NATIONAL STATISTICAL SERVICE OF GREECE (NSSG)

QUALITY IMPROVEMENTS OF THE SURVEY PROCESSES

Prepared by Ioannis Nikolaidis*

PIRAEUS, JANUARY 2009

European Commission grant agreement: No. 10300.2005.021-2007.331

* Head of the Methodology, Analysis and Research Section of the NSSG, e-mail: giannikol@statistics.gr

Contributors to this project:
The employees of the Methodology, Analysis and Research Section of the NSSG,
Ioannis Nikolaidis (Head of the Methodology, Analysis and Research Section)
Ilias Virras
Maria Orfanou
Irene Sarantou
Vasiliki Spiliopoulou

Contents

1. INTRODUCTION..5

2. DEFINITION, ANALYSIS AND DOCUMENTATION OF THE SURVEY PROCESSES....7

2.1. DEFINITION OF THE SURVEY PROCESSES 7
2.1.1 Planning process..*7*
2.1.2 Operational process..*8*
2.1.3. Evaluation process..*8*
2.2 ANALYSIS OF THE SURVEY PROCESSES..8
2.2.1. Classifications, concepts, definitions and harmonization..*8*
2.2.2. Frames and Coverage..*9*
2.2.3 Methods for the data collection..*10*
2.2.4 Questionnaire design and testing..*11*
2.2.5 Statistical editing and imputation *12*
2.2.6. Statistical estimation and analysis..*12*
2.2.7 Presentation, publication and dissemination of statistical data..*12*
2.3 DOCUMENTATION OF THE STATISTICAL SURVEY..14
2.3.1 Relevance of statistical data..*14*
2.3.2 Accuracy of data *14*
2.3.2.1 Sampling errors 15
2.3.2.2 Non-sampling errors..15
2.3.3. Timeliness and punctuality..*18*
2.3.4 Accessibility and Clarity *19*
2.3.5. Comparability *20*
2.3.6 Coherence *21*
2.3.7. Completeness *21*
2.3.8 Cost and Burden..*22*
2.4 CASE STUDY ON THE DOCUMENTATION OF THE GREEK LABOUR COST SURVEY 22

3. PROCESS VARIABLES AND QUALITY PROCESS INDICATORS 24

3.1 MAIN PROCESSES AND SUB-PROCESSES OF THE SURVEYS 24
3.2. MAIN PROCESS VARIABLES..26
3.3. QUALITY PROCESS INDICATORS 28

4. PLANNING PROCESS QUALITY INDICATORS 30

4.1 IDENTIFICATION OF USERS' NEEDS..30
4.1.1 Recording the users' needs for statistics..*30*
4.1.2 Users' demands of product quality *30*
4.2 CIRCUMSCRIPTION AND RECKONING SURVEY FINANCING 31
4.3 SAMPLE DESIGN..32
4.3.1 Coefficient of correlation between the stratification variable and the main survey variables..*32*
4.3.2 Effect of the applied sampling scheme on the sampling errors of the survey variables.......*36*
4.3.3. Conclusions regarding the effect of the sample design on the sampling errors.................*37*

5. OPERATIONAL PROCESS QUALITY INDICATORS..39

5.1 THE EFFECTS OF WEIGHTING, CLUSTERING AND STRATIFICATION ON SAMPLING ERRORS AS PROCESS QUALITY INDICATORS 39
5.2 THE EFFECT OF MEASUREMENT ERRORS ON THE RESULTS..40
5.2.1 Extra variability due to measurement errors *40*
5.3 THE EFFECT OF NON-SAMPLING ERRORS ON THE SAMPLING ERRORS..43
5.3.1 Effect of frame errors on the sampling errors..*43*
5.3.2 Effect of all non-sampling errors on the sampling errors..*46*
5.3.3 Effect of miscoding on sampling errors *46*
5.3.4 Conclusions regarding the effect of non-sampling errors on sampling errors *47*

6. EVALUATION PROCESS INDICATORS..48

6.1 ANALYSIS ON WHETHER THE SURVEY RESULTS COMPLY WITH OTHER SIMILAR SURVEY RESULTS.48

6.2 COMPARING POTENTIAL AND ACTUAL DISSEMINATED PRESENTATION FORMS 49
6.3. SURVEY DOCUMENTATION SYSTEM 50

7. GENERAL REMARKS **52**

8. REFERENCES.. **55**

ANNEX 1: QUALITY EVALUATION ON THE GREEK LABOUR COST STATISTICS FOR THE YEAR 2004 .... **57**

ANNEX 2: CASE STUDY ON THE RELATIONSHIP BETWEEN RELEVANCE INDICATORS **73**

ANNEX 3: CASE STUDY ON THE RELATIONSHIP BETWEEN INDICATORS OF USERS' DEMANDS OF PRODUCT QUALITY.. **75**

ANNEX 4: METHODOLOGICAL NOTE FOR THE GREEK STRUCTURAL BUSINESS SURVEY **80**

ANNEX 5: METHODOLOGICAL NOTE FOR THE JOB VACANCY SURVEY **88**

ANNEX 6: RELATIONSHIP BETWEEN THE DEGREE OF URBANIZATION (STRATIFICATION VARIABLE) AND THE BASIC VARIABLE (INCOME) **91**

ANNEX 7: THE EFFECTS OF STRATIFICATION, WEIGHTING AND CLUSTERING ON SAMPLING ERRORS AS PROCESS QUALITY INDICATORS **93**

ANNEX 8: GREEK SURVEY OF GOATS... **125**

ANNEX 9: COMPLETE ANALYSIS OF ASSESSING THE COHERENCE BETWEEN GREEK EU SILC AND GREEK STRUCTURAL EARNING SURVEY (SES) RESULTS.. **127**

1. Introduction

Product quality is the quality of the output (data and services provided by statistical services). These products are created by a sequence of processes and as a result, the product quality is affected strongly by the process quality. The processes are monitoring through key process variables, which are factors that have the largest effects on critical product characteristics i.e. those characteristics that best indicate the quality of the product, as accuracy, timeliness etc (Jones and Lewis, 2003).

The point of departure for systematic quality work and for compiling quality indicators is the "users' needs", because (a) they form the research problem in each statistical survey and (b) users' needs, process quality and product quality are linked. As the analysis of processes is a precondition for quality improvements, developing a process flow chart should be firstly carried out, so that the key process variables to be determined for monitoring the quality of the processes at every survey stage (Jones and Lewis, 2003).

The improvements on the quality of the produced statistics is achieved through recording and analyzing the processes included in the statistical surveys on the basis of the relationship among quality of processes and quality of the produced statistics. The effect of each separate survey process on the quality of the final data should be investigated in detail (e.g. identification of users' needs on the relevance, sample design on accuracy and timeliness, data collection on the accuracy) for improving the quality of processes and products. The strategic and operational objectives for quality improvements of statistics can be based on:

- Documentation and analysis of the survey processes
- Examination of the relationship between the quality of the survey processes and the quality of data
- Establishment of a complete set of process variables (indices)
- Development of methods for measuring the process variables, which will determine the quality of the survey processes
- Measuring the effect of the quality of the survey processes on the quality of data

In this handbook, the strategic and operational objectives for the quality improvements of statistics are presented and analysed through process variables and process quality indicators, which are applied in case studies based on data from statistical surveys conducted by the NSSG.

The statistical survey is regarded as a total process consisting of three main stages or processes (planning, operation and evaluation) (Statistics Finland, 2002), where the following key actions may be performed for recording and improving the quality of statistics.

I. Planning process
(Identification of users needs, definition of the content, strategic decisions on data collection methods and planning of data collection, establishment of resources)
- Analysis of the users' needs and the users' demands on product quality.

- Circumscription and reckoning of an indicator concerning survey financing (needed funds and available funds). This indicator, if taken into account with other quality figures, will show how money can affect the decisions while designing the survey and the quality of the final results.

II. Operational process

(Construction of the sample frame, selection of the sample, compilation of the questionnaire, data collection, data processing, production of final data, estimation and analysis, publication of final results, dissemination of statistical information, archiving)

- Assessment of the sample design with the investigation on the way stratification and clustering (sample design) influences the precision of the results, creation and computation of relevant indices
- Compiling indicators showing the degree of coverage of the target population
- Examination of the effect of non response and frame errors on the survey results by using specific mathematical expressions
- Synthesizing of a mathematic formula so as to explore the effects of the variable biases of interviewers on responses to a variety of items
- Estimation of a rate concerning miscoded variables

III. Evaluation process

(Final products conformance with the definitions adopted at the planning stage)

- Designation and adoption of an index denoting whether the survey results comply with other similar survey results (coherence)
- Composition of an indicator which will help comparing potential and actual disseminated presentation forms (tables, variables, documentations, methodological reports) of results
- Development of an index for comparing the documentation on processes with the total number of the survey processes which could be documented

The main body of the handbook is split into five main sections:

- Definition, analysis and documentation of the survey processes
- Process variables and quality process indicators
- Planning quality process indicators
- Operational quality process indicators
- Evaluation quality process indicators

2. Definition, analysis and documentation of the survey processes

2.1. Definition of the survey processes

2.1.1 Planning process

The planning of a statistical survey is based on the customers' needs for social or economic or scientific research. These data needs should be identified and specified in the planning stage and according to these needs the research problem is formed. After forming the research problem the target population and the content of the survey must be specified and defined (e.g. definition of concepts and the survey characteristics, applied classifications, geographical level of the produced statistics). Additionally, the strategic decisions on data collection methods (census or sample methods, administrative sources) and the planning of the data collection are carried out. Of course, the choice of the data collection method is an important strategic decision in the planning stage of the survey, because it has to do with the nature of the data collection, the design of the questionnaire (setting of questions and other aspects of the questionnaire), the duration of the statistical survey and its costs.

The data processing methods are also specified in the planning stage, together with the editing and imputation methods needed for the data productions. Other important elements in this stage are interviewer training and the organization of the fieldwork.

Taking into account (a) the choice of the data collection method and (b) the timeliness of the produced data, the resources, such as budget, personnel and other equipments must be also set up in the planning stage as well as the time schedule for conducting each separate stage of the whole statistical survey must be defined so that all the survey processes to be conducted in time ensuring all time criteria for the timely data production. All the stages of the survey must be taken into account in planning the production schedule. The number of required stages varies from survey to survey and the stages are placed differently with the given emphasis.

In the sampling surveys the central and important work in the planning process includes the definition both of the sampling design and the estimation methods. In the survey design of both census and sampling surveys, it is necessary to assess the suitability of the sampling frames and the degree of comparability and coherence of these frames with the target populations. The sample selection method, the sample size and the inclusion probabilities of units in different target population groups (e.g. size classes of enterprises) are defined so that the estimation of statistics to fulfil the reliability criteria set for them (e.g. coefficient of variation of the estimation of the variable "turnover" at the total country to be less than 2%). The choice of a sampling method also includes correction methods for non-response errors, formulae or the calculation of estimates and their standard errors.

The success of every work of the planning process is assessed examining (a) if the final products fulfil the quality criteria set for them and (b) the quality of each operational process which is strongly related with the stages of the planning process.

2.1.2 Operational process

The whole operational process of the statistical survey comprises the following sub-processes (Statistics Finland, 2002):

- Construction of the sample frame with the required auxiliary information
- Selection of the sample
- Creation of the questionnaire with the required useful instructions tot the interviewers for the correct filling in the questionnaire
- Data collection
- Data processing (coding, checking for the correctness of the completed questionnaire, data editing, data entry etc)
- Production of final data
- Estimation and analysis
- Publication of final results
- Dissemination of statistical information
- Archiving

The success of every work of the operational process is assessed examining how the quality of final products has been affected by the quality of each sub-process of the whole operational process.

2.1.3. Evaluation process

This process of the statistical survey examines how the final products defined in the planning stage have been produced and published and how the quality criteria have been achieved.

2.2 Analysis of the survey processes

2.2.1. Classifications, concepts, definitions and harmonization

One important stage conducted in the planning process is how (a) the data will be classified, (b) the concepts will be defined and (c) the produced statistics will be harmonised in order to be comparable over time and across countries.

a. Classifications

As the statistics are produced from a large set of individual observations, the collected empirical observations need to be classified by their similarities. Thus, classification is an essential part of statistics and standard classifications are the key instruments of the official statistics. Usually, a classification consists of named groups and their identifiers (codes) accompanied by their definitions. Examples of classifications are described as follows:

- Demographic classifications describing the surveyed or statistical units (persons, households, enterprises, local units)
- Classifications describing economic activities of the enterprises (e.g. NACE Rev1. or Rev.2)
- Classifications describing the professions (e.g. ISCO-88)

- Classifications describing the products (e.g. CPA or PRODCOM)
- Regional classifications (NUTS I, II, III, IV)
- Special classifications as classifications of the education levels, of the diseases etc

Generally, the classifications used in statistics must be relevant for the purpose of the key users, because the statistics have a significant role in the social research and the decision-making. The quality of the classifications strongly depends on how systematically and consistently they classify the observations collected by the surveyed units.

b. Concepts

The concepts used for the description and production statistics are based on a scientific and statistical frame of reference and are applied (a) to define the subject of the survey characteristics, (b) to describe the statistical units and generally to describe the population under study.

c. Harmonization for comparability of statistics

Applying classifications and concepts based on international modes, recommendations and agreements it is ensured that the produced statistics are comparable across countries and across countries. It is essential for comparisons of surveys results between countries and within countries, that the contents of the compared surveys to be similar concerning the definitions and the measurements of variables. Differences in definitions, methods and operational procedures should be examined, justified and if it is possible these differences should be adjusted in order the efficiencies of comparisons to be improved.

The definitions of the classifications and concepts must be documented carefully in the metadata in such a form that the provided information easily to be transferred to the users of the statistics.

2.2.2. Frames and Coverage

After the formation of the research problem in the planning process, it follows the determination of the target population (statistical units under interest). The purpose of all statistical surveys is to generalise the final statistical products (distributions, sums, averages, ratios etc) to the whole target population. (Levy and Lemeshow, 1991), The goal to generalise, the relevance and the accuracy of the produced statistics require the use of a correct sampling frame for (a) design of the survey, (b) the data collection and (c) the estimation of the survey characteristics in the sample surveys.

In both census and sample surveys the coverage, completeness, timeliness, information content and accuracy of the frame are essential and critical factors regarding its suitability for the needs of the survey. It is important to be conducted the evaluation of the connection between the frame units and the surveyed units belonging to the target population.

Generally, the nature of the survey has an effect on the suitability of the frame as follows:

- In the census surveys, the frame and the auxiliary information must correspond to the content requirements of the survey
- In the sample surveys, it is essential that the frame contains the required auxiliary information for the selected sampling methods (e.g. for the stratification of the units, for the definitions of the inclusion probabilities etc).

The whole target population target population must be covered by the survey, but in practice it is rare for complete coverage to be achieved. The target population includes all units of interest, while the frame population contains those units that can be reached. The most commonly used sampling frames based on data from administrative registers. The information included in the frame should be up-to-data, thus, when using registers the stratification variables, the units' identification variables and the classification of units should be updated according to the latest possible information.

2.2.3 Methods for the data collection

The methods applied for the data collection are (a) census surveys (b) sampling surveys and (c) registers

a. Census or sampling survey

Complete enumeration censuses (in which the statistical information is obtained from all surveyed units) presuppose the existence of a certain minimum of facilities, such as funds, professional personnel for planning methodology and the supervision of field operations, sufficiently qualified enumerators, mapping material, computers equipments etc. In contrary, in order to reduce costs and response burden, sampling rather than enumeration for census is used in statistical surveys. A properly designed sample survey can provide the users with accurate estimations of the survey characteristics.

Advantages of a census survey (Zarcovich, 1965)

- Data from a census can be tabulated by administrative and other area units, whatever size
- Sample surveys are inefficient methods of obtaining information on rare events, such as areas under some crops and yields, thereof, the number of persons of advanced age, their distribution by sex, age and area of residence, the number of persons having a specified physical disability etc.
- Data of a complete census can be widely exploited as a basis for various surveys (e.g. compiling sampling frames of city blocks, using census data as auxiliary information for improving sampling results etc).

Advantages of a sample survey (Kish, 1979)

- Sample survey is much cheaper than a census
- Sample survey can be designed to obtain wide varieties of complex data, rich and deep in content
- Sample survey produces timely and relevant results

b. Administrative and Statistical Registers

Administrative records refer to any data collected primarily for some other purpose than the production of the official statistics. As the administrative data have been collected or other purposes, using them for statistical purposes does not increase the response burden. Additionally, the costs for the producers of statistics are very low compared to direct data collection. If administrative records consist of unit-level data, then they can be called registers providing users with micro-data. A key concept in the use of registers is the code system, which refers to the identification code by which data unit can be identified unambiguously. Major examples of produced statistical micro-data obtained from registers are (a) vital statistics (births, deaths, migration, emigrants, immigrants, marriages and divorces), (b) main variables of the structural business statistics based on data from the business register etc.

Advantages of producing statistics through administrative registers

- Inexpensive data production
- Timely data
- Precise and complete data
- Detailed data concerning small area regions and rare populations

2.2.4 Questionnaire design and testing

Collected data from the surveyed units should be recorded on a standardised instrument (paper or electronic questionnaire).

a. Questionnaire design

A well-designed questionnaire in surveys (business or social surveys) should collect data efficiently, with a minimum number of errors. In addition questions should: (a) facilitate the capture and coding of data, (b) minimize the amount of the required editing and imputation and (c) lead to the overall reduction in the cost and time associated with data collection and processing. Other considerations that should be taken into account in designing questionnaires include (Gower, 1994):

- Consistency in terminology and response categories with standard concepts and definitions
- Nature of respondent population (e.g. enterprises, households, individuals)
- Response burden
- Complexity of the data to be collected
- Comparability of the produced statistics with other surveys
- Data reliability
- Non-response in order the questionnaire to be friendly to the respondents

b. Testing

The testing is designed to provide a statistical evaluation of how the questionnaire performs and the pilot study is the main type of testing method. A *pilot study* is conducted to observe how all the survey operations, including the administration of the questionnaire, work together in practice. It duplicates the final survey design on a small scale from beginning to end, including data processing and analysis. It helps the statistician to see how well the questionnaire performs in relation to all other parts of the survey. There are some problems that can only be identified when all phases of the survey are tested together. For example, problems with question wording or concepts that need further clarification may be identified during interviewer training. The data processing stage may reveal problems on coding or answer categories.

2.2.5 Statistical editing and imputation

Editing is the procedure for detecting and adjusting individual errors in data records resulting from data collection and capture. The checks for identifying missing, erroneous, or suspicious values in computer-assisted editing are called edit rules or edits. An editing change occurs when an item (question) value is adjusted as a consequence of an action taken when an error is identified. Editing may be defined as the process of changing a response to another value.

Imputation means creating plausible (but artificial) substitute values for all those missing, while preserving the original weights when estimates are calculated. There are several methods and software suitable for imputation. Imputation may be defined as the process of changing a missing value to a plausible, non-missing, value.

2.2.6. Statistical estimation and analysis

The parameters that are the focus of interest in descriptive statistical surveys are usually totals, averages or proportions at the total population level or relative to parts of it. In surveys where the emphasis is on analysis, the interest is concentrated on correlations and connections between phenomena. Parameters have to do with linear models, coefficient of correlations, regressions etc. A common stage in both types of surveys is the estimation of unknown parameters as reliably as possible.

Estimations use the available survey data to produce numeric estimates for unknown parameters. The reliability of the estimates is assessed through calculating the standard errors or the coefficient of variations of these estimates. Information on both confidence intervals and coefficient of variations is important for evaluating the reliability of statistics and must also be reported to the users.

Statistical analysis is a process in which the survey data are examined through methods such as regression analysis, comparisons of means through t-tests or analysis of variance. In the analysis, the parameters of the statistical models e.g. regression coefficients are estimated with their standard errors for assessing the reliability of the estimates.

Another important stage of the analysis is to monitor the quality of the survey processes for generating improvements detecting the sources of errors, evaluating their influence on the reliability of the data and proposing methods for their adjustments (Statistics Finland, 2002).

2.2.7 Presentation, publication and dissemination of statistical data

Presentation, publication and dissemination of statistical data together with archiving (which includes micro-data) are the last steps in the whole process of a statistical survey.

a. Presentation of the statistical outcomes

Presentation is a very significant step because it determines how the outcome of the previous steps (which is referred to the statistical products) can be made widely and usefully exploitable. The presentations may be electronic (e.g. Internet, CD ROM etc) and paper publications.

Generally, statistical data may be presented through three methods: (a) textual presentation, (b) tabular presentation and (c) chart or graph presentation. So far, tables are the most frequently used presentation methods, but the statistical charts or graphs are becoming more and more useful because they convey information visually, more illustratively, efficiently and quickly than a table or a text. Both charts and graphs are suitable for bringing out the central results and those that are the most important to communicate. A statistical chart is better than a table for displaying the structural aspects of data, summarising large amounts of data, demonstrating how things are connected, communicating ideas and conclusions. Text is a more subjective presentation form than a table or a chart and it allows an extensive analysis of the results. On the other hand, a table includes information more extensive than chart but extracting information from a table is slow. However one of the advantages of the tables is that the items in the cells are as homogeneous as possible putting the data into an easily comparable form.

As technological advances in hardware, software, data documentation and web-sites make access to and analysis of micro-data more and more practical and desirable, researchers and academic society expect a wide range of *micro-data* to support their research.

b. Publication of statistics

All the produced statistics shall be released for publicity, which serves to notify the users about the completion and availability of a new statistical product or service. The means of releasing include release calendars, press releases, various publications and press conferences. Release calendars serve the purpose of improved predictability in agency operations, making it easier for users to plan ahead according to the release dates.

c. Dissemination of statistical information

The dissemination of statistical information comprises the provision of both free-based and non-free-based statistical services and products, the analysis of customer information needs and guidance to customers in how to find the right sources of information. The analysis of users' needs is significant at the planning process of a survey and in the marketing aiming at promoting the use of statistical products. Information and communication about statistical products and services aims at providing a significant service to current users and at the same time publicizing about existing information so that potential users can find the information they need. The producer of statistics aims to notify users about (a) the available information, so that they can have easy access to relevant and timely information and (b) the means and formats by which the statistical information cab be transferred to the users easily (Statistics Finland, 2002).

2.3 Documentation of the statistical survey

Documentation refers to the description of a statistical survey including the concepts, definitions, classifications and methods of data collection and processing used for the production of statistics. Additionally, the documentation describes the quality of the statistical products through quality indicators. The statistics and the general description of their production process should comply with the checklist structured in accordance with the quality criteria of the Eurostat (Eurostat 2002a, b, c and Eurostat 2003a, b, c).

2.3.1 Relevance of statistical data

Relevance is the degree to which statistics meet current and potential users' needs. It refers to whether all statistics that are needed are produced and the extent to which concepts used (definitions, classifications etc.) reflect user needs. As relevance is not an inherent characteristic of statistical data, it can be evaluated and measured through (a) the analysis of the collected data from the users' satisfaction surveys and (b) the recording the data requirements of Commission Regulations, International Organizations (IMF, OECD, etc).

What should be reported to document the relevance?

- A description and classification of users*:* Classification of the users for each separate survey in small number (about 5) of meaningful classes
- Ranking of the classes of users according to their importance
- The number or percentage of unavailable results, compared to what should be available
- Reasons for incompleteness as well as the prospects for future solutions
- Systematic documentation of the methods currently used for the measurement of user satisfaction for those types of users.
- Follow-up of the user satisfaction assessment, i.e. the measures and actions taken to improve user satisfaction

2.3.2 Accuracy of data

Accuracy is defined as the closeness between the value finally produced (after collection, editing, imputation, estimation, etc) and the true, but unknown, population value. The difference between the finally produced value and the unknown true value is the error. The types of errors are:

- *Sampling errors,* which affect only sample surveys, due to the fact that only a subset of the population, usually randomly, selected is enumerated
- *Non-sampling errors*, which affect both sample surveys and complete enumerations, are (a) Coverage errors, (b) Measurement errors, (c) Processing errors and (d) Non-response errors

2.3.2.1 Sampling errors

Sampling errors arise from the fact that not all units of the targeted population are enumerated, but only a sample of them. Therefore, the information collected on the units in the sample may not perfectly reflect the information, which could have been collected on the whole population. A traditional quality indicator for sampling error is the coefficient of variation (CV), which is defined as:

Coefficient of Variation (%) = (square root of the estimate of the sampling variance / estimated value)* 100

What should be reported to document the sampling errors?

- Description of the sampling design
- Estimator used for the expected value
- Methodologies applied for the variance estimation
- Coefficient of Variation for the main variables
- Explanations for the little accuracy of less accurate variables

If non-probability sampling is applied, then the documentation should include***:***

- Type of sampling (e.g. cut-off, purposive or controlled) and the exact way, by which the sample selection was carried out in the field,
- Estimates of sampling bias and the coefficient of variations of the main survey characteristics
- Assumptions used in the estimations

2.3.2.2 Non-sampling errors

a. Coverage errors

The *frame* is a device that permits access to population units. *Frame population* is the set of population units, which can be accessed through the frame and the survey's conclusions really apply to this population. Coverage errors (or *frame errors) are* due to divergences existing between the target population and the frame population. We can distinguish the following types of coverage error:

- *Under-coverage:* there are target population units which are not accessible via the frame (e.g. the new enterprises may not be included in the frame)
- *Over-coverage:* there are units accessible via the frame which do not belong to the target population (e.g. closed enterprises may be listed in the frame, duplication of some units)
- *Misclassification*: auxiliary information provided by the frame may be inaccurate for some population units (e.g. wrong size of business establishments in a business register or wrong economic activity of enterprises)

Over-coverage and misclassification can be identified and through the data collection process. On the other hand, the under-coverage is difficult to be measured unless another complete source of statistical information is available. The quality of the frame can be assessed through the indicators: over-coverage and misclassification error rates, which are defined as:

Over-coverage rate = (Number of out-of-scope units) / (Number of in-scope units + Number of out-of-scope units)

Misclassification rate = (Number of units misclassified but still in-scope)/ (Number of in-scope units)

What should be reported to document the coverage errors?

- Over-coverage rates and misclassification rates
- Possible impact of errors on the results
- Actions taken for the assessment of under-coverage
- Information about the frame: reference period and updating actions

b. Measurement errors

Measurement errors are errors that occur during data collection and cause the recorded values of variables to be different than the true ones. Their causes are commonly categorized as:

- *Survey instrument:* the questionnaire or measuring device used for data collection may lead to the recording of wrong values.
- *Respondent:* respondents may give erroneous data.
- *Interviewer*, interviewers may influence the answers given by respondents.

What should be reported to document the measurement errors?

- The methods used to identify the errors
- Errors' impact on statistics (bias and increased variance of the results)
- Indications about the causes of measurement errors
- The efforts taken for questionnaire design and testing
- Information on interviewer training
- Information about mechanisms used for reducing measurement error

c. Processing errors

Once data have been collected, they pass through a range of processes before the final estimates are produced: coding, editing, weighting etc. Errors introduced at these stages are called processing errors.

A range of indicators, in the form of rates, can be calculated to monitor the different processes. *Editing and imputation rates* have been selected among others because of their importance in the management of non-response errors.

Editing rate

Editing rate refers to a single variable, it is calculated for the key variable(s) of the statistical product and it is a measure of the contribution of edited values to the final estimate. The editing rate is calculated as follows:

Editing rate (for the key variable y *)* = (Number of records edited for the variable y) / (Total number of records))

Imputation rate

Imputation rate refers to a single variable, it is calculated for the key variable(s) of the statistical product and it is a measure of the contribution of imputed values to the final estimate. The imputation rate is calculated as follows:

Imputation rate (for the key variable y) = (Number of imputed records for the variable y) / (Total number of records)

Processing errors cause bias and variation in the produced statistics, just like measurement errors do.

What should be reported to document the processing errors?

- Methods used for identifying the processing errors
- Calculation of editing rates and imputation rates
- Indications about the causes of processing errors
- Presentation of the processes put in place for controlling and reducing processing errors (coders' training, performance data of automatic coding software, data entry personnel's training, data editing used, imputation algorithms used).

d. Non-response errors

Non-response is the failure of a survey to collect data on all survey variables, from all the population units designated for data collection in a sample or complete enumeration. The difference between the statistics computed from the collected data and those that would be computed if there were no missing values is the *non-response error.*

There are two types of non-response:

- *Unit non-response,* which occurs when no data are collected about a designated population unit, and
- *Item non-response,* which occurs when data only on some but not all the survey variables are collected about a designated population unit.

Response rates

The extent of response (and accordingly of non response) is measured with *response rates,* as follows:

Un-weighted unit response rate= (Number of respondent units used in estimation) / (Number of units designated for data collection)

Weighted unit response rate = (Sum of the extrapolation factors of the respondent units used in estimation) / (Number of population units)

Un-weighted item response rate = (Number of units with a value for the item) / (Number of units in-scope for item)

Weighted item response rate = (Weighted number of units with a value for the item)/ (Weighted number of units in-scope for item)

The impact of non-response on the statistics is that it increases their variability and introduces bias. Variability increases because non-response simply reduces the available number of responses. Bias is introduced by the fact that non-respondents may be different than respondents in their values of some survey variables. Re-weighting of the units and auxiliary variables usually may be used in the estimation process for reducing bias.

What should be reported to document the non-response errors?

- Non-response (remaining after call backs or data collection from other sources, but before *imputation):* Calculation of unit and item non response rates for main variables, both un-weighted and weighted
- Imputation methods used,
- Statement of whether statistics and the variances have taken into account the non-response and estimation methods used (including methods for re-weighting)
- Findings about similarity or not between non respondents and respondents for the main survey variables
- Indications of remaining non-response impact on statistics (bias and possible extra variation not accounted for)
- Indications about the causes of non-response
- Information about mechanisms (e.g. incentives, legal obligations of respondents, interviewer training, randomized response) used for reducing non response

2.3.3. Timeliness and punctuality

a. Punctuality refers to the possible time lag existing between the actual delivery date of data and the target date when it should have been delivered, for instance, with reference to dates announced in some official release calendar, laid down by Regulations or previously agreed among partners. If both are the same, delivery is punctual. The punctuality is calculated as follows:

Punctuality of time schedule of effective publication =
(Actual date of the effective publication)- (Scheduled date of the effective publication)

b. Timeliness refers to the lapse of time between the delivery and the reference dates. The latter being the date (or the period) to which data mostly applies.

What should be reported to document the timeliness and the punctuality?

Calculation of punctuality of time schedule of effective publication**,**

Regulations, official timetables or other agreements for assessing the timeliness
The reasons for late delivery: bottle necks in the production phase, strikes, etc.

2.3.4 Accessibility and Clarity

Accessibility and clarity refer to the simplicity and ease for users to access the statistics using simple and user-friendly procedures, obtaining them in an expected form and within an acceptable time period, with the appropriate user information and assistance.

Accessibility refers to the physical conditions in which users can access statistics: distribution channels, ordering procedures, time required for delivery, pricing policy, marketing conditions (copyright, etc.), availability of micro or macro data, media (paper, CD-ROM, Internet...), etc.

Clarity refers to the statistics' information environment: appropriate metadata provided with the statistics (textual information, explanations, documentation, etc); graphs, maps, and other illustrations; availability of information on the statistics' quality (possible limitation in use...); assistance offered to users by the NSI.

Number and types of means used for disseminating statistics

The "*Number and types of means used for disseminating statistics*" can be simply assessed through two indicators as follows:

a. The "Number of means used for disseminating statistics" can be defined for a specific statistical product (e.g. the statistics from the labour force survey), as the total number of paper publications, diskettes, CD-ROMs, internet publications, databases provided to the users concerning the statistical product itself. It takes only integer values greater or equal to 0. For example, for one statistical product, they are produced:

- A thematic volume about the survey results
- Some tables published in a more general paper publication (i.e. Statistical Yearbooks)
- Provisional results available on internet
- A CD-ROM that contains micro-data

In the above example, as for the statistical product the means used for disseminating statistics are four, the value of this indicator is equal to 4.

b. The "Types of means used for disseminating statistics" is a qualitative information provided to improve the interpretation of the previous indicator. For each statistical product it is necessary to provide the number of means of each type and the relative type.

The possible types are 6: paper publications (Statistical Yearbooks, Monthly Bulletins, Newsletters, Statistics in Focus, Press releases, other thematic publications), diskettes, CD-ROMs, Internet publications, reference databases (incl.

other specialized databases), answers to ad hoc users' queries. Consequently it can vary from 0 to 6, and assume only integer values.

The *proportion of types of means used* could be also calculated by dividing the number of different types by the maximum number 6.

Considering the previous example the indicator will be: 2 paper publications, 1 Web site (internet) publication, 1 CD-ROM. The proportion will be 0.5.

What should be reported to document the accessibility and clarity?

- A summary description of the conditions of access to data: media, support, marketing conditions, possible restrictions, existing service-level agreement, etc.
- Mention of the number and types of means used for disseminating statistics
- A summary description of the information accompanying the statistics (documentation, explanation, quality limitations, etc)
- A summary description of the possible further assistance available to users
- A presentation of possible improvements, compared to the previous situation.

2.3.5. Comparability

Comparability aims at measuring the impact of differences in applied statistical concepts and definitions on the comparison of statistics between geographical areas or over time.

