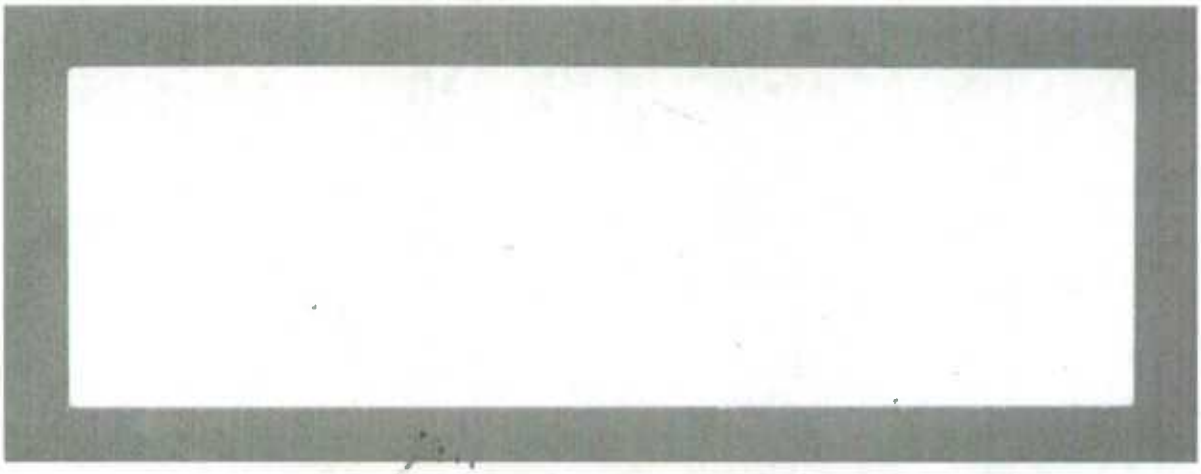
Statistics Canada Statistique Canada



Methodology Branch

Business Survey Methods Division

Direction de la méthodologie

Division des méthodes d'enquêtes-auprès des entreprises

Canada

WORKING PAPER NO. BSMD-2008-001E
METHODOLOGY BRANCH

CAHIER DE TRAVAIL NO. DMEE-2008-00E
DIRECTION DE LA METHODOLOGIE

Should Sub-Annual Surveys be Benchmarked to their Annual Counterparts? A Case Study of Manufacturing Surveys

by

Wesley Yung, François Brisebois, Charles Tardif,
Gordon Kuromi and Caroline Rondeau

BSMD-2008-001E

Should Sub-Annual Surveys be Benchmarked to their Annual Counterparts? A Case Study of Manufacturing Surveys

Wesley Yung, François Brisebois, Charles Tardif, Gordon Kuromi and Caroline Rondeau

Abstract

For many years, the results of the Monthly Survey of Manufacturing (MSM) and the Annual Survey of Manufactures (ASM) have been reconciled and then made equivalent through an annual benchmarking process. Amongst other things, the MSM used the benchmarking exercise to account for the contribution of the take-none portion of the sampling frame. After the 2004 benchmarking exercise produced some significant changes to the MSM annual growth rates, there were some questions as to whether it was still appropriate to force the results of the two surveys together. In reaction to these questions, a working group was created to answer the question 'Taking into account the many differences between the two surveys and the needs of the data users, should the MSM continue to be benchmarked to the ASM?'. This working paper presents the results of the analyses done by the working group and recommendations for the MSM and the ASM.

Executive Summary

This report presents an overview of the main findings and recommendations of the Benchmarking working group. The working group was created in the fall of 2006 to investigate the appropriateness of benchmarking the Monthly Survey of Manufacturing (MSM) to the Annual Survey of Manufactures (ASM). A coherence analysis of the two surveys was performed in an attempt to quantify the differences that the benchmarking process was adjusting for.

The main findings of the working group were:

- In general, the two surveys are very similar in terms of concepts, operations and methodology but there are several areas where they are very different. These areas are estimation of the take-none portion, the reference period covered and the sources of administrative data.
- The contribution of the take-none units in ASM is estimated using tax data, while the MSM depended on the benchmarking adjustment to account for the undercoverage coming from the take-none units. Note that this methodology is no longer used with the introduction of the restratified MSM design in January 2007.
- The reference period covered by the ASM is defined as the fiscal year of the respondent that ends in the period between April 1st of the reference year and March 31st of the reference year plus 1. Regardless of the reported reference period, ASM estimates are taken to represent January 1st to December 31st of the reference year. Annualized MSM estimates are benchmarked to these estimates from ASM. In 2004 and 2005, approximately 35% of all ASM units reported for the January to December period of the reference year. These units accounted for approximately 65% of the total shipments.
- For units in both surveys that reported for a common reference period, there was a 1% difference in data from the two surveys.
- Both surveys use tax data to supplement survey data. ASM uses annual tax data, while MSM uses monthly Goods and Services Tax (GST) data. Although there appear to be some differences between the two sources of tax data, they do not contribute significantly to the ASM or the MSM estimates.

Recommendations of the working group are:

- In order to reduce (or perhaps eliminate) the need for benchmarking, the working group recommends that coherence between sub-annual and annual surveys be built in at the design stage. While it is recognized that 100% coherence is unlikely, all efforts should be made so that the two surveys are equivalent in terms of concepts, methodology and operations. By designing surveys as such, benchmarking will be asked to account only for differences from sampling variability and inconsistencies that can be explained (and possibly quantified).
- Turning the ASM and the MSM in particular, given the issues with the possibly different reference periods and the difference observed in the reported data from the common units, it is recommended that the two surveys no longer be benchmarked.

- If the two series are no longer to be benchmarked, the importance of an annual reconciliation process increases. An annual reconciliation process should attempt to minimize, as much as possible, the differences observed in the two series.
- To implement the decision to remove the benchmarking, the working group recommends to use the latest ASM values for 2004 and 2005, but to converge the benchmark factors to 1 through the 2006 year. In addition, the MSM has recently gone through a restratification with a parallel run during the last four months of 2006. The results of this restratification should be integrated with the removal of the benchmark factors. More details are given in section 4.1 of the report.

