The Use of Tax Data in Official Statistics – The Canadian Experience

Yung, Wesley
Statistics Canada, Business Survey Methods Division
100 Tunney's Pasture Driveway
Ottawa, Ontario, Canada K1V 1T1
E-mail: wesley.yung@statcan.gc.ca

LeBlanc, Dave
Canada Revenue Agency, Statistical Services Division
410 Laurier Avenue West
Ottawa, Ontario, Canada
E-mail: david.leblanc2@cra-arc.gc.ca

St.-Louis, Gaétan
Statistics Canada, Tax Data Division
170 Tunney's Pasture Driveway
Ottawa, Ontario, Canada K1A 0T6
E-mail: gaetan.st-louis@statcan.gc.ca

1.0 Introduction

Canada's statistical system has evolved over the years to the point of now being quite dependent upon administrative data for the regular production of statistics. The computerization of tax administrative programs, including both income tax and commodity tax, over the past number of years has made it possible to increase usage of the resulting data files for statistical purposes. The motivation for this interest in using administrative data stemmed primarily from factors such as tightening budgets, which caused a search for alternatives to the relatively expensive collection costs of statistical surveys and censuses, and an increasing concern about the burden of statistical enquiries on respondents. In recent years, further technological advances have facilitated and reduced the cost of manipulating large administrative files and have encouraged the increased use of administrative data in official statistics.

In this paper, we present the Canadian experience of integrating tax data into official economic statistics. In the next section we discuss the partners involved and the framework in place to ensure a well functioning partnership which facilitates the use of administrative data for statistical purposes. In section 3 we discuss the data that are shared between the partners, while section 4 briefly presents how the shared tax data are used in official economic statistics. The paper concludes with some final remarks in section 5.

2.0 The Partners

Statistics Canada (STC) is Canada's national statistical agency with the mandate to "collect, compile, analyze, abstract and publish statistical information relating to the commercial, industrial, financial, social, economic and general activities and conditions of the people of Canada." The objective statistical information produced by STC provides a solid foundation for informed decisions by elected representatives, businesses, unions and non-profit organizations, as well as individual Canadians. In addition to conducting Population and Agricultural Censuses every five years, STC is responsible for approximately 350 surveys on virtually all aspects of Canadian life.

The Canada Revenue Agency (CRA) is the federal government organization responsible for the administration of tax laws for the federal and most provincial and territorial governments in Canada. The Agency also administers various social and economic benefit programs delivered through the Canadian tax system. The so-called “core” tax systems include personal income tax, corporate income tax, trusts and investment income and related taxes, and finally commodity taxes such as federal and provincial sales taxes. In performing this tax administration function, the Agency is also responsible for ensuring all Canadians pay their required share of taxes due, a function commonly referred to as compliance.

2.1 Statistics Canada and the Canada Revenue Agency Partnership

None of the work that is currently done at STC with tax information would be possible without very close co-operation between STC and CRA. Early in this relationship it was realized that the tax information collected by CRA contained large volumes of specific taxpayer information which could be used in an expanded number of statistical analysis projects. It was also apparent that the use of tax data would decrease the costs of collecting similar data in existing surveys and, in addition, would decrease response burden. Initial discussions between STC and CRA subsequently led to agreements in principle for sharing administrative data.

STC took a gradual approach to establishing a partnership with CRA. At first, STC was heavily involved in assessing the feasibility of using administrative data to meet the needs of the statistical programs. It was noted at the outset that just because there was an abundance of taxpayer information, commonly referred to as “administrative data”, this did not mean that the data were complete, coherent, and fit-for-use in a national statistical framework. Working groups were struck between the two organizations aimed at reviewing the various CRA databases and how they could meet the needs of STC. This work included reviewing and analyzing the data, understanding the definitions of the variables, evaluating any possible timing issues with the reporting and availability of the data, as well as comparing the “administrative tax data” with data collected under surveys to determine differences and possible explanations. Work progressed in this area and as the various analysis and studies performed by STC continued to show that the administrative data, with some adjustments for use in the statistical framework, were reliable and could be used, the two agencies continued to strengthen their relationship.

