Methodology for the Prepared Food and Beverage Sales Survey, 2009

Michèle Picard-Flibotte
Javier Oyarzun
Eric Pelletier

Business Survey Methods Division
Statistics Canada

May 21, 2010

Table of contents

Table of contents 2
1. Survey objective 3
2.Survey frame 3
3. Sampling and stratification 3
4. Edit and imputation 5
4.1 Definition of completed questionnaire 5
4.2 Convention for indicating responses 5
4.3 General selection criteria for donors 6
4.4 General selection criteria for donors to be used in imputation 6
4.5 Derived variables 6
4.5.1 Strata 6
4.5.2 Region 6
4.5.3 Industry group 7
4.6 Edit and Imputation Rules ..... **Error! Bookmark not defined.**
Question C0300: ..... **Error! Bookmark not defined.**
Question C0401: ..... **Error! Bookmark not defined.**
Question C0400: ..... **Error! Bookmark not defined.**
4.7 Imputation rate 7
5. Outliers ..... **Error! Bookmark not defined.**
6. Weighting 8
7. Response rate 9
8. Estimation and confidentiality ..... **Error! Bookmark not defined.**

1. Survey objective

This new survey is intended as a quick study of sales below $4 in the restaurant and prepared food and beverages industry in Ontario. It is in the form of a very short questionnaire (two questions) and allows for evaluating the proportions of sales that are currently non taxable. This is being done in order to eventually coordinate the sharing of a harmonized sales tax in that province.

2. Survey frame

The Business Register is used to create the survey frame. The statistical unit is the establishment.

3. Sampling and stratification

To select the sample, the following parameters were considered:

Target size of 1,500 units
Desired response rate of 60%
CVs targeted by domain
- 10% for industry groups 6, 11 and 12
- 5.8% for the other industry groups

The domains are the following:

Table 1: NAICS included in each domain

Domains: Industry Group	**NAICS included**
1- Bakeries	NAICS 311811 and 445291
2- Cafeterias; coffee shops; lunch counters and snack bars; doughnut and muffin shops; limited-service restaurants	NAICS 722210
3- Chip wagons; hot dog stands; coffee stands; mobile food services	NAICS 722330
4- Cocktail lounges, taverns and bars	NAICS 722410
5- Convenience stores	NAICS 445120 and 447110
6- Convention centres	NAICS 531120
7- Food service contractors	NAICS 722310
8- Full-service restaurants	NAICS 722110
9- Grocery stores	NAICS 445110
10- Hotels and motels	NAICS 7211
11- Lodging houses	NAICS 7213
12- Private or social clubs or legion halls	NAICS 813410
13- Snack bars	NAICS 512130, 7111, 7112, 71131, 712, 7131, 713210, 713299, 7139
14- Vending machines*	NAICS 454210

* Only vending machine operators that sold meals were included in the frame.

Each domain was stratified by size based on revenues to obtain one take-none stratum, two take-some strata and one take-all stratum. However, as this is the first time for this survey and the relationship between sales of prepared food and beverages under $4 and total sales of prepared food and beverages is unknown, a sample was selected from among the take-none strata. The generalized sampling system was subsequently used to allocate according to the total cost minimization and sample selection method.