1.0 Introduction

Statistics Canada conducts many annual and sub-annual business surveys, some of which cover the same target population. The main reason for having both annual and sub-annual surveys covering the same target population is that they are designed to produce different information. Annual surveys are designed to produce very detailed estimates of levels, such as estimates of annual revenue, expenses and manufacturing outputs. On the other hand, sub-annual surveys are typically designed to measure economic changes over the short term. In order to keep response burden to a minimum, they typically collect only limited information and can not produce detailed estimates like the annual surveys.

Despite the fact that the surveys cover the same target population, and probably estimate some of the same variables, rarely are the annual and annualized sub-annual estimates equal. Given that the surveys are designed for different purposes, perhaps one should not expect the surveys to produce identical estimates. However, many statistical agencies prefer that the published survey results are equivalent so that their data users, both internally and externally, get the same picture of the economy regardless of which survey results they are looking at. To ensure consistency between the annual and sub-annual surveys benchmarking methods are commonly used. For more on benchmarking techniques, we refer the reader to Dagum and Cholette (2006).

Statistics Canada has been using benchmarking to ensure the consistency of its Monthly Survey of Manufacturing (MSM) and its Annual Survey of Manufactures (ASM) for many years. As a result of the benchmarking exercise done in 2006 when the ASM values for reference years 2003 and 2004 became available, annual MSM growth rates for 2003 to 2004 and 2004 to 2005 were revised from 8.5% to 4.2% and from 2.9% to 0.6% respectively. These large revisions brought into question the practice of benchmarking. In particular should it be done and, if so, what is the best way to do it?

A working group, co-chaired by Wesley Yung of the Business Survey Methods Division (BSMD) and Kevin Roberts of the Manufacturing, Construction and Energy Division (MCED), was created to answer the first question. The members of the working group were François Brisebois, Jean-François Dubois, Susie Fortier, Charles Tardif, Gordon Kuromi, Caroline Rondeau, all of BSMD, and Michael Scrim and Rhonda Tsang of MCED. Guidance was provided by a steering committee consisting of Peter Lys (co-chair), Don Royce (co-chair), Richard Evans, Michael Girard, Andy Kohut, John Kovar, Jean-Pierre Simard and Wesley Yung. The working group decided to look at coherence issues between the surveys in an attempt to explain the difference in the estimates. The working group felt that if benchmarking was to be used, then the two surveys should be as coherent as possible and any incoherencies should be quantified before the benchmarking exercise takes place. The incoherencies were classified into one of three categories: conceptual, methodological and operational. All steps of the survey process were placed in one of these three categories and any differences between the two surveys were evaluated and, where possible, quantified.

The second question was investigated by the Time Series Research and Analysis Center of BSMD. Their findings are reported in Quenneville and Fortier (2006) where they discuss several different benchmarking techniques. The choice of technique depends on the particular circumstances.

This report documents the findings of the working group and is organized as follows. Coherence between the MSM and ASM is presented in section 2 and several analyses quantifying differences between the two surveys are presented in section 3. A short discussion and some recommendations are given in section 4.

2.0 Coherence between MSM and ASM

In theory, if the MSM and ASM were coherent in terms of concepts, methodology and operations, then estimates between the two surveys would be similar. Any difference between the two surveys would be only from sampling error. If this were the case, then benchmarking would simply be a cosmetic exercise to force the annualized MSM estimates to be equal to the ASM estimate or vice versa. In fact, the two survey estimates could be forced to some intermediary value if desired. Unfortunately, despite numerous attempts the MSM and ASM are not completely coherent. In this section, we look at the coherence between the two surveys in terms of conceptual, methodological and operational aspects. Under conceptual aspects we consider the target populations, the reference periods and the questionnaires of the two surveys. Under methodology, we look at the sampling frames, the stratification and allocation methods, the targeted CVs, the editing and imputation processes, the estimation methods (including that of the take-none portion) and the use of tax data. Under the operational aspects, we look at the collection process (including the follow-up strategy) and the reconciliation process. Many of the details mentioned above could have been placed in multiple aspects, but for simplicity we assigned each detail to one and only one aspect.

2.1 Coherence of Conceptual Aspects

2.1.1 Target Population

Looking at the definitions of the two target populations, both surveys cover all establishments engaged primarily in manufacturing industries defined by North American Industry Classification System (NAICS) sectors 31, 32 and 33. Note that the ASM also covers logging establishments identified as being in NAICS 1133. However, these establishments are not included in the benchmarking exercise.

The working group felt that the definitions of the target populations for the two surveys are conceptually the same and would not contribute significantly to any difference observed in the survey estimates.

2.1.2 Reference Periods

According to the ASM reporting guide, the reporting period is the business unit's 12-month fiscal period ending between April 1st of the reference year and March 31st of the following year. Regardless of the period covered by the reported data, estimates are considered as representing the period between January 1st and December 31st of the reference year. The MSM, being a monthly survey, covers the reference month. Both surveys ask the respondent to report the accounting period covered, so it is possible that some MSM data will not cover the appropriate reference month as well. When benchmarking is performed, the MSM data for the months of January to December are benchmarked to the ASM data despite the fact that the reference period of the ASM may not exactly match the calendar year. The magnitude of this difference will be reported later on in this document.

It is felt that the non-calendarization of the ASM may contribute significantly to the differences between the two surveys. If benchmarking were to continue, the development of a calendarization methodology for annual surveys should be considered. If benchmarking is no longer to be performed, then the different reference periods should be accounted for in the reconciliation process. One possible way of doing this is to match the MSM reference months to the reported ASM reference period.

