

**UNITED NATIONS STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE**

CONFERENCE OF EUROPEAN STATISTICIANS

WORK SESSION ON STATISTICAL DATA EDITING
(Ottawa, Canada, 16-18 May 2005)

Topic (v): Quality indicators and quality reporting

EVALUATING EDITING AND IMPUTATION PROCESSES: THE ITALIAN EXPERIENCE

Marco Di Zio, Orietta Luzi, Antonia Manzari

ISTAT, Italy

I. INTRODUCTION

Statisticians working in the context of data editing and imputation (E&I in the following) in National Statistical Offices (NSO in the following) agree that the two main goals of performing E&I on statistical survey data are, on one hand, improving the final data *quality* and, on the other hand, providing information about the overall survey process (including the E&I process itself). Therefore, any E&I process has to be designed and applied in order to meet as much as possible these requirements. The capability of an E&I process to reach these objectives can be verified through evaluation studies. The evaluation phase has been recognised as a central problem during last years, particularly because of the increasing demand of information about the characteristics and impact on data quality of E&I processes, and the need of keep under control the variability of final results due to the E&I.

In this paper we limit the discussion to the evaluation of those E&I activities performed at the data editing stage, i.e. excluding all the checks used in any previous stage of the whole survey process (data capture, data entry phase).

On the basis of some experiences performed or still in progress in ISTAT, in the paper we identify some general problems in the context of the evaluation of E&I procedures, and indicate the solutions adopted in ISTAT for dealing with different evaluation objectives.

The paper is structured as follows. In section 2 the evaluation problem is put into a general context taking into account the complexity of the E&I phase and the main different evaluation objectives. Sections 3, 4 and 5 deal more in depth with some particular evaluation objectives, with a discussion based on the Italian experience.

II. EVALUATION OF E&I PROCESSES: GENERAL PROBLEMS

The problem of the evaluation of the quality of E&I processes has been discussed in depth during last years in both a theoretical and operational context. The first step is to define the concept of *quality* of an E&I process.

For instance, Nordbotten (1999) defines the quality of statistical data editing as “a measure of how well the statistical process succeeds in satisfying users and producers needs”: users expect information about the data reliability, producers want information on the E&I process reliability for improving the resources allocation and the overall production performance.

A unique definition of the quality of an E&I process cannot be provided: it depends on the particular E&I aspect we are interested in. Hence the evaluation of the quality of an E&I process depends on the definition of the specific characteristic of the E&I process we want to highlight. Granquist (1997) performs a broad analysis of evaluation objectives through the description of several applications and studies presented in literature.

In our paper, we report the Italian experience on the following set of evaluation objectives:

1. verifying the statistical properties of a given E&I approach for a given E&I problem;
2. choosing the best set of techniques for a given survey application;
3. monitoring and optimising the performance of a given E&I methodology;
4. obtaining information on non-sampling errors sources;
5. measuring the impact of the E&I phase on the original raw data.

Objectives 1 through 4 are related to the producer's needs of information about the process reliability, while objective 5 aims at providing users with information about the reliability of the final data.

In practice, each objective corresponds to one of the phases that normally are to be carried out when building up and applying an E&I procedure. Depending on the phase, different quality and/or documentation needs are to be satisfied, as described in the following.

Typically, in the *planning* phase, objectives 1 and 2 are to be achieved in order to analyse and measure the potential performance, the possible benefits and the operational characteristics relating to the use of a new or a traditional methodology when applied to a given E&I problem or to a particular survey context. For example, objective 1 is pursued when the statistical properties of a new E&I method are to be tested and verified on real or simulated data reproducing the theoretical situation the method is designed to deal with. Objective 2 is typical of studies in which the evaluation aims at assessing the quality of an E&I process in terms of its capability of correctly deal with a real E&I problem.

Relating to this point, it is worthwhile to underline that the evaluation problem is generally a complex task because of the typical complex structure of any E&I procedure. In fact, an E&I process generally consists of many sub-phases, each of them performing a particular step of the whole data E&I process (Jong, 1996; Di Zio *et al.*, 2002b; Manzari, 2004). The simplest distinction refers to the error localisation and error imputation phases, but many other sub-processes can be identified, depending on the variables nature (ordinal, nominal, continuous), the error typologies (systematic or stochastic, influential or not influential, etc.), the survey characteristics (e.g. census, sampling surveys, panels), and so on. Therefore, an E&I process can be viewed as a completely integrated set of different techniques dealing each with a different E&I sub-problem but having the common aim of improving the overall quality of final results. Correspondingly, the problem of evaluating an E&I process can be split in the evaluation of simpler sub-problems, each of them focused on a specific E&I sub-phase. It is obvious that different evaluation criteria and priorities among them will be used depending on the particular sub-phase, on the features of the E&I method, and on the survey objectives. For example, if our goal is to evaluate the performance of a technique for treating influential data, the most relevant quality criteria could be the capability of the method of preserving aggregates reducing the costs and the respondent burden due to the interactive editing activities. Otherwise, if our aim is to verify the quality of an imputation method for data to be used to develop forecasting models, the most important quality criteria could be the preservation of micro data.

