

## **The Data Warehouse and analytical tools to facilitate the integration of the Canadian Macroeconomic Accounts**

Alistair Macfarlane<sup>1</sup>

### **Abstract**

The Data Warehouse has modernized the way the Canadian System of Macroeconomic Accounts (MEA) are produced and analyzed today. Its continuing evolution facilitates the amounts and types of analytical work that is done within the MEA. It brings in the needed element of harmonization and confrontation as the macroeconomic accounts move toward full integration. The improvements in quality, transparency, and timeliness have strengthened the statistics that are being disseminated.

### **1. Introduction**

The MEA is a highly integrated macroeconomic accounting system which requires the integration of concepts, methods, and data across a series of economic accounts dealing with production, consumption, incomes, financial flows and stocks. This inte

One of the ways Statistics Canada has addressed these challenges is through the development of a Data Warehouse which facilitates this integration process. This paper outlines the integrated nature of the MEA and how the Data Warehouse facilitates both the compilation of the data and acts as a multidimensional analytical tool that provides Business Intelligence (BI) Analytics to the compilers of the Canadian System of National Accounts at Statistics Canada (CSNA).

### **2. The Canadian macroeconomic accounts**

The Canadian Macroeconomic Accounts provide economic information to the Canadian public and economic stakeholders by using a highly integrated central framework which validates the interrelationship between production, consumption, and accumulation of income and wealth<sup>2</sup>. Each component, in itself, reveals an important aspect of the Canadian economy at a particular time in history or provides a time series perspective. Yet the value added in the Canadian context is the integrated relationship approach which requires the underlying data to be in sync with each component. Therefore a dollar of spending is matched by a dollar of income at any given point in time.

Any suggestion that this process is done with relative ease or is considered the norm for a National Statistical Agency is far from truth. Statistics Canada's commitment to improving the quality and timeliness of its macroeconomic statistics by the integration and reconciliation of the framework is unique. It began in earnest with the 1986 SNA Historical Revision, when a common GDP aggregate was introduced. According to Kishori Lal, a former Director General of the National Accounts, "integration requires a strong and long term commitment by the senior management, a willing cooperation by various parts of the CSNA and a dedication to develop and implement a common set of

---

<sup>1</sup> Alistair Macfarlane, Statistics Canada, 100 Tunney's Pasture Driveway, Ottawa, Ontario, Canada, K1A 0T6  
([alistair.macfarlane@canada.ca](mailto:alistair.macfarlane@canada.ca))

<sup>2</sup> Tebrake, J.(2008)- Director General, Macroeconomic Accounts Branch, Statistics Canada, CSMA Foundations  
Module 1\_Essential Concepts\_3-Powerpoint Presentation

classification and aggregation rules.....”<sup>3</sup> True to these words, some 30 years later, the integration of the macroeconomic accounts is still evolving.

### **3. Advantages to more integration**

At the outset many of the different macroeconomic outputs produced by the then System of National Accounts (now the MEA) were produced in relative independence of each other. Although there are underlying common inputs; working in independence or at a lower level of integration led to different results. There were different production schedules, accounting standards and revision cycles. There were known sector and universe differences, different levels of granularity, reporting frequency and data availability issues. The less integrated the framework the greater the likelihood that statistical differences could exist. When two outputs are conceptually measuring the same phenomenon but producing two different values, it begs the question: which one is right? Reconciling data from within the accounts is an integral piece of the validation and an essential part of any statistical process. Yet as the macroeconomic accounts grew in size and complexity so did the ability to reconcile and explain, in a timely fashion, the differences. The time had come to move the macroeconomic accounts to a more integrated approach and the need for a corporate tool to reconcile the differences became greater. Therefore the Data Warehouse was developed to alleviate the conceptual and IT effort required to confront data from different sources.

The advantages of a more integrated macroeconomic accounts are numerous, the most significant being the quality of the data is strengthened. The interrelationship between production, consumption, and accumulation of income and wealth is more evident with the integration. A movement in one aggregate is reflected immediately in other components due to connectivity between the accounts.

There will be more actual data available to be used in each production cycle. With the integration of the macroeconomic accounts reduces the need for estimation of macroeconomic components that were previously out of sync due to a less integrated system. Actual data is more readily available and used in the statistical output, bringing a greater level of confidence in the information.

The integrated approach brings in other data sources to confront, reassuring users that the statistical message being relayed is correct. Conceptually, the greater the number of sources of data that can be used in any confrontation or as indicators, the greater the validity of the data.

Another advantage would be the interchangeability of the data sources. A single feeder can supply several outputs with being reprocessed or adapted to the specific needs of the consumer.

### **4. More integration means more co-dependence**

Yet as the integration evolves, so does the inter- dependency and complexity of producing statistics in a timely fashion. In order to build a more integrated framework there is a need for a relational platform that would allow source data to be transformed into multiple classification frameworks and a reconciliation between those frameworks. As well, it would allow any feeder data sources not being directly produced by the macroeconomic accounts itself to be linked and visible to all.

