

## **Challenges and Benefits of Producing Business Statistics within a Highly Centralized Model**

**Daniela Ravindra, Statistics Canada**

**Jean Talon Building, Tunney's Pasture Driveway, Ottawa, Ontario**

**Canada**

### Abstract:

In 2009, Statistics Canada launched the development of the Integrated Business Statistics Program (IBSP), a highly centralized and standardized model for producing business statistics. The objective was to develop an infrastructure comprising mostly of generic systems as well as a set of statistical methods and data processes which are common across different business statistics programs. The main drivers for the initiative were the need to become more cost efficient and the desire to improve data quality (coherence, timeliness, relevance). Reducing response burden, specifically for small and medium sized enterprises, was an important Government of Canada priority and, therefore, became a secondary objective of the project. Since the scope of the project was rather extensive, the plan was to start with a group of programs (58 annual surveys) and then, gradually, over a period of five years, complete the integration XX programs).

Development work was completed in spring of 2014 and the first wave of programs has been in production under the new model ever since. More than a dozen programs have collected processed and disseminated data up to now. This has allowed Statistics Canada to make a first determination of how the model works and what adjustments are required to address elements which did not perform as expected. This paper will present the challenges associated with managing this type of program, as well as the benefits that Statistics Canada has realized since the implementation of the IBSP. There are many aspects that will be reviewed: content management, interdependencies (internal and external), risk management, internal and external relationship management, human resources management, production schedules and financial management. A number of solutions have already been put in place to address challenges and to build on the more successful aspects of the project and they will also be covered in the paper.

Key Words: centralization, generalized model, business statistics, Statistics Canada

## 1. Introduction

In 2010, Statistics Canada launched the Corporate Business Architecture initiative which entailed a review of Statistics Canada's business processes, methods and systems infrastructure. The objectives were to:

- Achieve efficiencies,
- Enhance quality assurance and
- Improve responsiveness in the delivery of statistical programs.

The initiative spawned a number of transformational projects in the area of administrative processes, collection infrastructure as well as survey data processing and analysis. At the time, the business statistics program at Statistics Canada comprised approximately 250 surveys and administrative-based programs. Although the majority of these programs had been developed in isolation, there was a great deal of commonality between the processes and methods employed to produce the information. This determination led to the recommendation that the organization move towards the development of a generalized model for producing business statistics. The Integrated Business Statistics Program (IBSP) was conceived. It makes use by and large of shared and generic corporate services and systems for collecting, processing, disseminating and storing statistical information. In addition, the surveys' content is harmonized wherever possible, and the approach to data analysis is fairly similar across the different programs. After four years of development and testing, the new model was deployed in production for 60 annual business surveys<sup>1</sup> in January 2014. Two cycles of the annual surveys have been processed and disseminated so far and more surveys are being integrated into the model currently.

This paper discusses the experience with the outcome of the program so far: the benefits already realized as well as the challenges that are being faced and for which solutions are being developed. This massive initiative entailed an important cultural shift across all areas involved in economic statistics production: roles and responsibilities changed and complex interdependencies were created. The paper will examine the solutions put in place in terms of governance, training and human resource management.

---

<sup>1</sup> The surveys cover manufacturing, retail, wholesale and services sectors as well as capital expenditure activities.

## 2. Background and Overview of the IBSP

### **2.1 Building on a Solid Base**

The precepts of the IBSP model are not new at Statistics Canada. They are based on the Unified Enterprise Survey<sup>2</sup> (UES program Brodeur et al. 2006) which was implemented 20 years ago and in a sense pioneered the generic approach to business statistics production. Under the UES 58 surveys were integrated, to a certain extent, with respect to content, collection and data processing. The UES experience proved that an integrated and consistent model for producing business statistics is not only possible but it is much more efficient than stand-alone customized solutions. The IBSP is a scaled up version of its precursor, both in terms of pushing the level of integration, as well as in the number of statistical programs in its scope.

### **2.2 IBSP Pillars**

The model was developed based on a number of fundamental principles which addressed the corporate objectives of the Corporate Business Architecture outlined in the Introduction. The principles are:

- The Business Register (BR)) is the common frame for all economic programs
- Electronic collection is the primary (but not unique) mode of data collection
- Use of common processing methods and systems
- Maximize use of administrative data either as replacement for collected information, as additional information used for data validation or auxiliary data for estimation.

Coherence among the various economic programs was an important quality objective as well as ease of cross-training of staff. In order to achieve the ultimate objectives as described in the introduction, as well as abiding by the clearly-stated fundamental principles, the economic programs using the IBSP model have in common the following:

1. A harmonized content, where common variables have the same definitions although may have different names in different questionnaires in order to address respondents' needs. As programs are integrated into the IBSP, their content is reviewed. This provides the opportunity to update it in order to maintain relevance but also to ensure coherence with other business programs (e.g. operating revenue has the same definition across all business surveys as well as the same mnemonic for data processing). As part of this exercise nearly thirty thousand variables have been reviewed and assigned to a concept, properly named and numbered.
2. All IBSP programs are hooked up to the BR which contains information for industry as well as activity surveys. It is the responsibility of each program to ensure that all information relevant to the frame is fed into the BR on a timely basis: this includes feedback from collection as well as from other sources such as administrative information and research carried out by analysts.
3. Electronic questionnaires are used as the primary collection mode; paper and telephone collection is also still being offered but the aim is to reduce over time the reliance on these options, as the take-up rates of the electronic collection go up. All IBSP electronic questionnaires are modular: 70 or so modules have been

