

Technical Details on Estimation in the Monthly Survey of Food Services and Drinking Places (MFS)

For the MFS, total sales are estimated using two types of estimation methodologies. The first is the Horvitz-Thompson estimator and amounts to multiplying each data response by its sampling weight. For the take-all strata, the weight is one since all enterprises in the stratum are selected in the sample. For the take-some strata, a random sample is selected and the weight is equal to the inverse of the probability of selection.

The second type of estimation methodology is ratio estimation. It is used in the population represented by the take-some strata of North American Industry Classification System (NAICS) 7221 and 7222, in Quebec, Ontario, Manitoba, Saskatchewan, Alberta and British Columbia. It is also used for the take-none strata for all NAICS and provinces.

In the MFS, ratio estimation improves the quality of the estimate by taking advantage of the high correlation between survey data and auxiliary information. Survey data is the sales as reported by respondents and auxiliary information is the revenue as reported to the Goods and Service Tax (GST) program that is administered by the Canada Revenue Agency (CRA). To calculate the ratio estimator:

- Both the GST revenue and survey sales data are obtained for a sample of units. As well, the GST revenue data are available for all of the non-sampled units in the population.
- An estimate of sales is calculated from the survey data using the sampling weights (Sales_est). This is the Horvitz-Thompson estimate of sales.
- An estimate of GST revenue is calculated from the GST data, based on sampled units only, using the sampling weights (GST_est). This is the Horvitz-Thompson estimate of GST revenue.
- The GST revenue data are summed for all units, both the sampled and non-sampled, to obtain the known total of GST revenue (GST_total).
- The ratio (GST_total / GST_est) is calculated and is called the g-weight.
- The ratio estimate of sales is equal to the Sales_est multiplied by the g-weight.

That is, the Ratio estimator (RE) is given by,

$$RE = Sales_est * (GST_total / GST_est)$$

From the sample, the estimate of the (known) GST_total is GST_est. If GST_est is larger than the GST_total then we expect (since there is a strong relationship between sales data and GST data) Sales_est to be larger than the (true and unknown) total sales. This results in a g-weight, (GST_total / GST_est), of less than 1 and an RE value less than Sales_est. If instead the GST_est is smaller than GST_total, then the g-weight is larger than 1 and the RE value would be greater than Sales_est.

The ratio estimation approach is currently being used to produce estimates for the MFS beginning with the January 2009 estimates, and replaces the ratio model approach used previously.

The advantage of the new ratio estimator is that it allows for an earlier detection of businesses no-longer operating ("deaths") since business closures for sampled units are detected immediately and are used in the calculation of the ratio estimate. These survey units would contribute a value of zero, thereby lowering the overall estimate and the weight of these units would represent other business closures that were undetected in the GST auxiliary data.

The old ratio model approach used preliminary or early GST data where timely information on business closures was lacking. This is because non-reporting businesses are initially assumed to be "alive" on the monthly GST file to account for late remissions of monthly remitters and remissions from quarterly and annual remitters. After a pre-determined amount of time and no further remissions are received or expected, the business is eventually considered "closed" as of the last remittance date.

While early versions of GST data do not reflect true deaths immediately, as later and more updated versions of GST data become available from CRA, those data more accurately reflect survey responses by finally recognizing business closures. Therefore, because the previously published revised estimates prior to 2009 - which are based on the former ratio model approach - use the latest and most updated version of GST data that reflect business closures, those older estimates are compatible with the estimates that are derived from the new ratio estimation approach and, thus, need no further revision.

Take-none:

There is no sample for the take-none strata. Instead, sales are estimated using the ratio estimation approach for all provinces and NAICS based on the data from the take-some strata.

Measures of accuracy:

The standard error and coefficient of variation (CV) of the estimates are derived from the sample design and estimation method using the collected survey data.