate low-utility questions in the interests of limiting response burden.

Another option is to divide prohibitively long schedules into several shorter ones, and draw different samples for each one. There are disadvantages of this approach, in terms of thoroughness and analytical utility. There may be other ways of preventing excessive response burden through carefully crafted survey design approaches (see Chapter 6).

Mediating increased response burden

The GIFI for all businesses will be implemented for all legal entities by the end of the decade, and its implementation is fundamental to our approach for minimizing response burden. However, even with the advantages implicit in the availability of the GIFI, some individual simple enterprises selected in samples will have to report, under the new program, a greater amount of detail than at present, due to the need for more data than is currently collected.

Another factor increasing response burden for these firms will be the increased sampling rates necessitated by the goal to provide better, more equal quality provincial estimates. This will have a disproportionate effect on the smaller provincial economies (e.g. the need to improve provincial estimates will lead to higher sampling rates in Prince Edward Island than in Ontario). Determining the impact of this factor will require an estimation of the proposed sample sizes within the discussion of overall sample design for this project.

Creative means will need to be found to minimize burden on small enterprises. Examples include use of franchising links to avoid burden on franchisees, sampling strategies which spread burden thinly across the small enterprise population, rotation strategies, consultations with industry associations to find alternative sources of information, etc.

Building on our relationships with other federal departments, other levels of government and business and industry associations will be a logical starting point in exploring administrative data sources other than Revenue Canada. One example of a potentially useful external data source is the Multiple Listing Service for real estate agencies. There may also be opportunities for obtaining the data we need from sources other than our respondents. For example, can we obtain data on retailers' class of customer by conducting 'exit surveys' of their clientele? Could non-business surveys add questions? For example, could FAMEX incorporate questions on where household purchases were made?

Investigating these possibilities should be a priority. The effort to ask difficult questions in simpler ways (e.g. percentage distribution estimates, rather than 'hard' dollar values; developing qualitative questions to replace more difficult quantitative ones) applies to elements of Parts I and II of the survey strategy as well.

Electronic data collection methods offer particular opportunities for reducing response burden. Software bookkeeping application packages that can now be used for tax filing can append similar files permitting the user to complete Statistics Canada survey reports. Considering the advantages of electronic reporting for respondents and Statistics Canada, current efforts at developing practical means of electronic reporting will be intensified.

Chapter

6

Sample Design and Stratification Issues

Introduction

This chapter describes the basic elements of the sample design and estimation strategy for the Unified Enterprise Statistics Program. In keeping with the unified approach, we will implement an integrated sampling and estimation strategy for annual surveys for all industries covered by the UESP. While ambitious, this approach offers the best long-term opportunity for coordination of data collection, data integration, management of response burden and eventual cost-effectiveness of sample maintenance. The first section gives a brief overview of the integrated approach and how it will function. The second section presents the advantages of this approach, while the sections which follow go into further detail on some particular considerations that must be addressed during the development of this approach. The final section describes some components of a Data Quality Program for the UESP.

Overview of the integrated sampling and estimation approach

The main features of the integrated strategy are the following:

There will be only one integrated sample design for all industries covered by the UESP, covering both the complex and simple universes, which will incorporate the requirements of both the Census of Complex Enterprises and the Parts 2 and 3 surveys of establishments. This integrated design will consist of a number of samples, each corresponding to a questionnaire module.

The UESP samples will be selected using a single Survey Universe File (SUF) created at a specific point in time once per year. This SUF would be enterprise-based but would contain all the linkages to statistical units lower in the hierarchy.

The basic sampling unit will be the statistical enterprise. That is, whenever a statistical enterprise is selected, all of its component companies, establishments and locations would be automatically included.

The Census of Complex Enterprises and the Survey of Establishments Within Complex Enterprises will be conducted on a census basis. However the sampling and estimation strategies will be designed to be sufficiently flexible to permit sampling for both surveys in future years if budgetary or response burden concerns so dictate.

Common geographical and industrial classifications and size measures from the Business Register will be used for stratification purposes. Where additional stratification variables are available from subject matter sources and are required by the survey, these variables must be incorporated into the sample design by appending them to the SUF. In the longer term such variables should be added to the Business Register.

Standard sample design and estimation methodologies will be used across industries for all annual surveys. This will include standard definitions of target populations (e.g., in terms of reference periods for defining the in-scope population, geographical and industrial coverage, and size thresholds), standard stratification methods, standard sample selection methods, standard methods for treatment of births, deaths and stratum jumpers, and standard estimation methods, e.g. domain estimation, handling of duplicates on the frame, etc.

Coefficients of variation for key variables required by PIPES will be equalized across provinces, subject to reasonable constraints on response burden in the smaller provinces and territories.

The integrated sample design will be sufficiently flexible to permit separate samples to be selected for different sets of questionnaire modules directed to the same target population. Furthermore, the sample selection process will permit control over the amount of sample overlap among these various samples. This is described more completely later in this chapter.

The Generalized Sampling System (GSAM) will be used to select the annual samples and The Generalized Estimation System (GES) will be used for estimation.

The framework given above will be implemented in a single sample selection system which will be used to select the annual samples for all UESP industries. Within the overall framework, subject matter divisions, supported by their servicing methodology staff, will be free to design their particular samples to meet their industry-specific needs. They will be responsible for determining the particular domains of interest for their programs, and the associated stratification, sample sizes and sample allocations required. They will then specify their sample design parameters, for example in the form of sampling fractions, to the integrated sample selection system and the final set of samples will be selected.

A considerable amount of work will be required to define the details of this general framework, to build and test the systems required to implement it, and to redesign existing survey samples under the new approach. The complete framework and the sample selection system will not be in place until late 1998. Sample designs for the new surveys for reference year 1997, which must be done by the end of 1997, will therefore be implemented using "throwaway" systems. At the same time, however, these new surveys will try to respect as many of the principles of the integrated approach as possible, in order to learn from the 1997 experience.

Advantages of an integrated approach

At present, each survey designs and selects its own sample based on its individual requirements. There are wide variations in definitions of target populations, the frames used, the timing of frame creation from the BR, the sampling units, and the sample design and estimation methods. In comparison, the integrated approach has a number of advantages:

The use of a single source (the Business Register) and a single annual SUF has attractive benefits from the point of view of integrating data across industries. The use of a single frame for selection of all annual business surveys has apparently been quite successful at

the Australian Bureau of Statistics. The single SUF approach does not imply that only those enterprises alive at a certain point in time are included. The single SUF can include enterprises which were alive at any time during a certain time interval (e.g., six months) during the reference year. Unless there are good reasons to the contrary, however, the definition of the interval should be common across industries.

The use of the enterprise as the common sampling unit facilitates the coordination of data collection for enterprises and establishments within multi-establishment enterprises, both complex ones and simple ones. More importantly, it facilitates the integration of enterprise-level and establishment-level data, since all of the data for a given enterprise are collected.

Having common sample design and estimation approaches for the annual samples will make integration of micro-data for different industries much more feasible.

Having only one sample design to maintain will make it easier to respond to external changes. During the next few years there will be enormous changes in the availability and quality of administrative data sources, both in terms of variables available on the Business Register for stratification, sample selection, and estimation, as well as sources of administrative data which will replace direct data collection. This will require continuing changes to the sample design and estimation procedures to adapt to these developments, as has been the case with SEPH over the past five years. Such changes will be much easier to make to an integrated design and a single sample selection system.

An integrated design will address the problem of undercoverage resulting from incorrect SIC codes on the frame. Under current procedures, if a sampled unit is found during collection to be out of scope for the survey due to an error in the frame SIC, it is simply dropped from the sample. Such units also have no chance of being sampled under their correct SIC, so they have no chance of contributing to any survey estimates. With an integrated approach, one can address this problem, for example by redirecting such units to the correct data collection instrument, or by making adjustments at the estimation stage. At a minimum, one will be able to produce a comprehensive measure of undercoverage due to SIC errors.

While the resulting sample design might not be optimal for any specific industry, due to the need to design the sample within a specific framework, the overall approach will be more cost-effective for Statistics Canada, since it allows concentration of our methodology resources on the development and maintenance of a single design for annual surveys. In fact, given the number of surveys to be redesigned and the number of new surveys to be designed under PIPES, this approach is probably the only way for BSMD to cope with the demands on its experienced staff.

A specific framework for sample design and estimation, embodied in concrete systems, will help maintain the integrity of the sampling and estimation methodologies for annual business surveys.

Some specific issue that need to be addressed during the development of such an integrated sampling and estimation approach are considered next.

Quality of the Business Register

The integrated sampling approach requires a Business Register of high and consistent quality across industries, in terms of its coverage and other information needed for sample design, most importantly industry codes, geographic codes, and size measures. The quality of the frame will be a major determinant of both the quality of the final survey results and the costs of carrying out surveys. Plans for improving the Business Register under PIPES are described elsewhere in this document. At the same time, however, it must be recognized that no frame will ever be perfect, and the sample design and estimation methodology must be therefore be designed to deal with frame defects.

A regular program to provide quantitative measures of the quality of the Business Register is also essential. Such a program already exists, but PIPES will necessitate improvements and extensions of its scope. The types of measures required include errors in SIC and geographic codes, duplication on the frame due to imperfect linkages among the various administrative sources used to create and maintain the Register, measures of undercoverage, and measures of the accuracy of the size measures used for stratification. Such measures could be used to improve estimates through corrections at the estimation stage, to provide information on the amount of error to users, and to identify areas where improvements are needed. This measurement program would be one component of the overall UESP Data Quality Project.

Sampling of Complex Enterprises

As mentioned above, both the Census of Complex Enterprises (Part 1 Survey) and the Survey of Establishments Within Complex Enterprises (Part 2 Survey) will be censuses. The target population for Part 1 is all statistical enterprises with establishments in more than one province and/or more than one industrial classification, at the most detailed level (currently four digits under the 1980 SIC-E).

However there will be some "small" complex enterprises, such as a business with one establishment in Ottawa and another in Hull. Including such enterprises and their establishments in an annual census will pose a response burden problem. Therefore the complex enterprises to be included in the Part 1 survey on a census basis will have to exceed a certain size threshold, which is yet to be determined. Complex enterprises below the threshold will be subject to sampling. A study will be required to determine what the size thresholds should be and whether they should vary by province and industry. It should be noted, however, that with the current Business Register rules these enterprises would have to meet the IP revenue thresholds, which vary by province and industry, to have any chance of being identified as complex.

In the case of the Part 2 survey, there are approximately 56,000 establishments within complex enterprises. The plan is to include all of them, but if in future years this becomes impossible, either for response burden or for financial reasons, the sampling and estimation methodology will be flexible enough to permit sampling them. However any sampling plan will be based on the enterprise, so that if an enterprise is selected, all of its establishments will also be selected.

Note that if sampling of complex enterprises is done for both the Part 1 survey and the Part 2 survey, these samples need not be the same. That is, one could have one sample of complex enterprises for Part 1 and a different sample of enterprises for Part 2. The Part 2 sample could be either a subsample of the Part 1 sample or partially overlap with it.

Sampling of Simple Enterprises

The Part 3 survey will be a sample survey of the simple enterprises, i.e., those that operate in only one province and industry. This group can include quite large enterprises, as well as enterprises with more than one establishment. According to the Business Register, there are approximately 3,500 such multi-establishment simple enterprises, covering some 11,000 establishments.

The Part 3 survey may serve a number of objectives, all of which may imply different sample designs and different sample sizes. These potential objectives include:

1. The production of industry-specific detail, both financial and non-financial;

2. Validation and calibration of the GIFI data;

3. Building models to be used to estimate for other characteristics not available from GIFI data but which are correlated with it, similar to the SEPH approach; and

4. Producing preliminary estimates of establishment-level financial data in advance of the GIFI data.

The objectives for the Part 3 survey are likely to change and evolve as the availability and linkability of administrative data evolves over the next several years. For example, objective 4 may be important until the GIFI data become available on a timely and reliable basis, but may no longer be required once the GIFI data program is fully in place. A key factor in the design will be to ensure the flexibility to respond quickly to such changes.

These different objectives and data needs may imply very different sample designs and sample sizes. For example, the industry-specific needs for commodity detail may require a different design, and much larger sample sizes, than required for establishment-level financial data. At the same time, however, response burden concerns dictate that we not impose any more burden on small businesses than necessary. Therefore, the integrated sampling strategy will provide for individual but coordinated samples to be selected for different groups of variables (in the form of questionnaire modules) sent to the same target population. The sampling system will be designed to permit whatever degree of overlap is desired among these different samples, including completely nested sub-samples, partially overlapping samples, or completely separate samples. The degree of overlap will depend on considerations of data integration and response burden. In practice, of course, there will always be some degree of overlap for the largest enterprises, since these will be take-all units under virtually any sample design. It might be noted that experience with surveys such as P13 and the Retail Commodity Survey has illustrated the need to build in the flexibility to select different samples from the same frame to meet different objectives. This strategy is analogous to the use of the Labour Force Survey and the General Social Survey in the Social, Institutions and Labour Statistics Field. These sampling frames for these vehicles are used as "platforms" for selecting samples for a wide variety of different surveys.

Stratification will be, at a minimum, by Industry, Province and Size. The industrial classification used will be the North American Industrial Classification System. Work will be required by subject matter divisions, together with the System of National Accounts Branch, to define the exact set of domains for which estimates are required, in order to establish the industry strata to be used in stratification. The implementation schedule for NAICS will also have to be taken into account. Current plans are for a "Mass Conversion" of the Business Register to a NAICS/Business Number basis at the beginning of July 1997, followed by a "Nature of Business" survey over the following eighteen months to improve the NAICS codes for these units which were not converted manually by subject matter or which were not converted automatically as one-to-one conversions. The quality of the NAICS codes at the time of sample selection will be a consideration, and it may be necessary to regroup some NAICS codes which are individually unreliable but are reliable in combination.

Stratification by province is routinely done for most annual surveys. In some case, such as retail, geographic stratification also extends to sub-provincial areas, such as the largest CMAs. In some cases, there will be some simple enterprises containing establishments in more than one sub-provincial area. A simple way of handling this would be to make all 3,500 multi-establishment enterprises (with their 11,000 establishments) take-all units for the Part 3 survey. These enterprises are likely to be large anyway, and most would likely be take-alls on the basis of size alone. They are also likely to be the enterprises that have the greatest chance of becoming complex enterprises in the future. The possibility of making all such enterprises take-alls will be studied, including response burden considerations, and a final decision taken during the sample design process.

Within each industry and province cell, stratification by size will be used to reduce the variance of the estimates. Generally there will be a "take-all" stratum of the largest units, with two to four "take-some" strata of the medium and smaller sized units. Statistical methods exist for determining optimum stratum boundaries, so this does not need to be specified by subject matter divisions. These boundaries depend only on the desired reliability (coefficient of variation) for that particular province and industry, the sample allocation method being used, and certain characteristics of the population being sampled.

As well as defining take-all units on the basis of (a) multi-establishment enterprises and (b) enterprise size, there may be special considerations which dictate that certain enterprises should be take-alls. For example, in the trucking industry arrangements exist with the trucking association to supply Statistics Canada with the data for all of their members. Since the data are provided at no cost to Statistics Canada (other than processing costs), it makes sense to include all such enterprises as take-alls.

At the other end of the size scale, there will be many very small businesses who may have limited capacity and willingness to respond to our surveys. This will especially be the case when small non-employer businesses are added to the Business Register. It may be necessary to develop methods to identify a "take-none" stratum, below which we would not survey, but for which we would estimate using other sources such as tax data. Statistical methods could be developed to identify such optimum lower thresholds on the basis of size. However this strategy needs careful study, due to the potential for introducing an unknown bias into the estimates. This bias may vary among provinces and, unlike sampling variance, it cannot be measured directly from the sample.

Further work is also required to define in more detail the coefficients of variation (cvs) required for the establishment-level estimates from the combined Part 2 and Part 3 surveys. The same considerations will apply to the combined Part 1/Tax Data estimates for enterprise-level data. While the latter may eventually be available on a census basis, in the short-term estimates will be based on the Tax Estimates Program, which uses a sample. The possibility of sampling of complex enterprises was also mentioned above.

The objective of PIPES is to move to a situation where provincial data used for allocation of the HST are of equal reliability. If by this it is meant that the percentage error in the amount distributed to the provinces should be the same, regardless of the size of the province, then this implies equal coefficients of variation for each province. However further work is required to interpret this general principle in sufficient detail to allocate the sample by province and industry. First, the key characteristics for which equal cvs are required must be defined. In addition, we need to define the level of industrial detail, or other disaggregations, at which equal cvs are required. Specifications from the System of National Accounts and subject matter divisions will be required.

The actual level of the coefficients of variation that will be required will be constrained by financial limitations on the sample sizes and by respondent burden considerations. However an absolute requirement for lower cvs in the smaller provinces will almost certainly require a census of all enterprises for some industries. It will be necessary to modify the requirement for equal cvs to reflect practical considerations of response burden.

It is important to note that the sample design for the Part 3 industries must take into account the establishments covered on a census basis by the Part 2 survey, since the coefficients of variation will depend on the combined results from both Parts. The decision to conduct Part 2 on a census basis, while having the advantage of simplicity, may result in a larger total sample size, and thus greater respondent burden, than would be required if sampling of the Part 2 establishments were permitted. While at this point the increase in sample size and response cannot be quantified, it is worth noting that the Part 2 establishments are concentrated in manufacturing, retail and wholesale, and for these sectors virtually all establishments within complex enterprises are take-alls. Nevertheless, it is prudent to build in the flexibility to permit sampling of complex enterprises at some point in the future if this should turn out to be desirable.

The UESP sample will be strictly cross-sectional in nature. There will be no attempt to control the overlap of samples from one year to the next. Naturally, units which are large will tend to be in the sample every year, as will the complex units. Stratification and sample selection would be carried out independently each year. This has the advantage of using the most recent information available to select the sample. The disadvantage is that there is no control over the distribution of response burden over the smaller enterprises.

Use of Administrative Data in Estimation

With the advent of GIFI and GST data in electronic form on a census basis, Statistics Canada will soon have a wealth of financial data available on Canadian businesses. At the same time, however, PIPES requires greatly improved provincial data for many other variables not available from GIFI or GST. In order to meet these latter demands while also minimizing the increase in response burden, it will be vital to exploit the administrative data as auxiliary data in the estimation process. This is in addition to its role as the sole source of financial data for simple enterprises.

One method for doing so has been successfully developed by the Survey of Employment, Payrolls and Hours. Statistical models are used which exploit the correlations between variables available from the administrative source (PD-7 in the case of SEPH) and the other variables of interest. The statistical models can be built using relatively small samples, certainly much smaller than would be required if estimates had to be based on survey data alone. A key component of the estimation strategy for the UESP will be to utilize such estimation techniques wherever possible.

The use of such techniques will not be immediate, however. GIFI data will not be available until 1998, and experience with other administrative data suggests that the administrative data have to be treated before being suitable for use in estimation. As well, the potential gains that can be achieved by using administrative data must be established quantitatively so that an appropriate sample design for model-building purposes can be developed. There also must be a clear link between the survey units, which are based on the Business Register, and the administrative data units. Such a link will exist for corporations via the Business Number, but not for T1 data. Methods for using the T1 data in estimation must be researched and developed. Finally, there will some variables required by PIPES, perhaps even the majority, where this approach does not provide any benefits. For such variables the estimates will continue to be based on survey data alone.

Data Quality

This chapter has described the main elements of the sample design and estimation strategy for the UESP, which are designed to provide estimates with the low possible levels of sampling error within cost, response burden and operational constraints. However sampling error is only one source of error in statistical estimates. Other sources include coverage errors, non-response errors, response errors and processing errors. Experience has shown that these errors are often more important to the quality of the data than sampling errors.

Because of the important use to which the data from the UESP will be put, there is a need for a comprehensive Data Quality program to be put in place. Such a program can be thought of as having three components which, while interrelated, can be distinguished from each other on the basis of their objectives, who carries them out, when in the survey process they are done, and how they are done. These three components are Quality Assurance, Certification, and Data Quality Measurement. A brief description of each of these follows.

Quality Assurance

The principal goal of quality assurance is to ensure that specific measures are implemented during both the design and implementation stages in order to minimize the level of error in the final results, given budgetary and other constraints. Here, the emphasis is on prevention and control of errors. Starting with a comprehensive review of the entire statistical process, from formulation of objectives to final output, the various stages of the process at which errors can occur are identified, and steps are taken to reduce these errors. During the development of the questionnaire, for example, testing of concepts and wording with potential respondents is advisable (and is in fact Bureau policy) to reduce errors due to misunderstanding of concepts or inability to respond as requested. During the implementation, statistical quality control of operations such as data capture and coding is used to control errors. Immediate feedback to operations personnel is important

to improve processes while they are still underway. Follow-up of non-respondents and editing of data are other simple examples of quality assurance measures. Typically, such measures would be designed jointly by methodology, survey management, subject matter and operations personnel.

Certification

Despite the best quality assurance program, however, the data from any survey will never be perfect. Therefore it is important that the data be reviewed to determine their quality prior to being disseminated to the public. In most cases, such a review has to take place under very tight time constraints, between the day when the data first come off the database to the date of release in The Daily. Thus, the objective here is not to conduct a comprehensive audit of data quality, but to identify any gross errors in the data and to take appropriate action. Such action can range from including warnings to data users about data quality problems to a decision not to release the data. Guidelines on this topic are available in a draft policy on data quality criteria in the release of statistical data.

Typically, certification is primarily a subject-matter activity, due to their familiarity with the data. Methods typically used for certification of data include comparisons to other related sources of data, reviewing non-response rates, coefficients of variation and other available measures of quality, and examining top contributors to survey estimates. Methodology also plays a role by suggesting approaches and advising on statistical techniques available to explore the data. Experience has also shown the value of sharing data certification practices across different subject matter domains and of developing standard tools and methodologies (e.g., consistently defined non-response rates) for carrying out this function.

Data Quality Measurement Studies

The Certification process, because it is carried out under strict time constraints, is typically based on incomplete information and depends heavily on subject matter knowledge. In most major survey programs there is also a need for more comprehensive, and quantitative, measures of specific sources of error. These sources of error typically include rates of undercoverage and overcoverage, measures of non-response and response errors, measures of errors introduced during coding, data capture and imputation processes, and measures of sampling error. In most cases these require specific measurement studies to be built into the survey process. However to avoid biasing the results it is also important that these studies retain a certain independence and objectivity in their design and execution.

The results from such studies have a number of uses, including identifying problem areas which must be improved in the next round of the survey, informing users of the quality of data (a requirement under Statistics Canada policy) and, if the results are available in time, for actually modifying the estimates from the survey. An example of the latter is the use of estimates of net undercoverage in the Census of Population to adjust the Population Estimates Program. Because these studies are usually fairly technical in nature, they are usually designed and carried out by methodologists.

Interrelationships of the Components of a Data Quality Program

The three components of Quality Assurance, Certification and Data Quality Measurement, while having distinct objectives, timings and participants, are of course inter-related. For example, a problem found in a Data Quality Measurement study may result in the addition of a new Quality Assurance operation in the next iteration of a survey. Similarly, a problem noted during the Certification may trigger a special Data Quality Measurement study designed to provide quantification of the problem and an explanation of its cause. Another example is the use of Data Quality Measurement results that can be produced quickly, such as response rates, in the Certification step. These interelationships make it essential that the three components be planned and executed in a coordinated fashion.