2.1.3 Questionnaires

Comparing the two questionnaires (available online at www.statcan.ca) one quickly sees that the MSM questionnaire is significantly shorter than that of the ASM. The MSM questionnaire consists of one page and four questions covering Shipments, Inventories, Orders and the period covered. The ASM questionnaire consists of approximately 20 pages, is Chart of Accounts[1] (COA) compliant and covers topics such as the Production and Cost Report, Destination of Sales, Sales of Goods Manufactured, Purchases of Raw Materials and Salaries and Wages. The questions asked by the MSM roughly correspond to the section on the Production and Cost Report, but the wording of the questions still shows some slight differences and up until 2007 was not COA compliant. In the ASM questionnaire, starting in 2004, shipments are now called sales and many more details are requested. For instance, in the ASM sales are requested for the following categories:

- Sales of goods
  - Manufactured
  - Logs and wood residue (for logging operations only)
  - Goods purchased for resale
  - Progress billing
- Revenue from repair work
- Revenue from manufacturing service fee and/or custom work
- Revenue from stumpage sales (for logging operations only)
- Revenue from other sales

[1] The Chart of Accounts is a standard for reporting on the financial position and performance of a business. For more information, please consult the Statistics Canada website at http://stds.statcan.ca/english/coa/coa_main.asp.

- Other revenue

In the MSM questionnaire, values for shipments are requested for

- Goods of own manufactures
- Goods purchased for resale

The definition of goods of own manufacture given in the MSM reporting guide matches closely to the sum of the first five categories of sales as requested by the ASM. Differences include the logging operations categories (logs and wood residue and stumpage sales), progress billing and revenue from other sales. Despite the fact that logging operations are excluded from the benchmarking process, care should be taken to ensure that sales from logging operations are removed for units that are included in the benchmarking and reconciliation processes.

As long as the contribution from logging operations is excluded from the benchmarking and reconciliation processes, the working group felt that differences in the questionnaires had minimal impact on the coherence of the resulting estimates. Following a restratification exercise in January 2007, the MSM questionnaire has been COA compliant. The working group sees this as a positive change towards improving the coherence of the two surveys.

2.2 Coherence of Methodological Aspects

2.2.1 Sampling Frames

Both the MSM and the ASM use Statistics Canada's Business Register (BR) as their sampling frame, however there are some differences that could possibly cause some incoherence. The ASM uses a frame that is created in October of the reference year. Given that the frame is created in October, any changes that occur in November or December are not reflected in the ASM frame. The ASM frame contains all active units in the population at that time but also contain any units that were active during the period between January and October of the reference year. This concept of being "ever-alive" ensures that a unit that was active during the reference period but may not be active when the frame is created is on the sampling frame. From a methodological point of view, these 'ever-alive' units should be on the frame but they can cause problems further in the survey process. If an 'ever-alive' unit is contacted, will it report information corresponding to the period that it was in-scope? If no response is obtained from the unit, will the edit and imputation system take into account the period that it was in-scope or not? Questions such as these need to be investigated further.

On the other hand, the MSM uses a monthly frame to update an existing sample with a sample of births. The frame is created at the beginning of each month and contains only those units that are active at the time of creation. Thus, if a unit was active for a portion of the month but becomes inactive just before creation of the frame, it is not included in it.

The working group felt that differences in the creation of the two sampling frames may exist but the effect on the resulting estimates were most likely minimal due to the small number of 'ever-alive' units and units that became inactive just before creation of the MSM frame. Nevertheless, these units should be investigated to get a better idea of the magnitude of the possible problems.

2.2.2 Stratification

As with most business surveys at Statistics Canada, both the ASM and the MSM use a stratified sampling approach with a take-none (TN) stratum, one or more take-some (TS) strata and a take-all (TA) stratum within an industry/province cell. Boundaries between the TS and TA strata are identified by the Lavallée-Hidiroglou algorithm (Lavallée and Hidiroglou 1988) applied to a measure of size. For the ASM, this measure of size is the maximum of the shipment variable from a previous survey occasion and the Gross Business Income (GBI) from the BR. If the shipment variable from a previous ASM survey is not available then it is estimated from a model based on GBI, revenue from Goods and Services Tax (GST) data and the number of employees. For the MSM, shipments are also used as a measure of size but its value comes from one of several sources. If available, an annualized value of shipments coming from previous occasions of the MSM is used. If that is not available, then the maximum value of ASM shipments, annual tax data revenue or GST revenue is used. Finally, if none of these sources is available, the GBI from the BR is used.

Both the ASM and the MSM have an industrial and geographical component to the stratification. ASM first stratifies by 3-digit NAICS by province (6-digit NAICS by province for the logging industry), while MSM stratifies by 4, 5 or 6-digit NAICS by province.

The ASM and the MSM both use take-none strata but have applied the concept in slightly different ways. The MSM defined the TN boundary at the province level such that units that represent the bottom 2% of total manufacturing shipments are placed in the TN stratum. ASM on the other hand excludes units based on thresholds within the industry/geography cells that are defined such that less than 10% of GBI is excluded. If a unit below the TN threshold has a GBI greater than $1M, then it is removed from the TN stratum and is placed in the appropriate size stratum within that cell. Note that since the MSM restratification that took place in January 2007, the MSM TN boundaries are now calculated in a similar fashion to those of the ASM. That is, the threshold is defined such that a maximum of 10% of shipments is excluded at the industry/geography cell level. Given that the TN boundaries are different between the MSM and the ASM, the TS and TA boundaries will also be different. Estimation of the contribution of the TN portion is discussed later in this document.

The working group was encouraged to see that the MSM will now calculate the TN boundaries using a methodology similar to the ASM. Given the improved coherence in determining the TN boundaries, the working group feels that any remaining incoherencies will contribute minimally to the differences between the two survey estimates.