Once a given technique or an integrated set of techniques have been selected as elements of the overall E&I strategy, the strategy itself has to be *designed*, *implemented* and *tested* through an iterative evaluation process aimed at tuning and improving its performance (objective 3). In this case, the evaluation criteria and the corresponding measurements will be defined and implemented in order to facilitate the continuous monitoring of the effects of each E&I sub-process.

A very important objective in the evaluation context relates on obtaining useful information on the non-sampling errors nature and patterns from an E&I process (objective 4). Once the final E&I procedure has been defined through the above mentioned tuning process, it is possible to identify the most likely sources of non-sampling errors, due to lacks in one or more survey data processing phases, by analysing the errors type and structure. A typical product of objective 4 is the so-called *error profile*, i.e. a description of the characteristics of all identified errors and their internal structure (Cirianni *et al.*, 2000). This information is generally used

either for improving the survey organization in future survey repetitions, or when specific studies on particular error types or mechanisms are to be carried out (objectives 1 and 2).

Objective 5 is the minimal requirement that each statistical survey has to accomplish. In fact, the evaluation of the effects of an E&I procedure on a given set of survey data allows to obtain basic products like:

- the documentation of the E&I processing flow;
- the analysis of the statistical effects on data of the E&I process;
- an assessment of the quality of collected data to be provided to data users.

For each evaluation goal, a critical aspect is represented by the problem of the identification of *standard indicators* quantifying the corresponding quality criteria independently from both the particular application and method under evaluation. Many researches have been done with respect to the different evaluation contexts (Chambers, 2000; Stefanowicz, 1997; Madsen *et al.*, 2000; Fortini *et al.*, 1999; Nordbotten, 1997; Nordbotten, 2000). The need of defining generalised measures is increasing due to many reasons:

- the increasing need of internal and external comparability of the impact of E&I processes;
- the increasing need of international comparability of data E&I processes;
- the development and application of new techniques requiring the assessment of their usefulness and statistical properties;
- the increasing demand of information on data processing coming from end users.

The assessment of quality of data and data processing, as well as quality reporting are central problems not only at NSO level, where the effort is concentrated in the definition and dissemination of *statistical guidelines* and *quality reports* (National Center for Education Statistics, 1992; Statistics Canada, 1998), but also in an international perspective: as an example, great attention is devoted to these aspects at Eurostat, where increasing importance is given to the problem of non-sampling error documentation and measurement (LEG, 2001).

Given the evaluation objective and the corresponding quality criteria, a crucial decision relates to how the evaluation has to be carried out, i.e. which particular evaluation approach has to be adopted. A general distinction can be made between *observational* and *experimental* studies (Biemer *et al.*, 1995). In the first context, the investigators study the phenomenon as they find it. For the purpose of this paper, we mean that the assessment of the quality of an E&I process is based only on observed data, eventually enriched with all available information on surveyed units and variables (e.g. coming from historical or administrative sources of information, or collected through re-interviews or complete data revision processes) (Granquist, 1995; Granquist, 1997; Biemer *et al.*, 1995; Poulsen, 1997). In the experimental approach, we generally investigate the phenomenon having a certain control on the mechanism ruling some key aspects of the phenomenon itself. In our context, it means that we simulate the situations we would investigate, and make inference from the simulation results obtained for assessing the quality of a procedure (an E&I procedure, or a single technique) (Kovar *et al.*, 1996).

Given the evaluation objective, the choice is sometimes obliged. For instance, if the objective is to choose the *best* technique for dealing with a specific E&I problem, an experimental approach is more appropriate than an observational one. On the contrary, measuring the effects on data of an E&I process is a typical observational study. In other cases, however, there are other elements influencing the choice of the approach; the most important are the available resources (budget, human, tools, other costs), the timeliness requirements, the available external information.

III. EVALUATING THE QUALITY OF E&I METHODS: THE SIMULATION APPROACH

In this paragraph the objectives 1) and 2) mentioned in the previous section are synthetically discussed and a general description of the evaluation approaches used until now in ISTAT for dealing with them is provided. As already mentioned in paragraph II, both these objectives aim at providing information to data producers in the planning phase of an E&I process: this information is generally used for improving the efficacy (in terms of quality of results) and the efficiency (in terms of costs) of each E&I phase and hence of the overall E&I process.