Chapter 7

Characteristics of the Improved Business Register

This chapter deals with the considerations and issues related to the business register to support the requirements of PIPES. Specifically, it deals with the requirements of the business register to support the survey program changes arising from the Unified Enterprise Survey Program (UESP).

The requirement for a single business register

The goal of having all economic survey programs fully "hooked" to a single business register (BR) has been a long-term goal of Statistics Canada. The BR is the essential tool to ensure complete and unduplicated coverage of businesses.

Population coverage

The current business register is a list of all employer businesses in Canada. Businesses that have no employees such as some unincorporated farms, some corner stores and small owner-operated service businesses are not on the business register today.

The business register's coverage will be augmented in July 1997, coincident with the implementation of the Business Number (BN), to include non-employers having a BN account. About 150,000 non-employer businesses having revenue over the $30,000 GST threshold will be added in July with the remaining non-employer GST registrants to follow at a later date.

An improved business register should contain a complete and unduplicated list of all complex enterprises with information about their legal and operating structures. The goal is to enable the proper delineation of business entities into component statistical entities necessary to enable the efficient conduct of economic business surveys. We need to systematically identify multi-provincial enterprises, including multi-province/same-SIC and multi-province/multi-SIC units, as well as provincially significant businesses. The enterprise register should also contain government business enterprises, government departments, agencies and institutions and not for profit entities.

Other frame sources used by individual surveys will have to be reconciled with the Business Register. For example, mining entities, which are presently obtained from Natural Resources Canada, and cultural entities required by Services Division's joint surveys with Education, Culture and Tourism Division will have to be reconciled Extensions to BR coverage that will be considered are the inclusion of all GST accounts below $30,000 (700K units) as well as self-employed tax filers that do not have a Business Number yet report T1 income from farming, fishing and trapping, commissions, or professional activities. Extending the BR to cover these smaller units could be important for accurate measurement in certain smaller provincial economies. The character of the population of these small units will not be known until after further analysis is completed.

Besides private for-profit businesses, the current business register also includes government business enterprises, government agencies, departments and institutions and not for profit entities. It is intended to separately flag the not for profit entities, to increase the coverage of government entities such as municipalities, and to improve the timeliness of their updating on the register.

The business register does not cover some types of units that are currently used in some surveys. An example is the "mutual fund" unit, which is used for the IOFD quarterly survey of financial statistics for enterprises. Such financial funds will be added.

Business Register Boundaries

Under the UESP, complex enterprises will be surveyed on a census basis and the simple firms on a sample basis. The BR's current IP/NIP threshold design for processing units manually (IP) or in batch (NIP) will no longer be required, and the BR design can move to the maintenance of one continuous frame list. The data model will no longer depend on thresholds, but rather on organizational complexity, and will support the two basic UESP sub-populations, i.e. the complex and simple enterprises. Frame considerations for the Part I, II and III surveys will be addressed in the generation of the associated master survey universe files.

Primary attributes of the Business Register

Currently the business register classifies all establishments and locations to the 1980 SIC-E at the four-digit level. The company and enterprise units are classified to the 1980 SIC-C at the four-digit level. The full geographical code is applied to establishments and locations only.

There are a number of potential frame requirements raised by the PIPES program which need to be investigated to determine the best way of dealing with them. The major issues include (1) industrial coding of the "stock" of businesses existing prior to BN conversion, (2) coding of non-BN accounts (full 6-digit NAICS or something less), and (3) new stratification variables such as commodity information, exporter/importer status, specialization levels, foreign ownership and origin and destination characteristics. The size of the residual mentioned in (1) will not be known until July 1997, when the BR completes the process of shifting to the BN as the single business identifier. The non-BN-registered businesses mentioned in (2) will be largely related to T1 business income tax filers who have not opened a BN account. If a linkage between the BN and the Social Insurance Number is not established, this will be a difficult population to delineate.

When the need for new stratification variables is clearly established, the BR should be able to handle them within its data model. However, the extent of the systems changes that may be required to accommodate and maintain these additional variables suggests the investigation of alternatives. The question has been raised whether all of the stratification variables need to be resident on the BR itself. Instead, they might be located on other databases, tightly linked to the BR. Certainly the systems technology exists to allow the implementation of this strategy. Whether it would result in any efficiencies in terms of processing, flexibility or ease of access is an issue to be investigated.

Importance of timely updating frame information (such as GBI)

It will be important to process the new sources of tax-based data into the business register in a timely fashion. Gross Business Income (GBI) is presently updated on the basis of a ratio estimation model. Ratios of revenue (from T2 sources) to wages and salaries (from T4 sources) by industry are used to derive GBI from remittances data on the PD files. GIFI forms in future will directly provide actual tax data for the GBI variable for all simple enterprises. For establishments within complex businesses, the census of complex enterprises data can be used. For establishments of new complex units not yet surveyed or not responding, GIFI data could be used in connection with GST data for allocation to establishments. Similarly, consideration might be given to replacing the current employment size stratification variable with PD-7 employment data, now available and already linkable to the enterprise units via the BN.

The current BR design places considerable emphasis on maintaining the structural integrity of the information on the database, particularly across a multiplicity of time periods and this must continue to be a key element of the new information model. However, the application of the "high water mark" will be reviewed with a view to speeding up the processing of updates to the frame.

Reconciliation of the BR with the IOFD and Agriculture Division frames

The IOFD-BRD reconciliation initiative is crucial to the strategy of interlacing the current quarterly survey of financial statistics for enterprises within a new enterprise survey and to the conversion of the corporate tax data series to NAICS. Given the presence of the BN linkage and past reconciliation efforts, the reconciliation of enterprises with only one legal entity should be achieved with relative ease. For the remaining 1100 or so multi-legal-entity enterprises, reconciliation must be done on a case-by-case basis by the members of the newly constituted Reconciliation of Statistical Enterprises Project (RSEP) group. Once the reconciliation is established resources will be needed to maintain the IOFD/BR links.

Linking the existing IOFD quarterly survey to the business register is not expected to imply major changes to the coverage of the business population, although a number of SIC code re-alignments are likely to be identified. The linkage should result in an improvement in the identification of business entities made up of consolidations. With the necessary feedback mechanisms in place, another gain will be more current legal and operating structures for the businesses in-sample for the quarterly survey.

The reconciliation of the farm register to the BR should be greatly facilitated by the availability of GST records. At present, many farm units are operated without paid employees and therefore are not on the current BR. Most farmers do appear to have a GST account and therefore will be included in the BN registration process. Indications of sales activity through the GST records will result in the inclusion of an operating farm unit on the register. Preliminary analysis has indicated that for farms with significant activity there is close mapping of GST and employer account businesses with the farm register. The reconciliation will be greatly assisted if the next census of agriculture is able to collect GST or BN identifiers directly from census respondents.

Levels of statistical units

The UESP strategy sees an evolution toward the enterprise and establishment (possibly with a modified definition) as the primary statistical units for the majority of business surveys. The business register currently maintains four levels of statistical units — statistical enterprises, statistical companies, establishments and locations. We need to re-examine the practical definitions used for the statistical units, and consider whether any changes are required. We need to determine whether we need four levels of statistical units. We also need to consider whether supplementary units such as regulatory units and financial funds should be introduced into the register for specialized purposes. Many surveys now use a combination of establishments and locations, some use units set by regulators, few use statistical companies There also appears to be a growing need for an expanded list of activities undertaken by enterprises not covered by the standard model. Examples of these are private trucking fleets, in-house training facilities, computer services, and research and development activity within establishments. These issues need further study before they can be resolved.

Linking operational entities to legal entities

The current business register links legal entities to the business entity but not to their corresponding establishments. For most businesses in the simple population the relationship is one legal to one operating entity. For the case where there is one legal entity but multiple operating entities, these links are not currently maintained by the BR. The introduction of the BN as the single business identifier will facilitate this kind of linkage in the future. However, there are about 1,100 business entities on the register that comprise more than one legal entity. In these cases the legal entities are only linked to the top business entity (the so-called "top op"). There is no attempt to link these legal entities to the establishments in the operating structure.

Linking the establishments of the multi-legal-entity enterprises directly to their legal entities is an area worth further investigation. Such linkages could improve the quality of business statistics by providing more information. Where establishments are not responding to the survey, temporarily or otherwise, the linkages would facilitate imputation based on tax data. Even more importantly, the links could be used to substitute tax-based data for survey questions, thereby reducing the number of questions that need to be asked to all but the unlinked establishments.

The results were not promising when a study was conducted ten years ago on the possibility of linking establishments with legal entities in multi-legal entity enterprises. However, the anticipated availability of GIFI data has dramatically increased the potential for integrating survey data with tax data. Further the planned use of the BN as the single business identifier for both Revenue Canada's tax files and Statistics Canada's business register has altered the landscape considerably. A fresh study is needed, perhaps as part of the Reconciliation of Statistical Enterprises Project.

Two other linkage issues also need further consideration: the Franchiser/Franchisee link and the possibility of a link between the Canadian enterprise reflected on the BR and the associated CALURA global enterprise. Reflection of the latter linkage would extend the current register practice of including only one level of foreign parent to indicating all significant foreign ownership links.

Interfaces with surveys

Currently the business register provides a single frame to business surveys by providing the complete population of statistical entities which represent the universe for the survey. This file of statistical entities is used to conduct sampling. Also the register maintains a historical file of all units sampled for a survey along with the collection entities for the in-sample units and a file with data to support the mail-out for the survey. The register also records the log-out and log-in of questionnaires needed for the company contact system and controls the automated survey feedback process.

Most major business surveys provide feedback to the business register on updates to frame information, and this is largely carried out in a batch mode. For some surveys, the current data collection software will not support the automated transfer of frame updates to the register. Resorting to a variety of manual interfaces introduces a number of complexities, all prone to delays and errors. The UESP initiative will have to concentrate on resolving this situation to ensure that accurate and timely updates are made in all situations.

Additional potential issues to be addressed are: (1) the continuity of collection entities for all surveys linked to the BR, (2) the need for additional data elements pertaining to the collection entity and (3) the possibility of determining confidentiality restrictions globally within the BR versus the current survey by survey practice.

Time travel

The current business register was developed to support "time travel". This means that every frame change that occurs within a business entity is date stamped with an effective date (when the event occurred), a processing date (when the update was made to the register) and a committed date (the date the update was made viewable to the user of the business register).

The important concept of time travel means we must process all changes to a business in chronological order. While it appears there may be some events which could be processed out of order, the processing of changes out of sequence can jeopardize database integrity. An example would be "deathing" a legal entity before dispersing the operating units (liquidation or selling the assets). The "high water mark" logic which demands the processing of events in chronological order makes the updating process particularly unwieldy for complex enterprises. An investigation is needed to determine the best possible compromise in processes and systems changes to preserve the integrity of the register while bringing improvements in the timeliness of updates and a reduction in the manual workload of maintaining the register.

Systems

The existing informatics system supporting the BR was designed about a decade ago. A need has been expressed to permit users of the BR to operate more effectively within an NT client-server environment, with friendlier "point and click" features. The register could then be updated or queried from desktop PCs. Such a system could produce more timely results with fewer errors. Survey managers should be able to browse the BR easily, or to produce customized analytical demographic reports simply.

Because of time and resource constraints, the existing mainframe system for the BR will likely be used with minor enhancements for at least the next two years. However, consideration will be given to creating a new "windows-based" front end for access to the existing mainframe system, to gain some of the benefits described above.

Timetable

NAICS and the BN are to be implemented in July 1997. A preliminary universe file is to be generated to begin testing the sampling strategy for the Unified Enterprise Statistics Program in October 1997. A production version of the universe file must be prepared by early January 1998 for industries to be covered by the new survey approach for reference period 1997. The survey feedback system should be ready for operation by the target mail-out date in mid-March 1998.

Summary profiling information from the quarterly survey of enterprises

The IOFD quarterly survey already collects legal entity profile information from enterprises along with financial statement information. Respondents are asked to confirm or update a list of legal entities, sent with the survey questionnaire. Consideration should be given to asking respondents to update an added list or diagrammatic representation of operating entities. The availability of such information could lead to more timely updating of the Business Register.

This step would not be a significant increase in response burden. The respondent would be providing information only about changes, and when they are fresh in mind. These changes coincide with changes in the financial statements. When legal entities are sold or acquired, corresponding changes may occur in the operating profile. Such information collected for all 8,000 complex enterprises would provide more frequent updates of the business register. For the 1,100 multi-legal-entity enterprises it would facilitate better linkages between establishments and legal entities when corporate reorganizations occur and could provide timely identification of new enterprises that are carved out of existing ones. The workload of updating the profiles for both STC and the respondent would be spread out over the year. Promptly incorporating quarterly updates to the BR will require significant changes in the updating process.

Concerns

The concept of *one and only one* business register database is fundamental to the strategy described in this paper. Response burden cannot be easily monitored and managed, for example, unless all survey activity is channeled through the same register. We must avoid encouraging or facilitating the development of a multiplicity of partial frame data bases — for example, frames of large or small enterprises, tax-based frames, frames for particular industries, frames of legal entities or a special frame for enterprises managed by key provider managers.

For the business register program to be in a position to adequately service expected PIPES requirements it is essential that NAICS and BN implementation be given high priority.

In addition a survey "hook-up" to the register may appear to be an onerous task to a survey manager and his staff. The transition to a fully "hooked" status is often perceived as difficult given a divergence in concepts employed by a particular survey and the register "standards." More often the greater problem is that the transition to the register takes considerable time in order to provide staff adequate time to become knowledgeable about the register. Perhaps the transition period for a survey staff might be eased by the design of user friendly register interfaces but it should also be noted that a substantive investment will have to be directed to staff training. Experience has underscored that this kind of training is critical to the successful "hook-up" of a survey program to the register.

Chapter

8

The Integrated Questionnaire

The purpose of this chapter is to describe the conceptual model of the integrated questionnaire, and to present an outline of the activities needed in the next two years to bring this vision to reality in time to meet the target dates of the Unified Enterprise Statistics Program (UESP).

There are more than 200 business surveys conducted by Statistics Canada, with more than 600 questionnaires. Roughly half of the surveys and more than two thirds of the questionnaires are annual. Several mechanisms now in use help us integrate questionnaires, beginning with the basic economic concepts outlined in the System of National Accounts. The BR provides the common frame and identifiers for integrating data about the same enterprise from different sources. There are classification standards such as the Standard Classification of Goods providing frameworks for questionnaire content. There are also several conceptual models, including the production, financial and employment models, which are intended to guide the development of surveys and questionnaires. Finally, questionnaire design principles have evolved over time which provide guidance for questionnaire development.

However, our questionnaire set remains far from integrated in a comprehensive sense, and we could do much more to bring together the concepts, definitions and questions on our surveys. The Unified Enterprise Statistics Program offers us an unparalleled opportunity to integrate and consolidate our surveys and questionnaires. In the process, we will simplify and standardize the wording on our questionnaires where appropriate, and reduce the amount of duplication in our survey activities. We need to make responding to business surveys easier for our respondents by ensuring that we use normal business language that businesses understand. The integrated questionnaire approach will yield added benefits in better quality data from respondents, and a more integrated set of micro and macro data for economic analysis. In the longer run, we can also look forward to a more efficient data collection process.

Integrated questionnaire - the conceptual model

The new integrated questionnaire will be a database containing all questions on financial and operating or other characteristics which might be asked of any enterprise or establishment in Canada. The questions will be grouped in "modules" covering the major categories of information normally requested on business surveys: revenue, expenses, balance sheet items, capital expenditures, employment, inputs/outputs, class of customer, destination of shipments/sales, and "industry-specific questions". The questions will also be "flagged" to indicate whether they are aimed at the enterprise as a whole, or aimed at production units. Blocks of questions will also be identified to indicate whether they are best answered normally by the comptroller, the payroll manager, the production manager or some other officer within the business.

The goal is to create a database system which can be used to generate a custom set of questionnaires for each responding enterprise with:

(a) The Enterprise-Level Questionnaire - a set of standard questions on consolidated income, expenses (including purchased services), balance sheet items, investment and employment for the entire enterprise.

(b) The Enterprise Portrait - a diagrammatic presentation of the enterprise structure, generated from the BR profiling sub-system, showing the legal and operating structure of the enterprise as currently recorded, to be confirmed or updated by the respondent as required.

(c) The "middle" questionnaire designed to collect, from the enterprise-level, the 5 or 6 basic variables for each establishment needed to estimate value added.

(d) The Establishment Schedules - a series of "tear-off" schedules to collect commodity input and output, class of customer, destination of shipments and other detailed "industry specific" information from establishments.

"Simple" enterprises with only one establishment will not need to receive the enterprise portrait or the "middle" questionnaire. They may however, be asked a few questions on organizational structure, to determine if a profiling follow-up is required. They will receive the "tear-off" establishment schedule(s) for the one industry in which they operate.

Initially, a set of "harmonized default schedules" based on the existing questionnaires will be loaded on the database and used to generate industry-specific establishment schedules. However, the ultimate aim is to generate "personalized" or "customized" questionnaires for each responding enterprise. The set of establishment schedules for each enterprise will be individually printed. Initially, the schedules sent to an enterprise will be selected based on the industry codes assigned to the establishments identified on the profile. Subsequently, after development work is complete, the selection of the schedules will still be based on the industry codes, but the selection of input and output questions on each schedule will be triggered by the firm's previous responses to establishment schedules.

Customized questionnaires

The questionnaire database will be linked to Collection Entity (CE) structures maintained on the BR by the Collection Entity processes. The CE arrangements related to every questionnaire that is to be mailed out will be generated from the business register, as is currently done for more than 20 business surveys.

For each in-sample unit, a specific questionnaire is associated with a Coverage Record and the targeted Statistical Unit(s) which correspond most closely with the production unit(s) for which the questionnaire is intended. The Questionnaire Record is linked to a Contact Point Record which indicates who is to receive the questionnaire, the address where it is to be mailed, etc. Collection Entities are constructed and updated automatically on the business register, and maintained over time through survey feedback processes. The survey manager initiates the process by defining the rules for selecting statistical entities which correspond as closely as possible to the production entities from which the data will be sought. The survey manager is provided with a set of tools to allow for the customization of CE's for exceptional cases when the general selection rules do not apply.

Using the questionnaire database and the CE system, it will be possible to generate a customized set of questionnaires and schedules for each enterprise, and to direct the various schedules to selected contact points within the enterprise. For the UES, a choice will be made whether to build one CE driver to cover construction of CE's for all schedules or to build a separate simple CE driver for each applicable schedule. Using one or the other of these approaches, a full set of questionnaires will be constructed for each enterprise based on the total CE information on the BR. Combined with previous-period survey information, the questionnaire could be further "personalized" to request only the information applicable to each enterprise.

The "personalized" portion of the system will not be feasible for "the 1998 industries" (for the 1997 reference year), but should be ready for use for the 1999 survey (for the 1998 reference year). This will be a challenge, because of the need to link not only the BR and profiling information, but also the survey response database needed to identify the specific input and output commodity questions which have been filled in for each establishment in the 1997 survey. Where personalized questionnaires are already in use, as in the Annual Survey of Manufactures, their use will continue next year as well as in the future.

Development plan – outline of activities

Under the UESP schedule, developments in Year 1 will focus on implementing the Part I survey for all 8,000 complex enterprises, while implementing Part II and Part III surveys for only a few selected industries - "the 1998 industries" - where no annual business survey program now exists. The specific industries to be included in the Year 1 development include Construction; Real Estate Operators; Real Estate Agents and Brokers; Restaurants, Caterers and Taverns, the Taxi Industry, the Courier Industry and Aquaculture. The major implementation effort is now planned for Year 2, when the Part II and Part III surveys will be expanded to cover the remaining industries. This pilot or "mini-UESP" will combine subject matter from diverse areas within the integrated questionnaire, to be shared as a common tool of the subject matter areas participating in the pilot. The efforts to harmonize concepts and formats in the pilot will help prepare us for the full-scale implementation of the UESP in the second year.

The major impact this revised plan will have on the development of the integrated questionnaire database will be (a) to reduce the immediate need for "default questionnaires" to the selected "1998 industries" and (b) to force a number of activities to proceed in parallel. Work is already underway on development and testing of the Part I enterprise questionnaire. Work on the Part II and Part III establishment questionnaires for "the 1998 industries" will commence immediately. Work on content harmonization, as well as work on the data model, database specifications and database construction and loading will have to proceed in parallel with these activities.

To support the harmonization process, a prototype database will be built initially from the full set of existing annual business questionnaires. The questions on these questionnaires will be studied, classified by module, target level, variable and industry, and analyzed to identify areas of duplication. Where different surveys ask for the same information, special working groups, organized by module, will be set up to develop common concepts and definitions for the variables.

Major activities

Briefly, the vision of the future development of the integrated questionnaire covers the following major activities:

- We need to develop a comprehensive UES data requirements model (DRM) starting from work done over the last few years in the Business and Trade Statistics Concepts Committee. This will involve bringing together both physically and conceptually the previous work done on the industrial production, financial and production models, and incorporating any additional requirements arising from intended use of our data in the HST allocation formula.

- Several draft versions of the new questionnaire for "complex" enterprises at the enterprise level have been developed. One of these drafts, an Annual Supplement to the IOFD Quarterly Financial Statistics Survey, is being tested now with the enterprises in scope for the quarterly survey. The results of this test should give us a good indication of the ability and willingness of enterprises to provide the kind of revenue and expenditure data we will need to collect in the Enterprise-Level Questionnaire.

- The draft questionnaires will be reviewed by a working group including representation from the SNA branch and other subject matter and operational areas. The mandate of this group will be to construct a definitive enterprise questionnaire (incorporating the ideas of operations, methodology and forms management) providing a clear picture of the content required at this level of the enterprise. The working group will ensure that the questionnaire requests all of the income, expense, balance sheet and profile information required for the calculation of true value-added. Particular attention must be paid to questions on **purchased services**. Selected sets of questions will be tested with a small group of complex enterprises during the second quarter of 1997.

- The set of existing business surveys which are to be encompassed in the full implementation of the UESP has been identified. These surveys are listed in Annex 4 of this paper. The full set of **annual** business questionnaires associated with these surveys will be studied and the information collected will be cross-classified and analyzed to identify areas of duplication. This analysis will initially segregate surveys/questionnaires by target units (enterprise, company, establishment and location), and by "module" (revenue, expenses, etc.) with subsequent analysis of overlap and duplication among the questions to see if this overlap is justified.

- A separate working group will develop the content and wording of the "new" questions on class of customer and origin/destination of shipments. The intent is that these questions be asked in a standard way wherever possible. However, innovative arrangements may be required to collect class of customer data in certain industries, where the respondents cannot be expected to know a great deal about the class of customer. The working group will be expected to identify these cases (which are probably numerous) and to recommend innovative - and sensible - approaches for trial.

