

# THE USE OF THE GOODS AND SERVICES TAX BY THE MONTHLY RESTAURANTS, CATERERS AND TAVERNS SURVEY

Zak Pritchard<sup>1</sup> and Charles Tardif<sup>2</sup>

## ABSTRACT

The Monthly Restaurants, Caterers and Taverns Survey (MRCTS) collects sales data from a sample of restaurants, caterers and taverns in Canada. The sales estimates, an input to the System of National Accounts, are used by various agencies to develop national and regional economic policies and programs. Since May 2004, data from selected companies in the sample have been modeled using Goods and Service Tax (GST) data provided by the Canada Revenue Agency (CRA). We will discuss the usage of GST data by the MRCTS and its effects on the survey in terms of data quality and response burden.

KEY WORDS: Data quality; response burden; tax data.

## RÉSUMÉ

L'enquête mensuelle sur les restaurants, traiteurs et tavernes (EMRTT) recueille les ventes d'un échantillon de restaurants, traiteurs et tavernes au Canada. Les estimations de l'EMRTT sont partie intégrante des comptes nationaux et sont utilisées par divers membres de l'industrie pour fins de planification. Depuis mai 2004, les données de la Taxe sur les produits et services (TPS) provenant de l'Agence de revenu du Canada (ARC) sont utilisées pour modéliser les ventes des compagnies échantillonnées. Nous discuterons de l'utilisation de la TPS par l'EMRTT et de son impact sur l'enquête au niveau de la qualité des données et du fardeau de réponse.

MOTS CLÉS : Données fiscales, fardeau de réponse, qualité des données.

## 1. THE MONTHLY RESTAURANTS, CATERERS AND TAVERNS SURVEY (MRCTS)

### 1.1 Survey Description

The Monthly Restaurants, Caterers and Taverns Survey (MRCTS) collects data on total sales and the number of locations by province, territory and industry from a representative sample of restaurants, caterers and taverns in Canada. The estimated data from MRCTS are used as input to the System of National Accounts (SNA). It is used by the federal government to estimate provincial taxation shares and by food service agencies for planning purposes.

Each month, Statistics Canada's Business Register Division produces a survey universe file (SUF), used by the MRCTS to select the sample. The SUF is a list of all establishments that are in the business as of the beginning of the month. The MRCTS currently samples at the company level, where a company consists of at least one establishment. The population consists of employing companies that belong to one of the following North American Industrial Classification System (NAICS) codes:

- 722110: Full-Service Restaurants - Restaurants where patrons order while seated and pay after eating.
- 722210: Limited-Service Restaurants - Restaurants where patrons order at a counter, and/or order by phone and pay before eating.
- 722310: Food Service Contractors – Companies that supply food services under contract for a specific period of time.
- 722320: Social Caterers – Companies that provide food services for social or business events.
- 722330: Mobile Food Services – Companies that serve food and beverages, from motorized vehicles or non-motorized carts.
- 722410: Drinking Places - Prepare and serve alcoholic beverages and limited food services. (e.g., bars, night clubs, taverns)

---

<sup>1</sup> Zachary Pritchard, Statistics Canada, Ottawa, Canada, K1A 0T6, pritzac@statcan.ca

<sup>2</sup> Charles Tardif, Statistics Canada, Ottawa, Canada, K1A 0T6, tardcha@statcan.ca

The size of the frame, as of the May 2006 reference period, is 76,600 companies. Note that some of the companies in the frame are "dead." That is, they are no longer in business or are no longer in the scope of the survey (e.g. they may have become non-employed companies). Dead units remain in the sample to represent other dead units that remain in the frame.

## 1.2 Sampling

The MRCTS sample is a stratified random sample, with each company belonging to one of 234 possible strata (13 provinces/territories \* 6 NAICS codes \* 3 revenue cut-offs). The MRCTS sample is updated with births at the start of each month. That is, the sample is the same from month to month, with the exception of the births. Some companies are always in the sample, including large or complex companies (e.g. companies that operate in more than one province) companies that operate in Yukon, Nunavut and the Northwest Territories. The current size of the MRCTS sample, as of the May 2006 reference period, is 4,100 companies. As is the case in the frame, some of the companies in the sample are dead. These dead companies remain part of the sample until the next re-stratification to represent dead companies in the population.

The sample size quoted above is the complete sample size, which is larger than the collected sample size. The former includes (1) companies that report sales data for several companies in the sample; and (2) dead companies, which represent dead companies in the population. The companies in the latter group have been a cause for concern, as deaths in sample have increased at a faster rate than have births since the beginning of 2006. This fact, coupled with the fact that some companies have refused to submit data to the MRCTS, has lowered the quality of MRCTS sales estimates.

A mix of long- and short-term solutions has been proposed to deal with the problem of maintaining data quality. In the long-term, the survey will undergo a complete redesign, which will include changes to its sampling and estimation systems. Agreements with some restaurant chains to report data for all of their franchisees have recently been reached, and such agreements with other chains are in progress. In the short-term, the survey is increasing its usage of administrative data, with one source being data submitted by business for the collection of Goods and Services Tax (GST).