- *Geographical comparability:* it refers to the degree of comparability between similar surveys that measure the same phenomenon, are conducted by different statistical agencies and are referring to populations in different geographical entities
- *Comparability over time:* it refers to the degree of comparability between two survey instances

What should be reported to document the geographical comparability?

- Description of all concepts and methods that can affects the comparability of the results across regions (e.g. countries).
- Differences between national practices and European standards (if such standards exist)

What should be reported to document the comparability over time?

- The period of the survey where the break occurred either (a) whether the difference reported is an “once-off adopted policy” with limited implications for the time series or an adopted policy for the future or (b) if the reported change led to a harmonization with any standards
- The difference in concepts and methods of measurement before and after the break.
- A description of the difference (e.g. changes in the classification, in the applied statistical methodology, in the target population, in the methods of data collections and manipulation, etc.).

– Assessment of the magnitude of the effect of the change in a quantitative way.

2.3.6 Coherence

Where similar statistics from various sources exist, they should be identified and any differences should be quantified and explained. A discrepancy between two sets of statistics produced by different surveys may be due to differences in the data collection processes or differences in reporting units resulting in different estimates.

When originating from different sources, and in particular from statistical surveys of different nature and/or frequencies, statistics may not be completely coherent in the sense that they may be based on different approaches, classifications and methodological standards. Usually, the types of coherence are as follows:

– Coherence of annual and short-term statistics
– Coherence of statistics in the same socio-economic domain
– Coherence of statistics with national accounts

What should be reported to document the coherence?

a. Coherence of annual and short-term statistics

A comparison on an annual basis of the growth rates of the produced statistics is conducted. If differences are not fully explained differences should be investigated and assessed (e.g. different data collection methods, different target populations)

b. Coherence of statistics in same domains

– Annual differences for the common characteristics according to accuracy component and differences in concepts
– Estimation of asymmetries due to the differences in concepts and in accuracy.

c. Coherence with National accounts

A summary contains the results of the comparisons which are conducted for the study of the coherence.

2.3.7. Completeness

Completeness is the extent to which statistics are available - compared to what it should be available - for meeting the requirements of the European Statistical System. There are clear relations between completeness and relevance. Furthermore, the availability of statistics is, in some cases, limited by accuracy and confidentiality reasons. The completeness is assessed through the "*rate of available statistics*" as follows:

Rate of available statistics= (Number of values provided) / (Number of field applicable)

2.3.8 Cost and Burden

Cost and respondent burden are aspects of the quality assessment task in the sense that quality of statistics cannot be regarded as isolated from them. The assessment of the *cost* associated with a statistical product is a rather complicated task since there must exist a mechanism for appointing portions of shared costs (for instance the business register or shared IT resources and dissemination channels) and overheads (office space, utility bills etc) and must be detailed and clear enough so as to provide for international comparisons among agencies of different structures. However, a proposal for measuring the cost of a survey may be as follows:

a. Number of staff involved:

Total number

Of which, professional and managerial

b. Cost for the statistical authority

- Staff cost
- Data collection costs (Printing, Mailing, Interview Costs = Number of fully completed questionnaires x Cost per questionnaire)
- Costs for treatment of non-response (Post, Telephone, Interview)
- Other costs

Regarding the *response burden***,** it can not be easily materialized in financial terms, but rather in time spent for filling up questionnaires or responding to an interviewer. The response burden involves two components:

R=the number of respondents

T= total time required to provide the information (including and the time spent assembling information prior to taking part in the interview), and the time taken up by any subsequent contacts after receiving of the questionnaires

The estimation of the response burden is: *Response burden = R x T*

What should be reported to document the cost and burden?

Costs supported by National Statistical Institutes
An evaluation of the response burden, only in physical terms (time required for response)

2.4 Case study on the documentation of the Greek Labour Cost Survey

The documentation of the Greek Labour Cost Survey 2004 accompanied with the applied quality criteria are presented in the *annex 1.* Through this documentation which is the form of a quality report: (a) the users can have access to a range of relevant quality measures and indicators so that they can understand the strength and limits of the statistics and know to use them properly and (b) the producers to have a picture on the product quality in order to identify the points of further improvements. Additionally, this documentation provides information on the main quality characteristics of the product, so that the user to be able to assess the product quality and the quality manager to monitor the product quality and to improve processes.

3. Process variables and quality process indicators

The production of statistics is achieved through the planning and operational processes. Thus, for improving the product quality it is necessary to study, monitor and improve the quality of the underlying processes through measuring the key process variables. As key process variables are defined the factors linked with the processes and also have the largest effect on main survey characteristics, i.e. characteristics indicating the product quality (Jones and Lewis, 2003). Therefore, for measuring the process quality, primarily it is necessary quantitative process indicators to be defined after (a) identifying the critical products characteristics and (b) developing a process flow map depicting the processes, which will be applied for the statistical outputs.

The identification of the critical products characteristics is achieved recording the users' needs through the study and analysis of data collected from the users' satisfaction surveys. After forming the research problem on the basis of the of users' needs, all processes which will be applied for the production of statistics are recorded, in order to determine key process variables. It is important to determine which variables are essential and critical given product requirements and also to evaluate the measurement capability of these variables (*Körner, Bergdahl, … [et al.], 2007).*

Quality management should demand not only quality output indicators but also quality process indicators assessing the quality of processes, in order to improve planning and operational processes of the surveys. The processes indicators can set up a system of continuous monitoring of processes, which is required to be sure the management system remains high. Of course this system always is an objective for improvements over time in order the process control system to ensure best processes for output production of high quality.

Statistical services have always measured some process variables. Examples of process variables are (a) resources used, (b) time used, (c) response rates, (d) response burden and (e) error rates in data collection and processing. However, a systematic approach is required to identify and measure these variables. A method for systematic and continuous measurement of process variables mainly is based on recording analytically all processes and sub-processes followed in statistical surveys and next compiling process variables having strong correlation with the quality of the statistical output.

3.1 Main processes and sub-processes of the surveys

The usual processes and sub-processes applying in the surveys are as follows (Jones and Lewis, 2003):

a. Survey design

- Performing the research problem based on users' needs
- Determine target population
- Selection of method for the data collection (census, accessing to administrative data, sample survey)
- Develop measurement instruments (e.g. questionnaires) including testing

- Develop field procedures, including testing
- Develop edit / imputation strategies
- Develop data management strategies
- Develop dissemination strategies
- Document

b. Accessing to administrative data

- Arrange access (including any legal issues)
- Document data that is accessed

c. Sample design in the case of sample surveys

- Determine and created the frame
- Determine the sampling scheme
- Determine the estimation methods
- Design and allocate the sample in domains and strata
- Sample selection
- Allocation of sampling units to interviewers for interview based surveys and dispatch of workloads

d. Data collection

- Dispatch of mail based or electronic questionnaires to respondents
- Interviewing for interview based surveys
- Resolution of queries relating to selected units
- Management of collection, including quality assurance processes and monitoring of progress
- Follow up procedures, including re-issue of sample to interviewers and reminders to mail based respondents
- Document procedures and outcome of processes

e. Editing and validation, imputation and coding

- Unit level editing and validation
- Imputation and construction
- Derivation of variables
- Quality assurance of processes
- Document procedures and outcome of processes

f. Weighting and estimation

- Weighting and grossing-up
- Outliers
- Special adjustments
- Sampling errors and quality assurance based on sampling errors
- Document procedures and outcome of processes

g. Analysis of primary outputs

- Macro editing for detection further errors at unit level based (a) on spatial and time comparisons, (b) on coherence with statistics of the same socio-economic domain and (c) on unexpected high sampling errors
- New corrections of the detected errors
- Tabulations

h. Further analysis

- Quality assessment of the results based on spatial and longitudinal analysis
- Validate the results
- Document including the quality report

i. Confidentiality and Disclosure

- Identify user requirements for output and priorities
- Identify potentially disclosure information
- Apply methods for protection the confidentiality
- Document

j. Dissemination (data and metadata)

- Dissemination of standard aggregated outputs including text, diagrams, numbers, through (a) paper and electronic publications and (b) web sites
- Dissemination of tailor made outputs
- Dissemination of micro-data in a form in which confidentiality is protected
- Dissemination of metadata and quality reports
- Document processes

3.2. Main process variables

Taking into consideration the processes and sub-processes followed in the statistical surveys, process variables are defined connecting the processes with the survey results and especially with the indicators assessing the quality of the statistical output. The following table presents a set of process variables which mainly influence the quality indicators: "Relevance and completeness", "Accuracy of the results", "Timeliness", and "Comparability", "Coherence" and "Accessibility and Clarity".

Table: Process variables

Process	Process variable
	Completeness according to the users' needs
	Expected sampling errors and under coverage errors (e.g. due to cut off sampling)
	Sampling fraction
Survey design	Number of stratification criteria
	Number of final strata
	Expected comparability over time and across countries

Process	Process variable
	Expected coherence with other surveys data according to the definitions and coding of the variables
	Expected cost of the survey
	Expected working staff for the survey
Questionnaire design	Length of the questionnaire: (Number of questions / Number of surveyed variables)*100
	Percentage of open questions with respect to the total numbers of the questions
	Expected average time for filling in the questionnaires according to the data from the pilot survey conducted for testing the questionnaire
Training	Number of seminars for the training of the interviewers
	Time spent for seminars aiming at the training of the interviewers
	Number of handbooks with guidelines on the correct filling in the questionnaires and the timing on survey procedures
Sampling frame updating	Ineligibility index due to non-effective sampling frame updating: Percentage of sampling units found not eligible in the total sample={Ineligible units / (Eligible units + Ineligible units)}*100
Data collection	Percentage of unit non-response (gross response rate): Percentage of responses in total unit={Responses / (Eligible units +Ineligible units)}*100
	Net response rate: Percentage of responses in eligible units= (Responses / Eligible units)*100
	Refusal rate: Percentage of refusals in the total eligible units= (Refusals / Eligible units)*100
	"Temporary away" rate for household surveys: Percentage of temporary away units in eligible units = (Temporary away units / Eligible units)*100
	Percentage of item non-response for key variables
	Percentage of measurement errors for key variables
Data collection (interviewing activities)	Interviewing time for filling in the survey's questionnaires
	Travel time of interviewers
	Total interview time of the survey
	Planned total interview time of the survey
	Working hours per man for the data collection
	Planned working hours per man for the data collection
Data processing (editing)	Percentage of measurement errors detected and corrected
	Time spent for editing
	Time planned for editing
	Working hours per man for the data processing
	Planned working hours per man for the data processing
Data processing (imputation)	Percentage of imputed values per variable
	Percentage of questionnaires in which all the values are imputed in the case of the unit non-response
	Time spent for imputation
	Time planned for imputation
	Working hours per man for the imputation
	Planned working hours per man for the imputation
Data processing (coding)	Miscoding rates
	Time spent for detecting and correcting miscodes
	Time planned for detecting and correcting miscodes
	Working hours per man for detecting and correcting miscodes

Process	Process variable
	Planned working hours per man for detecting and correcting miscodes
Weighting and estimation	Criteria used for weighting and estimation (probability selections, auxiliary information, post-stratification, calibration etc) after taking into account the non-response rates and the frame errors
	Time spent for weighting, estimations and tabulations
	Planned time spent for weighting, estimations and tabulations
	Working hours per man for weighting, estimations and tabulations
	Planned working hours per man for weighting, estimations and tabulations
Analysis of primary outputs (tabulation, checking and corrections)	Percentage of errors detected and corrected at macro-level with respect to the total number of records
	Time spent for macro-editing
	Planned time spent macro-editing
	Working hours per man for macro-editing
	Planned working hours per man for macro-editing
Confidentiality and Disclosure	Percentage of loss of information of the basic variables due to disclosure control methods applied for protected the confidentiality of the tabular data
	Percentage of the number of suppressed cells in tabular data due to primary and secondary suppression with respect to the total numbers of cells
	Percentage of loss of information of the basic variables due to disclosure control methods applied for protected the confidentiality of micro-data
	Time spent for protecting the confidentiality of tabular data
	Expected time spent for protecting the confidentiality of tabular data
	Time spent for protecting the confidentiality of micro data
	Expected time spent for protecting the confidentiality of micro data
Disseminations	Number of means used for disseminating and documenting aggregated data
	Number of means used for access to and documenting the micro-data

Monitoring the statistical processes through the process variables it is possible to find solutions and implement corrective measures for improving both the quality of process and quality of the produced statistics.

3.3. Quality process indicators

The process variables are linked with the processes and have largest effect on the statistical output. However, measuring only the process variables, it is not possible to assess the effect of the quality of the processes on the quality of the results. For example measuring the response rate or the measurement errors, we do not know the effect of the non-response or the measurement errors on the quality of the results. Thus, it is required process quality indicators to be defined, measuring the effect of the process quality on the quality of the produced statistics. More analytically, the quality process indicators should have strong relationship with the indicators applied for measuring the quality of the statistical output. In conclusion: *"As quality process indicators are defined the indicators measuring the effect of the process quality on the quality of the produced statistics."*

In the next capitals 4, 5 and 6, some important process quality indicators are presented and calculated using data from statistical surveys conducted by the NSSG in order to show how a survey process can affect the quality of the results. The quality process indicators are presented according to the three main survey stages (planning, operation and evaluation).

4. Planning process quality indicators

4.1 Identification of users' needs

The starting point of the planning process the starting point every statistical survey has to do with recording the users' needs. Conducting users' satisfaction surveys and taking into account the requirements of Commission Regulations, International Organizations (IMF, OECD, etc), the users' needs and the users' demands on product quality (especially timeliness, accuracy and accessibility of statistics) are identified.

The analysis of the collected data of the users' satisfaction surveys is based on both recording the users' needs for statistics and users' demands of product quality, which encompasses desired attributes of timeliness, accuracy and accessibility of statistics.

4.1.1 Recording the users' needs for statistics

This is achieved through compiling the relevance indicators as well as the analysis of the relationship between relevance indicators, which provides more composite picture of the whole relevance.

a. Relevance indicators (Eurostat 2003a and 2003b)

i. The percentage of statistics which meet the users' needs stated by a scale of satisfaction (e.g. completely, partially, not at all)
ii. The fields of the statistics that do not meet the users' needs (etc. the variables $\chi_1, \chi_2, \ldots$, the regional statistics NUTS III)
iii. The main core users (This indicator is indispensable, because the questions of the questionnaire for the data collection are formed after taking into account the requirements of the International Organizations and the consultation with the main national users)
iv. The used categories of statistics
v. The main purposes for which the users need the statistics (e.g. analysis of current developments for short-term decision making, analysis of trends for longer-term decision making, forecasting, research purposes etc).

b. Relationship between relevance indicators

The relationship between relevance indicators provides a more complete picture of the relevance, because the combination of two or more indices can be examined and studied, simultaneously. This relationship is possible to be depicted on the factor level, using *Correspondence Factor Analysis* (Benzecri, 1992), as in the examples in the annex 2, because this analysis is based on non-continuous data, without supposing that the population approximately follows the normal distribution. The investigation and the analysis between relevance indicators is a useful tool for the survey design, the questionnaire design and especially for setting up the working groups for the design and conducting the survey.

4.1.2 Users' demands of product quality

The analysis of users' demands of product quality is based on the opinions of users especially about the data accuracy, timeliness and accessibility of statistics. The users' demands are recorded through compiling suitable indicators, as well as the relationship between these indicators.

a. Indicators of users' demands of product quality

These indicators are usually the percentage rates of "positive", "negative" or "no opinion" of users on accuracy or timeliness or accessibility of statistics separately by "type of users" or "category of statistics" or "purpose for which statistics are used".

b. Relationship between indicators of users' demands of product quality

The relationship between indicators of users' demands of product quality is possible to be depicted on factor level using *Correspondence Factor Analysis,* because and in this case the analysis is based on non-continuous data, as in the examples of the *annex 3.*

4.2 Circumscription and reckoning survey financing

After recording and analyzing the "users' needs" and the "users' demands on the product quality" the formation of the research problem is carried out. Taking into accounts all the requirements of the research problem, the strategic decisions on the data collection methods and the planning of the data collection are carried out in connection with the resources, such as budget, personnel and other equipment. These strategic decisions investigate the needed funds and the available funds and thus an indicator is necessary to be compiled showing how the available funds (and generally the available resources) can affect the decisions while designing the survey and the quality of the final results. This indicator may be defined as:

Financing = available funds / needed funds

As the cost of a survey is strongly dependent on the number of respondents (sample size), the indicator "*Financing"* is approximately equal to the fraction:

Sample size defined according to the available funds / optimum sample size

An example on **c**ircumscription and reckoning survey financing is presented for the trade section of the Greek Structural Business Surveys (year 2006). According to the users' needs the research problem was performed, which requires accurate data (relative sampling error less than 5%) at four-digit level (NACE Rev1) for each separate geographical region (NUTS II). If the needs of the research problem are met, then the sample size is 8.650 enterprises (optimum or needed sample size). So, the optimum sample size determines the needed funds of the survey. Taking into consideration the available funds the research problem was modified in order to require accurate data (relative sampling error less than 8%) at three-digit level (NACE Rev1) for each separate geographical region (NUTS II). Now, the new modified research problem needs sample size 6.000 enterprises (sample size defined according to the available funds). Thus:

"Financing" =available funds / needed funds ≅ Sample size defined according to the available funds / optimum sample size =6000 / 8650.

4.3 Sample design

The design of the survey is conducted on the basis of the results coming from the analysis of the "users' needs" and the "users' demands on the product quality". The decisions about the three modes of data collection (sample, census, administrative sources) are based on criteria, which are related to the cost, coverage, geographical details, timeliness, relevance, credibility etc (Kish 1979). The analysis of the sample design process is presented in order to show how the quality of this process can affects the accuracy of the produced statistics.

Having been defined the target population, the sample frame is constructed and subsequently, the sample design is carried out (strategic decision on the applied sampling scheme, stratification of the population units, determination of sample size in strata, selection of the sample units).

As an important part of the products quality is dependent on the sample design, it is required process variables to be defined for monitoring and assessing the quality of the sample design process. As process variables may be considered:

- The co-efficient of correlation between the stratification variable and the main survey variables (characteristics), which shows the correct decision on the stratification variable
- Indicator measuring the total effect of the applied sampling scheme on the sampling errors of the survey variables
- The design effect for the basic survey variables, measuring the effect of stratification, weighting and clustering on the sampling error (Kish, 1995)

4.3.1 Coefficient of correlation between the stratification variable and the main survey variables

In stratified random sampling, the stratification usually produces gains in precision of the estimation of the survey characteristics, because the stratification may divide a heterogeneous population into subpopulations (strata), each of which is internally homogeneous. If each stratum is homogeneous, in that the measurement vary little from one unit to another and a precise survey characteristics estimations of any stratum can be obtained from a small sample in that stratum. These estimations can then be combined into precise estimations for the whole population.

The best variable for the construction of homogeneous strata is the variable, which is strongly correlated with the main survey variables (characteristics). Thus, the successive choice of the stratification variable is determined on the basis of the correlation coefficient values of the stratification variable with the main survey characteristics. The largest correlation coefficient values, the largest gains in precisions of the survey characteristics estimations are achieved due to stratification. So, the coefficient of correlations between the stratification variable and the main survey variables may be used as a process variable assessing the quality of the stratification conducted in a stratified sample design, because is a key process variable linked with the stratification, which has the largest effect on the accuracy of the survey characteristics estimations. For this reason, the coefficient of correlation between the stratification variable and the main survey characteristics is required to be accompanied the coefficient of variations (%) of the main survey characteristics in order to show the gains in precision because of the stratifications. Additionally, it will be proved how the coefficient of correlation r can be related to the coefficient of variations CV (%).

a. Continuous stratification variable

When the stratification variable is continuous the coefficient of correlations r are calculated applying the "traditional" formula: $r = \dfrac{S_{xy}}{S_x \cdot S_y}$, where S_{xy} is the covariance between two variables x and y and S_x and S_y are the standard deviations of the variables x and y.

Example 1: In the Greek Structural Business Survey with reference year 2005 (one stage stratified sampling scheme has been applied, see annex 4), using sampling data (a) the coefficient of correlations of the stratification variable (turnover) with the main survey characteristics and (b) the coefficient of variations of the main survey variables were calculated as they are depicted in the following tables 1 an 2.

Table1. Coefficient of correlations between the turnover and the main survey characteristics

Economic activities	Main survey variables			
	Value added	Personnel costs	Gross investments at tangible goods	Number of employees
Constructions	0,927	0,925	0,132	0,924
Repair of motor vehicles and motorcycles	0,999	0,777	0,428	0,646
Wholesale trade	0,999	0,736	0,539	0,568
Retail trade	0,998	0,963	0,733	0,554
Hotels and restaurants	0,962	0,954	0,329	0,943
Land transports	0,878	0,860	0,578	0,867
Air transports	0,498	0,999	0,376	0,965
Supporting and auxiliary transport activities and Travel agencies	0,961	0,739	0,386	0,771
Post activities and Telecommunications	0,988	0,798	0,381	0,775

Table2. Coefficient of variations (%) of the main survey characteristics

Economic activities	Main survey variables			
	Value added	Personnel costs	Gross investments at tangible goods	Number of employees
Constructions	2,7	3,9	13,9	4,0
Repair of motor vehicles and motor cycles	1,5	4,0	5,5	4,4
Wholesale trade	1,2	3,0	6,9	2,8
Retail trade	2,8	4,7	4,8	4,6
Hotels and restaurants	2,0	1,9	9,8	2,4
Land transports	1,2	0,9	0,9	2,5
Air transports	0,7	0,6	10,7	1,2

Economic activities	Main survey variables			
	Value added	**Personnel costs**	**Gross investments at tangible goods**	**Number of employees**
Supporting and auxiliary transport activities and Travel agencies	1,6	2,0	70,1	3,4
Post activities and Telecommunications	0,1	0,0	0,0	0,1

According to data in the above tables 1 and 2, as the coefficient of correlations r increase the coefficient of variations $CV(\%)$ decrease. Some exceptions may be appeared in economic activities in which high percentage of turnover (stratification variable) is collected by few dominant large scale enterprises surveyed on a census basis (e.g. Air transport, Postal activities and Telecommunications).

Example 2: In the Greek Job Vacancy Quarterly Survey of the year 2006 (one stage stratified sampling scheme has been applied, see annex 5), using sampling data (a) the coefficient of correlation of the stratification variable (number of employees) and the "number of job vacancies" and (b) the coefficient of variations by quarter of the main survey variable "number of job vacancies" were calculated as they are depicted in the following tables 3 an 4. This example is different than the previous one, because it shows how almost negligible gains of the precisions of the results are appeared, when the coefficient of correlation values of the "number of employees" with the "number of job vacancies" by economic activities are small.

Table 3: Coefficient of correlations between the stratification variable "number of employees"and the main survey characteristic "number of job vacancies" by economic activities

Economic activities (Codes NACE Rev.1)	**Coefficient of correlations**
C	0,07
D	0,17
E	0,24
F	0,15
G	0,38
H	0,09
I	0,28
J	0,42
K	0,05
L	0,83
M	0,50
N	0,07
O	0,42
Total	0,21

Table 4: Coefficient of variations of the "number of job vacancies" by economic activities

Coefficient of Variations %				
Economic activities (Codes NACE Rev.1)	Quarters			
	1	2	3	4
C	54,5	39,5	50,8	42,4
D	13,8	17,3	18,8	18,2
E	20,0	9,0	7,4	13,0
F	56,8	55,2	45,9	61,1
G	17,8	18,8	25,1	19,5
H	31,9	35,0	26,1	37,0
I	31,7	12,3	9,9	24,5
J	53,6	34,1	29,9	51,6
K	49,7	63,7	47,4	71,3
L	0,0	0,0	0,0	0,0
M	0,4	1,5	1,2	6,3
N	86,7	11,1	36,3	42,4
O	28,8	29,9	27,9	28,6
C-O	13,6	7,7	9,0	17,0

The fact that the coefficients of variation at section level and total level appear to be quite large (table 4), is owing mostly to the small correlation that exists between the number of employees in a enterprise (which was used for the determination of the classes) and the respective number of job vacancies, as the values of table 4 depict.

In the section L, the $CV(\%) = 0$, because the statistics come from administrative sources. In the section M, values of $CV(\%) \leq 6{,}3$, although the coefficient of correlation is small ($r = 0{,}5$), because the construction of statistics is based on combination of sample and administrative data.

b. Non-continuous stratification variable

When the stratification variable is non-continuous variable (e.g. degree of urbanization: urban, semi-urban, rural areas), then the coefficient of correlation between stratification variable and the main survey variables is not possible to be calculated through the "traditional" formula applied for continuous variables.. In this case, the relationship between the stratification variable and the main survey variables can be depicted on the "factor level" through the application of the "*Correspondence Factor Analysis*". In the annex 6, an example of monitoring the relationship between the stratification variable and the main survey variables through the application of *Correspondence Factor Analysis*" is presented for the Greek EU-SILC survey.

Applying correspondence factor analysis, the relationship of non-continuous stratification variables with the main variables can be, approximately, monitored. In this case the picture of the "factor level" provides the quality manager with an important tool for monitoring and assessing the successive choice of the stratification variable applied for creating homogeneous strata, when the stratification variable is non-continuous.

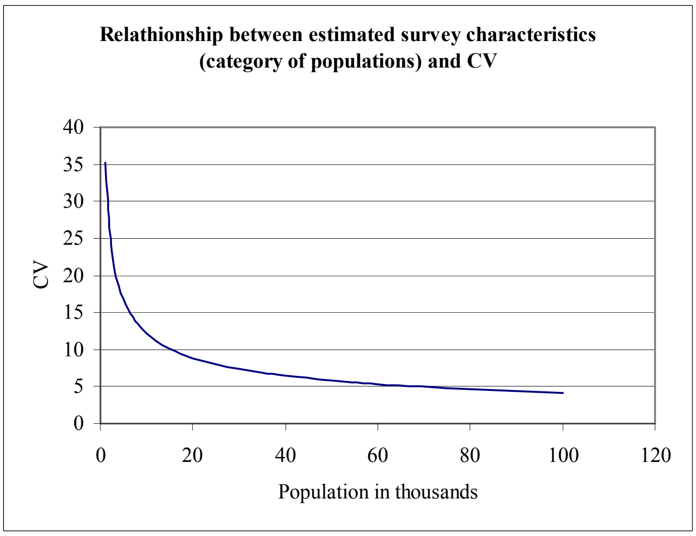
4.3.2 Effect of the applied sampling scheme on the sampling errors of the survey variables

The strategic decision on the applied sample design scheme of a survey takes into consideration two criteria: (a) the research problem (defined by the "users' needs" and the "users' demands on the product quality" and (b) the available resources, such as budget, staff and other equipment. After conducted the sample design based on the above two criteria above, it is necessary to be monitored the effect of the applied sampling scheme on the sampling errors of the survey characteristics. Thus, a process indicator is required compile for assessing how the quality of the process "sample design" affects the accuracy of the produced statistics. In other words, after the sample design (stratification, determination of optimum sample sizes in strata) it is necessary to measure and presented the expected sampling errors (in the form of coefficient of variations) of the survey characteristics through a formula which links the values of the survey characteristics (e.g. number of employees, area cultivated with cereals) with the expected coefficient of variations.

An example of a formula measuring the relationship between values of characteristics and the coefficient of variations is presented for the Greek Labour Force Survey, under a given sampling scheme (two-stage stratified sampling scheme, with sampling fraction 0,85% of the total size of households). Analysing the variance of the estimated characteristics from the results of a previous survey a formula connecting the estimated value $\widehat{Y}$ of the characteristic y with the coefficient of variation $CV(\widehat{Y})$ is compiled. This formula has the form:

$$CV(\widehat{Y}) = \frac{35{,}264864}{\widehat{Y}^{0{,}461818}}$$

The following diagram depicts the expected coefficient of variation (CV%), in relation to the frequency of the survey characteristics expressed as population in thousands.

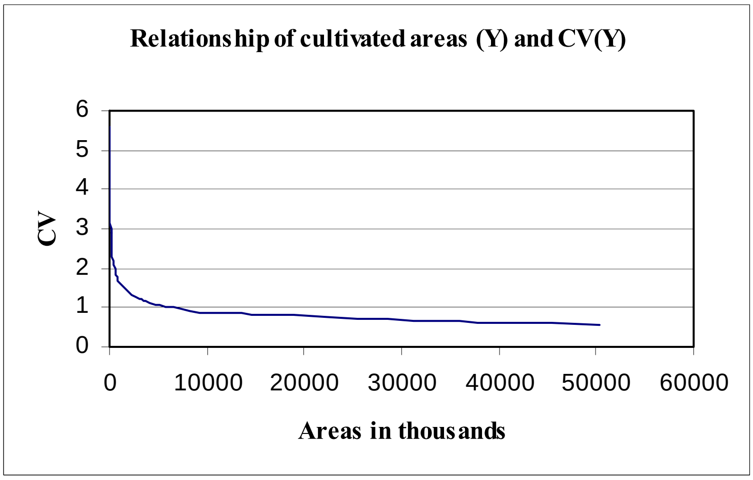


According to the diagram the population (e.g. employees in electricity), with frequency lower than 60.000 appears coefficient of variation larger than 5% and as the coefficient of variation decreases, the population frequency increases according to the data calculated using the above formula. Analysing the data coming from this formula, the quality manager has a complete picture regarding the effect of the applied sampling scheme on the coefficient of variations of all survey characteristics.

Another example of a formula measuring the relationship between values of characteristics and the coefficient of variations is presented for the Greek Farm Structure Survey, under a given sampling scheme (one stage stratified sampling scheme, with sampling fraction 11% of the total size of holdings and stratification criteria as follows: geography-NUTS III, type of holdings, economic size of holdings. Analysing the variance of the estimated cultivated areas a formula connecting the estimated cultivated area $\widehat{Y}$ of the characteristic y (e.g. area with olive trees) with the coefficient of variation $CV(\widehat{Y})$ is compiled. This formula has the form:

$$CV(\widehat{Y}) = \frac{69}{\widehat{Y}^{0,27}}$$

The following diagram depicts the expected coefficient of variation (CV%), in relation to the frequency of the survey characteristics expressed as cultivated areas thousands.



4.3.3. Conclusions regarding the effect of the sample design on the sampling errors

- The coefficient of correlation between the stratification variable and the main survey variables may be considered as process indicator assessing the successive choice of the stratification variables for creating homogeneous strata.
- The effect of the applied sampling scheme on the sampling errors of the survey variables may be considered as a process indicator which is compiled through a suitable formula which links the values of the survey characteristics with the expected coefficient of variations under a given sample design.

Apart from the above process indicators assessing the quality of the process “sample design”, there is one more indicator the design effect (Kish 1995), which assesses the quality not only of the sample design but also of the data collection process of the complex sample surveys. For this reason, the design effect as a process quality indicator assessing the stratification, weighting and clustering is presented and analysed separately in the next capital regarding the operational process variables.

5. Operational process quality indicators

5.1 The effects of weighting, clustering and stratification on sampling errors as process quality indicators

a. Stratification

Stratification may produce a gain in precision in the estimates of characteristics of the whole population. It may possible to divide a heterogeneous population into subpopulations, each of which is internally homogeneous. The gains in precision in stratified sampling for means or proportions or ratios may be quantified for assessing the effect of the stratification on the sampling error. The factor denoting the decrease of sampling error due to the stratification of the surveyed population can be used as quality indicator, measuring the positive effect of the stratification on the sampling error.

b. Weighting

Weighting adjustments are primarily used to compensate for unit non-response. The essence of all weighting adjustment procedures is to increase the weights of specified respondents, so that they represent the non-respondents. The procedures require auxiliary information on either the non-respondents or the total population. Moreover, the population weighting adjustments attempt to reduce the bias created by non-response and coverage errors. The non-response and population adjustment reduces the bias, but it results increased variances of survey estimates and the sampling errors due to random weighting. The factors denoting the increase of sampling error due to non-response and population adjustments can be used as quality indicators, indicating the effect of non-response and coverage errors on the sampling errors after conducting weighting adjustments. These indicators can be calculated both for one-stage and multi-stage stratified surveys.

c. Clustering

In multi-stage surveys, the sampling errors increase considerable due to clustering. The effect of clustering on the sampling errors of the survey estimates can be used as a process variable assessing the quality of the sample design.

In the *annex 7*, a complete analysis is presented on how the effects of stratification, weighting and clustering on sampling errors may be considered as process quality indicators. These indicators are defined and compiled through the components of the design effect, as the design effect represents the combined effect of stratification, clustering and unequal weighting (Kish 1995).

Additionally, in the annex 7, the above indicators are applied in the Greek Labour Force Survey with reference period 2nd quarter of 2007 as a case study.

d. Conclusions

- The non-response and the coverage errors create bias, the effect of which is unknown. Thus, weighting adjustments attempt to reduce the bias that non-response and coverage errors may cause in survey estimates. The weights for non-response and population adjustments are random variables provoking increases in the sampling errors of the survey estimates. These increases tend to persist undiminished for most subclasses and for all statistics. The indicators quantifying the effect of random weighting on sampling errors and design effects may be used as quality indicators assessing the data collection process and the weighting process targeting to remove the bias due to non-response and coverage problems.