- Detailed specifications for the integrated questionnaire database and the system to generate the various schedules will be developed. This will involve representatives from subject matter areas, the Business Register Division, the Forms Management Section and systems specialists. Most industry questionnaires are already generated

from machine-readable templates and part of the task of this group will be to determine whether and how this information can be used in the process to generate the required schedules. Ideally, the system should be able to produce machine-readable as well as paper versions of schedules. The machine-readable versions would be made available for those respondents who wish to respond electronically. If possible, the machine readable versions should incorporate edits on the data to be entered by the respondent.

How big is the job?

There are 206 Statistics Canada surveys, from 17 Divisions, going to business respondents in Canada each year. These 206 surveys use more than 600 separate forms or questionnaires. Some of the surveys are designed to collect data on the operations of the entire enterprise, while others are designed to collect information on specific operating units (statistical companies, establishments or locations) within the enterprise. Trying to construct one "integrated questionnaire" to encompass the content of all of these many forms seems a huge task.

However, these surveys have differing frequencies and most are directed at business units which are classified to specific industries, which will be identified by NAICS industrial classification codes.

Based on the SDDS, the 206 surveys presently going to business respondents break down in the following ways:

TABLE 8.1 Business Surveys by Division and Frequency

	*SDDS Total Surveys**	*Non-Business Surveys ** Admin or other Sources*	*Surveys going to Business*	*Annual Surveys*	*Quarterly Surveys*	*Monthly Surveys*	*Other Sub-Annual*	*Other Surveys*
Agriculture	35	3	32	16	2	7	3	4
BoP	25	2	23	17	4	2		
B&TS Field	1	1	0					
BRD	2	1	1		1			
ECTD	31	24	7	7				
IOFD	7	3	4	3	1			
Industry	58		58	17	9	30	1	1
International Trade	1		1			1		
I&CSD	4	1	3	3				
Labour	8	2	6	2	1	2		1
Micro Economic Studies	1		1	1				
National Accounts	4	1	3	1				2
Prices	19	6	13	2	4	4	3	
Science and Tech	13	5	8	5		2		1
Services	11		11	10	1			
Small Business	13		13					13
Transportation	26	4	22	6	7	6	3	
TOTAL	*259*	*53*	*206*	*90*	*30*	*54*	*10*	*22*

* The SDDS definition of "survey" includes programs which produce "derived statistics", where no survey activity per se is conducted.

** The second column excludes derived program and surveys aimed at non-business universe.

TABLE 8.2 Business Surveys by Division and Target Unit - Annual

	SDDS Total Surveys*	Annual Surveys	Enterprise	Company	Establishment	Location	Other
Agriculture	35	16			16		
BoP	25	17	17				
B&TS Field	1						
BRD	2						
ECTD	31	7		1	5	1	
IOFD	7	3	3				
Industry	58	17		5	10	2	
International Trade	1						
I&CSD	4	3			3		
Labour	8	2	2				
Micro Economic Studies	1	1	1				
National Accounts	4	1	1				
Prices	19	2		1	1		
Science and Tech	13	5		2			3
Services	11	10		6	1		3
Small Business	13						
Transportation	26	6		5	1		
TOTAL	*259*	*90*	*24*	*20*	*37*	*3*	*6*

* The SDDS definition of "survey" includes programs which produce "derived statistics", where no survey activity per se is conducted.

The 90 annual surveys are the primary focus for the Unified Enterprise Survey. The target statistical units for these surveys are shown in Table 8.2. The UESP will consider sub-annual and other surveys after the annual surveys have been integrated into the UESP. However, when we negotiate special reporting arrangements with large enterprises, we must be prepared to deal with the full range of Statistics Canada's surveys, monthly and quarterly as well as annual surveys, since these enterprises will want to discuss their reporting arrangements to *all* Statistics Canada surveys.

Enterprise-level surveys

At the enterprise level, there are five subject matter areas conducting surveys, but the two major players are Balance of Payments and Industrial Organization and Finance Divisions. The IOFD questionnaires concentrate on the collection of financial statement information, including income, expense and balance sheet items, while the BoP questionnaires are aimed at specific categories of international transactions of selected enterprises. Both divisions also conduct annual "profiling" surveys, in which they collect updates to information on the Canadian companies (legal entities) which form part of the reporting enterprises. Two of the 17 BoP surveys have broad coverage, i.e. they go to most Canadian enterprises. The remainder appear to be narrowly focused on very select groups of enterprises, such as banks and shipping firms, which are likely to have

significant international transactions. A working group (perhaps a sub-group of the working group on the Enterprise-Level Questionnaire) should be formed to recommend whether the data requirements of these two divisions should be consolidated in one enterprise-level questionnaire, or whether some of the BoP requirements might be better handled through one or more "schedules" attached to the enterprise survey.

The working group on the Enterprise-Level Questionnaire will also need to review the surveys listed below from Labour, Micro-Economic Studies and National Accounts and Environment Divisions to determine whether they should be incorporated into the Enterprise Questionnaire, handled as "schedules", or left to operate as independent surveys in the short run. In the long run they should be incorporated into the Unified Enterprise Statistics Program.

1. New Workplace and Employee Survey (Labour);
2. Annual Survey of Employer-Sponsored Pension Plans in Canada (Labour);
3. Survey of Operating and Financial Practices (Micro-Economic Studies); and
4. Environmental Industries Survey, 1995 (National Accounts and Environment).

Company-level surveys

At the company level, six subject matter areas are conducting 20 annual surveys. The major players are Industry, Services, Science and Technology and Transportation Divisions, with ECTD and Prices each conducting one annual survey. While these surveys will have to be examined further, some preliminary observations follow. Services Division has been planning to move their surveys to the establishment level. Their inclusion in the UESP will provide an opportunity to follow through on this plan. The Industry surveys should also be easily incorporated in the UESP. Four of the five surveys collect financial and operating data on energy companies. The fifth is the annual survey of the construction industry, one of the areas to be given special emphasis under PIPES.

The situation for the three Science and Technology surveys is not so clear. Two are collecting data on R&D spending, and special collection/contact arrangements are required to obtain these data. Further examination will be needed to determine how these surveys can be handled under the basic UESP collection model. The third company survey is in the regulated telecommunications industry. Regulated industries present particular problems, as discussed below.

The five transportation surveys are also in regulated industries. However, since the primary function of these surveys appears to be the collection of financial data, there is no a priori reason to exclude them from consideration for inclusion in the UESP program.

Establishment-level surveys

The thirty-seven annual establishment surveys are carried out by seven divisions, but the two major players are the Agriculture Division with sixteen surveys and Industry Division with ten. The agriculture surveys are generally aimed at the farm and differentiated by

major crop. The basic question with these surveys is not whether they should be included, but when. Integration of the Farm Register with the Business Register will be a prerequisite for their inclusion in the UESP. Given the priority associated with integrating the IOFD enterprise and legal entity files with the BR, the work on the FR-BR integration will be delayed until 1998.

Industry Division's annual establishment surveys, on the other hand, are high on the list for inclusion in the UESP, especially the annual surveys of manufactures. The other surveys should be studied in further depth before proceeding.

The three ICSD surveys cover capital expenditure intentions and preliminary estimate, the revised forecast and the estimates of actual expenditures. While there are definite timing issues to be resolved, all three of these surveys will be included in the UESP.

The remaining surveys in ECTD, Prices, Services and Transportation Divisions need further study, but the general approach for all annual establishment surveys will be to include them in the UESP.

Other units

The five surveys addressed to non-standard units are found in two areas: the Science and Technology Project, and Services Division. The two S&T surveys are aimed at licensed operators of radio and television stations, and of cable operating systems. The non-standard unit in each case is used because of specific regulatory requirements arising from the fact that these surveys are conducted jointly with the CRTC, and designed to fulfill particular regulatory requirements of that agency.

The three Services Division surveys were called non-standard because they are addressed to firms in professions where partnerships are the dominant form of organization. This may present no obstacles to their inclusion in the UESP, but some further investigation is warranted.

Survey content and concepts

Still more analytical work is needed before we can come to conclusions regarding similarities and differences in concepts for similarly named variables in different surveys. Work done in 1993/1994, examining the extent to which annual production surveys adhere to the concepts and definitions of the production model, will provide a valuable starting point.

Special situations

A number of special situations will need to be examined in detail before decisions are taken regarding the manner of their inclusion in the UESP.

1. The Balance of Payments program, with its specific focus on trans-border transactions, is one of these areas. As mentioned earlier, a special working group will be formed to examine the possibility of merging Balance of Payments data requirements with those of IOFD. Regulated industries such as telecommunications and transportation are subject to very specific data collection requirements because Statistics Canada surveys serve a dual purpose, collecting data jointly for statistical

and regulatory use. These arrangements have developed over the years as a way of reducing response burden. Statistical surveys have also benefited from high response rates because of the added leverage associated with regulatory demands for information. In some cases, the regulatory requirements also dictate the use of non-standard units for data collection. Because of these benefits and differences from the rest of Statistics Canada's business data collection program, further study is needed before decisions are taken regarding the manner of inclusion of these business units in the UESP.

2. By formal agreement, Mining Surveys in Canada are conducted by Natural Resources Canada and shared with Statistics Canada. NRC has a continuing need for technical as well as financial information on mining operations, and has developed excellent working relationships with provincial mining authorities who also need the same information. These formal working arrangements will have to be examined in detail.

Chapter 9

Classification Systems

The implementation of the Unified Enterprise Statistics Program (UESP) for the 1997 reference year will coincide with the implementation of the North American Industry Classification System (NAICS). Priorities of the UESP need to be reflected in the process of confirming and correcting NAICS codes on the Business Register.

On the commodity side, it is very important that coherent commodity classifications be used. For data comparability, and for economies in the design of questionnaires, standard classifications of commodities, both goods and services, will be used as far as possible on both the input and output sides of questionnaires.

The North American Industry Classification System (NAICS)

Goals

All annual business surveys will produce data on a North American Industry Classification System (NAICS) basis beginning with the 1997 reference year.

By early July 1997, all statistical location units on the BR will be coded to NAICS, either directly or by using defaults. Using the NAICS coded statistical locations, the BR will code to NAICS all statistical establishments, statistical companies and statistical enterprises.

The BR will conduct a nature of business survey of the locations whose NAICS codes were assigned by default or directly but with a low degree of confidence. Priorities for nature of business reports (NBRs) will be determined. By December 15, 1997, NAICS codes resulting from the NBR process will be available on the BR for all statistical units that may be covered by the UESP for the 1997 reference year. NAICS codes for the remaining statistical units will be on the BR by the end of November 1998.

UESP samples will be drawn based on NAICS. Sample selection will take place at the outset of the first quarter of 1998 for mail-out beginning at the end of the first quarter of 1998.

Application of NAICS to statistical locations and statistical establishments

NAICS is designed for the classification of statistical establishments, as was the classification that it replaces, the 1980 Standard Industrial Classification for Establishments (1980 SIC-E).

Statistical establishments (usually just called establishments) are standard statistical units which correspond to individual production entities or groups of them, depending on the availability of certain data about the production entities. (Establishments sometimes correspond to parts of a production entity - notably, when a production entity crosses provincial boundaries.) The establishment is the unit for which statistics of economic production are produced. Under the UESP, data may be collected about groups of

establishments. An example would be all the establishments of one enterprise which are in the same industry and province. If data are not collected for each individual establishment, they will be allocated to each establishment.

The industrial classification is also used as a proxy for an activity classification in classifying statistical locations according to their primary activity. The statistical location is the lowest level unit in the statistical structure, coming below the establishment.

NAICS codes will be applied to statistical locations and to establishments.

Application of NAICS to statistical companies and statistical enterprises

NAICS will also be used for the classification of statistical companies and statistical enterprises on the BR. The classification process works from the bottom of the statistical structure upwards and the classification of statistical companies and statistical enterprises depends on the classification of their establishments. There are about 4,000 enterprises on the BR with establishments in more than one industry (1980 SIC-E 4-digit). Statistical companies and statistical enterprises which have establishments in more than one industry will be classified to the industry which accounts for the largest share of value added. Although the word "industry" is used here, implying the lowest level in the statistical structure, research is currently being conducted by Standards Division to determine the level of NAICS to which statistical companies and statistical enterprises will be coded.

Although designed basically as a classification for establishments, NAICS is replacing the 1980 Standard Industrial Classification for Companies and Enterprises (SIC-C) as well. The SIC-C provided classes for vertically integrated statistical companies and statistical enterprises. The basic set of NAICS classes does not make such provision.

The issue for the UESP is that, for purposes of data presentation (though not for sampling), the UESP requires an alternative method of applying NAICS to statistical companies and statistical enterprises. A small number of NAICS special classes, say 10 to 20, representing combinations of NAICS industries, and a residual class, need to be created for statistical companies and statistical enterprises for which no one industry accounts for a specified percentage of value added (threshold to be determined). Standards Division is currently conducting research on this topic.

Application of NAICS to legal entities and related Revenue Canada accounts

In July 1997, each legal entity on the BR will be linked to its Business Number root (BN). Each BN will have an industry code. After an initial period of reconciliation, the industry code of a BN will be the code of the statistical enterprise.

The BR will provide Revenue Canada with changes to the industry codes of BNs that result from profiling and NBRs, so that Revenue Canada can update the industry codes held on its BN database. At any time the most up-to-date industry codes for BNs will be those derived from the BR rather than those on the Revenue Canada database or attached to administrative records generated by Revenue Canada.

Regardless of the number of legal entities / BNs in the legal structure of a business entity, the business may be involved in a variety of operations and its statistical structure may contain establishments belonging to more than one industry.

NAICS implementation issues

In early July 1997, all statistical units on the BR will be assigned an initial NAICS code. The NAICS codes of the statistical location units will be determined by one of the following methods:

- Where the 1980 SIC-E code corresponds to only one NAICS code, the NAICS codes will be assigned automatically, using a concordance table.

- Where the 1980 SIC-E code corresponds to more than one NAICS code, subject matter staff are in the process of assigning a NAICS code if they have sufficient information. They are also indicating the degree of confidence. Statistical locations whose NAICS codes are flagged with a low degree of confidence require follow-up.

- Where the 1980 SIC-E code corresponds to more than one NAICS code, and subject matter staff do not have sufficient information to assign a NAICS code, a NAICS code will be assigned by default. All statistical locations whose NAICS codes are assigned by default will require follow-up.

  Defaults will be established for all 1 to n situations. They will of necessity be more accurate in instances where a survey activity has taken place than they will be for situations where no surveys have taken place.

Most cases will be followed up using a Nature of Business Report (NBR). The 250 large and complex enterprises profiled by the Business Register Division on behalf of LESP on an annual basis will not be included in the NBR process but rather will have NAICS codes assigned on the basis of on-site profiles.

Current plans call for NBRs to be mailed out from July 1997 to December 1998, with completion of the process by March 1999, assuming a three-month lag in response and processing. However, as noted above, the UESP goal is that by December 15, 1997, NAICS codes resulting from the NBR process will be implemented on the BR for all statistical units that may be covered by the UESP for the 1997 reference year and that by the end of November 1998, NAICS codes for the remaining statistical units will be on the BR.

The issues for the UESP are to have priorities for NBRs assigned according to UESP survey priorities for the reference year 1997, and to have the results of the NBRs available on the BR when they are needed. The priority-setting task is underway. The demand for NBRs is being compared to the capacity to conduct them, and consideration is being given to accelerating the process.

In future years the issue of recoding units as their activities change will arise. Resistance rules are a set of conditions under which a unit's industry code is changed. Resistance rules are used to provide some assurance that a change will endure. The UESP needs to look at creating standard resistance rules to update or replace the various rules currently in use.

Commodity classification

The need for commodity detail under the UESP is derived from the needs of the subject matter divisions, the System of National Accounts and their clients, the administrators of the HST revenue allocation formula, in particular. Business surveys will have to produce commodity detail at a level that will meet new requirements of the SNA branch to produce improved provincial economic statistics, and also meet the ongoing needs of the SNA, the ongoing needs of the subject matter divisions and any cost recovery obligations.

It is very important that coherent commodity classifications be used. For data comparability, and for economies in the design of questionnaires, standard classifications of commodities, both goods and services, should be used as far as possible on both the input and output sides of questionnaires.

Statistics Canada policy stipulates that the Harmonized System (HS) become the definitional basis for all commodity statistics. The Standard Classification of Goods (SCG), an extension of the HS, is intended as a standard for application to all commodity (goods) surveys. The extent to which the SCG has been implemented is described below.

The UESP goal is that business surveys will use the Harmonized System/Standard Classification of Goods, or derivatives thereof, as the standard commodity classification for goods, and will use a standard commodity classification for services as far as possible, considering that a Canadian classification of services is still under development.

The SNA will provide its minimum requirements for commodity detail in the near future, and, for goods, they should be expressed in terms of classes in the Standard Classification of Goods, or aggregations of those classes.

Before redesigning questionnaires, an investigation will be needed of the commodity classifications being used on business survey questionnaires and in the System of National Accounts. This research will apply to inputs as well as to outputs. On the goods side, where the HS/SCG is not being used, any impediments to using the HS/SCG will be identified. On the services side, it will be necessary to identify: a) the services about which commodity questions are being asked on business survey questionnaires, and b) the types of breakdowns of services that are appearing, and how close those breakdowns come to the Central Product Classification or to parts of the Standard Classification of Services.

A PIPES exercise which covers more than the UESP is already working towards harmonization of the portions of commodity classifications which relate to personal expenditure. A committee chaired by Standards Division has been created with the mandate of improving the degree of harmonization among commodity classifications used by the System of National Accounts, FAMEX, the CPI, Retail Trade and Services. The time frame to achieve measurable progress is 6 to 12 months; the target reference year for improvements is 1997.

The material below provides background on the various commodity classifications and their use.

Commodity classification: goods

The Standard Classification of Goods (SCG) and its derivatives

The Standard Classification of Goods (SCG) is the standard used by Statistics Canada to collect commodity data and to uniquely identify such data in data banks. It is an extension of the international Harmonized Commodity Description and Coding System (HS). The SCG extends the six-digit HS coding system by an additional two digits. This extension was required to provide more statistical detail for goods that are manufactured in Canada, for the materials used in their manufacture, and for goods that are imported or exported.

The following table provides a status report on the extent to which the SCG was implemented by January 1, 1992. Notes at the bottom of the table give the current situation for some of the surveys. Standards Division is updating this table for inclusion in the 1996 version of the SCG, to be published in the first quarter of 1997.

Table 9.1 SCG Implementation in Commodity Surveys

Survey	Number of Survey Categories	Implementation Date
Imports	13,130 categories	January 1988
Exports	5,603 categories	January 1988
Annual Survey of Manufactures	7,430 categories of shipments; 3,494 categories of materials	1988 data year
Industry Division current surveys (approx. 54)	varies by industry	about 90% were SCG-based for 1992 data
Water transport*	483 categories	1991 data
Truck transport*	651 categories	SCG-based classification ready 1992, not implemented
Rail transport*	328 categories	SCG-based classification ready 1992, not implemented
Dairy product surveys**	varies by survey	draft classification ready 1991
Other Agriculture Division surveys**	varies by survey	preliminary work done 1991
Family expenditure surveys***	over 300 categories of goods	preliminary work done 1992
Consumer Price Index****	over 300 categories of goods	no action to date

* Water, truck and rail transport surveys will use the new Standard Classification of Transportable Goods, described below.

** Agriculture surveys have not yet implemented the SCG.

*** Family Expenditure Surveys have not yet implemented the SCG.

**** Implementation of the SCG in the CPI depends on its implementation in Family Expenditure Surveys.

The SCG is updated annually to reflect not only HS changes and basic changes resulting from commodity classification research but also the requirements of survey divisions for the incorporation of new detail from new surveys and changes in commodity detail covered by existing surveys.

In spite of the large number of commodity classes in the SCG, the commodity detail is insufficient to meet all requirements. This may be due to lack of space within the relevant range of codes or to the necessity to accommodate a different classification structure. In such cases, additional digits are allowed at the discretion of the subject-matter divisions.

Survey managers select from the standard the appropriate level of detail that is relevant for the nature of the universe surveyed and the purpose of the survey. Where the structure is not suitable, alternative structures may be developed, using classes from the standard as building blocks.

The Standard Classification of Transportable Goods (SCTG) is such an alternative structure. It is a new classification (October 1996) consisting of four levels that contain groupings based on HS or SCG "building blocks". It has been constructed by using a blend of transportation characteristics, commodity similarities, and industry-of-origin considerations, designed to create statistically significant categories.

The SCTG has not been implemented yet, however. The HS/SCG-based classification noted in the previous table was implemented for water transport, and could be replaced by the SCTG fairly easily. But the classifications used for rail and truck transport are not related to the HS/SCG and considerable work is needed to create concordances and code commodity libraries to the SCTG.

One other alternative structure has been created: an HS-based classification of machinery and equipment for the capital expenditure survey. The new structure has been field-tested and its full implementation will be considered as part of the changes to the capital expenditures program in conjunction with the UESP.

Retail trade

In the area of retail trade, a new classification of commodities has been created for implementation in the new quarterly Retail Commodity Survey beginning in January 1997. Following consultation with retailers and their associations, as well as with the SNA, the retail trade section has created approximately 110 retail trade commodities. The commodities are not simply HS / SCG classes or breakdowns of them, however.

The HS grew out of a need to simplify and expedite trade documentation of all types, including customs documentation. To facilitate the compilation of market data, Statistics Canada also implemented the HS for production statistics. The HS is designed primarily to facilitate the identification and coding of goods; therefore, it is principally organized by component material, a characteristic that is apparent from a physical examination of the goods. A corollary is that intended use is not required to classify a good.

The way in which retailers classify their inventories and sales is at odds with this principle since it is more along the lines of intended use than component material, and the new classification of retail commodities reflects this. A considerable amount of work will be required to complete a concordance between the HS / SCG classes and the retail trade classes.

Commodity classification in Input-Output Division

Input-Output Division uses 678 commodities at the working level. These commodities are aggregations of Principal Commodity Groups (PCGs) which in turn are aggregations of classes from the 1988 SCG, in the case of goods. In the case of services, the PCGs and working level commodities are based on a roll up of classes from the old Standard Commodity Classification (SCC), with provision made for new commodities.

The working level commodities are further aggregated to the M and S level commodities.

Commodity classification: services

A classification of services will be important for the collection of service inputs on head office questionnaires, for collecting information about intermediate service inputs on any questionnaire, and for the collection of outputs on questionnaires for service-producing industries. A minimal list of service outputs (such as rental and leasing) to be collected across all industries will be compiled.

Central Product Classification (CPC)

The CPC is an international classification of both goods and services. Nevertheless, it is being discussed here under the heading of services classifications since the services portion of the CPC is of interest for the UESP and the goods portion may not come into play very much, given the existence of the more detailed HS and SCG.

Classes in the goods portion of the CPC are aggregations of HS classes and, therefore, are more aggregated than SCG classes. The CPC publication shows the links from the CPC classes to the HS. The goods portion of the CPC is being updated to bring it in line with the 1996 version of the HS.

Redrafting of the services portion of the CPC was completed recently. It is still subject to some, probably minor, changes by the United Nations Statistical Office prior to being presented to the United Nations Statistical Commission.

The CPC has existed in a provisional version since 1989. The revised version, if approved, will be published as "Version 1". It will contain areas in the services portion which have not been used or tested by statistical agencies. The CPC will still be subject to revision.

