Rolling Estimates and the Dynamic Assessment of Quality: A Generic Approach and a Source of Efficiencies

EESW 2013 - Work Session on Process Design

Étienne Saint-Pierre and Serge Godbout, Statistics Canada
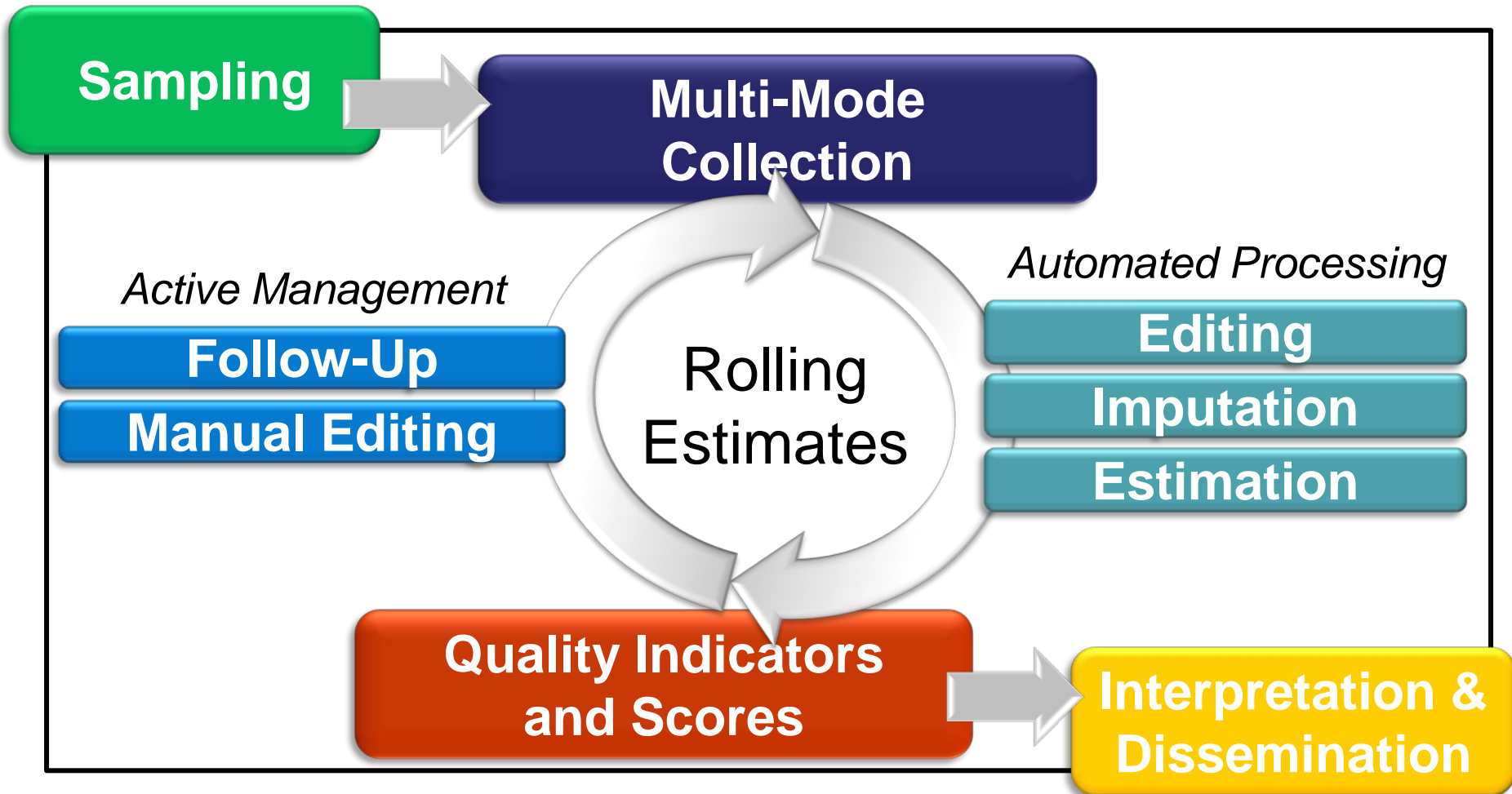
Nuremberg, Germany, September 9, 2013