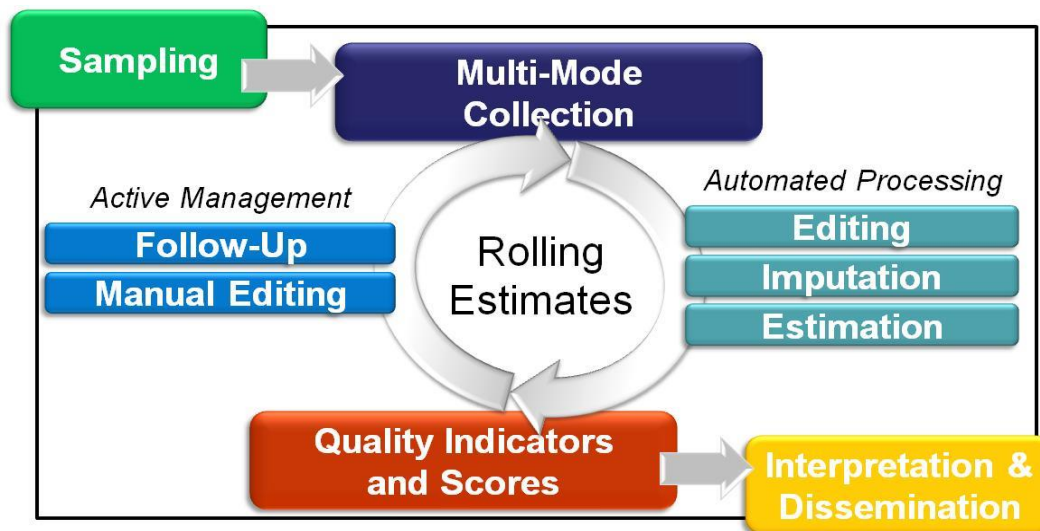
---

<sup>2</sup> The UES was a standardized approach for processing data from collection and administrative sources for 58 annual surveys covering the manufacturing, retail, services and aquaculture sectors.

developed and can be reused when assembling the collection vehicle for a given survey. They can be included on a frequency basis to some or all surveys

4. The processing systems are based on generic modules for sampling, edit and imputation, and estimation. Each one of these modules contain the different types of methods that would normally be used to process the data and programs select those that are appropriate for their needs. A generic system has also been developed for applying the confidentiality mask, however more work is required to render it more robust and efficient.
5. The most important change that was introduced with the IBSP was the iterative production model, named the Rolling Estimates (RE). Previously, data processing started only after collection was complete and analysts made manual interventions after each processing step (edit, imputation, allocation of national estimates to regions and estimation). This also encouraged a lot of microdata analysis and correction. Based on the RE model, estimates are produced iteratively as data is still being collected and analysts start by looking at the macro data, only delving into detail when necessary. **Figure 1** provides an illustration of the RE. Quality indicators and scores are employed to actively manage collection as well as to assess domain estimates.
6. A set of analytical tools as well as a common database serve all IBSP programs. This allows analysts to have easy access to data for all the economic programs, not only the one for which they are responsible thereby making their job a lot more effective. Using common tools will also ease movement of staff from one program to another as less training will be required.

**Figure 1. Overview of the IBSP Rolling Estimates Approach**



### 3. Results to date from the IBSP programs

The IBSP model was applied first to 60 annual surveys covering the manufacturing, retail, wholesale services and capital expenditure in January 2014. These programs were chosen as they, with the exception of the capital expenditure survey, had been part of the UES and therefore already subjected to a high degree of integration and standardization. They were seen as the low hanging fruit, with the more heterogeneous programs slated for later

integration. Although this proved to be the case, there was a need for some adjustments and additions to the systems and processes put in place as subsequent programs were integrated into IBSP. It could be argued that starting with more complex programs allows for a more exhaustive set of requirements, leading to the development of systems and processes that meet a wider set of needs.

### 3.1 Break in Series

Given the fundamental changes (methodology, content, collection tools, process, systems, etc.) being introduced with the IBSP, it was recognized from the onset that this will very likely result in breaks in series. A number of potential avenues were explored:

- Use a parallel run to estimate the impact and adjust the series
- Terminate existing series and start anew
- Wait until two years of IBSP data are obtained and use the new level to back cast the series

The option of doing a parallel run was not possible for the annual programs due to the prohibitive costs required to maintain two sets of systems. It was also decided that series continuity was paramount to users, especially the System of National Accounts. Therefore, for the first wave of annual programs, the strategy was to communicate to users the changes and their likely impact on the series and to warn them about making year over year comparisons until two years of data are obtained and adjustments can be applied to the old series. This strategy worked fairly well as only the usual amount and types of questions were received by subject-matter areas after the release of the first year of estimates.

It should be noted that some detailed-level estimates were held back from publication in the first year as analysts were not able to satisfactorily vet them. Once the second year of data were obtained they were able to ascertain the impact of the IBSP and make appropriate adjustments before disseminating them. A fairly rigorous set of vetting procedures were developed by subject-matter areas to vet the IBSP data (see Appendix A) and identify breaks in series. This was important since analysts were working with new analytical tools and they had to familiarize themselves with a new set of methods that were introduced such as calibrating surveys estimates to tax data. These procedures also ensured process uniformity and a very high degree of scrutiny.