Standard Classification of Services (SCS)

The Standard Classification of Services (SCS) is a Canadian classification of services which is under development by Standards Division.

It is likely that the SCS will not be constrained by the structure of the CPC. Although classes in the SCS will not cross the detailed CPC classes, the SCS classes may aggregate differently than CPC classes. The SCS will have more detail. There will be a concordance between the SCS and the CPC. As the SCS is developed, recommendations for changes to the CPC will be made where appropriate.

Some portions of the SCS are already in use in surveys - computer services, management consulting and accounting. Some other portions could be used in tests, and others still have to be developed. The suitability of the CPC and SCS for the UESP will be assessed in the months ahead.

Other classifications

Geographical classification

Geography Division has expressed an interest in facilitating the production of sub-provincial data, where feasible. Given the GIFI and the larger number of business contacts that will result from PIPES, there is considerable potential to produce sub-provincial estimates. However, sub-provincial estimates are not the focus of the UESP.

Chapter

10

The Use of Administrative Data

Introduction

The use of administrative data has become a cornerstone of Statistics Canada's business survey programs. Statistics Canada started using taxation data as a replacement for survey data in the mid-1960s. The reliance on this administrative source has increased considerably over the intervening years. Initially the use of administrative data was seen as a way to reduce response burden imposed on the small businesses. However, increasingly administrative data have also become vital to the very maintenance of important statistical programs. Apart from providing the basis of and signals for the BR, administrative data are extensively used in the Agriculture and Labour Statistics programs. As well, they are a major input source for analytical and cost recovery projects such as the production of small business profiles of income and expenses and selected balance sheet items. The PIPES initiative will define an environment in which our reliance on the use of administrative data, and in particular, tax data, will increase even further.

PIPES demands that Statistics Canada provide higher quality provincial data, but this cannot be accomplished without regard to the consequent impact on response burden. It is with this background that the role of tax data will be discussed in this chapter. In the interests of brevity, neither the existing role of tax data nor the problems of transition to the post-PIPES environment will be covered in detail.

Background

In the past, our ability to use administrative data has been impeded by problems associated with timeliness, linkages, data elements and cost.

The fact that companies are allowed six months after the end of their fiscal year to file T2 tax returns, introduces a significant delay in the process of acquiring tax data for publically incorporated companies. In addition, the traditional process of transforming tax data into statistical information has been a lengthy process of data analysis, capture and editing from photocopies of a sample of non-standardized businesses' financial statements. The introduction of a standardized GIFI schedule (previously known as the Financial Statement Information (FSI) form) should reduce the time lags associated with acquiring tax data, but there will still be inherent delays in this process.

Other constraints, however, have even more impact on data quality. Presently, the financial statement data received are in the format that the tax preparer wishes to use. The transformation of these data to a standard set of statistical items is fraught with problems of definition, interpretation and completeness. The most acute problem presently is the lack of linkages between administrative units and the units of interest to Statistics Canada, and among the various administrative data sources. A different registration and numbering system has been used for each of the administrative programs, resulting in data sets that often are not easily linked, thus sub-optimizing the use of the data for statistical purposes.

What has changed?

Business Number

In its February 1992 budget, the Government announced an initiative to establish a "single business registration number" among the measures for "making it easier to deal with the Federal Government". The budget papers stated:

"At present, different departments require as many as six different registration numbers from Canadian businesses. This means more work for businesses, more cost for government and, inevitably, poorer service. The government is committed to making the necessary changes to arrive at a single registration number in cases where this would be advantageous to the business concerned. This will mean less paperwork for business, greater efficiency and responsiveness by government, better economic statistics and a more effective system of revenue collection."

The proposal of a single number which businesses could use when dealing with the Federal Government was accepted and is now being implemented as the "Business Number (BN)[2]". The BN is now being issued by Revenue Canada to anyone requesting Payroll Deduction, GST, Importer-Exporter or T2 accounts. Revenue Canada is using the BN exclusively for account registration and updates for business activity descriptions effective January 1, 1997. It is Revenue Canada's plan to have the BN universe of existing business accounts converted and confirmed by the end of March 1997.

The BN does not link legal and operating entities. The BN root number does, however, facilitate the linkage of all Revenue Canada accounts. By using the T2 account of the BN a link can be made to the legal entity (LE) on the Business Register (BR). The LE is linked to the top operating unit or enterprise record on the Register. Accordingly, while the BN does not provide for direct links between legal and operating entities of complex enterprises with more than one legal entity. It does allow all of the administrative and survey records for a particular legal entity to be associated with one another.

What has really changed? Statistics Canada is now assigning an SIC to the BN root which effectively results in an SIC assignment to all associated Revenue Canada accounts. This SIC will be harmonized with the Business Register (BR) SIC. The BR SIC will be the same SIC used for the T2 tax data and for T1 businesses using a BN account and providing the BN when they file their tax return. In other words, after the BN is fully implemented, a new era will be ushered in where it will be far simpler than in the past to link together diverse sources of information about a particular business enterprise.

Standardized General Index of Financial Information (GIFI)

Revenue Canada has initiated a program to standardize the content and method of filing for financial statements, including the Income and Expenses statement, the Balance Sheet and associated schedules such as the Income Tax Reconciliation which is used to reconcile the book profits to profits subject to taxation. This initiative is referred to as the GIFI schedule.

[2] Starting in January 1997, all businesses required to participate in one (or more) of the following tax programs: Payroll Deduction, Corporate Income Tax, Goods and Services Tax and Importers and Exporters Tax, will have a single registration number with Revenue Canada.

The size of this undertaking prompted Revenue Canada to redesign its corporation taxation system and its personal income tax system as it relates to the filing of unincorporated business income. Many working groups have been formed, one of which was given the task of drafting the GIFI and all the associated schedules. Statistics Canada was represented on this group to speak on behalf of the statistical users of GIFI information. Extensive consultations with private sector groups (including accountants and software designers) and the provinces are ongoing to ensure that all interested parties are given the opportunity to comment on the new filing requirement. The GIFI schedule will define a standard set of Income Statement and Balance Sheet items that corporations and individuals will be required to file when reporting the results of their business activities to Revenue Canada. By and large, corporations will be asked to report the same detail that they would normally have included on non-standardized financial statements so it is reasonable to infer that this initiative will not add to the existing level of response burden. GIFI reporting will be required in electronic form for all but the smallest businesses, meaning that it will be a very timely source of data.

Goods and Services Tax (GST)

Every business in Canada with $30,000 in annual sales is required to register a Goods and Services Tax account with Revenue Canada in order to submit tax collected and to claim input tax credits. Smaller firms may also choose to establish an account, in order to claim credits. Taxi and limousine service businesses are required to file irrespective of the $30,000 threshold. There are about 2 million active GST accounts.

The GST file from Revenue Canada contains the firm's total sales of GST liable goods, the GST value collected and the GST on inputs purchased by the firm (and thus for which an input tax refund claim is being filed). Identifier information includes 4-digit SIC coding. Revenue Canada's GST files are available to Statistics Canada. Research on GST data quality is presently being undertaken by staff in the Tax Data, Business Register and Services Divisions, and by the Time Series Research staff.

Services Division work is considering the usefulness of GST remittances and sales data to support industry estimates of revenue. The work is not yet complete but preliminary results are very promising. The GST data could permit very timely and comprehensive analysis of business cycle trends, given its monthly availability, and the fact that it covers virtually all businesses. The fact that GST account information for firms will be linked to other taxation account data through the Business Number will make GST data useful for estimating enterprise revenues.

Payroll Deduction

The Payroll Deduction (PD) file obtained from Revenue Canada contains information on remittances, employment, and total payroll for the firm. The two latter items have been added to the PD form at Statistics Canada's request to provide data for the Survey of Employment, Payroll and Hours (SEPH). For roughly three years, these data were

captured by Statistics Canada staff for a sample of firms under what is called the "PD7 sample". Since August 1996, Revenue Canada has carried out the data capture of **all** of the forms. Revenue Canada transmits PD files to Statistics Canada on a weekly basis for the remittances filed in the previous week. Revenue Canada also provides a file containing identifier information (e.g. active or inactive, remittance periodicity, etc.); this file, known as PAYDAC, is transmitted for update purposes twice a month.

Presently the Labour Division makes extensive use of PD7 remittance data within its program. Employment and payrolls data have been used for small simple businesses since March 1993 and for all non-complex businesses since May 1996. PD records will be used for complex businesses beginning with the May 1997 reference month.

The availability of the single Business Number for linking PD to other taxation account data will enhance the present use of PD data for estimating employment (as used by the Survey of Employment, Payroll and Hours) and for determining whether a firm is active (as recorded on the Business Register).

Electronic filing

Revenue Canada will require that GIFI be filed electronically. This promises to largely eliminate the need for sampling and data capture while at the same time improving both the timeliness and the quality of the tax information made available to Statistics Canada.

A more complete description of the definitions, content, coverage, volume, periodicity, timing and quality of each of the administrative data sources mentioned above is provided in Annex 3 of this document.

Assumptions

The proposed uses of tax data in this report are predicated on the following assumptions:

1. Association of legal entities with their related operating units on the Business Register will be expanded to cover multi-legal-entity enterprises where possible.

2. The Business Number (BN) will be available in the Business Register by July 1997 and will provide a direct linkage between existing Revenue Canada accounts.

3. As of July 1997, the Business Register will cover all enterprises with a BN linked to at least one active Payroll Deduction (PD) account or to at least one active GST account showing at least $30K in Gross Business Income. The resistance rule that there must have been some remittance activity over the past 12 months will still apply.

4. Starting in April 1998, Revenue Canada will capture GIFI data for all the T2's with a non-zero Gross Business Income, even if most filers provide free-format paper statements. However, for tax year 1997, there will be no electronic data in the GIFI format for a significant portion of the T2 universe.

5. For tax year 1998, the GIFI form and electronic filing will be mandatory for the corporations with revenue or assets larger than $3 million. For smaller corporations the form will be available, but completion will not be mandatory. As a result, Revenue Canada will have to capture data from free-format statements for a portion of the T2 population.

6. For tax year 1999, Revenue Canada will introduce the GIFI form for the T1 businesses. For those not filing electronically, GIFI data will be captured from the paper forms or free-format statements if these T1 businesses have a gross income larger than $30,000 from any one source. The paper filers may be the majority. For the T1 farms filing with paper statements and having an income smaller than $30,000, Revenue Canada will capture GIFI data for a sample.

Proposed uses of tax data

As previously mentioned, the use of administrative data is a cornerstone of the business statistics program at Statistics Canada. The intent of the UES Program is to use administrative data, particularly GIFI data, to replace direct collection of financial information from simple enterprises in Canada. Expected improvements in administrative data sources will also permit more extensive use in frame maintenance, sample design, edit and imputation, and estimation stages. While many details remain to be worked out, the following pages present specific ideas which will be further investigated.

If the enterprise has only one legal entity (and therefore, in principle, one BN) and a set of production units in one industry and one province, the data can be used not just as an enterprise-level benchmark total, but also as a source of information about production. (Additional information may be required for T1 partnerships and T2 joint ventures in order that revenue may be allocated.) If the enterprise has only one legal entity and its set of production units covers more than one industry or more than one province, the administrative data can be used for enterprise statistics but cannot be used directly for production statistics. The data can be used for production statistics only if the profile of the enterprise is known and appropriate allocation variables have been provided. If the enterprise is composed of a family of legal entities and a set of production units in more than one province or more than one industry, the administrative data cannot be used directly for either the enterprise statistics or the production statistics. The data can be used only if a profile exists where consolidation and allocation variables are provided. In summary, regardless of the complexity of an enterprise it is possible to use the administrative data as long as a profile with appropriate allocation data is available.

Sample design

One use of the administrative data is to provide economic variables to be used for stratification and sample size allocation. Table 10.1 shows for each reference year the evolution in the use of administrative data, for sample design purposes, for each of three possible variables. In the table, the abbreviation SUF stands for Revenue Canada's

taxation universe file, which contains five variables for every T2 tax record processed by RC in a given tax year. COMSCREEN refers to Revenue Canada's computer screening process for unincorporated tax filers reporting business or professional income.

Table 10.1 Use of Administrative Data in Sample Design

Variable	*Sub-Population*	*RY 1997*	*RY 1998*	*RY 1999*	*RY 2000*	*RY 2001*
No. of Employees	PD account owners	Current PD-7 # of employees	Current PD-7 # of employees	Current PD-7 # of employees	Current PD-7 # of employees	Current PD-7 # of employees
Gross Business Income	GST account owners	Current GST data	Current GST data	Current GST data	Current GST data	Current GST data
	PD owners without GST	Current PD remittances	Current PD remittances	Current PD remittances	Current PD remittances	Current PD remittances
	T2 enterprises without GST or PD	TY 1995 SUF	TY 1996 SUF	TY 1997 SUF	TY 1998 GIFI	TY 1999 GIFI
	T1 enterprises without GST or PD	TY 1995 SUF	TY 1996 SUF	TY 1997 SUF	TY 1998 SUF	TY 1999 GIFI
Assets	T2's	TY 1995 SUF	TY 1996 SUF	TY 1997 SUF	TY 1998 GIFI	TY 1999 GIFI
	T1's	TY 1995 COMSCREEN (for ½ T1's with GI>$70K)	TY 1996 COMSCREEN (for ½ T1's with GI>$70K)	TY 1997 COMSCREEN (for ½ T1's with GI>$70K)	TY 1998 COMSCREEN (for ½ T1's with GI>$70K)	TY 1999 GIFI

For enterprises with GST accounts, the GST data will be used to estimate the gross business income (GBI). Currently, the GBI estimates on the BR are either model-based or come from a tax file that is at least two years old. For enterprises without a GST account, we will use GIFI data when Revenue Canada delivers this data source.

For the T2's without a GST or PD account, the Gross Business Income and Assets variables will be those on the SUF for tax year 1997. For tax year 1998, if Revenue Canada cannot capture the GIFI data for the enterprises with a revenue and assets smaller than $3M, it will be possible to use the SUF data as a contingency plan for these smaller businesses. A similar contingency plan will be used for T1's if the introduction of the T1 GIFI is delayed.

Edit and Imputation

The introduction of the Business Number will allow the greater use of administrative data for editing and the imputation of survey data. For reference year 1997, the sources will be: the Import/Export data, PD-7 sample data, PD remittances, T4 data, GST data, and various tax data files captured by Revenue Canada and Statistics Canada. For reference year 1998, data sources will be the same as for 1997 for T1 unincorporated businesses. However, for the T2s, the Statistics Canada samples will be dropped since the T2 GIFI data will then be available. From reference year 1999 onward, the proposed plan is to drop the T1 sample as the T1 GIFI data should become available. If Revenue Canada plans for the GIFI do not materialize, it will always be possible to use the other tax data files.

The administrative sources will provide data not only to edit and impute survey data but also to edit and impute data across sources. This way it should become possible to construct a complete set of administrative variables for the entire universe of enterprises.

Estimation

The use of administrative data at the estimation stage is the one for which the expectations are the greatest as its impact can be very beneficial. The benefits expected are to reduce respondent burden and survey cost as well as to increase the reliability of survey estimates. However, the full benefits are not for reference year 1997, but for a few years down the road. During the transition process, careful quality evaluations of the administrative sources will be required.

The SEPH program has set out a good model for making effective use of tax data in conjunction with survey data. That program now uses PD7 records exclusively to measure two key variables, labour income and employment. A small sample of enterprises is used to collect information on other labour characteristics – hours worked, types of employees, etcetera. Modelling techniques are invoked to relate these characteristics to the two variables for which comprehensive tax data are available. This permits the program to project the sample results across the entire population.

When GST and GIFI data are fully on-line and useable for statistical purposes, they will provide a detailed and comprehensive picture of the revenues, expenditures, assets and liabilities of simple enterprises. Additional variables that are required under PIPES – commodities purchased and sold, origin and destination of shipments, class of customer information and so on – will be measured with small sample surveys and then associated by models with the administrative data. In the short term – meaning the next year or two – these administrative sources will not be available, so it will be necessary to continue sampling and capturing T1's and T2's. However, by 1999 or 2000 this activity will cease and the improved administrative data from the GST and GIFI files will then be in use.

Risks associated with each tax source

Generally speaking, the risks associated with any administrative data source are difficult to measure. Responsibility for the content and implementation plan rests with another Department and, as such, are not under STC control. In the past few years, however, the climate has changed to the point where STC now is a significant partner in the design and

implementation of broadly based government initiatives, thus lessening the overall risk to our programs which use these external data sources. The following risk statements are made with the above caveat in mind.

1. Timing risks associated with the GIFI appear to be substantial. The data content has been established but this has yet to be put to the test of the Canadian Institute of Chartered Accountants (CICA), focus groups, or individual businesses in the form of a pilot test. In addition, mandatory compliance to a minimum level of detail is not part of the current plan and, depending on what the actual response pattern turns out to be, this could jeopardize the plan to use GIFI data as a direct replacement for survey data. Discussions between the CICA and Revenue Canada are ongoing and are expected to be completed by the spring of 1997. This will be followed by field testing during the rest of 1997 and voluntary introduction of GIFI by April 1998. Full introduction was originally scheduled for October 1997, but the target has now been postponed until October 1998.

2. Risks associated with the BN are seen as moderate. Presently, approximately 80% of legal entities have their accounts converted to the BN format. The remaining 20% will have a BN mechanically assigned, after matching on name and address, etc. to identify and remove duplicates. Given the limitations of record matching techniques, there could still be some marginal duplication among the BN identifiers which will be implemented on the BR by July 1997.

Chapter

11

Computer Systems and Database Design Requirements

Introduction

This chapter considers issues related to the computing environment (hardware and software) requirements for the Unified Enterprise Statistics Program. It portrays in broad terms the functional requirements to achieve the goals of the project.

The computer systems vision and database design are now in development, based on the UESP concept described in this paper. A more comprehensive and detailed paper focused on the long term systems vision and a practical short term implementation strategy will be released in Spring, 1997. This more detailed document will include discussions of the systems design and architecture, data models and process models for the UESP.

This chapter outlines general systems requirements based on the activities and functions of the UESP, as described in this paper. The systems staff on the project and elsewhere in Statistics Canada can use this outline to plan and specify in more detail, to acquire and to develop the components of a functioning system which achieves the goals of the project within tight time constraints. All of the foregoing must of course be accomplished within the Information Technology Framework (ITF) and current STC computing environment and respecting appropriate and relevant federal and STC policies regarding confidentiality, security, access, language and procurement.

"The challenge, and the opportunity, therefore, lies in moving to corporate rather than consortium data management."[3] In this initiative the BTS Field will make major strides towards implementing corporate data management. It is a unique opportunity, not to be missed, for bridging the technological islands which have evolved within business statistics programs. The improvements and harmonization made possible in concepts, definitions, methodology, use of common frames, classifications and integrated questionnaires must be carried through to the computing and communications environment in general and to the databases and supporting systems Statistics Canada will need to take maximum advantage of its information holdings. The strategic direction outlined in the ITF reinforces this theme of integration and the movement away from the former "stovepipe", and often stand-alone, approaches towards cooperating systems and common information repositories. Some of the important elements of the long term informatics vision for the UESP are:

- A fully redesigned Business Register with a graphical user interface, real-time updating and corporate-wide browse access.
- A redesigned and better integrated suite of generalized systems for the UESP.

[3] G. Priest, "Data Integration: The View from the Back of the Bus", Proceedings of STC's Symposium '95.

- Faster network servers and greater bandwidth for improved communications and file transfer capability; and computer hardware with much more processor power, memory and disk storage capacity.

- Computer security systems to ensure convenient access of tax and survey data to authorized Statistics Canada users, and only to authorized users.

- A well integrated set of tax and survey micro-databases based on a good data model, well developed standards, and a rich set of meta-data.

- A well developed set of analysis tools to tabulate and graphically present data from the infrastructure micro-databases.

- Revised output production systems built to exploit to maximum advantage the infrastructure databases.

- Multi-media informatics training available widely at Statistics Canada at the desktop.

A key to realizing the benefits that have been referred to in previous chapters will be the development and use of a suite of interconnected, reliable and matched software tools and services and their integration with new data structures. The goal will be to provide authorized analysts, service areas and operations staff with the range of information they need in an environment that has a consistent system image with respect to computational services, data accessibility and security in a seamless and transparent fashion from the perspective of the user. The envisioned systems approach is a set of comprehensive and coherent databases of enterprise information in a distributed client-server computing environment linking all major operational, service area, analytical and dissemination groups to relevant data with intuitive, user-friendly, reliable, secure and timely access. Another important element in the success of this project will be the implementation of a rigorous data model to support these new data structures and to use them efficiently in our definition of the corporate databases and in every step of the system development cycle. Clear lines of authority and responsibility, and some rules of precedence, will have to be established for version control over data, timing of certain operations and activities, and for aggregations, as well as revisions (to the extent permitted) in the distributed database environment. Essentially this requires working through the roles of subject matter, infrastructure and service divisions in terms of the data processes (sampling, collection and capture, imputations, estimation and aggregation, dissemination, etc.).

An open development process must be adopted and a number of teams formed representing each of the major stakeholders - developers, implementation and maintenance, and users. Specialists, knowledgeable about BTS systems, from Systems Development Division, Survey Operations Division, and Operations and Integration Division will work with key people from BTS divisions and staff from the PIPES implementation team in order to meet the tight timetables for the project.

An Intranet site will be developed where project and meeting schedules, "key decisions", related background documents, etc. can easily be stored, found and searched electronically. Regular "town hall" style meetings for raising problems and issues for discussion may be helpful in order to ensure an open, informed process to resolve issues. To support this goal the project team will investigate and recommend a document control software to support authoring, publishing and document management.

General principles shaping the UES systems environment

The design for the UES systems environment draws on departmental instruments already available. Individually such tools may not be completely adequate for specific requirements; however, as a package they go a long way towards satisfying program needs. Four broad sets of databases emerge namely, those relating to infrastructure, tax and other administrative data, survey micro-data and program output data. Each is described very briefly below (see Diagrams 11.1 and 11.2).

Infrastructure databases

The infrastructure databases comprise information supporting the conduct and management of business surveying. In terms of the ITF these databases are present in the reference layer whether as data, metadata or meta-information products. This category includes the business register database, a classification concordance database, a database of survey metadata, and databases holding respondent and questionnaire information. It is vital to ensure that such relevant information is not only available but accessible and understood. In terms of respondent burden the Survey Impact Database (SID) and the Company Contact System (CCS) track key items of interest. These two systems will be closely examined as Statistics Canada implements a comprehensive monitoring system capable of measuring response burden by enterprise. Corporate tools such as the Survey Data Documentation System (SDDS), Meta-Information Data System (MIDAS) and the Central Non-Response Data Base (CNRDB) are important. In addition, the new integrated questionnaire database will assist in generating, collecting and compiling the results from the various surveys.