2.2.3 Editing and Imputation

The editing and imputation methods used by the two surveys are very similar but their systems are very different. ASM uses Statistics Canada's generalized edit and imputation system BANFF, while MSM is using a custom system written specifically for that survey. For both surveys, consistency and deterministic edits are applied to ensure that totals are equal to the sum of their parts and that any missing values that can be derived are derived properly. In addition, outlier detection methods are applied to identify units that should not be used during the imputation process.

In terms of imputation, both surveys create imputation classes based on a combination of industry, province and a measure of size. The exact creation of these classes differs between the surveys, but the methodology is similar. Imputation is performed within each imputation class and uses methods such as historical or auxiliary data with or without a trend, ratio imputation based on a measure of size and donor imputation. Due to agreements with external clients, the ASM produces a census of financial information and, as such, significantly more imputation takes place in the ASM than the MSM. However, the majority of this imputation is based on current year tax data.

The working group felt that while the differences in the edit and imputation systems did not contribute significantly to the differences observed in the estimates coming from the two surveys, there is some room for improvement. For instance, the feasibility of redeveloping the MSM edit and imputation system in BANFF and the quality of the imputation based on tax data by ASM should be evaluated.

2.2.4 The Use of Tax Data

The use of tax data in Statistics Canada's business surveys has increased over the past few years. Tax data are now used in many steps of the survey process, but their largest impact on the estimates comes from their use in imputation and tax replacement. The ASM and MSM use different sources of tax data and thus have different ways of incorporating it into the survey. The MSM uses monthly GST data for a sample of simple units and for several hundred chronic non-respondents. Simple units are those businesses that are involved in a single activity within a single province. This definition of a simple unit allows tax revenue to be correctly allocated to the proper NAICS code within a province. If a unit is not simple, it is said to be complex. The ASM uses annual data for all non-sampled simple units above the take-none threshold and for non-responding units, thus creating a pseudo-census for the financial variables.

The MSM uses monthly GST data but, since data for the reference month m are not always available in time, data from month m-2 are used. A statistical model is used to account for differences between GST sales and the variable of interest (shipments) and for the time lag between the reference month and the reference period of the GST data. The parameters of the statistical model are estimated using a sample of simple units for which both GST and survey data are available. Once estimated, the model is applied to the GST value to obtain an estimated survey value for units identified for tax replacement.

The ASM uses annual tax data, but does not apply a statistical model to link the tax and survey data. The coherence between the two sources of data is ensured by the Chart of Accounts (COA), which was developed to ensure consistency of concepts and definitions between survey and tax questionnaires. Thus if a survey is COA compliant, tax data and survey data should be readily interchangeable and no statistical model is necessary to account for differences in concepts. In addition, annual tax data for reference year, t, are available in time for processing the annual survey for the same reference year.

Given that two different sources of tax data are used, one needs to investigate differences between them. These differences may include the definition of the variables, the reference period covered, the overall quality of the data and how missing values are handled. For example, annual tax data are typically from the business' financial statements and are most likely of very high quality. However, GST data are based on remittances which may occur monthly, quarterly or annually and may not correspond exactly to the sales of a particular business. On the other hand annual tax data may not cover the same reference period as the survey and, thus, should perhaps be calendarized. A more in depth analysis of the two tax data sources is presented later in this report.

Differences in the approach used to incorporate tax data were not investigated by the working group but should be analyzed in the future.

The working group felt that although the use of two sources of tax data had little effect on the differences observed between the 2004 ASM and MSM estimates, work should be done to ensure that tax data are being processed and used in a consistent manner by both the monthly and annual surveys. As Statistics Canada's surveys continue to use more tax data, the tax sources, the processing of the tax data and the approaches of using it must be evaluated.

2.2.5 Estimation

In typical probability based surveys, sampling weights are obtained through the sample design and weighted estimates are produced as the sum of the weighted responses. While the methods of estimation for the annual and sub-annual surveys are similar, many adjustments can be made to the weights to account for nonresponse or estimation for the take-none portion. In the case of the MSM, it uses a count adjustment to the weights to account for unit nonresponse. That is, based on counts the respondents weights are adjusted to account for the nonrespondents. This assumes that the behavior of the nonrespondents is similar to that of the respondents.

On the other hand, the ASM produces a pseudo-census for financial variables using tax data. For all non-sampled units above the take-none threshold, tax data are imputed and are then treated like survey data. Thus, the ASM has a census of financial information for units above the take-none threshold. Sample weights are used by the ASM to produce estimates of commodities.

To account for the take-none portion, the MSM uses a factor based on the last benchmarking exercise. The idea behind this methodology is that if the MSM and the ASM are measuring the same quantity, any difference between the two level estimates can be attributed to the take-none portion. Thus, the percentage difference between the two level estimates is carried forward to account for the take-none portion when producing MSM estimates. This factor is applied until the next benchmarking exercise. As part of the restratification exercise in January 2007, this methodology was replaced by one that is similar to that used for a sample of simple units above the take-none threshold (see section 2.2.4). That is, an estimated survey value will be obtained from GST data and a statistical model.

In the ASM, for units below the take-none threshold a combination of T1 and T2 annual tax data is used to calculate totals of selected financial variables. Unincorporated businesses (sole proprietors and partners) file a T1 tax form, while incorporated (corporations) file a T2 tax form. T1 data are not available for all units in the T1 tax universe, so a sample of these records is drawn and an estimate is produced based on this sample. A census of T2 information is available and data from this census is used for T2 units below the take-none threshold.

The working group felt that the take-none methodology did not significantly impact the 2004 benchmarking exercise. However, it also feels that the use of benchmark factors to account for the take-none portion of the MSM is not appropriate and are encouraged by the change in the take-none estimation methodology. In addition, the working group would like to see some evaluation of the quality of the pseudo-census produced by the ASM. Perhaps weighted estimates of financial variables based on sampled units could be produced and compared to those obtained from the pseudo-census. These weighted estimates would also allow the calculation of quality measures.

