Topic (i) Selective and macro-editing

Introduction

Session Organizers: Emmanuel Gros, INSEE, France and Pedro Revilla, INE, Spain

Topics of interest
• Selective editing (score function approaches, significance editing, optimization methods) and macro-editing (output editing, top down editing) focus on identifying a limited amount of data with influential errors
• Used before or after data aggregation
• Particularly appropriate in the context of business surveys, where a few of the largest units have a decisive influence on results
• Increasingly used since they allow to save time and optimize manual review resources, while maintaining the quality of the statistical data products
• Especially focus
  - Unification of selective editing approach
  - Selective editing for classification variables
  - Quality assessment of selective editing
  - Complementarity between selective editing – throughout the data collection process – and macroediting after the data collection process
Topic (i) Selective and macro-editing Papers

Long presentations

• **Italy.** *Selective editing of business investments by using administrative data as auxiliary information.* Marco Di Zio, Ugo Guarnera, Massimiliano Iommi, Antonio Regano.

• **France.** *Output editing based on winsorization in the French SBS multisource system Esane.* Thomas Deroyon, Emmanuel Gros.

Short presentations

• **Spain.** *Developing a theoretical framework for selective editing based on modelling and optimization.* Pedro Revilla.

• **Netherlands.** *Changes in macro-editing and score functions for Dutch STS.* Mark van der Loo.

• **Italy.** *Model-based selective editing procedures for agricultural Price indices.* Tiziana Pichierri, Daniela Ichim, Maria Liria Ferraro, Ugo Guarnera.

• **USA.** *Selective and macro-editing of a large business based administrative data set.* David R. H. Hiles.

• **UK.** *Method for reviewing selective editing thresholds at ONS, RSI pilot study.* Charlotte Gaughan.
Topic (i) Selective and macro-editing

Discussion
Overall summary

Session papers provide a wide-ranging overview of current research and development activities in the area of selective & macro editing carried out at various NSIs with the aim of enhancing the efficiency of survey editing strategies.

• Links and interactions between selective & macro editing and administrative data are widely discussed, whether it is the question of the editing of administrative sources or the issue of using administrative data as auxiliary variables for selective & macro editing.

• There is also a focus on the application of selective editing methods on short term statistics and indices, with specific problems linked to times series (e.g. seasonality).

• Development & implementation of model-based selective editing procedures, where selection of units is based on model based estimates of error rather than traditional score functions.

• Issues related to quality assessment of selective editing are also discussed, with specific attention to those relating to maintenance of the implemented methods/models.
Points for discussion

Selective & macro editing and administrative data

- Difficulties of editing administrative or multisource data...
- ... But a very valuable source of information for data editing, which implies to develop new approaches.
- Which role for the administrative data:
  - Direct use for estimation purpose: feasibility, costs, role of selective & macro editing...
  - Auxiliary information for data editing of survey data: for error modelling, error detection, error correction? Advantages compare to other auxiliary information
  - Some mix of both: to what extent, conditions to decide to use administrative data as primary or auxiliary data source...

⇒ Papers from France, Italy and USA
Points for discussion

➤ Model-based selective editing

• Gives theoretical framework for data editing

• Modeling of the error distribution allow to use powerful tools, but modeling problems raise:
  – how to estimate parameters ?
  – convergence issue
  – robustness of the model to errors in the auxiliary information ?
  – sensitivity of the result of selective editing to misspecification of the model ?

• More powerful but more complex selective editing procedure
  ➔ can this black box side prevent the acceptation of new selective editing procedure by analysts/survey staff

⇒ Papers from Italy and Spain
Points for discussion

- Selective editing of STS / indices / time series
  
  - Selective editing and output editing seems even more complementary as for structural statistics
  
  - Specific issues:
    - how to deal with seasonality?
    - choice of the “level” for analysis?
    - historical information versus external information as auxiliary variables for selective editing

⇒ Papers from Italy, Spain and Netherlands
Points for discussion

➢ Quality assessment & maintenance of selective editing

• Importance of assessing quality of a selective editing procedure both on statistical and organizational point of view
• Which indicators, quality measures, evaluation approaches?
• Quality assessment under selectively edited data
• For maintenance or improvement, need of “over-edited” data:
  – Fully edited data versus sample of non-edited data
  – Size & sampling design of the sample of non-edited data?
  – Frequency of such evaluations and maintenance operations?

⇒ Papers from UK, Netherlands and Italy