As data are increasingly shared or become the inputs of other estimates, then an enhanced level of co-dependence will exist. In order to produce more timely statistics for stakeholders there must be a high degree of harmonization within the production schedules of the macroeconomic accounts. This would allow for data to flow as quickly as possible through the production system to and be disseminated in a more timely fashion.

---

<sup>3</sup> Kishori, L. (2003), “Canadian System of National Accounts – An Integrated Framework”, in Collected Articles of Kishori Lal, Statistics Canada Library, Catalogue no. 13-598-XIE, page 324

In the same way as envisioning the most efficient transfer of a baton between track and field relay runners as the way to have the lowest race time, the integration of the macroeconomic components must be efficient and seamless in order to produce timely statistics.

## **5. The Data Warehouse facilitates integration**

The Data Warehouse allows for the compilation and production of macroeconomic statistics. The source data are presented in multiple classification frameworks with a reconciliation between them made possible. Sources of data such as CAPEX<sup>4</sup>, wage and benefits data and tax data are made available to MEA analysts for use in data confrontation.

Therefore the MEA Data warehouse is, conceptually, the “roads and sewers”<sup>5</sup> that allows the flow of data from multiple divisions and different output systems (i.e. INCOMM, FAME, KEYNES)<sup>6</sup>, to be compared and confronted, leading to a more integrated set of macroeconomic accounts. It also contains the ability to data mine down to the lowest level of the data collection that allows for a greater ease during analytical work or reconciliation exercise.

## **6. Steps towards integration**

The first challenge for the Data Warehouse was to harmonize conceptual issues. A lot of work was done by the Data Warehouse team (senior management, methodologists, accountants, IT resources, and subject matter experts) to harmonize frameworks, accounting concepts, production schedules. If, in theory, the same data were being used to build two or more macroeconomic outputs with only the classification system differing, then at the aggregate level the numbers must equal. Example: once by industry and once by commodity. Although the classification system is different, the aggregate sum should equal. Yet in many cases this was not true; and the reasons for this were quite valid. Highlighted below are some of the challenges that were encountered during the beginning stages of the MEA harmonization.

The Data Warehouse accomplished harmonization by

1. Co-ordinating production schedules. Different Statistics Canada divisions produced and revised data on very different schedules. In order to be able to confront the data the synchronization of the outputs became necessary. Not only for the final product but also when the product is an intermediate statistical input to another process. Considerable time and effort over several cycles was needed to harmonize the survey and production schedules in order for the Data Warehouse model to work.
2. The alignment of frameworks and classification, as well as the maintenance of versioning. The streamlining and standardization of divisional frameworks. As the integration further develops so does the inter-dependency on each classification needing to be uniformed. Standardization of classifications between MEA Divisions became essential. This also meant the versioning also had to be in sync. One cannot compare one division’s data coded to NAICS<sup>7</sup> 2012 and with those of another division’s coded to NAICS 2007.
3. Accounting concepts needed to be harmonized. Unfortunately the world of finance does not always afford comparability. Businesses and governments can and do have different year-end dates (calendar vs. fiscal). Accounting frameworks can be mixed between cash and accrual basis. Bringing uniformity to be able to compare was required.

---

<sup>4</sup> CAPEX is the acronym for the Capital Expenditure Survey, produced by Statistics Canada.

<sup>5</sup> Sabourin, J.D. (2015) reference made by a former MEA Data warehouse Content Manager and Statistics Canada employee, Jordan Sabourin to describe his view of the structure of the Data warehouse.

<sup>6</sup> INCOMM is the production system used by the Industry Analysis Division (IAD), FAME is the production system used by the National Expenditure Accounts Division (NEAD), and KEYNES is the production system used by Public Sector Statistics Division (PSSD).

<sup>7</sup> NAICS 2012, a recognised North American Industry Classification System (NAICS) plus version year.

4. The frequency and granularity of reports had to be comparable. One could not directly compare quarterly data with monthly or annual. Neither is it appropriate to compare preliminary with final data and expect them to be equal.
5. Building concordances to link classifications together. Ultimately the biggest achievement was the ability of the Warehouse to showcase two or more sources of data on the same page and allow microanalysis to be done. This ability arises because of Data Warehouse concordances that allow different classifications to be transformed into other classifications. Thus an Industry Analysis Division (IAD) industry classification can be transformed into a National Economic Accounts Division (NEAD) commodity classification using the same underlying data. As an example, we can present business sector surplus from an IAD, NEAD and tax perspective, all equal in terms of dollar value but with 3 different classification choices.

## **7. How the Data Warehouse empowers analysts**

One of the most important roles of the Data Warehouse is to remove the operational processing complexity through automation. As the macroeconomic accounts have become more integrated the complexity linking those components has also risen. There is an enhanced possibility of error as subject matter experts must rely on their own IT expertise or the expertise of fellow analysts to compile the proper data to be analyzed. The MEA Data Warehouse returns to the analyst the time and effort they would use formulating these linkages and validating the data in the Warehouse as fit for use. A second gain is Business Intelligence Analytical tools are available to all analysts and not just a selected few. More often than not analysts have, by need, created their own BI Analytical tools for their own exclusive use. The Warehouse makes these tools available to all, allowing for more in-depth analysis to be done by more people.