2.3 Coherence of Operational Aspects

2.3.1 Collection Processes

The ASM is a mail-out/mail-back survey with telephone follow-up for nonrespondents and for edit failures. The MSM uses both mail-out/mail-back and computer assisted telephone interviewing collection methods, depending on the preferences of the respondent. For the ASM, a score function is used to manage the follow-up process. The score function takes into account the importance of a unit in terms of its contribution to the total value of shipments at an industry/province level and to the total value of shipments for each commodity produced by the unit. On the other hand, the MSM does not use such a function and their follow-up is not managed in any specific fashion.

Both surveys use similar collection processes with only minor differences and, as such, it was felt that the process had very little effect on the differences observed between the two survey estimates.

2.3.2 Reconciliation

Reconciliation is the process of comparing estimates from the two surveys, at different levels of detail, in an attempt to identify large differences. If large differences are found, then analyses at lower levels may be performed. One type of analysis is a micro-level analysis where data from the two surveys are compared for the same unit. If large differences are found at the micro-level, the data are investigated and if an error is found it is then corrected. The goal of this exercise is, as much as possible, to ensure consistency between the two surveys before benchmarking.

Reconciliation is an important step in ensuring coherence between sub-annual and annual surveys and should be done regardless of whether benchmarking is to be carried out or not. Although reconciliation is a potentially time consuming process, it should be built into survey processing schedules in order to ensure coherence between surveys.

3.0 Analysis of Coherence between MSM and ASM

Some of the differences outlined above were studied in more depth and these analyses are presented in this section. The first analysis looked at the effect of the two surveys using different sources of tax data (GST versus T1/T2). A second analysis looked at the effect of the different reference periods as a result of the non-calendarization of the ASM data.

3.1 GST versus T1/T2

As mentioned earlier, ASM uses T1/T2 annual tax data, while MSM uses monthly GST data. Each of these tax data sources has its own processing system. Since the source and processing of tax data used by each survey are different, it could potentially be a source of variability that contributes to incoherent estimates. A major difference between the two annual tax data sources is that T1 data are not available for all units in the T1 universe, while T2 data are available for all units in the T2 universe. Estimates from the T1 data are obtained at an aggregated level. T2 data are used for units above the take-none threshold and comparisons were made only between T2 and GST tax data, in the context of the manufacturing population.

According to the BR, in 2004 there were 98,951 T2 records that were active in the manufacturing industry. Of these records, 82,430 were found on the T2 data file. Further investigation showed that the 16,521 units that were not on the T2 data file were all inactive. Of the remaining 82,430 records, a further 21,812 were dropped because they did not report for a full period (that is, a start date which was the first day of a month and an end date that was the last day of a month) or were not on the GST file for the full time period reported on the T2 file. With this population of units, the total annual revenue used by the ASM (as defined through the COA) was compared to the sum of the twelve months of GST revenue corresponding to the fiscal period of the corporation as declared on the T2 return. Over all industries at the Canada level, the T2 revenue was 20.5% higher than the GST revenue. The relative difference by industry varied from 3.5% to 37%, with T2 always higher than the GST. This difference is a little misleading as it includes all units, complex

and simple, in the manufacturing industry while the two surveys use tax data for simple units only. If the complex units are removed, the number of units reduces to 57,044 and the relative difference becomes 6.2% at the Canada/All industry level. While the T2 total is still larger at this level, there are several industries where the GST is larger than the T2. Finally, only those simple units that had tax used in both the MSM and the ASM surveys were investigated. Of the 57,044 units, only 1,110 had tax data used in both surveys. Looking at these units, the relative difference at the Canada/All industries level was 6%.

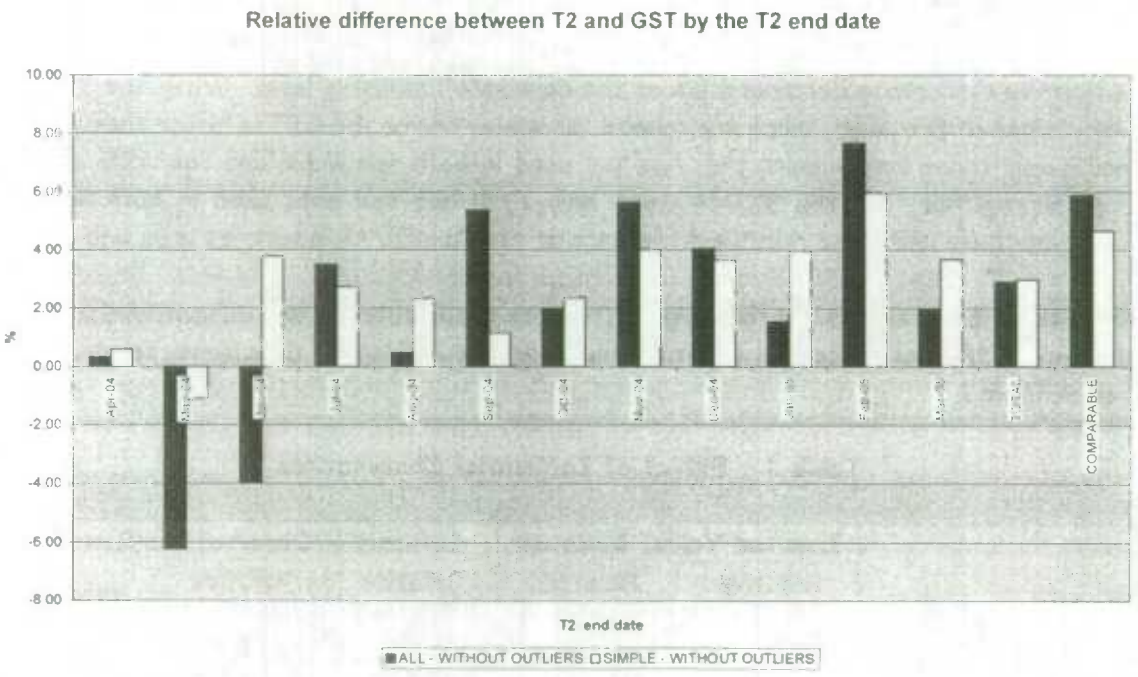
Two possible explanations for this large difference are imputation and influential data. The relative differences, once the imputed values or the influential data were removed, are given in Table 1.

