

# A Tool for Managing Product Quality: the Case of Statistics Sweden

Paul Biemer

RTI International and  
University of North Carolina at Chapel Hill

Heather Bergdahl

Lilli Japiec

Åke Pettersson

Statistics Sweden

Dennis Trewin

Australian Statistician



# Acknowledgements

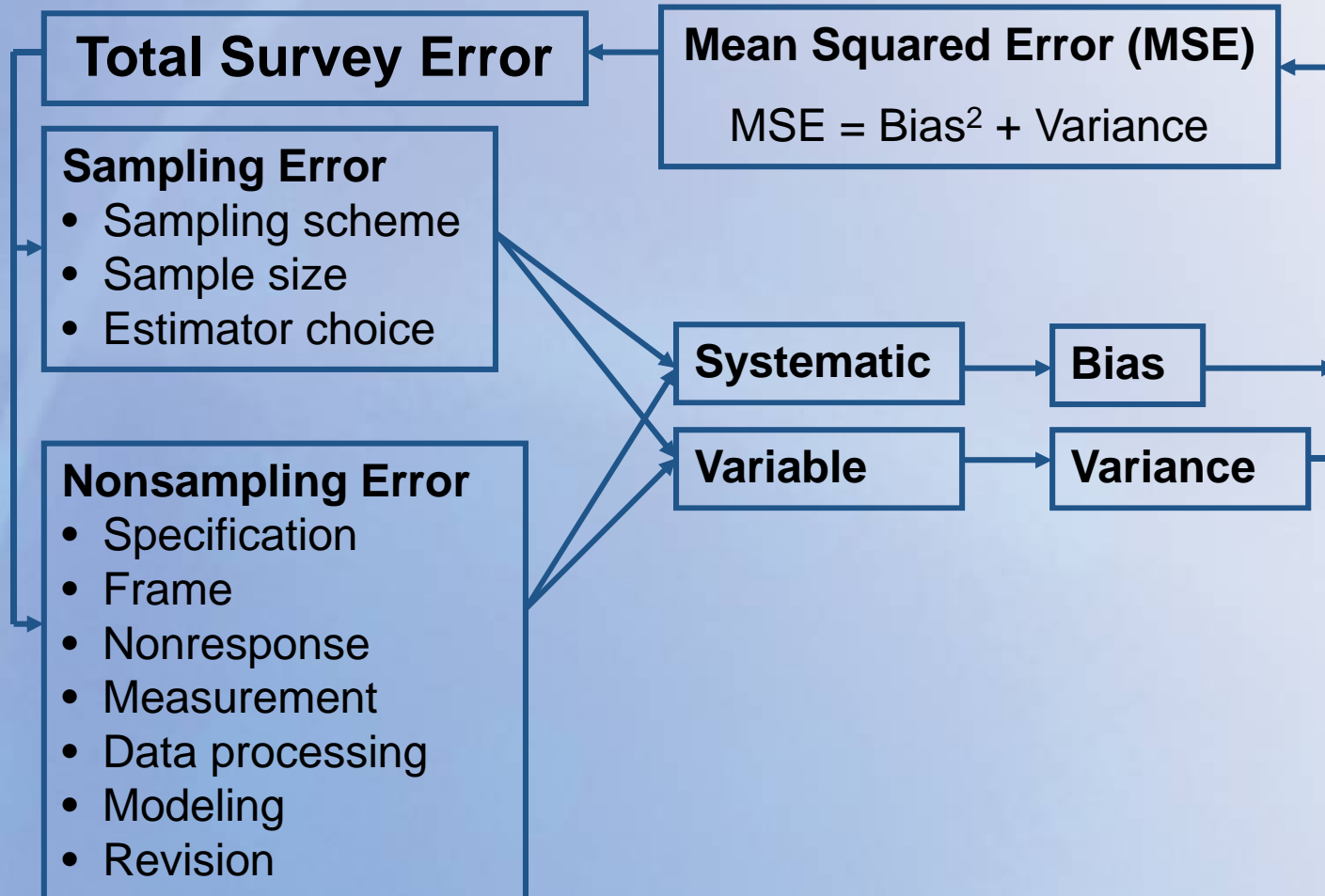
- This work would not have been possible without the assistance of many staff members at Statistics Sweden who commented on the work at various stages of its development

# A General System for Evaluating TSE

## The Case of Statistics Sweden

- Background
  - Need for a quality evaluation system and process  
Ministry of Finance to use results to monitor quality improvements over time
- Quality to be assessed for many surveys, registers, and programs within the SCB
- The process should be thorough, the reporting simple, and the results credible
- Requires external reviewers because internal review failed.
  - Paul Biemer and Dennis Trewin asked to develop and implement a quality evaluation system

# Total Survey Error Model



# Products to be Reviewed

Survey Products	Error Sources
Foreign Trade of Goods Survey (FTG)	Specification error
Labour Force Survey (LFS)	Frame error
Annual Municipal Accounts (RS)	Nonresponse error
Structural Business Survey (SBS)	Measurement error
	Data processing error
	Sampling error
	Model/estimation error
	Revision error

## Products to be Reviewed (cont'd)

Registers	Error Sources
Business Register (BR) Total Population Register (TPR)	Specification error Frame: Overcoverage Undercoverage Duplication Missing Data Content Error
Compilations	Error Sources
National Accounts (NA) Consumer Price Index (CPI)	Specification error Missing Data Content error Sampling error Model/estimation error Revision error