The evaluation experiences performed until now in ISTAT can be classified in both the observational and experimental context.

In the observational context, the evaluation of E&I approaches has been performed essentially in the following area: for a given E&I problem, starting from raw data, verifying the usefulness of a new methodology by comparing its results with those obtained by using the traditional approach (Di Zio *et al.*, 2002a), or evaluating the best technique among different new approaches (Cirianni *et al.*, 2001).

As already mentioned the experimental approach is the natural framework where the evaluation of the quality of E&I methods can be embedded. In ISTAT the simulation approach (Della Rocca *et al.*, 2000; Barcaroli *et al.*, 1997) is at present adopted in many evaluation studies. This approach consists of the following steps:

- i. corrupt original true values with errors and/or missing values;
- ii. apply the E&I strategy in order to obtain the edited data;
- iii. compare some quantities computed on the original true values with the same quantities computed on the final edited values.

Before a more detailed discussion on the theoretical aspects related to the use of this approach, we want to stress some practical aspects.

The simulation approach is effective for both the evaluation objectives 1) and 2) in section II. In fact, the first objective consists of verifying the behaviour of a technique with respect to some theoretical situations of a given E&I problem; for example, if we want to evaluate the performance of two techniques with respect to the presence of random errors, we can simulate random errors in data and comparing the results obtained with the two techniques.

For the second evaluation objective, that is to choose the best set of techniques for a given survey application, we try to simulate real situations and evaluate the method (or set of methods like an integrated E&I procedure) with respect to them. To reach this goal it is important to analyse the *error profile* of the survey we are going to investigate. In this case, it would be possible to reproduce, at least for the main aspects, the error mechanism affecting data.

The main advantages of this approach with respect to other ones are low costs and the possibility of controlling some critical aspects of the applications. Drawbacks mainly relate to the effort required for modelling the error patterns and characteristics.

Some interesting experiences in this area have been carried out in ISTAT in different contexts: comparing the performance of different approaches to a given E&I problem (Manzari *et al.*, 2002), assessing the capability of a new approach of dealing with some specific kinds of errors (Manzari, 1999; Coppola *et al.*, 2002) and, in the context of the European EUREDIT project (www.cs.york.ac.uk/euredit/), evaluating the best method in an overall survey context. In the latter case, the evaluation criteria and the corresponding performance indicators are those described in Chambers (2000).

In the following of this section, we will stress some theoretical aspects related to the use of the above cited simulation approach.

Let us suppose we have observed p variables on n units. Let $X=(X_1, \dots, X_p)$ be the vector of original true values, let $Y=(Y_1, \dots, Y_p)$ be the vector of corrupted values and let $X'=(X'_1, \dots, X'_p)$ be the vector of edited final values. As so far introduced, for evaluating the E&I procedure, we generally compare the generic quantities $Q(X)$ and $Q(X')$ through a suitable generic loss function $D(Q(X), Q(X'))$. One of the possible choice is for instance $Q(X) = X$ and $D(Q(X), Q(X')) = (Q(X) - Q(X'))^2$ the square of the difference between original and edited values.

It is worthwhile noting that, the X' values can be thought as a realisation of a random vector from a multivariate probability distribution depending on different random mechanisms: the corruption of the original values with probability law f (error mechanism), the error localisation method (with probability law g when it

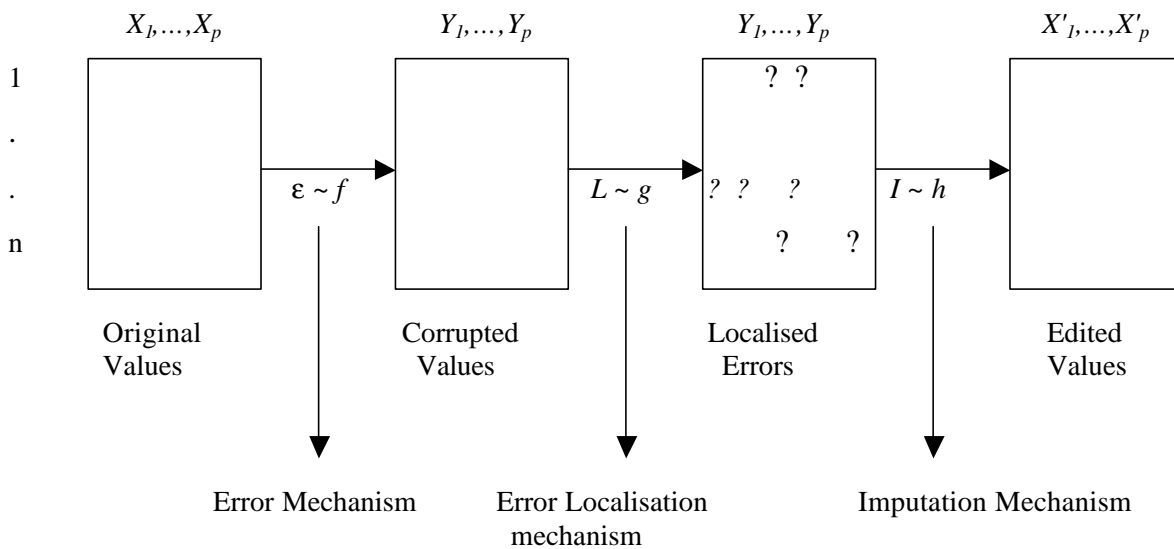
is not deterministic) and the imputation method (with probability law h when it is not deterministic); this process is shown in the following Figure 1.