Diagram 11.1

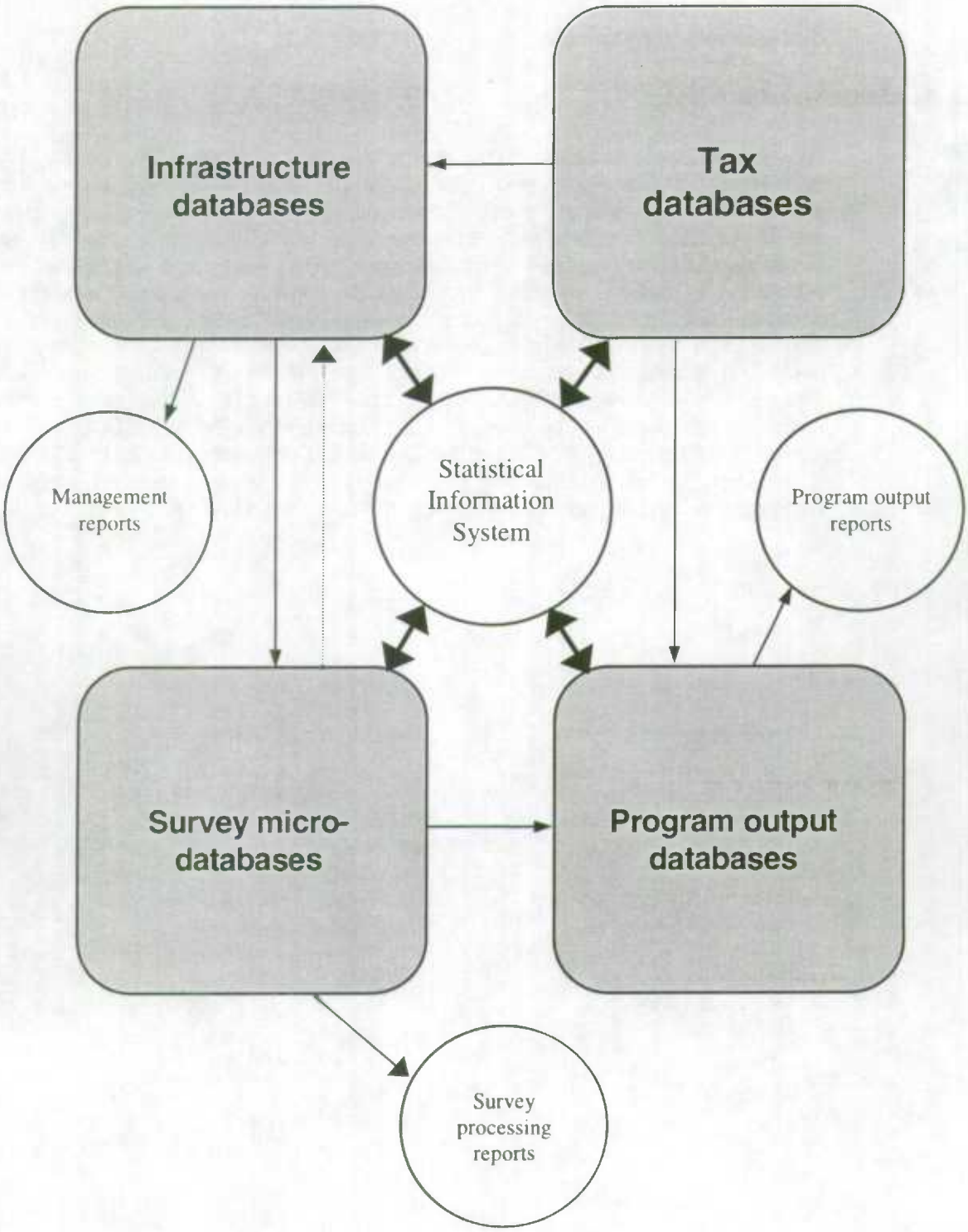

Infrastructure databases
Tax databases
Management reports
Statistical Information System
Program output reports
Survey micro-databases
Program output databases
Survey processing reports

Diagram 11.2

Infrastructure databases

Business Register	CCS, SID	Metadata & quality	Consolidated questionnaire	CMS: SIC, SGC, CPC

Administrative databases

GIFI	T1/T2	T4	GST	PD7	EI	Others

Survey micro-databases

Complex HQ census	Complex P.E.s VA	Complex P.E.s other	Simple other	Sub-annual surveys

Program output databases

BTS division outputs	SNA Outputs	Social outputs	CANSIM II

Administrative databases

The administrative databases represent STC holdings of tax and other externally collected data from a variety of sources. Many of these sources are discussed in detail in Chapter 10. The ability to link these disparate data sets and to access them simply, in an inexpensive and timely manner from the desktop of any authorized user, is an explicit requirement of this plan. Appropriate security policies and procedures, as defined by Statistics Canada, Revenue Canada and other sources, will be put in place. In order to reduce respondent burden and in some instances to improve data quality, tax databases are used in preference to survey information, so tax microdata bases must be available easily and consistently.

Survey micro-databases

At present, individual survey divisions manage their own survey databases, using a variety of database formats and computer hardware and software systems to do so. This makes it very difficult for users in other areas of Statistics Canada to access or to cross-reference the micro-data sets. The UESP will replace these local databases with an integrated set of linked micro-databases, using standard formats and key fields, that can be accessed directly by any authorized subject matter analyst for purposes of building program outputs. Access privileges and feedback arrangements will be defined in conjunction with subject matter areas. The categories of micro-databases shown in Diagram 11.2 correspond to the parts of the UESP discussed in earlier chapters.

Program output databases

Four broad groupings of output databases are shown in Diagram 11.2. Each will be affected in the proposed restructuring of the computing environment. The principal impacts relate to the first two databases. First, there are the subject-matter-related divisional outputs associated with existing BTS programs that must be maintained. There will be a wide range of new data outputs to be developed flowing from more detailed information on commodities, provincial origins and destinations, and class of customer.

Second, there are the existing and to-be-developed outputs from the Systems of National Accounts Branch, particularly related to value added and input-output.

Software integration tools

Operations and systems staff working with the databases will be employing existing software tools from the STC toolbox, but applying them in a more coordinated and inter-related fashion as a consequence of the integrated framework. All new software will be developed using technologies that have been tried and proven in Statistics Canada. In addition to the software to manage the interconnections of the databases across common data structures, there will be a need to develop new software and modify existing tools to perform operations newly required by the integrated data sets. The common data structure and integrated data sets can only be accomplished if they are supported by an approved Information Architecture defined by an active metadata and meta-information processing system. The overall goal is meaningful information which can shared by approved users at any phase of a survey process. Our challenge is to develop a common understanding of our data and processes.

These databases will be inter-connected to enable sharing of relevant data for operational, analytical and dissemination purposes using a database management interface modeled after the Statistical Information System (SIS), but enhanced and adapted to the distributed client-server environment of the UESP.

An integrating framework such as that currently fulfilled by the Statistical Information System is vital to the project. The interface must provide for integrated data storage and access. Its functions must include definition of surveys and variables; loading of, and linkage with survey files; queries by master key elements, covered subjects and statistical identifiers; selective data file extraction in data point, observation and time series formats; and archival and data base maintenance functions. A data base security system for password controlled access by user authorization level and requested function is also an integral component.

In terms of structure SIS is comprised of four distinct modules.

The Classification Maintenance System (CMS) ensures historical consistency. In a static environment bridging sets of information is relatively straightforward. Concordance is achieved through matching common variables and thereafter merging the data sets of interest. When examining diverse survey results over an extended period of time the probability of a change in concepts or coverage becomes greater and therefore progressively more difficult to navigate the longer the period of interest. In the SIS environment this time dimension concordance is achieved through the Classification Maintenance System (CMS). At its most basic level the CMS stores the various classifications used by the loaded surveys including English and French names. A survey might change coding systems across time, however it will use coding systems from the same domain for the same classification purpose. For example, a survey that in the past used SIC-80 and SGC-81 to classify data by industrial activity and geographic location, now might use NAICS and SGC1986 for the same purpose. For each coding system applicable to specific data points or surveys a link to the corresponding CMS category and concordant coding systems values is generated. An automated concordance system ensures that disruptions due to changing codes are minimized and continuity is maintained over time. The CMS will need to be updated to handle the coming conversion to NAICS.

Historical multiple versions of individual data points are an important consideration and must be maintained for purposes of data continuity. The immediate requirement is related to backcasting selected data series and was described in some detail in Chapter 9.

A data interchange facility that permits the extraction, processing of data and related definitions into suitable SAS, Excel, or other software is also required. It should be stressed that this moves data to the analytical tools in several steps, from extraction through conversions, definitions, and loading. In an ideal setup the data analysis and presentation products would access the database directly. A common look and feel in the form of a Windows-type graphical user interface (GUI) is needed. The guidelines identified in the departmental Information Technology Framework provide a basis for determining the basic needs.

Analysts using the system will be equipped with a user-friendly Windows screen driven by powerful classifications-aided navigational tools and other thematic search tools. The aim is to enable authorized staff to query the distributed databases and to specify and assemble micro-data or aggregated data, quality indicators, view different time-stamped versions of the same estimate, and other attributes of interest, in multi-dimensional table views which include geography and time. It would include the ability to progressively view, virtually instantly, more detailed information along any of the axes presented. An integrated graphics capability would likely be desirable. There could be a multi-media outputs dimension for table, graphics and text objects to be produced in various electronic formats such as SGML, HTML and Acrobat.

Producing a functioning networked system along the lines described will be a major challenge to stakeholders in the process but it must be accomplished for the PIPES project to achieve its goals, and for the agency to realize the other major improvements which clients, respondents and the public will expect.

Next Steps – The Short-term Strategy

The short-term UESP strategy will focus first on conducting the "mini-UESP" in the first year and on launching the full UESP the following year. The first-year UESP will use existing systems to the greatest extent possible. The first year system will be an interim system, functional but limited and inelegant. It will incorporate elements of the long-term system where possible but elements of it will be ad hoc and temporary. For the full UESP in the second year, a more permanent and efficient system will be built and tested. This new system will be further refined in the years to follow.

Fortunately Statistics Canada has much experience in developing new computer systems to meet the requirements of survey programs. It has an excellent pool of highly qualified informatics professionals and it is they who will be depended upon to work together with subject matter and service area staff to deliver the new systems required for the UESP. With little more than a year's worth of development time available for many of the interim systems, subject-matter and service area specification work and the information technology framework must begin immediately. This framework will define the building blocks that will address the problems related to the construction of the network of hardware, software and people.

Many new CS professionals will be hired over the coming two years. Some of these new employees will undoubtedly make a valuable contribution to the project. However, it will be the more experienced staff already on-site and working on core programs that will play the key role. PIPES will not succeed unless a large number of these key people are diverted to work on the project. Inevitably, this will mean that development work on existing systems will suffer, but that cannot be helped. In the long run, all BTS programs will benefit greatly from the investment.

It will be one of the most ambitious systems development projects ever undertaken by Statistics Canada.

Chapter

12

Implementation issues and timing

The principal stimulus for launching the Unified Enterprise Statistics Program is to satisfy commitments to Finance Canada and the provincial governments for better quality data with respect to provincial economic accounts and input/output tables. These obligations involve improving provincial data starting in reference year 1997. Meeting this commitment would be difficult enough if the approach taken was simply to increase sampling rates and improve coverage by initiating new surveys for missing industry or commodity data. However, having chosen - for very good reasons - to restructure the core program to make it more centred on the enterprise necessitates the development of new infrastructure and makes the task more complex in terms of planning and implementation timing.

This chapter describes the current plan (as of March 1997) for implementing the UESP. It lays out some indicative dates to put into perspective the key implementation issues. A more rigorous critical path will be developed during the more detailed planning stage that lies immediately ahead.

To arrive at the indicative milestones, some broad working assumptions have been made:

- The program will undertake to disseminate better quality data for reference year 1997 toward the end of calendar year 1998. This will involve the development of various computer systems that will allow the UESP staff to carry out their operations while facilitating development and linkages of the relational databases that will be used to access the data for compilation, analysis and dissemination in the responsible subject matter divisions.

- While survey-taking activity will shift to the new integrated survey vehicles, the responsibility for subsequent data development, analysis and dissemination will remain fully with the subject matter divisions. The transition to the new framework will require a great deal of cooperation, collaboration and team effort. Divisional sign-offs will be needed at various stages.

- The program will be as much as possible enterprise-centred. However, it will be constrained by respondent preferences with respect to data provision and by resource constraints during the developmental stage. For these reasons it will not be possible to move 100% to the new approach in the first year. The plan foresees a phased pace of implementation.

The planning horizon in perspective

The UESP will be implemented over the next three years. In 1998 for the 1997 reference year, a "mini-UESP" will be conducted consisting of the Census of Complex Enterprises (Part I) and a prototype test of the other two parts of the UESP for a small number of industries where survey programs do not now exist. PIPES objectives will also be addressed in the first year by some enhancements to existing surveys.

The annual Census of Complex Enterprises will be based on the integrated BR frame. The content of the CCE will be based on the existing IOFD survey but augmented to include: (1) sufficient income statement detail to calculate enterprise value added, (2) enterprise capital expenditures and planned expenditures, (3) enterprise employment and (4) profiling information. In 1998 for the 1997 reference year, we will also implement (1) a new unified survey (or census) of establishments of complex enterprises and (2) a new unified survey of simple enterprises. These two components of the UESP will be implemented for a selected set of industries only. The specific list of industries to be covered in the mini-UESP is construction, couriers, taxis, real estate agents, real estate operators, restaurants caterers and tavems, and aquaculture. In 1999, for the 1998 reference year, we will expand the UESP to cover all remaining industries. This means that agriculture, mining, fishing, forestry, manufacturing, services, financial services, transportation and communications will be brought into the survey. Mining will have to be considered in relation to current working arrangements with Natural Resources Canada and the provincial mining departments.

In the year 2000 we will complete the process by expanding the Unified Enterprise Survey Program to include where appropriate sub-annual business surveys as well as annual ones.

To realize these goals, the following supporting goals will also have to be achieved.

Once the detailed plans are completed and implementation work has begun, the organization can discuss the UESP in detail with its stakeholders (other departments, provincial govemments, etc.) by means of a conference and/or other consultation mechanisms. This is planned for May 6 1997, early enough to be able to make changes if so required, but late enough to allow sufficient time to develop a detailed and well thought out plan.

The integrated Business Register, combining the information in the IOFD frame with that of the BR, must be ready by the fall of 1997. This integrated register must be expanded to reflect the information in the farm register as well, by the fall of 1998. The BR must be expanded to include significant non-employer businesses and the quality of the BR frame for the 1998 industries must be assessed. The BR must also implement NAICS codes and begin the NBR process by July 1997.

The current "survey impact database" and "company contact system" must be enhanced or replaced and be operational by early 1998, to track total response burden by enterprise. By early 1998, the system must have comprehensive coverage of the 1998 industries. By early 1999 it must have comprehensive coverage of all industries. In both cases, "comprehensive coverage" includes sub-annual as well as annual surveys, and non-BTS surveys (e.g. BOP or SEPH) as well as BTS surveys. The system must be efficient from an informatics point of view, and it must be supported by clearly articulated corporate policies requiring all survey contacts to be registered in the system.

Clean tax databases must be developed and made easily available on-line for the T1-GIFI, T2-GIFI, GST and PD7 sources. In the first year, plans must be developed to capture a sample of tax data since GIFI data will not be available for the entire 1997 tax year.

A new integrated questionnaire must be conceived, that integrates all existing business questionnaires, avoids duplication, enforces consistent terminology and limits the amount of information requested by recognizing what is already available from administrative sources.

Methodologies and computer systems must be developed and put in place within one year that will facilitate:

- Storing and accessing microdata from the unified surveys and the tax databanks;
- Capturing, editing, imputing, estimating and aggregating the survey data that are collected;
- Generating customized, consolidated questionnaires for complex enterprises;
- Interfacing easily with the BR from the desktop;
- Implementing the "survey impact database" or "company contact system".

The "key provider manager" program must be put into place, involving the hiring and training of 7 senior company contact people and the commencement of discussions with 50 selected enterprises.

Some Key Issues

There will be many issues to resolve as the UESP is implemented. Feedback to the draft proposal (December 20, 1996 paper) has helped to identify and elaborate some of the most significant challenges we will face.

1. There will be a close relationship of the IOFD quarterly financial survey to the CCE and many details on the relationship between these two surveys in terms of survey content, sample design, editing, and imputations need to be worked out.

2. The implementation of the UESP on a limited basis in 1998 will create some coordination issues vis-à-vis existing survey programs. For example, construction enterprises may have some non-construction establishments already surveyed.

3. The use of tax estimates in relation to survey estimates raises issues of consistency of frames, especially in the short term. The Business Number is expected to reduce the importance of these issues in the longer term by facilitating the linkage of GIFI data to the BR.

4. Many elements of the UESP need to be tested before a final detailed vision can be formulated. Some of the major elements requiring testing are the "spreadsheet" questionnaire, the KPM approach, new data elements such as class of customer, the common sampling approach, and the many new systems and processes.

5. The completion of NAICS implementation on the BR is scheduled for March 1999. This creates transition issues related to sampling for the first two years of the UESP.

6. The BR faces many transition issues (in addition to NAICS conversion) in response to the UESP. These include: changing processes to focus on complex and simple enterprises rather than IP/NIP, expanding the BR to incorporate the BN source, reconciling the IOFD enterprise frame and the Farm Register, hooking up remaining sector surveys to the BR, creating SUFs for the first year UESP, the development of a comprehensive company contact system and finding methods of detecting newly complex enterprises.

7. The UESP integrated questionnaire raises issues of harmonization of concepts, adherence to standards and integration of data.

8. The roles of subject matter, infrastructure and service divisions will have to be carefully developed.

Some of these issues have been continuing challenges in the existing survey approach. Others arise from the shift to the UESP either as transition issues or long term challenges to be overcome.

In summary, some key milestone planning dates for the coming two years are:

1. Completion of detailed planning ..... March 1997
2. Begin field test of the new enterprise HQ questionnaire ..... April 1997
3. Formally release this paper ..... April 1997
4. Discuss plan with key stakeholders ..... May 1997
5. Start implementing "key provider managers" ..... May 1997
6. Implement mass conversion NAICS codes on BR ..... July 1997
7. Complete IOFD-Business Register Reconciliation ..... October 1997
8. Finalize Consolidated questionnaire and sample selection for pilot-UES ..... December 1997
9. Complete computer design - front end ..... December 1997
10. Implement company contact system ..... January 1998
11. Initiate mail-out ..... March 1998
12. GIFI data for corporations starts to become available ..... April 1998
13. Complete computer design - back end ..... July 1998
14. Tax databases implemented ..... July 1998
15. Survey data ready for compilation /analysis /dissemination ..... Fall 1998

Chapter 13

Conclusion

The HST allocation formula will use Statistics Canada Input-Output and Income and Expenditure data at the provincial level. To do so, broadly based improvements in the quality of provincial data are needed. The principal objective of the Unified Enterprise Statistics Program is comprehensive quality improvement, as explained in this paper. The UESP involves fundamental changes in processes and therefore will have a major impact on the organization of work in Statistics Canada.

Organizational Impact

The UESP will fundamentally change the way Statistics Canada conducts business surveys. It is a further step in the direction of organizing the work of data gathering along functional rather than subject matter lines.

Up to the 1970's, business surveys were conducted virtually entirely as individual surveys, with all processes from determining data content to disseminating estimates controlled and performed by one subject matter unit. Increasingly, control of some survey processes has been transferred in varying degrees to infrastructure divisions. Standards Division has created a number of classification and other standards that influence the data content of business surveys leading to greater harmonization of concepts and adherence to common standards. The business register has been increasingly the source of business survey frames. The roles of methodologists in a centralized Business Survey Methods Division has led to increased commonality of business survey design approaches and data quality measurement. The development of a standard set of generalized survey processing software tools has moved the Bureau some way toward a standard processing system. The creation of Operations and Integration Division in the 1980's created a central area responsible for data collection, capture and collection edits, in addition to the Regional Offices survey collection activities.

The UESP will extend these trends by further integrating the core data collection activities. Subject matter divisions will become increasingly focused on managing statistical programs, making estimates from a set of infrastructure databases from survey and administrative sources. The principal functions of subject matter divisions will be in the areas of determination of data content, tabulation of estimates, data analysis, respondent relations, product development, data dissemination and client relations. Infrastructure divisions will have the principal responsibility for core survey and administrative data processes such as frame development and maintenance, sampling and estimation, data collection and capture edits, statistical edits, imputation, calendarization, assessment of data quality and database maintenance. Subject matter divisions will of course heavily influence many of these processes.

Advantages of the Organization Changes

There are significant potential advantages from the organizational changes arising from the UESP.

1. The enterprise-focused approach to data collection will permit much greater potential for data integration. The ability to analyze the entire range of data for particular business enterprises will lead to greater internal data coherence as well as a rich source for addressing more complex analytical issues.

2. The shift in focus of subject matter divisions to managing statistical programs rather than managing survey programs will lead to better and more extended analysis, more interaction with clients and more relevant products and services.

3. The increasing specialization of infrastructure divisions on specific data collection functions will lead to greater effectiveness. Best practices can be extended across the board and ineffective or inefficient practices can be replaced on the basis of economy-wide experience of what tends to work well for business surveys.

The UESP will of course have major transitional challenges. Chief among them will be to implement the UESP in a way that maintains most of the advantages of direct subject matter participation in data collection activities through appropriate means of influencing these activities while permitting the advantages of greater specialization in roles.

Long Term Impact of PIPES and UESP

PIPES presents major opportunities to increase the usefulness of the business statistics program to users. Over the past decade and more, budget restrictions led to selective cuts to the business statistics programs. PIPES will restore the coverage of the business statistics program to areas formerly covered such as construction and extend coverage by filling gaps in the measurement of important sectors such as business services. Provincial input-output tables will be compiled annually. Commodity level production estimates will be extended to sectors where commodities have rarely or never been measured before.

These and other data enhancements arising from PIPES will be useful far beyond the allocation of HST revenues.

Provincial Input-Output tables can be used in a wide variety of policy analysis applications at the regional level. Some examples are: impact analyses of events such as the Saguenay floods, the PEI-NB bridge, the price of prairie grain, U.S. quotas on lumber, a strike in the automobile industry or a proposal to host the Olympic games in a specific province. New light will be shed on long standing issues such as regional underdevelopment, inter-provincial trade barriers and labour immobility. The I-O tables will add to the understanding of technological diversity across Canada and the reasons for differential rates of technological progress by region. I-O tables also provide a cross-checking framework to improve the quality of provincial economic statistics.

The services economy will be measured in a comprehensive manner at the provincial level. Small business dominates in the services sector and its concerns and special problems can be given closer attention with better service sector data. Service industries

tend to be highly dynamic. Changes will be measured more effectively and in greater detail than ever before.

The area of distributive trades will receive special attention. These statistics will provide valuable information on trends in consumer demand by province and will be especially helpful in analyses of the regional impact of the business cycle.

Both public and private sector users will benefit. Business will find the increased data detail useful in connection with marketing, production and investment planning. Increased usefulness of Statistics Canada data to private sector users is particularly important as a factor in maintaining and improving response rates. As well the potential to produce sub-provincial data will increase dramatically due to larger survey samples and comprehensive administrative data due to the GIFI and the Business Number initiatives. Sub-provincial data are in great demand by provincial governments and private sector users.

The UESP will realize added benefits, including greater data integration from increased standardization of data content and a high degree of data consistency. There will be more equal data quality across provinces, within a framework that maximizes use of administrative data and minimizes response burden.

