

CONFERENCE OF EUROPEAN STATISTICIANS

UNECE Work Session on Statistical Data Editing

(27 – 29 May 2002, Helsinki, Finland)

TOPIC (III): EDITING OF ADMINISTRATIVE DATA

Issues for discussion

Discussant: Claude Poirier (Statistics Canada)

Invited Papers:

- Use of administrative data in population censuses – definition of main type of activity as an example (Finland)
- CHYPRES, A prototype of data tool for hybrid surveys (France)
- “Smart editing” of administrative categorical data (Israel)

Contributed Papers:

- Edit and imputation of the General Index of Financial Information (Canada)
- Use of administrative data by the Czech Statistical Office (Czech Rep.)
- Editing and imputation of Tax Return File – evaluation of applied methods (Finland)

This topic will deal with the major differences between editing of typical survey data and administrative data, like the amount of information, the difference in concepts, and the intended use. Issues to be discussed are:

- A. **What are the problems related to data definitions?** Data are defined differently for each administrative source. They also refer to various reference periods. Are models appropriate to « convert » data from one source to another one?
- B. **Is there any interest in correcting for the concept differences?** Editing may focus on differences between collected data and true values corresponding to the administrative concepts. In a second stage, it may also include differences between administrative and survey concepts.
- C. **When should we stop editing?** Administrative files are generally large and may then require more resource to reach a given level of quality. Macro-Editing or selective editing approaches may be a solution to this problem.
- D. **What editing methods can be used when dealing with administrative data?** Typical editing and imputation methods may be used depending on the available information: Manual editing, historical, ratios, models, balance editing, outlier analysis, etc. Discussion will target recommendations that depend on the target use: Frame development, sampling information, editing of other sources of data, estimation purposes, etc.
- E. **What are the operational constraints?** Linkage of micro-data, whenever needed, is one important issue. How can we deal with it? Are there any other constraints?
- F. **What information should be kept/deleted/added?** Administrative data may include duplicates or multiple entries for specific units. It may also under-cover the population of interest. What are the impacts and how to deal with this?
- G. **What are the “good practices/procedures” in editing administrative data?** How to use information from sample surveys to improve the administrative registers and vice versa.
