# T1 Estimates - SAT - 2006 

Taxi and Limousine Service (TL) NAICS : 4853**

by<br>Javier Oyarzun<br>BSMD

Statistics Canada
2007-12-12

## 1. Introduction

### 1.1 Use of administrative data

Over the last few years, Statistics Canada (STC) has been accentuating the use of administrative data in its survey programs. Administrative data can be used for various purposes: creation and updating of sample frames, direct tabulation or analysis (when administrative data are used instead of a survey), estimation (e.g., administrative data can be used as auxiliary variables in a regression model), evaluation (survey data are compared with administrative data) and direct replacement of survey data.

There are several sources of administrative data. Good examples include import/export files, registers of vital statistics such as births and deaths, income tax and Employment Insurance files, and administrative records of educational, health care and justice institutions.

One of main benefits of using administrative data consists in reducing the burden for survey respondents. That is why they are used in the unincorporated business estimates program, more commonly referred to as the T1 program.

### 1.2 Definition of an unincorporated business

Every Canadian who earns income on which tax must be paid to the government is required to complete an income tax return (a T1 form) at the end of the year in which the income was earned. If professional, business, commission, rental, farming or fishing income is reported on the T1 form for a given year, the individual is considered an unincorporated business. If a business has employees and has to make salary deductions, or if a business has to collect goods and services tax (GST) because it makes more than $\$ 30000$ per year, it must have a Business Number (BN) issued by the Canada Revenue Agency (CRA). Otherwise, the business is uniquely identified only by the individual's social insurance number (SIN).

### 1.3 Information provided by the CRA

Each year, STC receives two major files from the CRA. One of them, the Assessed Record File (ARF), contains information about all individuals who reported an amount other than zero for at least one of the six types of income listed in section 1.2. The main variables in the file are gross income and net income for the six income sources. In 2006, there were just over 3.6 million unincorporated businesses in Canada. The other file, known as the E-File, is a subset of the ARF. It contains the information for all unincorporated businesses that submitted their data to the CRA electronically. About 50 variables are available in the file. In 2006, about $56 \%$ of respondents reported their data electronically.

### 1.4 The T1 estimates program

### 1.4.1 Estimates for the UES

The T1 estimates program has two components. One component is related to the Unified Enterprise Survey (UES). For this component, the aim of the T1 estimates program is to provide estimates of totals for about 30 variables in a number of different domains for about 40 surveys. The UES is based on several principles. The UES uses common generic processing systems and methods and an integrated questionnaire with harmonized variables and concepts, and it is supported by the Business Register (BR), a central database created and maintained at STC. To be included in the BR, a T1 must meet one of the following conditions: make salary deductions, or have a GST account. In 2006, just over 725000 T 1 s satisfied one or both of the conditions. Units that are included in one or more of the 40 surveys are identified, and a SIN is associated with each one. Units for which there is no SIN are excluded from the target population. In general, we manage to find a SIN for about $95 \%$ of the units. The estimates produced for this component are based on the BR and hence do not provide a complete picture of all T1s in Canada.

### 1.4.2 Satellite estimates

The second component is referred to as satellite estimates. They are produced for three surveys: Taxi and Limousine Services Industry, Couriers and Local Messengers Industry, and Real Estate Rental and Leasing and Property Management. A large proportion of the T1s for these industries have neither employees nor GST accounts and therefore are not included in the BR. As a result, the estimates produced by the UES do not take these units into account even though they make up a significant part of each industry's total revenue. To remedy this problem, we use the ARF (see section 1.3) to define the target population and not the BR. Then, using the E-File (electronic taxfilers), we create statistical models to estimate the proportion for which information is unavailable (paper taxfilers). The resulting estimates are known as "satellite estimates". Since 2005, satellite estimates have also been generated for all NAICS codes. Those estimates are used by Industry Accounts Division (IAD), Service Industries Division and Transportation Division.

### 1.4.3 Comparability of UES estimates and satellite estimates

For a given NAICS code, the estimates produced for the UES and the satellite estimates are never comparable. The main reason is that the target population is defined completely differently in the two estimation systems. In general, we expect that satellite estimates will be higher than UES estimates. Yet, because the estimation methods are different and because the ARF and the BR can have different NAICS codes for the same unit, it is possible that for certain variables estimated for small domains, the satellite estimates will be lower than the UES estimates.