The UESP shifts the focus of the statistical program to the enterprise, which is increasingly the level at which important decisions are made. The enterprise is a better basis for analysis of globalization issues. As well, the UESP will permit improved modeling of the firm with valid linkage of financial and production data.

When the GIFI and UESP are fully implemented, a census database of enterprise-level income statement and balance sheet data and establishment-level value added data will yield unprecedented capacity for detailed distribution and geographical analysis.

There are powerful new possibilities offered from developing computing technologies. The UESP will use new technologies such as client-server computing, relational databases, corporate access to databases, the Internet and other developing electronic data collection methods. These technologies will empower the individual analyst to extend the scope and range of statistical analysis at Statistics Canada like never before.

Data quality measurement and improvement is fundamental to the UESP strategy. Data quality is multi-faceted and cannot be measured as a single number. Therefore, a vector of quality indicators will be developed across the business statistics program to monitor quality change over time in a comprehensive and consistent manner. The UESP will foster an environment of continuous improvement to data quality, based on consistently applied and measured quality criteria.

The UESP will increase program flexibility, such as reacting to respondents' demands for customized treatment, modifying questionnaire content, changing data edit specifications, applying systems or methods improvements and exploring micro-analytical and cross-cutting issues.

Statistics Canada is committed to bringing about the expected benefits of PIPES and the UESP. The result will be a program of national and provincial economic statistics that will be among the very best in the world.

Annex 1

Glossary of Terms

Business Number (BN)	Starting in January 1997, all businesses required to participate in one (or more) of the following tax programs: Payroll Deduction, Corporate Income Tax, Goods and Services Tax and Importers and Exporters Tax, will have a single registration number with Revenue Canada., called the Business Number or BN.
Business Entity	A Business Entity represents an economic transactor having the responsibility and the authority to allocate resources in the production of goods and services, thereby directing and managing the receipt and disposition of income, the accumulation of property, borrowing and lending, and maintaining complete financial statements accounting for these responsibilities.
Business Register (BR)	A database consisting of a complete list of all statistical enterprises in a defined universe, plus a list of their associated legal and production entities, classified into industrial categories based on their principal business activities, and the hierarchical relationships among these entities, plus addresses, telephone numbers and other contact information necessary for conducting surveys.
CANSIM	Canadian Socio-Economic Information Management System, Statistics Canada's electronic database or repository, and retrieval system, for Time Series and other data.
CANSIM II	The proposed replacement for CANSIM which is currently under development.
CCS	Company Contact System maintained by the Business Register provides selected response burden information for those surveys drawing a frame from the Central Frame Data Base (CFDB).
CFDB	Central Frame Data Base of the Business Register.
CMS	Classification Maintenance System which was developed by the Prices Division to track concordances between classification systems.

CPC	Central Product Classification is an International standard for the classification of goods and services. Classes for goods are aggregates of classes in the HS.
CQ	Consolidated Questionnaire represents an omnibus approach to collecting survey data while recognizing varying degrees of business complexity
CSNA	Canadian System of National Accounts
Company	See Statistical company.
Corporation	A legal entity.
Distributed PC-server computer environment	A network of computers organized in a dispersed fashion throughout an organization using a Local Area Network (or Wide Area Network) in which there is a client-server relationship between the devices in the sense that one device (the "server") is a repository of information from which the other devices ("the clients") obtain copies of all or portions of the data requested using data communications protocols (software). The "server" may be a seemingly ordinary PC or more often it is a more powerful PC which has added storage capacity and has been set up (optimized) for faster communications and internal processing of instructions (higher speed rating in Hertz).
Enterprise	See Global enterprise and Statistical enterprise.
Establishment	As defined in the 1993 System of National Accounts, this is an enterprise, or part of an enterprise, that is situated in a single location and in which only a single non-ancillary production activity is carried out, or in which the principal productive activity accounts for most of the value added. The management of an establishment normally has significant short-term discretion over production and employment decisions, but may have little or no discretion over financial management, investment planning or other broad strategic matters. While establishments are situated in a single location, that location might be, say, a province, and there might be several physical location units within the definition of that broader location. See also Statistical establishment.
General Index of Financial Information (GIFI)	The GIFI is a list of standard accounting variables that defines a standard set of Income Statement and Balance Sheet accounts that corporations and individuals will be

	required to use (starting April 1998) when reporting the results of their business activities to Revenue Canada.
Global Enterprise	Two or more statistical enterprises assembled together under a common control group but operating independently.
Goods and Services Tax (GST)	A federal tax on goods and services sold which is structured so as to be a value added tax on final consumption.
HS or Harmonized System	Harmonized Commodity Description and Coding System is an international standard for the classification of goods.
ICN	Statistics Canada's Internal Communications Network, similar to the Intranet but using different software and usually offering more of a broadcasting function (one-way communications) than a two-way communications capacity inherent in Internet/Intranet.
Institutional unit	As defined in the 1993 System of National Accounts, this is an economic entity that is capable, in its own right, of owning assets, incurring liabilities and engaging in economic activities and transactions with other entities.
Intranet	The use of a LAN (or WAN) and specialized software to provide data communications (but may include data, audio, video on an interactive basis) within the boundaries of an organization from specific usually more powerful Personal Computers called "servers", this being a variation on Internet which provides the same services over the public switched telephone system.
Legal Entity (LE)	An incorporated business or company, technically defined as an entity granted by law the power to own assets, borrow resources and transact business (buy and sell goods and/or services).
LAN	Local Area Network, a relatively small grouping of personal computers (PC) served by specialized communications software, usually also confined to a relatively small geographic area, whereas a WAN is a Wide Area Network in which the distribution is much larger either in terms of geography or numbers of PC's attached together.
LESP	Large Enterprise Statistics Project
Metadata	Information or data about the data or information in question
MIDAS	Meta Inventory of Data Assets is a corporate listing of STC data sets

NAICS	North American Industrial Classification System is to be implemented in 1997 as the North American and Canadian standard for the classification of statistical establishments by industry. It is to be used for the classification of statistical locations by activity and for the classification of statistical companies and statistical enterprises by industry. Contains a level for Canadian detailed classes. Replaces both the SIC-E and the SIC-C.
NBR	Nature of Business Report is a series of forms used by the Business Register Division to determine the industrial classification of a business.
Payroll Deduction (PD) Account	PD accounts are established by employers to remit to RCT employer and employee contributions towards Canada/Quebec Pension Plan and Unemployment Insurance contributions, as well as source deductions for income tax.
PCG	Principal Commodity Groups are aggregations of SCG classes which are used in the publications of the Input-Output Division
PIPES	Project to Improve Provincial Economic Statistics
Production entity	The smallest distinct component unit of an enterprise, consisting of one or more physical locations, with resources (labour and capital) dedicated to the production of a specific set of goods and/or services. Dedicated means, directly assigned to the production of the goods and services of a particular place of business" Dedicated Labour is human resources directly assigned to the process of the production entity. These may be composed of either or any combination of full or part time employees, contract labour or volunteers. Dedicated Capital is a combination of cash, land and buildings, plant, machinery and equipment, or inventories of raw materials. The economic resource land represents the natural resources available such as, land, plants, streams, fish, forests, minerals and so on. Process is said to take place when resources such as capital, labour and entrepreneurship are combined to produce a set of specific goods and or services. Dedicated Manager to provide on site management to manage the 'Process' on a daily basis.
RCT	Revenue Canada Taxation

SCG	Standard Classification of Goods is the Canadian standard for the classification of goods. An extension of the HS, it is more detailed than the HS.
SCS	Standard Classification of Services is the Canadian standard for the classification of services. (It is currently under development.)
SCTG	Standard Classification of Transportable Goods is the Canadian standard for the classification of transportable goods. Based on the HS and the SCG, it is intended for use in water, rail and truck transport surveys.
SIC-C	1980 Standard Industrial Classification for Companies and Enterprises is the Canadian standard for the classification of statistical companies and statistical enterprises by industry. (In 1997 it will be replaced by NAICS.)
SIC-E	The Standard Industrial Classification for Establishments which was developed in 1980. It is the Canadian standard for the classification of statistical establishments by industry. Also used for the classification of statistical locations by activity. (In 1997 it will be replaced by NAICS.)
SDDS	Statistical Data Documentation System describes statistical activities undertaken by STC.
SID	Survey Impact Database describes in considerable detail survey contacts for the 270 businesses under the purview of the Large Enterprise Statistics Program. It comprises information by survey on completion times, contact history and coverage as well as an internal bulletin board for the free exchange of latest corporate developments.
SIS	The Statistical Information System is a multi survey integrated database comprising micro and aggregate data.
STC toolbox	A suite of software tools in use or being evaluated for use in Statistics Canada, and which will be corporately supported in various ways, and to specifically defined service levels.
STATGEN	A "fully automated process" used to standardize the delineation and classification of the four levels of statistical entities (enterprise, company, establishment and location). It is based on the structure and data elements within the operating structure of a business entity.

Statistical Enterprise	The complete organizational unit of a business entity a) whose activities are as industrially homogeneous as possible; and b) that independently direct and controls the allocation of resources and economic activities relating to operations in Canada; and c) for which accounting records provide a complete set of financial accounts from which international transactions, international investment position and a consolidated financial position can be derived, as well as all consolidated elements of revenues and expenses required to measure profit.
Statistical Company	The smallest organizational unit a) whose activities are as industrially homogeneous as possible; and b) for which revenue and expense elements are available that allow the calculation of an operating profit; and c) which has assets and liability elements to measure capital employed in the production of goods and/or services. Management normally has significant short- and medium-term discretion over production and employment levels, and possibly investment decisions, but little or no discretion over financial management or other broad strategic matters.
Statistical Establishment	One production entity or the smallest grouping of production entities a) which produces as homogeneous a set of goods and/or services as possible; and b) which does not cross provincial boundaries; and c) for which records provide data on the value of output together with the cost of principal intermediate inputs used and cost and quantity of labour resources used to produce the output.
Statistical Location	An operating entity, specifically a production entity a) conducting economic activity at or from a single physical location or group of locations; b) within the smallest standardized geographical area; and c) for which, at a minimum, employment data are available. At the location level, management may have little or no significant short- or long-term discretion over production and employment levels.
T1 (Individual) Income Tax Return	A T1 return is filed annually by individuals who are reporting income any source, including business, profession, or a farm. Of the 15 million T1 forms filed in any given year, approximately 2 million report income from one of these sources.
T2 (Corporate) Income Tax Return	A T2 return is filed annually by corporations who report income from any source. There are approximately 1 million T2 forms filed on an annual basis.
UESP	Unified Enterprise Statistics Program

Annex

2

Data Requirements and Priorities for PIPES from the System of National Accounts

Background

In the fall of 1996 a committee was struck to evaluate PIPES proposals and assign priorities to provide input to the annual planning process at Statistics Canada. Members of this committee were Yusuf Siddiqi (Chair), Philip Smith, Ross Vani, Dave McDowell, Michel Vallières, Richard Martel, Hans Messinger, Roger Jullion, Steve O'Brien and Dave Leblanc (Secretary). This process was subsequently replaced by the PIPES secretariat and the project management structure of the UESP. It was felt that there was still a need for a statement of requirements and priorities from the viewpoint of the System of National Accounts (SNA). This paper was drafted by Yusuf Siddiqi with comments from member of the original committee.

This paper outlines some basic criteria for determining the requirements of a complete set of Provincial Accounts (Provincial Input-Output Accounts and Provincial Income and Expenditure Accounts. This is followed by a short discussion of data collection issues related to these needs. The main body of the paper contains a list of data needs with priorities assigned to each, for each of the major industry groups in the Provincial Input-Output Accounts.

Criteria for Data Requirements

The Revenue Allocation Formula under the HST specifies the following tax bases: consumer expenditures, residential construction (consisting of new housing construction, alterations and improvements, and transfer costs), financial institutions, public sector bodies (MUSH[4]) and providers of other exempt commodities[5] Therefore, data improvements directly affecting these areas must be assigned a higher priority (see discussion on p. 122).

Economic accounts, including input-output tables, present an interdependent system of accounting. Therefore, requirements for data improvements and coverage are assessed within this context. Though the need for some source data or improvement may not be immediately apparent, it may have a significant impact on GST-generating commodities. SNA's data requirements for PIPES are looked at from a perspective, which confronts the supply and demand sides of the equation. Then GST and, for that matter all other commodity taxes, must be estimated and balanced. For this purpose, absolute levels of various data sources are as important as are movements from period to period.

[4] These public sector bodies, known as the MUSH sector, consist of Municipal Governments, Universities, Schools, Hospitals and Residential Care Facilities, Non-profit and Charitable Organizations, and Urban Transit.

[5] Producers of other exempt commodities are financial institutions (including banks, credit unions, etc.), the residential rent industry, producers of health care (e.g., physicians, dentists, chiropractors, etc.), unions, the child care industry, the ferry, road and bridge tolls, and legal aid.

The required improvements in the provincial dimension of the source data follow from the principles contained in the PIPES plan put forward by the Bureau, namely, that of "greatly improved annual provincial economic accounts estimates"[6]. The required improvements help ensure that the quality of provincial estimates is comparable across provinces and this has direct implications for the way in which surveys are conducted.

For input-output accounts, annual data for the reference year 1997 must be delivered by March 1999, at the latest. Lags for reference year 1998 and beyond must not exceed 14 to 15 months from the reference year. However, National Accounts and Environment Division will incorporate improvements in advance of I/O tables publishing preliminary estimates of reference year 1997 in May/June 1998 and preliminary estimates of reference year 1998 in May of 1999 etc. To be able to effect these improvements in NAED's more current provincial Income and Expenditure Accounts, data would be required by April/May 1998 implying some improvements to sub-annual data or P13 data. Data for subsequent years should follow the same lag although the steady state schedule for publication of the preliminary estimates has not yet been finalized and may be changed to align with the flow of data from the new annual survey system proposed by UESP.

For the SNA production accounts, data is generally required on an *establishment basis*. The establishment as a statistical unit may cover secondary activities in addition to its principal activity. The establishment concept, as defined in the 1980 SIC, may not be applicable to all industries. In such cases, the definition of an establishment should be approximately met. In all cases, data ***must be additive across industries***. Since the data are required for ten provinces and two (eventually three) territories, the establishment unit and ancillary activities must not cross provincial or territorial boundaries (see page 25 "Standard Industrial Classification 1980").

The annual input-output (I/O) tables for ten provinces and two territories are balanced with 254 industries and 674 commodities (goods and services). At present, the industry classification in the I/O tables is based on the 1980 SIC but, starting with the reference year 1997, it will be based on the NAICS. The I/O classification of goods is fully concorded with the Standard Classification of Goods (SCG). In the absence of a Canadian standard classification of services, I/O Division uses its own classification of services, which are defined in terms of characteristic products of the industry. For instance, the output of the accounting industry is called accounting services. Where a breakdown of several characteristic products is available, as in the case of the computer industry, the output has been broken into computer lease and rental services, software products, professional and processing services, repairs and margins on goods purchased for resale (GPRS). Data availability has been a limiting factor in industry and commodity classification of services industries.

From 1997 onward, national and provincial input-output tables will be based on the NAICS, so that all basic data from subject matter divisions in Business and Trade Statistics and Social Institution and Labour Statistics Fields will be required on the NAICS basis. In terms of industries, as a minimum, we must receive commodity output, input and capital formation data for each four-digit NAICS industry. A parallel data run should be provided both on the NAICS and the 1980 SIC bases for the reference year 1997. The 1997 data on the NAICS basis is expected to be backcasted to 1992, which means that

[6] See "A Proposal for a Unified Enterprise Statistics Program", Statistics Canada, December, 1996.

the SNA could potentially receive all basic data sets on two classification bases. For time-series analysis, I/O Division will create a historical link series, which will extend back to 1981.

In production accounts, *data needed to calculate value added is not the be-all and end-all.* proper commodity detail is required for outputs, intermediate inputs of goods and services and components of industry value-added. As a general rule, for ***production or output***, data are required for each industry, detailed by commodity. The manufacturing industry, for example, should be accorded special treatment so far as the data on commodity details of inputs and outputs are concerned. The present SCG level of commodity detail should be used for data collection using the long form approach. Only through this approach will technological changes (commodity inputs related to commodity outputs) in the industry be captured including the emergence of new commodities. For goods-producing industries (except wholesale and retail trade) production is defined in terms of shipments adjusted for inventories. For services-producing industries, the output is equal to revenue derived from sale of services, be it amusement services, computer rental, etc. For wholesale and retail trading activity, *production* is defined as sales less cost of sales. Hence, for each kind-of-business in the trade area, information is required on sales and cost of sales together with details on revenue from secondary activities, such as revenue from leasing, repair service, etc. In addition, the commodity composition of sales is essential for calculation of consumer expenditures and trade margins. For goods, we require at least six-digit details of SCG. Prices Division might require more detail for adequate pricing. For services, the International Product Classification could be used as a guide. On the ***use or input*** side, data are required on cost of materials, fuel, electricity, other utilities, purchased services and property taxes in full commodity detail (SCG level for goods). Information is required on operating expenses of head offices and ancillary units with a commodity detail comparable to that of production units.

On the revenue side, for each industry, information is required on sales by class of customer. The class of customer data should be at least in terms of sales to households, governments, businesses and exports. Wherever feasible, sales to the business sector of the SNA should be broken out by industry.

Finally, subject matter divisions should devote more resources to the analysis of the data both at the national and provincial levels. This process, along with a close interaction with the SNA, would promote micro and macro data integration and improvement.

Data Collection Problems

In some cases, different data sets are available at different levels of the enterprise hierarchy where reporting units and statistical units are not the same. For our establishment-based production accounts, some data may be obtained from the establishment directly and some from the company (which may be an aggregation of multi-industry establishments) when these variables are available only at higher levels of the company hierarchy. For example, an establishment of a complex company is usually able to provide information on detailed commodity production, shipments, inventories, purchases of raw materials, fuel and electricity, employment and wages and salaries but may not be able to provide data on purchased services. These services (like advertising,

travel, professional fees, etc.) are, in most cases, purchased at the company level on behalf of constituent establishments. In such cases, the data for these items can be collected at the company level and an appropriate variable can be used to allocate purchased services at the company level to constituent establishments. The same principle should be applied to head office and/or ancillary activities, which may be serving diverse establishments, say, in mining or manufacturing for an integrated company.

Provincialization of the 1990 I/O tables revealed that data of multi-provincial corporations are a major complication in compiling estimates for provinces. *For economic accounts, the value added for an enterprise* (adjusted for National Accounts) *must be translated into a set of establishments' value added by SIC and by province.* To accomplish this, data would be required on

(i) Services purchased by a complex enterprise for itself and on behalf of its legal entities, or establishments;

(ii) Administrative expenses of the enterprise;

(iii) Services purchased by constituent legal entities and their administrative expenses.

All these expenses must be allocated back to constituent establishments to derive establishment-based true value added. However, a reliable commodity detail for these expenses is equally important, because industries' demand for each commodity must balance with supply.

A complete profile of all the statistical entities of complex enterprises should be provided separately, together with data for each of the head offices and ancillary units serving constituent establishments classified to different industries in the same province, or in the same industry but located in different provinces. This will require a well-maintained Business Register. More work needs to be done in the service producing industries, as the survey universe of services has not been firmed up compared to goods producing industries.

Specification of Requirements and Priorities

In this section, items for which data are required to improve provincial economic accounts are presented and ranked in order of priority. Priorities are ranked on a scale of 1 to 3 as follows:

1 **Critical Impact** expected: item affects personal expenditure estimates or estimates for industries subject to GST.

2 Significant **Impact** expected on the entire system of estimates.

3 **Data** will be **useful**.

All source data discussed here are needed by province and on a calendar year basis even if not so specified.

While items with a ranking of 1 have first priority, the following items take precedence over the rest due to their direct bearing on the Revenue Allocation Formula and merit immediate attention. In addition to their ranking of 1, they are identified with an asterisk in the text.

- Retail and Wholesale trade with a true provincial breakdown of sales and commodity content of sales and class of customer information;

- New motor vehicle sales also covering fleet sales, lease revenue, value of vehicles at termination of lease, inventory for lease, sale of used vehicles and revenue from parts, repairs and body work;

- Direct sales,

- New housing sales (including the value of land), by type of structure and value, regardless of which industry has produced it,

- Alterations and improvements of housing,

- Improvements to the Home Repair and Renovation Survey and the Landlord Repair and Renovation Survey, as present provincial CV's are very large,

- FAMEX,

- Real estate commissions separately for residential and non-residential properties and land and other transfer costs such as property assessment and inspection fees,

- Revenues and expenses of landlords,

- Telecommunication services including resellers: provincial estimates of revenue for telephone service, cablevision (including pay and specialty TV), cellular telephone services, other non-telephone services such as pager, fax, and internet services,

- Accommodation and food service, with class of customer information by origin (residence),

- Canadian Travel Survey and International Travel Survey.

Data Requirements by Industry

Agriculture and Fishing

- 2 At present, Agriculture Division does not cover agriculture activity in the Yukon and NWT.

- 2 The Agriculture Division is responsible for measuring aquaculture which is partially being covered by Fisheries and Oceans Canada.

- 2 The data on fish landings comes from Fisheries and Oceans Canada but there are no data on inputs.

- 2 Destination of shipments for agricultural products;

- 2 Destination of shipments for fish and seafood products;

- 3 Services incidental to agriculture should be measured;

- 3 Improved estimates of crop and livestock sector.

Forestry and Manufacturing

- 1 Data on purchased services are not at present collected. This is essential for calculating true value added. Purchased services usually represent 20% of value added. It should be ensured that expenses of captive ancillary activities are included with the relevant industries.

- 1 The class of customer data are required in terms of sales to households, wholesale, retail (however defined), other businesses, governments, and exports. This information is required separately for own-manufacture and goods purchased for resale.

- 1 The Survey of Destination of Shipments should be improved to include commodity-specific details, preferably at the same level as that of shipments. The commodity detail for destination of shipments may be collected at less frequent intervals than shipments.

- 2 Destination of shipments for forestry products to estimate inter-provincial movements of logs;

- 2 The Survey of Destination of Shipments should include goods purchased for resale (GPRS) by commodity.

Mining, Quarrying and Oil Wells

- 1 Purchased services are not at present collected for mining, quarrying, crude petroleum and natural gas industries.

- 1 Captive ancillary activities must be covered in the relevant industries.

- 2 At present, the 'Services related to mineral extraction' (Major Group 09 of 1980 SIC) is not surveyed. This industry is quite sizable in terms of output and GDP. These variables for 1993 are estimated at $4.2 and $1.9 billion respectively.

- 2 For the Crude Petroleum and Natural Gas industry, reliance is placed on the Net Cash Expenditure Survey which does not provide enough detail annually to enable a reasonable estimate of true value added. The expense data has in the past been collected in cooperation with the Petroleum Monitoring Agency (PMA) which has now been eliminated. Hence there is a need for an annual survey of this industry. This is important as it impacts on estimates of personal expenditure via the intermediate use of services and production of energy commodities. The 1993 output and value added of this industry were $22 billion and $16 billion, respectively.

Electric Power Systems Industry, Gas Distribution Systems Industry

- 1 The data on sales to households does not cover apartments.