Table 1 - Effect of Influential Observations

Population	Imputed Values Removed		Influential Observations Removed		
	Relative Difference Before	Relative Difference After	Relative Difference Before	Relative Difference After	Number Removed
All (60,618)	20.5%	10.7%	20.5%	5.9%	129 (0.4%)
Simples (57,044)	6.2%	4.8%	6.2%	4.7%	251 (0.4%)
MSM/ASM (1,110)	6%	4.4%	6%	2.7%	29 (2.6%)

As one can see, the removal of imputed or influential units has a large effect on the population containing both complex and simple units (All) but less on the population of simples. Despite the removal of imputed and influential units in the population of simples, the relative difference between the two sources remains rather high.

Another aspect of the tax data that was investigated was the effect of having different reference periods between the two data sources as corporations filing a T2 tax return are free to choose their own fiscal year end. The analysis presented above matched the GST to the T2 reference period. Another analysis was performed where the T2 data were compared to the GST data that covered January to December 2004. Including all complex and simple units but excluding outliers, there were 66,870 records for which T2 and GST data were available. At the Canada/All manufacturing industries level, the difference between the two sources was 2.9% with T2 again being larger than the GST value. If we restrict the analysis to simple units only (62,235), the difference increases to 3% with T2 being higher than GST. Looking at the relative differences by the end month of the fiscal year as reported on T2 data, we see that there is no discernable pattern.



In general, the T2 totals are higher than the GST totals except for units that had fiscal years end in May or June. At an industry level, the results are quite varied. Most industries have several months where the T2 is higher than the GST, while one or two industries have the T2 always higher than the GST. However, there does not appear to be any pattern within any industry.

In summary, it appears that there are some differences in values obtained from the two sources of tax data. Regardless of whether or not the reference periods line up, the T2 total appears to be slightly higher than the GST total. One thing to note is that the differences illustrated here may not necessarily translate to similar differences in the MSM and the ASM for several reasons. First of all, the GST data that were used in this analysis were close to final and, thus, had very little imputation. In practice, MSM uses GST data that has much more imputation. Unfortunately, the actual files used during 2004 were not available for this analysis. In addition, when GST is used by MSM, there is a statistical model that is applied to adjust for possible differences in concepts and timeliness issues. This model was not applied in this analysis. Finally, by design, the units for which the MSM uses GST data are the smaller units in terms of revenue. Thus, it is felt that the difference observed between the GST and T2 data should not significantly contribute to the gap between the two survey estimates.

3.2 Non-Calendarization of the ASM

As previously mentioned, respondents to the 2004 ASM were asked to provide information on their own 12 month fiscal year that ended between April 1st, 2004 and March 31st, 2005. During the benchmarking process, the MSM values for January to December 2004 are summed up and then adjusted to match the 2004 ASM value regardless of the true

reference period of the ASM units. The distribution of ASM units according to their reporting periods is given in Table 2.

Table 2 - Distribution of ASM units by Reporting Period

Reporting Period	Number of Units	Percentage of Units	Percentage of Total Shipments
Jan 04 - Dec 04	11,911	35.3%	65.4%
Start 03 - End 04	16,085	47.6%	22.7%
Start 04 - End 05	5,059	15.0%	10.3%
Not 12 months	726	2.1%	1.6%
Overall	33,781	100%	100%

From Table 2, we see that only 35% of the ASM units reported for the 2004 calendar year, but these units represented 65% of the total shipments (unweighted). That is, the units that reported for the calendar year tended to be the larger units. In terms of units that didn't match the calendar year, we see that approximately 48% of all units reported for a 12 month fiscal period that started in 2003 and ended in 2004. These units represented 23% of the total shipments. In comparison, 15% of the units that reported for a 12 month period that started in 2004 but ended in 2005 represented only 10.3% of total shipments. Based on these results, it appears that units that started their fiscal period in 2003 were smaller, in term of shipments, than those that started in 2004. The remaining units reported for a period that did not correspond to twelve months.

To investigate the effect of the non-calendarization of the ASM, the ASM data were linked to the MSM data covering the January to December 2004 time period. Of the approximately 34,000 units in the ASM 8,892 were found in the MSM. The relative differences of Total Shipments (unweighted) between the two surveys by reporting period are given in Table 3. Note that the relative differences are defined as the MSM shipments minus the ASM shipments divided by ASM shipments and are expressed as a percentage.

Table 3 - Total Shipments for Common Units

ASM Reporting Period	Number of Units	Percentage of ASM Total Shipments	Relative Difference in Total Shipments
Jan 04 - Dec 04	4,155	71%	1%
Start 03 - End 04	3,430	18%	5%
Start 04 - End 05	1,102	10%	4%
Not 12 months	205	1%	20%
Total	8,892	100%	2%

In terms of total shipments (suppressed for confidentiality reasons), these 8,892 represent over 80% of the overall ASM total shipments. This implies that the MSM is covering the 'important' units in its sample. Approximately 47% of the linked ASM units report for the January to December calendar year and they represent over 70% of the ASM total shipments for the linked units. Units with fiscal periods starting in 2003 and ending in 2004 represent approximately 18% of total shipments, with the units with fiscal periods starting in 2004 and ending in 2005 making up approximately 10%. Units that did not report for 12 months represent the remaining 1%.

For units where the fiscal period matches the calendar year, there is only a 1% difference in the total shipments between the two surveys with the total from MSM being larger than the ASM total. As could be expected, units whose fiscal periods do not match the calendar year show a larger discrepancy (4.5% versus 1%). These numbers seem to indicate that the non-calendarization of the ASM does have an effect on the coherence between the two surveys.