### 1.4.4. BR's undercoverage for T1's

The business register (BR) is a structured list of businesses that produce goods and services in Canada. The list is used by statistical programs to determine the in-scope population, to select samples, support collection activities, monitor respondent burden and support/perform business demographic analysis. The BR undercovers the T1 population, because it contains only enterprises that have at least one employee or/and that have at least $\$ 30000$ in GST sales. In other words, the BR is missing the units that have no employees and a GST sales of under $\$ 3000$.

Figure 1.4.4. : T1 coverage portions (the dark grey is the portion not covered by the BR)


## 2. Processing the ARF and the E-File

### 2.1 Imputing NAICS codes

An important variable in the ARF is the NAICS code, i.e., the industry code for the T1 business as reported by the taxfiler. When STC receives the ARF, about $30 \%$ of the NAICS codes are missing or incorrect (the majority of them are coded 000000). Since estimates are required for all NAICS6 x PROVINCE combinations, a valid NAICS code is needed for each unit to avoid underestimation. If the code is missing or incorrect, it will be imputed. The imputation process consists of two steps. First, an attempt is made to find a valid NAICS code (at least two digits) with information from other sources, and then the third, fourth, fifth and sixth digits are imputed probabilistically if necessary. The process is described below.
(a) E-File

It is often the case that an individual has more than one unincorporated business in more than one industry. If so, the individual may file as many returns as he or she has businesses. Each one will appear as a separate entry in the E-File, and a valid NAICS code will generally be associated with each return. When that occurs, the NAICS code in the ARF is 000000. In that case, the ARF's NAICS code is imputed with the NAICS code of the tax return with the highest total gross income. At this point, the imputed NAICS code may have two, three, four, five or six digits. Just over $50 \%$ of the missing NAICS codes are imputed in this manner.

## (b) Previous years’ ARFs

When a valid NAICS code cannot be found in the E-File, the ARF for the previous year is used. If no valid NAICS code is found, the ARF from two years earlier is used. This method results in the imputation of nearly $2 \%$ of the missing NAICS codes (two, three, four, five or six digits).
(c) BR

Step 3 is to use the BR. With a SIN-BN concordance file generated by Tax Data Division (TDD), a BN can be associated with a number of units in the ARF. With the BN, the ARF can be matched with the BR to obtain a NAICS code for a number of units. About 7\% of the ARF's units are imputed with a NAICS code from the BR.
(d) Primary activity

Indirectly, the type of income reported in the ARF may provide a good indication of the business's industry. For example, more than $94 \%$ of T1s in the ARF whose main source of income is rent have a NAICS6 code of 531111. It is therefore reasonable to assume that $94 \%$ of the units in the ARF that have a missing NAICS6 and have reported rental income as their primary activity have a NAICS6 code of 531111. The imputation method used involves assigning a random number between 0.0000 and 100.0000 to the units in the ARF whose main source of income is rent and whose NAICS code is unknown. We also generate a cumulative frequency table (see below) showing the distribution of NAICS codes for units that reported rent as their main source of income.

Table 2.1: Distribution of units whose main source of income is rent

| NAICS | Frequency | Percentage | Cumulative <br> frequency | Cumulative <br> percentage |
| :---: | ---: | ---: | ---: | ---: |
| 110000 | 404 | 0.047 | 404 | 0.047 |
| $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ | 2.5188 |
| 530000 | 4 | 0.0005 | 21536 | 2.5196 |
| 531000 | 7 | 0.0008 | 21543 | 2.5211 |
| 531100 | 13 | 0.0015 | 21556 | 2.5216 |
| 531110 | 4 | 0.0005 | 21560 | 96.6575 |
| 531111 | 804880 | 94.1359 | 826440 | $\ldots$ |
| $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ | 100.000 |

For example, if a random number between 2.5188 and 2.5196 is chosen for a given unit whose NAICS code has to be imputed, the imputed NAICS code will be 531000 . If the random number is between 2.5216 and 96.6575 , the imputed NAICS code will be 531111. Thus it is easy to see that the distribution of the imputed NAICS codes will be the same as that of the non-imputed NAICS codes.

The process is repeated for the other five types of income (farming, fishing, commission, business, professional).