- The total weighting (random weighting and weighting due to disproportionate allocation, unequal selection probabilities) usually increases the variances. The percentages of these increases may be used as quality process indicators for assessing (a) the quality of the sample design and (b) the effect of random weighting on the results.
- The intra-class correlation measures the effect of clustering on the sampling errors and it may be used as a design process indicator assessing if the homogeneity of the sampling elementary units in clusters influences the precision of the results. Additionally, this indicator evaluates the efficiency of the cluster sizes, and shows if the elementary final sampling units have been spread in an efficient number of sample clusters. The importance of the intra-class correlation as an indicator is significant, because the intra-class correlation in combination with the average cluster sizes and the cluster effect links the sample design with the sampling error.
- The stratification effect is an important process quality indicator assessing the gains in precisions due to stratification of surveyed units.

5.2 The effect of measurement errors on the results

Measurement errors are errors that occur during data collection and cause the recorded values of variables to be different than the true ones. Their causes are commonly categorized as:

- *Survey instrument:* the questionnaire or measuring device used for data collection may lead to the recording of wrong values.
- *Respondent:* respondents may, consciously or unconsciously, give erroneous data.
- *Interviewer,* interviewers may influence the answers given by respondents.

Measurement errors may cause both bias and extra variability of the produced statistics. In order to assess instrument or interviewer effects on the produced statistics repeated measurements would have to be taken with different instruments (e.g. alternative phrasing of questions) or different interviewers. Alternatively an experiment should be carried out with sub-samples being randomly allocated to different instruments and /or interviewers. Respondent effects are even harder to assess, requiring independent sources of information about the same respondent (Eurostat 2003a and b).

5.2.1 Extra variability due to measurement errors

Besides sampling errors (those arisen in selection or estimation procedures), survey results are also affected by errors, which occur in the course of the observation (measurement), recording and processing of the data. From these errors important are the variable (measurement) errors, because these errors cannot be distinguished from sampling errors among respondents unless replicate measurements are taken on the respondents. In general, they can be regarded as random errors, which increase the variance with contributions, which enter automatically the computations of the variance. However, this extra variance (interviewer variance) due to measurement errors is important to be measured in order to assess the effect of measurement error on the accuracy of the results. This extra variance can be considered as a process quality indicator assessing the quality of data collection process, because the extra variance (interviewer variance) is linked with the data collection process and also has large effect on the accuracy of survey characteristics.

The interviewer variance s_a^2 (extra variance due to measurement errors) should be viewed as a component of the total population variance s^2, denoted as (Kish 1962):

$$s^2 = s_a^2 + s_b^2 \qquad (5.2.1)$$

Where, the quantity s_b^2 is the variance, without any effect of measurement errors.

The proportion of the effect of measurement errors is given by:

$$\rho = \frac{s_a^2}{s^2} \qquad (5.2.2)$$

In other words the quantity s_a^2 is the "between interviewer" and s_b^2 the "within interviewer" component. The definitions of s_a^2 and s_b^2 are as follows:

Symbolisms:

i : The order of the respondent i ($i = 1,2,...,m$)
m : The number of the respondents in which repeated interviews have been carried out
k : The number of the repeated interviews for each respondent
y_{ik} : The value of the variable y of the respondent i which in the k response
$\bar{y}_i$: The average value of the k repeated responses corresponding to the i respondent
s_i^2 : The variance of the y_{ik} values
s_a^2 : The average of the s_i^2 variances
s_b^2 : The variance of the $\bar{y}_i$ values

The component s_a^2 takes into account the mean variance appeared from the different values of the same respondent, and the quantity s_b^2 takes into account the variance of the mean values of the respondents. The values s_a^2 and s_b^2 are calculated using the data only from the respondents in which repeated interviews have been carried out.

a. Increase of the sampling error due to measurement errors in the one stage stratified random sampling

Symbolisms:

h : The order of the stratum
Y : The sum of the values of the characteristic y of all population units under the survey
N_h : The number of population units belonging to the stratum h
n_h : The sample size in the stratum h
s_h^2 : The sample variance in the stratum h

The variance estimation of $\hat{Y}$ is given by:

$$V(\hat{Y}) = \sum_h \frac{N_h \cdot (N_h - n_h)}{n_h} \cdot s_h^2 \qquad (5.2.3)$$

The coefficient of variation (%) of total estimation $\hat{Y}$ is given by:

$$CV(\hat{Y})\% = \frac{\sqrt{V(\hat{Y})}}{\hat{Y}} \cdot 100 \qquad (5.2.4)$$

If $V_a(\hat{Y})$ stands for the extra variance due to measurement errors, then according to the formula (5.2.2), the extra variance $V_a(\hat{Y})$ is given by:

$$V_a(\hat{Y}) = \sum_h \frac{N_h \cdot (N_h - n_h)}{n_h} \cdot \rho_h \cdot s_h^2 \qquad (5.2.5)$$

Due to the measurement errors, the increase (%) of the sampling error of the estimation $\hat{Y}$ is given by: $IS(\hat{Y})\% = \sqrt{\frac{V_a(\hat{Y})}{V(\hat{Y})}} \cdot 100$ (5.2.6)

In the case that the there were not measurement errors in the surveys, then the coefficient of variation would be equal to:

$$CV'(\hat{Y})\% = CV(\hat{Y})\% - IS(\hat{Y}) \cdot CV(\hat{Y})\% \qquad (5.2.7)$$

In the *annex 8,* an example on the extra variability due to measurement errors for the Greek survey of goats is presented.

As a conclusion, the percentage of the increase of the sampling error (or coefficient of variation) due to measurement errors measure the effect of the measurement errors on the accuracy of the produced statistics. This extra variability of the selected data can be considered as a process variable assessing the quality of the data collection process.

b. Bias due to measurement errors

Using data from independent sources of information about the same respondent, it should be investigated to what extent differences in basic characteristics are appeared between these sources of information. The analysis for detecting important differences may be based on t-tests for comparing means of continuous survey variables.

In the annex 8, an example on possible appearing bias due to measurement errors for the Greek survey of goats is presented.

As a conclusion, possible bias on the results due to measurement errors can be considered as an additional process variable assessing the quality of the data collection process.

5.3 The effect of non-sampling errors on the sampling errors

Sampling errors arise from the fact only a subset of the target population is enumerated (usually randomly selected). The sampling errors are dependent on (a) the population and the sample size and (b) the variability of the survey characteristics. The sampling errors increase, when either the sample size decreases or the variability of the survey characteristics increases.

The non-sampling errors, apart from the fact that they may create bias, increase the sampling errors because:

- The non-response and the over coverage (inclusion of non-population units) reduce the sample size
- The misclassification, the measurement errors and the processing errors (miscoded units, data entry errors etc) increase the variability of the survey characteristics.

5.3.1 Effect of frame errors on the sampling errors

a. Over coverage and misclassification

In the one stage stratified sampling surveys, the most important factor for increasing the variance of the results is the over coverage and misclassification and especially the misclassification of the surveyed units according to their sizes. Taking as an example the business surveys the wrong sizes (number of employees or annual turnover) of the enterprises in the business register increase the variability of the survey characteristics. This happens because at the time the sample was drawn the size classes of the enterprises were derived from the business register data of the year before ($t-1$). After data collection and data processing the size classes of the sample units may have been changed creating increase in the variability of the survey characteristics. Usually, after the data collection business register data of the year t (the survey's reference year) is available. In this case the following two events may be happen:

a. Adjusted strata

According to collected data (after data collection data) for every enterprise it is known in which one of the adjusted size class h it is placed. To adjust the weighting factor $w_k = \frac{1}{\pi_k}$ an extra factor c_h is used for every adjusted stratum h. The value of the correction factor c_h follows the formula $\sum_{k \in h} \frac{c_h}{\pi_k} = N_h$, where N_h the number of enterprises in size class h for the year t. The expression for the correction weight is $c_k = \frac{N_h}{\sum_{k \in h} \frac{1}{\pi_k}}$ and the adjusted weight is $w_k = \frac{1}{\pi_k} \cdot c_k$. The factor c_k also may be calculated applying the calibration equations (Deville and Sarndal, 1992) and in the same stratum it may take different values for the units k. The factor c_k usually increases the variability of the survey characteristics, as it is a random factor having been created after the data collection. The effect of the additional weighting c_k on the sampling error has been examined and analyzed in the capital 5.1.

b. Post stratification

After the data collection the units are post-stratified according to their new size class or post stratum h (post-stratification), under the condition that the size classes N_h are updated (Kish, 1992). That means the data of the size classes and the sample have the same reference year t stratification. The post-stratification produces unbiased estimators with increased variances due to deviations from the optimum sample allocation in the size classes (post-strata). The effect is given by the formula (Cochran, 1977):

$$\frac{V - V_{\min}}{V_{\min}} = \sum_h \frac{\hat{n}_h}{n} \cdot g_h^2 \qquad (5.3.1)$$

Where:

$\hat{n}_h$: the actual sample size in the post-stratum h

n_h: the optimum sample size in the post-stratum h

n : the total sample size

g_h: the absolute difference in the in the post-stratum h between the actual sample size and the optimum sample size expressed as a fraction of the actual sample size. That is: $\frac{|\hat{n}_h - n_h|}{\hat{n}_h}$.

The formula (5.3.1) may be considered as a quality indicator assessing the effect of coverage errors (over coverage and misclassification) on the variance of the estimated survey characteristics.

b. Under coverage of the sampling frame

Under-coverage is the most serious problem, because it can neither be detected from the frame nor from the sample. Population totals will almost always be underestimated, because part of the target population is out of the observation scope and as a result these totals are negatively biased.

For deducing the under-coverage biases in the two stage-stratified sampling surveys methods, in each primary sampling unit (area=one or more unified city blocks), before selecting the final sampling units (households or individuals) from the sampling frame (list), the sampling frame is updated.
In the agricultural surveys (although the one stage stratified sampling scheme are applied), in each municipality or commune to which at least one sampling unit belong, the lists of holdings are updated in order to reduce the biases, correcting the estimators after the elaboration of the data from the updated lists.

In business surveys the underestimated population of enterprises and employees due to under-coverage may be corrected from the Labour Force Survey data (two stage stratified survey in which the sampling areas are updated before the selection of households), considering that the number of self-employers with employees approximately coincides with the number of enterprises employing at least one person.

Sample surveys for checking the under coverage

In censuses (population or agricultural holdings) for checking the under-coverage and errors concerning the content of the census questionnaires, a special sample survey (population or agricultural holdings coverage survey) is conducted. Concerning the population coverage survey, the dwellings within an enumeration district, can be considered as belonging to (a) those which were missed at the census and (b) those which were actually visited by the enumerator. In the first case the coverage error is due to the failure on the part of the enumerator to list the dwelling whereas, in the second case, the error is due to the fact that the enumerator could not include those and only those persons who should have been enumerated. For checking the cases (a) and (b):

- Firstly, we select a sample of areas (e.g. enumeration districts) with probabilities proportional to their sizes (number of households according to the previous population section). The dwellings within the selected area are re-canvassed to examine if any were missed at the census, for those apparently missed at the interviewer to fill in a questionnaire, In this way, by means of the area sampling, all missing dwelling are surveyed for recoding all individuals having been missed in the census.
- Secondly, in each selected area, a systematic sample of dwellings, which were included in the census is selected and surveyed in order to see if any missing individuals are in these dwellings.

Applying the estimators used in the multistage stratified sampling surveys in which the primary sampling units (areas) are selected with probabilities proportional to their sizes and in each selected area the dwellings are selected with equal probabilities, the number of missed persons is estimated. This estimation of missed people in comparison with the census population provides the under-coverage rate of the census population.

In the agricultural holding coverage survey, the same sampling method is applied with the following differences:
- Instead of dwellings, we have holdings
- Instead of individuals, we have cultivated areas and number of animals

5.3.2 Effect of all non-sampling errors on the sampling errors

Taking into consideration the above cases (a) and (b) of the capital (5.3.1), the misclassification in size strata increases the variance of the estimated survey characteristics. The effect not only of the misclassification but also all the non-sampling errors on the sampling errors may be considered as quality process indicator assessing the quality of the total operational process (data collection and data processing). This effect can be estimated easily measuring the extra variance (or sampling error) of the estimated characteristics due to non-sampling errors as follows:

- According to the sample design the expected sampling errors of the main characteristics have been calculated (e.g. in the Structural Business Statistics and in the section of trade the sample design was conducted so that the coefficient of variation of the variable "turnover" to be less than 2% and of the variable "number of employees" to be less than 3%).
- After data collection and data processing the actual sampling errors (in the form of coefficient of variation) are usually greater than the expected sampling errors (e.g. the coefficient of variation of the variable "turnover" is 2,5% and of the variable "number of employees" is 3,8%)
- The percentage increase between the actual and the expected sampling error measures the effect of not sampling errors on sampling errors:

  Effect on non-sampling errors (%) =

  *[(Actual sampling error – Expected sampling error) / Expected sampling error] * 100* (5.3.2)

According to the above data used as an example, the effect of non-sampling errors on the variable "turnover" is: $\frac{2,5-2}{2} \cdot 100 = 25\%$ and on the variable "number of employees" is: $\frac{3,8-3}{3} \cdot 100 = 27\%$.

5.3.3 Effect of miscoding on sampling errors

Miscoding of the survey variables produce bias and increase the variability of the survey characteristics. The extra variability due to miscoding is not possible to be calculated easily, but it can be estimated indirectly from the following formula:

Total extra variance due to non-sampling errors =

Extra variance due to measurement errors + Extra variance due to non-response and frame errors + Extra variance due to miscoding (5.3.3)

Where:

- Total extra variance due to non-sampling errors = Actual variance – Expected variance according to the sample design (see 5.3.2)
- Extra variance due to non-response and frame errors = Extra variance due to reducing the sample size and random weighting, because of non response and over coverage + Extra variance due to misclassification estimated through the formula (5.3.1)
- Extra variance due to measurement errors is estimated through the formula (5.2.5)

The formula (5.3.1) for the estimation of the extra variance due to misclassification is applied only in the one stage stratified sampling survey.

The extra variance due to miscoding expressed as a fraction of the actual variance may be considered as a quality process indicator assessing the effect of the quality of coding on the variance of the estimates. The square root of the above fraction expresses the effect of the quality of coding on the sampling errors.

5.3.4 Conclusions regarding the effect of non-sampling errors on sampling errors

- The percentage increase between the actual and the expected sampling error due to non-sampling errors can be considered as a process quality indicator assessing the quality of operational process (data collection and data processing), because the extra sampling error is linked with the operational process and also has large effect on the accuracy of survey characteristics.
- In the case that the one stage stratified sampling scheme is applied with post-stratification for reducing the bias due to non-response and removing the bias due to frame errors (over-coverage and misclassifications), the increase in variance of the estimated survey characteristics can be measured through the formula (4.3.1). This may be considered as a quality process indicator assessing the effect of non-response, over-coverage and misclassifications on the precision of the survey characteristics.
- In the case that the one stage stratified sampling scheme is applied with application of adjusted strata, then the extra random weighting increases the variance. This case in details is examined and analysed in the capital 4.1.
- The effect of miscoding on the sampling errors can be estimated indirectly and may be used as a process quality indicator.
- The under coverage rates in the censuses are estimated through special sample coverage surveys

6. Evaluation process indicators

In the evaluation stage of the survey an analysis is conducted assessing if the final statistical products have been produced according to the concepts, definitions and quality criteria defined in the planning stage. This analysis is mainly based on the calculation of the quality indicators assessing not only the quality of the products but also the quality of the processes followed for the final statistical output. The values of the quality indicators accompanied with comments on these values are included in the quality report by means of which (a) the whole documentation of the survey is presented and (b) an auditing can be carried out evaluating the total process of the statistical survey and identifying areas for improvements. Additionally, the purpose of the product descriptions and the quality reports is to enable the end-users of statistics to be able to judge the factors, which restrict the quality of the statistics.

Although approximately the whole documentation of the survey is fulfilled through the quality report however some extra indicators may be compiled on:

- Whether the survey results comply with other similar survey results (coherence)
- Comparing potential and actual disseminated presentation forms (tables, variables, documentations, methodological reports) of results
- Comparing the documentation of the processes with the total number of the survey processes which could be documented

6.1 Analysis on whether the survey results comply with other similar survey results

The coherence of statistics is a basic indicator included in the quality report. However this indicator is usually limited to comparisons of the survey characteristics belonging to the same socio-economic domain but coming from different sources (e.g. surveys, registers, administrative data) without examining (a) if the differences in the comparisons of statistics are statistically important and (b) if the relationship of two or more survey characteristics coming from different sources coincide (e.g. if the relationship of employment and wages of the structural earning survey and the EU-SILC coincide).

a. Differences in comparisons for assessing the coherence

If the comparisons made for assessing the coherence of common characteristics are conducted using data coming from different sampling surveys, then it may be required to examine if these comparisons are statistically important through t-tests for continuous variables and χ^2 for non-continuous variables.

b. Coherence of the relationships of the survey characteristics

Apart from assessing the coherence of common characteristics from different surveys, the **c**oherence of the relationships of common survey characteristics may be assessed based on the analysis of data from different surveys belonging to the same socio-economic domain.

The analysis of the relationship of two or more continuous characteristics from different sources is based on the comparisons of the coefficient of correlations between common characteristics from different sources (e.g. calculation of the coefficient of variation between employment and wages separately for the structural earning survey and the EU-SILC and afterwards analysis on the comparisons of these coefficients of correlations). In the case that the analysis regards differences in the relationships of two or more non-continuous survey characteristics from different sources then the correspondence factor analysis is applied and for each source of information the relationship among survey characteristics is depicted on the factor levels, and afterwards the appearing differences on the factor levels through *hierarchical cluster analysis* are examined.

c. Case study for a complete analysis on whether the survey results comply with other similar survey results

In the annex 9, it is presented a case study regarding a complete analysis on whether the Greek EU-SILC results comply with similar survey results from the Greek Structural Earning Survey. This example is a part of a study conducted in the NSSG for studying the detail coherence of the results of the above surveys.

6.2 Comparing potential and actual disseminated presentation forms

In the context of disseminated statistical products, the users' data requirements are at the centre of the official statistics actions. The Statistical Services to cover the relevance of statistics more efficiently, they identify and analyze the users' needs and the users' demands on product quality and they also categorize the users in groups addressing customers in term of target groups. As customer interests are manifold, the range of statistical information offered has to be varied (Official Statistics of Federal Republic of Germany, 2005), too:

- For the public's general need for information, official statistics ensures free basic provision of information
- With its chargeable standard publications, official statistics meets the regular and recurring information demand of customers
- By offering custom-made product and services, the Official Statistical Services react to customers' individual needs, to cover, for example, either special analyzes or special tabulations or contract work carried out at the Statistical Services.

The diversity of statistical products through the above three types of offer is supplemented by diversity types of transmission. In order to disseminate the statistical results to the customers, the Statistical Services use a mix of print (paper) products, electronic offline products (diskettes, CD-ROMs etc), online products (websites), databases, telecommunication (info service, fax and email retrieval) and personal information (Official Statistics of Federal Republic of Germany, 2005).

Additionally, the statistical products should be fully documented with product descriptions (through metadata) and quality reports. As some end-users may not be in a position to asses the quality of the produced statistics through the quality indicators, it is required the Statistical Services to carry out all the necessary quality evaluations and make the results of these evaluations available to the users in a friendly way (in a format which is easy to use). Compiling and maintaining a public and easily accessible both statistical product and product description (through metadata and quality reports) is one of the main responsibilities of the statistics producer. The information contained in the product descriptions adds to the value of the statistics, because (Statistics Finland, 2002):

- Support and facilitate scientific research and appropriate interpretation
- Provide the users with a guarantee of the high quality of the produced statistics
- Increase the number of users
- Support and facilitate improvements in ongoing and repeated surveys

Taking into consideration all the types of accessible the products through the analysis of the results from the users' satisfaction surveys and all types of descriptions of the products through the users' demands and the scientific recommendations based on scientific studies, the potential disseminating presentation forms of results (paper tabular data and publications, offline and on line electronic tabular data and publications, micro-data in public use files, paper publications, electronic publications, press releases and press conferences, safe-centers for access to confidential micro-data etc) and potential documentations (metadata, methodological reports, quality reports, brief quality reports etc) are created. So, for improving the dissemination process of the surveys, it is required to be compiled two process quality indicators comparing:

Actual and potential disseminated presentations forms
Actual and potential types of statistics documentations

The above indicators are process quality indicators, because they assess the effect of the dissemination process quality on the quality of the produced statistics, which is measured by the indicator "accessibility and clarity of the results".

6.3. Survey Documentation System

As the quality of the produced statistics is strongly depended on the quality of the processes followed in the surveys, the improvement of the produced statistics quality is achieved only improving the quality of the processes. Therefore, for improving the product quality it is necessary to study, monitor and improve the quality of the underlying processes through measuring both the key process variables and the quality process indicators.

The process variables, the quality process variables, the quality reports and the metadata reports are the tools of the Survey Documentation System (SDS) developed to support the quality control activities for monitoring and improving the quality of both the survey processes and the statistical products. The main purposes of the system are (Official Statistics of Federal Republic of Germany, 2005):

- To allow the survey managers to monitor their production processes, to assess the quality of statistics over time
- To support the users to analyze the survey characteristics and to be able to know the factors which restrict the quality of the statistical products

- To provide the top management with qualitative and quantitative information on quality of statistics (processes and products) for decision making purposes

The efficiency of the developed Survey Documentation System depends on the number of processes, which are monitored, analyzed and documented. Thus, it is required an indicator to be compiled for evaluating the efficiency and the effective functionality of the Survey Documentation System. This indicator may be defined as:

Efficiency of SDS=Number of documented survey processes / Total number of survey processes

The quality improvement focuses on identifying, measuring and analysing process variables, because only a process quality approach allows early identification of problems that occur during a statistical operation. Thus, it is necessary the process and sub-processes of the surveys to record and measure and document. The indicator "*Efficiency of SDS*" informs on how effective is the Survey Documentation System regarding the evaluation and documentation of the survey processes

7. General remarks

i. The quality of the statistical products is dependent on the quality of the statistical processes. As a result, improving the quality of the processes it is achieved the quality of the statistical products. The processes are monitored through the process variables and the quality of the processes are studied and analysed through the quality process indicators, which measure the effect of the process quality on the quality of the products.

ii. The statistical products are required to be documented to monitor their production processes, to assess the quality of statistics over time, to support the users to analyze the survey characteristics and to provide the top management with qualitative and quantitative information on quality of statistics (processes and products) for decision making purposes. This documentation is conducted through the process variables, the quality process variables, the quality reports and the metadata reports, which are the tools of the Survey Documentation System (SDS) developed to support the quality control activities for monitoring and improving the quality of both the survey processes and the statistical products. The efficiency and the effective functionality of the Survey Documentation System is monitored through a suitable indicator (*Efficiency of SDS indicator).*

iii. Through case studies from surveys conducted by the NSSG, the relationship between the quality of the processes and the quality of results was proved.

iv. Applying *Correspondence Factor Analysis,* on data from the user satisfaction survey, it was proved that the relevance indicators used in the quality reports have a strong relationship among them. For example, the "type of users" has strong relationship with the required "categories of statistics", the "purposes for the use of official statistics", and the opinion on "accuracy", "timeliness" and "accessibility of statistics", etc.

v. The sample design, as a process, has strong relationship with the relevance, the costs, the personnel employed in the survey and the as well as with the quality of the produced statistics (accuracy, completeness, comparability, and coherence). The proposed quality process indicators are the design effect, the coefficient of correlation between stratification variable and the survey variables, the effect of the sampling scheme on the sampling errors. In the case that, the stratification variable is non-continuous variable, then the depicting on the factor level relationship between stratification variable and the main survey variables can be considered as a quality variable monitoring the degree of the successful stratification of the surveyed units. Additionally, as the whole sample design is dependent on the available budget and personnel, the effect of sample scheme on the sampling errors directly denotes how the availability of resources can affect the quality of the results. This relationship is possible to be appeared with the creation of a suitable formula connecting the needed funds for achieving the users' requirements with the available funds.

vi. The non-response and the coverage errors create bias, the effect of which is unknown. Thus, weighting adjustments are applied to reduce the bias that non-response and coverage errors may cause in survey estimates. The weights for non-response and population adjustments are random variables provoking increases in the sampling errors of the survey estimates. These increases tend to persist undiminished for most subclasses and for all statistics. The indicators quantifying the effect of random weighting on sampling errors and design effects may be used as quality process indicators assessing the data collection process and the weighting process, which aims at removing the bias due to non-response and coverage problems.

vii. The total weighting (random weighting and weighting due to disproportionate allocation, unequal selection probabilities) usually increases the variances. The percentages of these increases may be used as quality process indicators for assessing (a) the quality of the sample design and (b) the effect of random weighting on the results.

viii. The intra-class correlation in multistage surveys measures the effect of clustering on the sampling errors and it may be used as a design process indicator assessing if the homogeneity of the sampling elementary units in clusters influences the precision of the results. Additionally, this indicator evaluates the efficiency of the cluster sizes, and shows if the elementary final sampling units have been spread in an efficient number of sample clusters. The importance of the intra-class correlation as a quality process indicator is significant, because the intra-class correlation in combination with the average cluster sizes and the cluster effect links the sample design with the sampling error.

ix. The stratification effect is an important quality process indicator assessing the gains in precisions due to stratification of surveyed units.

x. The measurement errors have to do with the quality of the data collection process. The effect of the measurement errors on the sampling errors can be measured through a suitable process indicator. Through a case study based on data from a surveys conducted by the NSSG, the effect of measurement errors on the results was assessed (a) through the percentage of the extra existent variability, which increases the sampling error and (b) through the study for detecting possible bias on the results.

xi. The non-sampling errors (non-response errors, measurements errors, data process errors) provoke increase in the sampling errors of the survey characteristics. The percentage of increase between the actual and the expected sampling error can be considered as a process quality indicator assessing the quality of the operational process (data collection and data processing), because the extra sampling error is linked with the operational process and also has large effect on the accuracy of survey characteristics.

xii. In the case that the one stage stratified sampling scheme is applied in which the sampling were post-stratified for reducing the bias due to non-response and removing the bias due to frame errors (over-coverage and misclassifications), the initial optimum allocation in strata has been vanished provoking increase in variance of the estimated survey characteristics, which can be measured through a suitable formula. This increase of the variance may be considered as a quality process indicator assessing the effect of non-response, over-coverage and misclassifications on the precision of the survey characteristics.

xiii. The extra variance due to miscoding expressed as a fraction of the actual variance may be considered as a quality process indicator assessing the effect of the coding quality on the variance of the estimates. The square root of the above fraction expresses the effect of the quality of coding on the sampling errors.

xiv. The under coverage rates in the censuses are estimated through special sample coverage surveys. The effect (bias) of the under coverage on the census results may be considered as a quality process indicator assessing the quality of the process “sampling frame updating”.

8. References

Benzecri, J.P. (1992). Correspondence Analysis Handbook. *New York: Marcel Dekker, Inc.*

Cochran G.W. (1977). Sampling techniques. *John Wiley and Sons*. *New York.*

Deville J. C. and Sarndal C.E (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association, 87, 376-382*.

Eurostat (2002a). Standard Quality Report: May 2002. *Eurostat Working Group on Assessment of Quality in Statistics, Luxembourg, May 2*-3.

Eurostat (2002b). Definition. *Eurostat Working Group on Assessment of Quality in Statistics, Luxembourg, May 2-3*.

Eurostat (2002c). Glossary. *Eurostat Working Group on Assessment of Quality in Statistics, Luxembourg, May 2-3.*

Eurostat (2003a). How to make a Quality Report. *Methodological Document for the Working Group on Assessment of Quality in Statistics, Luxembourg, October 2-3.*

Eurostat (2003b). Quality Report. *Methodological Document for the Working Group on Assessment of Quality in Statistics, Luxembourg, October 2-3.*

Eurostat (2003c). How to make a Quality Report: September 2003. *Eurostat Working Group on Assessment of Quality in Statistics, Luxembourg, October 2-3*.

Gower, A. R. (1994): Questionnaire Design for Business Surveys. Survey *Methodology, 20, 125-136*.

Jones, N and Lewis (eds, with Aitken A, Horngren J, Zilhao M, J) (2003). Handbook on improving quality by analysis of process variables. *Final Report. Eurostat*

Kish, L. (1962). Studies of interviewer variance for attitudinal variances. JASA, 92-115

Kish, L. (1979). Samples and Censuses. International Statistical Review, 47, 99-109

Kish, L. (1992). Weighting for Unequal P_i. Journal of Official Statistics, 8, 183-200.

Kish, L. (1995). Methods for design effects. Journal of Official Statistics, 11, 55-77

Körner T., Bergdahl M., Elvers E., Lohauss P., Nimmergut A., Sæbø Hans Viggo, Szép K., Zilhao Maria João (2007). Handbook on Data Quality Assessment Methods and Tools (DatQAM), *European Commission grant 10300.2500.021-2005.708.*

Levy, P.S. and Lemeshow, S. (1991). Sampling of populations: methods and applications. *New York: John Willy and Sons*

Official Statistics Federal Republic of Germany (2005). Quality Standards in German Official Statistics. *Statistische Ämter des Bundes und der Länder, Wiesbaden*

Statistics Finland (2002): Quality Guidelines for Official Statistics. *Finland*.

Zarckovich, S.S. (1965): Sampling methods and Censuses. *FAO. Rome*

ANNEX 1: Quality evaluation on the Greek Labour Cost Statistics for the year 2004

Introduction

This documentation summarizes the collection of structural data on labour cost in Greece, and it presents the quality of the produced statistics, according to the quality evaluation criteria included in the Commission Regulation (EC) No 698/2006/5.5.2006 (Official Journal of the European Communities L 121, 6.5.2006, page 30).

1. Relevance

The relevance is the degree to which statistics meet current and potential users' needs. As the relevance is not an inherent characteristic of the statistical data, it can be measured only with the help of user satisfaction survey. The NSSG conducts a user's satisfaction survey twice a year for the selection of information on the relevance of the produced statistics. This survey is limited to the customers visiting the library of the NSSG.

The results of the 'user satisfaction survey; of the year 2004 conducted by the NSSG, showed that the labour market statistics (employment and labour cost) met the users' needs as stated by the following scale of satisfaction:

- Completely by 78,3%
- Partially by 16,7%.
- Not at all by 5%

The fields of the labour market statistics that did not meet the users' needs were the following:

The variable "wages and salaries in shares of company"
The regional statistics (NUTS II and NUTS III)

As the questionnaire is the indispensable tool for the data collection, the survey questionnaire of the labour cost survey was designed in accordance with the European Union requirements as well as the national ones. The Community obligations represent the 95% of the questionnaire, and the rest questions were formed, after consultation with the *main national core users*, which are:

- The academic and research community
- The banks and business
- The government agencies
- The national Parliament and
- The media.

The main purposes for which the users need the labour cost statistics are:

- Analysis of current developments for short-term decision-making
- Analysis of trends for longer-term decision- making
- Forecasting
- Research purposes

2. Accuracy

2.1 Sampling errors

2.1.1 Sample design

The labour costs survey covers the sections C-K and M-O of the NACE Rev.1 and the enterprises with average annual employment equal to or greater than 10 employees. The single stratified random sampling method was applied, employing the enterprise as a surveyed unit and obtaining statistical information from *each separate local unit* of the enterprises included in the sample.

The sampling frame used for the sample design was based on the Business Register (BR) of the NSSG. This BR is based on the VAT Register of the Ministry of Economy and Finance and it is updated through the statistical surveys of the NSSG and the register of the Social Insurance Foundation.

The statistical data for the public services of the sections M (Ministry of education and public schools) and N (Ministry of Health, public hospitals and public health centres) were collected from the Ministry of National Education and Religion Affairs, and the Ministry of Health and Social Solidarity, respectively.

2.1.2 Stratification

The enterprises with 10 or more employees included in the survey were stratified as following:

a. By geographical region – NUTS I,
b. By Division (two – digit NACE Rev.1 code) within each geographical region (Geography x Economic activity = Major stratum), and
c. By size class of the enterprise. In each of the major strata, the enterprises were stratified into H=7 size classes, according to their size, determined by their average annual number of employees in the business register, as follows.

Class 1	10-19	Employees
Class 2	20-49	"
Class 3	50-99	"
Class 4	100-249	"
Class 5	250-499	"
Class 6	500-999	"
Class 7	1000 or more	"

The enterprises that belong to the 7^{th} size class were surveyed exhaustively.

2.1.3 Sample size

The sample size is 4.313 enterprises (sampling fraction=20,1%) and the response rate =61,2%. The sample size of the enterprises was defined, so that the relevant standard error (co-efficient of variation CV) of the variables “annual labour costs” and “hourly labour costs” at 2-digit code level of economic activity at the whole country does not exceed 5%. The sampling units (enterprises) were distributed to size strata applying the method of optimal (Neyman) allocation.

The population (N) and the sample size (n), broken down by section and by size class of enterprises, are presented in the following table.