Initially, there was not a coordinated approach at STC for all of the uses of tax data. Each statistical program area acted independently to meet their own needs. This led to a lot of duplication and a lack of coherence in the ways in which various areas used tax data. To deal with these issues, in 1997, STC set up a separate tax data coordination group called the Tax Data Division. (TDD). In the early years, the various groups at STC who used tax data dealt separately with their own contacts at CRA. As the use and potential of tax data proliferated at STC, TDD became the focal point for all communication and liaison with CRA. Currently, when a STC division needs to contact CRA, be it to obtain new data, for explanation of existing data, or to explore the potential of making changes to some tax forms to include specific data requirements of STC, they must go through TDD. CRA also established a similar client relations area as the single point of entry to the agency for data requests and delivery of products. This centralized liaison approach allows for a more coordinated and efficient communication and delivery system, for both agencies. The two agencies communicate on a weekly basis through a client liaison group to help ensure that requirements are met and that changes can be managed quickly and effectively. This close co-operation and existing partnership between STC and CRA has led to benefits for the Canadian public.

2.2 Legal Framework Governing the Partnership

The partnership between STC and CRA is governed by a formal agreement, specifically a Memorandum of Understanding (MOU) signed by both parties. The MOU references the legal instrument that gives STC its authority to collect information, which is the Statistics Act. The MOU also references the legal authority under which CRA is able to share taxpayer information, namely the Income Tax Act and the Excise Tax Act, both of which outline the mandate of CRA and its ability to share data with other federal departments and agencies, including STC. In particular, the Statistics Act states that the Chief Statistician, or his designate, may inspect and have access to any returns, certificates, statements, documents or other records obtained on behalf of the Minister of National Revenue as long as this information is subject to the same secrecy requirements to which it was subject when collected.

The MOU between STC and CRA outlines the roles and responsibilities of the partners. Specifically, it includes:

- legislation under which the sharing of information can take place
- type of information authorized to be shared
- the purpose for which the information will be shared
- who has access to the information
- the level of protection required for the information
- the sharing of the information with a third party
- dispute resolution mechanisms
- amendment process and
- data transmittal process.

The overriding principles of this MOU are that STC will use these data exclusively for statistical purposes and that the confidentiality of the data will be protected. The confidentiality requirements are the cornerstone of the relationship between STC and any partner departments. It has always been emphasized at STC that protecting the confidentiality of data providers, be they respondents to surveys or tax filers, is the only way to ensure the public trust and co-operation without which a statistical agency cannot exist.

As tax data at the taxpayer-identifiable level are considered highly sensitive, added measures of protection are employed by STC. Employees are sworn in under the Statistics Act which clearly references the employees' responsibilities to protect the confidentiality of data. Access to tax information at STC is granted on a "need to use" basis only and is closely monitored. As well, the physical and electronic storage facilities for tax data are of the highest standards thus preventing disclosure or access by unauthorized employees. There is also a tracking system in place which enables STC to take a snap shot at any point in time and report to CRA all users of the tax data and the intended use of the data. In addition, STC conducts regular audits on its internal procedures for safeguarding the administrative data in order to ensure that all conditions in the MOU are fulfilled.

In the interest of transparency, whenever possible, private citizens and business owners are informed of the fact that their personal and business tax information may be shared with STC for statistical purposes only. This is done either by CRA which indicates on their web site that they will be sharing the data with STC or, in the case of taxation information, by STC indicating on questionnaires and on user guides that taxation data will also be used for statistical purposes. Based on consultation over the years, STC has learned that the vast majority of Canadians support the use of administrative information by STC. This is mostly based on two factors; trust in STC's ability to protect the confidentiality of the data and the resulting decrease in response burden. As long as these continue to be true, along with the rigorous protection of such information by STC,

administrative data from CRA will continue to be a valuable source of information for STC.