Table 2: Population and sample sizes by NAICS code and stratum

GroupeInt	NAICS(NAICS)	Sizeclasslh								Total	
		0		1		2		3			
		Take-None		Small Take-Some		Medium Take-Some		Take-All			
		N	n	N	n	N	n	N	n	N	n
	311811	437	39	2	0	52	10	11	11	502	60
	445291	285	19	151	12	45	10	10	10	491	51
1	Total	722	58	153	12	97	20	21	21	993	111
2	722210	5840	30	5767	58	1099	52	18	18	12724	158
3	722330	283	13	98	8	23	7	7	7	411	35
4	722410	540	12	441	20	178	22	18	18	1177	72
	445120	1741	19	3265	32	145	5	4	4	5155	60
	447110	912	8	49	1	481	8	19	19	1461	36
5	Total	2653	27	3314	33	626	13	23	23	6616	96
6	531120	17223	87	5048	52	488	50	47	47	22806	236
7	722310	667	8	0	0	238	8	13	13	918	29
8	722110	5615	28	5097	52	1453	52	10	10	12175	142
9	445110	2297	23	713	33	271	35	6	6	3287	97
	721111	537	7	3	0	201	7	48	48	789	62
	721112	69	1	94	0	18	0	0	0	181	1
	721113	160	2	73	2	32	1	7	7	272	12
	721114	339	2	544	7	8	0	0	0	891	9
	721191	432	5	27	0	0	0	0	0	459	5
	721192	410	5	83	1	2	0	0	0	495	6
	721198	111	0	26	0	7	0	0	0	144	0
10	Total	2058	22	850	10	268	8	55	55	3231	95
11	721310	37	7	51	7	21	7	9	9	118	30
12	813410	770	8	1532	25	316	33	55	55	2673	121
	512130	122	1	0	0	108	7	0	0	230	8
	711111	194	4	210	2	8	0	2	2	414	8
	711112	56	2	9	0	8	0	2	2	75	4
	711120	31	0	37	0	1	0	0	0	69	0
	711130	484	5	604	6	17	3	0	0	1105	14
	711190	17	1	27	1	0	0	0	0	44	2
	711211	182	2	5	0	19	1	3	3	209	6
	711213	668	4	303	3	29	0	7	7	1007	14
	711218	215	0	13	1	3	0	0	0	231	1
	711311	106	0	23	1	7	0	2	2	138	3

	711319	79	0	24	0	9	0	1	1	113	1
	712111	36	1	30	1	3	0	1	1	70	3
	712115	3	0	8	0	3	0	2	2	16	2
	712119	21	1	32	0	8	0	1	1	62	2
	712120	6	0	13	0	1	0	0	0	20	0
	712130	24	1	7	0	2	0	2	2	35	3
	712190	22	0	18	0	5	1	1	1	46	2
	713110	51	1	6	0	7	1	5	5	69	7
	713120	114	0	11	0	6	1	0	0	131	1
	713210	2	0	2	0	2	0	4	4	10	4
	713299	72	1	27	0	31	2	2	2	132	5
	713910	349	7	255	4	217	21	4	4	825	36
	713920	35	0	5	0	17	1	0	0	57	1
	713930	136	0	229	5	29	1	0	0	394	6
	713940	917	11	599	9	113	11	4	4	1633	35
	713950	66	0	137	3	10	1	0	0	213	4
	713990	820	6	754	11	51	6	2	2	1627	25
13	**Total**	4828	48	3388	47	714	57	45	45	8975	197
14	**454210**	386	7	18	18	97	14	28	28	529	67
	Total	43919	378	26470	375	5889	378	355	355	76633	1486

4. Edit and imputation

This part contains the specifications for the edit and imputation. Actions (edits and imputation) are listed in the order which they should be carried out. Further directions are given with respect to the creation of a series of derived variables. These derived variables form a key part of the imputation strategy and therefore their creation is the first step in the process.

Edit rules are used to identify and correct records with inconsistent, incomplete or invalid responses. Where no direct correction is possible, the unacceptable values will be replaced by imputed values. All edits are indicated with a circular bullet.

Imputation indicates the manner in which the record that was not corrected through edits will be corrected. As a general rule, it has been determined that we will not engage in imputation methods which might introduce bias. This means that donor imputation will be used applying the general selection criteria below, either alone or in the context of a specified value or nearest neighbour search using some other variable.

Edit and imputation will be carried out on completed questionnaires only.

4.1 Definition of completed questionnaire

A questionnaire is considered to be complete if one of these questions has been filled out.

4.2 Convention for indicating responses

The questionnaire contains two questions eliciting an accepted range of responses. The following provide an explanation of how they appear in this text:

> C0300) Please report your total sales of food and non-alcoholic beverages in 2009 in dollars $: We expect a number ranging from 0 $ to 999 999 999 997 $.

For the second question the respondent has the option answering either at question 1a or 1b.

C0401) Of that total, please report the amount of sales that were PST-exempt in 2009 in percentage %: We expect a percentage ranging from 0 % to 100 %.