In this context, to statistically evaluate the quality of the E&I process we have to consider all the stochastic mechanisms affecting the results. Thus if we want to evaluate fairly the performance of the E&I process through a suitable generic function $D(Q(X), Q(X'))$, we should need to compute it with respect to the distributions of the stochastic mechanisms present in the process. One natural way, allowing to take into account the different random mechanisms, is to compute the expected value of D :

$$E_{e,L,I} [D(Q(X), Q(X'))] \quad (1)$$

where (e, L, I) represent the random mechanism as in Figure 1. It is clear that choosing $D(Q(X), Q(X')) = (Q(X) - Q(X'))^2$, we obtain the well known Mean Square Error.

Figure 1: Editing and imputation process in a simulative context



Generally, we are not able to compute analytically the expected value (I) because of the difficulty of modelling the joint distribution of the random variables X' . On the other hand, we are often able to draw observations from the joint distribution of the random variable $X'(e,L,I)$. In this context we can use a Monte Carlo integration method to approximate the expected value (I) (Rubinstein, 1981). If we repeat m times the E&I process, we obtain $(X^{(1)}, \dots, X^{(m)})$ that can be thought as a m sample drawn from the distribution of X' , hence for m sufficiently large

$$\frac{1}{m} \sum_{i=1}^m D(Q(X), Q(X^{(i)})) \cong E_{e,L,I} (D(Q(X), Q(X')))$$

As illustrated in Figure 1, since the process is hierarchical, it could be interesting to assess the impact on the final data of each single stochastic mechanism. This can be achieved through the usual formulae with respect to the natural conditioning arising from the stream depicted in Figure 1. For example if we suppose that the imputation mechanism is not stochastic, we can decompose the total variance in a part induced by the error localisation mechanism and by the corruption method used, i.e.

$$Var_{e,L} (D(Q(X), Q(X'))) = E_e [Var_{L|e} (D(Q(X), Q(X')))] + Var_e [E_{L|e} (D(Q(X), Q(X')))]$$

where $E_e [Var_{L|e} (D(Q(X), Q(X')))]$ gives the part of total variance expressed by the variability of the error localisation random mechanism, and $Var_e [E_{L|e} (D(Q(X), Q(X')))]$ quantify the part of total variance expressed by the variability of the error introduction random mechanism. All these quantities can be computed through simulations by Monte Carlo methods.

IV. TUNING AN E&I PROCESS

Activities performed at the data E&I stage are recognised as costly and time consuming. The importance of optimising the effectiveness and efficiency of E&I processes has been continuously underlined in literature, and possible ways of improving the performance of E&I activities in terms of quality and resources allocation have been discussed by many authors (Granquist, 1995; Lepp *et al.*, 1993). Some general problems relating to the design and application of E&I processes have been highlighted in Granquist (1995) together with possible solutions: in particular, the Author identifies three types of costs connected with editing (producer costs, respondent cost and losses in timeliness) and suggests some possible actions the survey managers can do in order to improve the cost-benefit efficiency of editing processes (evaluation methods, developing improved edits, improving questionnaire design, limiting manual follow-up to those flagged records with the heaviest potential impact on estimates).

It is well known that E&I strategies are complex integrated processes consisting of several sub-phases: therefore, there is not a unique way to apply a given E&I strategy to a given set of data, depending on the specification of a series of elements as the hierarchy among units/errors/variables, the edit rules (logical, mathematical, statistical), the stratification variables, the imputation algorithms, and so on. Therefore, once the components of the E&I process have been defined (in terms of approaches, methods, algorithms to be used for dealing with the different E&I problems), the actual operational characteristics of the process itself are to be *designed* and *implemented*, and the performance of the E&I strategy has to be *tested* and *tuned* on the actual observed data. These tasks are generally accomplished through an iterative evaluation process in which the subject matter expert deals with the above-mentioned E&I elements for improving the performance of the integrated set of techniques involved in the overall E&I strategy.