2.3 Maintaining the Partnership

Once a partnership has been established and an MOU is in place, it falls on the two partners to maintain this relationship. To ensure that the partnership is well maintained, a structure of governance has been established. Senior managers from both departments meet on a regular basis to discuss the broad issues affecting the partnership, while working-level staff maintain close communication on a weekly basis to ensure that operational issues are dealt with and to plan for future activities. As well, directors of stakeholder divisions on both sides meet several times a year to keep the communication flowing.

Both agencies keep each other informed of any planned changes. In certain cases, when changes are planned, STC is invited to review the plans and contribute an opinion. When forms used to administer programs at CRA are being redesigned, efforts are made to include STC requirements. This is done in recognition of the fact that, as part of the Government of Canada, all departments and agencies have the obligation to work together to ensure that public resources are used in the most efficient way for the benefit of all Canadians.

It is important that both partners are equally involved in maintaining the relationship and that, in times of need, one partner is ready to accommodate the other. STC has prepared comprehensive contingency plans for all of its outputs which rely heavily on administrative information. These contingency plans are shared with CRA, who have agreed to treat them as a priority. In summary, establishing and maintaining partnerships between STC and CRA is a complex process which requires dedication and perseverance on both sides, and a deep belief in the public good.

3.0 Administrative Data

CRA provides a large amount of tax administrative data to STC. The following list shows the wide range of information shared:

- Taxpayer-identifiable information related to **Personal Income tax**, which includes several variables containing personal information as well as information such as employed and self-employed earnings;
- Taxpayer-identifiable information related to **Corporate Income tax**, including several variables showing details of revenues and expenditures by corporation;
- Information on Trusts, Investment Income, Registered Retirement Savings Plans;
- Information on Labour Income and Payroll Deductions, including paid earnings from employment by employer, various deductions at source, as well as employment data;
- Benefit information including child care, Old Age Security, Canada Pension Plan, registered charities and non-profit information,
- A complete list of all businesses operating in Canada and registered with CRA, including their Business Number; and
- A wide-range of commodity tax micro information including data for the federal sales tax (GST) and federal/provincial harmonized sales taxes administered by the federal government (includes variables such as total sales, taxes collected, credits offset, taxes remitted)

CRA has a number of databases containing the information referenced above. There are several processing and data validation steps used in each of the administrative data systems to check the various data

provided by taxpayers against information provided by employers, financial institutions, and other federal departments and agencies. Each system has its own set of edits (or verification rules) in place to help validate the integrity of the data. When CRA prepares the data files to be provided to STC, there is also a set of additional edits in place to help identify data quality issues and provide explanations to STC. CRA works jointly with STC on developing and implementing numerous edit routines to help ensure the data are complete and of sufficient quality to be subsequently transmitted to STC. There are continual discussions between analysts in both agencies to address data definitional and quality issues, changes to the file and variables, including sharing edit routines and best practices so that when files are received by STC they can be quickly loaded into the STC data repositories for subsequent processing. It should be noted that although the data supplied by the CRA meet their purposes for tax collection, compliance, and some benefit payments, it does not always meet the data needs of the statistical system.

In the next subsection we present the steps that STC takes to make the data received from CRA ready for use in their economic statistics program. We concentrate on the Corporate Tax (T2), Personal Tax (T1) and Federal Goods and Services Tax (GST) data as they are the most used by STC's economic statistics program.

3.1 Processing of Administrative Data at STC

3.1.1 Corporate Income Tax (T2)

The T2 program refers to all tax information originating from incorporated businesses. The T2 universe consists of approximately 1.4 million corporations. There are two sources of T2 data:

- General Index of Financial Information (GIFI), which is a standardized framework of balance sheet and income statement data (which include variables such as assets, revenue, expenses, etc); a total of 700 variables can be recorded on GIFI but they would never all be applicable to a single corporation.
- Supplementary tax schedules (CORTAX); there are nearly 100 different types of schedules, but only schedules applicable to an individual corporate tax filer are required.