**Table 1** shows where the IBSP was deemed to have caused breaks in series for the Services Industries programs. In most of these instances the breaks in series were caused by a change in coverage and by the introduction of the calibration method for estimation. Some of these results have led to the need for adjustments to the calibration method as it was deemed to provide unexplainable results in some isolated instances.

Of course a parallel run is the best method for assessing if the introduction of new methods or processes are responsible for breaks in series and for deriving an adjustment factor when necessary. The parallel run became an option of convenience, rather than by design, for some of the monthly programs being migrated to the IBSP. Because the first set of annual IBSP programs experienced significant delays in publishing (up to 8 months) due to many difficulties, a mitigation strategy had to be devised for monthly surveys which have fairly tight production schedules and for which significant delays are not acceptable. Therefore, it was decided that when content changes were not significant the old as well as the IBSP processing systems were going to be run in parallel. This not only mitigated the risk of delays but it also allowed subject-matter to assess the impact of the new model on the estimates. In most cases the parallel runs were short lived, only two to three months. The disadvantage was that staff took much longer to buy into the IBSP model and fully commit

to it as they had the old production model to fall back on. This is not to be underestimated since, as it will be explained in the next section, culture change was and still is the most significant challenge of the IBSP implementation.

**Table 1. Impact of the IBSP on Services Industries Series**

| Surveys Released                                                         | Break or No Break |    |
|--------------------------------------------------------------------------|-------------------|----|
|                                                                          | Yes               | No |
| <b>Accounting Services</b>                                               |                   | ✓  |
| Advertising & Related Services                                           | ✓                 |    |
| Amusement and recreation                                                 | ✓                 |    |
| Architectural Services                                                   | ✓                 |    |
| Automotive equipment                                                     |                   | ✓  |
| Commercial and industrial machinery                                      |                   | ✓  |
| Consulting Services                                                      | ✓                 |    |
| Consumer goods rental industry                                           | ✓                 |    |
| Employment service                                                       | ✓                 |    |
| Engineering services                                                     | ✓                 |    |
| <b>Film and video distribution</b>                                       |                   | ✓  |
| Film, television and video post-production                               |                   | ✓  |
| Film, television and video production                                    |                   | ✓  |
| Food Services & Drinking Places                                          | ✓                 |    |
| Head Office                                                              | ✓                 |    |
| Periodical publishing                                                    |                   | ✓  |
| Real Estate agents, brokers, appraisers and other real estate activities | ✓                 |    |
| Real Estate Rental & Leasing & Property Management                       | ✓                 |    |
| Repair & Maintenance Services                                            | ✓                 |    |
| Software Development & Computer Services                                 | ✓                 |    |
| Sound recording and music publishing                                     | ✓                 |    |
| Specialized Design                                                       | ✓                 |    |
| Spectator Sports, Event Promoters, Artists & Related Industries          | ✓                 |    |
| Surveying & Mapping                                                      | ✓                 |    |
| Travel arrangement services                                              | ✓                 |    |
| Traveller Accommodation                                                  | ✓                 |    |

### 3.2 Managing Changes in Roles and Responsibilities

As mentioned in the introduction, the IBSP was part of a much wider corporate initiative at Statistics Canada aimed at achieving efficiencies. Key to this corporate initiative was the centralization of statistical services (collection, IT, dissemination, and data processing) across the organization. The idea was that this will allow for better standardization in terms of quality for the provision of these services and will allow analysts to focus on what they do best, vetting the data, liaising with users and developing new products to ensure that programs stay relevant. The transition to these centralized services was made simultaneously with the transition to the IBSP for the first wave of annual surveys. As it can be seen from comparing Figure 2 and Figure 3 Below. This shift caused a significant change in the roles and responsibilities of subject-matter analysts. They went from performing most, and in some cases all survey processes, to focusing mostly on the book-end of the process: specification of requirements and validation and dissemination of output.

**Figure 2. Before IBSP: subject-matter responsible for most processes**

| Quality management          |                                        |                                           |                                |                             |                               |                                          |                          |
|-----------------------------|----------------------------------------|-------------------------------------------|--------------------------------|-----------------------------|-------------------------------|------------------------------------------|--------------------------|
| Metadata management         |                                        |                                           |                                |                             |                               |                                          |                          |
| Specify needs               | Design                                 | Build                                     | Collect                        | Process                     | Analyse                       | Disseminate                              | Evaluate                 |
| Identify needs              | Design outputs                         | Build collection instrument               | Create frame and select sample | Integrate data              | Prepare draft output          | Update output systems                    | Gather evaluation inputs |
| Consult and confirm needs   | Design variable descriptions           | Build or enhance process components       | Set up collection              | Classify and code           | Validate output               | Produce dissemination products           | Conduct evaluation       |
| Establish objectives output | Design collection                      | Build or enhance dissemination components | Run collection                 | Review and validate         | Interpret and explain outputs | Manage release of dissemination products | Agree on action plans    |
| Identify concepts           | Design sampling frame and sample       | Configure workflows                       | Finalise collection            | Edit & impute               | Apply disclosure control      | Promote dissemination products           |                          |
| Check data availability     | Design processing and analysis         | Test production system                    |                                | Derive new variable & units | Finalise outputs              | Manage user support                      |                          |
| Prepare business case       | Design production systems and workflow | Test statistical business process         |                                | Calculate weights           |                               |                                          |                          |
|                             |                                        | Finalise production systems               |                                | Calculate aggregates        |                               |                                          |                          |
|                             |                                        |                                           |                                | Finalise data files         |                               |                                          |                          |
| Center of expertise         |                                        |                                           |                                |                             |                               |                                          |                          |
| Subject-Matter Division     | Methodology Service                    | Collection Services                       | IT Service                     | Dissemination Service       |                               |                                          |                          |

