United Nations Economic Commission for Europe
Conference of European Statisticians

Workshop on Statistical Data Editing
18-20 September 2018
Neuchatel, Switzerland

Topic (iii): Reducing cost, and increasing efficiency of data editing

Topic organizers: Jeroen Pannekoek, CBS, Netherlands and Simona Rosati, Istat, Italy
Topic (iii): Reducing cost, and increasing efficiency of data editing - Introduction

• Reducing cost and increasing efficiency of data editing have been a central theme in the development of data editing methodology since the mid 1990’s.

• Selective and automatic editing have been successful in increasing the efficiency of editing traditional surveys.

• The theme remains urgent, not only to further increase the efficiency of traditionally collected data but even more so to obtain efficient editing strategies for alternative data sources.
Topic (iii): Reducing cost, and increasing efficiency of data editing - Introduction

• The papers in this topic have several points in common:
  – They all refer to applications of E&I in the domain of business statistics.
  – In all papers there is a central role of selective editing methods to reduce the costly manual editing.
  – Selective editing is usually combined by one or more automatic editing procedures to efficiently produce the necessary data quality.

• There is some attention for efficiency gains in operational aspects of E&I procedures by standardization of methods and software.

• One paper is on the editing of administrative data (VAT), all others on business surveys.
Topic (iii): Reducing cost, and increasing efficiency of data editing - Papers

1. Selective editing in the Integrated Business Statistics System (SINTESI) – Italy

2. Implementation of selective editing at SURS – the Republic of Slovenia

3. The strategy to treat unique enterprises in the Swiss survey on production and value added – Switzerland.

4. Investigating methods of efficient detection of errors in VAT data – UK.
Enjoy the presentations!
Topic (iii): Reducing cost, and increasing efficiency of data editing

Discussion
1. Selective editing in the Integrated Business Statistics System (SINTESI) – Italy
   - Selective editing with mixture model methodology (SeleMix)
   - Applied to Survey on Turnover and Orders in Industry

2. Implementation of selective editing at SURS – the Republic of Slovenia
   - Implementing selective editing in statistical process.
   - Reduced manual editing. But redesign of the statistical process and new tool for manual editing.
3. The strategy to treat unique enterprises in the Swiss survey on production and value added – Switzerland.
   ▪ Prioritizing unusually large units for both data collection and editing.
   ▪ Selection on score function based on number of employees.

4. Investigating methods of efficient detection of errors in VAT data – UK.
   ▪ Using a score function at unit-level (compares with historic mean).
   ▪ Combine with outlier detection (using a time-series model) on the macro-level.
Points for discussion

1. Preferred process flow? Automatic editing is often applied to the records that are not treated manually. But manual editing may also be applied after automatic editing (UK paper): high scoring values are first corrected by automatic imputation and thereafter manually validated.

2. A difficulty in implementing selective editing techniques reported by Slovenia is the acceptance by subject-matter statisticians. This problem came up in many previous Data Editing Workshops. Is it still a problem in other countries as well? What can we do about it?
Points for discussion

3. Standardization of E&I processes (Italy paper). What are the major obstacles encountered in trying to achieve this object? The advantages are clear; are there any disadvantages?