Statistics Canada's Integrated Business Statistics Program

- Major transformational project for business surveys aimed at:
  - Enhancing quality assurance
  - Achieving efficiencies
  - Improving responsiveness in the delivery of new programs
- Covers all aspects of the business surveys
- 150 surveys in 10 different programs integrated by 2017
  - Annuals, sub-annuals, industry specific, economy wide
- Key feature to achieve the goals: Rolling Estimates

Overview of the IBSP Rolling Estimates Approach

Sampling

Multi-Mode Collection

Active Management

Follow-Up

Manual Editing

Rolling Estimates

Automated Processing

Editing

Imputation

Estimation

Quality Indicators and Scores

Interpretation & Dissemination

Active Management – Strategy Settings

- A subset of all Key Estimates is selected
- All Key Estimates are:
  - Ranked from the most to the least important
  - Weighted relatively using an importance factor
  - Assigned a Quality Target
    - Targets are set in line with the importance factor.
    - Active Collection ends for a Key Estimate when the Quality Indicator meets the Quality Target.

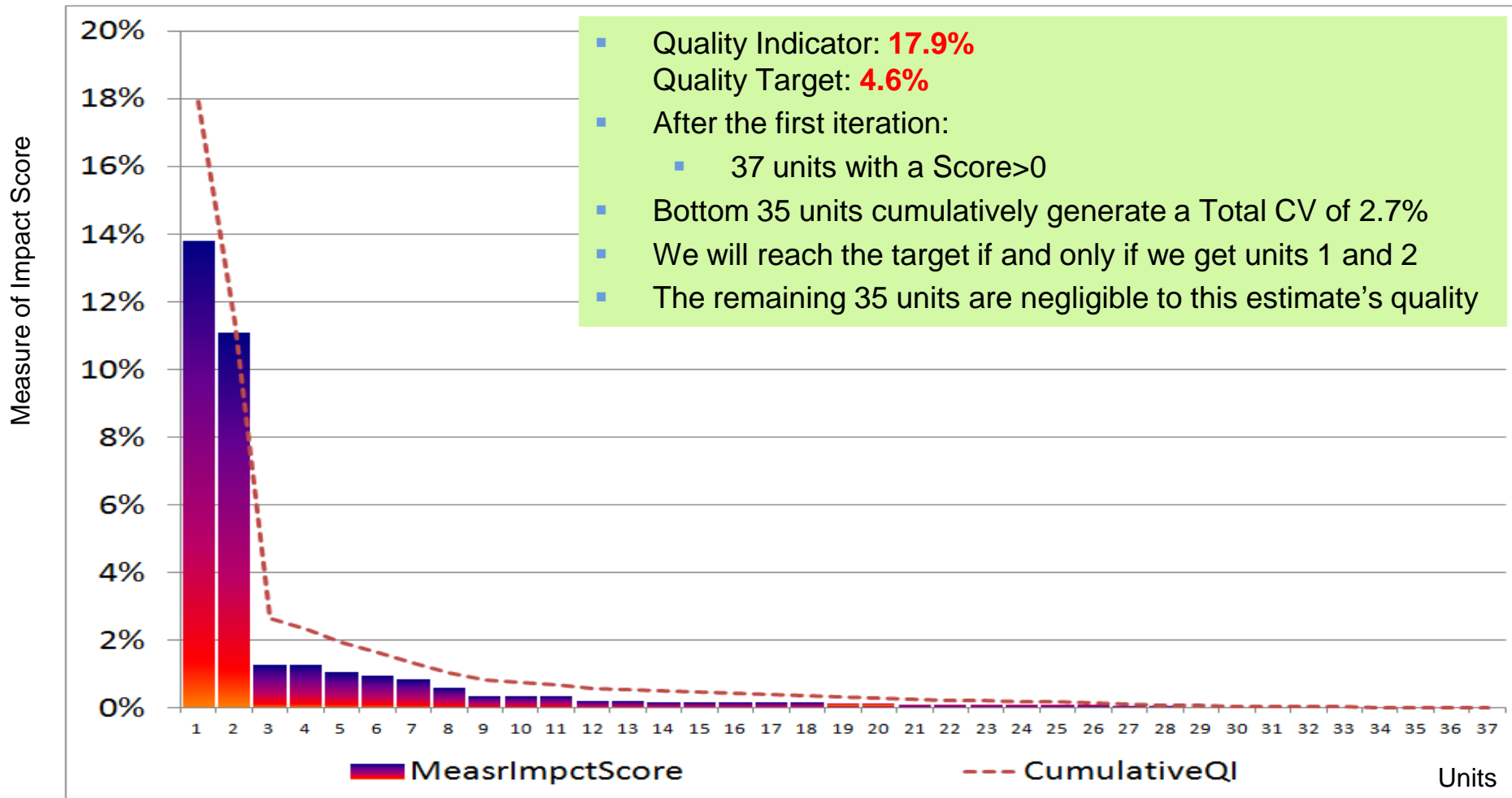- Active management and sampling strategies are coherent by design.