Table 1: The population (N) and sample size (n) by section and by size class

NACE Rev.1	Total		Size classes									
			10 –49		50-249		250 - 499		500 - 999		1000+	
	N	n	N	N	N	n	N	n	N	n	N	n
Total	21506	4313	17865	976	3080	2847	327	273	145	128	89	89
C	148	47	124	24	19	17	4	5	0	0	1	1
D	5955	1465	4.754	401	981	880	142	117	52	41	26	26
E	75	41	46	24	23	13	3	1	1	1	2	2
F	1112	221	861	22	222	178	15	9	12	10	2	2
G	5897	836	5.163	112	644	641	51	46	21	19	18	18
H	3302	390	2.947	37	322	318	21	19	9	13	3	3
I	1325	312	1.096	93	189	184	18	14	9	8	13	13
J	134	73	65	18	43	32	9	5	3	4	14	14
K	1689	441	1.323	103	314	293	35	29	13	12	4	4
M	191	58	152	25	31	28	6	3	2	2	0	0
N	367	144	211	19	124	99	16	12	12	10	4	4
O	1310	285	1.123	98	168	164	7	13	10	8	2	2

Selection of the sampling units (enterprises)

In each of the final strata (let h), a sample of n_h enterprises was selected. The enterprises to be surveyed were selected from the total of the N_h enterprises with equal probabilities and by applying systematic sampling. The sampling units (enterprises) were selected from the sample frame based on data from the Business Register of the NSSG.

2.1.5 Survey characteristics estimation

a. Symbols

Defining with index i the selection order of an enterprise from the sampling frame in the stratum h and symbolizing with the y one of the survey characteristics, we can define the following:

y_{hi} : the value of the survey characteristic y of the enterprise of order i in the stratum h

$\boldsymbol{Y_h}$: the sum of the values of the characteristic $\boldsymbol{y}$ for all enterprises falling into the survey and belonging to the stratum $\boldsymbol{h}$

$\boldsymbol{Y}$: the sum of the values of the characteristic $\boldsymbol{y}$ of all enterprises under survey belonging to one economic activity with two digit code . That is:

$$Y = \sum_h Y_h \qquad (2.1)$$

Estimation process

The estimations of the magnitudes $\boldsymbol{Y_h}$ and $\boldsymbol{Y}$ come from the following relations:

$$\widehat{Y}_h = \frac{N_h}{n_h} \cdot \sum_{i=1}^{n_h} y_{hi} \qquad (2.2)$$

$$\widehat{Y} = \sum_h \widehat{Y}_h \qquad (2.3)$$

Generally, in order the estimations of the survey characteristics to be produced at any level, we add up the estimations of the (final) strata, which form the level under survey. The estimates of totals are produced using the Horvitz-Thompson estimator of the relation (2.2), which is *unbiased*.

There are quantities being produced through the ratio of two variables (as the annual labour costs per employee or the hourly labour costs). The estimations are produced using the ratio estimator, which is usually *slightly biased*.

We assume that the population parameter to be estimated is the ratio:

$$R = \frac{\sum_{i=1}^{N} y_i}{\sum_{i=1}^{N} x_i} = \frac{Y}{X} = \frac{\overline{Y}}{\overline{X}}$$

where y_i and x_i are the values for the each unit of order i in the population of size N (e.g. the variable y is the total labour cost and the variable x is the number of hours actually worked). If the stratified random sampling scheme is applied, then the combined estimation of R is:

$$\widehat{R} = \frac{\sum_h \frac{N_h}{n_h} \sum_{i=1}^{n_h} y_{hi}}{\sum_h \frac{N_h}{n_h} \sum_{i=1}^{n_h} x_{hi}} = \frac{\widehat{Y}}{\widehat{X}} \qquad (2.4)$$

The ratio estimator $\widehat{R}$ is biased. In general, the ratio estimation has a bias of order $1/n$. Since the standard error (s.e.) of the estimation $\widehat{R}$ is of order $1/\sqrt{n}$, the quantity $Bias/s.e$ is also of order $1/\sqrt{n}$ and it becomes negligible, as the sample size n becomes large. In practice, this technical bias is usually unimportant in samples of moderate and large size.
As the technical bias of $\widehat{R}$ occurs because the denominators x of $R = y/x$ are random variables, one can use the $CV(\widehat{X}) < 20\%$ (CV: Coefficient of variation of $\widehat{X}$), as an indicator examining if the effect of bias on the accuracy of $\widehat{R}$ to be neglected. Thus, the $CV(\widehat{X})$ serves as a critical control on the validity of combined ratio estimations and it is a useful and safe-check on the bias of ratio statistics.

c. Variance estimation

The estimations of the variances of $\widehat{Y}_h$ and $\widehat{Y}$ come from the following relations:

$$Var(\widehat{Y}_h) = \frac{N_h(N_h - n_h)}{n_h(n_h - 1)} \cdot \left[\sum_{i=1}^{n_h} y_{hi}^2 - \frac{\left(\sum_{i=1}^{nh} y_{hi} \right)^2}{n_h} \right] \quad (2.5)$$

$$Var(\widehat{Y}) = \sum_h V(\widehat{Y}_h) \quad (2.6)$$

The coefficient of variation (%) of the quantity $\widehat{Y}$ is given by the following relation:

$$CV(\widehat{Y}) = \frac{\sqrt{Var(\widehat{Y})}}{\widehat{Y}} \cdot 100 \quad (2.7)$$

The estimation of the variance of $\widehat{R}$ is calculated from the following relation:

$$Var(\widehat{R}) = \frac{1}{\widehat{X}^2} \cdot \sum_h \frac{N_h \cdot (N_h - n_h)}{n_h} \cdot \left[S_{yh}^2 + R^2 \cdot S_{xh}^2 - 2 \cdot R \cdot Cov(y_h, x_h) \right] \quad (2.8)$$

where: $S_{yh}^2 = \frac{1}{n_h - 1} \cdot \left[\sum_{i=1}^{n_h} y_{hi}^2 - \frac{\left(\sum_{i=1}^{nh} y_{hi} \right)^2}{n_h} \right]$, $S_{xh}^2 = \frac{1}{n_h - 1} \cdot \left[\sum_{i=1}^{n_h} x_{hi}^2 - \frac{\left(\sum_{i=1}^{nh} x_{hi} \right)^2}{n_h} \right]$,

$$Cov(y_h, x_h) = \frac{1}{n_h - 1} \cdot \left[\sum_{i=1}^{n_h} y_{hi} \cdot x_{hi} - \frac{\left(\sum_{i=1}^{n_h} y_{hi}\right) \cdot \left(\sum_{i=1}^{n_h} x_{hi}\right)}{n_h} \right]$$

The coefficient of variation of $\hat{R}$ is calculated from the following relation:

$$CV(\hat{R}) = \frac{\sqrt{Var(\hat{R})}}{\hat{R}} = \frac{1}{\hat{Y}} \cdot \sqrt{\sum_h \frac{N_h \cdot (N_h - n_h)}{n_h} \cdot \left[S_{yh}^2 + R^2 \cdot S_{xh}^2 - 2 \cdot R \cdot Cov(y_h, x_h)\right]} \quad (2.9)$$

The coefficient of variation of the variables "annual labour costs" and "hourly labour costs" are shown in the following tables according to the structure of tables A (national data), B (regional data) and C (national data by size class of enterprise).

Table A: Coefficient of variations (%) by economic activity (Sections)

NACE Rev. 1	Annual labour costs	Hourly labour costs
Total	0,8	0,7
C	4,5	3,8
D	1,4	1,0
E	0,1	0,1
F	6,1	4,1
G	3,3	3,2
H	7,5	5,0
I	1,5	1,2
J	1,0	1,0
K	4,8	4,5
M	0,2	0,2
N	0,6	0,4
O	4,9	4,2

Table B: Coefficient of variations (%) by NUTS I

NUTS1	Annual labour costs	Hourly labour costs
Total	0,8	0,7
North Greece	1,7	1,4
Central Greece	1,9	1,4
Attica	0,9	0,8
Islands of Aegean and Crete	4,4	2,8

Table C: Coefficient of variations (%) by size class of enterprises

Size classes	Annual labour costs	Hourly labour costs
Total	0,8	0,7

Size classes	Annual labour costs	Hourly labour costs
10-49	3,3	2,4
50-249	1,2	1,0
250-499	2,1	1,7
500-999	2,7	1,9
1.000+	0,0	0,0

In the section with code E, the coefficient of variations of the variables "annual labour costs" and "hourly labour costs" are equal to 0,1%, because 90% of the total statistical information was collected from two large enterprises (average annual employment higher than 1000 persons) belonging to the census (take –all) stratum.

In the sections M and N, the coefficient of variations of the variables "annual labour costs" and "hourly labour costs" are less than 1%, because the statistical information for the public services was collected from administrative sources (Ministry of National Education and Religion Affairs, Ministry of Health and Social Solidarity).
In the sections with codes D, I and J, the coefficient of variations of the variables "annual labour costs" and "hourly labour costs" are less than 1,6%.

In the sections with codes C, G, K and O, the coefficient of variations of the variables "annual labour costs" and "hourly labour costs" are ranged between 3% and 5%.

In the sections with codes F and H, the coefficient of variation *only* of the variable "annual labour costs" is higher than 5%, because in these sections strong seasonality is appeared. As a result, two different types of enterprises belong to the same size classes, as follows:

- Enterprises operating all the year
- Enterprises operating only a time of period less than one year (approximately, half a year)

The enterprises of the first type have annual labour costs higher than the enterprises belong to the second type, and as a result, in the same size stratum the annual labour costs of the enterprises are not homogeneous due to the different types of enterprises. Thus, although in the same size stratum, internally homogenous enterprises exist, according to their number of employees, however internally heterogeneous enterprises are appeared, according to their values of the "total annual labour costs". This increases the variance of the total annual labour costs of the enterprises, reducing the gain in the precision from the stratification that was introduced initially in the sample selection.

Concerning, the problem of the internal heterogeneity in the size strata was not appeared in the variable "hourly labour costs", because the variables "annual labour costs" and "hours actually worked" are strong correlated. As a result, there was not any high variability in the hourly labour costs due to the different types of enterprises.

2.2 Non-sampling errors

2.2.1 Coverage Errors

There were problems of over-coverage, under-coverage and miss-classification.

The *over-coverage* problems mainly have to do with enterprises that were included in the business register, they were selected in the sample, but they were not actually existed at the time of the survey (closed enterprises). These enterprises actually reduced the initial sample size of primary units, n_h. The decrease of the number of sampling units from n_h to m_h in each stratum inflates the variance of the estimated statistics. In this case the estimator is unbiased under the condition that the death rate of enterprises is equal to their birth rate.

The *under-coverage* refers to units missing from the sampling frame. The probability of selection of each missing unit of order i is equal to zero ($p_i = 0$) and thus, the extrapolation factor w_i of the missing unit cannot be defined ($1/p_i = 1/0$). As a result, the under-coverage problem underestimates the produced statistics. Corrections and weighting for *non-coverage* is difficult, because the under-coverage rates cannot be obtained from the sample itself, but only from external sources.

Due to *miss-classification* problems of the register, some sampling units changed design strata after data collection. These units were allocated to the new strata, *retaining* their initial probabilities of selection. This event changes the initial element variance, destroys the initial allocation of the enterprises of the sample and as a result inflates the variance of the estimations. Consequently, the co-efficient of variation of the produced statistics is *higher* than the co-efficient of variation based on the initial sample design.

2.2.2 Measurement and processing errors.

The data collection method used was face-to-face interview completing paper questionnaires. The collection method applied ensured the high quality of the information gathered, since the interviewers assisted the respondents, and carefully checked the filled in questionnaires, before leaving the enterprise.

The interviewers participated in the survey were experienced permanent staff of the National Statistical Service of Greece (NSSG), as well as private collaborators. Before launching of the survey, the interviewers attended a one-day training seminar. The scope of the seminar was to enable the interviewers to: a) fully understand the definitions of the survey characteristics in order to avoid the respondent bias, (b) correctly fill in the questionnaire, and (c) efficiently check for errors by applying logical controls.

The structure and the size of the questionnaire were designed to be user-friendly for the interviewers and the questions were formulated in a clear and simple language, using appropriate vocabulary. Additionally, documents containing useful instructions were compiled, analyzing all the questions of the questionnaire. This activity targeted at collecting fully completed questionnaires, with no missing variables.

The support and supervision of the data collection and the data processing were decentralized in the regional offices of our Service. In regional offices were carried out coding, checking for the detection of measurement errors, logical controls and comparisons with other sources of statistical information.

After performing all final controls for discovering non-sampling errors, the database was ready for the extrapolation weighting process and the plausibility checks after tabulation. These checks included comparisons of data with relevant data of previous years and other surveys.

2.2.3 Non-response errors

The following table shows the unit response rates (%), total and broken down by section and size classes of enterprises.

Table 2: Unit response rates (%) by section and size classes

NACE Rev.1	Total	Size classes				
		10-49	50-249	250-499	500-999	1000+
	%	%	%	%	%	%
Total	61,0	74,2	52,6	80,1	84,4	100
C	68,1	66,7	64,7	80,0	100	100
D	67,3	76,6	59,8	80,3	100	100
E	100	100	100	100	100	100
F	47,1	100	33,1	100,0	80,0	100
G	66,5	100	58,0	84,8	94,7	100
H	61,3	100	54,4	84,2	30,8	100
I	63,1	72,0	51,6	92,9	100	100
J	54,8	61,1	53,1	40,0	50,0	100
K	41,5	47,6	31,4	82,8	100	100
M	53,4	32,0	42,9	100	100	100
N	52,8	63,2	46,5	58,3	60,0	100
O	49,8	55,1	45,1	38,5	75,0	100

In the census (take-all strata), in which all population units are included in the sample, the unit response rate is equal to 100%. In the sampling strata, in which only a part of population is included in the sample, the re-weighting method was applied for statistical adjustments of the produced statistics.

The re-weighting method amends suitably the extrapolation factors taking into account the response rates in all final strata. This method compensates for non-responses, and reduces the absolute bias in the estimation of $\hat{Y}$. If $\overline{Y}_{rh} = \overline{Y}_{mh}$ (where $\overline{Y}_{rh}$ and $\overline{Y}_{mh}$ are the means for respondents and non-respondents in stratum h for the variable y), as it occurs in expectation when the non-respondents are missing at random, then in stratum h the bias of non-response is equal to zero. Generally, the total bias due to the non-response is approximately equal to zero, if either the response rates or the respondent means do not vary between strata.

Any imputation method was not applied for the item non-response, as the item non-response was not appeared in the enterprises included in the sample.

3. Timeliness and punctuality

3.1 Punctuality

The multiple operations of the Labour Cost Survey were carried out in four phases, as detailed below:

Phase 1: Organization and preparation of the survey

The first phase was carried out from 1st January 2005 to 31st March 2005, and it comprised the organization activities and the preparatory work for the survey. More precise the following actions were carried out:

- Issuing of a special decision of the Ministry of Finance and Economy, which sets out the time schedule, the organization and the cost of the survey
- Sample design
- Design and printing of the questionnaire (paper and pencil),
- Printing of the manual with the instructions for the data collection
- Creation of software program for the data entry and automatic controls
- Programming for the creation of database files
- Selection and appointment of the interviewers for the conduct of the survey
- Training seminar of the interviewers for the effective data collection
- Delivery to the regions (prefectures) of the questionnaires and the questionnaires' instructions
- Transmission of information letters to the enterprises belonging to the sample

Phase 2: Data collection

The second phase was carried out from 1st June 2005 to 31st December 2005. In this phase the following operations were carried out:

- Distribution to the interviewers of the questionnaires and the lists with the sample units and other necessary documents
- Collection of the statistical data
- Monitoring and supervision of the operation from the beginning to the end by the supervisors (heads of the regional offices and the head of the competent department of the Central Office)
- Delivery by the interviewers of the questionnaires to their supervisors

Phase 3: Data processing

The third phase was carried out from 1st January 30th June 2006

- The following operations were carried out:
- Checking for the completeness of the questionnaire
- Logical and consistency controls of the data
- Coding
- Data entry and automatic data editing
- Creation of a database with the survey data

- Qualitative controls of the data in the database
- Calculation of the extrapolation factors
- Estimation of the survey characteristics
- Tabulation of the estimated statistics for qualitative analysis
-

Phase 4: Evaluation of the results-Publication and Dissemination

This phase was conducted in July 2006 and the following operations were carried out:

- Qualitative analysis and documentation of the results
- Production of national tables with the final results
- Transmission to Eurostat tabular data through the appropriate technical format for the transmission of the results

The *punctuality* of data transmission to Eurostat is evaluated according to delays stated in Council Regulation (EC) No 530/1999, in which the results are forwarded to Eurostat within a period of 18 months from the end of the reference year.

The data elaboration had to be completed, not later than 15^{th} June 2006, in order the produced statistics to have been transmitted to Eurostat not later than the end of June. The qualitative controls of the data in the database were out of schedule and as a result the whole survey was extended by one month.

3.2 Timeliness

The length of time between the release of data and the reference period of data is equal to 19 months.

4. Accessibility and clarity

4.1 Accessibility

a. There is a publication in Greek containing:

- A short description of the methodology applied for the data collection
- Tables with the results of the survey

b. Tables with the results of the survey are available in the website of the NSSG.

In the case that the users need more detailed information, they can ask for it in the NSSG and special tables can be produced on request. Moreover, in some cases (i.e. for research purposes) anonymised individual data can also be provided to the users. The format of the anonymised data is so that the confidentiality to be protected and the respondents not to be revealed.

5. Comparability

5.1 Geographical comparability

The definition of the statistical units, the reference population, the classifications and the definitions of the observed variables in the transferred results to Eurostat were determined according to the Council Regulation (EC) No 530/1999/9.3.1999 (Official Journal of the European Communities L 63, 12.3.1999, page 6). Thus, the produced statistics are comparable between the member-states of the European Union.

5.2 Comparability over time

The labour cost surveys with reference periods the calendar years 1969, 1973, 1974, 1976, 1977 and 1978 produced statistics only for section "Manufacturing" and the reference population was limited to establishments with 10 or more employees.

The surveys of the years 1981, 1988 and 1992 apart from manufacturing, were extended to the sections "Mining and quarrying" and "Electricity, gas and water supply". Moreover, the surveys were designed and conducted, in order the produced statistics to be harmonized and comparable with the corresponding data of the rest member-states.

The survey of the year 1996 widened the economic activity coverage to include the sections "Construction", "Wholesale and retail trade; repair of motor vehicles, motorcycles and personal household goods", "Hotel and restaurants", "Transport, storage and communication" and "Financial intermediation". However, only enterprises with 10 or more employees participated in the survey. Furthermore, the classification applied for economic activities corresponded to NACE Rev.1.

The survey of the year 2000 widened the economic activity coverage to include the section "Real estate, renting and business activities". However, only enterprises with 10 or more employees participated in the survey.

The survey of the year 2004 widened the economic activity coverage to include the sections with NACE Rev.1 codes M, N and O. However, only enterprises with 10 or more employees participated in the survey.

6. Coherence

a. Coherence with statistics from the labour force survey

The number of hours actually worked per employee of the labour cost survey (LCS) and the labour force survey (LFS) are appeared in the following table.

Table 3: Hours actually worked per employee by section and survey

NACE Rev.1	LCS	LFS	Difference (%)
Total	1.583	1.957	-19,1
C	1.722	2.095	-17,8
D	1.708	2.067	-17,4
E	2.096	2.052	2,1
F	1.655	2.081	-20,5

NACE Rev.1	LCS	LFS	Difference (%)
G	1.756	2.093	-16,1
H	1.466	2.202	-33,4
I	1.898	2.197	-13,6
J	1.589	1.963	-19,1
K	1.630	2.002	-18,6
M	929	1.136	-18,3
N	1.881	1.937	-2,9
O	1.665	1.950	-14,6

Some significant differences in both surveys are appeared due to the fact that, the micro-enterprises (1-9 persons employed) were not included in the surveyed population of the labour cost survey. The micro-enterprises in Greece represent a significant share of the production structure and have a relatively high impact on the overall employment.

b. Coherence with structural business statistics

The variable "wages and salaries" per employee of the labour cost survey (LCS) and the Structural Business Survey (SBS) are appeared in the following table.

Table 4: Wages and salaries per employee by section and survey

NACE Rev.1	SBS	LCS	Difference (%)
Total	17.030	18.056	6,0
C	30528	21.315	-30,2
D	17981	17.305	-3,8
E	31963	42.482	32,9
F	12894	15.306	18,7
G	13419	15.198	13,3
H	11359	12.501	10,0
I	21569	26.022	20,6
J	30.349	29.261	-3,6
K	22847	17.624	-22,9

The deficiencies in coherence between SBS and LCS are due to the following reasons:

- In the SBS, the surveyed unit is the enterprise, whilst in the LCS the local unit. So, in the SBS, one enterprise may contain local units belonging to different economic activities and to different geographical regions
- In the SBS, the values of D11 do not contain the values of variables "payments to employees saving schemes" and "wages and salaries in kind".

c. Coherence with Labour Cost Index

Table 5: Hourly Labour Costs of the LCS for the years 2000 and 2004 by economic activity

NACE Rev.1	YEARS		Difference (%)
	2000	2004	
Total	10,8	15,3	41,1
C	12,7	16,9	33,4
D	10,3	14,0	36,8
E	17,0	28,0	64,5
F	7,6	13,1	72,3
G	8,2	12,0	46,5
H	7,7	11,3	45,7
I	13,8	19,7	43,4
J	17,8	26,9	51,2
K	11,6	14,9	28,4

Table 6: Average annual LCI for the years 2000 and 2004 by economic activity

NACE Rev.1	YEARS		Difference (%)
	2000	2004	
Total	100,0	127,9	27,9
C	100,0	120,7	20,7
D	100,0	129,5	29,5
E	100,0	125,1	25,1
F	100,0	127,9	27,9
G	100,0	126,5	26,5
H	100,0	127,8	27,8
I	100,0	125,4	25,4
J	100,0	129,1	29,1
K	100,0	131,5	31,5

Table 7: Annual growths of hourly labour costs of the LCI and the LCS by economic activity

NACE Rev.1	Growth Rates		(LCS/LCI)-1
	LCI	LCS	%
Total	27,9	41,1	47,2
C	20,7	33,4	61,6
D	29,5	36,8	25,0
E	25,1	64,5	157,0
F	27,9	72,3	159,6
G	26,5	46,5	75,6
H	27,8	45,7	64,6
I	25,4	43,4	70,6
J	29,1	51,2	75,8
K	31,5	28,4	-9,7

The reason for differences between the two sets of statistics (growth rates from LCI and LCS) is the different time schemes of the production of statistics for a given year. The annual statistics are collected after the year, whilst the short-term statistics are collected during the year. The population being surveyed changes during the year (births and deaths, mergers and break-ups etc). Such changes are better known when producing the annual than the short-term statistics. Hence, even if the target population is the same, the frames may be different for the two surveys.

6.3 Coherence with national accounts (NA)

In the following table the variable "compensation of employees" expressed per employee are appeared by section and source.

Table 8: "Compensation of employees" per employee by section and source

NACE Rev.1	NA	LCS	Difference (%)
Total	17.547	18.056	2,9
C	24.857	21.315	-14,2
D	15.818	17.305	9,4
E	23.195	42.482	83,2
F	14.047	15.306	9,0
G	14.739	15.198	3,1
H	15083	12.501	-17,1
I	18.811	26.022	38,3
J	25.462	29.261	14,9
K	13.899	17.624	26,8
M	22.027	14.699	-33,3
N	20.512	19.880	-3,1
O	17.146	16.846	-1,7

The National Accounts build first on the short-term statistics and later on annual statistics, when the annual statistics are available. The values coming from NA are provisional data and they have been produced using the changes over time from the values of the Labour Cost Index. As a result, some significant differences are appeared between the compensations of employees between the NA and the LCS. Additionally, in national accounts the variables D1 and D11 do not contain the values of variables "payments to employees saving schemes" and "wages and salaries in kind".

References

Bellhouse (1988). Systematic sampling. In Handbook of Statistics, Vol. 6, (Eds. P.R. Krishnaiah and C.R. Rao). Amsterdam: Elsevier Science, 125-145

Cochran, W.G. (1977). Sampling Techniques, New York: John Wiley and Sons

Dalenious T., and Hodges, J.L (1959). Minimum variance stratification. JASA, 54,88-101

Deming, W.E. (1953). On a probability mechanism and the bias of non-response. JASA, 48, 743-772

Evans, W.D. (1951). On stratification and optimum allocation. JASA, 46, 95-104

Hansen, M.H., Hurwitz, W.N., Madow, W.G. (1953). Sample Survey Methods and Theory. Vol. I, New York: John Wiley and Sons

Hess, I, Sethi, V.K., and Balakrishnan,T.R (1966). Stratification: A practical investigation. JASA, 61, 74-90

Holt, D., and Elliot, D. (1991). Methods of weighting for unit non-response. The Statistician, 40, 333-342

Kalton, G. and Kasprzyk, D. (1986). The Treatment of Missing Survey Data. Survey Methodology, 12, 1-16.

Kalton, G. and Flores – Gervantes, I. (2003). Weighting Methods. Journal of Official Statistics, 19, 81-97.

Kish, L., (1965). Survey Sampling, New York: John Wiley and Sons

Kish, L., and Frankel, M.R. (1974). Inference from complex samples. Journal of the Royal Statistical Society, A, 139,80-95

Kish, L., and Anderson, D.W. (1978). Multivariate and multipurpose stratification. JASA, 73, 24-34

Kish, L., (1987). Statistical Design and Research, New York: John Wiley and Sons

Kish, L., (1988). Multipurpose Sample Designs. Survey Methodology, 14, 19-32

Kish, L., (1992). Weighting for Unequal Pj. Journal of Official Statistics, 8, 183-200

Little, R.J.A. (1982). Models for non-response in sample surveys. JASA, 77, 237-250

Little, R.J.A. (1986). Survey non-response adjustments for estimates of means. International Statistical Review, 54, 139-157

Madow, L. H (1946). Systematic sampling and its relation to other sampling designs. JASA, 41, 207-214

Murthy, M.N. and Rao, J.T.(1988). Systematic sampling with illustrative examples. In Handbook of Statistics, Vol. 6, (Eds. P.R. Krishnaiah and C.R. Rao). Amsterdam: Elsevier Science, 147-185

Nikolaidis B. I.(2000). Variance Estimation in Business Statistics. Invited paper for the Eurostat working group on assessment of quality in statistics (4^{th} and 5^{th} April 2000). Doc. Eurostat/A4/Quality/00/General/Variance Estimation/NSSG

Royall, R.M and Herson, H.J. (1973). Robust Estimation in Finite Populations I. JASA, 68, 880-889

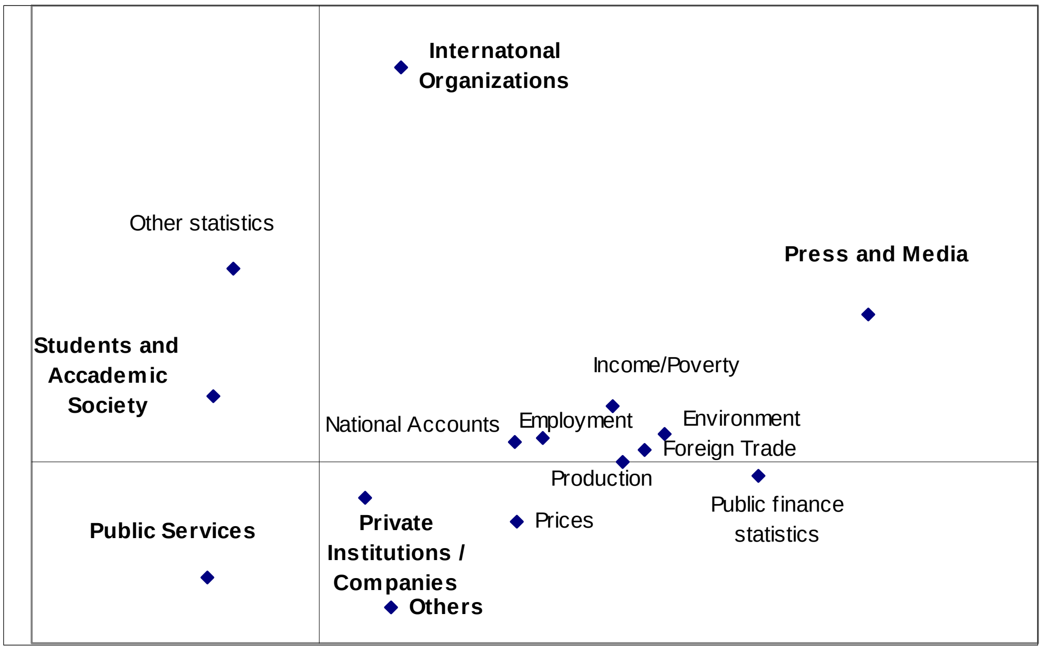
Royall, R.M and Herson, H.J. (1973). Robust Estimation in Finite Populations II: Stratification on a Size Variable. JASA, 68, 890-893

Sarndal, Swensson, and Wretman (1992). Model Assisted Survey Sampling. New York: Springer-Verlag

ANNEX 2: Case study on the relationship between relevance indicators

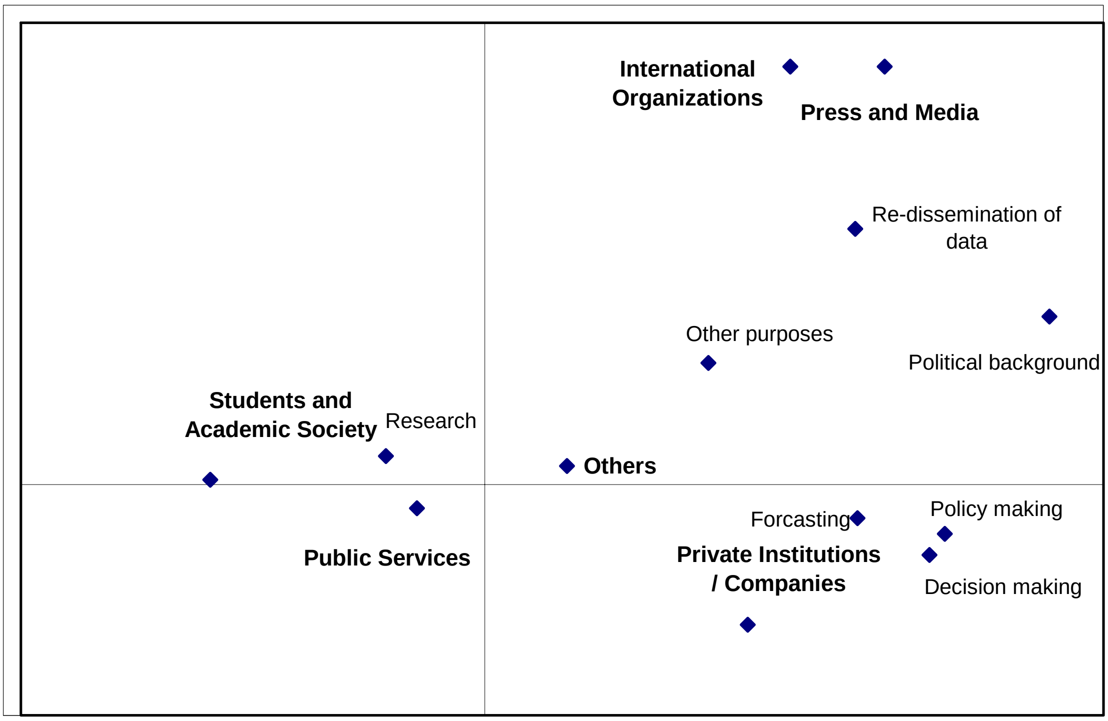
The relationship between relevance indicators is possible to be depicted on factor level using *Correspondence Factor Analysis* as in the following examples. The used data were collected from the users' satisfaction survey conducted by the NSSG in the framework of the "peer review" for the implementation of the European Statistics Code of Practice.

Example 1: Factor level of the variables "type of user" and category of "official statistics"



At the factor level, the preferences of each type of users separately for each category of official statistics can be depicted. Specifically, "private institutions/ companies" request mainly statistical data on "production", and "prices". In addition, "press and media" show interest in "income and poverty statistics", "employment" and "environment". This manner is totally expected from media, as these kinds of statistics are very useful on social and political analysis. As for "international organizations", it seems that their attention splits to plenty of categories of official statistics. However, as the sample from the International Organizations was low strong correlation between International Organizations and categories of statistics is not appeared. We should probably mention that "students and academic society" show interest in "other statistics", which are stated to be vital statistics, population censuses, culture/ entertainment/ sports statistics, education statistics, justice and public order statistics and more.

