INSEE
DIRECTION GÉNÉRALE
INSTITUT NATIONAL
DE LA STATISTIQUE
ET DES ÉTUDES
ÉCONOMIQUES

Guidelines for finding a balance between accuracy and delays in the statistical surveys

This project has been funded by the European Commission
Eurostat Grants for 2005 in the field of quality management and evaluation
Grant agreement number : 10300.2005.021-2005.707

Version V1 - june 2007

Table of contents

Introduction

In order to facilitate the implementation of the European Code of Practice for quality, Eurostat supports the development of a set of handbooks on current best methods. These handbooks are intended to present a review of the state-of-the-art on a defined subject, as well as a review of the applied current practices of the Member States.

This handbok is relative to the question of the **balance between accuracy and delays in the statistical surveys**. Timeliness and accuracy are two of the components that have been defined in the Eurostat quality report (Eurostat, 2003). They are often conflicting : on the one hand, one wants to produce statistics as soon as possible (and in some cases, the deadline is very strict) ; on the other hand, different phases of the survey process, as for example the work of data editing, may take a long time, and the longer it is the most accurate statistics are. Also, in some cases (especially mail surveys), statisticians have to wait for the returns of questionnaires, and the delays for producing estimates are considerably increased. Finally, one may also consider that, generally, delays are linked to cost aspects (the longer the delays, the higher the costs for the statistical institute), and also, in some cases, to an increase of the burden for the respondents (for example, if the units having sent questionnaires for which variables are suspicious are contacted again).

From the user's point of view, there is more and more pressure to "improve" the timeliness of statistics, especially for short-term economic statistics. This pressure does take place at the national level, but is also coming from international institutions as the European Central Bank, the OECD or the IMF (see for example Bier and Ahnert, 2001, or McKenzie, 2006).

But all kinds of statistics are concerned : short-term as well as structural statistics, social as well as business statistics. This handbook deals with the preparation and execution of statistical surveys. First, as mentioned before, an assessment of the practices of the National Statistical Institutes (NSI) of the Member States on the subject was necessary. It has been made by sending a questionnaire to these NSIs.

The results of this survey, as well as the questionnaire, are presented in two annexes of this handbook, and a summary of these results is presented in **chapter 1** of the handbook.

Chapter 2 is relative to the general implementation and execution of a survey : is it possible, before implementing a survey, to study, for each of the different steps (sampling, preparation of the questionnaire, data processing, …), how to reduce the time required ? An overall strategy might be considered to use the total budget in an optimal way.

The question of the delays is also to be tackled during the execution of the survey : observing some indicators (as response rates relative to different categories, "quality" of the variables of the questionnaires, time remaining before the deadline for publishing statistics), is it possible to make choices among the different possible actions to make ? **Chapter 3** gives some elements on this question.

Then, **chapter 4** leaves the strict field of statistical surveys and gives a few elements on methods of production of preliminary estimates that are particularly used for short-term statistics.

Chapters 2, 3 and 4 are followed by executive summaries, and give references of papers used for this handbook (most of the references are available on websites, but not all of them).

The present handbook does not intend to present a complete review of all methods that may be used considering the subject of balance between timeliness and accuracy, since other documents may exist, that are more specialized on a specific subject. For example, the project EDIMBUS (Luzi, De Waal, Hulliger, 2006) is aimed to produce a recommended practices handbook on the subject of edit and imputation in business surveys. Some data bases, as the data base of OECD dedicated to short-term statistics (McKenzie, 2006), or the K-Base, may also be very interesting. The present handbook focusses mainly on the overall approach of methods useful for timeliness aspects.

One main outcome of the handbook is that the subject does concern business surveys in a more important way than household surveys, because of different reasons that are the more quantified aspects of the questionnaires of business surveys, the skew distribution of enterprises, the focus on the production of aggregates for business surveys, compared to the studies of "distributions" that are more frequent for household surveys (and not very well fitted to the methods presented here to reduce the delays). So, some survey managers (for example of household surveys) may find here less interesting elements than business surveys managers.

A last point to mention is that the organizational aspects (including the "human" component) are important concerning timeliness aspects. They are not always studied in the academic papers relative to the methods presented in this handbook : some elements on this subject are given in the following chapters.

The elements presented in this handbook are mainly based on the results of the survey of the NSIs and on some personal experiences. Unavoidably, they may "reflect" a partial view, and leave for example some important references behind ; the present version has thus to be considered as a first version, that could be improved by the comments of colleagues from different NSIs.

At last, people having provided informations useful for this handbook, especially through the survey of the NSIs, have to be thanked.

References :

Eurostat (2003) : Definition of quality in statistics - Eurostat working group « Assessment of quality in statistics", Luxembourg, 2-3 october (http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47141301/DEFINITION_2.PDF)

Bier W., Ahnert H. (2001) : Trade-off between timeliness and accuracy - ECB requirements for general economic statistics (http://dsbb.imf.org/vgn/images/pdfs/bierahnert.pdf)

K-BASE : The knowledge base on statistical editing http://amrads.jrc.ec.europa.eu/k-base/

Luzi, O., De Waal, T., Hulliger, B. (2006) : Recommended practices for editing and imputation in the European statistical system : the EDIMBUS project, UN/ECE Work Session on Statistical Data Editing, Bonn (http://www.unece.org/stats/documents/ece/ces/ge.44/2006/wp.24.e.pdf or http://edimbus.istat.it)

McKenzie R. (2006) : Improving timeliness for short-term economic statistics (paper presented at the Conference Q2006, Cardiff) (http://www.statistics.gov.uk/events/q2006/downloads/W00_mckenzie.doc)

Chapter 1 - The survey of the National Statistical Institutes about their practices concerning the balance between accuracy and delays in the statistical surveys

A questionnaire was sent to all Member States of the European Union in september 2006 to make a review of the real practices of the National Statistical Institutes concerning the question of balance between accuracy and delays in the statistical surveys.

21 answers were received, some with additional papers giving information about the methods used in Member States[1].

1.1 - Design of the questionnaire of the survey

The objective of the survey was to collect informations about the practices of Member States relatively to the questions of balance between timeliness and accuracy in the statistical surveys.

A questionnaire has been prepared and, after having been discussed with people in charge of statistical surveys within INSEE, sent to all Member States. It consists of four parts:

- a first introductive part is relative to the statistics produced at different times using statistical surveys, and makes a distinction between household surveys and business surveys on the one hand, and on infra-annual (i.e. monthly or quarterly) surveys and annual, or "one-shot" surveys, on the other hand ;

- the second part is relative to the methods of data editing being used, distinguishing data editing methods for one single questionnaire, selective editing methods (generally used to give priorities within the set of questionnaires to check, for example to edit manually just a part of them), and output editing (of aggregated data). The questionnaire asked for information concerning the time saved by the use of such methods ;

- a third part is relative to the methods used for the follow-up of the non-respondent units ;

- a fourth part is about the existence of an overall strategy used by the NSI concerning the choices to make to reduce the delays, and to produce statistics before the deadline : particularly, does the NSI use some indicators to define priorities for actions to be undertaken, do some "stopping criteria" exist ? This part of the questionnaire was the most interesting one for this handbook, and probably the most difficult to answer.

Annex 1 gives the final version of the questionnaire and the accompanying letter that has been sent with it.

1.2 - Main results of the survey

Annex 2 gives detailed results of the survey.

From a general point of view, the first parts of the questionnaire were filled in a better way than part 4 ("overall view of the survey") for which there was an important proportion of item non-response. Also, some respondents noticed that some words might be confusing (infra-annual surveys instead of monthly or quarterly, for example). However one may consider that the answers of the NSIs give a good overview of their real practices.

[1] Some of these documents have been used for the parts 2 to 4 of this handbook.

First element, **all NSIs of the Member States are concerned with the subject of delays and of production of results at different times, with different kinds of "products"** (statistics more or less detailed, definitive files, etc.), **probably more for business surveys**.

There is a strong difference between household surveys and business surveys : for the majority of the latter ones, there are strict deadlines, especially for short-term statistics, where the pressure for quick economic statistics is strong. The case of some household surveys (but not all of them) is somewhat different : there is often an objective to produce different kinds of statistics, some about distributions, rather than focussing on a few aggregates, and it is difficult to define a few "target" statistics. Also, files of individual data are produced, that will often be used by researchers for econometric analysis. Global statistics may be published first, and then, (generally six monthes to one year after), a complete individual data file is produced.

Also, the **methods of data editing** are different according to the distinction between household and business surveys, more than relatively to the distinction between infra-annual and annual (or one-shot) surveys.

For household surveys, generally speaking, the question of logical checks is important (relatively to the more important proportion of qualitative variables). This may be made through computer-assisted personal interviewing (CAPI), or computer-assisted telephone interviewing (CATI), so that the checks of data are close to their collection. Some softwares of data editing mentioned use the Fellegi-Holt method.

Business surveys use mainly consistency checks, outliers detection, and also logical checks. First, most of the informations collected through these surveys are quantitative, with strong links between them. Moreover, samples generally include the same units (at least for large enterprises) for successive periods, and the use of historical data (from the previous period : month, quarter, year, …) is possible. And the skew distribution of the universe of enterprises makes it important to detect outliers.

For both kinds of surveys, the edits applied in the method used are fitted regularly, using data from the previous period.

Very few Member States did consider that using data editing methods reduced the delays for the results of the survey : they generally said that these methods led first to improve the quality of the statistics produced, or to reduce the size of the data editing staff.

Selective editing is used more often for business surveys than for household surveys : as mentioned before, the skew distribution of the universe of enterprises is probably one of the main reasons for this situation. The most used way of selecting the units to check in a detailed way is to give priority to large units. In some cases, the calculation of a score (giving an estimate of the impact of the unit on a given statistic) is mentioned.

For household surveys, there is generally no selection of some units to check, but priority may be given to some variables ("core variables") that will be checked in a first time (this method may be used also for some business surveys). However, for some household surveys (for example wealth surveys), the situation may be close to that of business surveys, because of the skew distribution of the studied variables.

As for the question concerning the time saved by the use of "general" methods of data editing, there were few answers for the number of days, or monthes, saved by the use of selective editing : NSIs rather mentioned that these methods allowed to reduce the size of editing staff.

Output editing (checks of aggregated data) is also used more for business surveys than for household surveys. The comparison of aggregated data with those of the previous period is the most popular method, but comparison with other sources is also used.

For the **follow-up of the non-respondent units**, the situation is very different concerning business surveys, where the follow-up is focussed on some units (large units, or units having a large impact on estimates, or "long-term" non-respondents), and household surveys, where there is no distinction among the units of the sample. This question might even be a bit confusing concerning household surveys, that are very often face to face surveys, compared to business surveys that are more often mail surveys.