Active Management – Definitions

- Quality Indicator (QI)
  - QI= Sampling CV & Imputation CV & Pseudo Relative Bias

- Measure of Impact (MI) Score
  - Impact of a unit on the QI for a given estimate
  - Units imputed from a poor model or with reported/imputed values far from their predicted values will have high MIs.

Active Management– Example for a Key Estimate

- Quality Indicator: **17.9%**
  Quality Target: **4.6%**
- After the first iteration:
  - 37 units with a Score>0
- Bottom 35 units cumulatively generate a Total CV of 2.7%
- We will reach the target if and only if we get units 1 and 2
- The remaining 35 units are negligible to this estimate's quality

Empirical Study – RY2011 Prototype

- Parallel run for 47 Business Surveys
- Four Rolling Estimates iterations
- Total CV calculated for all key estimates (8,600) at each iteration

		Mthly Iteration 1 (July)	Mthly Iteration 2 (October)
Current Approach	Key Estimates with Quality Target met	76%	85%
	Number of units – Follow-up for non-response	13,200	27,900
IBSP Approach	Key Estimates with Quality Target met	N/A	98%
	Number of units – Follow-up for non-response	N/A	18,400

Statistics Canada • Statistique Canada

	Pre-IBSP	IBSP
Primary mode of collection	Paper, CATI	Electronic questionnaire
End of Collection Period:	Based on Target Dates and Response Rates	Based on Target Quality
Adaptive Collection Design:	Maximization of the weighted collection response rate	Minimization of the Total CV and Pseudo Bias
Collection Edits:	Could be applied on all variables on the questionnaire; Minimize outstanding cases with failed edits	Almost exclusively limited to key variables; resolved only for units with a high MI score
Processing:	Collection, Processing and Analysis executed sequentially	Collection, Processing and Analysis executed simultaneously, iteratively
Data Analysis	Conducive to micro-editing	Conducive to top down analysis

Questions from the EESW Scientific Committee

- Administrative data and the Active Management approach
  - The IBSP will maximize the use of administrative data.
    - An alternative to collected data
    - To impute non-response within a good imputation model
    - Considered in the Adaptive Collection Design
    - Used as a key component to produce the early estimates
- The Active Management design and the Social Statistics
  - Principles could be exported
    - Use adaptive collection design
    - Minimize non-response variance and bias as quality indicator
    - Early stage
  - Methods would have to be adapted to Social Statistics
    - CATI, Re-weighting, Non-Mandatory Survey

References

- **Godbout, Serge, Yanick Beaucage and Claude Turmelle (2011).** *Quality and Efficiency using a Top-Down Approach in the Canadian's Integrated Business Statistics Program. Conference of European Statisticians, Work Session on Statistical Data Editing. Ljubljana, Slovenia (9-11 May 2011).*

- **Mills, Fraser, Serge Godbout, Keven Bosa, Claude Turmelle (2013**), *Multivariate Selective Editing in the Integrated Business Statistics Program, Proceedings of the Joint Statistical Meeting (JSM) Conference. Montreal, Canada.*

- **Ravindra, Daniela (2012**). *The Integrated Business Statistics Program: An Overview. Proceedings of the ICES-IV Conference, Montreal, June 2012.*

- **Turmelle, Claude, Serge Godbout, Keven Bosa (2012**), *Methodological Challenges in the Development of Statistics Canada's New Integrated Business Statistics Program, Proceedings of the International Conference on Establishment Surveys (ICES) IV Conference. Montreal, Canada.*

- For more information, please contact:

- Pour plus d'information, veuillez contacter :

Étienne Saint-Pierre (613) 951-2025
Etienne.Saint-Pierre@statcan.gc.ca