

Administrative Data in Statistics Canada's Business Surveys: the Present and the Future



Wesley Yung, Eric Rancourt and
Mike Hidioglou,
Statistics Canada

Outline



- Introduction
- Tax Data
- Current Uses
- Possible Future Uses
- Combined Framework
- Summary

Introduction



- Statistics Canada's (STC) business surveys use both annual and monthly admin/tax data
- Annual tax data - T1 and T2
- Monthly tax data – GST, Payroll Deductions
- Collected by Canadian Revenue Agency (CRA) and provided to STC

Introduction



- STC ‘processes’ the data
- Use in the past has been mostly for data confrontation, quality studies, frame updates, etc.
- More recently, tax has been used to ‘replace’ survey data
- In the future, tax will be used...

Tax Data, cont.



- Annual data – T2
 - Covers incorporated businesses
 - Corporations choose fiscal year, but must file with 6 months of year end
 - Entire universe provided to STC in GIFI format
 - STC receives data continuously
 - In October of RY+1, full universe of micro-data made available to survey programs

Tax Data



- Monthly data – GST
 - Goods and Service Tax of 6%
 - Variables available: Sales and total GST
 - Micro-data available 7 weeks after the end of reference month

Current Uses



- Sampling frame
 - For most business surveys at STC, sampling frame comes from the Business Register (BR)
 - BR uses signals from CRA for births
 - Tax data used to confirm business activity
 - Size measure from tax or modeled using other admin data

Current Uses, cont.



- Collection
 - Tax data not used in collection, but collection is reduced by use of tax data
 - Tax data is used instead of contacting respondent for certain ‘simple’ units
 - ‘Simple’ units are businesses for which there is a one-to-one link with tax data
 - Note that there are many ‘simple’ units, but they tend to be small businesses

Current Uses, cont.



- Annual surveys use T2 data and direct replacement
 - Coherence is ensured by Chart of Accounts (COA) that link survey questionnaires to GIFI format
- Monthly surveys use GST and models
 - Models account for conceptual issues
 - Sales on GST may not correspond exactly to sales in survey

Current Uses, cont.



- Edit and imputation
 - Tax data used for data confrontation
 - Used to impute for non-respondents in some surveys

Current Uses, cont.



- Estimation
 - In addition to units identified for tax replacement, tax data used for estimation below a threshold
 - As of now, annual and monthly tax data not used for calibration purposes

Current Uses, cont.



- Summary
 - Tax data used in many steps of survey process
 - Tax data has been incorporated into a **survey based** framework
 - Incorporated into existing surveys and systems
 - Treating tax data as data from another source
 - Use of tax data has been constrained
 - It has not been used to its fullest extent possible

Possible Future Use



- **Tax based framework**
 - What if all constraints were removed?
 - Admin data from CRA would define the target population
 - All financial data would come from tax and would have a census
 - Seems easy and logical, but there are some outstanding issues

Possible Future Use, cont.



- Outstanding issues
 - BR and tax frames not 100% coherent
 - Differences should be minimized to ease transition
 - No more feedback to BR
 - Need a mechanism to obtain up-to-date information (status, activity, changes, ...)
 - Perhaps a Nature of Business Report (NBR)?

Possible Future Use, cont.



- Contact of ‘non-simple’ (complex) units
 - Complex structures cause problems for tax data
 - Information may not be available at level required
 - Need to collect information at level required or information to perform allocation
- No sampling error but need to quantify non-sampling error
 - Imputation of tax data, tax data not observed without error, etc.

Possible Future Use, cont.



- No characteristics data from tax
 - Use models?
 - Combine characteristics questions with NBR?
- STC becomes dependent on CRA
 - STC would not have direct control over data
 - CRA uses data for different purpose
- A switch of framework **will** introduce a break in series
 - STC and users must be willing to accept this

Combined Framework



- Until issues are resolved, can use a combination of survey and tax frameworks
- One possible way:
 - First stratify population into simple and complex units
 - For complex units, use existing survey framework

Combined Framework, cont.



- For simples, use tax for financial variables
 - Annual surveys, use direct replacement
 - For monthlies, a sample maybe necessary to account for time and conceptual differences
- A sample of simples could be contacted for NBR
 - This information may be used to estimate death rates, misclassification rates, etc. for adjustment purposes

Combined Framework, cont.



- If characteristics needed, combine with NBR
 - Weighted estimates could be used – might necessitate a large sample
 - Use models to relate characteristics from sample with variables on tax file

Summary



- Tax data now used in many survey steps
- Use is still constrained by existing sample designs, systems and operational practices
- Several issues need to be resolved before moving to a tax based framework
 - Surveys would probably still be needed for NBR and characteristics
- In the meantime, a combination of survey and tax can be used



Pour plus d'information, veuillez contacter:
For more information, please contact:

Wesley.Yung@statcan.ca

(613) 951-4699