Of the 700 fields on GIFI, only eight fields are mandatory so it is possible that many variables needed by STC's survey programs would not be available which would reduce the relevance of T2 financial data. These non-mandatory variables are commonly known as detailed variables. Fortunately for STC, corporations usually report more than the eight mandatory variables on their financial statements. On average, approximately 30 financial variables on the balance sheet and 25 variables on the income statement are reported.

During initial processing at STC, the financial statements are balanced; capture errors are identified and corrected; detailed fields are checked to ensure they add up to totals and financial statements and schedules are verified for coherence. These checks increase the accuracy of T2 data, since these quality rules are always respected. In addition, reassessment records received from CRA are used in order to improve the quality of the information. Most of the problems identified in initial processing are corrected automatically. However, some problems are more serious (approximately 500 cases per month) and need to be addressed manually by accounting specialists. These adjustments also improve the accuracy of T2 data.

As discussed previously, corporations do not have to report all of the detailed variables on their financial statements provided to CRA, even if some variables are essential for T2 data users at STC. Three variables in particular have been identified as essential by STC survey programs using tax data: inventories, depreciation, and salaries and wages. Due to the importance of these variables, additional sources of

administrative data are used. For example, attempts are made to find salaries and wages information from payroll deduction data obtained from CRA. Examples of other sources of administrative data include the many different supplemental tax schedules included in the CORTAX program. By carefully analyzing these other sources of data, STC has been very successful in finding the required information

The Generic to Detail Allocation (GDA) process allocates amounts reported in generic fields into more detailed fields. The process first examines all businesses in a block which have responded to details only and groups them in clusters such that the distributions are homogeneous. Discriminatory models are then created to assign businesses with an unknown data distribution to clusters, after which the ratio distribution for that cluster is applied to the business.

On average, after the GDA, the number of different GIFI fields available increased from 30 to 50 for the balance sheet and from 25 to 31 for the income statement. These results show that the GDA process improves the relevance of T2 data for STC users by increasing the number of variables available.

Twice a year, T2 data are imputed for missing and erroneous records of corporations. Two principle imputation methods are employed: historical and donor imputation. In historical imputation, data from the previous year are copied to the current year for the same record. A trend, calculated from GST or T2 data, which reflects the growth rate from one year to another is calculated and applied to this imputed data. Donor imputation is done by finding the nearest neighbor of the recipient record using matching variables. One source of information used to determine the nearest neighbor is GIFI data from current and previous reference years. When previous years' data are used for a recipient, some of the matching variables need to be trend adjusted to become comparable with those of its donors received in the current year. This trend reflects the growth rate from a previous year to the current year. After imputation is performed the T2 data become available to STC surveys wishing to use them.

3.1.2 Personal Income Tax (T1)

The T1 unincorporated business universe consists of 3.6 million self-employed tax filers reporting income from unincorporated businesses wholly or partially owned by the tax filer, according to the most recent tax year data available. The Assessed Records File (ARF) is received from CRA and covers the entire universe of the 3.6 million filers. This file contains a very limited amount of information for STC use. However, STC also receives all financial information for businesses who file electronically representing approximately 80% of the total population. The information obtained from the electronic filers corresponds to the business' income statement. Balance sheet information is not available for T1 units.

During initial processing, the financial statements are balanced; capture errors are identified and corrected; and detailed fields are checked to ensure that they add-up to totals. These checks increase the accuracy of T1 data, since these quality rules are always respected.

Most of the identified problems are addressed automatically. However, for a certain number of cases, the problem is more serious and has to be addressed manually by accounting specialists. This improves the accuracy of the T1 data.