Imputation results
Table 2.1.2.: Results of steps $a, b, c$ and $d$ (cases where the NAICS code was missing or invalid in the ARF)

| Method | Frequency | Percent |
| :--- | ---: | ---: |
| E-File | 535636 | $52.52 \%$ |
| ARF 2005 | 21161 | $2.07 \%$ |
| ARF 2004 | 4620 | $0,45 \%$ |
| BR | 45260 | $4.44 \%$ |
| Primary act. | 413140 | $40.51 \%$ |
| TOTAL | 1019817 | $100 \%$ |

(e) NAICS3 to NAICS6

At this stage, all records have at least a valid two-digit NAICS code. We now have to make sure that the third, fourth, fifth and sixth digits are present and valid; if not, they will be imputed.

We start with the third digit. We proceed as in step d; that is, we assign a random number to all records with a missing third digit. We also generate a cumulative frequency table showing the distribution of the NAICS3 for each NAICS2. This ensures that the imputed NAICS3s will have the same distribution as the non-imputed NAICS3s.

The operation is repeated for the NAICS4s, NAICS5s and NAICS6s.
It is important to note that information about the imputation results is retained so that we can always find out whether a NAICS code was imputed, what source was used, and which digits were imputed.

Imputation results
Table 2.1.3.: Results of step e (cases where digits $3,4,5$ or 6 are missing following steps $a, b, c$ and d)

| Digits imputed | Cases where a NAICS2 <br> was initially invalid in <br> the ARF | Cases where a NAICS2 was <br> initially missing or invalid in <br> the ARF |
| :--- | :--- | :--- |
| Digits 3, 4, 5 and 6 were imputed | 143435 | 67839 |
| Digits 4, 5 and 6 were imputed | 258500 | 171651 |
| Digits 5 and 6 were imputed | 91029 | 38543 |
| Digit 6 was imputed | 115635 | 58256 |
| TOTAL | 608599 | 336289 |

Note: The records in the last four columns of Table 2.2a are also in the last column of Table 2.2b above.

In 2006, there were 3685191 records in the ARF. For 1050204 of those records (28.5\%), the entire NAICS code was imputed; for 143435 records ( $3.9 \%$ ), only the third to sixth digits were imputed; for 258500 records ( 7.0 \%), only the fourth to sixth digits were imputed; for 91029 records ( $2.5 \%$ ), only the fifth and sixth digits were imputed; and for 115635 records ( $3.1 \%$ ), only the sixth digit was imputed. In all, at least one digit of the NAICS code was imputed for 1658 803records (45\%).

### 2.2 Identifying partnerships

A partnership exists when two or more individuals are partners in the same unincorporated business. For tax purposes, each partner must report the partnership figures and not the figures for his or her share of the partnership. For example, if a couple owns a dwelling that brings in $\$ 5000$ a year in rental income, each partner must report $\$ 5000$ in gross income on his or her tax return. Each partner's share of the business is reflected only in his or her net income. To avoid overestimating the revenue and expenditures of unincorporated businesses, it is important to accurately identify such partnerships. However, not all of them can be identified with the information available in the CRA's files. A simple method was used to detect them. First, the NAICS code, region and total gross income variables were compared; if two or more individuals reported the same values for all three variables, they were deemed to be members of the same partnership. Similarly, if two or more persons reported the same values for the NAICS code, region and gross income from primary activity (a primary activity - farming, fishing, rental, business, commission or professional - is identified for each NAICS code), it is deemed to be a partnership.

To avoid detecting false partnerships, we consider only those records with a gross income of more than $\$ 10000$. Below that threshold, there is a high risk of finding partnerships that are not really partnerships.

The partnership identification process is currently under review. A better method that uses more information is expected to be introduced within a few months.

### 2.3 Detecting outliers

Outliers are values that are erroneous or inconsistent with other data. A good example of an outlier is \$999 999 999, which occurs in about 50 records in the "total gross income" field in the ARF. Outliers are excluded from the estimates and are not imputed. There are various methods of detecting them. In the case of the satellite estimates, a value is deemed to be an outlier if it is greater than \$5 000000 .

A more robust outlier detection system is in development. Studies will be carried out to identify the best methods for use in subsequent years.

## 3. Estimates

Estimates are produced for about 60 variables (for a list of estimated variables, see Appendix 1). Variables beginning with the letter L are available directly in E-File. Variables beginning with C or D are not available directly in E-File but are a combination of variables beginning with L .

