

JOINT EUROPEAN COMMISSION – OECD WORKSHOP ON INTERNATIONAL
DEVELOPMENT OF BUSINESS AND CONSUMER TENDENCY SURVEYS

BRUSSELS

14 – 15 NOVEMBER 2005

Task Force on Harmonisation of Survey Operation and Technical Design

Efficient sample design and weighting methodologies

Analysis of Key Issues and Recommendations

Marco Malgarini

- November 2005 -

Contents

Introduction

- A. Business Tendency Surveys
 - A.1. Key Issues in Efficient Sample Design
 - A.1.1. Reference Population
 - A.1.2. Sample Frame
 - A.1.3. Sampling methods
 - A.1.4. Treatment of missing data
 - A.2. Key issues in weighting methodologies
- B. Consumers Surveys
 - B.1. Key Issues in Efficient Sampling design
 - B.1.1. Sample Frame
 - B.1.2. Sampling Methods
 - B.2. Key Issues in weighting methodologies
- C. Minimum Requirements and Recommendations
 - C.1. Business Tendency Surveys
 - C.1.1. Sample Design
 - C.1.2. Weighting
 - C.2. Consumer Surveys
 - C.2.1. Sample Design
 - C.2.2. Weighting

Introduction¹

Business Tendency Surveys (BTS) and Consumers Surveys (CS) are qualitative economic surveys asking companies and consumers' opinion about their own economic situation and that of the country. BTS and CS are often considered to provide relevant information, complementary to that obtained with quantitative data; moreover, survey results are generally available earlier than quantitative data and are provided with higher frequency. For these reasons, they are considered very useful for short-term analysis and for the early prediction of turning points.

However, the usefulness of BTS and CS strongly depends on their statistical quality; more precisely, according to the recent Handbook on BTS published by OECD, quality of survey data may be measured in terms of reliability, timeliness of release, comparability over time, transparency and accessibility to the users (OECD, 2003). In this respect, this OECD Task Force has the goal of identify the key issues at stake in the fields of sample design and weighting methodologies for both BTS and CS, in order to be able to draft some first recommendations aimed at increase the reliability, and therefore the overall quality, of survey data.

In fact, surveys are generally carried at the national level. However, in the European Union they are co-ordinated and harmonised by the European Commission, while outside Europe the OECD has recently started to develop a system of metadata, with the aim of providing a common starting point for the realisation of BTS and CS worldwide, and ultimately increase their reliability and quality. In this sense, this taskforce aims at building on previous work in order to derive "best practice" recommendations and discuss minimum requirements for efficient sample design and weighting methods. In the following, section A and B investigate on the key issues at stake for sample design and weighting methodologies, respectively for both BTS and CS, providing an overview of the main methods used by the Institutes conducting the surveys, both in the EU and in OECD countries in general; section C concludes providing a first draft of minimum methodological requirements for the optimal design of BTS and CS.

¹ Task Force Members are: Richard Curtin (University of Michigan, United States), Isabelle De Greef (National Bank of Belgium), Richard Etter (KOF / ETZ, Switzerland), Christian Gayer (European Commission), Marie Hormannova (CZSO, Czech Republic), Marco Malgarini (Institute for Studies and Economic Analysis – ISAE, Italy), Rony Nilsson (OECD), Raymund Petz (GKI Research, Hungary), Takashi Sakuma (ESRI, Japan), Philippe Scherrer (INSEE, France), Anna Stangl (Ifo Institute for Economic Research, Germany), Andres Vertes (GKI Research, Hungary), Peter Weiss (European Commission), Jonathan Wood (Confederation of British Industry – CBI, United Kingdom).

According to the report, the ideal survey would utilize probability methods for sample selection, adjust the results for differential non responses, and weight the data so that the results are representative of the total reference population or volume of economic output for business sector surveys. The size of the sample should be determined by the width of the largest confidence intervals allowable for key results. All surveys should clearly document the sampling frame, sample selection procedures, response rates, imputation methods for missing data and weighting procedures.

A. Business Tendency Surveys

In this section, key issues in efficient sample design and weighting methodologies are discussed for BTS, i.e. for the surveys on the manufacturing, services, retail and construction sectors. Once key issues are identified, for each of them we present an analysis of the current practices of the Institutes conducting the survey. Information for EU countries has been taken from the metadata published by the European Commission, available on the internet²; data for countries outside EU are taken from the OECD database. On the whole, for the manufacturing survey we gathered information from 43 institutes worldwide, while for the other surveys only Institutes of the 25-countries wide European Union are considered.