In this evaluation area, the problem of measuring the performance of an E&I process can be re-formulated in terms of what is happening to data *during* its execution (Nordbotten, 2000; Thomas, 1997). In particular, the evaluation objectives can be resumed in the following points:

- verifying the effectiveness of the adopted editing methods in terms of results reliability;
- verifying that the E&I process results correspond to those expected in the planning phase in terms of costs and timeliness;
- verifying that resources are allocated in the appropriate way through the E&I process sub-phases.

To these aims, a great help in specifying and setting the E&I process comes from the acquired knowledge on the matter obtained by previous *quality evaluation studies* or *past experiences* in the application of the selected methodology in a survey process similar to the specific current one. The existing documentation can assist the statistician in choosing some solutions and in discarding some others. However, the subject matter expert has to deal with new collected data, so, although the hints provided by the existing documentation, he needs information suitable to analyse alternative settings and to choose the one which provides the *best* E&I process in his production context, that is to evaluate the E&I process.

Therefore, it is good practice to supply each E&I process with E&I diagnostics in order to continuously monitor its execution and be able of identifying as soon as they appear possible problems. The diagnostics are helpful tools whatever the survey features and the E&I system used.

These diagnostics represent performance indicators to be used for monitoring the reliability and suitability of the applied techniques. Performance criteria that can be used in this evaluation context, in which generally only raw and edited data are available, are, for example, data processing costs, timeliness, amount of creative and over-editing, respondent burden, preservation of micro and macro data, edits consistency and usefulness, data coherence with other sources of information. It is obvious that while some of these criteria are common to many E&I sub-phases, like data processing costs and timeliness, other ones are specifically used in some phases: for example, preservation of both micro data and data distributions are particularly critical in automatic data processing, while respondent burden and aggregates preservation mainly relates to selective editing and interactive editing phases (De Pol, 1997; Engstrom, 1995; Latouche *et al.*, 1995).

During last years ISTAT researchers essentially have worked in two main directions: improving the trade off between timeliness and data quality in automatic editing (in both households and business surveys) and reducing costs and time related to the manual/interactive editing (particularly in business surveys).

In the latter area, the problem of improving the identification of relevant units (in terms of their impact on estimates) has been faced through the approaches of *macro editing* (Barcaroli et al, 1995) and *selective editing* (Luzi et al., 1999; Di Zio et al., 2002b; Cirianni et al., 2000). Encouraging results have been obtained in all the performed studies and applications. Further research is needed in this area.

In the field of automatic data processing, lot of experience has been gathered in ISTAT in the use of performance measures produced by generalised software dealing with random not influential errors.

As known, generalised software facilitates the standardisation of approaches and methods used for specific E&I problems, giving the user guarantees about quality requirements through the use of theoretically consolidated algorithms, and allowing to save resources and time by eliminating the need of developing ad hoc code.

A further advantage of using these systems relates to the possibility of checking the redundancy and the consistency of edits. Redundancy between edits does not affect the correctness of their definition but can affect the efficiency of the E&I process. On the contrary, inconsistency between edits means that can be errors in the definition of some edits, and this generally affects the efficacy of the E&I process. An overall description of possible problems and drawbacks related to the incorrect use of edits is done in Granquist (1995). Here we refer to the use of validity checks and consistency edits used in automatic data E&I for identifying fatal errors.

Another important advantage related to the use of generalised software is generally the standardisation of evaluation and documentation measures through the production of minimal sets of indicators and reports useful in monitoring and calibrating the effects of the data processing actions performed on data. Examples of these reports are those produced by systems like SCIA (Riccini et al., 1995) and GEIS (Kovar et al., 1996), actually used in ISTAT for dealing with random errors respectively in categorical and continuous variables. Both software are based on the Fellegi and Holt approach (Fellegi et al., 1976), thus they give some guarantees to the user about the quality of results (minimum change of original data, final data plausibility with respect to the user-defined edits, preservation of original distributions under specific conditions). A new software, called DIESIS (Bruni et al., 2001; Bruni et al., 2002), has been developed for dealing with categorical and continuous data, implementing the *data driven* and the *minimum change* approaches.