The estimates for the variables beginning with $L$ were derived with statistical models. The models are constructed on the basis of electronic taxfilers (for whom all variables are available) and then applied to paper taxfilers. The models were constructed at the NAICS6 level. That means that the parameters of the various models were derived for every possible NAICS6 grouping (NAICS6 codes beginning with 31, 32, 33 and 91 were modelled at the NAICS2 level because there were too few records for a number of NAICS6 codes).

No model was built for NAICS6 codes with fewer than three electronic taxfilers. For those NAICS6s, it was possible to derive only variables L8299 and L9946 directly with the ARF.

Two different models were used in the estimation process: simple linear regression estimation, and ratio estimation. The details are provided below.
(a) Variables L8299 and L9946

The following model was selected for this variable:

$$
y_{i}=\beta_{o}+\beta_{1} x_{i}+\varepsilon_{i}
$$

where $\beta_{0}, \beta_{1}$ and $\varepsilon_{i}$ are the usual regression model parameters and $x_{i}$ is the "total gross income" variable for L8299 (for L9946, $x_{i}$ is the "total net income" variable) which is available for all units in the ARF. We want to estimate

$$
Y=\sum_{i \in U} y_{i}
$$

using the predictive approach.
Let $\hat{Y}$ be an estimator of $Y$. Thus we have

$$
\hat{Y}=\sum_{i \in s} y_{i}+\sum_{i \in U-s} \hat{y}_{i}
$$

with $\hat{y}_{i}=\hat{\beta}_{o}+\hat{\beta}_{1} x_{i}$. An estimator of the variance of $\hat{Y}$ under the linear regression model is derived as follows:

$$
V(\hat{Y})=\frac{(N-n)}{f} \hat{\sigma}^{2}\left[1+\frac{\left(\bar{x}_{s}-\bar{x}_{U}\right)^{2}}{(1-f) \sum_{i \infty}\left(x_{i}-\bar{x}_{s}\right)^{2} / n}\right]
$$

where

$$
\hat{\sigma}^{2}=\sum_{i \in s}\left(y_{i}-\hat{y}_{i}\right)^{2} /(n-2)
$$

and

$$
f=n / N
$$

(b) Other variables beginning with L

A ratio model was preferred for these variables. In mathematical form, the model can be written as follows:

$$
y_{i}=\beta_{1} x_{i}+\varepsilon_{i}
$$

where $\beta_{1}$ and $\varepsilon_{i}$ are the usual ratio model parameters and $x_{i}$ We want to estimate

$$
Y=\sum_{i \in U} y_{i}
$$

using the predictive approach.
Let $\hat{Y}_{\text {be an estimator of Y. Thus we have }}$

$$
\hat{Y}=\sum_{i \in s} y_{i}+\sum_{i \in U-s} \hat{y}_{i}
$$

with

$$
\hat{y}_{i}=\hat{\beta}_{1} x_{i}
$$

An estimator of the variance of $\hat{Y}$ under the ratio model is derived as follows:

$$
V(\hat{Y})=N^{2} \hat{\sigma}^{2}\left(\frac{\bar{x}_{u}}{\bar{x}_{s}}\right)^{2} \frac{1-f}{n} \sum_{i \in s}\left(y_{i}-\hat{y}_{i}\right)^{2} /(n-1)
$$

where

$$
\hat{\sigma}^{2}=\sum_{i \in s}\left(y_{i}-\hat{y}_{i}\right)^{2} /(n-2)
$$

Another way of obtaining $\hat{Y}$ is the following equation:

$$
\hat{Y}=\sum_{i \in U} x_{i} \frac{\sum_{i \in s} y_{i}}{\sum_{i \in s} x_{i}}
$$

(c) Variables beginning with C or D

Variables beginning with C or D are simply derived using variables beginning with L (see Appendix 2).

## 4. Résults

In this section we'll present the results obtained using the satellite estimates for the Taxi and Limousine Service survey (TL: 4853**) :
4.1. Number of T1 units

| Year | Number of units |
| :---: | ---: |
| 2006 | 38943 |
| 2005 | 37090 |
| 2004 | 36778 |
| 2003 | 36206 |
| 2002 | 34128 |