Three times a year, T1 data details are imputed for missing and erroneous records. Total Revenue and a derived calculation of Total Expenses are available for all records from the ARF. Details are then imputed by using a donor's ratio of detailed items to totals on the recipient to derive the recipient's detail.

3.1.3 Goods and Services Tax (GST)

This database is established from sales activities and the resulting federal sales taxes collected and remitted for all registered businesses in Canada. In many countries this tax is known as the Value Added Tax (VAT). The GST database covers all registered businesses in Canada that either remit GST payments for taxes collected on their sales or receive GST rebates on some taxes they have paid. The only financial information contained on the database is Total Revenue and Total Sales. In order to make this information available to the monthly survey programs, a system has been developed at STC for editing, imputing and calendarizing the GST data. Not all GST filers are required to file on a monthly basis and even those who are may not report for a complete calendar month. The current STC system produces edited and imputed data on a calendar month basis for all records within 24 to 48 hours of receipt of the file from CRA.

Upon receiving data for reference month ‘m’ from CRA, the coherence of the records is verified and then the edit and imputation process is performed. Imputation is done by either using historical data with a trend or by imputing an average within the imputation class. After that, calendarization is performed so that there is a revenue value for each month. This may involve breaking and rebuilding a transaction if the filing frequency of that business is longer or shorter than the calendar month, e.g. quarterly, annual or multi-weekly transactions. Finally, allocation is done to assign the proper revenues to the actual business establishment, which could be a disaggregation of a company that has more than 1 ‘establishment’, based on a set of allocators derived from STC’s Business Register.

Each month, between 800,000 and 1,000,000 transaction records are received from CRA. The monthly update of the GST data processes approximately 6 million transactions, as all of the transactions in the months ‘m-6’ to ‘m’ are reprocessed at every production cycle.

4.0 Use of Tax Data in Statistics Canada’s Economic Statistics Program

As previously mentioned, CRA provides a wide range of administrative data to STC. In this section we present some of the uses of tax data in STC’s economic statistics program. STC’s economic statistics program uses a mix of annual and sub-annual surveys which produce timely indicators for the Canadian System of National Accounts (CSNA) and many data users such as federal and provincial departments, business associations, etc. Tax data use in STC’s economic statistics program is restricted to so-called ‘simple’ units. Businesses in Canada are classified as either simple or complex according to their structure. Simple units are those businesses that have simple structures (i.e., single industry within a geographic region) which allows a direct one-to-one link between tax data and the unit as represented on STC’s business register. By having such a link, STC is able to assign the tax data to the appropriate industry and region combination. In contrast, complex units are those businesses that are involved in multiple industries or have revenue generated in multiple regions. Tax data are not used for these units due to the difficulty of allocating them down to the industry within geography level, a requirement of many economic surveys at STC.

In the following sub-sections we describe how tax data have been integrated into some of the annual and sub-annual business survey programs. We close out the section with a few words on the challenges that have been encountered during the integration.

4.1 Annual Surveys

In the late 1990’s, STC developed the Unified Enterprise Survey (UES) which now incorporates over

60 annual business surveys into a single framework that applies similar concepts and methodologies to all of the surveys. When first developed, the UES used annual tax data (T1 and T2) to estimate the contribution of the so called 'take-none' stratum only. The take-none stratum is made up of the smallest units who collectively represented at most 10% of the economic activity of the survey population. The take-none stratum contribution is estimated using tax data for both T1 and T2.

In the early 2000's, with the increased availability and quality of tax data, the decision was made to use tax data for some units who are in the so called 'take-some' strata. Units in the take-some strata are medium to large units who have a non-zero probability of being selected in the sample. For a portion (approximately 50%) of the simple units (one business activity) selected in the take-some strata, no contact is attempted by STC and their 'responses' are obtained directly from tax data for financial variables. For survey variables that have an equivalent tax counterpart, tax data are used and are treated as if they came directly from the respondent. The linking of survey and tax concepts is greatly aided by STC's Chart of Accounts (COA), which provides mappings between survey and tax concepts. For variables that do not have equivalent concepts, imputation methods are used to complete the 'tax replaced' record. While treating tax data as respondent data is not entirely correct from a methodological point of view, this approach was adopted since it allowed for the increased use of tax data with minimal changes to existing systems and processes. Integrating tax data this way into the annual business surveys has generated significant savings in collection costs and reductions in response burden with very little effect on the quality of the resulting estimates. In addition to using tax data as a 'replacement' for survey data, they are also used during the editing process and as values for imputation for nonresponding units.