The part 4 of the questionnaire (**overall view of the survey**) was probably the most important. The results were not as meaningful as expected, since this part was rather less filled in by the NSIs than the previous ones.

Around one half of NSIs declare that they use a global list of indicators for the follow-up of their statistical surveys. This proportion is stronger for business surveys than for household surveys.

One NSI mentions a centralized system of standard quality indicators, including non-response rates, coding and imputation rates, size of revisions, etc. The response rate (in some cases weighted) is the most widespread indicator, the second one being the “coverage” coefficient of the sample. Some NSIs declare that indicators supply information on accuracy of the results that it is possible to produce at a given moment, by using only the questionnaires received and edited.

Concerning priorities for the actions to undertake at a given moment of the survey, taking into account delays and cost aspects, half of the NSIs declare to have such priorities for business surveys, whereas it does exist only for one third of the NSIs for household surveys. The priorities are generally given according to the size of the units, or to the impact on the coverage of an economic sector (in this case the priority may be given first to a more global level, i.e. a sector, and then to enterprises belonging to this sector). One NSI mentioned the fact that, for some surveys and on the last days of data processing, response chasing can be reduced and focus is put on clearing errors. Another one uses thresholds, for selective editing, that are reviewed considering the available human resources and time.

The last question, about stopping criteria, got the same kind of answers than the question concerning the list of indicators for the following of the survey : less than half of the NSIs mentioned their existence, and the most popular method is the follow-up of the response rate.

1.3 - Conclusion

Despite the fact that the answers to the questionnaire were more or less complete (and that they were maybe dependent on the “understanding” of some words or questions), some general comments may be formulated.

The question of the delays is very significant for NSIs. Data editing methods are used by most of them, in a more or less sophisticated way. Particularly, selective editing is used, but in a rather basic way compared to some methods presented in “academic” papers (see chapter 2 for a description of some of them). Generally speaking, the follow-up of the non-response rates is the most usual method.

Considering a global approach of the problem, few countries mentioned using a centralized system of indicators which gives elements to produce choices between the different actions to undertake, taking into account the available resources (human resources, time, …). In most cases, only some basic indicators are used.

Chapter 2 - Methods to use when preparing and carrying out a survey

Implementing methods to produce accurate statistics with taking into account deadlines for the delays is studied in this chapter. Its three parts will be relative to three questions :

- For each component of the survey process, is it possible to save time ?
- How to use methods of data editing, which are a very powerful tool to reduce costs and delays ?
- Is it possible to have a global approach concerning the balance between delays and accuracy ?

2.1 - Redesign of the overall organization of the survey process

A first approach may be to search for excellence concerning every component of the survey relatively to timeliness aspects. When preparing the survey, its process may be shared in different steps. Looking at the quickest way to conduct every step may be a way to to minimize the global delays to produce statistics.

The following components of the survey process may be considered as important to save time :

- Selection of a sample of units ;
- Data collection through the questionnaire, and through the data collection method ;
- Follow-up of the non-respondents ;
- Data entry ;
- Data editing ;
- Production of estimates.

More detailed "breakdowns" of the survey process exist, for example the one of OECD given in the STES Timeliness Framework reference table[2] (OECD, 2005). This is not the objective of this handbook to review all elementary steps of a survey, and the following considerations will be focussed on the ones giving the opportunity to reduce the delays.

The sequence of components given above might appear as a sequential one. It is not the case, since some of these components may be "overlapped". Saving time has to consider this aspect.

In this sequence, the correction of non-response is approached in two different steps : when using imputation methods (data editing), and through the production of estimates (especially for the total non-responses).

Selections of a sample of units :

One factor to reduce the delays may be to reduce as much as possible the size of the sample (that means in particular that data processing will thus be lightened), but also to focus on the units considered as important for the estimates to produce, for example by using cut-off criteria for the sampling frame: this is the case for some short-term statistics as the industrial production index or the retail trade turnover and volume of sales indices (Eurostat, 2006), for which some European countries do not cover the whole population of enterprises, excluding small enterprises which may be numerous but represent a small economic share (for example concerning the total turnover). This has to be made by studying the impact of the part of the population that is covered by the survey on the estimates. Obviously, making this choice will be at the cost of not being able to produce detailed results for some groups, for example small enterprises.

[2] This database provides a very interesting bibliography about the different steps that are detailed.

Data collection :

From a general point of view, it is possible to save time on different components of the data collection.

The easiness to fill the questionnaire, as the publicity made for the survey, for example through trade unions, associations, etc., may be incentives for the units to give a response quickly to the survey.

Then, the use of new technologies may shorten the delays : for example CAPI (computer assisted personal interviewing) or CATI (computer assisted telephone interviewing), where it is possible to use some checks very close to the respondent. But generally additional checks (more complex) are necessary, and lead to the recontact of some of the respondents.

.

On-line surveys are also very promising. The use of electronic questionnaires reduces the delays of transmission, and allows introducing some basic checks carried out when the unit is filling the questionnaire. However, the question of the edits to use in these surveys is not an easy one. In their paper, Nichols et al. (2005) present some interesting issues (how many edits ? should failed data be allowed to be submitted ?, etc.) and guidelines on the subject. Another difficulty, with the use of this technique, is that it is often combined with classic data collection (for example a mail survey), because the proportion of units using internet to answer to the survey is far to be 100%. Producing statistics early may generally favour internet answers (compared to mail answers), favouring then ICT-consuming units which have specific characteristics. This has to be taken into account in the estimates.

For some surveys, switching the observation period is also possible, and may improve timeliness : this is for example the case for output prices indexes, for which some NSIs use the mid-period rather than the end-of-a-period (Werkhoven). However, this cannot be done for other statistics, as for example turnover index, and is very specific to a limited number of domains.

In some cases, the willingness to shorten the delays may have some negative consequences on the quality of estimates, by introducing biases : this is for example the case of household surveys conducted on a sample of "main homes" at a period when some households (particularly those composed of retired persons) may be in "second homes" ; it may be necessary, in such cases, to adapt the observation period to that category of units.

Follow-up of the non-respondents :

In some cases, especially mail surveys, the delays may be enlarged because some units take a long time to send back the questionnaires. It is possible to give priorities in the follow-up of these units, and to avoid the publication of statistics to be postponed : generally, priorities will be given to "large" units having an important influence on the statistics that will be produced. This question will be examined in more details in chapter 3.

From a general point of view, having conducted studies on the response behaviour may be profitable in a long-term standpoint. For regular surveys, remind of "periodic" non-respondents may be, for example, launched earlier than for other units. Also, studying the distribution (daily distribution, hourly distribution) of the efficiency of the calls (in case of a telephone follow-up) may be interesting : for example, some hours of the day lead to a more successful contact rate.

Data entry :

Data entry using some techniques of optical character recognition may be interesting. Also, automatic coding methods (see for example Conference of European Statisticians, 1997, chapter 6) may be time-saving, for example when the survey uses a CAPI data collection.

Data editing :

This part of the survey is one of the most important for the question of timeliness, and will be detailed in part 2.2 of this chapter. Many papers have been produced on this subject. The aim of the methods of data editing is to check the plausibility of the data, including the coherence between data within one questionnaire, and to produce, if necessary, a "correct" value through methods of imputation. The item non-response is also dealt with at this step.

This work may be done in a manual way, or in a more or less automatic way, and does not necessarily search for a perfect final database.The time that this part of the work will take is thus depending on the chosen method.

The efficiency of one method is generally improved when applying it for a long period (for example, several consecutive years) : the parameters used may be adjusted, taking into account the experience of the previous years, for example to improve the hit rates of the edits that have been implemented.

Production of estimates :

It is often possible to produce preliminary estimates based on an incomplete set of data for estimates at a “high level” of aggregates, if their quality is considered sufficient enough. Definitive results at a more detailed level will be published afterwards, and revisions in estimates have to be studied retrospectively and regularly. Chapter 4 examines this subject with more details.

The step of the production of estimates is also the step where, having a deadline for the results, the time remaining has to be used in the most efficient way. Very often, “global” statistical methods (compared to individual editing) are used, particularly concerning the total non-responses (for example, reweighting methods), and the choice of the method has consequences on the quality (bias, variance) of the estimates.

2.2 - The data editing process

Data editing is known as one of the most time-consuming parts of the survey process. Many studies have shown that it is not necessary to correct all errors to ensure that statistics will be of sufficient quality to be published (see for example Granquist and Kovar, 1997). Those studies suggest the use of different methods to save resources (time, human resources).

As mentioned before, the recommended practices handbook produced by the EDIMBUS project (Luzi, De Waal, Hulliger, 2006) is dedicated to the subject of edit and imputation in business surveys, and gives useful and detailed information on this topic.

The elements presented here will only give an overall approach of data editing, and focus on some methods or points that will be very useful for timeliness aspects (especially selective editing).

From a general point of view, as presented for example in De Waal and Coutinho (2005), there may be different steps in editing :

- receipt of observed data (with sometimes a first elementary control of data) ;
- selective editing, that shares the flow of data in two streams : one that will be checked manually, one that will be edited and imputed with an automatic method (in some cases, for this second flow of data, only raw data are used) ;
- manual or automatic editing, at a micro level (depending on the previous step, manual editing for one stream, automatic for the other one) ;
- final macro-editing, used as a validation of aggregates.

This combination of methods may be considered as an efficient one. From the point of view assumed for data editing and particularly in this handbook, the question is to shorten the delays while preserving the quality of the statistics. So, the objective is not to get the "perfect" final data, but rather to have the most efficient possible process to produce these statistics. Using selective editing, the work of manual editing, that will be focussed on a subset of units, will be conducted in a more efficient way.

In a first part, principles relative to automatic edit and imputation methods will be presented. Then the handbook will mention some selective editing methods that focus on units having the biggest impact on target statistics.

Final macro-editing (sometimes called output editing) will not be studied in a detailed way. This kind of method consists in studying the "quality" of aggregates by comparing them to "external" statistics (that may be, for example, the same statistics for the previous year). It may allow to detect some errors that were not found by other methods, for example for variables that were not considered as "target variables" when implementing the editing (including selective editing) process. Also, some systematic misunderstandings of questions (for example, concerning what has to be included in wages) may be detected at this step, leading to additional treatments, or to the decision of not publishing estimates for one variable.

2.2.1 - Automatic edit and imputation methods

The aim of these methods is to detect errors in the data, and to propose corrected data (imputation). Using these methods may be very efficient.

Different kinds of edits may be defined :

- logical edits (generally between categorical variables);
- accounting balance edits (generally for totals) ;
- statistical edits (for example the ratio between two variables has to belong to a given interval) ; that kind of edit is often used to control the evolution of the value of a variable compared to that of the previous period.