4.2. NAICS imputation

| Year | Number of units (without imputation) | Number of units (with imputation) | Imputation E-FILE | Imputation Historical ARF | Imputation BR | Imputation Activity |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 2006 | $\begin{array}{r} 27791 \\ (71,36 \%) \\ \hline \end{array}$ | 38943 | $\begin{array}{r} 4436 \\ (39,78 \%) \\ \hline \end{array}$ | $\begin{array}{r} 161 \\ (1,44 \%) \\ \hline \end{array}$ | $\begin{array}{r} 739 \\ (6,63 \%) \\ \hline \end{array}$ | $\begin{array}{r} 5816 \\ (52,15 \%) \\ \hline \end{array}$ |
| 2005 | $\begin{array}{r} 25945 \\ (69,95 \%) \\ \hline \end{array}$ | 37090 | $\begin{array}{r} 4112 \\ (36,90 \%) \\ \hline \end{array}$ | $\begin{array}{r} 109 \\ (0,98 \%) \end{array}$ | $\begin{array}{r} 1276 \\ (11,45 \%) \\ \hline \end{array}$ | $\begin{array}{r} 5648 \\ (50,68 \%) \\ \hline \end{array}$ |
| 2004 | $\begin{array}{r} 24944 \\ (67,82 \%) \\ \hline \end{array}$ | 36778 | $\begin{array}{r} 3765 \\ (31,82 \%) \\ \hline \end{array}$ | $\begin{array}{r} 1515 \\ (12,80 \%) \end{array}$ | $\begin{array}{r} 852 \\ (7,20 \%) \end{array}$ | $\begin{array}{r} 5702 \\ (48.18 \%) \\ \hline \end{array}$ |
| 2003 | $\begin{array}{r} 26539 \\ (73,30 \%) \\ \hline \end{array}$ | 36206 | $\begin{array}{r} 2428 \\ (25,12 \%) \\ \hline \end{array}$ | $\begin{array}{r} 2207 \\ (22,83 \%) \\ \hline \end{array}$ | $\begin{array}{r} 445 \\ (4,60 \%) \\ \hline \end{array}$ | $\begin{array}{r} 4587 \\ (47,45 \%) \\ \hline \end{array}$ |
| 2002 | $\begin{array}{r} 24802 \\ (72,67 \%) \\ \hline \end{array}$ | 34128 | $\begin{array}{r} 2022 \\ (21,68 \%) \\ \hline \end{array}$ | $\begin{array}{r} 2567 \\ (27,52 \%) \\ \hline \end{array}$ | $\begin{array}{r} 241 \\ (2,58 \%) \\ \hline \end{array}$ | $\begin{array}{r} 4496 \\ (48,21 \%) \\ \hline \end{array}$ |

4.3. Proportion of E-FILER

| Year | E-FILER | P-FILER | Number of units |
| :---: | ---: | ---: | ---: |
|  | 17819 <br> $(45,76 \%)$ | 21124 <br> $(54,24 \%)$ | 38943 |
| 2006 | 16007 | 21083 | 37090 |
| 2005 | $(43,16 \%)$ | 14347 | $(56,84 \%)$ |

4.4. Proportion of partnerships

| Year | Within a partnership | Not in a partnership | Number of units |
| :---: | :---: | :---: | :---: |
| 2006 | $\begin{array}{r} 10304 \\ (26,46 \%) \end{array}$ | $\begin{array}{r} 28639 \\ (73,54 \%) \end{array}$ | 38943 |
| 2005 | $\begin{array}{r} 9771 \\ (26,34 \%) \\ \hline \end{array}$ | $\begin{array}{r} 27319 \\ (73,66 \%) \\ \hline \end{array}$ | 37090 |
| 2004 | $\begin{array}{r} 9643 \\ (26,22 \%) \\ \hline \end{array}$ | $\begin{array}{r} 27135 \\ (73,78 \%) \\ \hline \end{array}$ | 36778 |
| 2003 | $\begin{array}{r} 9486 \\ (26,20 \%) \end{array}$ | $\begin{array}{r} 26720 \\ (73,80 \%) \end{array}$ | 36206 |
| 2002 | $\begin{array}{r} 8932 \\ (26,17 \%) \\ \hline \end{array}$ | $\begin{array}{r} 25196 \\ (73,83 \%) \end{array}$ | 34128 |

4.5. Outliers

| Year | Number of units | Outliers | $\begin{array}{r}\text { Number of units } \\ \text { (ARF) }\end{array}$ |
| :---: | ---: | ---: | ---: |
| 2006 | 38943 | $\begin{array}{r}157 \\ (0,40 \%)\end{array}$ | 39100 |
| 2005 | 37090 | $\begin{array}{r}87 \\ (0,23 \%)\end{array}$ | 37177 |
| 2004 | 36778 | $\begin{array}{r}101 \\ (0,27 \%)\end{array}$ | 36879 |
| 2003 | 36206 | 448 |  |
| $(1,22 \%)$ |  |  |  | \(\left.\begin{array}{r}364 <br>