In a stable economy, the non-calendarization of the ASM should not contribute to a difference between the results of the two surveys. For units whose ASM reporting period started in 2003, their ASM total shipments will include several months from 2003. In comparison, their annualized MSM total shipments will include several months from 2004 that were not covered in their reported ASM data. However, if the economy is stable then their monthly total shipments would be fairly stable and replacing values for these months in 2003 by the corresponding 2004 months should not lead to significant differences. On the other hand, in an economy that is increasing, one would expect to see the MSM totals for units that started in 2003 and ended in 2004 to be higher than the ASM totals but for the MSM totals to be lower than the ASM totals for those units that started in 2004 but ended in 2005. The opposite behavior would be expected if the economy was showing a downward trend. In 2004, the MSM showed a positive trend but Table 3 does not show the expected behavior. That is, one would expect a positive relative difference for units that started in 2003 and ended in 2004 and a negative relative difference for those units that started in 2004 and ended in 2005. An explanation for this apparent discrepancy is being investigated.

An additional analysis broke down Table 3 in terms of the source of the data for the units that reported for the same time period and is given in Table 4. Note that the 'source' is based on the ASM and does not necessarily reflect the MSM source. Comparisons of units that used tax data for both surveys were not performed because of the small number of units that would be identified.

From Table 4 one can see that while the differences between the tax reporters are larger (relatively) than the difference between respondent data, one must take into account the contribution of tax to the total overall shipments (last column). That is, even though units that match the calendar year have a 4.2% difference for tax units, this difference represents only 0.1% of the total shipments. On the other hand, respondent data has a relative difference of 0.9% but this represents 0.6% of the overall total. Based on this,

even if ASM were to be calendarized it appears that there would be a difference in the data, be it reported or coming from tax.

Table 4. Total Shipments for Common Units by Source - MSM values Adjusted to ASM Reporting Period

ASM Reporting Period	Source	Number of Units	Relative Difference	
			To ASM Reporting Period Total	To Overall Total
Jan 04 - Dec 04	Respond	3,180	0.9%	0.6%
	Tax	975	4.2%	0.1%
Start 03 - End 04	Respond	1,829	4.4%	0.7%
	Tax	1,601	7.2%	0.2%
Start 04 - End 05	Respond	571	3.4%	0.3%
	Tax	531	5.3%	0.0%
Not 12 months	Respond	132	18.5%	0.3%
	Tax	73	32.2%	0.1%
Total	Respond	5712	2.0%	1.9%
	Tax	3180	6.6%	0.4%
	All	8892	2.3%	2.3%

One possible solution to the non-calendarization of the ASM would be to match the MSM reference period to the ASM period. This was done with the linked units and the results are given below in Table 5. Note that units that did not report for a full 12 months have been removed.

Table 5 - Results for Linked Units

ASM Reporting Period	Number of Units	Relative Difference of Shipments
Jan 04 - Dec 04	4,155	1%
Start 03 - End 04	3,430	2%
Start 04 - End 05	1,102	4%
Total	8,687	1%

As one can see from the Table 5, the matching of reference periods has reduced the differences between the units that start in 2003 and end in 2004, but has not affected the units that start in 2004 and end in 2005. Overall, the difference has been reduced from 2% to 1%. Thus, even if the MSM was forced to match the reference period of the ASM during benchmarking, it appears that there would still be a difference between the two surveys.

4.0 Discussion

The goal of the working group was to recommend whether or not the MSM should be benchmarked to the ASM. The strategy of the working group was to look at the two surveys in terms of conceptual, methodological and operational differences and to see if these differences could be identified and quantified. It was felt that if the effect of these differences could be eliminated or at least minimized, the remaining differences in the estimates would be small and that the mathematical exercise of benchmarking could look after them. However, after performing the coherence analysis between the two surveys there does not appear to be many differences between the two surveys that could be eliminated quickly. In terms of concepts, the two surveys appear to be very close. There might be some minor differences in the way that the questions are asked but they appear to be targeting the same concepts. It is unknown whether respondents from one survey are systematically reporting differently than respondents from the other and this will likely remain unknown. There is a problem with the different time periods, but it is doubtful that a calendarization method could be developed quickly for the ASM. Even if a method could be quickly developed, an analysis showed that there might still be a difference between the data reported for a unit for the same time period. A preliminary analysis of 2005 data showed a similar behavior as the 2004 data.

In terms of methodological aspects, there are some areas where the two surveys could be made more coherent. The sampling frames are both coming from the BR, but units that have been active for only a portion of the target reference period may be handled differently. Exactly how these units are handled should be evaluated and methods to allow them to be treated the same way by both surveys should be investigated. In terms of stratification, the same measure of size should be used by the two surveys as well as the level of industry grouping. With the newly restratified MSM it appears as though the take-none strategies are similar and should result in some reduction of the differences between estimates from the two surveys. Although the tax data used by the two surveys are different, it does not appear that this contributes significantly to the difference in the estimates. For the ASM, tax units contribute only 14% to the overall total and even less for the MSM. Thus although our analysis showed that there were some differences in the two tax sources, it is felt that overall they were insignificant. In terms of estimation, it is hard to quantify the true contribution to the overall difference because the ASM produces a pseudo-census and the MSM produces weighted estimates. We are currently working on a method to quantify the effect of using the MSM weights. One point that came out of the analysis was that for units that reported for the calendar year, there was still a 1% difference in the unweighted totals coming from the two surveys.

In terms of operational aspects, there does not appear to be many differences. Although the ASM uses a score function for follow-up, it is unlikely that it significantly contributes to the difference between the surveys. The reconciliation process is performed by staff from both the MSM and the ASM and it is possible that changes could be made to either survey during the process. Results from recent reconciliation exercises show that there are only minimal changes (approximately 1% in terms of shipments in 2004 and 2005), but it

remains an important step that needs to be performed regardless of whether or not benchmarking will be done.