Fellegi and Holt (1976), in their landmark paper, proposed a method for solving the error localisation problem, by generating implicit edits (explicit edits being defined by the statistician) and, then, considering the set of explicit and implicit edits needed to correct the data, for finding the minimum number of variables to modify. Bankier proposed another method (NIM, New Imputation Methodology, that became then Nearest-Neighbour Imputation Methodology, see for example Bankier, Luc, Nadeau, Newcombe, 1996 or Bankier, Poirier, Lachance, 2001) that searches first for nearest-neighbour donors and then determines the minimum change imputation action associated with these donors.

These methods are used in softwares developed by different NSIs, as DISCRETE (Bureau of Census), CANCEIS (Statistics Canada), GEIS (Statistics Canada), SPEER (Bureau of Census), CHERRY PIE (Statistics Netherlands), etc. This list is of course not exhaustive, since the objective of this handbook is not to have a detailed presentation of all these methods. It should be noticed that the choice of the method is depending on the main objective of the survey concerning the estimates : if aggregates are the main expected results, rougher methods may be used than when the results are concerning distributions, or disparity statistics (for one example of an approach on this topic, see Gautier, 2006). More information about these methods and softwares may be found in the EDIMBUS handbook, on K-Base, or on the Euredit website for the associated softwares.

But other methods may be implemented : for example, French enterprise annual surveys use a self-developed software that is based on the following method (Rivière, 1997). Different kinds of checks are defined for every variable of a questionnaire : internal-account edits, time-likelihood edits, internal-likelihood edits. Every check concerning one variable will produce a “mark”, and combining all marks relative to this variable, the value of the variable will be “qualified” or not. Then, if the value of the variable is considered as “not qualified”, the software tries different methods of imputation (using available data, for example historical data), until one is found to work. If no imputation is possible, there is an editing failure. The principle is to have a process as automatized as possible, to avoid overediting, rather than to search for perfect data.

In this procedure, one has to define a hierarchy of variables, and to rank them into groups. The first group that is controlled is composed of six variables. This group is edited and imputed first, then it is the turn of the lower-ranking groups. Then, the whole questionnaire gets a global quality code. The value of this code, and the messages produced for variables concerned with edit failures, give information to survey clerks on the work to do for this qestionnaire.

During all this process, internal meta-data are produced : mark of each control, quality code of each variable, quality code of each questionnaire, rough value of each variable, “modified” value (value proposed by the survey clerk, if necessary) of the variable, definitive value of the variable. All these data are very important : they may be used to check the efficiency of the data editing process : for example by seing the differences between rough, modified and definitive values, or by observing the distribution of the marks concerning the different variables.

Whatever the method used, an evaluation, conducted at regular periods, of the efficiency of the data editing process is very important. Di Zio, Luzi and Manzari (2005) give interesting ideas on the subject, by suggesting to study “*the number of passing and failing records, the edits failing frequencies, the imputation frequency and the edits mostly involved in the imputation process with the corresponding frequency of such involvment, 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)*”.

Martin & Poirier (2002) present what has been done for the Canadian unified enterprise surveys to evaluate the survey processing, especially relatively to the edits. In general, adjusting the parameters of the data editing

process may be very profitable. It has to be done by using the data observed on the survey of a previous period. So, it is necessary that the data editing software “stores” some meta-data relative to the results of the different edits, to be used for further studies.

Time should be kept for this assessment work, that should be conducted in a systematic way.

As mentioned in many papers, an important “by-product” of data editing does exist : the “reassessment” of the questionnaire of the survey. The results of some edits may mean that the way one question was asked was not adequate, and sometimes lead to the deletion of this question, or to the change of the wording used. **The role of data editing has thus to be considered as broader than the strict error localisation and correction.**

A difficult question is the one of the definition of the “best” edits. According to the litterature, it seems that there is no “perfect” method to define appropriate edits, and that these edits generally rely on the opinion of subject-matter specialists. However, through statistical analysis of links existing between variables, it is possible to get elements, for example to choose one variable that will be considered as interesting to control the plausibility of the value of another one.

2.2.2 - Selective editing methods

Selective editing is a very efficient way to save time while maintaining a good accuracy. Different methods exist. As seen with the survey of the NSIs, they are used in a more intensive way for business surveys than for household surveys. **For most of the answers of the NSIs, the priority given to the units to be edited in a detailed way is given according to their size, the biggest units being considered as the most influential ones. However, some more refined methods have become popular for the last years**, based on the computing of "scores" for each unit depending on the data it has provided.

A first kind of method was presented by Latouche & Berthelot (1992) as the DIFF function and then developed by other authors, and relies on the difference between the raw data provided by the unit and an "expected" value, that will be used as an estimate of the potential impact of the unit on an aggregate. This kind of method will be presented in a first step, and then other methods will be presented. In a third part, practical aspects of the implementation of these methods will be tackled.

2.2.2.1 - DIFF function and related methods :

The basic idea of this kind of method of selective editing is to give priorities to some units for being edited in a detailed way by ranking first the units and determining a "threshold" dividing the whole set of units into two parts, one to be edited in an automatic way (or, in some cases, to be used with the raw values that have been provided), the other one being edited manually.

Different papers are relative to this method, especially those of Latouche & Berthelot (1992), Lawrence & McKenzie (2000) (using the wording "significance editing" for selective editing), Hedlin (2003).

The approach consists in different phases :

- definition of target statistics ;
- computation of a score for each target statistic (local score, sometimes called item score, if corresponding to a statistic produced with one variable) ;
- computation of a synthetic score (global score) that will be used to rank the units ;
- choice of the value of a threshold for the global score.

How to produce local scores ?

The basic idea, if the target statistic is the estimation of the total $T(Y)$ of a variable Y, is to compute the following value, giving the "expected" impact of editing on the estimation, in case of a survey using a sample :

$$w(i)(z(i) - y(i))$$

where $w(i)$ is the weight of the unit i
$z(i)$ is an "expected" value of the variable Y for unit i
$y(i)$ is the raw value of the variable Y for unit i given on the questionnaire.

The "difficulty" is to get a value for $z(i)$. This value will be used to compute the score, and so to rank the unit (compared to other units), but not to compute statistics. So, the best "proxy" for the expected value has to be searched, but it does not need to be a perfect one. It may be, when available, the value of the previous period, or an estimation using an auxiliary variable (by applying a ratio computed on a set of units), or a rougher estimate, as the mean or the median of the variable (computed on a set of units for which the value is available). It may also be given by an automatic data editing software that will produce an imputed value (if the raw data is considered incorrect, otherwise the expected value may be equal to the raw value).

The target statistics may be more complex that the estimation of a total (for example a function of totals). In this case, a linearization method will be used. For example, if the estimation is relative to a ratio (for example turnover per employee, estimated as the ratio $T(Y)/T(X)$, where $T(Y)$ is the estimation of the total of the turnover, and $T(X)$ the estimation of the total of employees), the local score will be computed as (Lawrence & McKenzie, 2000) :

$$\frac{1}{T(X)} w(i)\left[(ye(i) - yr(i)) - \frac{T(Y)}{T(X)}(xe(i) - xr(i)) \right]$$

where $ye(i)$ is the expected value of the variable Y for unit i (the same for X)

$yr(i)$ is the raw value of the variable Y for unit i (the same for X)

For the values $T(X)$ and $T(Y)$, values from the previous period may be approximates of sufficient quality.

How to produce a global score using the local scores ?

Since units (questionnaires) need to be "treated" (automatically or manually) on a "unit by unit" basis (and not item by item), it is necessary to combine the local scores in a global score. Scaling variables that may use different kinds of "measures" (for example, monetary variables) is important, otherwise some local scores might be too prominent in the global score.

One kind of scaled local score proposed[3] is :

$$local\_score = \frac{w(i)(z(i) - y(i))}{T(Y)}$$

where $T(Y)$ is an estimation of the total of Y, that may be obtained in some cases with the value of the previous period (if, for example, too few questionnaires have been received by the Statisical Institute to produce a robust estimate).

Some papers (for example Lawrence&McKenzie, 2000) propose to use, for the denominator, the standard-error of the estimate of the total, instead of the estimate of the total (in order to refer to the accuracy of the estimate rather than to the estimate itself).

Then, local scores have to be combined. One may use different methods of aggregation, for example :

- maximum of the local scores
- Euclidean distance, with the possibility of introducing weights α_j for each local score j, depending on the relative importance given to each local score :

$$S_i = \sqrt{\sum_j \alpha_j S_{ij}^2}$$

S_{ij} being the local score for unit i and item j , and S_i its global score.

[3] One may note that this value gives the « contribution » of the unit to the evolution in case the data from the previous period is used as expected value.

How to use this global score ?

The basic idea is to calculate the value of a threshold (cut-off score) using a simulation study. For this, one needs data coming from a former survey to be available, giving raw and definitive data (after editing).

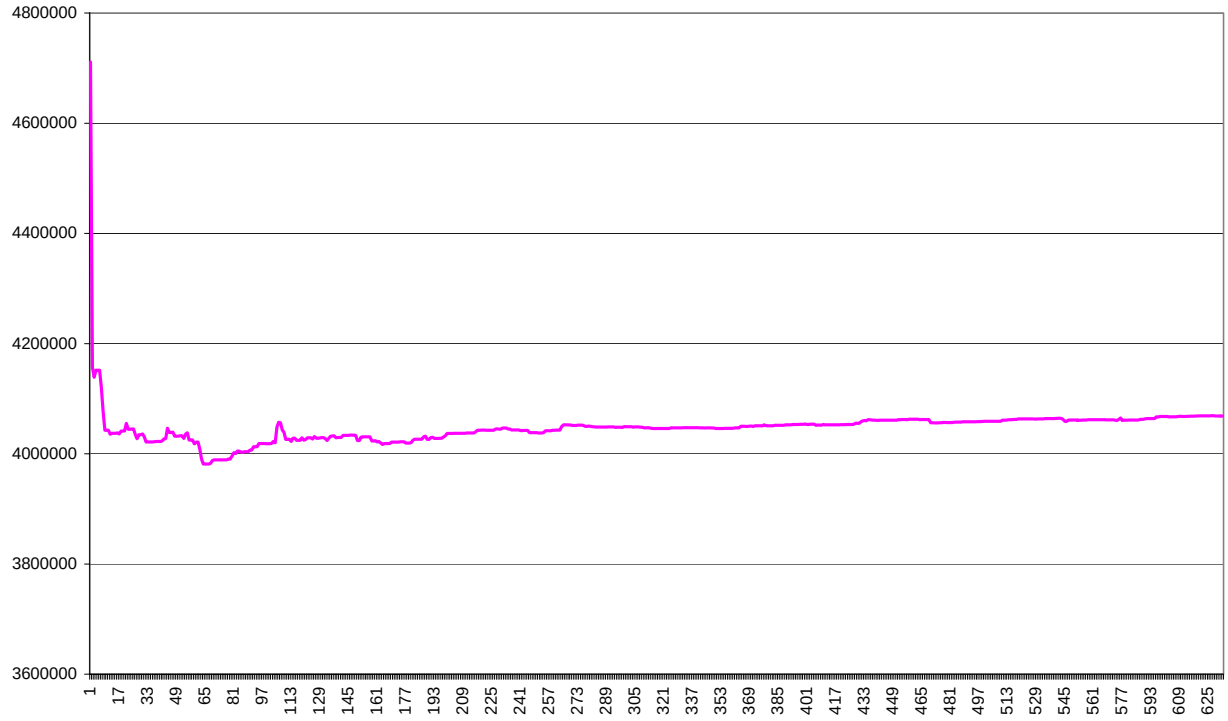
Figure 1 presents the result of a simulation conducted on French annual enterprise survey data. It gives the evolution of the estimate of total turnover for an economic sector (retail trade), the units (enterprises) being ranked according to the value of a score that has beeen computed, and raw data being used for the non-edited units whereas definitive data are used for the manually edited units. As a result, on the left part of the figure, few units are edited and the estimate is mainly based on raw data, while on the right part most of them are edited. More precisely, the point of rank k on the graph (position k on the abscissa) has for ordinate the value of the statistic obtained when calculating it with data manually edited for the k first units (according to the order given by the score) and the raw data for all other units.