**Figure 3. Specialized Service Providers are Core to the IBSP Model**

| Quality management          |                                        |                                           |                                |                             |                               |                                          |                          |
|-----------------------------|----------------------------------------|-------------------------------------------|--------------------------------|-----------------------------|-------------------------------|------------------------------------------|--------------------------|
| Metadata management         |                                        |                                           |                                |                             |                               |                                          |                          |
| Specify needs               | Design                                 | Build                                     | Collect                        | Process                     | Analyse                       | Disseminate                              | Evaluate                 |
| Identify needs              | Design outputs                         | Build collection instrument               | Create frame and select sample | Integrate data              | Prepare draft output          | Update output systems                    | Gather evaluation inputs |
| Consult and confirm needs   | Design variable descriptions           | Build or enhance process components       | Set up collection              | Classify and code           | Validate output               | Produce dissemination products           | Conduct evaluation       |
| Establish objectives output | Design collection                      | Build or enhance dissemination components | Run collection                 | Review and validate         | Interpret and explain outputs | Manage release of dissemination products | Agree on action plans    |
| Identify concepts           | Design sampling frame and sample       | Configure workflows                       | Finalise collection            | Edit & Impute               | Apply disclosure control      | Promote dissemination products           |                          |
| Check data availability     | Design processing and analysis         | Test production system                    |                                | Derive new variable & units | Finalise outputs              | Manage user support                      |                          |
| Prepare business case       | Design production systems and workflow | Test statistical business process         |                                | Calculate weights           |                               |                                          |                          |
|                             |                                        | Finalise production systems               |                                | Calculate aggregates        |                               |                                          |                          |
|                             |                                        |                                           |                                | Finalise data files         |                               |                                          |                          |
| Center of expertise         |                                        |                                           |                                |                             |                               |                                          | Shared responsibilities  |
| Subject-Matter Division     | Metadata Service                       | Content Specifications Service            | Methodology Service            | Processing Service          | Collection Services           | IT Service                               | Dissemination Service    |

This required a fairly complex and in some cases delicate strategy of communication and support from management and training and re-organization for those impacted. In some cases staff had to be re-assigned to different duties or moved to a different part of the organization where their skills could be used. For example staff responsible mostly for data processing who were not good candidates to take on analytical responsibilities were encouraged to move to the area where data production was centralized. This was more complex in the case of staff doing collection as collection was centralized in regional offices away from Ottawa. The transition was also difficult for subject-matter analysts who were generally implicated in data processing as well. Subject-matter analysts were directed to spend all their energies on data validation and dissemination as well as on producing more analytical output. A training and development path was developed to support these objectives. This push dove-tailed into an updated set of priorities for the Economic Statistics Field, the organization within Statistics Canada responsible for all economic programs: last fall senior management met for a day to discuss and put in motion activities that will increase the analytical capacity and output of the Economic Statistics Field.

The transition is still a work in progress as it will take some time before the organization fully re-aligns the skill set of the staff with its objectives.

### 3.3 Governance

Over 70 programs are now using the IBSP model and many more are in the process of being moved over. In the next few years Statistics Canada will have over 100 programs using the IBSP. Given the sheer size of this endeavour, as well as the fact that the IBSP is



based on a service provider model, the governance structure is paramount to ensuring effective and efficient operations. Good governance has to oversee the following:

- Ongoing production with its many interdependencies
- Periodically re-assess the transition schedule of programs slated to migrate to the IBSP
- Prioritize development work needed to improve the existing model
- Prioritize work required to maintain programs relevant (content development, sample selection, etc.)