The UES has been using tax data in this manner successfully now for many years but its systems are starting to show their age. This has led to a decision to redesign the UES and STC plans to use tax data more extensively than it does today. While plans are not finalized, STC is looking at using tax data for all 'simple' units in the population and not just for half of the sampled simple units.

4.2 Sub-Annual Surveys

At STC, three mission critical surveys (the Monthly Wholesale, Retail and Manufacturers surveys) use monthly GST tax data. In these surveys, similar to the annual survey approach, a sub-sample of simple units is selected for tax "replacement". However, when used in monthly surveys, tax data are modified through a model to account for possible conceptual differences and timeliness issues. While the annual tax data are linked to survey concepts through the Chart of Accounts, no such link exists for monthly tax data. In addition, monthly tax data are not always available for the reference month in time to meet STC's tight production timeframes for monthly surveys. Thus, monthly tax data from month m-1 or even m-2 are used for reference month m. A sample of simple units is contacted in order to have some units that have both GST tax and survey data from which a model can be developed. Once obtained, the model is applied to the GST tax values to generate "survey" data for the units identified for tax replacement. These units are then processed as if the values were reported by the respondent. That is, the modified tax data go through the production process (edit and imputation, review by analysts and weighted estimation) as usual.

It is recognized that, from a methodological point of view, this approach is not the most efficient use of tax data as they are used for only a small number of units. In addition, by treating them as respondent data, variances are more than likely underestimated. However, this approach has allowed the introduction of tax data into an existing survey program with only minor changes to the processing system and review process. Since the introduction of tax data into the Monthly Wholesale, Retail and Manufacturers surveys, two other surveys have either undergone a redesign or have been developed. STC has capitalized on these opportunities

to significantly increase the use of GST data. These two surveys are the Monthly Food Services Survey (MFS) and the Quarterly Industry Revenue Indices (QIRI).

4.2.1 Monthly Survey of Food Services

The Monthly Survey of Food Services (MFS) produces estimates of the value of sales and the number of locations of restaurants, caterers and drinking places as defined by the North American Industry Classification System (NAICS), which is a classification system broken down by detailed business activity. The MFS sample design combines elements of a typical design used for business surveys at STC with a strategy that takes full advantage of tax data. In both designs, the first step is to stratify by NAICS and geographic region (provinces and territories). In some business type/geographic region groups a further stratification is performed according to whether or not the unit is complex or simple. Units are selected in both strata but during estimation a ratio estimator is used in the stratum containing simple units with population totals coming from the GST file. This methodology is an improvement over the one employed in the other monthly surveys since tax data are being used for the entire population of simple units: tax data are not being treated as respondent data and quality indicators produced for the ratio estimator are methodologically correct. In addition, the strong correlation between the GST data and the survey variables has allowed a significant decrease in the number of simple units needing to be sampled while still maintaining quality estimates.

4.2.2 Quarterly Industry Revenue Indices

As in most countries, the service sector represents an increasingly large portion of the economy; however, our knowledge of this sector remains somewhat incomplete. While STC has a strong program of annual services surveys, there is a lack of quality data at the sub-annual level. Until now STC has relied heavily on employment data to produce sub-annual measures of output for the services sector. Given that many services industries are dominated by businesses that tend to be small non-employers, the use of employment data as a proxy for output data poses a risk of bias. In response, STC has recently developed the Quarterly Industry Revenue Indices (QIRI) which relies heavily on the use of GST data to produce a quarterly index for revenues for 22 service industries. In order to be relevant these indicators are published 90 days after the end of the reference quarter at a semi-aggregated business activity level, by region and at the national level. In addition, revised estimates are available 90 days later (i.e., 180 days after the end of the reference quarter).