Example 2: Factor level of the variables “type of user” and “purposes for the use of official statistics”



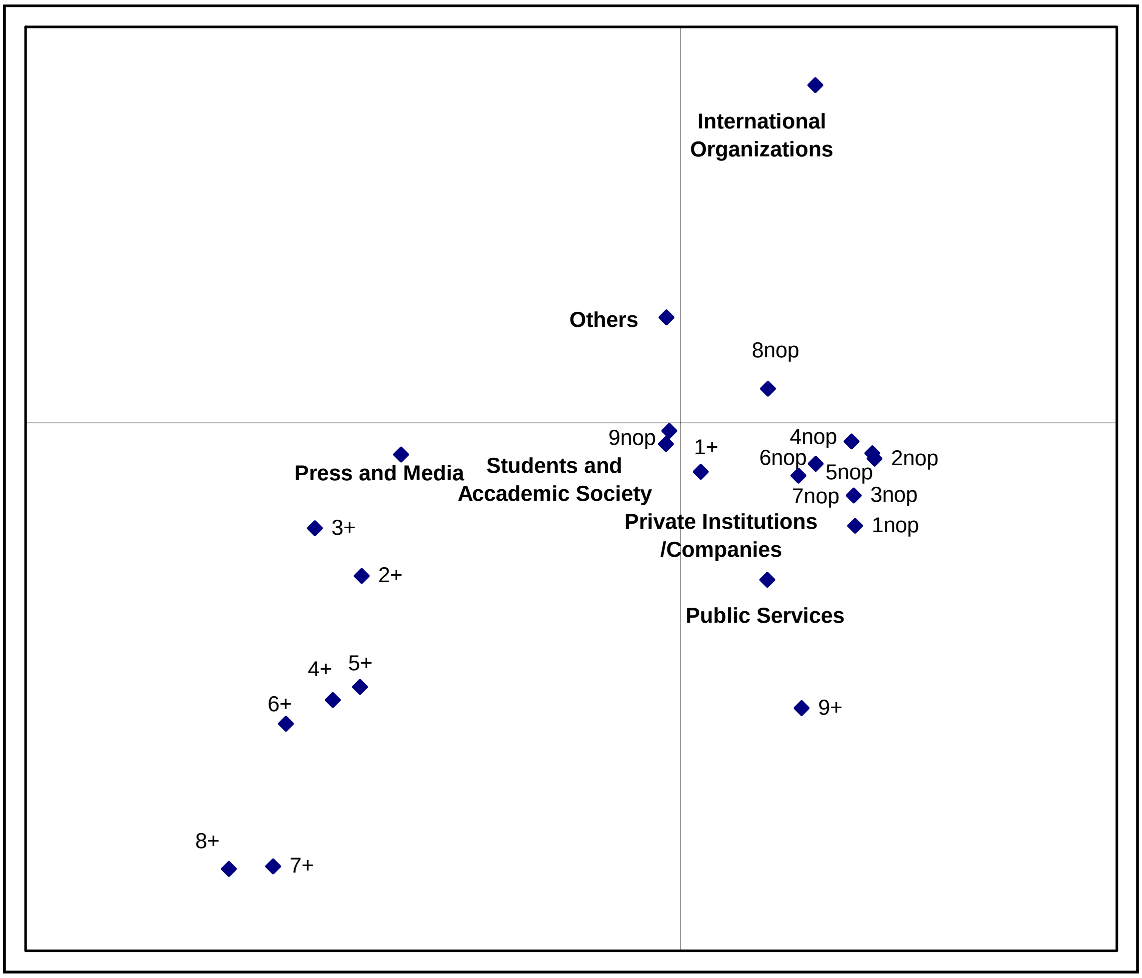
The types of users “Students and Academic Society”, “Public Services” and “Others” mainly utilize the official statistics mainly for “research purposes”.
On the contrary, “Private Institutions / Companies” use official statistics mainly for conducting “Forecasting”, “Policy and Decision Making”. In this type of user, “research purposes” take the second place of preference.
Besides, the International Organizations” use official statistics for “re-dissemination” “political background” and “other purposes”.
The media also use official statistics for “re-disseminating”, “political background” and “other purposes”. In other words, international organizations and media behave similarly, as their points at the factor level are very close to each other.

ANNEX 3: Case study on the relationship between indicators of users' demands of product quality

The relationship between indicators of users' demands of product quality is possible to be depicted on factor level using *Correspondence Factor Analysis* as in the following examples in which the data were collected from the users' satisfaction survey having been conducted by the NSSG in the framework of the "peer review" for the implementation of the European Statistics Code of Practice.

Example 1: Opinion on Accuracy (In your opinion, is the underlying methodology of official statistics sound and appropriate?)

Factor level of the variables "type of user" and the answers "Yes" and "No opinion"



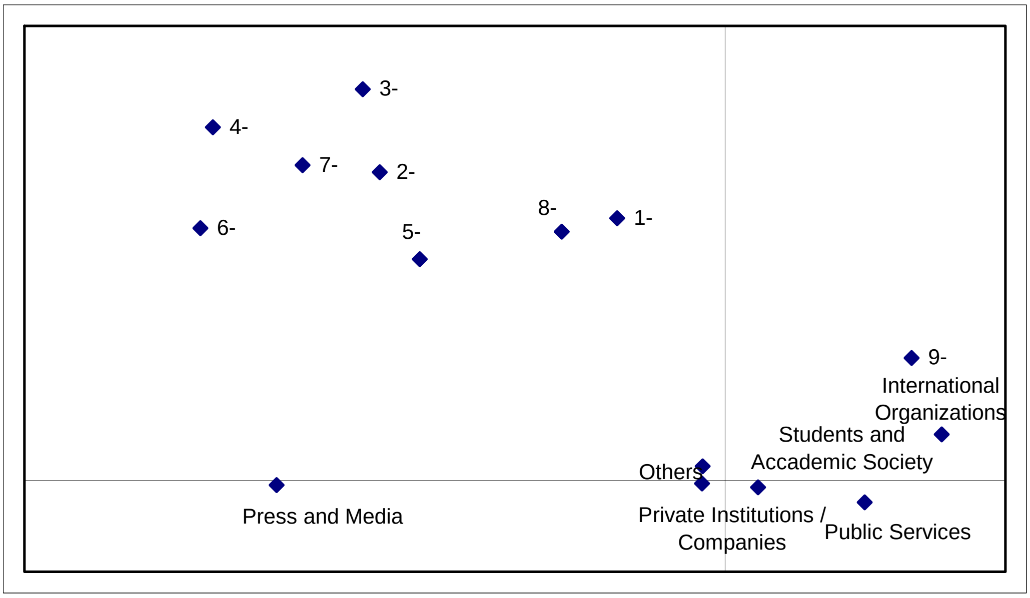
Where

1+	National Accounts	Yes
2+	Prices	Yes
3+	Public Finance Statistics	Yes
4+	Production	Yes
5+	Employment Statistics	Yes
6+	Foreign Trade Statistics	Yes
7+	Income & Poverty Statistics	Yes
8+	Environment Statistics	Yes
9+	Other Statistics	Yes

and

1nop	National Accounts	No opinion
2nop	Prices	No opinion
3nop	Public Finance Statistics	No opinion
4nop	Production	No opinion
5nop	Employment Statistics	No opinion
6nop	Foreign Trade Statistics	No opinion
7nop	Income & Poverty Statistics	No opinion
8nop	Environment Statistics	No opinion
9nop	Other Statistics	No opinion

Factor level of the variables "type of user" and the answer "No"



Where

1-	National Accounts	No
2-	Prices	No
3-	Public Finance Statistics	No
4-	Production	No
5-	Employment Statistics	No
6-	Foreign Trade Statistics	No
7-	Income & Poverty Statistics	No
8-	Environment Statistics	No
9-	Other Statistics	No

As the points of "no opinions" are close to almost users, important percentage of all users (average 70%) has no an opinion about the methodology of official statistics. This probably means that most users are not familiarized with the statistical science in a satisfactory degree. Thus, the methodology of official statistics of Greece is well appraised from the rest of the users (around 24% on average) and only about 4% of them have a negative opinion. So this situation stands for all the kinds of official statistics.

In detail, the academic society follows the general pattern described above, while this is quite interesting as it is a type of user we would expect not to go with the current at least in the terms of "no opinion". We can also say that the methodologies of "employment statistics" and "prices" are more highly graded than the other kinds of official statistics and concurrently "public finance statistics" gather the less negative opinions.

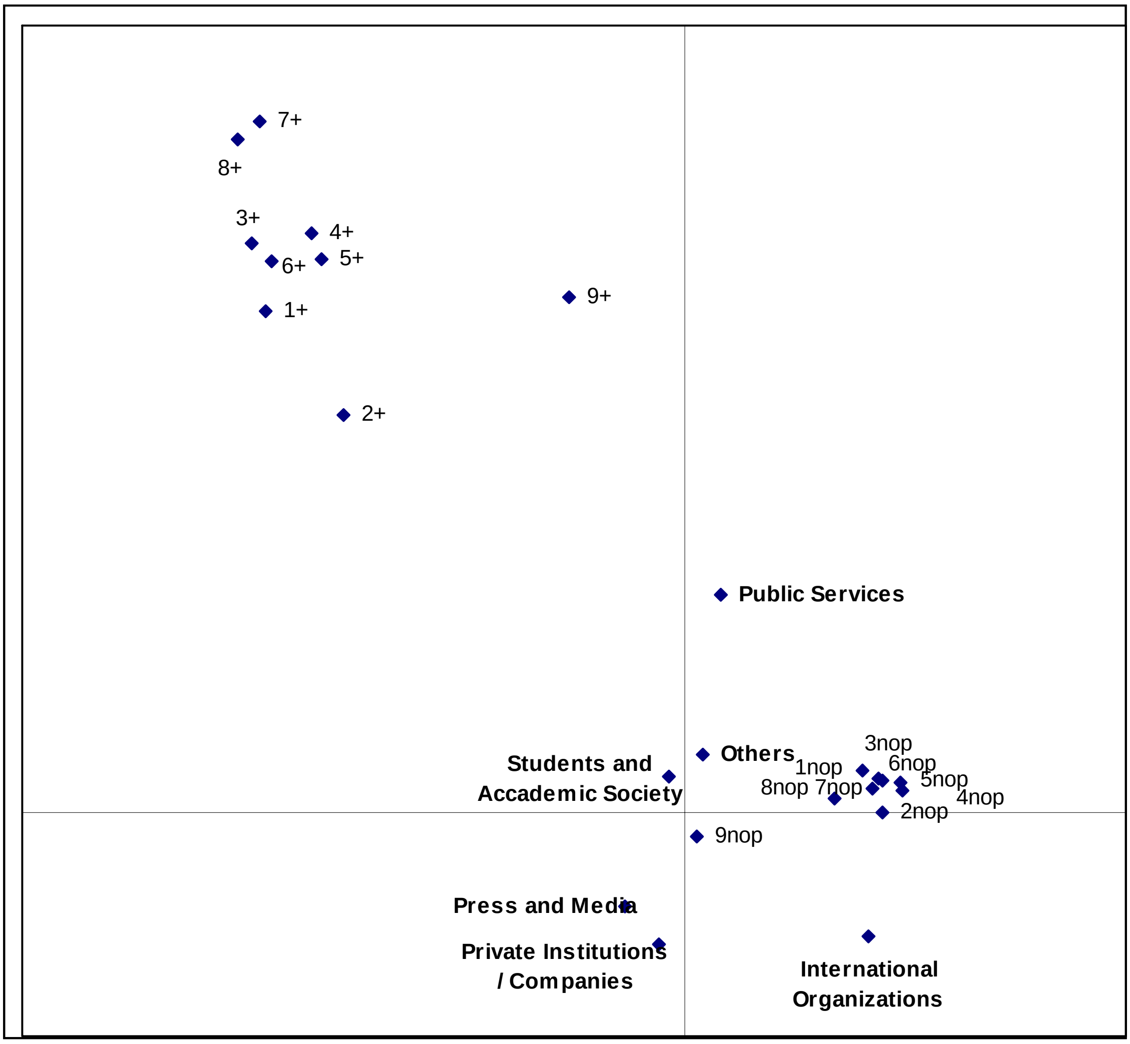
Additionally, companies put higher grades on the methodologies of "prices" and "employment statistics" and at the same time they present null negative opinions on "environment statistics".

Media seem to have a shared positive opinion on almost all kinds of official statistics, apart from "other statistics".

As for public services, they present an interesting situation on negative opinions. Concretely, they have null negative opinions on the methodologies of “national accounts”, “prices”, “public finance statistics”, “production” and “income and poverty statistics”.
The type of user “International Organizations” operates in a quite different way compared to the other users. Actually, this type of user is the only type of user where we find that in some kinds of official statistics the negative opinions are more than the positive. This stands for “national accounts” and “income and poverty statistics”. In the cases of “public finance statistics” and “environment statistics” the positive and negative points of view are equally shared. On the other hand, “production” and “foreign trade statistics” are considered to have better methodology and they gather null negative opinions.
Furthermore, the type of user “other” shows null and very low negative opinions on “other statistics” and “environment statistics” respectively. At the same time, they more highly approve the methodology of “prices” and “production”.

Example 2: Opinion on timeliness (Question: Do you consider that the official statistics are disseminated sufficiently in time for your purposes?

Factor level of the variables “type of user” and the answers “Yes” and “No opinion”



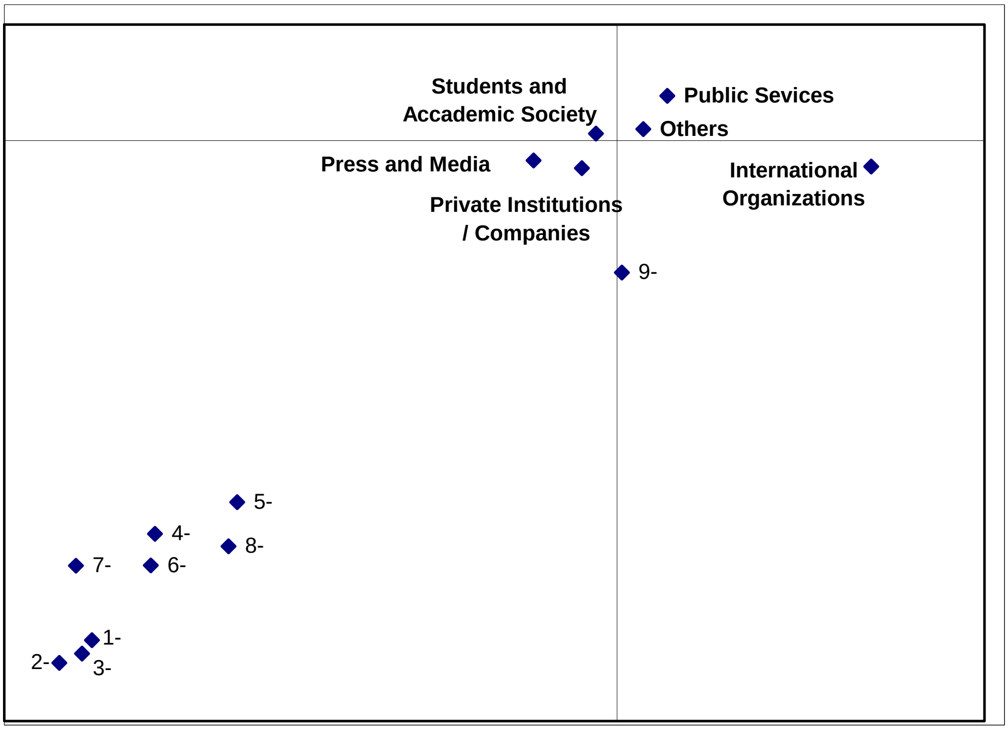
Where

1+	National Accounts	Yes
2+	Prices	Yes
3+	Public Finance Statistics	Yes
4+	Production	Yes
5+	Employment Statistics	Yes
6+	Foreign Trade Statistics	Yes
7+	Income & Poverty Statistics	Yes
8+	Environment Statistics	Yes
9+	Other Statistics	Yes

and

1nop	National Accounts	No opinion
2nop	Prices	No opinion
3nop	Public Finance Statistics	No opinion
4nop	Production	No opinion
5nop	Employment Statistics	No opinion
6nop	Foreign Trade Statistics	No opinion
7nop	Income & Poverty Statistics	No opinion
8nop	Environment Statistics	No opinion
9nop	Other Statistics	No opinion

Factor level of the variables “type of user” and the answer “No”



Where

1-	National Accounts	No
2-	Prices	No
3-	Public Finance Statistics	No
4-	Production	No
5-	Employment Statistics	No
6-	Foreign Trade Statistics	No
7-	Income & Poverty Statistics	No
8-	Environment Statistics	No
9-	Other Statistics	No

At the factor level, as the points of almost users are close to “no opinions” and far from the positive and negative opinions important percentage of users (average 69%) has no an opinion on the timeliness of the releases of official statistics for their purposes.

Public services present an interesting state of affairs, as they appear to be fully pleased with the dissemination time of official statistics in most cases. At the case of "employment statistics" the most positive and at the same time the most negative opinions are gathered. Another case worth mentioning is that of "other statistics" where the positive opinions are equal to the negative ones. In all the other cases, the negative opinions are null or very limited.
In general, the academic society follows the average pattern of users as the point of this type of users is very close the origin of the coordinates.
On the whole, companies seem to be less satisfied by the timeliness of the dissemination of several official statistics, as almost all the percentages of negative opinions are quite higher than those of the total of the users. So, at the factor level the point of "Private Institutions/ Companies" are far from positive opinions and closer to negative opinions.
Also, in "other statistics" the negative opinions are more than the positive ones as the point 9^{-} is close to all types of users.
As far as it concerns the media, the opinions are shared on "yes" and "no", while "no opinion" is quite high.
Moreover, the type of user "International organizations" assembles a very high percentage on "no opinion" and for more of the statistics the opinions are shared on "yes" and "no"
Additionally, regarding the type of user "others" the path of the type of users "students and academic society is followed as the factor level the points of these two types of users are very close.

ANNEX 4: Methodological note for the Greek Structural Business Survey

1. Annual survey on constructions (Section F)

1.1. Sample design

The survey covers the section F of NACE rev. 1. The sampling scheme that was applied is the *one-stage stratified sampling* with sampling unit the enterprise. The sample design (determination of sample size, selection of sampling units) was based on the updated business register of the NSSG.

1.2 Stratification

The enterprises included in the survey were stratified as following:

- By geographical region – NUTSII
- By class of NACE Rev.1 (4-digit level of economic activity), within each geographical region
- By size class of the enterprise. In each of the major strata (major stratum = geography x economic activity), the enterprises were stratified into L=5 size classes, according to their size, determined by their annual turnover in the business register, as follows:

Size Class	Turnover (in €)
Class 1	1 – 89.999
Class 2	90.000 – 249.999
Class 3	250.000 – 1.499.999
Class 4	1.500.000 – 9.999.999
Class 5	10.000.000+

Let h be one of the final strata (Final stratum = Geography X Economic Activity X Size Class). The final stratum that contains size classes with L=5, is census stratum (take-all).

The variable used for the construction of size classes, the size class boundaries and the number of classes were determined as follows:

- The variable used for the creation of the size classes of the enterprises belonging to the register of the NSSG is the annual turnover y of the enterprises, as the value of y in combination with the economic activity is highly correlated with all the survey characteristics. If we could stratify the enterprises by the value of y in regions and economic activity (4-digit code of NACE Rev.1), there would be no overlap between strata, and the variance within strata would be much smaller than the overall variance, particularly if there are many strata.

- Given the number of strata, for the determination of the best size class boundaries, the Dalenius-Hodges rule was applied, which is roughly equivalent to making $W_h \cdot S_h$ constant (W_h is the weight of the size class h, S_h is the standard deviation of y in the size class h, $h = 1,2,...,5$).

- The question relevant to a decision about the number of size classes is at what rate does the variance of $\widehat{Y}_{st}$ decrease as L (number of size classes) is increased? ($\widehat{Y}_{st}$: The estimated value of y in stratified sampling, given the sampling size). So, applying the Dalenius-Hodges rule, the holdings were stratified in L=4 to 7 strata, and subsequently, given the sampling size, in each separate case the variance $V(\widehat{Y}_{st})$ of $\widehat{Y}_{st}$ was calculated. As L was increased, the values of $V(\widehat{Y}_{st})$ were decreased. As very little reduction in variance appeared beyond L=5, we decided that the ideal number of the size classes should be equal to 5.

1.3 Sample size

The sampling size is 3.141 enterprises (sampling fraction =2,8%) and the response rate is equal to 46,9%. The sample size of the enterprises was defined, so that the relevant standard error (co-efficient of variation CV) of the variable "number of employees" and "turnover" at 2-digit code level of economic activity and at the whole country does not exceed 3%. The sampling units (enterprises) were distributed to the size strata by applying the method of optimal (Neyman) allocation.

1.4 Selection of the sampling units

In each of the final strata (let h), a sample of n_h enterprises was selected. The enterprises to be surveyed were selected from the total of the N_h enterprises with equal probabilities and by applying systematic sampling. The sampling units (enterprises) were selected from the sample frame based on data from the Business Register of the NSSG.

1.5 Survey characteristics estimation

a. Symbols

Defining with index i the selection order of an enterprise from the sampling frame in the stratum h and symbolizing with the y one of the survey characteristics, we can define the following:

y_{hi}: the value of the survey characteristic y of the enterprise of order i in the stratum h

Y_h: the sum of the values of the characteristic y of all enterprises falling into the survey and belonging to the stratum h

Y : the sum of the values of the characteristic y of all enterprises under the survey.

That is: $Y = \sum_h Y_{hi}$

N_h : the number of all enterprises falling into the survey and belonging to the stratum h

n_h : the sample size in the stratum h

m_h : the number of respondent units in the stratum h

r_h : the response rate in the stratum h ($r_h = \frac{m_h}{n_h}$)

w_{hi} : the extrapolation factor of the enterprise of order i belonging to the stratum h. That is: $w_{hi} = 1/$(Probability of selected unit i in stratum h) $\cdot r^{-1} = \frac{N_h}{n_h} \cdot \frac{n_h}{m_h} = \frac{N_h}{m_h}$

b. Estimation process

The estimations of the magnitudes Y_h and Y come from the following relations:

$$\hat{Y}_h = \sum_{i=1}^{m_h} w_{hi} \cdot y_{hi} \qquad (1.1)$$

$$\hat{Y} = \sum_h \hat{Y}_h = \sum_h \sum_i w_{hi} \cdot y_{hi} \qquad (1.2)$$

Generally, in order to make the estimations of the survey characteristics at any level, we add up the estimations of the final strata, which form the level under survey.

c. Variance estimation

The estimations of the variances of $\hat{Y}_h$ and $\hat{Y}$ come from the following relations:

$$V(\hat{Y}_h) = \sum_{i=1}^{m_h} w_{hi} \cdot (w_{hi} - 1) \cdot \left(y_{hi} - \bar{y}_h\right)^2 \qquad (1.3)$$

where:

$$\bar{y}_h = \frac{1}{\sum_{i=1}^{m_h} w_{hi}} \cdot \sum_{i=1}^{m_h} w_{hi} \cdot y_{hi}$$

$$V(\hat{Y}) = \sum_h V(\hat{Y}_h) \qquad (1.4)$$

The coefficient of variation (%) of the $\hat{Y}$ is given by the following relation:

$$CV(\hat{Y}) = \frac{\sqrt{V(\hat{Y})}}{\hat{Y}} \cdot 100 \qquad (1.5)$$

2. Structural survey on the trading enterprises (wholesale-retail sale)

2.1. Sample design

The survey covers the section G of NACE rev. 1, and the sampling scheme that was applied is the *one-stage stratified sampling* with sampling unit the enterprise. The sample design (determination of sample size, selection of sampling units) was based on updated the business register of the NSSG.

2.2. Stratification

The enterprises included in the survey were stratified as following:

a) By geographical region-NUTS II
b) By 4-digit code economic activity
c) By size class of the enterprises.

In each of the major strata (geography X economic activity), the enterprises were stratified into H=5 size strata, according to their size, determined by their annual turnover, as follows:

Class	Turnover description (amounts in Euros)		
1	1	Through	99.999,9
2	100.000	Through	399.999,9
3	400.000	Through	1.399.999,9
4	1.400.000	Through	4.999.999,9
5	5.000.000	Through	Highest

The enterprises that belong to the 5th turnover class were surveyed exhaustively.
The variable used for the creation of the size classes of the enterprises belonging to the register of the NSSG is the annual turnover y of the enterprises, as the value of y in combination with the economic activity (4-digit code level) is highly correlated with all the survey characteristics.
Given the number of strata, for the determination of the best size class boundaries, the Dalenius-Hodges rule was applied, which is roughly equivalent to making $W_h \cdot S_h$ constant (W_h is the weight of the size class h, S_h is the standard deviation of y in the size class h, $h = 1,2,...,5$).
We decided that the ideal number L of the size classes should be equal to 5, because very little reduction in variance $V(\hat{Y}_{st})$ appeared beyond L=5 (The same method was applied as in the survey on construction, paragraph 4.2).

2.3 Sample size

The sample size is 8.672 enterprises (sampling fraction 2,8%) and the response rate=61,6%. The sample size of the enterprises was defined, so that the relevant standard error (co-efficient of variation CV) of the variable "number of employees" and "turnover" at 2-digit code level of economic activity at the whole country does not exceed 3%. The sampling units (enterprises) were distributed to size strata applying the method of optimal (Neyman) allocation

2.4 Selection of the sampling units (enterprises)

In each of the final strata (let h) a sample of n_h enterprises was selected. The enterprises to be surveyed were selected from the total of the N_h enterprises with equal probabilities and by applying systematic sampling. The sampling units (enterprises) were selected from the sample frame based on data from Business Register of the NSSG.

2.5 Estimation of the survey characteristics

An expansion is used for the estimation of population totals. The estimation and the variance estimation processes are similar to the annual survey on constructions.

3. Structural business survey in the tourism sector

3.1. Sample design

The survey covers the section H of NACE rev. 1 and the economic activities were grouped as follows:

- 55A: 55.1+55.2
- 55B: 55.3+55.4+55.5

The sampling scheme that we applied was the *one-stage stratified sampling*, with surveyed unit the *enterprise*. The enterprises were selected from the sampling frame that was based on the updated business register of the National Statistical Service of Greece.

3.2 Stratification

The stratification of the enterprises was carried out on the basis of the following criteria:

- Geographical Regions (NUTS II
- Economic Activity (2 classes): 55A, 55B
- Turnover classes

Group	Classes	Turnover Classes
55A	1	0€ - 110.000€
	2	110.000€ – 300.000€
	3	300.000€ - 600.000€
	4	600.000€ - 1.200.000€
	5	1.200.000+

Group	Classes	Turnover Classes
55B	1	0€ - 60.000€
	2	60.000€ – 130.000€
	3	130.000€ - 300.000€
	4	300.000€ - 900.000€
	5	900.000+

The enterprises that belong to the 5th turnover class were surveyed exhaustively.

The variable used for the creation of the size classes of the enterprises belonging to the Business Register of the NSSG is the annual turnover y of the enterprises, as the value of y in combination with the economic activity (groups 55A and 55B) is highly correlated with all the survey characteristics.
Given the number of strata, for the determination of the best size class boundaries, the Dalenius-Hodges rule was applied, which is roughly equivalent to making $W_h \cdot S_h$ constant (W_h is the weight of the size class h, S_h is the standard deviation of y in the size class h, $h = 1,2,...,5$).
We decided that the ideal number L of the size classes should be equal to 6, because very little reduction in variance $V(\widehat{Y}_{st})$ appeared beyond L=5 (The same method was applied as in the survey on constructions, paragraph 4.2).

3.3 Sample size and selection of sampling units

The sample size is equal to 3.008 enterprises (sampling fraction=3,1%) and the response rate=63,9%. The sample size of the enterprises was defined, so that the relevant standard error (co-efficient of variation CV) of the variable "number of employees" and "turnover" at groups 55A and 55B of economic activity, at the whole country, does not exceed 3%.
The allocation of the number of the sample units to the various strata was carried out using the "optimal allocation" method. In each stratum (economic activity x geographical region x turnover class), the sampling units were selected with equal probabilities and the application of the systematic sampling scheme.

3.4 Survey characteristics estimation

a. Symbols

In each stratum let be:

y_{hi}: the value of the characteristic y of enterprise of order i belonging to the stratum h
N_h: the total number of enterprises belonging to the stratum h
n_h: the number of the respondent enterprises
Y_h: the total of the variable y for all enterprises in stratum h

Y: the total of the variable y for all enterprises in all strata. That is: $Y = \sum_h Y_h$

b. Estimation process

The estimation of Y_h and Y is given by the following formulas:

$$\widehat{Y}_h = \frac{N_h}{n_h}\sum_{i=1}^{n_h} y_{hi} \qquad (3.1)$$

$$\widehat{Y} = \sum_h \widehat{Y}_h \qquad (3.2)$$

c. Variance estimation

The variance estimation of $\widehat{Y}_h$ and $\widehat{Y}$ is given by:

$$V\left(\widehat{Y}_h\right) = \frac{N_h\left(N_h - n_h\right)}{n_h} s_h^2, \quad (3.3)$$

where:

$$s_h^2 = \frac{1}{n_h - 1}\left[\sum_{i=1}^{n_h} y_{hi}^2 - \frac{\left(\sum_{i=1}^{n_h} y_{hi}\right)^2}{n_h}\right],$$

$$V\left(\widehat{Y}\right) = \sum_h V\left(\widehat{Y}_h\right) \qquad (3.4)$$

The coefficient of variation (%) of total estimation $\widehat{Y}$ is given by:

$$CV\left(\widehat{Y}\right) = \frac{\sqrt{V\left(\widehat{Y}\right)}}{\widehat{Y}} * 100 \qquad (3.5)$$

4. Survey on Structural Business Statistics (SBS) on transport sector (Section I)

4.1. Sample design

The survey covers the section I of NACE rev. 1. The sampling scheme that was applied is the *one-stage stratified sampling* and the sampling unit is the enterprise. The sample design (determination of the sample size, selection of the sampling units etc) was based on the updated business register of the NSSG.

4.2. Stratification

The enterprises included in the survey were stratified as following:

a) By geographical region-NUTS II

b) By size class of the enterprises. In each of the major strata (geography x size class), the enterprises were stratified into H=6 size strata, according to their size, determined by their annual turnover, as follows:

Class	Turnover description (amounts in Euros)		
1	1	Through	49.999
2	50.000	Through	199.999
3	200.000	Through	599.999
4	600.000	Through	1.999.999
5	2.000.000	Through	4.999.999
6	5.000.000	Through	Highest

The enterprises that belong to the 6th turnover class were surveyed exhaustively.

The variable used for the creation of the size classes of the enterprises belonging to the Business Register of the NSSG is the annual turnover y of the enterprises, as the value of y in combination with the economic activity is highly correlated with all the survey characteristics.
Given the number of strata, for the determination of the best size class boundaries, the Dalenius-Hodges rule was applied, which is roughly equivalent to making $W_h \cdot S_h$ constant (W_h is the weight of the size class h, S_h is the standard deviation of y in the size class h, $h = 1,2,...,5$).
We decided that the ideal number L of the size classes should be equal to 6, because very little reduction in variance $V(\widehat{Y}_{st})$ appeared beyond L=6 (The same method was applied as in the survey on constructions, paragraph 4.2).

4.3 Sample size and selection of sampling units

The sample size is equal to 3.238 enterprises (sampling fraction=7,5%) and the response rate=54,7%. The sample size of the enterprises was defined, so that the relevant standard error (co-efficient of variation CV) of the variable "number of employees" and "turnover" at 2-digit level of economic activity at the whole country not to exceed 3%.
The allocation of the number of the sample units to the various strata was carried out using the "optimal allocation" method. In each stratum (economic activity x geographical regions x turnover class), the sampling units were selected with equal probabilities and the application of the systematic sampling scheme.

4.4 Survey characteristics estimation

An expansion is used for the estimation of population totals. The estimation and the variance estimation processes are similar to the annual survey on constructions.

ANNEX 5: Methodological note for the Job Vacancy Survey

1. *Scope of the survey*

The survey covers the sections C-O of NACE Rev.1 and the enterprises with average annual employment equal to or greater than one employee. The one- stage stratified sampling method was applied, employing the enterprise as a surveyed unit. The sampling frame used for the sample design was based on the Business Register (BR) of the NSSG. This BR is based on the VAT Register of the Ministry of Economy and Finance and it is updated through the statistical surveys of the NSSG and the register of the Social Insurance Foundation.
The statistical data for the section L and for the public services of sections M (Ministry of education and public schools) and N (Ministry of Health, public hospitals and public health centres) were collected from the Ministry of Interior, Public Administration and Decentralization

2. Stratification

The enterprises included in the survey were stratified as following:

a. By geographical region – NUTS II
b. By 2-digit code NACE Rev.1 economic activity, within each geographical region
c. By size class of the enterprise. In each one of the major strata (major stratum = geography x economic activity), the enterprises were stratified into L=6 size classes, according to their size, determined by their average annual number of employees in the business register, as follows.

Size class	Number of employees
1	1-9
2	10-49
3	50-249
4	250-499
5	500-999
6	1000+

Let h be one of the final strata (Final stratum = Geography X Economic Activity X Size Class). The final stratum that contains size classes with L=6, is census stratum (take-all).

3. Sample size

The sampling size is 4.266 enterprises (sampling fraction =2,1%) This decision of the sample size was based on financial criteria and on the existing experience as far as the accuracy of the resulting statistics is concerned. The response rate is equal to 76,5%, 79,4% and 80,9% for the 1st, 2nd and 3rd quarter of the year 2006, respectively. The sampling units (enterprises) were distributed to the size strata by applying the method of optimal (Neyman) allocation.