## 2. THE GOODS AND SERVICES TAX (GST) AND THE MRCTS

### 2.1 The GST

The GST, a value-added-tax that has been collected by Canadian businesses since January 1, 1991 is remitted to the Canada Revenue Agency (CRA). Approximately seven weeks after the start of each month, the CRA sends an administrative data file corresponding to that reference month to Statistics Canada's Tax Data Division (TDD). TDD then edits the data as required. At the time of remission, businesses, uniquely identified by 15-digit business numbers, report additional information, including the operating revenue on which the tax has been calculated. The frequency of remission, ranging from annually to monthly, increases with the amount of GST remitted. Although few businesses remit GST revenue on a monthly basis, most GST revenue is remitted on this frequency (See Table 1). For more information on the GST, please refer to Brodeur and Pierre (2003).

**Table 1: GST Revenue Remission Statistics**

Reference Month: April 2006	Frequency		
	Monthly	Quarterly	Annually
Percentage of businesses remitting GST revenue at this frequency	7.28%	57.28%	35.44%
Percentage of GST revenue remitted at this frequency	81.79%	15.48%	2.73%

The administrative data file cannot be used directly as a replacement for the MRCTS, with the main reason being that there is not a direct correspondence between an MRCTS company number and a BN. Certain enterprises may remit GST under several business numbers, or several MRCTS companies may share the same BN.

In the case of most MRCTS companies, however, there is a direct link between a company number and a BN. There is a high correlation between a company's reported sales and the operating revenue corresponding to its BN on the GST data file (see Table 2). The MRCTS has been making use of this fact since the May 2004 reference period, using the revenue to impute sales for some companies and to model sales based on GST for others. The former is necessary when a company has refused to submit sales data to the MRCTS during a particular month or its data are not yet available. The latter has permitted the MRCTS to reduce response burden. Approximately 600 companies now have their sales figures determined by the application of this model. In both cases, the MRCTS employs a ratio method to estimate sales revenues reported by a group of companies during the current period in the numerator and sales revenue from the GST data file from a previous period in the denominator. As a measure of the amount of modeling, around one fifth of the companies in the sample had their sales data modeled during the February 2006 reference period, and modeled sales accounted for around 30 % of the national MRCTS sales estimate.

**Table 2: Correlations between revenue from the GST file and sales (MRCTS companies only)**

Reference Month	Correlation
April 2006	0.89
March 2006	0.91
February 2006	0.94

## 2.2 Imputation of sales using revenue from the GST data file

This type of ratio imputation, also known as GST Trend Imputation, imputes sales for companies that are directly surveyed, but have not submitted data to the MRCTS. It is the third choice for imputation of surveyed companies, after imputation based on sales reported 12 months ago ("last year imputation") and during the previous month ("previous month imputation"). There are five levels of imputation, listed below in order of preference. Imputation will only take place at the highest level with at least 10 donors in the corresponding imputation group.

1. Imputation initially takes place at the stratum level (province-NAICS-revenue size). Although there are 234 possible imputation groups at this level, the actual number of groups will be smaller. As there must be a minimum of 10 companies in an imputation group before this type of imputation will take place, its usage is limited (most of the time) to restaurants in Alberta, British Columbia, Manitoba, Ontario and Quebec.
2. The second level of imputation is the modified stratum level, with 156 possible imputation groups (the two smaller revenue cut-offs in each province-NAICS combination are combined). Missing data from restaurants in the smaller provinces and drinking places in most provinces tend to be imputed at this level.
3. The third level is the Province \* NAICS, with 78 possible imputation groups. Missing data from caterers and drinking places in most provinces are often imputed at this level.
4. The fourth level is the Combined Province \* NAICS level. There are now seven "regional groupings," Alberta, British Columbia, Ontario, Quebec, the Atlantic Provinces the Prairie Provinces and The Territories, resulting in a total of 42 imputation groups. Missing data from larger restaurants in the territories are often imputed at this level.
5. The final level is the NAICS level, with six possible imputation groups. Missing data from smaller restaurants, caterers and drinking places in the territories are often imputed at this level.

## 2.3 Modeling of sales using revenue from the GST data file

As mentioned earlier, most companies in the sample are directly surveyed. When the MRCTS began modelling sales using revenue from the GST data file, it restricted the modelling to restaurants in Alberta, British Columbia, Ontario and Quebec, the provinces that account for over 80% of Canada's population in the food services and drinking places industry on Statistics Canada's Business Register (BR). Starting in May 2004, one month after the most recent re-stratification of the sample, 244 companies (labelled as "S2" companies) were randomly selected to have their sales modeled using revenue from the GST data file. At the same time, 170 companies (labelled as "S1" companies) were randomly selected

to be directly surveyed, with their sales data used in the modelling of the sales of the S2 companies. Additional S1 and S2 companies were randomly selected in February 2005.