C0400) Of that total, please report the amount of sales that were PST-exempt in 2009 in dollars: We expect a number ranging from 0 $ to 999 999 999 997 $.

4.3 General selection criteria for donors

Throughout this document there is reference to 'general selection criteria'. This refers to the general method to be applied for donor selection during imputation. Donors are to be selected using ranked criteria, where rank 1 is considered the optimal donor, followed by rank 2, etc. Each rank has two possible levels of matching based on the industry detail.

4.4 General selection criteria for donors to be used in imputation

A. Industry group x Strata

B. Industry group

4.5 Derived variables

A series of derived variables are to be created for the purpose of implementing the imputation strategy, for the creation of statistical tables and for use by researchers and analysts.

4.5.1 Strata

This revenue size variable is based on the revenue variable derived from Business Register information. A strata use for sampling was developed using this information.

Strata	Criteria
Take-All	The take-all stratum which would consist of the largest enterprises.
Take-Some – Large	The take-some large stratum is composed of medium-size enterprises put into a substratum with a higher sampling fraction than the one containing the smallest take-some enterprises.
Take-Some – Small	The take-some small stratum is composed of medium-size enterprises put into a substratum with a smaller sampling fraction than the one containing the larger take-some enterprises.
Take-none	The take-none stratum is composed of smallest enterprises and has a smallest sampling fraction.

4.5.2 Region

Region	Criteria
Onta	Ontario

4.5.3 Industry group

Industry Group	Criteria
Bakeries	NAICS 311811 and 445291
Cafeterias; coffee shops; lunch counters and snack bars; doughnut and muffin shops; limited-service restaurants	NAICS 722210
Chip wagons; hot dog stands; coffee stands; mobile food services	NAICS 722330
Cocktail lounges, taverns and bars	NAICS 722410
Convenience stores	NAICS 445120 and 447110
Convention centers	NAICS 531120
Food service contractors	NAICS 722310
Full-service restaurants	NAICS 722110
Grocery stores	NAICS 445110
Hotels and motels	NAICS 7211
Lodging houses	NAICS 7213
Private or social clubs or legion halls	NAICS 813410
Snack bars	NAICS 512130, 7111, 7112, 71131, 712, 7131, 713210, 713299, 7139
Vending machines	NAICS 454210

4.6 Imputation rates

Table 3: Weighted imputation rate per question and per industry group

NAICS Groups		Weighted		
		C0300	C0401	C0400
	Total - Ontario	9.1%	9.8%	9.8%
1	Bakeries	5.8%	5.6%	5.6%
2	Cafeterias; coffee shops; lunch counters and snack bars; doughnut and muffin shops; limited-service restaurants	14.3%	19.4%	19.4%
3	Chip wagons; hot dog stands; coffee stands; mobile food services	0.0%	0.0%	0.0%
4	Cocktail lounges, taverns and bars	0.0%	13.8%	13.8%
5	Convenience stores	22.2%	27.4%	27.4%
6	Convention centres	1.9%	0.0%	0.0%
7	Food service contractors	0.3%	0.6%	0.6%
8	Full-service restaurants	6.7%	6.4%	6.4%
9	Grocery stores	40.5%	16.0%	16.0%
10	Hotels and motels	0.0%	6.0%	6.0%
11	Lodging houses	1.6%	0.0%	0.0%
12	Private or social clubs or legion halls	4.0%	5.3%	5.3%
13	Snack bars	6.1%	8.1%	8.1%
14	Vending machines	0.0%	1.2%	1.2%

Table 4: Unweighted imputation rate per question and per industry group

NAICS Groups		Unweighted		
		C0300	C0401	C0400
	Total - Ontario	6.2%	11.4%	11.4%
1	Bakeries	7.0%	5.3%	5.3%
2	Cafeterias; coffee shops; lunch counters and snack bars; doughnut and muffin shops; limited-service restaurants	14.6%	16.7%	16.7%
3	Chip wagons; hot dog stands; coffee stands; mobile food services	0.0%	0.0%	0.0%
4	Cocktail lounges, taverns and bars	0.0%	20.0%	20.0%
5	Convenience stores	17.3%	30.8%	30.8%
6	Convention centres	0.8%	0.8%	0.8%
7	Food service contractors	8.3%	16.7%	16.7%
8	Full-service restaurants	5.3%	9.3%	9.3%
9	Grocery stores	28.2%	30.8%	30.8%
10	Hotels and motels	0.0%	16.7%	16.7%
11	Lodging houses	5.9%	0.0%	0.0%
12	Private or social clubs or legion halls	1.2%	6.0%	6.0%
13	Snack bars	5.2%	10.4%	10.4%
14	Vending machines	0.0%	6.8%	6.8%