Quarters				
	1	2	3	4
Sample size	4.266	4.266	4.266	4.266
Responses	3.264	3.386	3.451	3.357
Response Rate	76,5%	79,4%	80,9%	78,7%

4. Estimation of the survey characteristics estimation

a. Symbols

Defining with index i the selection order of an enterprise from the sampling frame in the stratum h and symbolizing with the y one of the survey characteristics, we can define the following:

y_{hi} : the value of the survey characteristic y of the enterprise of order i in the stratum h

Y_h : the sum of the values of the characteristic y of all enterprises falling into the survey and belonging to the stratum h

Y : the sum of the values of the characteristic y of all enterprises under the survey. That is: $Y = \sum_h Y_{hi}$

N_h : the number of all enterprises falling into the survey and belonging to the stratum h

n_h : the sample size in the stratum h

m_h : the number of respondent units in the stratum h

r_h : the response rate in the stratum h ($r_h = \frac{m_h}{n_h}$)

w_{hi} : the extrapolation factor of the enterprise of order i belonging to the stratum h. That is: $w_{hi} = 1/$(Probability of selected unit i in stratum h)$\cdot r^{-1} = \frac{N_h}{n_h} \cdot \frac{n_h}{m_h} = \frac{N_h}{m_h}$

b. Estimation process

The estimations of the magnitudes Y_h and Y come from the following relations:

$$\hat{Y}_h = \sum_{i=1}^{m_h} w_{hi} \cdot y_{hi} \quad (1)$$

$$\hat{Y} = \sum_h \hat{Y}_h = \sum_h \sum_i w_{hi} \cdot y_{hi} \quad (2)$$

c. Variance estimation

The variance estimation of $\hat{Y}_h$ and $\hat{Y}$ is given by:

$$V(\hat{Y}_h) = \frac{N_h \cdot (N_h - m_h)}{m_h} S_h^2, \ (3)$$

where:

$$S_h^2 = \frac{1}{m_h - 1}\left[\sum_{i=1}^{m_h} y_{hi}^2 - \frac{\left(\sum_{i=1}^{m_h} y_{hi}\right)^2}{m_h}\right],$$

$$V(\hat{Y}) = \sum_h V(\hat{Y}_h) \quad (4)$$

The coefficient of variation (%) of total estimation $\hat{Y}$ is given by:

$$CV(\hat{Y}) = \frac{\sqrt{V(\hat{Y})}}{\hat{Y}} * 100 \quad (5)$$

ANNEX 6: Relationship between the degree of urbanization (stratification variable) and the basic variable (income)

Greek EU-SILC Survey

There are two levels of area stratification in the sampling design. The first level is the geographical stratification based on the partition of the total country area into thirteen standard administrative regions corresponding to the European NUTS II level. The two major city agglomerations of Greater Athens and Greater Thessalonica constitute separate major geographical strata.
The second level of stratification entails grouping municipalities and communes within each NUTS II administrative region by the degree of urbanization, i.e., according to their population size. The scaling of urbanization was finally designed in four groups:

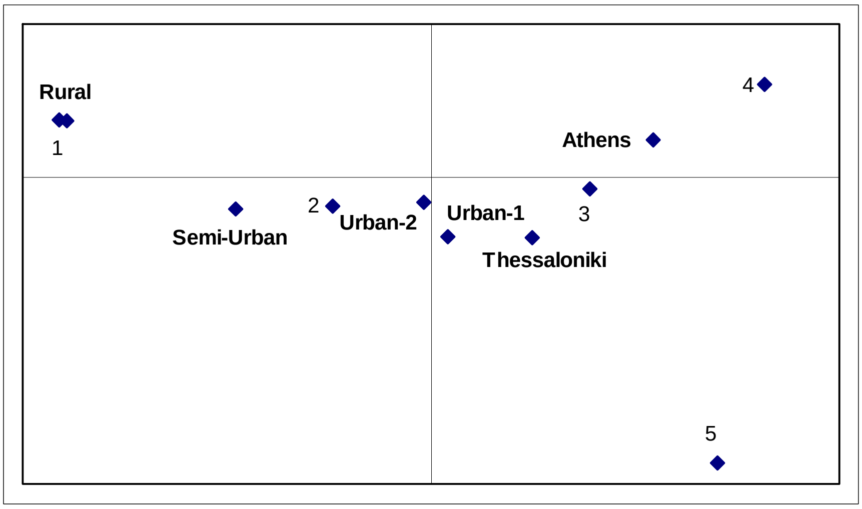
- >= 30.000 inhabitants (Urban-1)
- 5.000 - 29.999 inhabitants (Urban-2)
- 1.000 - 4.999 inhabitants (Semi-urban)
- 0 - 999 inhabitants (Rural)

The Greater Athens Area was divided into 31 strata of about equal size (equal number of households) on the basis of the lists of city blocks of the Municipalities that constitute it and taking into consideration socio-economic criteria. Similarly, the Greater Thessaloniki Area was divided into 9 equally sized strata.

In our example we study the relationship between degree of urbanization and income applying Correspondence Factor Analysis. As this type of analysis works with non-continuous variables the income of persons was divided in five size classes, as follows:

Size Class	Income intervals
Class 1	$[x_1, x_2)$
Class 2	$[x_2, x_3)$
Class 3	$[x_3, x_4)$
Class 4	$[x_4, x_5)$
Class 5	x_5+

Factor level of the variables “degree of urbanization” and “income”



At the factor level, the points of “Rural” and income class “1” are very close to each other. This means that there is a strong correlation between this type of urbanization and the 1st class of income.
The point of income class “2” is between the points of Semi-urban and Urban-2. This means that the relation of these types of urbanization with the 2nd class of income is strong.
The point of income class “3” is between the points Urban-1, Thessaoniki and Athens, and the point of income class “4” is close to the point of Athens. This means that strong relationship exists between the income classes “3” and “4” and the urbanization. The relatively big distance between the point of income class “5” and the points of Urban-1, Thessaloniki and Athens shows that little relationship exists the between the income class “5” and these types of urbanization. However, the income class “5” has no relationship with the types of urbanity “Rural” and “Semi-Urban”.

Applying correspondence factor analysis, the relationship of non-continuous stratification variables with the main variables can be, approximately, monitored. In this case the picture of the “factor level” provides the quality manager with an important tool for monitoring and assessing the successive choice of the stratification variable applied for creating homogeneous strata, when the stratification variable is non-continuous.

ANNEX 7: The effects of stratification, weighting and clustering on sampling errors as process quality indicators

1. Introduction

The design effect of the survey estimates and the evaluation of the design effect's components can be used as tools for assessing the quality of the sample design and the data collection process of the complex sample surveys. The design effect is defined as the ratio of the variance of an estimate under the complex sample design to the variance of the same estimate that would have been obtained from a simple random sample of the same size (Kish, 1965, 8.2). The design effect represents the combined effect of a number of components such as stratification, clustering, and unequal weighting. Thus, the measurement of the design effects for each of these components individually assesses their effects on the precision of the survey estimates.

The multi-stage surveys often employ complex sample designs typically involving a number of design features, such as stratification, clustering, and unequal weighting (e.g. weighting due to unequal selection probabilities or due to adjustments for non-response). The efficiency of the complex sample design may be evaluated for each design feature through decompositions of the design effects. Kish (1987) proposes to model the overall design effect for a weighted sample mean $\hat{\bar{Y}}$, $deft^2(\hat{\bar{Y}})$, as a product of two individual components $deft_W^2(\hat{\bar{Y}})$ and $deft_C^2(\hat{\bar{Y}})$ associated with the unequal weighting and the clustering, respectively. This Kish's two-factor decomposition model is written as:

$$deft^2(\hat{\bar{Y}}) = deft_W^2(\hat{\bar{Y}}) \cdot deft_C^2(\hat{\bar{Y}}) \quad (1.1)$$

The formula (1.1) holds not only for estimated population means $\hat{\bar{Y}}$, but also for population proportions and ratios of two estimated variables. Additionally, expression (1.1) ignores the stratification component, and thus $deft_C^2(\hat{\bar{Y}})$ includes not only the effect of clustering but also the effect of the stratification. Thus, as the effect of stratification decreases the quantity of $deft^2$, the real value of $deft_C^2$ is underestimated applying the two-factor decomposition model .

Park, Winglee, Clark, Sedlak and Morganstein (2003) consider the three-factor decomposition model, (it is an extension of Kish's type production model), according to which the design effect is decomposed in one more component the $deft_S^2(\hat{\bar{Y}})$, which quantifies the stratification effect on the design effect. The three-factor decomposition model is written as:

$$deft^2(\hat{\bar{Y}}) = deft_S^2(\hat{\bar{Y}}) \cdot deft_W^2(\hat{\bar{Y}}) \cdot deft_C^2(\hat{\bar{Y}}) \quad (1.2)$$

The design effect and the individual components $deft_S^2(\hat{\bar{Y}})$, $deft_W^2(\hat{\bar{Y}})$ and $deft_C^2(\hat{\bar{Y}})$ are presented and analyzed for the "unemployment rate" of the Greek Labour Force with reference period 2nd quarter 2007. The analysis is based on both Kish's two-factor and three-factor decomposition model and its target is to examine if the design effect and the design effect's components may be used as quality process indicators assessing the sample design process (clustering, stratification, weighting) and the effect of the additional weighting due to non-response and coverage problems on the sampling errors of the results.

2. Stratification

Stratification may produce a gain in precision in the estimates of characteristics of the whole population. It may possible to divide a heterogeneous population into subpopulations, each of which is internally homogeneous. The gains in precision in stratified sampling for means or proportions or ratios may be quantified for assessing the effect of the stratification on the sampling error and on the design effect.

The stratification effect on the sampling error and on the design effect is measured directly as follows (Park, 2004):

$$deft_S^2(\hat{\bar{Y}}) = \frac{V(\hat{\bar{Y}}_{SCW})}{V(\hat{\bar{Y}}_{CW})} \quad \textbf{(2.1)}$$

where:

(a) $V(\hat{\bar{Y}}_{SCW})$: the final variance of the mean or ratio in which stratification, clustering and weighting has been applied

(b) $V(\hat{\bar{Y}}_{SC})$:the variance of the mean or ratio in which only weighting and clustering has been applied

3. Weighting

3. 1 Reasons for weighting

In many of the households surveys, regional and separate urban / semi-urban / rural estimations are important objectives, often requiring sampling fraction at different values so as to obtain adequate sample sizes for different domains. Departures from self-weighting (all sampling units have the same chance to be selected) occur targeting either to increase the sample sizes in some domains (regions, urban, semi-urban, rural strata etc) because of the increased emphasis on these domains or to decrease the costs in some others or sometimes to decrease the element variance of the sample data. These different sampling rates (fractions) require weighting of the sample data for the estimations of the survey characteristics. Additionally, weighting adjustments are commonly applied in surveys to compensate for non-response and non-coverage, and to make weighted sample estimates conform to external values (values from censuses, administrative sources etc). Weights are assigned to respondent units in a survey data file in order to make the weighted sampling units represent the population of inference as closely as possible. The weights are usually developed in a series of stages to compensate for unequal sampling fractions, unequal selections probabilities, non-response, non-coverage, and sampling fluctuations from known population values (Kalton and Flores-Cervantes, 2003).

The weights concerning the weighting adjustments (compensation for differential non-response, shortcoming in sample implementation etc) are essentially random variables, not related to differences in domain variances and tend to inflate the sampling errors and the design effects of the estimates. It is important the effect of random weights on sampling errors and design effects to be examined, as this effect tends to persist undiminished on subclasses and uniformly inflate all design effects (Verma and Le, 1996).

In the first stage of weighting and applying probability sampling, the selection probabilities are known and the base weights are generally readily determined.

In the second stage of weighting, the weight development attempts to compensate for non-response. The base weights of responding elements are adjusted to compensate for the non-responding elements. Respondents and non-respondents usually are sorted into homogeneous weighting cells or classes or strata or post-strata (e.g. cell = age groups x sex), which are formed from available auxiliary information. The weights of the respondents in each cell are increased by a multiplying factor, so that the respondents represent the non-respondents in theses cells (Kalton and Kasprzyk, 1986).

In the third stage of weighting, the weight development involves a further adjustment to the weights to make the resultant weighted estimates from the sample conform to known population values for some key variables. This stage of adjustments serves two purposes: (a) to compensate for non-coverage and (b) to improve the precision of the survey estimates. Statistical adjustments for improved estimates have diverse names: post-stratification, calibration etc. (Kalton and Flores-Cervantes, 2003).

The weighting due to non-response and non-coverage is used to remove the bias. However, this weighting usually increases the sampling errors due to random weighting. The effect of weights on the design effect and sampling errors should be quantified and used as a quality indicator assessing the weighting process.

Concerning the Greek Labour Force, the weighting adjustment due to non-response is carried out at the level of clusters (primary sampling units composed from one or more unified blocks), and in each cluster the initial weights of the respondents (inverse of probability selections) by the inverse of response rates recorded at the clusters. In the third stage of weighting of this survey, the respondents (individuals) are post-stratified into homogeneous post-strata (post-stratum= age groups x sex) and in each post-stratum the interim weights (initial weights x inverse of response rates of clusters) are multiplied by a factor created so that the final estimates from the sample data to conform to known NUTS II region total population values for the key variables sex and age groups.

3.2 Weighting process

The general and most useful form of weighting is to assign the weights w_i to the sample units i, with $w_i = \frac{1}{P_i}$, where P_i is the selection probability of sample unit i ($i = 1,...,n$=sample size). The selection probabilities for all sample units must be known for all probability samples by definition. Usually, the determination of selection probabilities is carried out after stratified the population units into homogeneous strata decreasing the element variances. Disproportional sampling fractions can be introduced not only decreasing variances but also the costs. Additionally, the weights are used to compensate for non-responses, so that $w_i = \frac{1}{(P_i \cdot r_h)}$, where r_h is the response rate often calculated for classes (h) of response or post-strata or strata for stratified sampling or for clusters. It is also possible to incorporate adjusted weights W_i making the weighted respondent distributions for certain variables conform to population total for these variables, so that:

$$w'_i = \frac{W_i}{(P_i \cdot r_h)} \quad (3.1)$$

The adjusted weights, which correct non-sampling errors arisen from non-response and frame problems, are calculated applying one of the following methods:

Cell weighting: "The standard cell weighting procedure adjusts the sample weights so that the totals conform to population totals on a cell-by-cell basis. The assumption, which underlies the cell weighting adjustments for non-response, is that respondents within a cell represent the non-respondents. Unlike other methods, cell weighting makes no assumptions about the structure of the response probabilities across cells. A potential disadvantage of cell weighting is that it can lead to a large variability in the distributions of the weighting adjustments, thereby inflating the variances of the survey estimates" (Little 1986; Kalton and Kasprzyk 1986). This happens especially, when the sample size is not large.

Raking: "This operates on the marginal distributions of the auxiliary variables. Raking is an iterative proportional fitting procedure, according to which the sample row totals are forced to conform to the population row totals and afterwards the sample adjusted column totals are forced to conform to population column totals; then the row totals are readjusted to conform and so on until convergence is reached. Reasonable convergence is usually reached fairly rapidly. However, in some cases the convergence is reached slowly"(Kalton and Flores-Cervantes 2003; Ireland and Kullback 1968; Oh and Scheuren 1987).

Linear weighting: "It adjusts the weights to make the sample marginal distributions agree with the population marginal distributions. This method is like raking except a different distance function is used" (Deville and Sarndal 1992; Deville and Sarndal and Sautory 1993). Linear weighting has the undesirable feature, as in some cases negative weights are produced.

GREG weighting: "This weighting adjustment derives from the standard regression estimator in survey sampling. This method extends to cover several auxiliary variables in the regression model and to incorporate unequal weights" (Deville and Sarndal 1992; Fuller, McLoughlin and Baker 1994).

Concerning the Greek Labour Force the 'cell weighting' method is applied for the final weighting adjustment of the sample weights, so that the totals conform to population totals on a cell-by-cell. This method is applied after the weighting adjustment due to non-response.

3.3Design effect's weighting component $deft_W^2(\hat{\bar{Y}})$ **or** D_w^2

The unequal selection probabilities between strata with disproportionate stratification result in the need to use weights in the estimation and analysis of survey data. The weighting due to disproportionate sampling and the non-response and non-coverage compensations inflate the design effects. Thus, design effect's component $deft_W^2(\hat{\bar{Y}})$ (or D_w^2) represents the increment in the design effect and the variance due to unequal weighting.

The effect unequal weighting inflates the design effect and the variance of the estimated weighted population means by a factor (Kish 1965; 1990; 1992):

$$D_w^2 = \frac{n \cdot \sum_i k_i^2}{\left(\sum_i k_i\right)^2} = 1 + cv^2(k_i) \qquad (3.2)$$

where:

k_i: the final weight of the unit i ($k_i = \frac{W_i}{(P_i \cdot r_h)}$, when the effect of the unequal weighting (disproportionate stratification sampling, random weighting to compensate non-response, and frame problems) on design effect is measured.

$cv(k_i)$: the coefficient of variation of the weights

3.3.1 Remarks on the formula 3.2

i. It is in fact, the ratio of the variance of $\hat{\bar{Y}}$ (estimated mean of the variable y) under disproportionate stratified sampling to that under proportionate stratified sampling.

ii. Apart from means and population proportions, it holds approximately and for ratios.

iii. It is based on the assumptions of equal strata means and unit variances. In household surveys, the assumption of equal, or approximately equal, within-stratum variances is often reasonable. One type of estimate for which the within-stratum variances may be unequal is a proportion. However, if the unit variance in the stratum h with a proportion P_h is variation in $S_h^2 = P_h \cdot (1 - P_h)$ then the variation of S_h^2 is only slight for proportions between 0.2 and 0.8, from a high of 0.25 for $P_h = 0.25$ to a low 0.16 for $P_h = 0.2$ or $P_h = 0.8$.

iv. It does not work well in the case that the weights are post-stratified or calibrated to known control totals from an external source and the variable y is highly correlated with one or more control totals. For example assume the weights are post-stratified to control totals of the number of persons in the country by sex and age groups. Consider the extreme case where the survey data are used to estimate the proportion of men or women belonging to one age group in the population. In this case of a perfect correlation between the variable y and the control variable, the estimated proportion is not subject to sampling error and hence has zero variance and design effect. When the correlation is sizeable, post-stratification or calibration to known population totals can improve the precision of the survey estimations, but this improvement will not be shown through the application of the formula (3.2). On the contrary, this formula will indicate a loss in precision (increment of the variance and the design effect) (Kalton, Brick and Le, 2005).

3.3.2 Direct estimation for the effect of weighting

For avoiding the problems mentioned on the remarks (iii) and (iv) of the above capital (3.3.1), a direct and completely accurate estimation of the effect of weighting ($deft_W^2(\hat{\bar{Y}})$ or D_w^2) on the design effect and the variance is achieved as follows (Park 2004):

$$deft_W^2(\hat{\bar{Y}}) = \frac{V(\hat{\bar{Y}}_{SCW})}{V(\hat{\bar{Y}}_{SC})} \quad \textbf{(3.3)}$$

where:

a. $V(\hat{\bar{Y}}_{SCW})$: the final variance of the mean or ratio in which stratification, clustering and weighting has been applied

b. $V(\hat{\bar{Y}}_{SC})$:the variance of the mean or ratio in which only stratification, clustering has been applied

3.4 Effect of random weighting on design effect and sampling errors

In the formula (3.1), the initial weights $w_i = \frac{1}{P_i}$ are not random variables, as their values have been determined through the selection process. The probabilities are calculated from data coming from the sampling frame on which the sample design has been based. However, the additional weights $\frac{1}{r_h}$ and W_i are random variables, because their values are determined after the data collection. As the weights $\frac{1}{r_h}$ and W_i are random, not related to differences in domain element variances, they tend to inflate the sampling errors. The effect of essentially arbitrary weights is to *uniformly* increase the variances for all survey characteristics, but reducing the bias due to non-response. (Verma and Le, 1996).

a) The effect of arbitrary (random) weights is to inflate the design effect and the variance of the estimated characteristics (ratios and population means) by a factor D_{wrand}^2 , which is estimated as follows:

$$D^2_{wrand} = \frac{D^2_{wfin}}{D^2_{win}} \quad (3.4)$$

In the formula (2.4), the quantities D^2_{win} and D^2_{wfin} and are calculated using the formula (3.2) with k_i the initial and final weights, respectively. In the case that, k_i are the initial weights and the interim weights applied for compensating the non-response, then through formula (3.4), it is measured the effect on non-response on the sampling errors (Nikolaidis, 2006).

For avoiding the problem mentioned on the remark (iii) of the above capital (3.3.1), a direct and estimation of the effect of random weighting is achieved as follows:

$$V_{wrand}(\hat{\bar{Y}}) = \frac{V_{wfin}(\hat{\bar{Y}})}{V_{win}(\hat{\bar{Y}})} \quad (3.5)$$

where: V_{win} and V_{wfin} stand for the variances of $\hat{\bar{Y}}$ applying the initial and the final weights, respectively.

b) The value of the D_{wrand} (square root of the D^2_{wrand}) or the direct estimation value $V_{wrand}(\hat{\bar{Y}})$ quantifies the inflation of (a) the square root of the design effect and (b) the sampling error due to unequal random weighting. Thus, the effect of random weights on sampling errors is quantified as:

$$D_{wrand} = \sqrt{D^2_{wrand}} \quad (3.6.1) \text{ or } ste_{wrand}(\hat{\bar{Y}}) = \sqrt{V_{wrand}(\hat{\bar{Y}})} \quad (3.6.2)$$

d) The random weighting usually increases the variances, but it removes approximately the biases. The quantities of the biases which are removes are expressed through the relative bias and the bias ratio as follows:

$$\text{Relative bias} = \frac{\hat{\bar{Y}}_{win} - \hat{\bar{Y}}_{wfin}}{\hat{\bar{Y}}_{win}} \quad (3.7)$$

$$\text{Bias ratio} = \frac{\hat{\bar{Y}}_{win} - \hat{\bar{Y}}_{wfin}}{ste(\hat{\bar{Y}}_{win})} \quad (3.8)$$

where: (a) $\hat{\bar{Y}}_{win}$ and $\hat{\bar{Y}}_{wfin}$ stand for the estimates of $\bar{Y}$ applying the initial and the final weights, respectively and (b) $ste(\hat{\bar{Y}}_{win})$ stands for the standard error of $\hat{\bar{Y}}_{win}$.

3.4.1 Remarks on the formula 3.4

- The value of D^2_{win} quantifies the increment of the design effect due to unequal weighting arisen only from the disproportionate stratification (unequal sampling fraction or/ and unequal selection probabilities). In self-weighting samples $D^2_{win}=1$.
- The value of D^2_{wfin} quantifies the increment of the design effect due to unequal weighting arisen from the disproportionate stratification and the random sources such as non-responses or frame problems.
- The value of D^2_{wrand} quantifies the increment of the design effect due to unequal weighting arisen only from the random sources such as non-responses or frame problems.

3.4.2 Quality process indicators on weighting

3.4.2.1 Random weighting

*i) **Relative bias (RB) and Bias ratio (BR)***

The non-response and the frame problems create non-sampling errors (biases), which can be measured (approximately) through the sample data by calculating (a) relative bias and (b) bias ratio applying the formulae (3.7) and (3.8), respectively. The relative bias and the bias ratio are be measured and considered as indicators, because (a) they assess the data collection process and (b) they provide us with important information on the magnitude of the bias, which should be eliminated through the additional (random) weighing. In case that, the absolute bias ratio is not less than 0.1, then the additional random weighting (due to non-response and coverage problems) is required and is completely justified.

ii) Inflation of the design effect due to random weighting

For approximately removing (or more accurately for reducing) the bias due to non-response and frame problems, weighting adjustments are applied which usually inflate the design effects and the sampling errors. The percentage of these inflations can be considered as *quality process indicator,* expressed by D_{wrand} or $ste_{wrand}(\hat{\bar{Y}})$, which are calculated through the formulae (2.4) and (2.6.1-2.6.2), respectively. It is required the values of D_{wrand} calculated by Kish's formula to be accompanied by the values of the $ste_{wrand}(\hat{\bar{Y}})$, because in some cases the D_{wrand} may not work well (see iii-iv of 3.3.1).

3.4.2.2 Total weighting

The effect of the total weighing (initial weighting due to unequal probabilities selections plus random weighting due non-response and frame errors) on the design effect and the variance is quantified through the value of the quantity D^2_{wfin} or the direct estimation $deft^2_w(\hat{\bar{Y}})$. Both D^2_{wfin} and $deft^2_w(\hat{\bar{Y}})$ may be considered as quality process indicators assessing simultaneously (a) the initial sample design (disproportional allocations in domains, unequal selections probabilities) and (b) the random weighting process. Thus, it is important, these process indicators to accompany the values of the design effect, so that the share of the weighting component on the total design effect to be measured and assessed.

The advantages of the RB, BR, D_{wrand}, $ste_{wrand}(\hat{\bar{Y}})$, D^2_{wfin} and $deft^2_W(\hat{\bar{Y}})$ as quality indicators, are as follows:

- They can be estimated using only sample data and not external sources of statistical information
- They incorporate all weighting processes for reducing the bias due to non-response and non-coverage
- They can be calculated easily.

4. Clustering

The second major component of the design effect in most household and population surveys is the effect of clustering on the variance and the design effect in the multistage surveys. Samples are clustered to reduce the data collection costs, because it is uneconomical to list and sample households spread across an entire country or region. Usually, two stages of sampling are employed, where the first-stage or primary sampling units (PSUs) are defined geographical areas (e.g. one or more unified city blocks) that are selected with probabilities proportional to their size (usually the number of households from the last general dwelling and population census). In each selected PSU, updated sampling frame (list) with dwellings is compiled, and the sampling households living in the dwellings are selected with equal probabilities from the updated frame. In the case that, two or more households are included in one selected dwelling, then all households are surveyed. For a survey of persons, a list of persons from the sampled households is compiled and either all or a sample of persons belonging to the target population is surveyed.

4.1. Design effect's clustering component $deft_C^2(\bar{\bar{Y}})$

The value of the design effect $deft^2(\bar{\bar{Y}})$ of the weighted mean $\bar{\bar{Y}}$ for a given sample design tends to inflate as the cluster size increases. For measuring this inflation, which is called effect of clustering and symbolized as $deft_C^2(\bar{\bar{Y}})$, Kish (1965) introduced the second the design effect's component as follows:

$$deft_C^2(\bar{\bar{Y}}) = 1 + (\bar{b} - 1) \cdot \rho \quad (4.1)$$

where:

a) ρ is the intra-class correlation coefficient (or rate of homogeneity: roh), which measures the homogeneity of the y- variable in the PSUs. In practice, units within a PSU tend to be somewhat similar to each other of nearly all variables, although the degree of similarity is usually low. Thus, ρ is almost always positive and as a rule greater than 1. However, there are extreme cases in which $\rho \leq 0$. The value of Roh is a synthetic measure introduced with the aim of measuring the average degree to which values of particular variable are homogeneous within the PSUs.
b) $\bar{b}$ is the average cluster size. That is: $\bar{b} = \frac{\sum b_i}{a}$ (b_i is the number of elementary sample units in the PSU of order i and a is the total number of sampling PSUs).The expression (3.1) has been developed in the absence of large variations in the cluster sizes.
c) In the present of large variations in cluster sizes, it is more appropriate to compute the "average size" as $\bar{b}' = \frac{\sum b_i^2}{b_i} = \bar{b} \cdot (1 + cv(b_i)$ in the place of the simple average $\bar{b}$ (Verma and Le, 1996).

The expression (4.1) shows that the design effect from clustering the sample within the PSUs depends on two factors: (a) average the sample size $\bar{b}$ (number of the sampling households or individuals) within the selected PSUs and (b) the intra-class correlation ρ for the statistic y in question.

Additionally, the expression (4.1) measures the increment in the variance $\bar{\bar{Y}}$ due to clustering for multistage equal probability samples. This expression originates from the design effect for equal probability cluster sampling of equal size clusters. In reality, PSUs are not of equal size and they are not selected by simple random sampling. In almost the households sample designs, stratified samples of PSUs are selected with probabilities proportional to their sizes. Hence, the expression (4.1) does not apply directly. However, it serves as an approximation of the design effect's second component considering that: (a) the ρ is the average within-stratum measure of homogeneity, provided that the homogeneity within each stratum is roughly of the same magnitude and (b) the size b_i of the sub-sample in PSU of order i is replaced by the average sub-sample size $\bar{b}$ (Kalton, Brick and Le, 2005).

4.1.1 Calculation of clustering effect $deft_C^2(\hat{\bar{Y}})$

The value of $deft_C^2(\hat{\bar{Y}})$ can be calculated as follows:

a) Kish's two-factor decomposition model

$$deft_C^2(\hat{\bar{Y}}) = \frac{deft^2(\hat{\bar{Y}})}{deft_W^2(\hat{\bar{Y}})} \quad (4.2)$$

where:

$deft_w^2(\hat{\bar{Y}}) = D_{wfin}^2$ (Kish's formula) or $deft_W^2(\hat{\bar{Y}}) = \dfrac{V(\hat{\bar{Y}}_{SCW})}{V(\hat{\bar{Y}}_{SC})}$ (direct estimation)

b) Three -factor decomposition model

1[st] Method: The above expression (4.2) has been developed in the absence of the effect of stratification (implicit or explicit). To isolate the effect of stratification, we divide the $deft_C^2(\hat{\bar{Y}})$ on the left hand by $deft_S^2(\hat{\bar{Y}})$.

2[nd] Method (Direct estimation): The clustering effect on the design effect is estimated directly as follows (Park, 2004):

$$deft_C^2(\hat{\bar{Y}}) = \frac{V(\hat{\bar{Y}}'_C)}{V(\hat{\bar{Y}}'_{srs})}$$

where:

(a) $V(\hat{\bar{Y}}'_C)$: the variance of the mean or ratio in which only clustering has been applied

(b) V($\hat{\bar{Y}}_{srs}$) :the variance of the mean or ratio in which only the simple random sampling scheme has been applied

4.1.2 Quality indicators on clustering

Both the quantity $deft_C^2(\hat{\bar{Y}})$ and the rate of homogeneity (Roh) or intra-class correlation ρ may be used as a quality process indicators assessing not only the effect of clustering on the results but also the of sample design on the results, because these two indicators are mainly related with the number of sample clusters in the strata. Clusters of moderately large size may be efficient sampling units when the intra-class correlation between elements within clusters is low positive or negative ($|\rho| \leq 5\%$) and will be less efficient when the correlation is positive and high. If the $|\rho| > 5\%$, the number of sample clusters should be increased so that the average sample size and the homogeneity within clusters to be decreased producing results more accurate, as the final (ultimate) elementary sample lies in more clusters and thus is more representative.

5. Case study on Greek Labour Force data

The effect of weighting on the sampling errors and the components of the design effects (two-model and three-model) are studied and presented for the Greek Labour Force with reference period the second quarter of the year 2007. The two-stage area sampling was adopted for the survey. The primary units are the areas (one or more unified blocks), and the ultimate (final) sampling units selected in each sampling area are the households in which all the individual are surveyed (detail description of the sample design is presented in the annex 7.4).

According the sample design in each final stratum (crossing of Prefecture NUTS III x Degree of Urbanity) the sample is self-weighting (equal probabilities for final units to be selected). As the sampling fractions in all strata of the Geographical Region (NUTS II) are the same (proportional allocation of sample in strata), the sample of all final units is self-weighting with constant sampling fractions.

5.1 The effect of random weighting on the precision of the unemployment rate

Tables 1.1 and 1.2 in the annex 7.1

The data of the tables 1.1 and 1.2 depict the relative bias, the bias ratio of the unemployment rates by geography, urbanity and classes.

At the total country, the relative bias is negative (-2.78%), which means that the total unemployment rate calculated using the initial weights is less than the respective value calculated applying the final weights by 2.78%. Although the absolute relative value is small, the absolute bias ratio is significant (=1.13). This value is so high that it exceeds the value of the sampling error of the unemployment rate calculated applying the initial weights. Thus, the produced bias on the results due to non-response and coverage errors is high and it is required to be removed or reduced using additional random weights. So, the extra weighting process is completely justified.

Concerning the geographical regions and the degrees of urbanity, except for Attica and Crete, the relative biases are negative and in most of the regions (with exceptions of Islands of Southern Aegean and Crete) the absolute biases are sizeable (mot less than 0.1). Thus, the bias persists on almost regions and the extra weighting is fully justified improving the precision of the results.

Regarding the bias on classes (sexes, age groups, education levels), in most of the classes the relative bias is negative and the absolute bias is significant (especially in females in which the value 1.12 exceeds the value of the sampling error of the unemployment rate calculated with the initial weights). Thus, the bias dominates on the results of all classes and the extra weighting is required for improving the accuracy of the results.

Tables 3.1 and 3.2 in the annex 7.1

The data of the tables 1.3 and 1.4 depict the coefficient of variation and the effect of random weighting on the sampling errors and the $deft$ of the unemployment rates by geography, urbanity and classes.

At the total country, the effect of random weighting inflates the sampling error and the $deft$ of the unemployment rate. These inflations approximately coincide (+2.4% and +2.1%, respectively).