- 2 At present "cost of purchased services" and "cost of contract repair and maintenance services" are separately collected for all provincial electric utilities except for those of Ontario. Ontario accounts for 50% of the output of the industry. In addition, no details of service inputs are available from the annual electric power census.

- 2 Gas distribution, expense data is collected on a functional classification basis, which is not compatible with SNA commodities.

- 3 Establishments producing and selling electric power may not be accounted for. According to Industry Division, approximately 150 entities operate power generating stations for own-use and which sell their excess power. The census of manufacturing and mining do not capture this information. This includes the rapidly growing co-generation facilities (combined heat & power).

- 2 Other Utilities: for 1994, 1995 and 1996, NAED has introduced a Waste Management Industry Survey, which is a census for SIC 4999 collecting 15 data items for operating expenses and revenues by province. We understand that funding for continuation of this survey is uncertain after the 1996 data year.

Construction

Final estimates of provincial GST-HST paid on investment in residential construction rely totally on STC estimates as currently specified by the Revenue Allocation Formula.

Annual data on work put in place by industry and by province are available from the Capital Expenditure Survey. However, what is required for the GST/HST tax base is new housing sales regardless of which industry (e.g., general contractors, special trade contractors, real estate and land developers, own account construction) has produced the housing, as well as data on alterations and improvements. Land is subject to GST and must be included in the tax base. Information on inventory change is important because GST does not apply until a house is sold. The survey should cover services industries incidental to construction, which include such activities as land assembly and servicing for resale to builders.

- 1* Purchases of new housing are eligible for GST rebates. As rebates vary with the house price range, it is imperative that sales are categorized by price range. Value of sales and unit sales by structure type (e.g., single, multiple) are needed. The price must include the cost of land, as well as other costs such as developer fees, architect and engineering fees, commissions, etc. These expenses are undoubtedly included in the sale price of a house. Sales to government authorities, by level of government, must be identified separately because different rebate patterns apply.

- 1* Homeowner Repair and Renovations Survey and Landlord Repair and Renovations Survey are at present funded by CMHC and must be continued and improved to become provincially more reliable. Provincial CV's are at present very large. Since both paid rent and imputed rent are exempt services; those who effect the repairs pay non-recoverable GST-HST on inputs used for maintenance and repair. Any alterations and improvements expenditures (i.e. capital expenditures) made by these industries represent non-recoverable GST-HST, which should be shown in investment in residential construction.

- 2 Information is not available even on aggregate value of materials, services and primary inputs. The survey of construction could be used as a vehicle to collect this information, especially in the context of the need to improve provincial economic accounts.

Banks and Other Finance

Banks and other finance industries are an exempt industry group and therefore pay GST on their purchased inputs. This area is weak in the provincial economic accounts. Data for this area are on an enterprise basis. For the provincial economic accounts, it is imperative that the data on output, expenses and GDP components be made available by province and by industry (SIC). The data on head offices are also important in the provincial context. IOFD is collecting data on destination of sales of financial institutions for the reference year 1996 and plans to cover the reference year 1997. An analysis of 1996 results might suggest changes to the proposed survey for the 1997-reference year.

Banks & Other Deposit Accepting (e.g., Trusts) Institutions

- 1 Income from fees (1992 value: $5 billion) including fee income earned by subsidiaries in businesses such as brokerage, underwriting, mutual fund, insurance and real estate is needed. Class of customer and origin and destination information is needed on service fees charged by these institutions.

- 1 Operating Costs: Banks total costs and a breakdown of current expenses by province and commodity are needed. While some national expense breakdown is provided by the Report of the Superintendent of Financial Institutions, no data are available at the provincial level.

- 2 Other revenues: Data are required on gross interest earned on assets (e.g., loans, etc.), gross interest paid on liabilities (e.g., deposits) so that net interest revenue can be calculated by province. Net interest at the national level was about $15 billion in 1992.

Mutual Funds

- 1 Data are required on the output ($4.1 billion in 1996) of this industry by province. Additionally, the cost of service of mutual funds should be estimated by province of residence of beneficiaries (persons paying the said fees). These data are also is needed for segregated funds managed by insurance companies. Both are critical for estimating personal expenditure by province.

- 1 Operating expenses and a breakdown of expenses (principally portfolio management expenses, accounting expenses and legal expenses) is also needed at the provincial level. Same details are needed for segregated funds managed by insurance companies.

Insurance

Property & Casualty Insurance:

- 1 For insurance carriers, premiums received, and claims paid, are required by province and class of customer. For trade flows, data are required on premiums paid and claims received by province of residence of the insured.

- 1 Total operating expenses and a breakdown comparable to what is available at the national level from the Superintendent of Financial Institutions is needed for each province where these expenses are incurred. IOFD has prepared, and planned to execute, a P&C insurance survey for the 1996 data year, which could generate our required provincial detail. The status of this survey is unknown.

- 2 Premium and claims data should identify at least the following broad class of customers for all auto and property insurance: persons (households), businesses (auto rental companies and real estate companies separate from other businesses), governments, non-profit institutions.

Life Insurance:

- 1 Premiums received and other income received are needed for life insurance carriers by province. Data are required based on the residence of insured persons.

- 1 Similarly, operating expenses and detail similar to what is available from OSFI at the national level is needed for each province.

Securities Dealers/Brokers

- 1 All revenue data are required by province. Stock and bond commissions, fees and interest income received by brokerage houses are needed by class of customer and province of residence of clients.

- 1 Commodity breakdown of operating expenses is required by province.

Consumer/Business Finance, etc.

- 1 Revenue data are required by province. Interest and fee data are needed by province of residence of client/borrower and by class of customer.

- 1 Operating expense data is needed for multi-establishment enterprises, by province.

Residential rent

- 1* Data on revenues and expenses of landlords are needed by province. A breakdown of expenses of landlords by province is needed as their expenses are subject to GST.

Transportation

Provincial location of mobile capital should correspond to the associated provincial output. To achieve this correspondence, provincial data on output and capital expenditure must be developed collaboratively by the Transportation and Investment and Capital Stock (ICS) Divisions. Data are required on the head offices and other ancillary activities for all modes of transport. Specific proposals made by Transportation Division are listed below with our ranking.

- 1 For air transport, the fair basis survey should be improved to provide a better person/business split of passenger revenue as part of class of customer information. Revenues are needed (at least for the personal expenditures and exempt industries) by origin and destination of the flight because this affects the tax status of the service. The origin must be linked to residency.

- 1 Conduct research on the feasibility of a business/persons split for rail transport revenues. To expand the coverage to include railways that are not federally regulated (short lines).

- 1 Cover the taxicab industry.

- 1 Improve provincial allocation of inputs of the water transport industry as well as class of customer data. Revenue by commodity should be collected as a counterpart of tonnage data at present collected.

- 1 Cover the courier industry on a regular basis. (An ad-hoc survey was carried out for reference year 1990). Transportation Division proposes to create a consortium to fund this survey (at least the courier portion). There will be two principal components, the national (long haul, big players) operations and the local messengers. Provincial allocation of inputs is expected to be an issue for the former group, but not the latter. The previous survey collected information on class of customer.

- 2 Bus (interurban & rural, school, other) and urban transit. Improvements in provincial allocation of inputs and class of customer data are required. For urban transit, details of revenues and expenses are needed by commodity and by province.

- 2 Along with output for all other transport industries, we need improved provincial allocation of inputs.

- 3 The current data gap in the 'other transportation industries n.e.c. ". ('80 SIC 4589')

  Include calèche (horse-drawn buggies) and limousine services (not airport limousines, rather the "prom night" type of service). Since most players are expected to be small businesses, Transportation Division plans to use tax sources for this survey.

- 3 There is no survey for warehousing (Grain Elevators and Other Storage & Warehousing) industries. These are important industries. These industries are estimated to produce over $600 million in revenues each in 1993.

- 3 To extend the coverage to firms classified to transport that do not usually transport goods themselves but are instrumental in linking shippers to transportation companies. According to Transportation Division this industry is also of interest to Industry Canada, which recently undertook an ad-hoc survey of freight forwarders.

- 3 In the specialty air transport, a variety of services including air ambulance, forest fire fighting and pilot instruction are at present covered. Some of these operations are inter-provincial and several have principally commercial clients. This is the only commercial air service covered under the NAFTA. According to Transportation Division, supplementary funding may be possible from other departments, such as Transport Canada and Foreign Affairs & International Trade.

- 3 Airports, which have been transferred to the private sector, generate significant revenues, not just from landing fees, but from services including concessions and parking. The larger privatized airports currently report labour and capital data to STC, but there is no existing source of financial and operating data. For air navigation services there will be one respondent (Nav Canada), but the revenues and expenses generated will be significant and will be spread over the country, requiring provincial allocation.

- 3 Revenue and expense data is needed by province for ferry and bridge facilities.

Telecommunication Services

- 1* Revenue data by province is required on regular and cellular phone services, cablevision (including pay and specialty television), and non-telephone services (such as pagers, facsimiles and internet service provision) provided by carriers and resellers;

- 1* The revenue data is required by class of customer and by destination of sales.

- 2 Data on operating expense detail are needed by province.

Distributive Trades (Wholesale and Retail)

- 1* The bulk of GST resides in personal expenditure and therefore this area represents an important part of the I/O demand equation. We understand that under the NAICS (although both retail and wholesale trade industries sell goods to the personal sector) the dividing line between the two is not whether one industry predominantly sells to the personal sector. Therefore, for both wholesale and retail industries (including automotive sales) data are required by province on sales and cost of sales by store- type and commodity composition of sales. The commodity detail collected should be harmonized with FAMEX and should identify GST-taxable versus non-taxable goods and services. To ensure proper coverage, estimates of wages and salaries derived from the survey must be reconciled with T-4 information. There is at present a large discrepancy between the two sets of estimates.

- 1* For retail trade, the class of customer information should cover sales to households, governments, businesses and export. For wholesale, the class of customer information should cover sales to household, other retail, governments and exports. All these data are required by province.

- 1* New motor vehicle sales, also covering lease revenue (monthly contract amount), value of leased vehicles returned at the termination of lease, inventory for lease, fleet sales, sales of used vehicles, and repair work are required (see also finance leasing discussion under Business, Personal and Other Services).

- 1* Trade data are also required by commodity on non-store sales to households, businesses and exports, i.e., the Direct Selling Survey.

- 2 Inter-provincial trade flows are required on origin and destination of sales by type of store. Additional improvements are needed to the wholesale trade origin and destination survey.

- 2 Provincial gross margin data by commodity should be collected if feasible.

- 2 Agents and Brokers should be covered by the wholesale survey to be complete. I/O tables will require only commission income and not gross sales for Agents and Brokers along with commodity detail and class of customer.

Business, personal and other services

Business services

- 1* Real estate commissions on the sales of buildings are subject to GST. The estimated amount of such commissions was $5.8 billion in 1992. The survey should collect real estate commissions by province and should separate the residential part from the non-residential component and separately collect revenue data on secondary activities (providing managed services, commissions on signing leases, etc.). Commissions should be reported separately for sales of land from land-and-structure real estate, and for signing leases. Ideally the survey should also collect sales value by structure type thereby permitting commission rates to be calculated. Other transfer cost data, such as property assessment fees and building inspection fees should be collected.

- 1* The finance leasing and operating leasing companies are combined in the same NAICS. The revenue detail obtained from finance leasing establishments will have to be specific. As reported revenues may include both the interest and repayment of principal, it would be essential to separate out the interest component of total payments because, according to the 1993 SNA, the repaid principal would be considered a balance sheet item and part of financial flow accounts;

- 1* Accommodation service industries

  The cost of sales should be broken down into the cost of sales of alcoholic beverages, sales of meals and GPRS. The class of customer data by origin (residence) are needed. The data on accommodation is also used for tourism Satellite Accounts.

- 1* Food and beverage service industries

  A breakdown of revenues into alcoholic beverages, meals, etc. is required. Expense detail by commodity for major intermediate inputs and value added components are required.

- 1* Improvements in the Canadian Travel Survey and the International Travel Survey: The Canadian Travel Survey should provide better detail (focusing on taxable/non-taxable commodities) of expenditure by class of traveler. This data is required for inter-provincial trade flows.

- 1 Accounting and legal services survey

  - Annual survey of accounting and legal services is required;

  - The survey should be on an establishment basis and should provide good estimates of provincial revenue.

  - Revenue from legal aid services should be broken out separately since it is GST exempt;

  - Class of customer information is essential.

- 1 Amusement and Recreational Services Survey

  The survey should be extended to cover those recreational services not at present covered, including non-profit and for-profit bingos, various types of lotteries, and casinos.

- 1 Municipal governments, 2 Universities, Schools and Hospitals (MUSH);

  MUSH are eligible for partial tax rebates. Commodity details of municipal government expenditures are not available. The commodity details of hospitals and education need to be improved.

- 2 Personal and household services

  Most of the output of these services flow to personal expenditure upon which the GST is levied. All establishments classified under the NAICS classification (including daycare and other forms of childcare) should be included such that we have appropriate revenue levels by industry.

- 2 Other business services such as management consulting should be surveyed and class of customer information obtained.

- 2 Membership organization industries, 1980 SIC

  At present membership organizations are not surveyed. Appropriate revenue and expenditure detail are required to estimate these industries at the national and provincial levels. These organizations pay GST-HST on the inputs they purchase.

- 2 Survey of non-profit and charitable organizations

  The last survey for some components was for data year 1980. Non-profit organizations are taxed differently. Data for revenue and aggregate expenditure for non-profit and charitable organizations are available from administrative sources. A suitable sample of these organizations would provide the commodity detail of expenditures. We must ensure that there is no duplication in the survey data.

- 2 Health Care

  This industry, consisting of physicians, dental practitioners, and other health services, produces tax-exempt services. There is no data on inputs bought by members of this industry while GST is paid on such purchases.

Family Expenditure Survey

1* Data from this survey are needed in order to confront personal expenditure estimates with data derived from the supply side of the I/O equation. The commodity classification of FAMEX surveys should be concordable to those of trade surveys and the various surveys of services.

Capital Expenditure Survey

2 Data on capital expenditures made by producers of exempt commodities should be identified and tabulated separately.

International Trade by Province

1 The concepts and definitions underlying measures of imports and exports of commodities as defined in the International Trade Division and the SNA are different. The SNA and the International Trade Division will be working together to arrive at a harmonized set of definitions to facilitate data compilation related to the administration of the HST.

Annex

3

ADDITIONAL INFORMATION ABOUT ADMINISTRATIVE DATA SOURCES

The major tax data sources that will be addressed in this Annex are:

1. Business Tax Returns with General Index of Financial Information (GIFI)
2. Goods and Services Tax (GST)
3. Payroll Deduction Accounts (PD)
4. Customs Documents - Import and Export Shipments

Business tax returns with General Index of Financial Information (GIFI)

Definition

The universe of tax returns includes both incorporated (T2) and unincorporated (T1) business information.

A T2 return is filed annually by corporations who report income from any source. STC receives, on an annual basis, a universe file containing "tombstone" information about the corporation and a basic set of financial variables - assets, equity, sales, profits and taxable income (with provincial breakout). There are approximately 1 million T2s filed each year and they must be filed within six months after the fiscal year end of the corporation. The Industrial Organization and Finance Division manages an SQL database with these data and uses them extensively in its program.

Approximately 20 million T1 returns are filed annually. Of these approximately 2 million report income from a business, profession, farming, fishing, commissions and rental operations. STC receives, on an annual basis, a universe file containing "tombstone" information about the individual tax filer as well as personal income and deductions and gross and net income from all self employed (unincorporated) sources.

To allow corporations in Canada to file tax returns electronically, there is a requirement to standardize the content and method of filing financial statements including the Income and Expense statement, the Balance Sheet and all of the associated schedules. This initiative is referred to as the GIFI schedule (previously called the FSI form).

Statistics Canada was represented on the working group formed by Revenue Canada to draft the GIFI and all the associated schedules for corporations. The GIFI defines a standard set of Income Statement and Balance Sheet items that corporations and individuals will be required to file when reporting the results of their business activities to Revenue Canada. These data will be captured electronically and made available to Statistics Canada annually.

Revenue Canada intends to make it mandatory for corporations with gross revenue or assets greater than $3,000,000 to file electronically using GIFI. Smaller corporations will retain the option of filing on paper. Recognizing the difference in complexity between these two groups as well as the distinctiveness of farm tax filers, three GIFI forms have been proposed[7]. The Complex GIFI is intended for large corporations and will also be the subject of a standard adopted by developers of accounting software. It is comprehensive and includes most of the items that are usually provided by the large filers.

The Short GIFI form is intended for those filers who are not likely use purchased accounting software to prepare their statements. It contains fewer than 100 data items and will be available on a form printed by Revenue Canada. Even though the level of detail on the Short form is limited when compared to the Complex GIFI form, all items on the latter (as well as the items on the Farm GIFI) can be mapped into the former. Finally, a third GIFI has been designed for farm corporations because many items that they may report are particular to their industry.

Content/Coverage/Volume

Lists of the items included in the Complex, Short and Farm GIFI forms are available on request. It is expected that 7% of all corporations (representing more than 80% of total revenue) will file the Complex GIFI while 90.5% will be eligible to use the Short and 2.5% will be using the Farm GIFI. However, among the potential users of the Short version, Revenue Canada estimates that 80% are already using a computerized method to prepare their financial statements and tax returns. This makes them more likely to voluntarily use the standardized statements and also to file electronically. In order to maintain equivalent reporting between paper and electronic filers, Revenue Canada will capture most of the information supplied by paper filers. This new process is scheduled to begin in April 1998 and will cover all of the 1997 T2 returns filed after that date (a corporation has up to 6 months after the fiscal year-end to file a T2) and the vast majority of the 1998 returns. As more corporations begin filing electronically or adopt the standard forms, the costs of the data capture process will decline.

Revenue Canada also intends to capture information from the T1 GIFI that is used by unincorporated businesses. The target for implementation is February 2000, for 1999 tax year data. The T1 program will align its GIFI form to the largest degree possible with the content of the T2 program.

Periodicity/Timing

Discussions with the Canadian Institute of Chartered Accountants (CICA) are taking place in 1997 regarding plans to use the GIFI as the method of communicating business results to Revenue Canada (RC). A pilot to test the GIFI will be carried out subsequently, and RC plans to introduce the GIFI for all corporations in April 1998. As well, RC will introduce mandatory electronic filing for all corporations, although for those with less than $3 million sales, there will be no enforcement action at the outset. RC intends to make the GIFI a

[7] In its present format the Complex GIFI form does not meet the all of the filing requirements for Financial Institutions, so a fourth GIFI form designed specifically for these corporations is being considered.

mandatory requirement by October 1998. They also plan to have a system in place to process GIFI data in April 1998 and as of that date, they will capture all GIFI documents electronically. In addition, any financial statements not filed using the GIFI format will be transcribed to GIFI forms and captured. It is noteworthy that although electronic filing is not presently an option for corporations, 80% of corporations actually file using computer prepared statements (the T2 returns or financial statements are prepared on computer and printed for RC filing purposes). It is expected that even for small corporations who do not file electronically in the future, many will use the GIFI form to report to RC, thus reducing the costs of transcribing data.

Unincorporated - Presently there are more than 3 million unincorporated T1 tax filers reporting business, professional, farming, fishing, commission or rental income to Revenue Canada. More than 1 million of these use the electronic version of the Standardized Financial Data (SFD) form, designed jointly by Statistics Canada and Revenue Canada to report their financial information. The remaining 2 million unincorporated tax filers continue to use paper filing to report this type of information. Each year, Statistics Canada captures a sample of these paper documents for T1 tax filers reporting business income and generates industry estimates which are used by several survey programs. This data capture program will stop when the T1 GIFI form is implemented by Revenue Canada. It was intended that the GIFI would replace the SFD form for tax year 1998, but this has been delayed. The intention now is to introduce the T1 GIFI form in 2000 for tax year 1999.

Note: For tax year 1996, STC's Tax Record Access (TRA) group and RC's Computer Screening (COMSCREEN) group will merge operations to transcribe current data requirements for all business and professional tax filers reporting a Gross Income above $70,000. This arrangement will continue until the GIFI form is introduced for T1 tax filers.

Goods and Services Tax (GST) Files

Definition

Statistics Canada now has access to the detailed records maintained by Revenue Canada on all businesses with GST accounts. The file contains information on each firm's total sales, GST collected, and GST paid. These data should provide a rich source for estimation of sub-annual industry sales (GST accounts are SIC coded) and for analysis of business cycle trends. Linkage with other tax and survey files will be facilitated by the fact that the Business Number is the key identifier on GST accounts.

There are potential problems to be overcome in using this data source. Some business entities – for example, local governments, hospitals, educational institutions – do not pay GST. Using GST data to estimate production is made further problematic because not all goods are taxed at the familiar 7% rate. Some goods are exempt and other goods, such as food and exports are not exempt but taxed at a 0% rate. However, research is underway to determine how this new data source can be used effectively, and how some of the problems can be overcome.

Content/coverage/volume

Every business in Canada with $30,000 in annual sales is required to register for a Goods and Services Tax account with Revenue Canada in order to submit tax collected and to claim input tax credits. Smaller firms may also choose to establish an account, in order to claim credits. Taxi and limousine service businesses are required to file irrespective of the $30,000 threshold. There are roughly 2 million active GST accounts on the Revenue Canada file.

The GST file contains the firm's total sales of GST liable goods; the GST value collected, and the GST on inputs purchased by the firm (and thus for which an input tax refund claim is being filed).

Periodicity/timing

Remittances are made on a variety of official filing periods which can be selected by the individual firm in balancing reporting burden with its request for input tax refund claims. Firms may file on a monthly (5% of remitters, usually the largest firms, accounting for 74% of total sales), quarterly (83% of remitters, not necessarily the calendar quarter, accounting for 25% of total sales) or annual basis (the remaining 12% of remitters, accounting for 1% of sales). Many of the monthly tax filer firms elect to file on a variety of multi-week bases.

Revenue Canada prepares a file for each month's remittances by roughly the third week of the following month.

Quality

GST data quality is presently being analysed by staff in the Tax Data, Business Register and Services Divisions, and by the Time Series Research staff. Preliminary results are promising.

The GST account normally represents a business unit virtually synonymous with the legal entity concept. GST account information for firms will be linked to other taxation account data through the Business Number. This will enhance the proposed use of GST data for stratification purposes during sampling design, and in developing revenue proxy data during edit and imputation stages. The latter will also replace the modeling method presently used to produce GBI data for the Business Register. These uses are restricted to single-establishment enterprises. In a broader sense, the GST data will permit very timely analysis of business cycle trends, given the relatively rapid access to monthly data, and the fact that all businesses are covered.

Payroll deduction

Definition

Payroll Deduction Accounts are established by employers to remit to RC employer and employee contributions towards Canada/Quebec Pension Plan and Employment Insurance contributions, as well as source deductions for income tax.

Content/coverage/volume

All employers are legally required to establish payroll deduction accounts. There are approximately 1 million active accounts in use.