This figure shows that, in this case, editing less than half of the units would be enough to produce a robust estimate. Comparing the evolution of the value of the estimate to the expected accuracy of the survey (when this information is available) is also important : if, on the right part of the figure, the estimate remains in an interval whose magnitude is, for example, less than 10% of the standard-error of the estimate, one may consider that continuing editing the questionnaires is not necessary. This kind of simulation has to be done for aggregates, but also for sub-aggregates (at a detailed level, for example) if necessary.

The value of the threshold that will be used for the following surveys is then deduced from the figure.

Figure 1

Example of a simulation conducted on data of the French annual enterprise survey for retail trade



2.2.2.2 - Other methods :

Two methods that do not need the production of an expected value of the variable to compute the value of a score are presented here.

• Sometimes, the contribution of a unit to a statistic is calculated as the value that the statistic would have if the unit was missing within the sample of the survey and the statistic was obtained with this sample "minus" this unit. The statistics for which the use of this method is possible are ratios, including the evolution of an aggregate compared to the one of the previous period. For example, the contribution of unit $i0$ is computed, for the evolution of an aggregate, as :

$$\left| \frac{\sum_{i \neq i0} w_i X_{i,t}}{\sum_{i \neq i0} w_i X_{i,t-1}} - \frac{\sum_i w_i X_{i,t}}{\sum_i w_i X_{i,t-1}} \right|$$

This method is sometimes called the **"drop-out" method**.
To calculate this score, one needs to have a raw data for every unit of the survey (in order to be able to calculate the aggregate), or at least an estimation of this raw data, whereas the DIFF function can be calculated as soon as the first questionnaire is sent back, even if all other raw data are not available.

• Another method was proposed, for periodic surveys, by **Hidiroglou & Berthelot** (1986).

The principle of this method is :

- to compare, for each unit, the ratio of the data of the current period to the data of the previous period : $R_i = \frac{Y_{i,t}}{Y_{i,t-1}}$ to the median of the ratio of the category of the unit, and to transform it into :

$$S_i = 1 - \frac{R_{med}}{R_i} \quad \text{if} \quad 0 < R_i < R_{med}$$

$$S_i = \frac{R_i}{R_{med}} - 1 \quad \text{if} \quad R_i \geq R_{med}$$

- to calculate the Hidiroglou-Berthelot (HB) statistic as :

$$S_i * \left[\max(w_i Y_{i,t}, w_i Y_{i,t-1}) \right]^u$$

- to study the distribution of the HB statistic, through its median and quartiles, and to define the interval, for this statistic, outside of which a unit will be considered as to check in a detailed way.

The HB statistic takes into account, at the same time, the evolution of the unit compared to the median one, and the size of the unit (multiplied by the sample weight). The parameter u, between 0 and 1, is adjusted to give more or less influence to this size.

The methods presented before may be combined in a more refined process.

For example, Jäder & Norberg (2005) propose for the Swedish foreign trade statistics a score combining a "suspicion" (using the Hidiroglou Berthelot method) for the unit price and a potential impact on aggregates.

Revilla (2002) presents a method of selective editing for the Spanish industrial production index based on the use of time series modelling (ARIMA and regression models) to give a forecast of the index and to compare it to the observed one.

It is also possible to use a two-step "drill-down" approach, studying first sub-levels (that means "estimations cells", for example at detailed levels of the NACE) and **giving priorities to some sub-levels to be checked first, and searching then for influential units within these sub-levels**.

2.2.2.3 - Practical aspects of the implementation of these methods :

From a general point of view, these methods are more adequate for business surveys, because of the skew distribution of enterprises. But for some household surveys (for example about wealth), the same kind of distribution may exist, and this kind of methods may be interesting.

Their implementation needs that data from a former survey are available, giving raw and definitive data for each unit of this former survey.

Methodological studies will enable to determine adequate values for the parameters that have to be tuned. In other cases, when former data are not available, or time is not sufficient to conduct these studies, rougher methods will be used, as for example giving priorities to large units (making this way the hypothesis that the potential impact of the raw values is linked to the size of the unit). This kind of practice seems to be used in many cases by NSIs, as observed in the survey that was conducted for this handbook.

The definition of target statistics is essential :

- **target variables ;**
- **level of publication (national / regional, global / detailed for categories of households or enterprises).**

The efficiency of the method depends on the level of the statistic to produce : at a sub-level (for example regional), one will have, very often, to check manually a proportion of units more important than at the global level. A unit classified as to be checked manually at a regional level will not necesseraly be classified in the same category at the national level.

Some technical difficulties when implementing this kind of method

Whatever the method used, tuning it is not immediate : for example for the DIFF method, studies have to be made to determine the "best" expected value for every variable, and then the "best" construction of a global score - by choosing a distance function -. Often, a lot of variables are concerned, and some authors recommend to use, for the global score, an Euclidean distance in this case (Farwell, 2005). They recommend also, sometimes, to group the variables that are correlated in order to limit the number of local scores.

Variables with a high frequency of "zero values" (concerning for example an expense that is rare) may be given, relatively, higher scores, and so are favoured, at the cost of disadvantaging other variables. The method is best adapted to "continuous" variables. "Volatile" variables, for example investments of enterprises, for which the value may be very different from one period to another, are more difficult to control.

The choice of the expected value might seem easy for regular surveys, by using the data of the previous year : but often, the sample has been partially renewed, or there were some non-responses on the previous period, so that, for one part of the units, expected values have to be "created" in another way (using a ratio applied to a variable available in the sampling frame, for example). Also, some choices have to be made concerning the raw data of the current period that may be missing because of item non-response ; in this case one may use, for example, the result of the imputation made by an automatic data editing software (imputed data will be used for the estimates traced on the graph for the units non edited manually).

Finally, studies preparatory to the use of selective editing are often time-expensive. But they generally do reveal the practical difficulties that will be encountered in the "real world" of the survey process.

One last point is concerning the continuity of the units. In some cases, there are discontinuities for units, for example for enterprises when they are restructuring. These cases cannot be taken into account with automatic procedures. They should be systematically treated in a manual way, at least for large units (especially large enterprises).

Organizational considerations

This kind of method gives an estimate of the proportion of units to be checked manually. However, this number may be variable from one year to the next one, and one has to adapt, when conducting a survey, to the characteristics of the answers (that may be of more or less good quality). Chapter 3 will consider this question.

The implementation of this method has also consequences on the organization of the staff working on data editing. When there is no selective editing, the decision of giving one category of units (for example, an economic sector) to a survey clerk is taken at the beginning of the survey process. Using selective editing requires some flexibility within the data editing staff, since the decision of checking manually a questionnaire or not may not be known in advance. These "cultural" aspects are very often difficult to deal with, and should not be considered as negligible. They are developed in part 2.3.2.

Using output editing

As said before, selective editing does not check all kinds of errors concerning the whole set of variables (except for surveys with a very limited number of variables). It happens that some variables are not taken into account in the selective editing process (otherwise, the global score would need to combine a number of variables so important that, finally, the threshold might classify most of the units as to be edited manually). Using final output editing (by checking aggregates, often by comparing them to former values) is always efficient. As mentioned in Rivière (2002), this part of the process is not well formalized at the present time, and may take time ; also, it seems difficult to propose methods to save time for this step of the process.

The necessity of a regular assessment of the method

Selective methods need to be assessed regularly, since otherwise there could be a risk that changes in the response behaviour might turn them into less efficient ones. To do this, it may be useful to keep a part of the sample (renewing it every year, for example) edited in a complete way, to get some data useful for methodological studies (with complete raw and edited data). Also, this option has a consequence on surveyed units : it sends a message from statisticians to the surveyed units about the "expected quality" of the answers (otherwise, the feeling that the data are not controlled could emerge among respondents who are never called back, with long-term consequences).

2.3 - Having a global approach

2.3.1 - Considering the question of timeliness as an optimization problem

What has been presented in the previous parts of this chapter is relative to a view of each component considered separately. **It is then necessary to have a global planning of the whole survey process**. A PERT chart may, for example, be used for this purpose. It gives the critical path method, that minimizes the total time spent to produce the results.

But, **minimizing the time may be viewed as a more general problem**, with parameters whose values may have consequences on the total time. In that sense, this problem is close to the economic question of producing statistics with an optimal use of available resources, and close to a "total survey error" approach.

The question of the application of this kind of approach for cost aspects has been considered by some authors : for example, is it possible to take into account the costs of the different components of the process, and to use a global approach to minimize the total cost, the mean square error of some statistics being defined, or to minimize the mean square error, the total cost being given ?

In practice, it is necessary to have information on the cost and time needed for each elementary component of the process. Few answers have been given to this question.

Destatis has developed IT-tools (Wein, 2004) for the planning of German statistical surveys, especially for the data editing process. The data editing process is broken down in different sub-processes, and the IT-tools give elements about the methods, but also about the estimated cost and time needed by every elementary activity. This information may be used by persons in charge of statistical surveys.