One aspect that the working group did not cover was an investigation of the actual benchmarking method that was applied. In 2004, the Denton method (Cholette 1984) was used and this method took the benchmarking factor from December 2004 and used it to adjust the MSM values from that point on. It was this factor that caused the 2004 to 2005 growth rate to be revised from 2.9% to 0.6%. A risk of using the Denton method is that the forecasted values are based on only the last month covered by a benchmark and that it is highly influenced by the differences between the last two years. Generally speaking, the Denton method assumes that if, based on the last two years, it appears that the differences are increasing then the next difference will be larger still. If on the other hand the differences are decreasing, then the next difference will be smaller. There are many other methods available to produce the forecasted values that would have led to a smaller revision. For instance the Canadian System of National Accounts (CSNA) uses an explicit forecast of the next annual discrepancy with the Denton method. If this method had been applied, the 2004 to 2005 growth rate would have been higher than the 0.6% obtained from the Denton method. For more on the technical aspects of benchmarking and other methods available, we refer the reader to Quenneville and Fortier (2006).

4.1 Recommendations

Like most sub-annual and annual surveys, the ASM and MSM were more or less developed independently. As such, inconsistencies in concepts, methodology and operational aspects have led to differences in the estimates produced. Benchmarking methods were most likely not considered until after the designs of the surveys were finalized. These methods were applied so that the survey estimates would be equal even though there may have been some fundamental differences that would imply that the estimates should have been different. The working group recommends that benchmarking be considered much earlier in the process and not simply as a 'tool' to deal with the differences in the estimates. Coherence should be built into the surveys as they are designed, thus minimizing or perhaps eliminating the need for benchmarking. It is recognized that it is not possible to have 100% coherence between the surveys but inconsistencies should be controlled as much as possible. Many of the inconsistencies identified in this document could be eliminated, while others are the result of conscious decisions due to operational or methodological reasons. The goal when designing these surveys should be to eliminate as many inconsistencies as possible so that the source(s) of the remaining differences are known and can be explained. If surveys were designed as such, then benchmarking would be used to account for only sampling variability and explainable (and possibly quantifiable) differences.

Turning to the ASM and MSM situation, regardless of the final decision to continue benchmarking or not efforts should go into making the two surveys as coherent as possible. As long as Statistics Canada publishes both sets of estimates, there will always be users who will compare the two estimates and will ask why there is a difference. Thus, it would be useful to perform a coherence analysis, such as the one described in this document, from time to time and that any differences that are identified be resolved if possible. Remaining

differences should be resolved at appropriate times, such as restratifications or redesigns. A reconciliation process should take place at the time of annual revisions and should become part of the annual survey process.

The question of whether Statistics Canada should continue to benchmark sub-annual surveys to annual surveys was posed to the Statistics Canada's Advisory Committee on Statistical Methods (ASCM) and there was some concern from the committee on the practice of blindly applying benchmarking to eliminate differences in the estimates. Two of the recommendations from the committee were:

1. If the estimated differences between ASM and MSM are larger than that attributable to sampling and measurement error, every attempt should be made to identify the causes. It is most important to identify reasons for period-to-period changes in the level of the difference between the two surveys. If causes, such as data collection changes and calendar effects, can be identified then procedures can be modified or variables introduced into the forecast equations.
2. The committee cautions against the extensive use of benchmarking. There are situations where benchmarking is not appropriate.

Based on the analyses performed, the reaction from the ACSM and in consultation with the Benchmarking steering committee, the working group recommends that the MSM no longer be benchmarked to the ASM. One of the key users, the CSNA, has stated that it can use the MSM series for estimates of change, as well as for their quarterly GDP estimates, and will use the ASM for their structural estimates. A communications strategy will be needed to inform other users of this decision and, on an ongoing basis, to have clear explanations of the differences between the two surveys and the objectives/uses of the two. For example, the MSM monitors monthly trends, while the ASM has more detailed structural analyses. The ongoing part of the communications strategy can be achieved by adding documentation and explanatory notes where appropriate to support users of these data.

When developing options to implement this decision, there was another aspect that had to be kept in mind. Starting with the January 2007 reference month, the MSM introduced a restratified sample. The restratified sample is essentially an updated sample to account for frame updates, updated stratification boundaries and some other improvements to the sample design and estimation. Given that both these changes needed to be introduced in approximately the same timeframe, it was decided that they would be implemented together. Several options were considered and investigated, with the following option recommended by the working group.

> The MSM series would be benchmarked to the revised ASM 2004 and to the preliminary ASM 2005 value using the modified Denton method that is currently being used in the CSNA. In addition, the MSM series would also be linked to the January 2007 restratified estimates and to the published December 2003 estimates at the same time. This method imputes a value for the ASM 2006 such that the annual growth rate from 2005 to 2006 is the same as that of the annualized MSM series. This method also uses the January 2007 restratified

estimates and the December 2003 MSM published estimates as benchmarks. Finally, the difference between the 'old' and the new restratified survey would be wedged back to January 2005. As a result of linking to the January 2007, the benchmarking would be removed. As a result of linking back to December 2003, no breaks would be introduced.

Acknowledgement

The authors would like to acknowledge Don Royce, Peter Lys, John Kovar, Andy Kohut and Claude Turmelle who have contributed significantly to improving the readability of this document.

References

Cholette, P. A. (1984). Adjusting Sub-Annual Series to Yearly Benchmarks, *Survey Methodology*, **10**, pp. 35-49.

Dagum, E. B. and Cholette, P. A. (2006). *Benchmarking, Temporal Distribution and Reconciliation Methods of Time Series*. Springer-Verlag, New York, Lecture Notes in Statistics, #186.

Lavallée, P. and Hidiroglou, M.A. (1988). On the Stratification of Skewed Populations, *Survey Methodology*, **14**, pp. 33-43.

Quenneville, B. and Fortier, S. (2006). *Benchmarking MSM Shipments*. Unpublished Statistics Canada document.