The PD account file contains information on remittances, employment, and total payroll for the firm. The two latter items were added to the PD form at Statistics Canada's request to provide data for the Survey of Employment, Payroll and Hours (SEPH). When these items were added to the form in 1993, Revenue Canada could not immediately change their system to capture the additional data. Therefore, these data were captured each month by STC staff for a sample of PD accounts. Beginning in August 1996, Revenue Canada began to capture these items from all payroll deduction remittance forms.

Periodicity/timing

Most payroll deductions come from small firms who file on a monthly basis, and generally have remittances of less than $15,000 per month. There are about 900,000 such PD accounts accounting for about 3,000,000 employees. Larger firms file remittances with the same periodicity as their payroll (i.e. firms which pay their employees weekly also remit weekly). This group comprises fewer than 70,000 businesses but accounts for more than 7,000,000 employees.

Revenue Canada transmits PD files to Statistics Canada each week for the remittances filed in the previous week.

Revenue Canada also provides a file containing identifier and related information (e.g. active or inactive, remittance periodicity, etc.). This file, known as PAYDAC, is transmitted for update purposes twice a month.

Quality

While PD Account data are of good quality, our ability to use the data is constrained by the nature of the file and the way accounts are structured within businesses, a difficulty shared with the GST files.

As with GST, There are cases where a business entity will have more than one account, e.g. there may be a PD account for salaried staff and one for hourly-rated workers. In other cases an account will be shared over a number of entities, e.g. the management staff across a number of plants or outlets may be covered by a single PD account. a further similarity to the GST files is the absence of establishment breakdown on the form. This information, however, can be estimated through linkages with other data sources.

Presently the Labour Division makes extensive use of PD7 remittance data within its program. Employment and payrolls data have been used from this administrative data source for small simple businesses since March 1993 and for all non-complex businesses since May 1996. PD records will be used for complex businesses beginning with the May 1997 reference month.

To date, the use of PD7 administrative records has worked well. Overall quality of SEPH program estimates (as measured by the degree to which SEPH diverges from the LFS) has improved and respondent burden has been significantly reduced, as have been program costs. The extension of the SEPH initiative to use administrative records for complex businesses is in line with the strategy outlined here. Because the administrative data are available for the enterprise and not the establishment, adjustment must be made if provincial or industrial breakdowns are required. To provide the estimates at the establishment level, the Business Register and other sources will be used to obtain information for allocation purposes.

The availability of the Business Number for linking PD to other taxation account data will enhance the present use of PD data for estimating employment (as used by the Survey of Employment, Payroll and Hours) and for providing a basis for determining whether a firm is active.

Customs Documents - Import and Export Shipments

Definition

The International Trade program has access to a continuous flow of Customs administrative data reflecting all Canadian import and export shipments. These data are compiled monthly to derive value and volume statistics by country, commodity and province of clearance/origin. Current efforts to structure these data by industry involve the assignment of SIC codes by employing a match of importer or exporter account numbers (GST or PD based) to the BR information or alternatively by utilizing name matching techniques. This approach has been used over several years to the point where 90% of annual imports and 87% of exports have been SIC allocated.

Revenue Canada is currently requesting that all businesses with a BN registration utilize this number as their importer or exporter number. When all businesses have been fully converted to the BN and when all Customs transactions carry the BN it will be possible to derive an SIC from the BR with a direct linkage based on this number.

A comprehensive compilation of trade data on an SIC basis should be of considerable assistance to the development of Input/Output and Provincial flow data.

Timing

The trade data are compiled on a monthly basis with the data for an annual reference period being available approximately by end March.

Quality

The utilization of the BN for import shipments will be quite comprehensive as the processing of import declarations is highly automated and the presence of the account number is basic to acceptance of a declaration for customs clearance. However, there will be a period of conversion (certainly all of 1997) where a mixture of old importer numbers and new BN account numbers are used and therefore the linkage to the BR will remain incomplete.

Import data suffer from a limitation in that the port (province) of clearance indicated on the transaction does not necessarily coincide with the province where the goods are consumed. Linkage to the BR will provide some additional provincial information if the goods are to be used by production entities identified for a given importer business. However, if the business happens to be involved in the wholesale or retail distribution of goods, little or no province of consumption information can be derived from the BR.

Export documentation for overseas shipments should carry the BN but as there is no confirmation of the number's presence by Custom's staff, the degree of reporting will be less than comprehensive. Furthermore, data for all US bound export shipments are compiled by the US Bureau of the Census and provided to the International Trade Division under an exchange agreement. These data will not carry the BN but rather a Canadian vendor number assigned by US Customs. An investment would have to be made in developing a concordance between the BN and the US assigned Canadian vendor number in order to derive the SIC from the BR. If the concordance was successfully maintained it might greatly assist in the attribution of exports to province of origin given a linkage to the BR and the business entity geographic characteristics.

Annex 4

Business Survey at Statistics Canada

SDDS	DIVISION	FREQUENCY	SURVEY NAME
3406	Agriculture Division	Annual	Harvested Area, Production and Value of Processing Vegetables
3407	Agriculture Division	Annual	Fruit and Vegetable Survey
3408	Agriculture Division	Annual	Fall Vegetable Survey
3411	Agriculture Division	Annual	Mushroom Growers' Survey
3412	Agriculture Division	Annual	Raspberry Growers' Survey
3413	Agriculture Division	Annual	Strawberry Growers' Survey
3414	Agriculture Division	Annual	Maple Products
3416	Agriculture Division	Annual	Annual Greenhouse, Sod and Nursery Survey
3419	Agriculture Division	Annual	Honey Production, Value and Colonies
3420	Agriculture Division	Annual	Special Potato Survey - Alberta
3426	Agriculture Division	Annual	Fur Farm Report - Mink and Foxes
3442	Agriculture Division	Annual	Forage Seed Usage Survey
3443	Agriculture Division	Annual	Miller's Annual Report
3444	Agriculture Division	Annual	Storable Vegetable Survey
3446	Agriculture Division	Annual	Potato Area and Yield Survey
3451	Agriculture Division	Annual	Seed Corn Trade Survey

SDDS	DIVISION	FREQUENCY	SURVEY NAME
3459	Agriculture Division	Biennial	Area Farm Survey
3450	Agriculture Division	Biennial (See NOTES)	The Farm Financial Survey 1994
3438	Agriculture Division	Quinquennial	Census of Agriculture
3462	Agriculture Division	Every 5 years to coincide with the Census of Agriculture	Progress of Seeding Follow-up (POS)
3403	Agriculture Division	Monthly	Miller's Monthly Report
3404	Agriculture Division	Monthly	Reports of Crushing Operations
3423	Agriculture Division	Monthly	Stocks of Frozen Meat Products (Monthly)
3425	Agriculture Division	Monthly	Inventory Statement of Frozen Eggs and Poultry Products
3430	Agriculture Division	Monthly	Dairy Factory Production and Stocks
3431	Agriculture Division	Monthly	Inventory Statement of Butter and Cheese
3436	Agriculture Division	Monthly/Quarterly	Farm Product Prices Survey (FPPS)
3401	Agriculture Division	Sub-annual (8/yr)	Field Crop Reporting Series
3434	Agriculture Division	Quarterly	Survey of Livestock Slaughter
3449	Agriculture Division	Quarterly	Survey of Wage Rates for Hired Farm Labour
3421	Agriculture Division	Semi-annual	Egg Producers Survey
3460	Agriculture Division	Semi-annual	July Livestock Survey; January Livestock Survey

SDDS	DIVISION	FREQUENCY	SURVEY NAME
1505	Balance of Payments Division	Annual	Great Lakes - St. Lawrence Seaway Shipping Transactions (BP-20)
1510	Balance of Payments Division	Annual	Investment in Canada by Non-Canadian Corporations (BP-22)
1513	Balance of Payments Division	Annual	Report of Cargo, Earnings and Expenses of Ocean Vessels Operated by Non-resident Interests (BP-24)
1514	Balance of Payments Division	Annual	Report of Cargo, Earnings and Expenses of Ocean Vessels Operated by Canadian Companies (BP-25)
1515	Balance of Payments Division	Annual	Report on Imports of Crude Petroleum and Petroleum Products and Other Shipping Operations (BP-26)
1516	Balance of Payments Division	Annual	Movement of Funds Between Canadian Incorporated Insurance Companies and Their Foreign Affiliates, Agencies and Bank Accounts and Other Companies or Persons Outside Canada (BP-27)
1517	Balance of Payments Division	Annual	Movement of Funds Between Canadian Branches of Foreign Insurance Companies in Canada and Head or Other Offices, Companies or Persons Outside Canada (BP-28)
1518	Balance of Payments Division	Annual	Report by Trust and Mortgage Loan Companies in Canada on Transactions with Non-residents (BP-29)
1522	Balance of Payments Division	Annual	Geographical Distribution of Capital (BP-52)
1523	Balance of Payments Division	Annual	Structure of Canadian Companies in the reporting enterprise (BP-53)

SDDS	DIVISION	FREQUENCY	SURVEY NAME
1525	Balance of Payments Division	Annual	Particulars of Selected Issues of Funded Debt and Foreign Bank Borrowing (BP-55)
1526	Balance of Payments Division	Annual	Geographical Distribution of selected long-term debt booked in Canada at Canadian Banks and Consolidated Canadian Subsidiaries (BP-56)
1528	Balance of Payments Division	Annual	Transactions of Foreign Airlines with Residents of Canada (BP-58)
1529	Balance of Payments Division	Annual	Capital Invested Abroad by Canadian Enterprises (BP-59)
1530	Balance of Payments Division	Annual	Capital Invested in secondary foreign companies by Canadian Enterprises (BP-59S)
1531	Balance of Payments Division	Annual	Canadian Investment in Non-Canadian Corporations (BP-60)
1532	Balance of Payments Division	Annual	Investment in Canada of Non-Canadian Partnerships (BP-61)
1506	Balance of Payments Division	Annual (See NOTES)	Transactions Between Canada and Other Countries (BP-21)
1533	Balance of Payments Division	Annual/Quarterly	International Transactions in Services (BP-21S, BP-21SQ, BP-17)
1501	Balance of Payments Division	Monthly	Gold and Silver Transactions, Assets and Liabilities (BP-11)
1519	Balance of Payments Division	Monthly/Quarterly	Security Transactions with Non-residents (BP-30, BP-30A, BP-30B, BP-30Q)
1507	Balance of Payments Division	Quarterly	Transactions Between Canada and Other Countries - Quarterly (BP-21A)

SDDS	DIVISION	FREQUENCY	SURVEY NAME
1511	Balance of Payments Division	Quarterly	Investment in Canada by Non-Canadian Corporations (BP-22A)
1520	Balance of Payments Division	Quarterly	Futures Trading with Non-residents (BP-31)
1105	Business Register Division	Quarterly	Business Register Profiling
2413	Education, Culture and Tourism Division	Annual	Film, Video and Audio-Visual Production Survey
2414	Education, Culture and Tourism Division	Annual	Film, Video and Audio-Visual Distribution and Videocassette Wholesaling Survey
2415	Education, Culture and Tourism Division	Annual	Motion Picture Laboratory Operations and Production and Post-Production Services
2416	Education, Culture and Tourism Division	Annual	Motion Picture Theatres Survey
3105	Education, Culture and Tourism Division	Annual	Survey of Book Publishers and Exclusive Agents
3139	Education, Culture and Tourism Division	Annual	Periodical Publishing Survey
3115	Education, Culture and Tourism Division	Biennial	Sound Recording Survey (Production, Distribution, Releases and Sales)
2511	Industrial Organization and Finance Division	Annual	Survey of the Insurance Business in Canada
4906	Industrial Organization and Finance Division	Annual	Survey of Leasing
2503	Industrial Organization and Finance Division	Annual	Corporations and Labour Unions Returns Act (CALURA), Part 1 - Corporations
2501	Industrial Organization and Finance Division	Quarterly	Quarterly Survey of Financial Statements

SDDS	DIVISION	FREQUENCY	SURVEY NAME
2001	Industry Division	Annual	Electric Utility - Financial Report Annual
2103	Industry Division	Annual	Annual Survey of Manufactures (ASM)
2107	Industry Division	Annual	Annual Survey of Forestry
2118	Industry Division	Annual	Annual Production of Soft Drinks
2139	Industry Division	Annual	Frozen Fish
2177	Industry Division	Annual	Coal Mines
2178	Industry Division	Annual	Crude Petroleum and Natural Gas Industry
2179	Industry Division	Annual	Oil Pipeline Transport - Annual
2180	Industry Division	Annual	Gas Utilities/Transportation and Distribution Systems (Annual)
2181	Industry Division	Annual	Electric Power Capability and Load Forecast
2193	Industry Division	Annual	Electric Power Generating Stations
2196	Industry Division	Annual	Electric Power Thermal Generating Station Fuel Consumption
2404	Industry Division	Annual	Vending Machine Operators
2405	Industry Division	Annual	Direct Selling in Canada
2409	Industry Division	Annual	Retail Chains and Department Stores
2433	Industry Division	Annual	Annual Retail and Wholesale Trade Survey
2807	Industry Division	Annual	Survey of the Construction Industry
2003	Industry Division	Monthly	Coke Monthly

SDDS	DIVISION	FREQUENCY	SURVEY NAME
2101	Industry Division	Monthly	Monthly Survey of Manufacturing (MSM)
2105	Industry Division	Monthly	Steel Pipe and Tubing
2106	Industry Division	Monthly	Steel Wire and Specified Wire Products
2110	Industry Division	Monthly	Mineral Wool Including Fibrous Glass Insulation
2116	Industry Division	Monthly	Steel Primary Forms, Steel Castings and Pig Iron
2117	Industry Division	Monthly	Electric Lamps (Light Sources)
2121	Industry Division	Monthly	Oils and Fats
2123	Industry Division	Monthly	Asphalt Roofing
2134	Industry Division	Monthly	Sawmills and Planing Mills
2136	Industry Division	Monthly	Pulpwood and Wood Residue
2138	Industry Division	Monthly	Construction Type Plywood
2140	Industry Division	Monthly	Cement
2141	Industry Division	Monthly	Particleboard, Oriented Strandboard and Fibreboard
2142	Industry Division	Monthly	Production and Disposition of Tobacco Products
2147	Industry Division	Monthly	Monthly Coal
2148	Industry Division	Monthly	Monthly Oil Pipeline Transport
2149	Industry Division	Monthly	Gas Utilities/Transportation and Distribution Systems (Monthly)

SDDS	DIVISION	FREQUENCY	SURVEY NAME
2150	Industry Division	Monthly	Monthly Refined Petroleum Products
2151	Industry Division	Monthly	Monthly Electricity
2183	Industry Division	Monthly	Industrial Chemicals and Synthetic Resins
2184	Industry Division	Monthly	Disposition of Shipments of Ingots and Rolled Steel Products
2191	Industry Division	Monthly	Monthly Oil Pipeline Statement
2401	Industry Division	Monthly	Wholesale Trade Survey (Monthly)
2402	Industry Division	Monthly	New Motor Vehicle Sales Survey
2406	Industry Division	Monthly	Retail Trade Survey (Monthly)
2408	Industry Division	Monthly	Monthly Department Store Sales and Stocks Survey
2428	Industry Division	Monthly	Motor Vehicle Dealers
2008	Industry Division	Quarterly	Retail Commodity Survey
2119	Industry Division	Quarterly	Footwear Statistics
2143	Industry Division	Quarterly	Sugar - Situation
2152	Industry Division	Quarterly	Business Conditions Survey (BCS)
2166	Industry Division	Quarterly	Quarterly Industrial Consumption of Energy Survey
2167	Industry Division	Quarterly	Natural Gas Disposition Quarterly
2168	Industry Division	Quarterly	Quarterly Survey of End-use of Refined Petroleum Products
2189	Industry Division	Quarterly	Shipments of Solid Fuel Burning Heating Products

SDDS	DIVISION	FREQUENCY	SURVEY NAME
2194	Industry Division	Quarterly	Electricity Supply/Disposition Quarterly
2199	Industry Division	Quarterly	Shipments of Plastic Film and Bags Manufactured from Resin
2011	Industry Division	Quinquennial	Water Use in Manufacturing Industries 1991
2131	Industry Division	Weekly	Steel Primary Forms Weekly
2201	International Trade Division	Monthly	Canadian International Merchandise Trade (Custom Basis)
2803	Investment and Capital Stock Division	Annual	Capital Expenditures, Preliminary Estimate for 19__, Forecast for 19__
2804	Investment and Capital Stock Division	Annual	Capital Expenditures, Revised Forecast 19__
2805	Investment and Capital Stock Division	Annual	Capital and Repair Expenditures, Actual for 19__
2615	Labour Division	Annual	Workplace and Employee Survey
2609	Labour Division	Annual	Pension Plans in Canada
2608	Labour Division	Biennial	Trusteed Pension Funds
2612	Labour Division	Monthly	Survey of Employment, Payrolls and Hours (SEPH)
2614	Labour Division	Monthly	Business Payrolls Survey (BPS)
2607	Labour Division	Quarterly	Quarterly Estimates of Trusteed Pension Funds
1650	Micro Economic Studies and Analysis Division	Annual	Survey of Operating and Financial Practices

SDDS	DIVISION	FREQUENCY	SURVEY NAME
1209	National Accounts and Environment Division	Annual	Environment Industry Survey, 1995
1903	National Accounts and Environment Division	Occasional	Environmental Protection Expenditure Survey, 1994
2009	National Accounts and Environment Division	Occasional	Waste Management Industry Survey
2333	Prices Division	Sub-annual	Price Report for Professional Computer Services
2331	Prices Division	Sub-annual 8 times a year	Average Prices of Selected Farm Inputs
2319	Prices Division	Annual	Canadian Telecommunications Plant Price Index (CTPPI), 1986=100
2328	Prices Division	Annual	Consulting Engineering Output Price Indexes (CEPI)
2301	Prices Division	Monthly (See NOTES)	Consumer Price Index (CPI)
2306	Prices Division	Monthly	Raw Materials Price Indexes (RMPI)
2310	Prices Division	Monthly	New Housing Price Indexes, 1992=100
2318	Prices Division	Monthly	Industrial Product Price Indexes (IPPI)
2317	Prices Division	Quarterly	Non-Residential Building Construction Price Indexes, 1986=100
2330	Prices Division	Quarterly	Apartment Construction Price Indexes, 1986=100
2305	Prices Division	Quarterly	Farm Input Price Index (FIPI)

SDDS	DIVISION	FREQUENCY	SURVEY NAME
2312	Prices Division	Quarterly (See NOTES)	Machinery and Equipment Price Indexes (MEPI)
2325	Prices Division	Three times per year	Electric Power Selling Price Indexes for Non-residential Customers, 1986=100
2722	Science and Technology Redesign Project	Annual	Telephone Statistics Survey - Annual
2724	Science and Technology Redesign Project	Annual	Radio and Television Broadcasting Survey
2728	Science and Technology Redesign Project	Annual	Annual Cable Television Survey
4201	Science and Technology Redesign Project	Annual	Research and Development in Canadian Industry
4205	Science and Technology Redesign Project	Annual	Energy Research and Development Expenditures
2721	Science and Technology Redesign Project	Monthly	Telephone Statistics Survey - Monthly
2726	Science and Technology Redesign Project	Monthly	Telecommunications Statistics Survey - Monthly
4218	Science and Technology Redesign Project	Occasional	Survey of Innovation in Services Industries
4220	Services Division	Annual	Survey of Real Estate Rental and Leasing and Property Management (1995)
2410	Services Division	Annual	Software Development and Computer Service Industry

SDDS	DIVISION	FREQUENCY	SURVEY NAME
2418	Services Division	Annual	Traveller Accommodation Survey
2420	Services Division	Annual	Architectural, Engineering and Scientific Services Survey
2423	Services Division	Annual	Survey of Miscellaneous Services
2424	Services Division	Annual	Survey of Personal and Household Services
2425	Services Division	Annual	Survey of Amusement and Recreational Services
2426	Services Division	Annual	Survey of Business Service Industries
2430	Services Division	Annual	Annual Survey of Professional Accountants
2432	Services Division	Occasional	Annual Survey of Management Consultants
2419	Services Division	Monthly	Restaurant, Caterer and Tavern Survey
2918	Small Business & Special Surveys Division	One time only	Survey of the Characteristics of Bankrupt Firms
2919	Small Business and Special Surveys Division	One time only	Survey on Aircraft Repair and Overhaul Sector
2920	Small Business and Special Surveys Division	One time only	Prairie Provinces Wage and Salary Survey
2921	Small Business and Special Surveys Division	One time only	Survey of Aboriginal Owned Businesses
2922	Small Business and Special Surveys Division	One time only	Survey of Technology Diffusion in Service Industries

SDDS	DIVISION	FREQUENCY	SURVEY NAME
2924	Small Business and Special Surveys Division	One time only	Survey of the Advanced Materials Sector
2925	Small Business and Special Surveys Division	One time only	Survey of Canada's Tourist Attractions 1996
2926	Small Business and Special Surveys Division	One time only	Survey of Information Technology Enterprises
2914	Small Business and Special Surveys Division	One time only	Survey of the Characteristics of Labour Demand in Quebec
2915	Small Business and Special Surveys Division	One time only	Survey of Small and Medium Sized Business in Atlantic Canada, 1994
2916	Small Business and Special Surveys Division	One time only	Survey of Practices in Support of Quality Services in the Public Service
2927	Small Business and Special Surveys Division	One time only	Canadian Acquaculture Industry Survey
2928	Small Business and Special Surveys Division	One time only	Survey of Communications and Electronic Banking Services
2713	Transportation Division	Annual	Canadian Civil Aviation - Annual Report
2734	Transportation Division	Annual	Railway Transport Survey - Annual
2742	Transportation Division	Annual	Annual Motor Carriers of Freight Survey
2753	Transportation Division	Annual	Financial Survey of Canadian Water Carriers
2795	Transportation Division	Annual	Annual Motor Carriers of Freight Survey, Private Trucking

STATISTICS CANADA LIBRARY
BIBLIOTHÈQUE STATISTIQUE CANADA
1010238733
Ca OOS

SDDS	DIVISION	FREQUENCY	SURVEY NAME
2798	Transportation Division	Annual	Passenger Bus/Urban Transit Survey - Annual
2751	Transportation Division	Continuous	Coastwise Shipping Survey
2701	Transportation Division	Monthly	Airport Activity Survey
2702	Transportation Division	Monthly	Air Passenger Origin and Destination - Scheduled Services
2705	Transportation Division	Monthly	Air Charter Statistics
2732	Transportation Division	Monthly	Railway Carloadings Survey - Monthly
2733	Transportation Division	Monthly	Railway Freight Traffic Survey
2735	Transportation Division	Monthly	Railway Operating Statistics Survey
2704	Transportation Division	Quarterly	System Passenger Origin and Destination Report - Other Unit Toll Services
2708	Transportation Division	Quarterly	Fare Basis Report
2712	Transportation Division	Quarterly	Air Carrier Operations in Canada
2714	Transportation Division	Quarterly	Air Carrier Fleet Report
2748	Transportation Division	Quarterly	Quarterly Motor Carriers of Freight Survey
2797	Transportation Division	Quarterly	Passenger Bus/Urban Transit Survey - Quarterly
2741	Transportation Division	Quarterly (See NOTES)	For-Hire Trucking Survey
2731	Transportation Division	Weekly	Weekly Railway Carloading Survey