Linacre & Trewin (1993) give some elements for one survey, the Australian construction industry survey. They pay a particular attention to the quality of the estimates of the unitary costs of each component of the process, and to the robustness of the conclusions that may be obtained using these elementary costs. They underline that simulations of the sensitivity of the results may be helpful.

In a more general paper, Platek & Särndal (2001) ask the question : "*Creating a functional total survey error model: an impossible dream?*". They explain that very often, statisticians focus on one single error, the sampling error, for which theoretical literature is available, and neglect other kinds of errors. For the subject of this handbook (timeliness), similar considerations do certainly exist. One may wonder if some researches could be started on this subject, by using information about the mean time needed for every elementary action, and referring to a global approach.

For example, from a timeliness point of view, is it better to have a large sample with a relatively light follow-up of non-respondents, or a smaller sample with a more intensive follow-up ? Is it better to spend time to continue the follow-up of the non-respondents, or to continue editing the questionnaires that have been received ?

This kind of studies shoud use, besides elements concerning the time spent for every elementary action, some information about the response rates and about the probabilities of success for the follow-up. Once having defined the parameters that can be modified (size of the sample, proportion of units to check manually, volume of follow-up with a visit, volume of follow-up by mail, etc.), one could specify a model as :

$$Time - spent = F(parameters)$$

and minimize the time spent for a given level of accuracy (that gives not necesseraly the minimization of the cost).

It is important that the “target statistics” are properly defined. Linacre & Trewin (1993) underline, for example, the fact that searching for statistics of evolution, or “level” statistics, will not lead to the same choices (for example concerning the rate of covering of the sampling frame).

At the present time, it seems that very few experiences do exist on that topic.

2.3.2 - Considering the organizational aspects

Organizational aspects are important especially concerning the data editing staff.

First, survey clerks are generally specialized on categories of units (for example economic sectors), having an extensive knowledge of these categories. Applying selective editing methods leads to a more flexible organization : observing the response rates and the “quality” of the answers of each category of surveyed units may lead to priorities given to some categories (for the follow-up of non-respondents, or for the manual editing). It is then necessary that the data editing staff gets a more extended knowledge of all sectors they can happen to work on, and, more generally, it is essential that this staff does support that kind of method. It is also necessary to implement practical tools giving priorities for the actions to undertake (for example through marks given to the questionnaires), these tools being clear (giving messages concerning the reasons why one unit is to be edited in a detailed way, for example), and the intermediate executive level having a good understanding of them.

Also, very often, survey clerks do not work on one single survey, and their time is shared between different operations. What could be considered as an optimal scheme for one survey is not necessarily the best one taking into account this aspect.

Then, it is essential to get information about the distribution of the “answers” of the units all along the year, in case of a mail survey : if, for example, the first questionnaires that are returned (at the beginning of the survey) are of “good” quality, they will not be considered for selective editing, and the quantity of work dedicated to the data editing staff will be very small, as it could be too important later on. This question of the “adjustment” of the quantity of work during the year is sometimes one of the most difficult. It may be considered by using old databases where the information about the time of delivery of the questionnaire is available.

During the survey process, time may not be considered as “linear” : very often, the NSI does not control the dates of return of the questionnaires and has to adapt to this situation, particularly a few days, or weeks, before the deadline.

Executive summary for chapter 2

1. Every component of the survey process has to be independently reviewed, to search for the most time-saving device.

2. As far as possible, an automatic data editing method has to be used. Information about existing methods may be found in some web sites, or in handbooks as the one resulting from the EDIMBUS project.

3. Selective editing methods are very powerful methods to produce results quickly by focussing on what is important for their accuracy. They are well fitted to some categories of units / variables, and less to other categories. Their implementation needs extensive methodological studies. By default, a rougher method may be applied in some cases, as giving priorities to large units, or focussing on "core" variables.

4. Time should be kept for a regular assessment of the methods used (automatic data editing, selective editing). Having meta-data available in the databases (about the data, particularly their change from raw data to definitive ones) is essential.

5. The support of the data editing staff to these kinds of methods is very important. The way these methods do include in the global organization has to be thoroughly studied.

References for chapter 2 :

Bankier, M., Luc, M., Nadeau, C., Newcombe, P. (1996) : Imputing numeric and qualitative variables simultaneously, proceedings of the section on survey research methods, American Statistical Association (http://www.amstat.org/sections/srms/proceedings/papers/1996_013.pdf)

Bankier, M., Poirier, P., Lachance, M. (2001) : Efficient methodology within the Canadian Census edit and imputation system (Canceis), proceedings of the annual meeting of the American Statistical Association (http://www.amstat.org/sections/srms/Proceedings/y2001/Proceed/00202.pdf)

Conference of European Statisticians (1997) : Statistical data editing - volume 2 - methods and techniques, United Nations, New York and Geneva (http://amrads.jrc.ec.europa.eu/k-base/)

De Waal, T., Coutinho, W. (2005) : Automatic editing for business surveys : an assesment of selected algorithms, International Statistical Review, volume 73, N° 1

Di Zio, M., Luzi, O., Manzari, A. (2005) : Evaluation editing and imputation processes : the Italian experience, UN/ECE Work Session on Statistical Data Editing, Ottawa (http://www.unece.org/stats/documents/2005/05/sde/crp.4.e.pdf)

EUREDIT project : http://www.cs.york.ac.uk/euredit/

Eurostat (2006) : PEEIs in focus - A summary for the industrial production index - A summary for the retail trade turnover and volume of sales indices

Farwell, K. (2005) : The general application of significance editing to economic collections, Methodology Advisory Committee, Australian Bureau of Statistics (http://www.oecd.org/dataoecd/24/52/34976185.pdf)

Fellegi, I., Holt, D. (1976) : A systematic approach to automatic edit and imputation, Journal of the American Statistical Association, N°71

Gautier, E. (2006) : Detection of inconsistencies and outliers and multivariate imputation in the French 2004 wealth survey, paper presented at the conference Q2006, Cardiff (http://www.statistics.gov.uk/events/q2006/downloads/T12_Gautier.doc)

Granquist, L., Kovar, J. (1997) : Editing of survey data : how much is enough ? In Survey Measurement and Process Quality, Eds Lyberg et al., Wiley

Hedlin, D. (2003) : Score functions to reduce business survey editing at the U.K. Office for National Statistics, Journal of Official Statistics, Vol. 19, n°2 (http://www.jos.nu/)

Jäder, A., Norberg, A. (2005) : A selective editing method considering both suspicion and potential impact, developed and applied to the Swedish foreign trade statistics, UN/ECE Work Session on Statistical Data Editing, Ottawa (http://www.unece.org/stats/documents/2005/05/sde/wp.12.e.pdf)

Hidiroglou, M., Berthelot, J.-M. (1986) : Statistical editing and imputation for periodic business surveys, Survey Methodology, n°12

K-BASE : The knowledge base on statistical editing http://amrads.jrc.ec.europa.eu/k-base/

Latouche, M., Berthelot, J.-M. (1992) : Use of a score function to prioritize and limit recontacts in editing business surveys, Journal of Official Statistics, Vol. 8, n°3 (http://www.jos.nu/)

Lawrence, D., McKenzie, R. (2000) : The general application of significance editing, Journal of Official Statistics, Vol. 16, n°3 (http://www.jos.nu/)

Linacre S., Trewin, D. (1993) : Total survey design - application to collection of the construction industry, Journal of Official Statistics, Vol. 9, n°3 (http://www.jos.nu/)

Luzi, O., De Waal, T., Hulliger, B. (2006) : Recommended practices for editing and imputation in the European statistical system : the EDIMBUS project, UN/ECE Work Session on Statistical Data Editing, Bonn (http://www.unece.org/stats/documents/ece/ces/ge.44/2006/wp.24.e.pdf or http://edimbus.istat.it)

Martin, C., Poirier, C. (2002) : Analysis of data slices and metadata to improve survey processing, UN/ECE Work Session on Statistical Data Editing, Helsinki (http://www.unece.org/stats/documents/2002/05/sde/2.e.pdf)

Nichols, E., Murphy, E., Anderson, A., Willimack, D., Sigman, R. (2005) : Designing interactive edits for U.S. electronic economic surveys and censuses, UN/ECE Work Session on Statistical Data Editing, Ottawa (http://www.unece.org/stats/documents/2005/05/sde/wp.22.e.pdf)

OECD, McKenzie R. (2005) : Improving timeliness for short-term economic statistics, working paper STD/DOC (2005) 5 (http://www.olis.oecd.org/olis/2005doc.nsf/809a2d78518a8277c125685d005300b2/7874b23a303d534ac12570df004e54ac/$FILE/JT00196395.PDF)

Platek, R., Särndal, C.-E. (2001) : Can a statistician deliver ? Journal of Official Statistics, Vol. 17, n°1 (http://www.jos.nu/)

Revilla, P. (2002) : Approaches for improving timeliness in production index : a selective editing method, presented at the Business Statistics Directors meeting, Eurostat

Rivière, P. (1997) : The new annual enterprise surveys in France, Courrier des statistiques English series n°3, Insee (http://www.insee.fr/en/ffc/docs_ffc/cs78e.pdf)

Rivière, P. (2002) : General principles for data editing in business surveys and how to optimize it, UN/ECE Work Session on Statistical Data Editing, Helsinki (http://www.unece.org/stats/documents/2002/05/sde/16.e.pdf

Werkhoven T., CBS : Improving the timeliness of short-term statistics, available on http://www.oecd.org/dataoecd/23/41/30045545.pdf

Wein E. (2004) : Improvement of data editing processes, Proceedings of the Q2004 Conference, Mainz

Chapter 3 - Using a dynamic approach

If chapter 2 was devoted to a static approach, while studying the best methods to use for the whole survey process, chapter 3 studies the question of what happens during the carrying out of the survey. Different results will be observed concerning the response behaviour (including the "quality" of the responses), that lead to a more or less important quantity of work.

Keeping in mind that deadlines exist for the production of the results, some choices have to be made.

3.1. Definition of priorities of some actions, especially the follow-up of the non-respondents :

The methods used for giving priorities in the data editing process have been presented in chapter 2, their fitting to the situation at a given time will be studied in part 3.2 of this chapter (adjusting thresholds).

This part 3.1 is focussed on **the following of non-respondents**. This question is particularly important for mail surveys, for which delivery times are not well controlled.

As said before, historical studies concerning the response behaviour are always profitable. Also, it is interesting to study the efficiency of the reminders of the non-respondents : is there an "optimal" choice concerning the number of reminders ?

Then different kinds of actions exist for the following of non-respondents : telephone remind, mail remind, visit, etc. Their cost may be very different, and if, in a first step, there is often a follow-up of all non-respondents, the next actions have to be limited to some units, especially for timeliness reasons.