In addition to restricting the GST modeling to restaurants in the most populous provinces, the MRCTS also conducted the modeling at one level, the province \* NAICS level. In the case of S2 companies, GST modeling is the first choice of modeling sales. In order for an S2 company to be modeled in this manner, its province \* NAICS combination must be the same as that for at least 10 S1 companies that have reported data during the month of interest. If the above condition does not hold, the company's sales will be modeled using other means, such as by Gross Business Income (GBI), currently available on the Business Register (BR).

As an example of this type of modelling, suppose that, as of the reference month February 2006, a limited-service restaurant in Ontario is an S2 unit. Then its modeled sales will be calculated as follows:

$$\begin{array}{rcl}
 \text{Modeled Sales} & = & \text{Revenue from the GST data file} \\
 & & \text{corresponding to the S2 company} \\
 & & \text{(January 2006 Reference Period)} \quad * \\
 & & \frac{\text{Mean Sales of Limited-Service Restaurants} \\
 & & \text{in Ontario that are S1 companies} \\
 & & \text{(February 2006 Reference Period)}}{\text{Mean Revenue from the GST data file for} \\
 & & \text{Limited-Service Restaurants in Ontario that} \\
 & & \text{are S1 companies} \\
 & & \text{(January 2006 Reference Period)}}
 \end{array}$$

Note that it is necessary to use the revenue from the previous reference month, as monthly MRCTS estimates are produced approximately two weeks before GST data corresponding to the same month become available.

### 3. PRESENT DEVELOPMENTS

The MRCTS has steadily increased its usage of administrative data since the beginning of 2006. Firstly, based on recommendations from the client division responsible for the survey, some companies in the sample that were to have their sales data collected but have refused to submit data have had their status changed to S2 and will now have their sales modeled using revenue from the GST data file. Prior to changing the status of a company, it must be verified that the business number corresponding to the company is not shared with another company, and it does not have more than one account on the GST data file.

In February 2006, to counterbalance an increase in sample deaths, around 300 companies were randomly selected and added to the sample some of which will be modeled using revenue from the GST data file.

GST Modeling, which had been restricted to restaurants in the most populous provinces, has been extended to other kinds of businesses and is now used in all provinces. As well, this new process uses the same five modeling levels as those corresponding to GST Trend Imputation. The extension of GST modeling was made possible partly by the increase in the sample size in February 2006 and partly by treating some of the units in the MRCTS sample that were surveyed but not in the S1 pool of companies as if they were S1 companies.

### 4. FUTURE DEVELOPMENTS

The way in which revenue from the GST data are used for modeling is currently under review. As discussed earlier, GST modeling makes use of data reported during the previous month. It is known that some restaurants, particularly those in resort areas, experience seasonal variation in sales, and it may be better to model sales using data reported 12 months ago rather than during the previous month. This pattern in terms of month to month variation in sales becomes very apparent each January where sales dip after the holidays shopping season has ended.

Furthermore, GST modeling may be extended to complex companies, now that GST revenue breakdowns by province \* NAICS \* establishment level are available.

Some new measures of GST data quality have been developed. For example, it is now possible to separate the percentage of the national sales estimate that has been calculated using revenue from the GST data file from that which has been reported by the respondent rather than being estimated.

A complete redesign of the MRCTS is currently in progress. The principal change to the survey will be at the level of the sampling unit. The sampling unit will change from the company level to the establishment level, putting the MRCTS in line with other monthly business surveys. Revenue thresholds will be established whereby establishments falling below the thresholds will have their sales estimated by revenue from the GST data file instead of being surveyed directly. The creation of the thresholds will result in smaller sampling weights for establishments that will be surveyed. Finally, the MRCTS population, which currently includes employing companies only, will include employing and non-employing establishments after the redesign.

## 5. CONCLUSION

The usage of administrative data has been a benefit to the MRCTS, resulting in higher data quality – (e.g. the weights assigned to each unit in sample have been reduced and reduced response burden. The survey will continue to make use of administrative data, such as GST, to improve data quality.

## REFERENCES

Cochran, W.G. (1977). *Sampling Techniques* (3rd. ed.). New York: John Wiley.

Rubin, D.B. (1976). “Inference and missing data”. *Biometrika*, **63**, 593-604.

Brodeur, M. and Pierre, L. (2003) “Use of Tax Data: An Application of Goods and Services Tax Data” Proceedings of the Statistics Canada Symposium 2003

Statistics Canada (2004). *Restratification and Methodology Changes in the Monthly Restaurants, Caterers and Taverns Survey (MRCTS)*. Survey documentation available as metadata on Statistics Canada website at [http://www.statcan.ca/english/sdds/document/2419\\_D3\\_T9\\_V1\\_E.pdf](http://www.statcan.ca/english/sdds/document/2419_D3_T9_V1_E.pdf)