Among the operational tools provided by these softwares for monitoring and evaluating the reliability of an automatic E&I process there are:

1. the *number of passing* and *failing* records;
2. the *edits failing frequencies*;
3. the *imputation frequency* and the *edits* mostly involved in the imputation process with the corresponding *frequency* of such involvement;
4. the *frequency distributions* before and after imputation and the *transition matrices* containing the frequencies of occurrence between values before and after imputation (for qualitative variables);

All these tools allow the subject-matter expert to analyse the usefulness and the reliability of the automatic strategy through the observation of its effects on data: for example, a too high difference between the number of *failing* and *passing* records provides a first insight that something is going wrong in data or in the E&I process; the *failure rates of edits* can reveal errors in the definition of edits, they can also help in checking the nature of the errors in data (random or systematic) and in reducing possible over-editing of data; the analysis of the *transition matrices* is useful to evaluate some aspects of the performed corrections: if concentration of frequencies are noted in cells outside the main diagonal and the marginal distribution does not show the same pattern, this can be due to errors in the definition of edits and/or to the presence of residual systematic errors.

V. ANALYSING THE IMPACT OF E&I PROCESS

Increasing attention has been paid to the definition of standard quality indicators and the production of standard quality reports at NSOs and Eurostat level. Many studies can be found in literature dealing with the problem of assessing and documenting the impact on data of E&I procedures (among others, Madsen, 2000; Nordbotten 2000; Lindell, 1997; Whitridge *et al.*, 1999; Cirianni *et al.* 2000), and several indicators have been proposed for evaluating the effects on raw data of E&I procedures. In this context, we can say that the term *quality* relates to the transparency, reliability and suitability of the E&I procedure with respect to the objectives of the planned and designed one.

Regardless of the adopted procedure, the production of standard quality indicators for evaluating and documenting the data processing activities performed on data is always recommended. Generally, these standard indicators are to be produced for satisfying both data users and producers needs. Such quality measures are in fact applied with the following main purposes:

- i. providing users with information about the characteristics of the E&I process, the quality of final data and the data modifications due to the E&I process;
- ii. providing subject-matter experts with information about the characteristics of errors in order to allow the identification of their possible sources;
- iii. providing the subject-matter experts with detailed documentation on the process impact on data (and hence on data quality) in each survey repetition in order to allow the data and process quality monitoring over the time;
- iv. providing NSO with information about the quality level of each survey in order to allow the comparison among *similar* surveys of the data and process quality.

According to Granquist and Kovar (1997), all these objectives aim at “gathering intelligence related to significant difference in the data for analytical purposes” and “providing feedback that can lead to improvements in data collection and processing”.

Researchers and subject matter experts agree on the importance of editing activities as source of information on the error sources: the analysis of the patterns of error and their characteristics generally allow the identification of possible lacks and problems in some survey phases (including E&I). This information can be used to plan the appropriate modifications of the survey organisation/methods in the future survey repetitions.

On the other hand, the analysis of the impact of E&I in terms of time, costs, and modifications produced on the statistical properties of raw data allows the statisticians to verify the reliability and the usefulness of the adopted methodologies and plan revised or new approaches for improving the performance of the E&I process.

Similarly, the comparative analysis over time of appropriate quality indicators for a given survey can highlight structural modifications in the surveyed phenomena or occasional organisational problems needing adjustments to the survey organisation or to the E&I procedure.

The comparison of standard measures among similar surveys in a same period is a powerful tool that survey managers and NSOs can use for identifying structural or methodological lacks in their surveys or data processing strategies, thus stimulating their revision and the improvement.

The common problems to all the objectives i.-iv. can be summarised in the following points:

- identifying a set of appropriate standard quality indicators;
- supporting data producers through the development of tools for computing indicators in a standard way;
- convince data producers about the need of the evaluation task and the advantages that can be obtained from the availability of standard quality indicators.

During last years ISTAT researchers have worked mainly in the following directions:

1. providing the survey managers with an information system especially dedicated to documenting and analysing the quality of the survey production process, including the E&I phase;

2. providing the subject-matter experts with a generalised tool for analysing the impact of their own data E&I processes on the statistical properties of collected data;
3. encouraging the production and the use of standard indicators and measures for documenting the E&I strategies and their impact on observed data, and monitoring the effects of E&I procedures over the time.

Concerning the first area, the centralised information system SIDI (Brancato et. al, 1998) has been designed and developed in Istat to support the quality control activity through the monitoring of the survey production process, the documentation of data production activities and the dissemination of suitable information on data quality to the final users. The adopted approach focuses on the production process according to the concept that the quality of final data benefits of improvements in the process. To this purposes, SIDI manages not only quantitative information (quality indicators) but also qualitative information (metadata on the survey production process characteristics). Relating to the E&I phase, the standard quality indicators and the metadata on the E&I phase included at present in SIDI (Fortini et al., 1999) are as much general as possible for allowing the comparative evaluations of surveys. All indicators use *raw* and *clean* data and allow the survey managers to analyse the impact of the editing activity on the statistical units and on the observed variables.