• Methods using the same principles than selective editing methods may be proposed. The difficulty is that data are not available for the non respondents of the current survey. McKenzie (2000) proposes to use the principle of the DIFF function to compute the following score for each non-respondent to a quarterly sample survey :

$$w(i)(z(i) - y(i))$$

where $w(i)$ is the weight of the unit i

$z(i)$ is the raw value of one variable for unit i

$y(i)$ is the value of the same variable that would result of an imputation process (using for example data of the respondents, and variables of the non-respondent that are available in the sampling frame) for unit i,

and to use, since the values $z(i)$ and $y(i)$ are not available, the values of the previous period (i.e. the previous quarter) as a proxy for the potential impact. The assumption made, that the result obtained with the data of the previous period is a good proxy, has to be checked regularly. Having defined a score function, simulations carried out on former files may help to define a threshold for the score, that will classify the non-respondents in two categories : those who need a follow-up and the other ones.

Very often, rougher methods are used : the follow-up is focussed on the larger units (moreover, it may be focussed on identified long-term non-respondents, so it is possible to send an early reminder). The underlying assumption used in this case is that the potential impact of the non-respondents is linked to their size. The follow-up may also be focussed on units considered as likely to have large evolutions, if some characteristics (for example, belonging to one category) are considered as good indicators of it.

Daoust (2006) proposes a method for the Canadian quarterly survey of financial statistics for enterprises, based on ranks defined within industry groups (using the importance of the enterprise to coverage targets), and an "integration" of the different industry goups to produce a score. This method is used in a dynamic way.

• For small-size units, another method may be applied[4], by distributing the follow-up of non-respondents not on identified units, but first on some strata of the survey (Brion, 2004). For example, for the estimation of an aggregate and if the questionnaires are not totally returned and edited, the loss of accuracy within a stratum h may be quantified as the increase of variance implied by non-response, under the assumption that non-respondents do not have a specific behaviour (so that no bias is expected) and if re-weighting procedures are used :

$$\Delta V_h = \frac{N_h^2 S_h^2}{n_h}\left(\frac{1}{x_h} - 1\right)$$

where : N_h is the number of units of stratum h

n_h is the number of sampled units in stratum h

$x_h = \frac{r_h}{n_h}$ is the rate of « usable » questionnaires (i.e. returned and edited) in stratum h

S_h is the standard-deviation of the concerned variable inside stratum h .

By studying this increase of variance stratum by stratum, it is possible to decide :

- if some strata have to be favoured for the follow-up of non-respondents (eventually by introducing rates of "success" of the follow-up) ;
- if, comparing the increase of variance added up on all strata to the expected variance (with no non-response), it may be considered as small and insignificant (for example, less than 5%) so that an aggregate may be published without calling back non-respondents.

This approach is more usable on strata of small and homogeneous units, compared to the first one. For the values S_h to use, it may be better to use information coming from the previous survey than computed with the first responses, at least at the beginning of the survey (these values might be unstable). As mentionned before, the use of this method is based on the hypothesis that there is no specific behaviour for the non-respondents. It is thus necessary to check this hypothesis by using variables available in the sampling frame.

3.2. Using a list of indicators :

At a given time, it is necessary to make choices concerning the use of the available resources, deadlines being fixed for the production of the results.

• Having an information system with a list of indicators, which may be used for a continuous follow-up of the survey, is necessary.

This information system has to provide a list of indicators giving at least :

- the response rates by categories of units (it is interesting, at least for business surveys, to have the weighted and unweighted rates) ;
- the coverage rate concerning some target variables ;
- the "item" response rates for categories ;
- the "item" imputation rates.

These information have to be given at regular times (daily, once a week, etc., depending on the frequence of the survey). For example, ISTAT, the Italian NSI, calculates different kinds of rates every day for the monitoring of

[4] For small-size units, the first method may also be used.

some household surveys. And a centralized system of indicators, SIDI (Della Rocca, Luzi, Scavalli, Signore, Simeoni, 2003) is available for business surveys.

• Estimates of the accuracy of provisional estimates that could be produced, at a given time, with the available data (responses that have been edited), are also very useful. They may help deciding if it is possible to produce early results, and also to observe what is the share of each category of unit to the error (for example, to the mean square error). To compute that kind of indicators is not easy, because they need statistical techniques that are difficult to fit in a rather quick approach : probably approximations have to be made at this step, to produce rough estimations. Also, for provisional estimates, it is important to check if the bias of the estimates may be considered as insignificant, accuracy being very often regarded only through variance aspects.

Being able to produce estimates at any time of the process survey is also important. Historical studies showing the "convergence" of estimates towards the final value may be helpful to improve the global understanding of the survey process.

• **Continuous information about the non-respondents** is also essential. Using variables available in the sampling frame may help to know if the response mechanism may reasonably be considered as ignorable, or if some specific techniques have to be used. This work should be completed by more detailed studies concerning non-respondents, with specific "post-surveys" that would contact some non-respondents to check if the response mechanism is ignorable or not.

The follow-up of the characteristics (in the sampling frame) of the non-respondents is important for mail surveys, for which the delivery time is not controlled, but it is not specific to them. It is sometimes observed that, for household surveys made by face-to-face interview, surveyors prefer to let the most "difficult" questionnaires for the end of the survey. Producing preliminary results with the first received questionnaires would lead to some "uncontrolled" biases, if a monitoring of the characteristics of the non-responses is not available.

• **How to use such a list of indicators ?**

As mentioned in the survey of the NSIs conducted for this handbook, **one may have a simple follow-up of the indicators**, for example by focussing on response and coverage rates, and giving priorities to large units to increase the value of these indicators. Generally, the remaining available resources (staff, time) are used as adjustment variables. In some cases, units raising a problem of "continuity" (especially, in case of business surveys, large enterprises restructuring) should systematically be reviewed. Also, priority may be given to non-respondent "births".

One may try to have a more refined follow-up, for example by using the methods presented in part 3.1 of this chapter (for the follow-up of the non-respondents), or by refering to a global model as :

$$Mean - square - error = F(parameters)$$

the parameters being relative to some actions to undertake, as the follow-up of the non-responses, or the data editing of questionnaires that have been returned. Cost and timeliness elements may be associated to each parameter. The principle is to accept a loss of accuracy, but in such a way that this loss is as small as possible.

This approach needs to have defined what are the target variables, and what is the level of publication of the results (global / aggregated).

In practice, few examples of implementations of such approaches seem to exist at the present time. One is given in Merad and Wagstaff (2005), who present an information system intended to be used by the British Office for National Statistics (ONS) for business surveys. This system registers return and validation failures rates. Since selective editing is used for these surveys (with the computation of a score for every questionnaire, and thresholds that are used to decide if the questionnaire will be edited manually or not), and since, for some

surveys, the volume of validation failures has been higher than expected, this information system was used by the ONS to adjust the quantity of work dedicated to data editing by adjusting the thresholds.

It was made by computing the expected number of questionnaires still “to arrive” (using an expected return rate), estimating the number of man-hours needed, and comparing it to the available resources. A new set of thresholds had to be computed (by minimizing a distance function between former and new thresholds, having constraints about the resources and the “new” accuracy, this being made on two groups of variables, considering it was computationally more efficient to reduce the number of variables used to determine this new set of thresholds).

Executive summary for chapter 3

1. An information system including a list of indicators (response rates, coverage rates, etc.) should be used, both for the follow-up of the non-respondents and the data editing process.

2. The follow-up of the characteristics of the non-respondents has also to be monitored in a continuous way.

3. The list of indicators may be used in a more or less "rough" manner, by studying the available resources at a given time (considering the deadline for the results). In many cases, priorities are given to large units to improve some indicators (response rates, coverage rates), but more refined approaches could be used.

4. Being able to calculate, at a given time, the expected accuracy of estimates helps to decide if some results may be published, at least at an aggregated level.

References for chapter 3:

Brion, P. (2004) : The management of quality in French business surveys, Proceedings of the Q2004 Conference, Mainz

Daoust, P. (2006) : Prioritizing follow-up of non-respondents using scores for the Canadian quarterly survey of financial statistics for enterprises, UN/ECE Work Session on Statistical Data Editing, Bonn (http://www.unece.org/stats/documents/ece/ces/ge.44/2006/wp.20.e.pdf)

Della Rocca, G., Luzi, O., Scavalli, E., Signore, M., Simeoni, G. (2003) : Evaluating, monitoring and documenting the effects of editing and imputation in ISTAT surveys, UN/ECE Work Session on Statistical Data Editing, Madrid (http://www.unece.org/stats/documents/2003/10/sde/wp.3.e.pdf)

McKenzie, R. (2000) : A framework for priority contact of non respondents, presented at the second international Conference on establishment surveys, Buffalo

Merad, S., Wagstaff, H. (2005) : A management information system for controlling editing quality in a survey with multiple requirements, UN/ECE Work Session on Statistical Data Editing, Ottawa (http://www.unece.org/stats/documents/2005/05/sde/wp.20.e.pdf)

Chapter 4 - The production of flash estimates : a brief overview

What has been presented in the former chapters is relative to the production of estimates considering the survey as one single operation. But, tackling the question of timeliness of statistical results, one may be led to use other approaches, which are more model-based (more than in the "models" used to deal with non-response, that need also some assumptions). Particularly, the "temporal" component of the short-term statistics provides informations about time series that are interesting to use.

The assumptions made in the model used may be of a more or less "strong" nature. Some examples of preliminary estimates are given in the first part of this chapter, while the second one is dedicated to the question of the analysis of revisions, that may lead to the improvement of the preliminary estimation methodologies. *However, this chapter gives only a few elements on the subject, and is not intended to present a complete overview of the existing methods.*

4.1. Preliminary estimates

Different methods do exist. The OECD database gives some examples of them, and it is not the objective of this handbook to present a complete overview of these methods. Only three examples will be given here.

• Bacchini, Gennari & Iannacone propose an indirect method using « input data » to compute the index of production for the construction sector. This method is based on the estimation of a production function :

$$LogY_i = \beta_0 + \sum_k \beta_k LogX_{k,i} + \varepsilon_i$$

where :

Y_i is the production

$X_{k,i}$ are input variables (hours worked, value of building materials, …).

If input variables are available quickly, advanced estimates may be produced. Depending on the "quality" of the production function, this method may be used to replace the former one, based on the direct observation of the production of selected enterprises.

• Aelen (2004) presents a method to improve the timeliness of industrial short-term statistics. He compares the "current" method and a trend-based method (used for the non-respondent units at a given time) to a more refined method using time series analysis. More precisely, in the latter method, units are classified into two groups : one for which data are available (group I), and one for which data unknown at a given time (group II).