Early in the development of the QIRI it was realized that the use of GST data for complex units would be problematic. Although allocated GST data existed, the allocation methodology was still under development and its resulting quality was unknown. On the other hand GST data for simple units, which did not require allocation, was a rich source of timely data. The decision was made to use traditional survey methods for the complex units and GST data only for the simples. Estimates for the complex and simple units would be calculated separately and then combined to produce a quarterly index. Given that no model would be used to adjust the GST data for potential conceptual differences, the resulting estimates of level might not be coherent with those produced by the annual survey. It was for this reason that an index is produced and not quarterly estimates of levels. While there was some concern with the levels produced by the GST, many studies have shown that the tax data track trends or movements very well.

4.3 Challenges of Using Tax Data

While STC has successfully integrated tax data into its economic statistics program, it has not been

without many challenges, some of which are still being addressed today. One challenge being faced by both annual and sub-annual surveys is the reduction of timely feedback that accompanies an increased use of administrative data. Survey feedback, information obtained during a survey contact with a respondent, is a major source of updates in terms of changes in industrial activity and business status. At STC, survey feedback is used to keep the Business Register information up-to-date, which leads to improved information for all surveys and not just the one that provides the feedback. With an increase in the use of tax data, many units will no longer be contacted and an important source of information will be lost. However, given the strong partnership with CRA, steps are currently being taken to address this issue by asking for relevant economic activity information through the tax filing process. This collaborative initiative is well under way but is not yet fully operational.

Another challenge that STC is facing with the use of tax data is the lack of non-financial information. Some STC surveys are required to produce estimates of commodities and characteristic variables (ex. inventories, origin and destination of shipments, etc.) that are not available through administrative data and are not well correlated to tax information. Current plans are to continue to perform surveys so that estimates of these quantities can be produced.

A challenge that all users of administrative data face is the potential to lose control of an important data source. While the likelihood of this occurring between STC and CRA is low, given their strong partnership, it is possible that a statistical agency could be so heavily dependent on administrative data that a change brought on by the source of the administrative data could have disastrous consequence on statistical programs. This possibility only highlights the importance of developing a partnership such as the one enjoyed by STC and CRA.

5.0 Summary

Like many statistical agencies around the world, STC is becoming increasingly more reliant on administrative data as more and more pressure is applied on it to reduce collection costs and respondent burden. STC has been fortunate to have developed such a strong partnership with CRA so that it has been able to take advantage of the numerous administrative data files available from CRA. This partnership did not happen overnight nor has it always been smooth, but with the support of very senior managers and the recognition that such a partnership is good for Canada as a whole, a very stable relationship has developed.

The use of tax data in STC's economic statistics programs discussed in this paper has focused mainly on the business statistics side of STC activity and is only the tip of the iceberg in terms of tax data use at STC. On the household survey side, tax data are being used as frame information and as replacement for respondent reported data. In surveys that require household financial information, respondents are given the option of granting STC permission to use their tax information directly from CRA instead of providing it on the questionnaire. This option is currently being offered to respondents of the National Household Survey which has replaced the Census of Population long-form questionnaire. Tax data are used in many other programs at STC such as the System of National Accounts, the Survey of Employment, Payroll and Hours, Census undercoverage studies, agricultural surveys and STC's Business Register, which provides the survey frames for essentially all business surveys at STC. Covering the details of these uses in this paper was not possible, but information is available from the authors for those who are interested.

As one can see, STC is making use of a significant amount of administrative data from CRA, none of which would be available without a well functioning partnership and strong legal framework that allows for the sharing of microdata.