

Quality evaluation for a statistical register: the Italian FRAME-SBS

Session 21 - Administrative Data Systems

2.6.2016

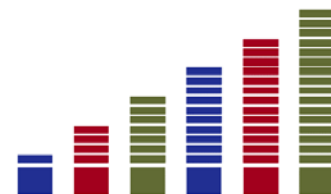
Orietta Luzi, Fabiana Rocci, Roberto Sanzo, Roberta Varriale, Giovanna Brancato

Italian National Statistical Institute – ISTAT

luzi@istat.it

Outline

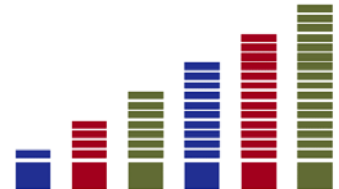
- Background
- The quality framework
- The quality indicators
- Conclusions and future work



Background

- In 2013 ISTAT has developed a **multi-source statistical register** for the annual estimation of economic accounts of Italian enterprises (Frame-SBS)
- Based on the **combined use of micro-data from admin and fiscal sources** (*Financial Statements, Sector Studies, Unico Model, IRAP, Social Security data*) and direct survey data
- Census-like data (**about 4,4 millions** enterprises) for the main SBS (*Revenues, Value Added, Wages, ...*) - Since reference year 2012 estimated based on the register

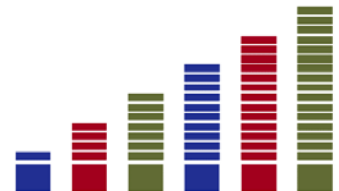
Luzi, O., et al. (2014), The new multiple-source system for Italian Structural Business Statistics based on administrative and survey data, European Conference on Quality in Official Statistics (Q2014).



Problem

Developing a **quality framework** in order to:

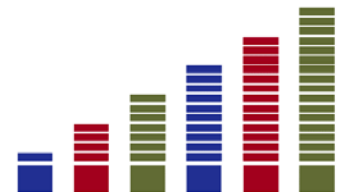
- assess, document and monitor the quality of all the components of the register production process (input data sources, data processing, and outputs)
- identify the sources of errors potentially affecting the data and the production process



The quality framework

Strategy

- 1) Mapping the life-cycle of the Frame-SBS production process
(Zhang L.C. (2012) *Topics of statistical theory for register-based statistics and data integration*)
 - Phase 1: Treatment of single-source micro-data w.r.t. admin targets
 - Phase 2a: Treatment of single-source micro-data w.r.t. SBS targets
 - Phase 2b: Treatment of integrated micro-data w.r.t. SBS targets
- 2) For each phase, identify the potential sources of error
- 3) For each potential source of error, develop a set of quality indicators

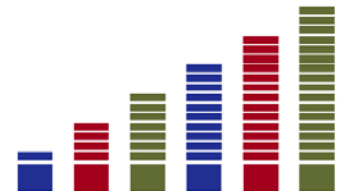
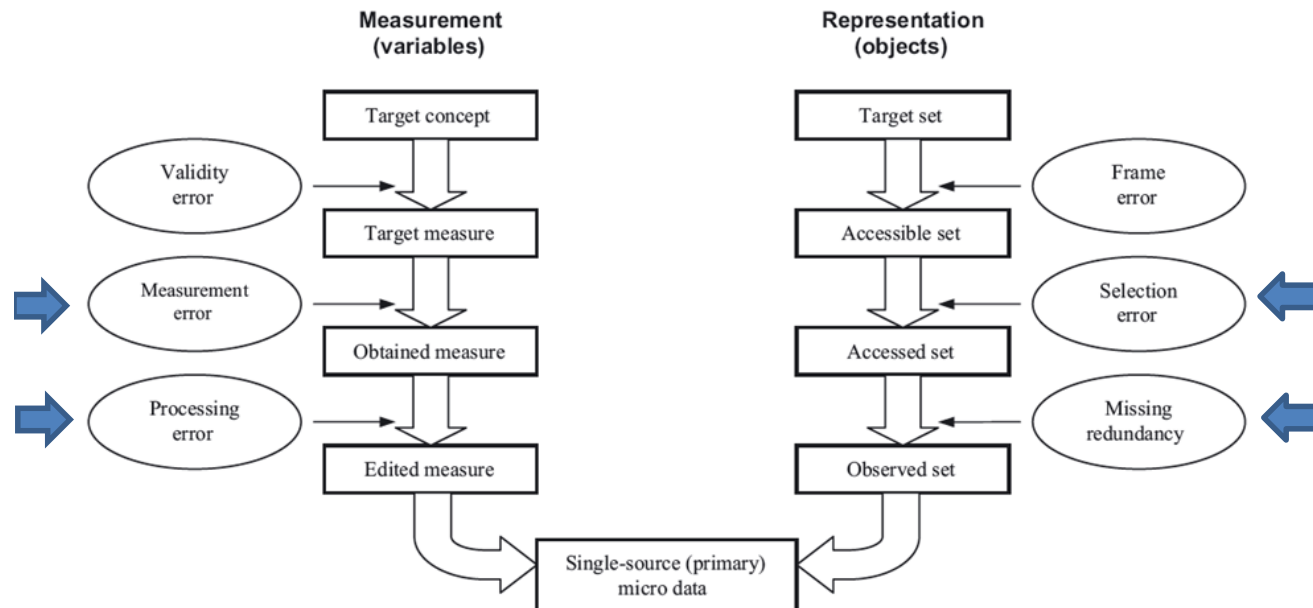