The model used for the aggregated turnover of group II is :

$$Y_{II,t} = \beta_0 + \beta_1 Y_{I,t} + Z_t$$

where :

$Y_{I,t}$ is the turnover of group I

β_0 and β_1 are regression coefficients

Z_t , the residuals, are modeled using ARIMA models.

This method gives better results for first estimates than the trend-based method and the current method. It allows reducing the delays while maintaining the level of accuracy of the previous method. Definitive results are produced using the "classical" method.

• Holmberg and Jäder (2005) propose to use, for rapid estimates of the Swedish foreign trade, a method based on an incomplete sample, built with companies for which data concerning two different periods are available at a given time. More precisely :

- on the 10th of month $t+1$, detailed results are produced for month $t-2$;
- on the 25th of month $t+1$, rapid estimates for month t and $t-1$ are produced.

If DET_{t-2} is the detailed estimate for month $t-2$, the rapid E_2 estimate for month t is :

$$E_2 = DET_{t-2} \frac{\sum_{i \in Bt} y_{i,t}^{t}}{\sum_{i \in Bt} y_{i,t-2}^{t}}$$

where Bt is the set of companies having reported for month t at the 25th of month $t+1$, and having also reported data for month $t-2$ at the 10th of month $t+1$.

The same principle is used to produce the rapid estimate E_1 for month $t-1$.

Holmberg and Jäder do insist on the fact that it is necessary to check if companies sending their data early are different, or not, from the "average" one. This work may be conducted by studying the discrepancies between rapid estimates and detailed results.

• One may also mention the fact that Eurostat launched a project concerning flash estimates for the Euro-zone and the EU-25th for three short-term economic indicators :

- monthly industrial production index ;
- quarterly GDP ;
- quarterly labour cost index.

This project is conducted by four NSIs, INSEE (France), ISTAT (Italy), ONS (United Kingdom) and DESTATIS (Germany), who will work in close cooperation. It is intended to make an inventory of available sources of information, constitute an historical database, develop the methodology relative to the question (particularly by a review of the literature existing on the subject), define quality indicators for the flash estimates, make some simulation studies and prepare a "toolbox" for the production of flash estimates.
Its results are expected for the end of 2008.

4.2. Analysis of the revisions

The analysis of revisions between, for example, preliminary and definitive estimates is important, since it gives a first assessment of preliminary estimates. It may also, in some cases, provide elements for improving the techniques that have been used.

Different methods may be used for this purpose (see for example Di Fonzo, 2005), only a few of them are presented here :

- descriptive analysis of the bias, if L_t is the later estimate and P_t the preliminary estimate :

mean revision $\overline{R} = \frac{1}{n}\sum_{t=1}^{n}(L_t - P_t) = \frac{1}{n}\sum_{t=1}^{n} R_t$

or mean absolute revision $MAR = \frac{1}{n}\sum_{t=1}^{n}\left|L_t - P_t\right| = \frac{1}{n}\sum_{t=1}^{n}\left|R_t\right|$;

- descriptive analysis of the variability by the use, for example, of the mean square revision:
  $MSR = \frac{1}{n}\sum_{t=1}^{n}(L_t - P_t)^2 = \frac{1}{n}\sum_{t=1}^{n} R_t^{\,2}$;
- use of statistical tests to check if the average revision is significantly different from 0 ;
- use of econometric analysis (Ladiray & Mazzi, 2001) : testing the value (compared to zero) of the coefficient of an ordinary least squares regression from the revision on the preliminary estimate, and also the coefficient of the current revision on the revisions of the previous period may give elements on the efficiency of the method.

Executive summary for chapter 4

1. Model-based approaches may be used to produce preliminary estimates.

2. Conducting some analysis of the revisions of these estimates may help to improve the accuracy of the estimates.

References for chapter 4 :

Aelen, F. (2004) : Improving timeliness of industrial short-term statistics using time series analysis, Statistics Netherlands, available on OECD website http://www.oecd.org/dataoecd/23/12/30044343.pdf

Bacchini, F., Gennari, P., Iannaccone, R. : A new index of production for the construction sector based on input data, available on OECD website http://www.oecd.org/dataoecd/21/49/30036154.pdf

Di Fonzo, T. (2005) : The OECD project on revisions analysis : first elements for discussion, paper presented at the OECD STESEG meeting, Paris, 27 - 28 june 2005, available on :
http://www.oecd.org/dataoecd/55/17/35010765.pdf

Jäder, A., Holmberg, A. (2005) : The growth rate method for production of rapid estimates of the Swedish foreign trade, Statistics Sweden, available on :
http://www.scb.se/statistik/_publikationer/OV9999_2005A01_BR_X100ST0509.pdf

Ladiray, D., Mazzi, G.L. (2001) : European short-term statistics : the state of the art, presented at he 11th CEIES seminar EU short term economic indicators : meeting new needs, available on :
http://www.oecd.org/dataoecd/22/58/30044960.pdf

OECD database on timeliness :
http://www.oecd.org/document/40/0,2340,en_2649_34257_30460520_1_1_1_1,00.html#register

Conclusion

The elements presented in chapters 2, 3 and 4 are not conflicting, but rather complementary.

Different conclusions may be drawn from all these elements.

1. The survey of the NSIs shows that NSIs have a rather pragmatic approach of the question of balance between delays and accuracy, and that few use methods presented in “academic” papers. It shows also that some NSIs use information systems in a regular way (that helps to take decisions relatively to the questions of timeliness) and are, from this point of view, better organized.

2. Improvement concerning delays aspects needs a long-term approach, and is easier for regular surveys than for one-shot surveys.

3. The methods presented in this handbook are well fitted for the production of aggregates (compared to the production of “individual” data bases). Their efficiency depends on the kind of variable, and on the level of publication, for the results to produce. Also, these methods are probably better fitted for business surveys than for household surveys.

4. An information system giving, in a continuous way, indicators about the survey is important. In particular, the characteristics of the non-respondent units have to be followed. Using the indicators helps to undertake some actions, at a given time, to respect the deadline for the results while preserving accuracy.

5. Concerning the survey data, getting metadata is very useful (for example results of checks, raw and definitive data for the same unit, etc.). It helps improving the efficiency of the used method for the following surveys. An important by-product of the studies that can be drawn from these metadata is the improvement of the edits, but also the improvement of the questionnaire itself.

6. The human component is very important. Since the approach that is presented needs some flexibility for the persons working on the survey, the tools that are implemented have to be rather simple and understandable, so that they are supported by this staff.

Annex 1

The « accompanying letter » and the questionnaire sent to the National Statistical Institutes for the survey about the practices concerning the balance between accuracy and delays in the statitical surveys

Direction des Statistiques d'Entreprises
Département "Système Statistique d'Entreprises"

INSEE
DIRECTION GÉNÉRALE
INSTITUT NATIONAL
DE LA STATISTIQUE
ET DES ÉTUDES
ÉCONOMIQUES

Dossier suivi par :
Philippe Brion
Tél. : 01 41 17 50 82
Fax : 01 41 17 65 20
Messagerie : philippe.brion@insee.fr

Paris, le 19 septembre 2006
N° 120/E210

Dear Colleagues,

In order to facilitate the implementation of principle 7 of the Code of Practice - “Sound methodology must underpin quality statistics. This requires adequate tools, procedures and expertise” -, Eurostat is supporting the development of a set of handbooks on current methods relating to some topics. In particular, this work is intended to capitalize on what is already done in Member states.

Insee is in charge of one of these handbooks, presenting guidelines for finding a balance between accuracy and delays in the statistical surveys. These two components of quality may be conflicting : the willingness of getting a great number of returns of questionnaires, and having an extensive data editing of them, leads to increasing the delays for the production of the results.

To finalize this handbook, it is important to collect information about the practices of the different National Statistical Institutes. The enclosed questionnaire has been prepared for this purpose, and intends to give information on strategies implemented in Member States on this subject. It aims at reviewing different kinds of surveys : infra-annual surveys, structural surveys, for households and businesses. Please feel free to provide any additional information that would not fit exactly to the framework of the questionnaire : for example, surveys relative to price indexes may need, if necessary, specific answers to the questionnaire.

The most important part of the questionnaire is part 4 : the answers to parts 1, 2 and 3 should be seen as elements to be set in the global framework presented in part 4 (questions asked for example in part 2 about data editing could, otherwise, be

seen as redundant relatively to other handbooks, as for example the one developed in the EDIMBUS project), and do not have to be very detailed.

I would be very grateful if you could fill this questionnaire and send it back to Insee (attention Philippe Brion - Insee timbre E210, 18 bd A. Pinard 75675 Paris cedex 14 France - philippe.brion@insee.fr) before the 15th of october 2006. Thank you very much in advance.

Best regards

Philippe Brion
Business Statistics Directorate
Insee

Handbook « *Guidelines for finding a balance between accuracy and delays in the statistical surveys* »

Study of the current practices of the NSIs

Name and function of the person in charge of filling the questionnaire :

Name of the National Statistical Institute :

Country :

Adress :

E-mail :

Phone :

Return adress :

Philippe Brion - timbre E210 - INSEE - 18 bd A. Pinard
75675 Paris cedex 14 - France

phone : *33 1 41 17 50 82*
fax : *33 1 41 17 65 20*
e-mail : *philippe.brion@insee.fr*

Thank you very much for your cooperation !

1. Introduction :

Do you produce different results at different times (preliminary results, standard aggregates, individual data files) for the surveys of your NSI ?

Yes / no

	Infra-annual surveys	Annual or one-shot surveys
Household surveys		
Business surveys		

Give, for each case, the kind of results and the delays :

2. Methods of data editing :

2.1. What kind of data editing method do you use for the surveys of your NSI ?

	Infra-annual surveys	Annual or one-shot surveys
Household surveys		
Business surveys		

If necessary, give some details for each method used :

If, for regular surveys, you use a data editing method, are the edits fitted regularly ?

Yes / no

	Infra-annual surveys	Annual or one-shot surveys
Household surveys		
Business surveys		

Did the use of data editing methods lead to reducing the delays for the results ?

Quantify the number of days (or weeks, or monthes) saved

	Infra-annual surveys	Annual or one-shot surveys
Household surveys		
Business surveys		

2.2. Do you use selective editing (giving « priorities » within the questionnaires to edit, and for example editing manually only a part of them) ?

Yes / no

	Infra-annual surveys	Annual or one-shot surveys
Household surveys		
Business surveys		

If you use selective editing methods :

- may you describe briefly their principles ?
- do you produce the results using raw data for the units that are not manually edited, or using imputed data ?