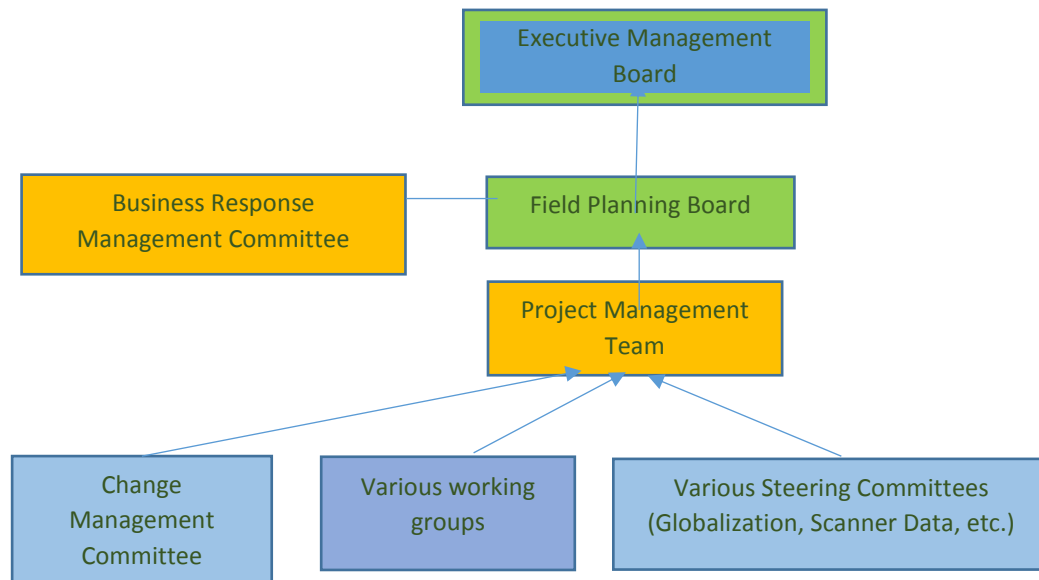
Statistics Canada, in general, relies heavily on committees for governance and decisions making. The IBSP governance is made up of some committees that were already in operation at Statistics Canada and a few specially created for the purpose. The important point to note is that the governance structure ensures that the IBSP operates within the larger Statistics Canada framework and that decisions regarding the program take into consideration overall priorities and constraints.

The Field Planning Board chaired by the Assistant Chief Statistician of the Economic Statistics Field approves the overall priorities for the economic statistics programs and is the ultimate decision maker for the IBSP unless there are fairly important budget and strategic issues that need to be approved by the Executive Management Board which is chaired by the Chief Statistician. The Field Planning Board receives input from the Project Management Team (PMT) and the Business Response Management Committee (BRMC). The BRMC reviews and approves all content decisions with a view of minimizing response burden while the PMT oversees all other aspects of the IBSP. As the BRMC is a corporate committee, all IBSP content is scrutinized in the context of Statistics Canada content.

Various working groups are tasked with specific issues, the Change Management Committee and a number of project Steering Committees provide input to the PMT. Steering Committees provide input and recommendations on new content, updates to standard classifications, new sources of data as well as the production process (collection, methodology and IT). The Change Management Committee is key to the smooth functioning of the program: the Committee is responsible for investigating the validity of a request, implications (resource, time, quality, etc.) of implementing the change, approval or refusal of the request and prioritization of implementation. Since the number of change requests has become extremely high, an additional level of scrutiny has been added: before submitting a request to the Change Management Committee, the request must be reviewed and approved by the director responsible for the program submitting the change. The rationale being that requests that could adversely affect the global priorities of the Agency can be filtered by the management as Directors have a deep knowledge of corporate priorities.

The integration of so many economic programs into the IBSP and the established governance have rendered the decision making process around these programs quicker and better informed. This is because all IBSP programs use the same infrastructure (collection, data processing and dissemination), the same basic methods for data processing and a common frame and harmonized content. While in the past decisions were made sometimes in silos or sometimes without fully appreciating their full impact, now each decision is made in the context of the entire economic statistics program. This advantage has already yielded a number of important achievements which are described in the next section.