The rates were obtained using the following formulas:

$$IR_{unweighted} = (ic)/nj$$

where ic:= imputed cells (if the cell is imputed, ic = 1, if not ic = 0) for respondent i to question j, nj:= number of cells responded for question j

$$IR_{weighted} = (ic_w)/nj_w$$

where ic:= imputed cells (if the cell is imputed, ic = initial weight, if not ic = 0) for respondent i to question j, nj_w:= sum of initial weights for cells responded for question j

5. Weighting

We used re-weighting to compensate for total non-response.

Correction for non-response

The units selected during sampling were grouped into four categories as follows:

Table 5: Response codes and their frequency in the sample

Grouped status code	Frequency
NR: non-response	465
OOS: out of scope	68
OOB: out of business	120
REP (INSCOPE): respondent in the scope of the survey	833

The weight of non-respondents is redistributed among respondents who are in-scope and the OOS and the OOB, using the strata from the sampling plan as a reweighting class. Normally, when there are not enough respondents in a given stratum, a regrouping of strata is necessary. In this survey, only two strata had only one respondent; one of these strata had no non-respondents and the other had a reasonable total of sales of prepared food and beverages. For these reasons, no strata were regrouped.

6. Response rate

To calculate the response rate, we calculated the unweighted and weighted response rates. The unweighted response rate calculated at the estimate is 56%, while the weighted response rate is 60%. The unweighted response rate gives an idea of the response rate at the sample level, and is often used during collection, while the weighted rate provides an overall idea, i.e., at the population level.

The rates were obtained using the following formulas:

a) $RR_{unweighted} = (c + p)/n = 833/1486 = 56.1\%$

Where c:= completed questionnaires, p:= partially completed questionnaires, n:= sample size

b) weighted: $RR_{weighted} = \dfrac{(c_w + p_w)}{(n_w - oos_w - oob_w - dup_w)} = \dfrac{36498}{(76633 - 5795 - 9945)} = 59{,}9\%$

where c_w:= weighted completed questionnaires, p_w:= weighted partially completed questionnaires, n_w:= size of weighted sample, oos_w:= weighted-out-of-scope, oob_w := weighted out-of-business, dup_w:= weighted duplicates.

Table 6: Unweighted response rates by industry group

NAICS Groups			pop	sample	RES	resrate_unweighted
	Total – Ontario		76633	1486	833	56.1%
1	Bakeries	311811 and 445291	993	111	57	51.4%
2	Cafeterias; coffee shops; lunch counters and snack bars; doughnut and muffin shops; limited-service restaurants	722210	12724	158	96	60.8%
3	Chip wagons; hot dog stands; coffee stands; mobile food services	722330	411	35	15	42.9%
4	Cocktail lounges, taverns and bars	722410	1177	72	35	48.6%
5	Convenience stores	445120 and 447110	6616	96	52	54.2%
6	Convention centres	531120	22806	236	126	53.4%
7	Food service contractors	722310	918	29	12	41.4%
8	Full-service restaurants	722110	12175	142	75	52.8%
9	Grocery stores	445110	3287	97	39	40.2%
10	Hotels and motels	721111, 721112, 721113, 721191, 721191, 721192 and 721198	3231	95	66	69.5%
11	Lodging houses	721310	118	30	17	56.7%
12	Private or social clubs or legion halls	813410	2673	121	84	69.4%
13	Snack bars	512130, 7111, 7112, 71131, 712, 7131, 713210, 713299 and 7139	8975	197	115	58.4%
14	Vending machines	454210	529	67	44	65.7%