(1,06 \%)\end{array}\right]\)
4.6. Results for revenues and expenses

| Year | Number of units | C2098 <br> Revenues | C4699 <br> Expenses |
| :---: | ---: | ---: | ---: |
| 2006 | 38943 | 1133495182 | 913439773 |
| 2005 | 37090 | 1096044003 | 909012259 |
| 2004 | 36778 | 1073743058 | 873263411 |
| 2003 | 36206 | 1004423362 | 818978343 |
| 2002 | 34128 | 939595636 | 748597435 |

4.7. BR's undercoverage
4.7.1. UES and SAT estimates

| Year | Number of units <br> (ARF) | Number of <br> units <br> (BR-SCF) | Différence of <br> units | Proportion of <br> undercovered <br> units |
| :---: | ---: | ---: | ---: | ---: |
| 2006 | 39100 | 12624 | 26476 | $67,71 \%$ |
| 2005 | 37177 | 12422 | 24755 | $66,59 \%$ |
| 2004 | 36879 | 11480 | 25399 | $68,87 \%$ |
| 2003 | 36654 | 10264 | 26390 | $72,00 \%$ |
| 2002 | 34492 | NA | NA | NA |


| Year | C2098 <br> Revenues <br> BR-UES | C2098 <br> Revenues <br> SAT | Différence of <br> revenues | Proportion of <br> undercovered <br> revenues |
| ---: | ---: | ---: | ---: | ---: |
| 2006 | 547624417 | 1133495182 | 585870765 | $51,69 \%$ |
| 2005 | 475837113 | 1096044003 | 620206890 | $56,59 \%$ |
| 2004 | 386410306 | 1073743058 | 687332752 | $64,01 \%$ |
| 2003 | 552301419 | 1004423362 | 452121943 | $45,01 \%$ |
| 2002 | NA | 939595636 | NA | NA |

4.7.2. Units under 30 000\$ (BR’s undercoverage)

| Year | Units reporting a revenue <br> below $\mathbf{3 0} \mathbf{0 0 0} \mathbf{\$}$ | Units reporting a revenue <br> above $\mathbf{3 0} \mathbf{0 0 0} \mathbf{\$}$ | Number of units |
| :---: | ---: | ---: | ---: |
| 2006 | 21766 | 17177 | 38943 |
| 2005 | $(55,89 \%)$ |  |  | | $(44,11 \%)$ |
| ---: |

## Appendix 1:

List of estimated variables.

| Nom de la variable | Définition |
| :---: | :---: |
| L4007 | Closing Inventories, farm |
| L4008 | Opening Inventories, farm |
| L8000 | Net Sales |
| L8141 | Real Estate Rental Income (transferred to L8000) |
| L8230 | Other Income |
| L8290 | Reserves deducted last year |
| L8299 | Total revenue |
| L8300 | Opening inventory |
| L8320 | Purchases |
| L8340 | Direct wages |
| L8360 | Sub-contracts |
| L8450 | Other costs |
| L8500 | Closing inventory |
| L8518 | Cost of goods sold |
| L8519 | Gross Profit |
| L8521 | Advertising |
| L8523 | Meals and entertainment |
| L8590 | Bad Debts expense |
| L8690 | Insurance |
| L8710 | Interest |
| L8760 | Bus. Tax, Fees, Licenses etc. |
| L8810 | Office expenses |
| L8811 | Supplies |
| L8860 | Legal, accounting and other professional fees |
| L8871 | Management and administration fees |
| L8910 | Rent |
| L8960 | Maintenance and repairs |
| L8963 | Boats Repairs |
| L9060 | Salaries, wages and benefits |
| L9062 | Crew shares |
| L9136 | Gear |
| L9137 | Nets and Traps |
| L9138 | Bait, ice salt |
| L9180 | Property taxes |
| L9200 | Travel |
| L9220 | Telephone and utilities |
| L9224 | Fuel costs |
| L9270 | Other expenses |
| L9275 | Delivery, freight and express |
| L9281 | Motor vehicle expenses |
| L9368 | Total business expenses |
| L9369 | Net income (loss) before adjustments |
| L9923 | Land addition |
| L9924 | Land disposition |
| L9925 | Details of equipment addition |
| L9926 | Details of equipment disposition |
| L9927 | Details of building addition |
| L9928 | Details of building disposition |
| L9931 | Total business liability |
| L9932 | Drawing |