# Quality Criteria were Applied to Each Error Source

## Criteria by Error Source

1. Knowledge of risks
2. Communication with users
3. Compliance with standards and best practices
4. Available expertise
5. Achievement toward risks mitigation and/or improvement plans

## Ratings by Criterion

Poor ( ● )

Fair ( ◐ )

Good ( ○ )






Very Good ( ◑ )

Excellent ( ● )

## Risks to Data Quality by Error Source

High, Medium, Low

# An Example of the Rating Guidelines – Knowledge of Risks

Poor 	Fair 	Good 	Very Good 	Excellent 
Internal program documentation does not acknowledge the source of error as a potential factor for product accuracy.	Internal program documentation acknowledges error source as a potential factor in data quality.	Some work has been done to assess the potential impact of the error source on data quality.	Studies have estimated relevant bias and variance components associated with the error source and are well-documented.	There is an ongoing program of research to evaluate all the relevant MSE components associated with the error source and their implications for data analysis. The program is well-designed and appropriately focused, and provides the information required to address the risks from this error source.
	<b>But:</b> No or very little work has been done to assess these risks	<b>But:</b> Evaluations have only considered proxy measures (example, error rates) of the impact with no evaluations of MSE components	<b>But:</b> Studies have not explored the implications of the errors on various types of data analysis including subgroup, trend, and multivariate analyses	



# The Evaluation Process

- Pre-interview activities
  - Background reading by the two evaluators
  - Self-assessments by each program area
- The Quality Interview
  - ½ day sessions involving 4-5 key product owners
  - Overview discussions of product processes
  - Detailed assessment of each of the 5 criteria
- Post-interview activities
  - Review of and comment on ratings by product owners
  - Ratings adjustments by evaluators to achieve equity

# Example of Rating Results Structural Business Survey

Error Source	Average score	Knowledge of Risks	Communication to Users	Available Expertise	Compliance with standards & best practices	Plan towards mitigation of risks	Risk to data quality
Specification	46						M
Frame	62						M
Nonresponse	74						M
Measurement	50						H
Data proc.	52						H
Sampling	80						M
Model/est'n	60						H
Revision	58						H
Total score	59						

# Summary of Results for All Products

Error Source	RS	CPI	FTG	LFS	NA	SBS	BR	TPR	Avg
Specification	74	<b>68</b>	62	66	<b>56</b>	46	62	44	60
Frame	36	42	62	58		62			49
Overcov.							48	<b>52</b>	
Undercov.							40	34	
Duplication							46	64	
NR/Miss. data	62	36	62	<b>66</b>	64	74	40	<b>60</b>	57
Meas/Content	<b>52</b>	<b>40</b>	<b>54</b>	<b>50</b>	<b>58</b>	<b>50</b>	<b>42</b>	50	50
Data proc.	46	70	46	54	<b>44</b>	<b>52</b>			52
Sampling		<b>54</b>		72	<b>44</b>	80			64
Model/est'n	54	<b>64</b>	66	46	<b>44</b>	<b>60</b>			56
Revision	74		62		62	<b>58</b>			64
Total	57	56	59	58	51	59	45	52	55

<sup>11</sup> **Red Bold = High Risk**, **Black Bold = Medium Risk**, **No Bold = Low Risk**

# Strengths and Weaknesses of the Process

## Strengths

- Comprehensive approach
- Easily understood by management
- Identifies important areas to improve within and across products
- Can be updated periodically to assess improvement

# Strengths and Weaknesses of the Process

## Weaknesses

- Does not quantify actual total MSEs of products
- Can be somewhat subjective – more specificity in the rating guidelines needed
- Highly dependent on knowledge and skills of the external evaluators
- Requires thorough documentation of processes and improvements (for e.g., at the level of a quality profile)

## Future Work

- Work to address areas of high risk and less than “good” ratings
- Extend model to include other quality dimensions; in particular, relevance, timeliness, accessibility, and comparability
- Perform second assessment of 8 products in November – December, 2012
- Continue to evaluate approach, especially its ability to improve quality in high risk areas across products

# Priorities

Score		Number of assessed criteria for 8 tested products	Of which concern error sources with high risk
Excellent	⊙	11	4
Very good	◒	90	18
Good	◯	103	45
Fair	◐	62	28
Poor	●	9	0
<b>Total</b>		275	95

1. Areas with high risk and lower scores

2. Low hanging fruit

Error sources	Knowledge of risks	Communication to users
Dataprocessing error	◯	◐

3. Documentation

Statistics Sweden  
EV/AKM  
Elisabet Andersson  
Linda Wahlman

DESCRIPTION OF THE STATISTICS  
2011-12-14  
1(1)  
AM0401

**Labour Force Surveys**  
2011  
AM0401

*In this description, the general and legal data on the surveys are first presented, followed by the objectives and background of the data. The content and accuracy of the surveys are then described, followed by how the survey is carried out and how the results are made available. By clicking on a heading on the contents page, you can move directly to the relevant section.*

Table of contents

A Administrative information ..... 2

B Quality declaration ..... 6

  B 0 Introduction ..... 6

  B 1 Contents ..... 6

    1.1 Statistical target characteristics ..... 6

      1.1.1 Objects and population ..... 6

## Possible areas of discussion:

- This approach is quite labor intensive. Are there ways of simplifying it for evaluating products on a mass scale?
- What criteria seem appropriate for evaluating relevance?  
Accessibility?
- What displays can you suggest for showing year to year change in scores for a product by error source?