	Infra-annual surveys	Annual or one-shot surveys
Household surveys		
Business surveys		

How many days (weeks, monthes) have been saved by applying these methods ?

Quantify the number of days (or weeks, or monthes) saved

	Infra-annual surveys	Annual or one-shot surveys
Household surveys		
Business surveys		

2.3. Do you use "output" editing (checks of aggregated data) ?

Yes / no

	Infra-annual surveys	Annual or one-shot surveys
Household surveys		
Business surveys		

If yes, may you give more details :

3. Follow-up of non-respondents :

Which method do you use for the non-respondents ?

1 = recall of all units
2 = focus on large units
3 = focus on units whith a bigger impact on estimates
4 = focus on « long-term » non-respondents
5 = other (specify)

remark : more than one answer is allowed

	Infra-annual surveys	Annual or one-shot surveys
Household surveys		
Business surveys		

4. Overall view of the survey :

4.1. Do you use a global list of indicators for the following of the surveys in your NSI (giving elements on different sub-processes of one survey) ?

Yes / no

	Infra-annual surveys	Annual or one-shot surveys
Household surveys		
Business surveys		

May you describe briefly these indicators ?

If yes for at least one category, go to question 4.2, otherwise go to question 4.4.

4.2. Do these indicators supply information on accuracy of the results that could be produced, at a given moment, using only the questionnaires that have been returned and edited ?

Yes / no

	Infra-annual surveys	Annual or one-shot surveys
Household surveys		
Business surveys		

4.3. Do you define priorities for some actions to be undertaken (particularly for the follow-up of non-respondents, and data editing), depending on the delays and on cost considerations ?

Yes / no

	Infra-annual surveys	Annual or one-shot surveys
Household surveys		
Business surveys		

4.4. Do you use stopping criteria for some surveys of your NSI ?

Yes / no

	Infra-annual surveys	Annual or one-shot surveys
Household surveys		
Business surveys		

Precise, if necessary, for each considered survey, the method (deadline, maximum number of used questionnaires, weight of these questionnaires, ...)

5. Other comments :

Annex 2 : Detailed results of the survey of the NSIs

21 questionnaires were received. There was some item non-response.

Part 1 (introduction)

Do you produce different results at different times (preliminary results, standard aggregates, individual data files) for the surveys of your NSI ?

	Infra-annual surveys	Annual or one-shot surveys
Household surveys	13 yes, 7 no	12 yes, 8 no
Business surveys	21 yes, zero no	20 yes, zero no

Business surveys are more concerned with the fact to produce results at different times.

Regulations concerning STS (short-term statistics) are often mentioned by many NSIs. For statistics as the industrial production index, the producer prices index, Prodcom statistics, or business tendencies, preliminary results are produced, followed by revised results (and sometimes definitive results after). The regulation relative to SBS (structural business statistics) is also mentioned, as the labour cost survey.

For Intrastat as extra-EU trade statistics, there are different releases.

Concerning household surveys, for some annual surveys first results are produced, and then (six monthes to one year after) an individual data file is often produced, particularly for the needs of researchers. For infra-annual household surveys, the labour force survey is often mentioned.

Part 2 (methods of data editing)

What kind of data editing method do you use for the surveys of your NSI ?

	Infra-annual surveys	Annual or one-shot surveys
Household surveys	CAPI (7) Logical checks (7) Consistency checks (5) Fellegi Holt (3) Completeness of the questionnaire (2) Outliers detection (1)	Logical checks (9) CAPI (7) Consistency checks (5) Completeness of the questionnaire (2) Visual control ((1) Fellegi Holt (1) Outliers detection (1)
Business surveys	Consistency checks (10) Logical checks (9) Outliers detection (8) Use of data of a previous period (5) Comparison with data from other sources (4)	Logical checks (10) Consistency checks (9) Outlier detection (7) Use of data of a previous period (5) Comparison with data from other sources (4)

	Completeness of the questionnaire (2) Check of the monetary unit (650<y(m)/y(m-1)<1350) (1)	Completeness of the questionnaire (4) Fellegi Holt (1)

The number of answers relative to each kind of method is indicated between brackets.
Remark : there may be some redundancy between some answers (for example, it is possible that some member states answering CAPI could also mention using logical checks).

There were few answers giving details about the methods of imputation used (if this kind of methods is used, reweighting being an alternative method). This is probably due to the way the question was formulated.
The answers the most frequent were :

- for household surveys, hotdeck method, and deductive (or deterministic) method
- for business surveys, model-based methods, and use of historical data.

If, for regular surveys, you use a data editing method, are the edits fitted regularly ?

	Infra-annual surveys	Annual or one-shot surveys
Household surveys	16 yes, 2 no	17 yes, 1 no
Business surveys	18 yes, 2 no	19 yes, 1 no

Did the use of data editing methods lead to reducing the delays for the results ?

Quantify the number of days (or weeks, or monthes) saved

This question was considered as difficult to answer by a lot of NSIs, especially concerning the quantification of the number of days saved.
However, some NSIs mentioned different aspects :

- rather than shortening the delays, the use of data editing methods may lead to improve the quality of statistics (5 answers) ;
- it may also allow to reduce the size of data editing staff (3 answers) ;
- in some cases, the use of more sophisticated data editing methods did rise the costs of data editing (1 answer).

Few answers gave quantified elements :

	Infra-annual surveys	Annual or one-shot surveys
Household surveys	20 days 3 weeks 20 days (labour force survey, quaterly results)	1 month 2 monthes
Business surveys	20 days 2 days (industrial production index) 30 days (foreign trade)	1 month

Do you use selective editing (giving « priorities » within the questionnaires to edit, and for example editing manually only a part of them) ?

	Infra-annual surveys	Annual or one-shot surveys
Household surveys	7 yes, 10 no	6 yes, 10 no
Business surveys	17 yes, 4 no	16 yes, 4 no

If you use selective editing methods :

- **may you describe briefly their principles ?**
- **do you produce the results using raw data for the units that are not manually edited, or using imputed data ?**

Different kinds of methods were mentioned :

- priority given to "large" units : 9 answers ;
- calculation of a score for each questionnaire (generally combining local scores taking into account the raw data, and plausible value for the same variable), and checking manually units above a threshold (for the score) : 4 answers ;
- priority given to "core" variables : 2 answers ;
- check of all values of large units and only high values of small units : 2 answers ;
- study of the contribution of each category (stratum) to the evolution of a statistic : 1 answer.

One NSI uses threshods, for selective editing, that are reviewed considering the available human resources and time.

Concerning the use of raw data or imputed data for units that are not manually checked, there were few answers : 2 for raw data, 3 for imputed data.

How many days (weeks, monthes) have been saved by applying these methods ?
Quantify the number of days (or weeks, or monthes) saved

Few responses were received for this question ; some NSIs pointed out that the implementation of these methods allowed to reduce the size fo data editing staff, rather to reduce the delays.
The quantified answers received are mentioned below.

	Infra-annual surveys	Annual or one-shot surveys
Household surveys	3 weeks some days to some weeks	1 month some weeks
Business surveys	3 weeks 5 - 10 days	1 month 10 - 20 days 3 monthes (farm structure survey) 4 monthes

Do you use "output" editing (checks of aggregated data) ?

	Infra-annual surveys	Annual or one-shot surveys
Household surveys	13 yes, 6 no	16 yes, 4 no
Business surveys	19 yes, 1 no	19 yes, 1 no

Mainly, three kinds of methods were mentioned :

- comparison of aggregated data with those of the previous year (14 answers) ;
- coherence of the statistics with other sources (9 answers) ;
- plausibility checks at the macro level, using defined ratios (3 answers).

Part 3 (follow-up of non-respondents)

Which method do you use for the non-respondents ?

1 = recall of all units
2 = focus on large units
3 = focus on units whith a bigger impact on estimates
4 = focus on « long-term » non-respondents
5 = other (specify)

remark : more than one answer is allowed

	Infra-annual surveys	Annual or one-shot surveys
Household surveys	1 (9 answers) 1,3 (2 answers)	1 (9 answers) 1,3 (2 answers) 4 (1 answer)
Business surveys	1,2,3 (9 answers) 1,2,3,4 (3 answers) 1,2 (3 answers) 2,3 (2 answers) 2,3,4 (2 answers) 1 (2 answers)	1,2,3 (8 answers) 1,2,3,4 (4 answers) 2,3 (3 answers) 1,2 (2 answers) 2,3,4 (2 answers) 1,4 (1 answer)

Remark : answers 5 were reclassified in categories 1 to 4 relatively to the explanations given on the questionnaires.
Also, some respondents found the question a bit confusing for household surveys. That is the reason why the response rate is lower for this category of survey.

Part 4 (overall view of the survey)

Do you use a global list of indicators for the following of the surveys in your NSI (giving elements on different sub-processes of one survey) ?

	Infra-annual surveys	Annual or one-shot surveys
Household surveys	8 yes, 6 no	9 yes, 6 no
Business surveys	11 yes, 9 no	9 yes, 8 no

May you describe briefly these indicators ?

The answers were (with the number of corresponding answers) :

- response rate / refusal rate (9)
- completeness / coverage coefficient (4)
- sampling error (4)
- timeliness indicators (3)
- validation rate (2)
- recontact rate (1)

- item non-response rate (1)
- rate of proxy answers (1)
- average time for an interview (1)
- interviewer preformance (1)
- use of a centralized system of indicators (1)

Do these indicators supply information on accuracy of the results that could be produced, at a given moment, using only the questionnaires that have been returned and edited ?

	Infra-annual surveys	Annual or one-shot surveys
Household surveys	5 yes, 3 no	6 yes, 2 no
Business surveys	9 yes, 1 no	7 yes

Do you define priorities for some actions to be undertaken (particularly for the follow-up of non-respondents, and data editing), depending on the delays and on cost considerations ?

	Infra-annual surveys	Annual or one-shot surveys
Household surveys	6 yes, 3 no	6 yes, 4 no
Business surveys	11 yes	11 yes

Three ways of giving priorities are mentioned :
- impact of coverage of economic sectors ;
- priority lists (for example using enterprise "size") ;
- decision to focus on clearing errors of one category of units, if the quality of its data is considered too low, rather to continue response chasing.

Do you use stopping criteria for some surveys of your NSI ?

	Infra-annual surveys	Annual or one-shot surveys
Household surveys	7 yes, 5 no	7 yes, 6 no
Business surveys	9 yes, 3 no	9 yes, 3 no

Precise, if necessary, for each considered survey, the method (deadline, maximum number of used questionnaires, weight of these questionnaires, ...)

The methods mentioned are :
- response rates, including weighted response rates (8 answers)
- coverage rate (2 answers) ;
- accuracy indicators (1).