| L9933 | Capital contribution |
| :--- | :--- |
| L9935 | Allowance on eligible capital property |
| L9936 | Capital cost allowance |
| L9943 | Other amounts deducted from share |
| L9945 | Business - use of home expenses |
| L9946 | Net income/loss |
| L9947 | Recaptured CCA |
| L9948 | Terminal loss |
| L9949 | Total personal portion of expenses |
| L9950 | Filer's share amount |
| C3040S | L8340 + L9060 = total salaries, wages and benefits |
| C4520 | L9935 + L9936 = depreciation and amortization |
| C4599 | L8299 - L9946 - L8710 = total operating expenses |
| C4699 | L8299 - L9946 = total expenses |
| D9857 | L9946 + L8710 Operating Profit |
| D9858 | L8340 + L9060 + L9935 + L9936 + L9946 + L8710 = value added |
| D9859 | L8299 - (L8340 + L9060 + L9935 + L9936 + L9946 + 8710) =intermediate <br> inputs |
| D9876 | L8710 = Non-operating expenses ( interest) |
| DV_FU | L9220+L9224=total fuel and utilities expenses |

## Appendix 2:

Mappings given by Transportation Division.

| Variable | Mapping |
| :---: | :---: |
| C2077 | L8230+L8141 + L9926 + L9928 |
| C2080 | L8000+L8290+L8230+L8141 + L9926 + L9928 |
| C2098 | L8000+L8290+L8230+L8141 + L9926 + L9928 |
| C2299 | L8000+L8290 |
| C3041 | L8340+L9060+L9062 |
| C3088 | L8360 |
| C3399 | $\begin{aligned} & \text { L8810 + L8811 + L9220 + L9136 +L9137 + L9138 + (L8300-L8500) + (L9925*0.5) + } \\ & \text { L9945 +(L8320*.40) } \end{aligned}$ |
| C4066 | L9224 |
| C4069 | (L9281*.65) |
| C4101S | $\begin{aligned} & \text { L8810 + L8811 + L9220 + L9136 +L9137 + L9138 + (L8300-L8500)+ (L9925*0.5) + } \\ & \text { L9945 +(L8320*.40) + L9224 } \end{aligned}$ |
| C4115 | L8910 + (L9281*.20) |
| C4140S | L8910 + (L9281*.20) |
| C4178 | L8960 + L8963 + (L9281*0.15)+ L9927 |
| C4220S | L8960 + L8963 + (L9281*0.15)+ L9927 |
| C4070 | L9275 |
| C4370S | L9200 + L9275 + L8360 + L8690 + L8521 + L8860 + L8523 + L8871 + (L8320*.20) |
| C4410 | L8760 + L9180 |
| C4520 | L9935 + L9936 + (L9925*0.5) |
| C4569 | L9270 + L8590 + L9943 + L9945 + L8450 + L9923 + L9924 +(L8320*.40) |
| D9875 | L9270 + L8590 + L9943 + L9945 + L8450 + L9923 + L9924 + (L8320*.40) |
| C4599 |  |
| C4630 | L8710 |
| D9876 | L8710 |
| C4699 |  |
| D9802 | $\begin{aligned} & \text { (L8000+L8290+L8230+L8141 + L9926 + L9928) - ((L8340+L9060+L9062)+(L9281*.65) } \\ & \text { + (L8810 + L8811 + L9220 + L9136 +L9137 + L9138 + (L8300-L8500)+ (L9925*0.5) + } \\ & \text { L9945 + (L8320*.40) + L9224)+ (L8910 + (L9281*.20)) + (L8960 + L8963 + } \\ & \text { (L9281*0.15) + L9927) + (L9200 + L9275 + L8360 + L8690 + L8521 + L8860 + L8523 + } \\ & \text { L8871 + (L8320*.20)) + (L8760 + L9180) + (L9935 + L9936 + (L9925*0.5)) + (L9270 + } \\ & \text { L8590 + L9943 + L9945 + L8450 + L9923 + L9924 +(L8320*.40) )) } \end{aligned}$ |