Relating to point 2, in order to facilitate and standardise the production of evaluation indicators measuring the statistical effects on data of E&I, a generalised software implementing both the SIDI indicators and other statistical measures has been designed and implemented (Della Rocca *et al.*, 2003). The available indexes correspond to quality criteria like preservation of elementary data, preservation of aggregates and distributions, preservation of data relations, exploiting the results on evaluation criteria done in the context of the EUREDIT project.

As regarding the third filed of activity, as mentioned in section II, detailed reports on the effect of automatic data E&I procedures are automatically produced by some generalised software available in ISTAT. These reports can be used for both documenting the automatic E&I activity impact on observed data and for analysing error typologies. For this reason, in ISTAT the use of such generalised software is encouraged. However, even in case of *ad hoc* software, the production of similar descriptive measures and reports is recommended, as well as for all the E&I sub-phases other than the automatic one. Relating to the latter, particularly in case of business surveys, future research will be concentrated in the analysis and development of impact measures in a generalised framework.

REFERENCES

- Barcaroli G., Ceccarelli C., Luzi O. (1995). An edit and imputation system for quantitative variables based on macroediting techniques. *Proceedings of the International Conference on Survey Measurement and Process Quality*, Bristol (UK), 1-4 Aprile, pp.12-17.
- Barcaroli G. and D'Aurizio L. (1997). Evaluating Editing Procedures: the Simulation Approach. *UN/ECE Work Session on Statistical Data Editing*, Prague, October 1997.
- Biemer, P., Fecso, R.S. (1995). Evaluating and Controlling Measurement Error in Business. In *Business Survey Methods*, eds. B.G. Cox, D.A. Binder, B.N. Chinappa, A. Christianson, M.J. Colledge, and P.S. Kott, New York: Wiley, pp.267-275.
- Brancato G., D'Angiolini G., Signore M. (1998). Building up the quality profile of ISTAT surveys. *Proceedings of the conference Statistics for Economics and Social Development*, Aguascalientes (Mexico), September 1-4.
- Bruni R., Reale A., Torelli R. (2001). Optimisation Techniques for Edit Validation and Data Imputation. *Proceedings of the XVIIIth Statistics Canada International Symposium on Methodological Issues "Achieving Data Quality in a Statistical Agency: a Methodological Perspective"*.
- Bruni R., Reale A., Torelli R. (2002). DIESIS: a New Software System for Editing and Imputation. *Proceedings of the Conference of the Italian Statistical Society*, Milano, 2002.

- Chambers R. (2000). *Evaluation Criteria for Statistical Editing and Imputation*. T001.05, EUREDIT report.
- Cirianni A., Di Zio M., Luzi O., Seeber A.C. (2000). The new integrated data editing procedure for the Italian Labour Cost survey: measuring the effects on data of combined techniques. *Proceedings of the International Conference on Establishment Surveys II*, Buffalo, June 17-21.
- Cirianni A., Di Zio M., Luzi O., Palmieri A., Seeber A.C. (2001). Comparing the effect of different adjustment methods for units with large amounts of item non-response: a case study. *Proceedings of the International Conference on Quality in Official Statistics* (First version), Stockholm, Sweden, May 14-15.
- Coppola L., Di Zio M., Luzi O., Ponti A., Scanu M. (2002). On the use of Bayesian Networks in Official Statistics. *Proceedings of the DataClean Conference*, Finland, May 2002.
- Della Rocca G., Di Zio M., Manzari A., Luzi O. (2000). E.S.S.E. Editing System Standard Evaluation. *Proceedings of the SEUGI 18*, Dublin, June 20-23.
- Della Rocca G., Di Zio M., Luzi O., Signore M., Scavalli E., Simeoni G. (2003). Documenting and monitoring Editing and Imputation Processes in the Italian Institute of Statistics, *UN/ECE Work Session on Statistical Data Editing*, Madrid, October 20-22.
- Di Zio M., Luzi O. (2002a). Comparing a purely deterministic and a semi-probabilistic approach for editing and imputation of Agricultural Census data. *Scritti di Statistica Economica*, Vol. 9, September 2002.
- Di Zio M., Luzi O., (2002b). Combining methodologies in a data editing procedure: an experiment on the survey of Balance Sheets of Agricultural Firms. *Italian Journal of Applied Statistics*, Vol. 14, No 1, pp. 59-80.
- De Pol F. (1997). Selective Editing in the Netherlands Annual Construction Survey. *Statistical Data Editing Methods and Techniques Vol. II*, Conference of European Statisticians, United Nations, 1997.
- Engstrom, P. (1995). A Study on Using Selective Editing in the Swedish Survey on Wages and Employment in Industry. *UN/ECE Work Session on Statistical Data Editing*, 1995.
- Fellegi I. P. and Holt D. (1976). A systematic approach to edit e imputation, *Journal of the American Statistical Association*, Vol.71, pp. 17-35.
- Fortini M., Scanu M. and Signore M. (1999). Measuring and Analysing the Data Editing Activity in ISTAT Information System for Survey Documentation. *UN/ECE Work Session on Statistical Data Editing*, Rome, June 2-4.
- Granquist, L. (1995). Improving the Traditional Editing Process. In *Business Survey Methods*, eds. B.G. Cox, D.A. Binder, B.N. Chinappa, A. Christianson, M.J. Colledge, and P.S. Kott, New York: Wiley, pp. 385-401.
- Granquist, L. (1997). An overview of Methods of Evaluating Data Editing Procedures. *Statistical Data Editing Methods and Techniques Vol. II*, Conference of European Statisticians, United Nations, 1997.
- Granquist L. and Kovar J. (1997). Editing of Survey Data: How Much is Enough? In *Survey Measurement and Process Quality*, eds. L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, and D. Trewin, New York: Wiley, pp. 415-435.
- Kovar J. and Winkler E.W. (1996). Editing Economic Data. *UN/ECE Work Session on Statistical Data Editing*, 1996.
- Latouche M., Bureau M., Croal J. (1995). Development of cost-effective edit and follow-up process: the Canadian survey of employment experience. *UN/ECE Work Session on Statistical Data Editing*, 1995.
- LEG (2001) *Proceedings of the International Conference on Quality in Official Statistics* (First version), Stockholm, Sweden, May 14-15.
- Lepp H. and Linacre S.J. (1993). Improving the efficiency and Effectiveness of Editing in a Statistical Agency, *Bulletin of the International Statistical Institute, Proceedings of the 48th Session*, Florence, 11-112.
- Lindell K. (1997). Impact of Editing on the Salary Statistics for Employees in County Council. *Statistical Data Editing Methods and Techniques Vol. II*, Conference of European Statisticians, United Nations, 1997.