## **8. Customization**

Although the Warehouse focuses on standardization and empowering the masses, it also realizes that there are specific wants and needs of the individual analyst. Thus the ability to customize and change queries with relative ease is an essential part of a good analytical tool. The Data Warehouse provides analysts with a set of flexible analytical tools that allows the data to be displayed in many ways (i.e. titles, classifications, vintages or cycles). Based on a pivot table structure, analysts can drag and drop from a list of categories (called dimensions) the variables they want to see into a Microsoft Excel based report. Dimensions can be added, removed, or juxtaposed on the fly specific to the analyst's wishes. These initial data are available at the aggregate or sub-aggregate level. But further analysis can be done with help from the generation of secondary SSRS<sup>8</sup> reports that provide data down to the lowest level of input. Other tools such as growth rates and deltas can be derived by the Data Warehouse, based on the parameters the analyst chooses within their dynamic pivot table or SSRS report. This allows quicker and more comprehensive analysis of the data.

## **9. Transparency**

Another very important advantage is in the transparency of the underlying data that sum to the aggregated totals. The Data Warehouse provides divisional data at its fit for use stage. It comes with the underlying data that is involved in producing the aggregates. This includes any adjustments made and those are identified as such. This transparency is ultimately the value added to any reconciliation process as it removes any bias or unknown from any reconciliation process. It also forces adjustments done to be explainable and not just accepted.

---

<sup>8</sup> SSRS- Sequel Server Reporting System- a table format of reporting underlying data.

## **10. Built for today and tomorrow**

The MEA Data Warehouse was built to facilitate the integration of the macroeconomic accounts today but also with the foresight to adapt to the changing needs of the analytical community going forward. The “plumbing” is there to accept more data, ‘big data’ or unstructured data; all can be incorporated with the foundation that is in place. Once there, the Data Warehouse becomes a multidimensional analytical tool that provides Business Intelligence Analytics so that the expertise of the analyst can take hold and meaning can be derived from the data. Bringing more output with less IT effort, more transparency and less statistical discrepancy. The Data Warehouse provides that ability of conversion with transparency by providing concordance tables in the metadata process.

There is still more harmonizing of the business processes and production systems needed to allow further integration with other parts of the macroeconomic accounts. Small steps continue to be made as users seek inputs and the needed resources become available, but full integration is still years away. There is little disagreement as to the benefit of more integrated macroeconomic accounts and the advantages that the MEA Data Warehouse affords to producing better statistics. Yet time and resources are equal considerations as to when integration can be expanded without compromise of what is in existence.

## **11. Methodology needs**

The Data Warehouse is an ever-evolving multidimensional analytical tool that is only limited by an analyst’s imagination. As the macroeconomic accounts continue to further integrate and bring in other data sources, the more opportunities there will be to look at data differently. Like the integration of the macroeconomic accounts themselves, the adaption of the Data warehouse as the default analytical tool is an ongoing process. There is substantial operational investment still required to bring more divisional data into the warehouse. Yet as more data become available for consumption there has to be an equal amount of methodological assurance that the statistical methods and theories being applied within the IT systems of the Data Warehouse are appropriate and, in the case of new developments, that the fundamental principles are respected while producing “state of the art” statistics. Methodology will need to play an important role in the future of the Data Warehouse.

## **12. Conclusion-Opening doors**

The Data Warehouse can and should be so much more. We (close to the product) see the immense potential of the system. We envision the enormous growth in analytical work that could be done if all of the macroeconomic accounts were fully integrated within. The improvements in quality, timeliness, and accountability that come with the integrated approach would be there in such a scenario, and the output would be better. Today’s users of the Warehouse are pushing for more. The casual users of yesterday are now the power users of today. Not only is the Data Warehouse able to do more than it could two years ago, it continues to integrate the changes the individual analyst wants while still aligning itself with the vision of integration the Corporation sees the future of macroeconomic statistics to be. The move from ASM to IBSP<sup>9</sup> and the addition of ten new cubes is a prime example of how the Warehouse has transformed the way macroeconomic statistics are being produced.

As the integration continues, the Data Warehouse will be called upon to contain, process and link with more economic information. As the number of user grows so will the variety and complexity of analytical tools they require. Users are already asking for more reports, more dimensions and more data. The Data Warehouse team must meet these challenges ahead on to allow analysts to go deeper into understanding the data. Exploring not only the data itself, but also how they interact with other pieces of the economic system, looking at both the Canadian and foreign context. Paving the way for a clearer picture offered to all.

---

<sup>9</sup> ASM (Annual Survey of Manufacturing) is now a much larger component of IBSP (Integrated Business Statistics Program)