Concerning the geography (regions and degrees of urbanity), apart from Western Greece and Crete, the effects of random weighting inflate the sampling errors and the $deft$ of the unemployment rates. The average increases of the sampling errors and the design effect are 5.1% and 2.4%, respectively. In the Western Greece and Crete the random weighting improved the precision of the survey estimates, but these improvements are not shown in the values of the D_{wrand}. This may occur for two reasons: (a) when the survey characteristics have sizeable correlation with the post-stratification variables used for adjusting the sample weights to conform to population totals on a cell-by-cell basis and (b) when the post-strata used for adjusting the sample weights "cut across" the sample clusters making the effect of clustering and the variance smaller, because the clusters with post-stratified units tend to be smaller in size than the clusters when the post-stratification has not been applied.

Concerning the classes, in most of the cases the effects of random weighting inflate the sampling errors and the *deft* of the unemployment rates, because the random weights add further variability, which adversely affect the precision of the survey estimates that are unrelated to the population variables employed in the adjustment post-stratification process. As for the Greek Labour Force, the age groups are used as population variables in the weighting adjustment, it is expected (it is not surprising) in some age groups (as 20-29, 30-44 and 45-64), the random weighting to improve the effect of the survey estimates. Additionally, the classes tend to "cut across" the primary selections (clusters) more or less evenly or randomly making the effect of clustering and the variance smaller.

5.2 Analysis of the design effect decompositions (two-factor model)

Table 2 in the annex 7. 2

The data of the table 2 depict the design effect $deft^2$ of unemployment rate, the weighting component D^2_{wfin}, the average sampled clusters $\overline{m} = \frac{\sum b_i^2}{b_i}$ and the ρ by regions and urbanity. The Kish's two-factor model has been applied and the weighting component D^2_{wfin} has been applied using Kish's formula.

Design effect

At the total country $deft^2$=1.79. However, the $deft^2$ averaged over all regions is 1.58, without considerable variation around the mean, $cv = 2.9\%$ .The average design effect across regions is less than the design effect at the total country, because at regions the selected elementary samples are self-weighting, and consequently only the random weighting influences the weighting component D^2_{wfin}. As the disproportionate allocation of the sample in the regions is applied, the total sample departs from the self-weighting design and the samples at the total country level are subject to larger weights (variations in selection rates across regions). Thus, the value $deft^2$ at the total country is greater than the respective average $deft^2$ across regions.

Except for Epirus, the design effects of regions lie in the relatively narrow range of 1.13-2.33 and this is related to the high degree of the standardization of the survey sampling procedures and designs across regions. In the extreme case of Epirus, $deft^2$ is less than 1. This occurs when the variance between clusters (city blocks) of the unemployment rate is small and the correlation between persons within a block is negative. This is evident, as the effect of clustering and stratification is equal to 0.73 (<1).

The pattern of $deft^2$ for the degrees of urbanity shows that the urban design effect (=1.71) is somewhat smaller than the semi-urban (=1.81) and the rural (=1.87) design effects. This is not surprising, because in rural areas compared to urban, the unemployment and employment tend to be more dependent on the local conditions of the Labour Market, and hence more homogeneity dominates in rural areas making the effect of clustering be somewhat larger.

Weighting effect

The value of the D^2_{wfin}, for the total sample, is equal to 1.23 (increase of the $deft^2$ due to weighting by 23%). This means that the variations in the weights are significant because the estimator applied for the estimations of the results at the total country is not self-weighting.

Except for Attica and Central Macedonia, in which the samples are not self-weighting, the values of D^2_{wfin} are less than the value 1.1 and they lie in the narrow range of 1.04-1.08. However, it is important that we found significant D^2_{wfin} for Attica region, implying sizeable variations in the weights not only across but also within the domains. Really, in the Attica the sampling fraction f for the Greater Athens is $f = 0.48\%$ and for the rest Attica (excl. Greater Athens) is $f = 1.29\%$ (the variation of sampling fractions around the mean is equal to $cv = 64{,}7\%$). In the Central Macedonia, the sampling fraction f for the Greater Thessaloniki is $f = 0.98\%$ and for the rest Central Macedonia (excl. Greater Thessaloniki it is $f = 0{,}57\%$ (the variation of sampling fractions around the mean is equal to $cv = 37.47\%$). As the cv (coefficient of variation of the mean) for Central Macedonia is much less than the respective value of cv for Attica, the value of D^2_{wfin} for Central Macedonia (=1.10) is much less than the respective value for Attica.

Concerning the degrees of urbanity, we found significant D^2_{wfin} implying sizeable variations of the weights within each separate degree of urbanity. This is expected because the degrees of urbanity cut across all regions in which the sampling fractions are deferent. The values of D^2_{wfin} for urban, semi-urban and rural areas are 1.22, 1.15, 1.14, respectively, with variations of sampling fractions around the mean equal to cv= 42.4%, cv=36.3% and cv=33.1%. It is evident that, the more fluctuations in the sampling fractions and the weights at domains the more increments in the values of D^2_{wfin} are appeared.

Intra-class correlation

At the total country, the intra-class correlation ρ is equal to 1.4%. The relative small value ($< 5\%$) indicates that the precision of the estimated unemployment rate is not influenced by the homogeneity of the population characteristics in the city blocks.

Except for Epirus, in which $\rho = -0.7\%$, in the geographical regions the values of ρ belong to the range of 0.3%-3.9%. Although the range of values is not large, the variation of the intra-class correlations across regions is large (cv=66.5%). This large variation reflects mainly the differences in population characteristics and Labour Market conditions in the regions and not any problems in the sample design in the regions (especially with the number the sample of clusters), because the variation of the average cluster sizes is not large (cv=12.5%). As in all geographical regions $|\rho| \leq 5\%$, the size of clusters (city-blocks) is not considered large for creating homogeneity among the elementary sampling units, and as a result clusters of the survey are efficient sampling units.

In Epirus in which the value of ρ is negative, the correlation between persons within a block is negative. In this instance, if we observe a person in a city block to reply unemployment, then it is more likely that the next person observed will reply somehing else (employment, not active), if he/she comes from the same block than if he comes from any other block.

Concerning the degrees of urbanity, the urban intra-class correlation (=1.35%) is less than the semi-urban (=1.9%) and rural (=2.1%). This has to do with the higher homogeneity of individual's behaviour on the labour force variables compared to the urban areas. The local conditions of the rural areas influence the individual's behaviour in a higher degree than in the urban areas.

5.3 Analysis of the design effect decompositions (three-factor model)

Table 3 in the annex 7.3

The data of the table 3 depict by regions the design effect $deft^2$ of unemployment rate, the three components of the design effect (stratification, weighting and clustering effect) as well the intra-class correlation calculated through the three and two-factor model in order possible differences to be detected. The three design effect's components have been calculated applying the direct estimation methods.

The design effects at the total country and the regions have the same values independently the applied model. Negligible differences in the values of $deft^2$ may be appeared due to rounding of the values of the design effect's components.

Stratification

At the total country the effect of stratification is $deft_S^2$=0.95. That means that due to stratification the variance of the unemployment rate has been decreased by 5%. The $deft_S^2$ of regions lie in the relatively narrow range of 0.85-1.0 and the $deft_S^2$ averaged over all regions is 0.96, without considerable variation around the mean ($cv = 3.9\%$).This is related to the high degree of the same stratification criteria (degrees of urbanity) applied across the regions.

In Peloponnesus and Crete the effect of the stratification is almost negligible (=0.99). In the extreme case of Attica the $deft_S^2$=1. That means no stratification effect exists, because the urbanity criteria used as stratification variables may not create homogeneous strata in Attica. Of course the urbanization strata are used not only for reducing the sampling error, but also for making the sample more representative as the sample is selected in each separate stratum.

Intra-class correlation

Calculation the intra-class correlations through (a) the application of the three-factor model with direct estimates of the design effect's components and (b) the two-factor model with Kish's formula D_{wfin}^2 for the estimation of the effect of weighting, we observe that the differences of the intra-class correlations values produced through these two different methods are not important (approximately negligible in some regions). This may occurs because Kish's formula serves as conservative value for the actual design effect (Gabler, Haeder, and Lahiri, 1999).

In Peloponnesus the application of the three-factor model produces ρ=0. In this instance, if we observe in a city block one unemployed person, then it is just as likely that the next person observed will be employed or not active person whether he comes from the same block or from any other block.

6. Conclusions

- The non-response and the coverage errors create bias, the effect of which is unknown. Thus, weighting adjustments attempt to reduce the bias that non-response and coverage errors may cause in survey estimates. The weights for non-response and population adjustments are random variables provoking increases in the sampling errors of the survey estimates. These increases tend to persist undiminished for most subclasses and for all statistics. The indicators quantifying the effect of random weighting on sampling errors and design effects may be used as quality indicators assessing the data collection process and the weighting process targeting to remove the bias due to non-response and coverage problems.

- The total weighting (random weighting and weighting due to disproportionate allocation, unequal selection probabilities) usually increases the variances. The percentages of these increases may be used as quality process indicators for assessing (a) the quality of the sample design and (b) the effect of random weighting on the results.

- The intra-class correlation measures the effect of clustering on the sampling errors and it may be used as a design process indicator assessing if the homogeneity of the sampling elementary units in clusters influences the precision of the results. Additionally, this indicator evaluates the efficiency of the cluster sizes, and shows if the elementary final sampling units have been spread in an efficient number of sample clusters. The importance of the intra-class correlation as an indicator is significant, because the intra-class correlation in combination with the average cluster sizes and the cluster effect links the sample design with the sampling error.

- Both Kish's two-factor and three-factor model produce approximately the same values of the intra-class correlations
- The stratification effect $deft_S^2$ is an important indicator assessing the gains in precisions due to stratification of surveyed units.

7. Annexes

Annex 7.1

Table 1.1 Relative bias, bias ratio and absolute bias by regions and urbanity.

Regions, Urbanity	Relative bias (%)	Absolute relative bias (%)	Bias ratio	Absolute bias ratio
Total Country	-2,78	2,78	-1,13	1,13
Eastern Macedonia &Thrace	-4,71	4,71	-0,52	0,52
Central Macedonia	-3,56	3,56	-0,60	0,60
Western Macedonia	-2,50	2,50	-0,20	0,20
Epirus	-6,34	6,34	-0,74	0,74
Thessaly	-4,48	4,48	-0,44	0,44
Ionian Islands	-10,17	10.17	-0,56	0,56
Western Greece	-1,28	1,28	-0,17	0,17
Central Greece	-8,28	8,28	-0,88	0,88
Attica	1,30	1,30	0,30	0,30
Peloponnesus	-9,28	9,28	-1,07	1,07
Islands of Northern Aegean	-11,64	11,64	-0,77	0,77

Regions, Urbanity	Relative bias (%)	Absolute relative bias (%)	Bias ratio	Absolute bias ratio
Islands of Southern Aegean	-0,31	0,31	-0,02	0,02
Crete	0,45	0,45	0,04	0,04
Urbanity				
Urban areas	-1,10	1,10	-0,38	0,38
Semi-urban areas	-6,28	6,28	-0,93	0,93
Rural areas	-6,87	6,87	-1,03	1,03

Table 1.2 Relative bias, bias ratio and absolute bias by class (sexes, age groups, education levels) at the total country

Sexes, Age groups and Education levels	Relative bias (%)	Absolute relative bias (%)	Bias ratio	Absolute bias ratio
Males	-3,55	3,55	-0,87	0,87
Females	-3,51	3,51	-1,27	1,27
15-19	-3,27	3,27	-0,31	0,31
20-24	-1,67	1,67	-0,33	0,33
25-29	0,33	0,33	0,07	0,07
30-44	0,34	0,34	0,09	0,09
45-64	0,54	0,54	0,10	0,10
65+	-12,31	12,31	-0,31	0,31
1 Received a post-graduate qualification	-12,81	12,81	-0,61	0,61
2 Received a university degree	-1,83	1,83	-0,27	0,27
4 Received a third-level technical - vocational inst. Degree	-0,59	0,59	-0,12	0,12
5 Completed secondary level education	-1,82	1,82	-0,50	0,50
6 Completed the third stage of 6-year secondary education	-2,83	2,83	-0,45	0,45
7 Completed primary education	-5,31	5,31	-0,90	0,90
8 Have not completed primary education	-6,90	6,90	-0,19	0,19
9 Attended no school at all	3,13	3,13	0,08	0,08

Table 1.3: Coefficient of variation (CV), percentage of the effect on sampling error and on square root of the design effect (*deft*) of the unemployment rate due to random weighting by regions and urbanity.

Regions, Urbanity	**CV (%)**	$(ste_{wrand}-1)$ %	$(D_{wrand}-1)$ **%**
Total Country	2,4	2,4	2,1
Eastern Macedonia &Thrace	9,2	6,9	1,9
Central Macedonia	5,8	0,2	1,0
Western Macedonia	12,4	2,6	2,1
Epirus	8,3	3,4	3,5
Thessaly	10,3	4,7	2,4
Ionian Islands	17,7	7,6	3,8
Western Greece	7,5	-1,1	2,2
Central Greece	9,3	7,3	3,5
Attica	4,6	2,3	4,2
Peloponnesus	8,5	7,0	1,9
Islands of Northern Aegean	15,8	16,6	2,8
Islands of Southern Aegean	18,0	3,7	2,6
Crete	11,8	-0,7	2,0
Urbanity			
Urban areas	2,9	1,4	1,9
Semi-urban areas	6,8	6,9	1,3
Rural areas	6,4	2,7	1,0

Table 1.4: Coefficient of variation (CV), percentage of the effect on sampling error and on square root of the design effect (*deft*) of the unemployment rate due to random weighting by classes (sexes, age groups, education levels) at the total Country.

Sexes, Age groups and Education levels	**CV (%)**	$(ste_{wrand}-1)$ %	$(D_{wrand}-1)$ **%**
Males	4,0	2,5	2,2
Female	2,8	3,0	2,0
15-19	10,8	5,6	0,0
20-24	5,0	0,1	0,3
25-29	4,4	-0,5	-0,5
30-44	3,7	-0,2	0,9
45-64	5,4	-0,4	2,5
65+	40,0	12,1	0,7
1 Received a post-graduate qualification	21,0	11,9	2,1
2 Received a university degree	6,8	1,3	2,5
4 Received a third-level technical - vocational inst. Degree	5,0	0,7	1,5

Sexes, Age groups and Education levels	CV (%)	(ste_{wrand} -1) %	($D_{wrand}-1$) %
5 Completed secondary level education	3,7	3,4	1,2
6 Completed the third stage of 6-year secondary education	6,3	3,2	0,3
7 Completed primary education	5,8	3,0	1,2
8 Have not completed primary education	33,8	2,0	0,2
9 Attended no school at all	38,2	-7,5	1,5

Annex 7.2.

Table 2 Design effect of the unemployment rate, design effect's components, average sampled clusters ($\overline{m}$) and intra-class correlation (ρ %) by regions and urbanity (two-factor decomposition model).

Regions, Urbanity	$deft^2$ of unemployment rate	D^2_{wfin} Kish's formula	Stratification and clustering effect	$\overline{m}$	ρ (%)
Total Country	1,79	1,23	1,45	33	1,4
Eastern Macedonia &Thrace	1,62	1,04	1,56	37	1,6
Central Macedonia	1,91	1,10	1,74	31	2,5
Western Macedonia	1,69	1,04	1,62	28	2,3
Epirus	0,78	1,07	0,73	39	-0,7
Thessaly	1,83	1,05	1,75	33	2,3
Ionian Islands	1,99	1,08	1.85	25	3,6
Western Greece	1,20	1,04	1,15	38	0,4
Central Greece	1,57	1,07	1,46	29	1,6
Attica	2,14	1,22	1,75	34	2,3
Peloponnesus	1,13	1,04	1,09	35	0,3
Islands of Northern Aegean	1,28	1,06	1,21	29	0,8
Islands of Southern Aegean	2,33	1,05	2,21	32	3,9
Crete	1,19	1,04	1,14	37	0,4
Urbanity					
Urban areas	1,77	1,22	1,45	35	1,3
Semi-urban areas	1,81	1,15	1,58	32	1,9
Rural areas	1,87	1,14	1,65	32	2,1

Annex 7.3.

Table 3 Design effect of the unemployment rate, the design effect's components (three factor model) and the intra-class correlations by regions.

Regions	Overall design effect $deft^2$	Design effect components			ρ (%) for three factor model	ρ (% for two factor model
		Stratification effect $deft_S^2$	Weighting effect $deft_W^2$	Clustering effect $deft_C^2$		
Total Country	1,81	0,95	1,15	1,66	2,0	1,4
Eastern Macedonia &Thrace	1,63	0,85	1,14	1,68	1,9	1,6
Central Macedonia	1,92	0,96	1,15	1,74	2,5	2,5
Western Macedonia	1,69	0,97	1,05	1,66	2,5	2,3
Epirus	0,78	0,97	1,07	0,75	-0,7	-0,7
Thessaly	1,83	0,96	1,10	1,73	2,3	2,3
Ionian Islands	1,98	0,96	1,16	1,78	3,3	3,6
Western Greece	1,20	0,93	0,98	1,32	0,9	0,4
Central Greece	1,57	0,96	1,15	1,42	1,5	1,6
Attica	2,14	1,00	1,12	1,91	2,8	2,3
Peloponnesus	1,13	0,99	1,14	1,00	0,0	0,3
Islands of Northern Aegean	1,28	0,98	1,36	0,96	-0,1	0,8
Islands of Southern Aegean	2,33	0,96	1,08	2,25	4,1	3,9
Crete	1,19	0,99	0,99	1,21	0,6	0,4

Annex 7.4.

Population under survey

For investigating the size and the composition of employment and unemployment all individuals aged 15 or over years constitute the survey population, with the exception of those serving their military service and those imprisoned.

Individuals permanently reside in collective houses (as hospitals, hotels, asylums, houses of old people, orphanages etc) are uncovered in the survey. These individuals are -as a rule- members of institutional households. If however we subtract from this population the conscripts and the imprisoned, the actual percentage not covered by the survey procedure, accounts for 2% of the total population, and in its major part concerns economically non-active persons.

Sample design

The two-stage area sampling was adopted for the survey. The primary units are the areas (one or more unified blocks), and the ultimate sampling units selected in each sampling area are the households. In each Department (NUTS III), the stratification of primary units was conducted by allocating the Municipalities and Communes according to the degree of urbanization (urban, semi-urban, and rural regions). Except for the two Major City Agglomerations (Athens and Thessaloniki) the produced strata according to the degree of urbanization are:

Urban	Stratum	1	Agglomerations and Municipalities with 10.000 inhabitants or more
Semi-urban	"	2	Municipalities and Communes with 2.000 to 9.999 inhabitants
Rural	"	3	Communes up to 1.999 inhabitants

The Greater Athens Area was divided into 31 strata of about equal size (equal number of households) on the basis of the lists of city blocks of the Municipalities that constitute it and taking into consideration socio-economic criteria. Similarly, the Greater Thessaloniki Area was divided into 9 equally sized strata. The two Major City Agglomerations account for 40% of total population and for even larger percentages in certain socio-economic variables.

Sampling fraction

In each quarter of the year, a sample of approximately 31.217 households is surveyed (sampling fraction≅0,85%). Furthermore, between consecutive quarters, partial sample overlapping exists with rate 5/6. This ensures better results in time comparisons.

The sampling fractions and the sample size by geographical regions (NUTS II) are depicted in the following table.

Geographical Regions (NUTS)(*)	**Sampling fraction (%)**	**Sample size (households)**
Eastern Macedonia &Thrace	0,98	2.013
Central Macedonia (excl. Greater Thessaloniki)	0,57	1.891
Western Macedonia	1,04	968
Epirus	2,14	2.408
Thessaly	0,77	1.835
Ionian Islands	1,08	768
Western Greece	0,93	2.030
Central Greece	1,18	2.151
Attica (excl. Greater Athens)	1,29	2.120
Peloponnesus	1,15	2.243
Islands of Northern Aegean	1,27	932
Islands of Southern Aegean	0,90	895

Geographical Regions (NUTS)[(*)]	Sampling fraction (%)	Sample size (households)
Crete	1,19	2.376
Greater Athens	0,48	5.726
Greater Thessaloniki	0,98	2.861
Whole Country	**0,85**	**31.217**

(*) *The geographical regions (NUTS II) in Greece are 13 in number. However, throughout this study the 2nd geographical region (Central Macedonia) was considered without Greater Thessaloniki and the 9th geographical region (Attica) without the Greater Athens area, while either of these two major agglomerations was treated as a geographical region.*

The calculations of the fractions and sample sizes depicted in the above table were based on data coming from the last general population census of the year 2001.

Sample selection

1st stage of sampling

In the this stage, from any ultimate stratum (crossing of Department with the degree of urbanization), say stratum h, n_h primary units are drawn (where the number n_h of draws is approximately proportional to the population size X_h of the stratum (number of households in the last population census of the year 2001).

Each area unit (primary unit) of the stratum has a probability of being selected proportional to its size. So, if X_{hi} be the number of households-according to the 2001 population census- of the unit in the sample of order i, then the probability of being drawn was:

$$P_{hi} = \frac{X_{hi}}{X_h}$$

Within each stratum the number n_h of draws is taken to be multiple of the number 6, because every quarter the sample of households changes in one sixth (1/6) of primary units

The total number of the primary sampling units is 2.640 areas.

2nd stage of sampling

In this stage from each primary sampling unit (selected area) the sample of ultimate units (households) is selected. Actually, in the second stage we draw a sample of dwellings. However, in most cases, there corresponds one household to each dwelling. If in the selected dwelling live more than one household, all of them are interviewed.

Let M_{hi} be the number of households during the survey period in the i selected area of the stratum h. Out of them a systematic sample of m_{hi} households is selected with equal probabilities. Each of m_{hi} households has the same chance to be included in the survey, equal to: $\frac{m_{hi}}{M_{hi}}$

In any selected primary unit, the determination of the sample size m_{hi} remains. The total number of households to be interviewed of the n_h selected primary sampling units will be $m_h = \sum_{i=1}^{n_h} m_{hi}$

i.e. finally and out of the whole procedure (1st and 2nd stages) is drawn from the stratum the percentage of households $\frac{m_h}{M_h}$.

In repeated sampling the numerator of this fraction will vary from sample to sample, in other words the fraction $\frac{m_h}{M_h}$ will be a random variable. Within primary sampling unit the calculation of sampling interval $\delta_{hi} = \frac{M_{hi}}{m_{hi}}$ will be carried out, so that the following two desired conditions to be satisfied.

a) The expectation of the fraction $\frac{m_h}{M_h}$ should be the predetermined over sampling fraction $\frac{1}{\lambda}$ in each geographical region (NUTS II): $E\left(\frac{m_h}{M_h}\right) = \frac{1}{\lambda}$

b) The estimator of the stratum total Y_h (for any characteristic) should be self-weighting. In other words, the estimation should be derived as product of the sum of the values of the characteristic over the m_h sample households by the overall raising factor λ, which is separate in each g geographical region.

The conditions (a) and (b) are satisfied when:

$$\frac{1}{n_h} \cdot \frac{1}{P_{hi}} \cdot \frac{M_{hi}}{m_{hi}} = \lambda \quad \Rightarrow$$

$$\frac{1}{n_h} \cdot \frac{1}{P_{hi}} \cdot \delta_{hi} = \lambda \Rightarrow$$

$$\delta_{hi} = \frac{M_{hi}}{m_{hi}} = \lambda \cdot n_h \cdot P_{hi}$$

Weightings

Design factor

The household and individual design weight is defined as the inverse of its probability of selection.

$$\frac{1}{n_h} \cdot \frac{1}{P_{hi}} \cdot \frac{M_{hi}}{m_{hi}} = DW_{hi}$$

M_{hi} = the number of households in the updated sampling frame in ***hi*** area (primary unit).

m_{hi} = the number of selected households in ***hi*** area (primary unit).

n_h = the sample size of primary units in ***h*** stratum.

P_{hi} = the selection probability of ***hi*** primary unit.

Non-response adjustments

Within each primary sampling unit (area), the non-response adjustment of the responding households is carried out by the inverse of the response rate, so as to "make up" for non-responding cases in that area (Primary Sampling Unit).

Adjustment to external population data

This involves the calibration of the personal weights in conjunctions with external sources (Projections to population totals for the year 2007). Thus, it enables the distribution of auxiliary variables on individual level.

The auxiliary variable used at personal level is the distribution of population by age (five years age groups) and sex in each Geographical Region (NUTS II). This adjustment of personal weights is conducted applying the method of cell weighting.

Sampling Errors

Estimation of survey characteristics

Let y_{hij} be the value of the characteristic **y** (of the sampling household of order j in case of a household survey characteristic or for the sampling member of order j in case of a household member survey characteristic, $j = 1,2,...,m_{hi}$) of the hi area. Moreover Y_h stands for the stratum total, which results when adding the characteristic y from all household members included in the stratum h.

The form of the estimator on the basis of the two-stage design is:

$$\hat{Y}_h = \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} \cdot y_{hij}$$

The w_{hij} represents the final weight (extrapolation factor) for the sampling household member hij. The weight w_{hij} is defined as follows:

$w_{hij} = DW_{hi} \cdot t_{hij}$, where $t_{hij} = R_{hi}^{-1} \cdot w'_{hij}$

R_{hi} = Response rate in hi area (primary sampling unit)

w'_{hij} = The weight for the adjustment to external data (level, variables used and sources). This weight was calculated using the cell (or class) weighting method. Evidently, the quantity t_{hij} is random variable (as it was determined after data collection) and it increases the variance of the survey estimates.

For estimating the characteristic $\mathbf{y}$ in country level, all stratum estimates $\hat{Y}_h$ should be added, that is:

$$\hat{Y} = \sum_h \hat{Y}_h$$

In order to estimate variances of the required characteristics, the following steps should be implemented.

(a) For every selected PSU i of the stratum h, we calculate the quantity T_{hi} using the following formulas:

$$T_{hi} = n_h \cdot \sum_{j=1}^{mhi} w_{hij} \cdot y_{hij}$$

(b) Since T_{hi} has been calculated for every PSU i ($i = 1,2,..., n_h$) of the stratum h, then:

$V\left(\hat{Y}_h\right)$ is calculated as (Rao, Wu and Yue, 1992):

$$V\left(\hat{Y}_h\right) = \frac{1}{n_h \cdot (n_h - 1)} \cdot \left[\sum_{i=1}^{n_h} T_{hi}^2 - \frac{1}{n_h} \cdot \left(\sum_{i=1}^{n_h} T_{hi} \right)^2 \right]$$

and

$V\left(\hat{Y}\right)$ (Country level) is calculated by adding $V\left(\hat{Y}_h\right)$ for all strata h, that is:

$$V\left(\hat{Y}\right) = \sum_h V\left(\hat{Y}_h\right)$$

The relative standard error of the estimate $\hat{Y}$ or else its coefficient of variation is defined as:

$$CV\left(\hat{Y}\right) = \frac{\sqrt{V\left(\hat{Y}\right)}}{\hat{Y}} * 100$$

For the estimation of the variance and the coefficient of variation of a ratio $R = \frac{\hat{Y}}{\hat{X}}$ (e.g. unemployment rate) additional steps should be implemented as follows:

(a) The estimation X_h in stratum h of the variable x is calculated using the formula:

$$\hat{X}_h = \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij}$$

while the estimation of the relevant characteristic in country level is calculated by adding all strata estimations, that is:

$$\hat{X}_h = \sum_h \hat{X}_h$$

(b) For every selected PSU i of the stratum h, we calculate the quantity F_{hi} using the following formula:

$$F_{hi} = n_h \cdot \sum_{j=1}^{mhi} w_{hij} x_{hij}$$

(c) Since T_{hi} and F_{hi} have been calculated for every PSU i ($i = 1,2,..., n_h$) of the stratum h, then :

$V\left(\hat{Y}_h\right)$ is calculated as:

$$V\left(\hat{Y}_h\right)=\frac{1}{n_h\cdot(n_h-1)}\cdot\left[\sum_{i=1}^{n_h}T_{hi}^2-\frac{1}{n_h}\cdot\left(\sum_{i=1}^{n_h}T_{hi}\right)^2\right]$$

and

$V\left(\hat{Y}\right)$ (country level) is calculated by adding $V\left(\hat{Y}_h\right)$ for all strata h, that is:

$$V\left(\hat{Y}\right)=\sum_h V\left(\hat{Y}_h\right)$$

Correspondingly, $V\left(\hat{X}_h\right)$ is given by:

$$V\left(\hat{X}_h\right)=\frac{1}{n_h\cdot(n_h-1)}\cdot\left[\sum_{i=1}^{n_h}F_{hi}^2-\frac{1}{n_h}\cdot\left(\sum_{i=1}^{n_h}F_{hi}\right)^2\right]$$

and

$V\left(\hat{X}\right)$ (country level) is calculated by adding $V\left(\hat{X}_h\right)$ for all strata h, that is:

$$V\left(\hat{X}\right)=\sum_h V\left(\hat{X}_h\right)$$

The variance of $\hat{R}$ can be calculated using the formula below

$$V\left(\hat{R}\right)=\frac{V\left(\hat{Y}\right)+\hat{R}^2\cdot V\left(\hat{X}\right)-2\cdot\hat{R}\cdot Cov\left(\hat{Y},\hat{X}\right)}{\hat{X}^2}$$

where

$$Cov\left(\hat{Y}_h,\hat{X}_h\right)=\frac{1}{n_h\cdot(n_h-1)}\cdot\left[\sum_{i=1}^{n_h}T_{hi}\cdot F_{hi}-\frac{1}{n_h}\cdot\left(\sum_{i=1}^{n_h}T_{hi}\right)\left(\sum_{i=1}^{n_h}F_{hi}\right)\right]$$

and

$$Cov(\hat{Y},\hat{X})=\sum_h Cov\left(\hat{Y}_h,\hat{X}_h\right)$$

For an estimate $\hat{R}$, the coefficient of variation is defined as:

$$CV\left(\hat{R}\right)=\frac{\sqrt{V\left(\hat{R}\right)}}{\hat{R}}*100$$

REFERENCES

Deville, J.C. and Sarndal C.E. (1992). Callibration Estimators in Survey Sampling. *JASA, 87, 376-382.*

Deville, J.C. and Sarndal C.E., and Sautory, O. (1993). Generalized Raking Procedures in Survey Sampling. *JASA, 88, 1013-1020.*

Fuller, W.A., McLoughlin, M.M., and Baker, H.D. (1994). Regression Weighting in the Presence of Non-Response with the Application to the 1987-1988 Nationwide Food Consumption Survey. *Survey Methodology, 20,75-85.*

Gabler, S., Haeder, S. and Lahiri, P. (1999). A model based justification of Kish's formula for design effect for weighting and clustering. *Survey Methodology, 25, 105-106.*

Ireland, C.T and Kullback, S. (1968). Contingency Tables With Given Marginals. *Biometrika, 55,179-188.*

Kalton G. (1983). Models in the Practice of Survey Sampling. *Int. Stat. Rev., 51, 175-188.*

Kalton G., Kasprzyk, D. (1986). The Treatment of Missing Survey Data. *Survey Methodology, 12, 1-16.*

Kalton G., Flores-Cervantes, I. (2003). Weighting Methods. *Journal of Official Statistics, 19, 81-97.*

Kalton, G., Brick, M and Le, T. (2005). Estimating components of design effect for use in sample design. *Chapter VI of the Household Sample Surveys in Developing and Transition Countries. United Nations, New York..*

Kish, L. (1965). Survey Sampling. John *Wiley and Sons. New York*.

Kish, L. (1987). Statistical Research Design. *John Wiley and Sons. New York.*

Kish, L. (1987). Weighting in $Deft^2$. *The Survey Statistician. June, 26-30.*

Kish, L. (1990). Questions / Answers (1978-1994). *Paris, INSEE, International Association of Survey Statisticians.*

Kish, L. (1992). Weighting for Unequal P_i. *Journal of Official Statistics, 8, 183-200.*

Little, R.J.A. (1986). Survey Non-response Adjustments for Estimates of Means. *Int. Stat. Rev., 54, 139-157.*

Liu, J. and Aragon, E. (2000). Sub-sampling strategies in longitudinal surveys. *Proceeding of the American Statistical Association, Section of Survey Research Methods, 307-312.*

Nikolaidis, I. (2006). The Effect of Weighting on Sampling errors as a Quality Indicator for Non-response and Frame Errors. *Paper for the meeting of the working group of Eurostat "Quality in Statistics", 27-28 February 2006.*

Park, I., Winglee, M., Clark, J., Rust, K., Sedlak, A., and Morkanstein, D. (2003). Design Effects and Survey Planning. *Proceedings of the American, Section on Survey Research Methods.*

Park, I. (2004). Assessing Complex Sample Designs Via Design Effects Decompositions. *Proceedings of the American, Section on Survey Research Methods.*

Oh, H.L and Scheuren (1987). Modified Raking Ratio Estimation. *Survey Methodology, 13, 209-219.*

Rao, J.N.K, Wu, C.F.J., and Yue, K. (1992). Some Recent Work on Re-sampling Methods for Complex Surveys. *Survey Methodology, 18, 209-217.*

Verma, V. and Le, T. (1996). An Analysis of Sampling Errors for the Demographic and Health Surveys. *Int. Stat. Rev. ,64, 265-294.*

ANNEX 8: Greek survey of goats

The sampling method applied is the single stratified random sampling. The livestock holdings with goats included in the survey are stratified as follows:

- By geographical region - NUTS II
- By size class of the holdings. In each geographical region, the holdings are stratified into L = 10 size classes, according to their size, determined by their number of goats in the updated livestock holding register, as follows.