**Figure 4. IBSP Governance**

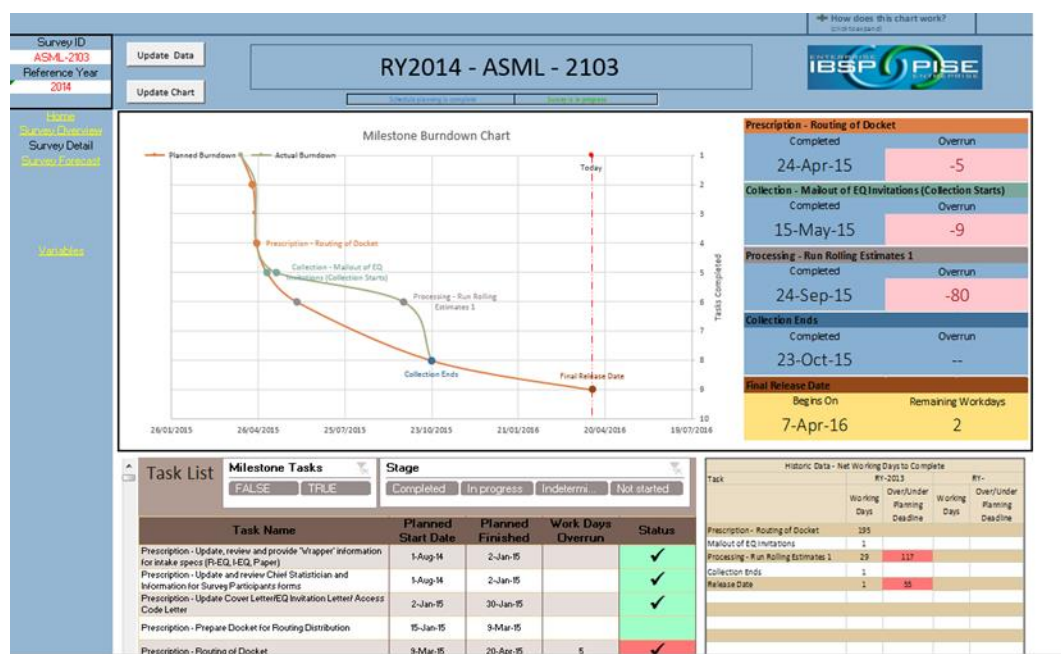


The process of managing the ongoing production work had to be re-thought given the sheer volume of programs that are currently in the IBSP. In the past most of the programs were managed separately by a manager who looked after the process from start to finish. With the centralization of the different survey processes in specific areas the need for a program management function became very important. One area was identified as being best placed to manage the planning as well as to monitor the execution of the production process. This is the area responsible for processing the data as they are received from collection and making the estimates available to subject-matter analysts and the Systems of National Accounts. Reports which track the Actual versus the Planned completion dates were developed for each IBSP program and they can be generated on demand. Figure 4 below illustrates one such high level report. Whenever required, a survey manager can drill down into each process to the task or the sub-task level. These report permit a quick assessment of the ongoing work and allow for effective decisions when mitigating strategies are required. They also allow the PMT to re-assign priorities while understanding all the impacts if there is an urgent need for a certain deliverable.

### **3.4 Early Successes of the IBSP**

Even though the program is only two years old, there have been already a number of benefits realized. One of the first was derived from the electronic questionnaire collection: it proved so successful with the first wave of annual surveys that last fall the PMT decided to eliminate paper collection for 18 surveys. Their electronic collection take up rates are detailed in Table 2 below. This has translated into important savings as well as a reduced risk of error. It is hoped that this year more surveys can move to 100% electronic collection.

**Figure 4. Example of Survey Production Report**



**Table 2. Electronic Collection Take up Rates**

| DIVISION | SURVEY                                | EQ TAKEUP RATE 2013 | EQ TAKEUP RATE 2014 | REPORT DATE |
|----------|---------------------------------------|---------------------|---------------------|-------------|
| RSID     | Advertising                           | 88.73               | 86.95               | 04-Jan-16   |
| RSID     | Architectural                         | 87.52               | 91.06               | 04-Jan-16   |
| RSID     | Automotive Equipment Rental           | 94.33               | 87.5                | 04-Jan-16   |
| RSID     | Consulting                            | 91.59               | 90.76               | 04-Jan-16   |
| RSID     | Consumer Goods Rental                 | 86.23               | 86.18               | 04-Jan-16   |
| RSID     | Database Publishers                   | 79.23               | 82.46               | 04-Jan-16   |
| RSID     | Employment                            | 88.75               | 92.27               | 04-Jan-16   |
| RSID     | Engineering                           | 88.25               | 89.18               | 04-Jan-16   |
| RSID     | Film Television Video Post-production | 83.62               |                     | 04-Jan-16   |
| RSID     | Film Television Video Production      | 87.03               |                     | 04-Jan-16   |
| RSID     | Film Video Distribution               | 78.43               |                     | 04-Jan-16   |
| RSID     | Machinery Equipment Rental            | 87.5                | 85.89               | 04-Jan-16   |
| RSID     | Real Estate Brokers                   | 89.19               | 91.43               | 04-Jan-16   |
| RSID     | Sound Recording Music Publishing      | 85.71               |                     | 04-Jan-16   |
| RSID     | Specialized Design                    | 87.23               | 83.25               | 04-Jan-16   |
| RSID     | Spectator Sports                      | 82                  | 84.03               | 04-Jan-16   |
| RSID     | Surveying Mapping                     | 86.54               | 89.05               | 04-Jan-16   |
| RSID     | Travel Arrangement                    | 89.01               | 87.05               | 04-Jan-16   |

A generic module on electronic commerce was added to the manufacturing, wholesale, retail and services surveys in time for the third annual production cycle. This was possible in a very short time and with minimal investment of resources because all these surveys use the same modular electronic questionnaire and the same processing systems and

methods. There was no need to consider different elements or issues for each individual survey given the standardization.

Providing additional commodity detail on a number of manufacturing industries, specifically those involved in logging and sawmills became an urgent priority late last year. In less than 3 months it was possible to test, add the content and adjust the processing systems for 23 manufacturing surveys. This process would have required a lot more investigation before a decision could be made as well as a lot more time to implement before the move to the IBSP.

Another example of benefiting from integration and standardization was the recent process of identifying the optimal date for moving to the next version of the industrial (NAICS 2017) and product (NAPCS 2017) classification systems. Since all IBSP surveys are NAICS 2012 and NAPCS 2012 compliant, the review of the impact in terms of disseminated output was isolated to a review of the NAICS and NAPCS changes between 2012 and 2017. As well, from an operational point of view implementing the changes is similar across all the programs impacted, therefore one plan fits all. The review and decision process was extremely quick as was the development of the implementation schedule.

One of the important objectives of the IBSP is to streamline the development process required when making changes to existing economic programs or initiating new ones. The move to using the electronic questionnaires as the primary collection instrument meant a lot of work in the area of developing these questionnaires, testing them and putting them into production. The questionnaires are modular and metadata driven. They are also connected to the Business Register. This connectivity makes the development process rather complex and the learning curve steep. The IBSP has set up a development and production team responsible for this work. Having acquired a lot of experience in this area the team is now able to quickly develop new questionnaires or to make adjustments to existing ones as needed. Their experience and know-how is important when conceiving new questionnaires as they are able to re-apply approaches that were successful while avoiding to repeat past errors. Given that the use of electronic questionnaires in economic surveys at Statistics Canada is fairly new, the fact that in just a few years the IBSP has put in production over 70 of them it shows the success of centralizing their development. In fact, a recent review of the process of developing electronic questionnaires at Statistics Canada has recommended that the social surveys adopt the same centralized approach as the IBSP.