Identify potential sources of error

Zhang's two-phase framework for quality assessment

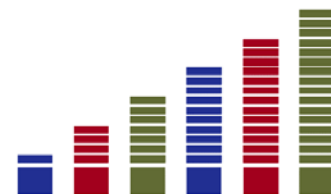
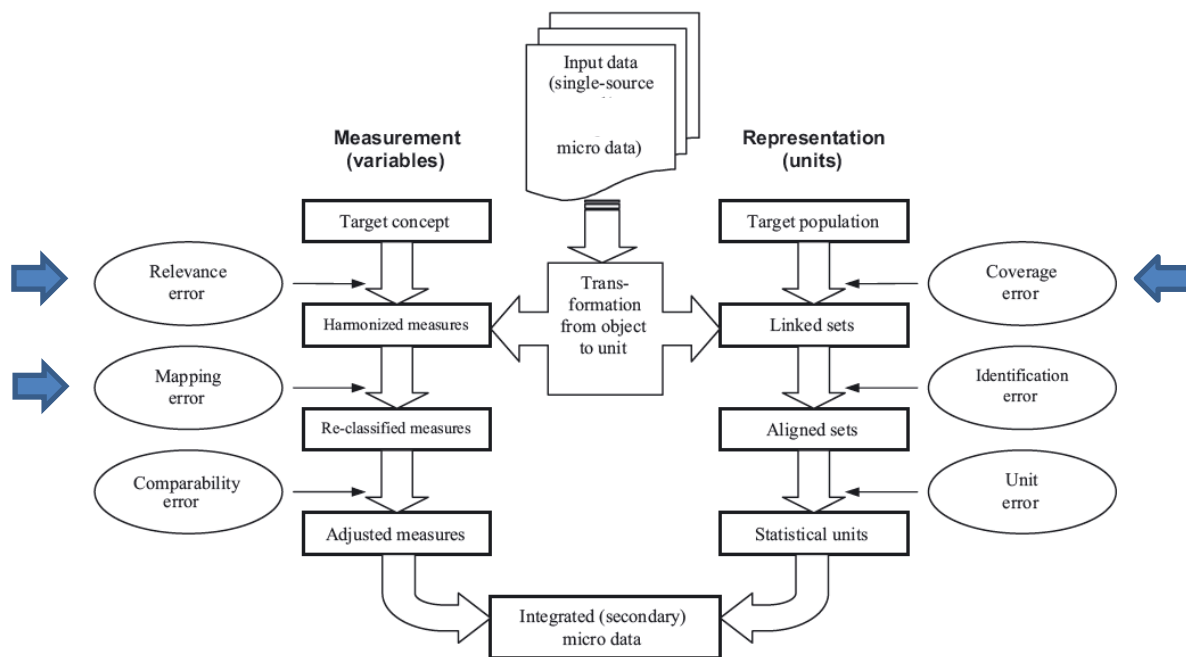
Provides an overview of potential error sources that can occur when statistics are based on combinations of various admin data and statistical data