Size Class	Number of goats
Class 1	1 - 4
Class 2	5 - 9
Class 3	10 - 19
Class 4	20 - 49
Class 5	50 - 99
Class 6	100 - 299
Class 7	300 - 499
Class 8	500 - 699
Class 9	700 - 999
Class 10	1000 +

Livestock holdings with goats belonging to classes 9 - 10 were surveyed exhaustively.

1. Extra variability due to measurement errors

Using available data (number of goats) from two independent sources of information regarding the common sampling holdings of the annual survey of goats and the farm structure survey of the same reference period (1st November 2005) and applying the formulae (4.2.3), (4.2.5) and (4.2.6),the quantity $IS(\widehat{Y})$ is:

$$IS(\widehat{Y}) = \sqrt{\frac{1602610}{2910240714}} \cdot 100 = 2{,}3\% \cdot$$

The value of $IS(\widehat{Y}) = 2{,}3\%$ means, that 2,3% of the total sampling error of the estimated number of goats has been created from the appearing measurement errors. Taking into consideration that the coefficient of variation of the $\widehat{Y}$ (estimated number of goats) is equal to $CV(\widehat{Y})\% = 1{,}1\%$, in the case that there were not measurement errors then the coefficient of variation would be according to formula (4.2.7) as follows:

$CV'(\widehat{Y})\% = 1{,}1\% - 0{,}023 \cdot 1{,}1\%$ or $CV'(\widehat{Y})\% = 1{,}07\%$.

2. Bias due to measurement errors

Possible bias on the results due to measurement errors should be investigated through student t-tests, according to which for two independent observations (sources) of the same respondents the means of the number of goats were compared. The t values were calculated in each stratum (NUTS II x size classes) and in all cases the t values were lower than the critical values with confidence level 95%. As a result, in each stratum the means of goats from the different sources of information did not differ significantly, and the bias due to measurement errors statistically was not important.

ANNEX 9: Complete analysis of assessing the coherence between Greek EU SILC and Greek Structural Earning Survey (SES) results[(1)]

1. EU SILC 2003[(2)] AND SES 2002

1.1. *Statistical Unit*

EU SILC 2003	SES 2002
It is based on private household: means a person living alone or a group of people who live together in the same private household and share expenditures, including the joint provision of the essentials of living and provide information on employees in enterprises independently of its size	Provide information on employees in enterprises with 10 or more employees classified by size and principal activity.

1.2. Classification economic activities

EU SILC 2003	SES 2002
The region in which the local unit is located was classified according to NACE rev.1.1: Statistical classification of economic activities in the European Community, Rev. 1.1. All classification economic activities	The region in which the local unit is located was classified according to NACE rev.1.1: Statistical classification of economic activities in the European Community, Rev. 1.1. covers all sectors in economic activities classified to Sections C to K and M to O of NACE Rev. 1.1 in enterprises with at least 10 employees. C. Mining and quarrying (10-14) D. Manufacturing industry (15-37) E. Electricity, gas and water supply (40,41) F. Construction (45) G. Wholesale and retail trade (50,51,52) H. Hotels and restaurants 55 I. Transport, storage and communication (60-64) J. Financial intermediation (65-67). K. Real estate (70-74)

(1) This example comes from the project "*Study of the impact on comparability of national implementations*"(Prepared by Giorgos Ntouros, Ioannis Nikolalidis, Ilias Lagos, Maria Chaliadaki), NSSG, Piraeus, July 2008.

(2) The income reference period concerns the year preceding the survey year i.e. 2002.

1.3. Size of the enterprise to which the local unit belongs

The size of the enterprise in terms of the number of employees was assigned to one of the following bands:

EU SILC 2003	SES 2002
• 1 — 10 (exact number if between 1 and 10) • 11 (if between 11 and 19 persons) • 12 (if between 20 and 49 persons) • 13 (if 50 persons or more) • 14 do not know but less than 11 persons • 15 do not know but more than 10 persons	• 10 to 49, • 50 to 249, • 250 to 499, • 500 to 999, and • 1 000 or more employees.

It is noted that the first band of EU-SILC (1 — 9 employees) has not been excluded.

1.4. *Employees*

EU SILC covers all employees living in private households and Structural Statistics on Earnings provide information on employees in enterprises with 10 or more employees classified by size and principal activity.

EU SILC 2003	SES 2002
Employees are defined as persons who work for a public or private employer and who receive compensation in the form of wages, salaries, fees, gratuities; non-conscripted members of the armed forces are also included.	Employees are all persons, irrespective of their nationality or the length of their working time in the country, who have a direct employment contract with the enterprise or local unit (whether the agreement is formal or informal) and receive remuneration, irrespective of the type of work performed, the number of hours worked (full-time or part-time) and the duration of the contract (fixed or indefinite). The remuneration of employees can take the form of wages and salaries including bonuses, pay for piecework and shift work, allowances, fees, tips and gratuities, commission and remuneration in kind. The employees to be included in the sample are those who actually received remuneration during the reference month. The definition of employees covers manual and non-manual workers and management personnel in the private and public sectors in economic activities classified to Sections C to K of NACE Rev. 1.1 in enterprises with at least 10 employees

1.5. Occupation in the reference month (ISCO-88 (COM))

EU SILC 2003	SES 2002
ISCO — 88 (COM): International standard classification of occupations (for European purposes), 1988 version at the two-digit level.	The occupation is to be coded according to the International Standard Classification of Occupations, 1988 version (ISCO-88 (COM)) at the two-digit level.

1.6. *Highest successfully completed level of education and training (ISCED 97)*

EU SILC 2003	SES 2002
International Standard Classification of Education, 1997 version (ISCED 97).	International Standard Classification of Education, 1997 version (ISCED 97).

1.7. Contractual working time (full-time or part-time)

EU SILC 2003	SES 2002
The distinction between full-time and part-time work should be made on the basis of a spontaneous answer given by the respondent. It is impossible to establish a more exact distinction between part-time and full-time work, due to variations in working hours between Member States and also between branches of industry. By checking the answer with the number of hours usually worked, it should be possible to detect and even to correct implausible answers, since part-time work will hardly ever exceed 35 hours, while full-time work will usually start at about 30 hours.	*Full-time employees* are those whose normal working hours are the same as the collectively agreed or customary hours worked in the local unit under consideration, even if their contract is for less than one year.

1.8. Net cash or near cash employee income

EU SILC 2003	SES 2002
Cash or near cash employee income is defined as the total remuneration in cash payable income by an employer to an employee income in return for work done by the latter during the income reference period.	Compensation of employees is defined as the total remuneration, in cash or in kind, payable by an employer to an employee in return for work done by the latter during the reference period.

EU-SILC 2003 collects information only on monetary income (and not income in kind) so we only used and compared it.

1.9. Taxes/ Compulsory social contributions

EU SILC 2003	SES 2002
Tax refers to the amount of all taxes on the employee's earnings withheld by the employer for the reference month and paid by the employer to the tax authorities on behalf of the employee.	Tax refers to the amount of all taxes on the employee's earnings withheld by the employer for the reference month and paid by the employer to the tax authorities on behalf of the employee.
Employers' contributions are defined as payments made, during the income reference period, by employers for the benefits of their employees to insurers (social security funds and private funded schemes) covering statutory, conventional or contractual contributions in respect of insurance against social risks.	Employers' contributions refer to the total amount of compulsory social contributions paid by the employer on behalf of the employee to insurance schemes authorities during the reference year.

EU-SILC 2003 collects information both on net income and taxes and social contributions, however taxes and social contributions' information is being considered unreliable.

SES collects information only on gross income. Taxes and social contributions were optional, thus only a few enterprises provided this information.

2. Differences in comparisons for assessing the coherence

2.1. AVERAGE EMPLOYEE INCOME BY ECONOMIC ACTIVITY

After checking the hypothesis that average net employee's income per economic activity does not present statistically significant difference among the two surveys (EU-SILC and SES), see table 1, the conclusion is that:

a) The hypothesis (since $T_{0,05}$=1,96) is not accepted for most of the economic activities, where A<1,96, that is in economic activities C, D, G, H, I, J and K.
b) On the contrary, keeping the same level of significance (=0,05) for economic activities *E. Electricity, gas and water supply* and *F. Construction* average yearly net employee's income presents statistically significant difference among the two surveys.

Table 1: Average employee income by economic activity

Economic activity	EU-SILC	SES	A<1,96
C. Mining and quarrying	11.393,46	12.433,18	0,47
D. Manufacturing industry	11.196,95	11.671,47	0,87

Economic activity	EU-SILC	SES	A<1,96
E. Electricity, gas and water supply	16.661,30	22.704,86	2,42
F. Construction	9.164,31	11.603,20	6,17
G. Wholesale and retail trade	9.938,92	10.426,61	0,99
H. Hotels and restaurants	8.007,43	8.575,63	0,89
I. Transport, storage and communication	14.822,53	14.598,51	0,23
J. Financial intermediation	18.441,36	17.127,12	0,63
K. Real estate	13.337,23	11.610,38	1,20

2.2 Average employee income by gender

After checking the hypothesis that average net employee's income per gender does not present statistically significant difference among the two surveys (EU-SILC and SES), see table 2, the conclusion is that:

a) As far as men are concerned the hypothesis can be accepted.
b) As far as women are concerned total average net income presents statistically significant difference among the two surveys.

Table 2: Average employee income by gender

Gender	EU-SILC	SES	*A<1,96*
Male	12.997,27	13.326,81	1,05
Female	11.112,90	10.308,97	2,47

2.3. Average employee income by age groups

After checking the hypothesis that average net employee's income per age group does not present statistically significant difference among the two surveys (EU-SILC and SES), see table 3, we conclude that we do accept the hypothesis for all age groups.

Table 3: Average employee income by age groups

Age groups	EU-SILC	SES	A<1,96
16-24	6.823,48	6.849,44	0,07
25-34	10.515,31	10.020,99	1,51

Age groups	EU-SILC	SES	A<1,96
35-44	13.423,58	12.970,67	1,02
45-54	14.818,03	15.570,20	1,23
55-74	15.072,60	16.355,95	1,12

2.4. AVERAGE EMPLOYEE INCOME BY EDUCATIONAL LEVEL

After checking the hypothesis that average net employee's income, as far as the educational level having been completed by the employee is concerned, does not present statistically significant difference among the two surveys (EU-SILC and SES), see table 4, we conclude that :

a) We can accept the hypothesis for persons having completed

- Primary education
- Upper secondary education and
- First or second stage of tertiary education

b) On the contrary, average net income presents statistically significant difference among the two surveys for persons having completed

- Lower secondary education and
- Post secondary non tertiary education

Table 4: Average employee income by education level

Education level	EU-SILC	SES	A<1,96
Never attended any level of education - Primary education	10.024,87	10.389,43	0,67
Lower secondary education	11.036,52	10.067,56	3,53
Upper secondary education	10.416,37	11.189,38	1,05
Post secondary non tertiary education	16.155,16	11.777,95	8,80
First and second stage of tertiary education	26.069,06	17.346,96	1,56

2.5. AVERAGE EMPLOYEE INCOME BY OCCUPATION

After checking the hypothesis that average net employee's income, as far as occupation is concerned, does not present statistically significant difference among the two surveys (EU-SILC and SES) and with significance level a=0,05 we conclude that we can accept the hypothesis for all occupations except for Extraction and building trades workers, other craft and related trades workers. Metal machinery and related trades workers. Precision, handicraft, printing and related trades workers for which average net income, among the two surveys, see table 5, presents statistically significant difference.

Table 5: Average employee income by occupation

Occupation	EU-SILC	SES	A<1,96
1. Legislators and senior officials-Corporate managers	24.608,76	22.061,14	0,76
2. Physical, mathematical, engineering science and other professionals	16.925,42	17.977,94	1,38
3. Physical, engineering science associate professionals and other associate professionals	13.815,13	14.369,75	0,73
4. Office clerks and customer services clerks	12.218,16	11.577,96	1,28
5. Personal and protective services workers, models, salespersons and demonstrators miscellaneous	9.654,90	8.944,06	1,64
6. Skilled agricultural and fishery workers	7.834,72	7.974,09	0,13
7. Extraction and building trades workers, other craft and related trades workers. Metal machinery and related trades workers. Precision, handicraft, printing and related trades workers	10.229,80	11.914,54	4,17
8. Stationary-plant and related operators, drivers and mobile plant operators, machine operators and assemblers	10.872,06	11.901,71	1,66
9. Sales and services elementary occupations, agricultural, fishery and related labourers in mining, construction, manufacturing and transport	8.092,91	8.516,41	0,84

2.6. NUMBER OF EMPLOYEES BY INCOME CLASSES AND GENDER

Applying X^2_v distribution tables with (κ-1).(λ-1) = (5-1).(2-1) = ν = 4 degrees of freedom and probability 95%, that is a=0,05, for both variables (gender / income category), in both surveys (EU-SILC and SES), see table 6 and 7, it arises that income category significantly depends on gender. The dependence degree in the SES is larger, that is 22,84% in relation to that of EU-SILC being 12,80%.

The dependence degree is calculated applying the formula:

$$C = \sqrt{\frac{x_v^2}{N + x_v^2}}$$

where N is the number of the surveyed units.
Note: if C=0 then there is completely independency.

Table 6: Number of employees by income classes and gender, EU- SILC

Income classes	**Gender**		
	Male	Female	Total
0-4.999	99.825	96.461	196.286
5.000-9.999	452.841	340.859	793.700
10.000-19.999	745.569	439.985	1.185.554
20.000-29.999	112.198	36.656	148.854
30.000+	48.546	10.401	58.947
Total	**1.458.979**	**924.362**	**2.383.341**

Table 7: Number of employees by income categories and gender, SES

Income classes	**Gender**		
	Male	Female	Total
0-4.999	39.192	33.567	72.759
5.000-9.999	170.769	162.852	333.621
10.000-19.999	296.984	133.463	430.447
20.000-29.999	64.592	10.955	75.547
30.000+	15.617	1.710	17.327
Total	**587.154**	**342.547**	**929.701**

2.7. NUMBER OF EMPLOYEES BY INCOME CLASSES AND ECONOMIC ACTIVITY

Applying X^2_ν distribution tables with $(\kappa\text{-}1).(\lambda\text{-}1) = (9\text{-}1).(5\text{-}1) = \nu = 32$ degrees of freedom and probability 95%, that is a=0,05, for both variables (income category/economic activity), in both surveys (EU-SILC and SES), see tables 8 and 9, it arises that income category depends on economic activity. The dependence degree of the two variables is strong for both surveys, 37,89% for EU-SILC and 44,25% for SES.

What's left to be studied is if in both surveys, the same economic activities depend on the same income categories. This will be done using corresponded factor analysis and Hierarchical cluster analysis and the results will be presented at factor level.

Table 8: Number of employees by income classes and economic activity, EU-SILC

Economic activity	**Income classes**					
	0-4.999	5.000-9.999	10.000-19.999	20.000-29.999	30.000+	Total
C. Mining and quarrying	1.158	5.699	11.112	468	0	18.437
D. Manufacturing industry	23.731	176.945	158.082	16.103	10.967	385.828
E. Electricity, gas and water supply	0	2.834	22.853	7.077	1.111	33.875
F. Construction	22.976	98.559	72.983	1.758	755	197.031
G. Wholesale and retail trade	27.968	177.125	109.719	13.422	3.478	331.712
H. Hotels and restaurants	37.350	66.237	34.981	671	942	140.181
I. Transport, storage and communication	6.600	49.258	103.344	24.881	9.714	193.797
J. Financial intermediation	3.412	12.278	37.964	12.104	8.514	74.272
K. Real estate	11.682	43.127	39.246	10.344	7.191	111.590
Total	**134.877**	**632.062**	**590.284**	**86.828**	**42.672**	**1.486.723**

Table 9: Number of employees by income classes and economic activity, SES

Economic activity	Income classes					
	0-4.999	5.000-9.999	10.000-19.999	20.000-29.999	30.000+	Total
C. Mining and quarrying	298	1.873	3.838	386	60	6.455
D. Manufacturing industry	18.902	110.271	143.000	17.287	3.497	292.957
E. Electricity, gas and water supply	83	282	15.273	15.485	5.155	36.278
F. Construction	4.687	18.364	23.586	3.686	440	50.763
G. Wholesale and retail trade	18.517	105.063	81.844	7.204	2.039	214.667
H. Hotels and restaurants	15.811	50.545	23.104	1.104	280	90.844
I. Transport, storage and communication	5.146	15.567	72.410	11.307	1.567	105.997
J. Financial intermediation	1.183	4.853	40.839	13.982	2.542	63.399
K. Real estate	8.132	26.802	26.552	5.106	1.747	68.339
Total	**72.759**	**333.620**	**430.446**	**75.547**	**17.327**	**929.699**

3. Coherence of the relationships of the survey characteristics

3.1. Factor Level: Correlation of NACE and income classes

Income has been split in five classes, as following:

a) The first class includes persons having yearly net income from 0 to 4.999 euros.
b) The second class includes persons having yearly net income from 5.000 to 9.999 euros.
c) The third class includes persons having yearly net income from 10.000 to 19.999 euros.
d) The fourth class includes persons having yearly net income from 20.000 to 29.999 euros and
e) The fifth class includes persons having yearly net income more than 30.000 euros.

The horizontal axe confronts employees having low income with employees having high income. Moving from left to right employees' income increases. The vertical axe confronts Mining and quarrying Manufacturing Industry (D and C) and Construction (F) plus Wholesale and retail trade (G) with the rest of economic activities (E, H, I, J, K).

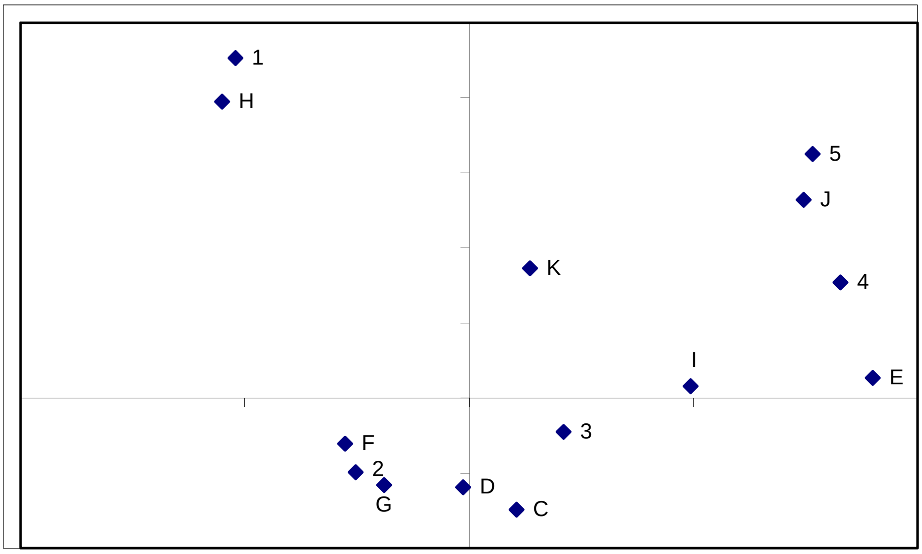
As far as the EU-SILC survey is concerned, according to their income and their economic activity employees have been split into four clusters (see graph 1), as following:

a) The first cluster consists of branch H with the employees reaching income of the first income class
b) The second cluster includes branches F and G with employees reaching income of the second income class
c) The third cluster includes branches D, C and K with the majority of employees reaching income of the third income class and
d) The fourth cluster includes branches I, E and J with employees with high income. Branch J is mostly correlated to income class 5, though.

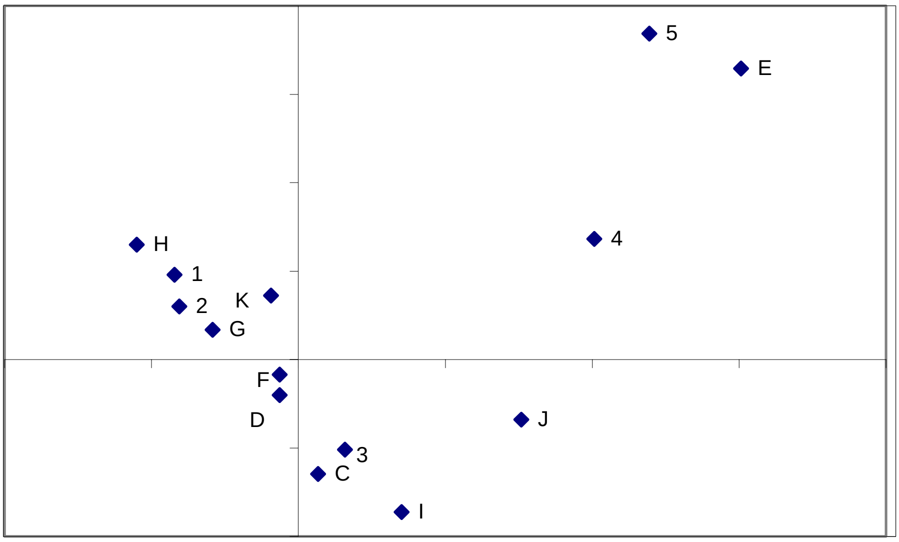
As far as the SES survey is concerned, according to their income and their economic activity employees have been split into four clusters (see graph 2), as following:

a) The first cluster consists of branch H with the employees reaching income of the first income class
b) The second cluster includes branches K and G *-and not F and G like the EU-SILC-* with employees reaching income of the second income class
c) The third cluster includes branches F, D and C – *and not K included in the EU-SILC-* with the majority of employees reaching income of the third income class and finally
d) The fourth cluster includes branches I, E and J with employees with high income. However, it should be noted that branches I and J of EU-SILC correspond to definitely higher incomes in relation to those of the SES. Additionally, it should also be noted that in the EU-SILC branch J is strongly correlated to income class 5, while in the SES branch E is strongly correlated to income class 5.

Graph 1. EU-SILC: NACE X Income classes



Graph 2. SES: NACE X Income classes



3.2 FACTOR LEVEL: CORRELATION OF GENDER, AGE AND INCOME CLASSES

Income has been split in five classes as following:

a) The first class includes persons having yearly net income from 0 to 4.999 euros.
b) The second class includes persons having yearly net income from 5.000 to 9.999 euros.
c) The third class includes persons having yearly net income from 10.000 to 19.999 euros.
d) The fourth class includes persons having yearly net income from 20.000 to 29.999 euros and
e) The fifth class includes persons having yearly net income more than 30.000 euros.

The horizontal axe confronts employees having low income with employees having high income. Moving from left to right employees' income increases. The vertical axe confronts age group 25-44 with the rest age groups for men and women.

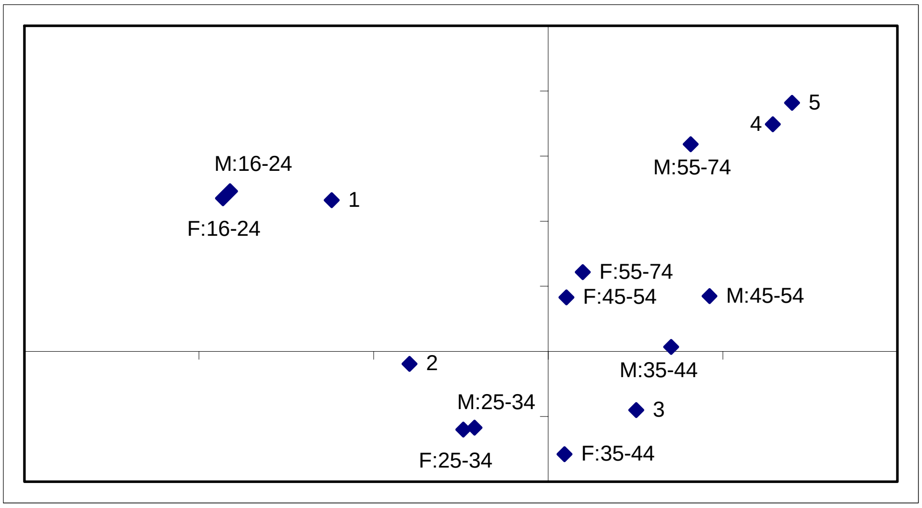
As far as the EU-SILC survey is concerned, according to their income four groups of persons are presented (see graph 3), as following:

a) The first group includes employees –men and women- aged 16-24 reaching income of the first income class
b) The second group includes employees –men and women- aged 25-34 reaching income of the second income class
c) The third cluster includes only women- aged 35-74 mostly reaching income of the third income class. However, women aged 35-44 mostly converge to income class 3.
d) Finally, the fourth group includes men aged 35-74 who mostly converge to income classes 4 and 5. Age group 55-75 mostly converges to income class 5.

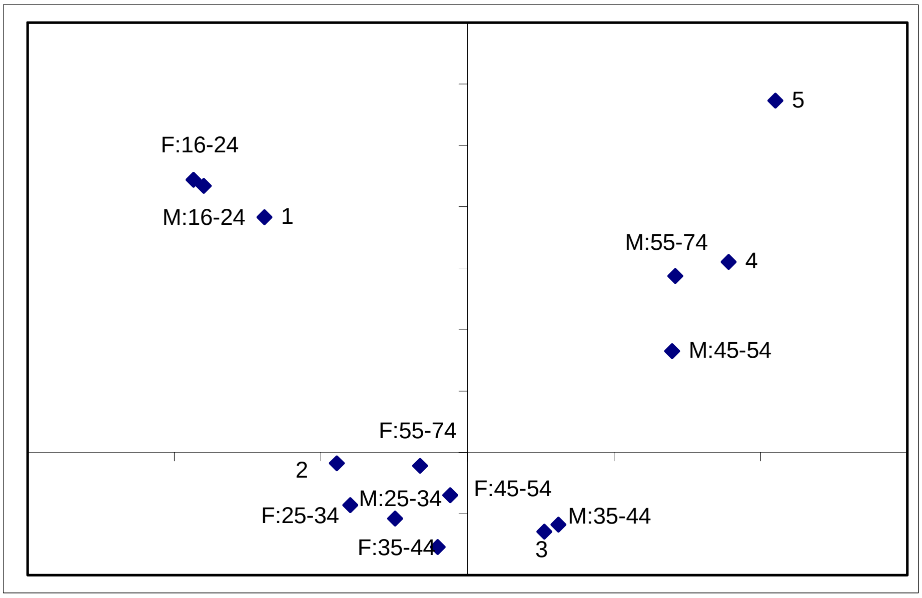
As far as the SES survey is concerned (see graph 4):

a) The first group coincides exactly with the relative first group of the EU-SILC.
b) The second group, though, includes women aged 25-74 and men aged 25-34 reaching income of the second income class.
c) The third group includes men aged 35-44 reaching income of the third income class and
d) The fourth group includes men 45-74 reaching income of the fourth and fifth income class.

Graph 3: EU-SILC: Gender (M, and F) +Age X Income classes (1-5)



***Graph 4*: SES: Gender + Age X Income classes**



3.3 FACTOR LEVEL: RELATIONSHIP BETWEEN PROFESSIONS (ONE DIGIT ISCO CODES) AND INCOME CLASSES

Income has been split in five classes, as following:

a) The first class includes persons having yearly net income from 0 to 4.999 euros.
b) The second class includes persons having yearly net income from 5.000 to 9.999 euros.
c) The third class includes persons having yearly net income from 10.000 to 19.999 euros.
d) The fourth class includes persons having yearly net income from 20.000 to 29.999 euros and
e) The fifth class includes persons having yearly net income more than 30.000 euros.

As far as the EU-SILC survey is concerned, the horizontal axe confronts employees having low income with employees having high income. Moving from left to right employees' income increases. The vertical axe confronts occupations with codes 2, 3, 4 and 8 with the rest occupations (see graph 5).

Employees according to their occupation and their income class are split in 3 big clusters.

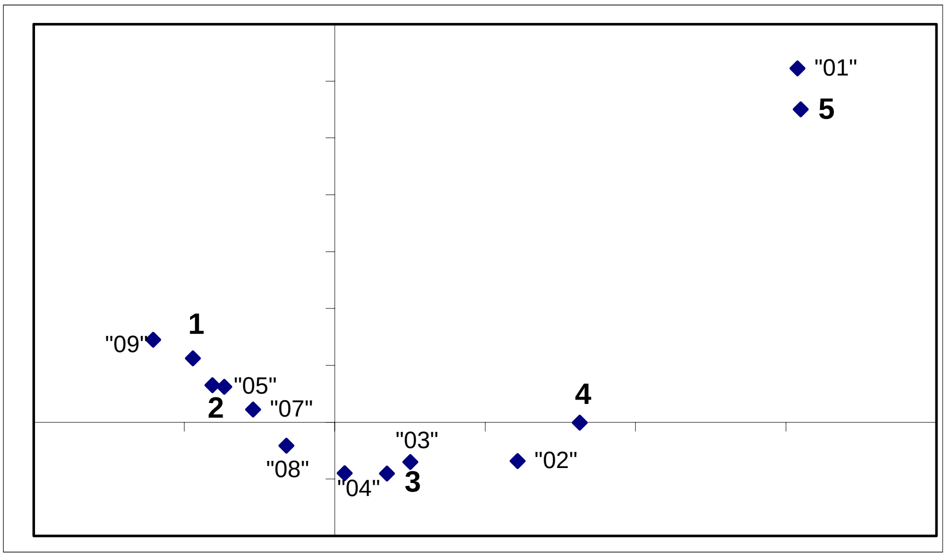
1. The first cluster includes employees in occupations with codes 5, 7, 8 and 9 reaching income of the first two income classes. Occupation 9, however, mostly relates income class 1, while the rest income class 2.
2. The second cluster includes employees in occupations with codes 2, 3 and 4 reaching income of classes 3 and 4. Occupations 3 and 4, however, mostly relate to income class 3, while occupation 2 to income class 4.
3. Finally, the third cluster includes employees in occupation 1 apparently related to income class 5.

As far as the SES survey is concerned, the horizontal axe confronts employees having low income with employees having high income. Moving from left to right employees' income increases. The vertical axe confronts occupations with codes 3, 4, 7 and 8 with the rest occupations (see graph 5).

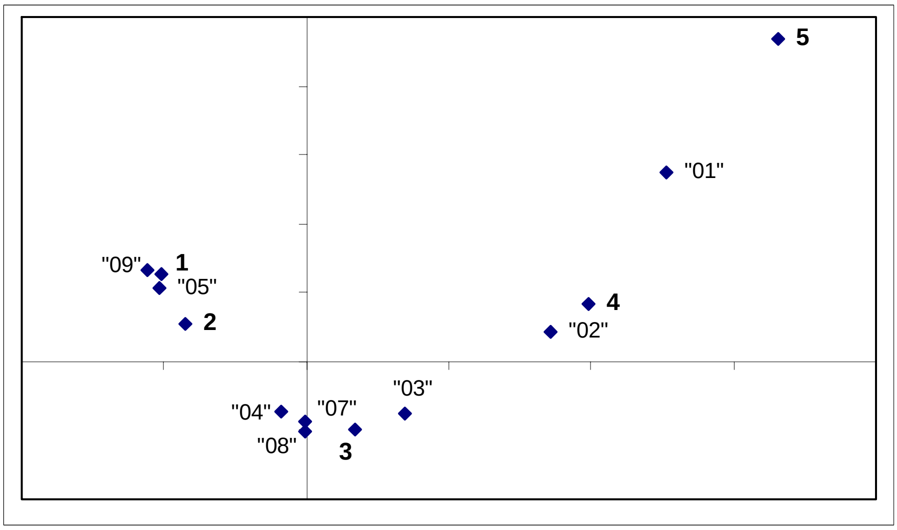
Employees according to their occupation and their income class are split in 3 big clusters.

1. The first cluster includes employees in occupations with codes 5 and 9 reaching income of the first two income classes.
2. The second cluster includes employees in occupations with codes 3, 4, 7 and 8 reaching income of class 3.
3. Finally, the third cluster includes employees in occupations 1 and 2 reaching income of classes 4 and 5. Occupation 2 however mostly relates to income class 4, while occupation 1 to income class 5.

Graph 5. EU-SILC: Occupation ("01"-"09" X Income (1-5)



Graph 6. SES: Occupation X Income



4. Conclusions

a. Average employee income:

- by economic activity does not present statistically significant difference, in most branches of economic activity with the only exception of branches E "*Electricity Gas and Water Supply*" and F "*Construction*",
- by gender for women presents statistically significant difference,
- by age groups does not present statistically significant difference,
- by education statistically significant difference is only observed in lower secondary education and post secondary non tertiary education,
- by occupation does not present statistically significant difference with only exception of "Extraction and building trades workers, other craft and related trades workers. Metal machinery and related trades workers. Precision, handicraft, printing and related trades workers".

b. Number of employees:

- by income classes, gender, economic activity and occupation are correlated. The correlation however is not presented with exactly the same standard in both surveys.