#### 4. Lessons Learned

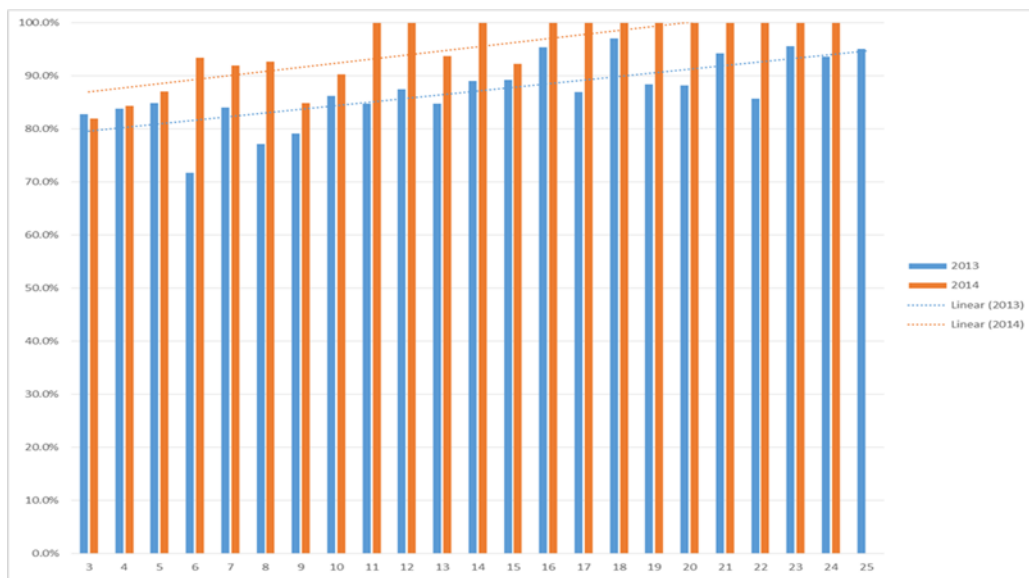
Two years into the new model have provided a wealth of information in terms of what worked well and what needs adjustment or better planning moving forward. Frequent consultations with users and service providers have been conducted to ensure that issues are identified and solutions developed and implemented as quickly as possible. This is crucial to ensure an efficient process but also to get buy in from all partners involved.

##### **4.1 Transition Work Cannot be Underestimated**

Fairly quickly into the transition process it became clear that the required effort on the part of client areas (surveys or admin programs being integrated) had been underestimated

when the project was planned. This was due to two major reasons. First, seasoned employees were required to provide specifications and they were also the same people delivering the ongoing programs. Having junior staff work on specifying needs proved to be less than optimal. The second reason was the longer than expected time and effort to bring the newly integrated programs into a regular production cycle with the expected level of quality and timeliness. Much more effort than initially estimated was required to handle numerous technical and methodological issues as each program went through the first year of production. Figure 3 shows the average (for 60 programs) success rate, defined as the Rolling estimate having completed and produced the expected output for reference years 2013 and 2014. For 2013 it took in some cases up to 30 iterations before achieving a successful one while in 2014 it was reduced to a third of that. Even though, as described in the next two sections, measures have been taken to improve the success rate, it does take initially more resources to get a program going than it will take to support regular production. Based on this we currently plan for an augmentation in resources needed for the first cycle of production and then a reduction to below what was required previous to IBSP, once the program is stable.

**Figure 3. Average Success Rate by Rolling Estimate Iteration**



#### 4.2 Test Now or Pay the Price Later

One of the most important lessons was the value of integrated testing. Because the development timelines for the project were fairly tight, a lot of testing was carried out on each component of the system but not necessarily on the entire process. As well representative surveys were selected for testing rather than systematically testing each program. This strategy worked for many of the less complex surveys but led to additional unplanned work for the others.

The testing strategy has now been adjusted and each program slated for integration is now fully tested before going into production. An important element of the current testing strategy is volume testing as it was determined that in many instances, the source of error was due to insufficient processing capacity.

#### **4.3 There needs to be a balance between complexity and quality**

The IBSP introduced a number of methodological changes geared towards improving data quality. In some cases these methods had never been used in economic surveys at Statistics Canada. Some examples were the calculation of variance due to imputation, the replacement of the Simple Random Sampling with Bernoulli sampling, two phase sampling and calibration to tax estimates. Testing and studies of these methods and approaches were conducted previous to implementation to assess their impact on data quality. However, once the first set of estimates were obtained it became quickly apparent that it was challenging for analysts to understand how these methods were applied and the estimates calculated. Additional training had to be provided. As well, in some cases, the methods had to be slightly modified in order to ensure more stable, year over year estimates. When implementing new methods it is important to assess the expected marginal increase in quality versus the additional work and expertise required to use these methods. This is difficult to undertake but it needs to be factored into the decision making process.