- Luzi O. and Pallara A. (1999). Combining Macroediting and Selective Editing to Detect Influential Observations in Cross-Sectional Survey Data. *UN/ECE Work Session on Statistical Data Editing*, Rome, June 2-4.
- Madsen B. and Solheim L. (2000). How to measure the effect of data editing. *UN/ECE Work Session on Statistical Data Editing*, Cardiff, October 18-20.
- Manzari A. (1999). A New Imputation Methodology: an Experimental Application to Italian Population Census Data. *Proceedings of the Conference of the Italian Statistical Society*, Udine, 1999 (in italian).
- Manzari A. and Reale A. (2001). Towards a new system for edit and imputation of the 2001 Italian Population Census data: A comparison with the Canadian Nearest-neighbour Imputation Methodology. *Proceedings Actes of the International Association of Survey Statisticians, Invited Papers, August 2001, Seoul*, pp. 634-655.
- Manzari A. (2004). Combining Editing and Imputation methods: an Experimental Application on Population Census Data. *Journal of the Royal Statistical Society Series A*, Vol. 167 (2), pp. 295-307.
- National Center for Education Statistics (1992). *NCES Statistical Standards*.
- Nordbotten S. (1997). *Metrics for the Quality of Editing, Imputation and Prediction*. Statistics Sweden Technical Report, SN, 3-Oct-97.
- Nordbotten S. (2000). *Evaluating Efficiency of Statistical Data Editing: A General Framework*. United Nations, 2000.
- Poulsen M.E. (1997). Evaluating Data Editing Process Using Survey Data and Register Data. *Statistical Data Editing Methods and Techniques Vol. II*, Conference of European Statisticians, United Nations, 1997.
- Riccini E., Silvestri F., Barcaroli G., Ceccarelli C., Luzi O., Manzari A. (1995). *The Methodology of Editing and Imputation by Qualitative Variables Implemented in SCIA*. Technical Report. Italian National Institute of Statistics.
- Rubinstein, R. Y. (1981). *Simulation and the Monte Carlo method*. New York: Wiley.
- Stefanowicz, B. (1997). Selected issues of Data Editing. *Statistical Data Editing Methods and Techniques Vol. II*, Conference of European Statisticians, United Nations, 1997.
- Statistics Canada (1998). *Statistics Canada Quality Guidelines*.
- Thomas, J. (1997). Statistical measurement and monitoring of data editing and imputation in the 2001 United Kingdom Census of Population. *Statistical Data Editing Methods and Techniques Vol. II*, Conference of European Statisticians, United Nations, 1997.
- Whitridge P. and Benier J. (1999). The impact of Editing on Data Quality. *UN/ECE Work Session on Statistical Data Editing*, Rome, June 2-4.