Phase 1) Treatment of single-source micro-data w.r.t. administrative targets



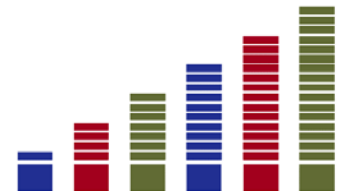
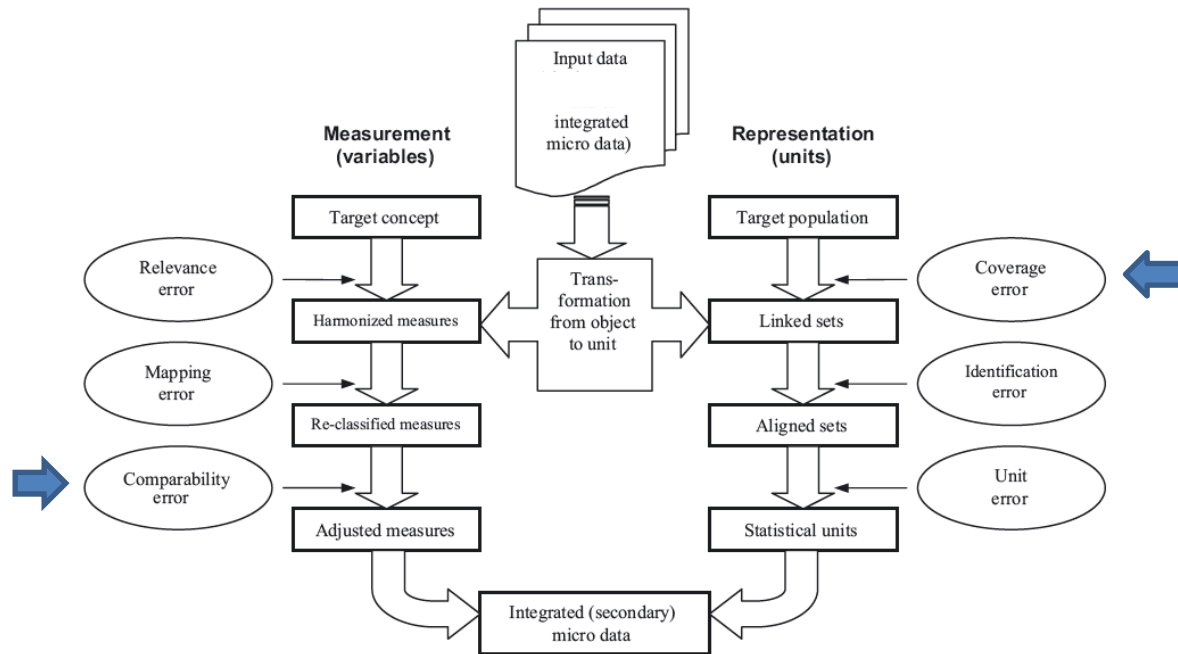
Identify potential sources of error

Phase 2a) Treatment of single-source micro-data w.r.t. SBS targets



Identify potential sources of error

Phase 2b) Treatment of the integrated micro-data w.r.t. SBS targets



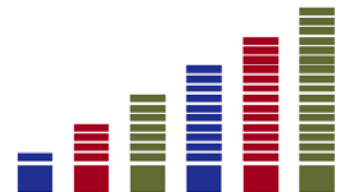
Defining quality indicators

- Quality indicators on both the measurements/variables and objects/units sides
- Qualitative and quantitative indicators

Zabala et al. (2013), Quality Measures for Statistical Outputs using Administrative Data, Stats New Zealand

EssNet AdminData; EssNet BLUE-ETS

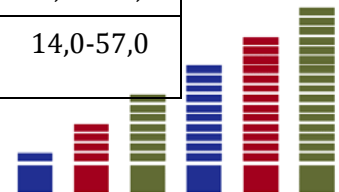
....



Example of quality indicators

Phase 2a - Errors and quality indicators (Year 2013)

	Indicator	Description	Source	Value
Variables. Target Concept Measures; -> Harmonized Measures; Relevance error	<i>Qualitative</i> indicators, by source	<i>Changes in definitions of all variables in each source and changes in definitions of SBS variables (Does definitions change over time?)</i>		
		<i>Conceptual scheme representing the re-classification of administrative concepts needed to produce the SBS variable definitions</i>		
Variables. Harmonized Measures -> Re-classified Measures; Mapping error	Quantitative indicators, by source	<i>Comparison of each harmonised variable with SBS benchmark variable (histograms, univariate statistics, statistical tests, etc.), to be repeated when variable definitions change</i>		
	Proportion of target variables which not require reclassification or mapping, by source	<i>No. variables captured directly from source S/Tot. no. variables x 100</i>	FS	100,0
			SS	71,0 – 85,0
		Unico	14,0-57,0	



Conclusions....

- The quality framework is **tailored** on the characteristics of the specific production process and statistical objectives
- In a longitudinal perspective, basis of information :
 - for monitoring **quality changes** over time
 - for the **continuous improvement** of the production process

...and future work

- Developing a **Quality Report** for documentation purposes
- Assessing the accuracy of estimates for variables with lower coverage rates in admin sources
- Measuring the uncertainty due to the imputation process



Thank you for your attention