#### **4.4 One Size does not Always Fit All**

The IBSP is premised on the utilization of generic systems, methods and processes to produce economic statistics in a much more efficient way. Of course this approach assumes that most economic programs are fairly similar. As we progress in the adoption of the IBSP for more and more economic programs we realize that there will be some instances where a generic solution will not work or will not give the desired outcome. So far each program that has been integrated into the IBSP required some adaptation of the generic process to ensure that requirements are met. However, we are now encountering instances where more than minor tweaking are necessary. The question that is now posed and needs to be answered is: to what degree is it worth adjusting what is already in place in order to accommodate some of the programs that do not fit the developed model well? To answer this question the benefits as well as the risks and draw backs have to be weighed. It will most likely turn out that the model is not suitable in some instances and that stand alone solutions or hybrid solutions where only certain components of the IBSP are used, will have to be put in place.

#### **5.0 Conclusion**

The last two years of IBSP production have proved both challenging and rewarding. The model works well for over 70 surveys already but a lot of additional unplanned work was required to get it to that point. The flexibility of the modular approach as well as the standardization of processes, content and methods have already yielded early benefits when the programs had to be adapted to respond to user needs.

There is still a lot of work remaining to be done to improve the performance and robustness of the system infrastructure but a lot of progress has been made towards that end. Integrating and processing unstructured and transactional data sets is a new challenge facing the program. Work is currently in progress on a corporate solution.

Recognizing that the model works well for a vast number of programs but potentially not for all is an important step to ensuring that the right decisions are made with respect to its adoption across the economic programs at Statistics Canada.

#### Data Validation Measures Checklist .

| Data Validation Measures Checklist                                                                                                           |          |    |      |
|----------------------------------------------------------------------------------------------------------------------------------------------|----------|----|------|
| Survey title:                                                                                                                                |          |    |      |
| Reference period:                                                                                                                            |          |    |      |
| Release date:                                                                                                                                |          |    |      |
| Initial the steps that have been undertaken for data validation of survey estimates.                                                         |          |    |      |
| Steps                                                                                                                                        | Initials |    | Comm |
| Key DV steps                                                                                                                                 | SMA      | UH |      |
| Top contributors reviewed, both weighted and unweighted ( <i>specify</i> )                                                                   |          |    |      |
| Used "Generalized" SAS <b>Macro</b> program to compare <u>variable relationships</u> (ratios) and identify problematic domains               |          |    |      |
| Used "Generalized" SAS <b>Macro</b> program to compare variables <u>over time</u> (growth rates) and identify problematic domains            |          |    |      |
| Identified and reviewed problematic domains using the pre-defined 5pp threshold                                                              |          |    |      |
| Identified and reviewed problematic domains using a user-defined (subject-matter specific) threshold                                         |          |    |      |
| Used alternative programs to compare and drill-down ( <i>specify</i> )                                                                       |          |    |      |
| If using alternative programs, estimates have been verified against official IBSP estimates                                                  |          |    |      |
| Used "Generalized" SAS <b>Micro</b> -outlier program to identify outlier records for pre-defined ratios for <b>key</b> variables             |          |    |      |
| Used "Generalized" SAS <b>Micro</b> -outlier program to identify outlier records for survey-specific ratios for <b>other characteristics</b> |          |    |      |
| <b>Preparation for DV meetings</b>                                                                                                           |          |    |      |
| DV tables include at least 3 years of historical data for Canada and provinces                                                               |          |    |      |
| DV tables include Customized tax data from <u>generalized</u> SAS program or <u>alternative</u> program ( <i>circle applicable</i> )         |          |    |      |
| DV tables include BSMD forecast data from <u>generalized</u> SAS program or <u>alternative</u> program ( <i>circle applicable</i> )          |          |    |      |
| Used Logic-check program to verify coherence in data maintained after macro-adjustments                                                      |          |    |      |
| DV tables include common units analysis derived from <u>IBSP</u> program or <u>alternative</u> ( <i>specify</i> )                            |          |    |      |
| DV tables include common units data for <u>Canada</u> and <u>provinces</u> ( <i>circle applicable</i> )                                      |          |    |      |
| DV tables include other auxiliary data ( <i>specify</i> )                                                                                    |          |    |      |
| Confrontation with other similar sources of data external to Statcan ( <i>specify</i> )                                                      |          |    |      |
| Confrontation with other similar sources of data published by Statcan ( <i>specify</i> )                                                     |          |    |      |
| All published domains have been analyzed prior to DV management meeting                                                                      |          |    |      |
| <b>DV meetings</b>                                                                                                                           |          |    |      |
| Meeting(s) to review estimates has been held with Unit Head & Chief                                                                          |          |    |      |
| Meeting(s) to review estimates has been held with SNA counterpart                                                                            |          |    |      |
| Consultation with external stakeholders (where applicable)                                                                                   |          |    |      |
| Meeting(s) to review interim/final estimates has been held with Management team                                                              |          |    |      |
| Table 1 "key" variables validated                                                                                                            |          |    |      |
| Table 2 expense breakdown validated                                                                                                          |          |    |      |
| Table 3 client base validated                                                                                                                |          |    |      |
| Other previously-disseminated tables validated ( <i>specify</i> )                                                                            |          |    |      |
| Coherence of results in context of known economic events explained by analyst at                                                             |          |    |      |
| Response rates reviewed                                                                                                                      |          |    |      |
| Quality indicators reviewed                                                                                                                  |          |    |      |
| <b>Certification / Signatures</b>                                                                                                            |          |    |      |
| Analyst                                                                                                                                      |          |    |      |
| Unit Head                                                                                                                                    |          |    |      |
| Chief                                                                                                                                        |          |    |      |