A.1 Key Issues in Efficient Sampling Design

Following and integrating the recent contribution by Donze, Etter, Sydow and Zellweger (2004)³, it is possible to identify the following key issues for an efficient sample design in the case of BTS surveys:

- Identification of the relevant Universe/reference population
- Identification of the sample frame
- The method used for sample selection
- The treatment of missing data

A.1.1 Reference population

² Eu metadata may be downloaded from the following address:
http://europa.eu.int/comm/economy_finance/indicators/business_consumer_surveys/metadata_en.htm

³ L. Donzé, R. Etter, N. Sydow and O. Zellweger, Sample Design for Industry Survey, ECFIN/2003/A3-03 Final Report, November 2004; they actually study survey design in the manufacturing sector only, but their analysis may easily be extended to the other sectors; on the other hand, a somewhat different approach should be considered for the consumers survey, see section 2.

The first step in the setting up of a survey is the identification of the Universe or population of reference; the reference population may be defined as the base population from which a sample is drawn at the time of initial sampling. Typically, in the case of BTS it is represented by all the firms operating in a given sector, as resulting from some official/statistical register. However, some of the firms may be excluded looking at their size or location, or on the basis of their structural characteristic (for instance, the exclusion of government bodies may reflect the definition of the reference population in terms of “private business”; on this, see also below, section 1.1.2, in particular the part concerning cut off strategies).

A.1.2 Sample Frame

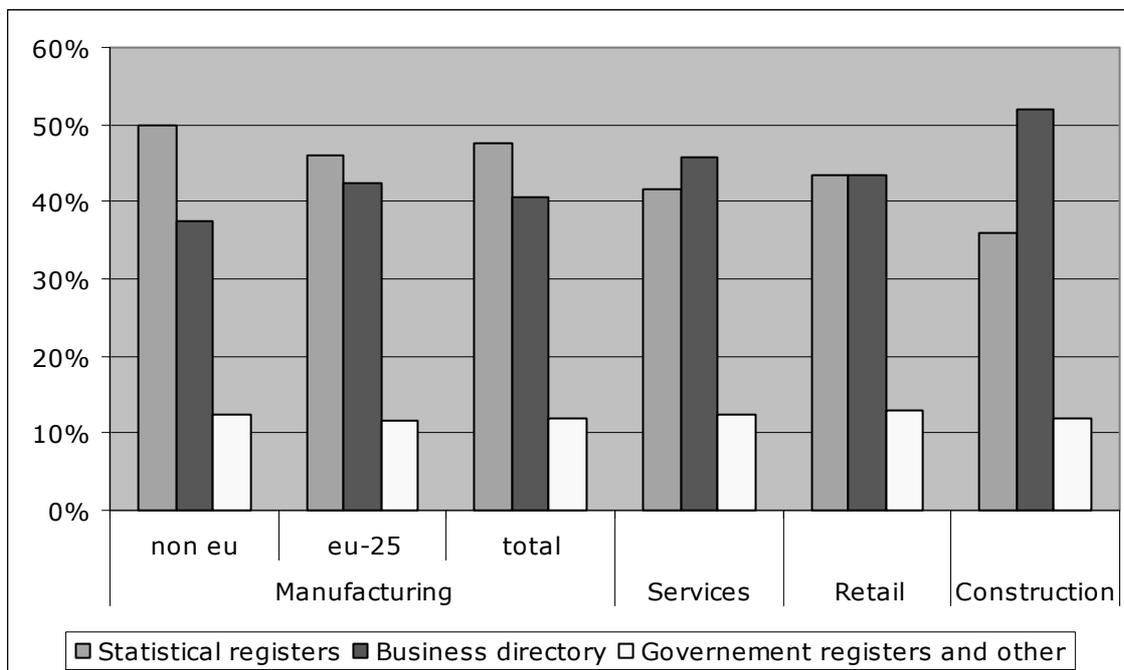
Once the reference Universe has been defined, the objective should be that of choosing a sampling frame able to maximise sample coverage and minimize coverage errors. BTS are usually based on a panel of responding firms, that are re-interviewed every month in order to ensure the stability of the panel and inter-temporal comparability of the results. Demographic and structural characteristics of the respondents have to be known in order to build up the sample, the construction of the sample frame implying the following steps:

- Identification of the appropriate frame list
- Eventual Adoption of a cut off-strategy
- Identification of the sample, reporting and response unit
- Updating of the frame list

The frame list. In starting the sample design, institutes have to choose the relevant frame list and the units that have to be considered inside the list. With respect to the first point, the list should contain an as comprehensive as possible account of the active firms operating in a given sector; in other words, the sample list should be as close as possible to the reference population, in order to ensure an adequate coverage of the relevant Universe. For this reason, official or statistical register are generally to be preferred to membership lists of business associations and chambers of commerce, and to business registers in general; moreover, as we will better see below in the section dedicated to the updating of the frame list, it is very important that the frame is updated regularly, in order to avoid undercoverage (for example, because of newly-born firms, not comprised in the original list) and ineligibility problems (firms that have been dissolved or simply incorrect entries). In the actual experience of the Institutes conducting the surveys, official and statistical registers do actually prevail; however, the use of more partial lists of active firms based on business registers is quite widespread, especially in the case of the services and construction surveys (see chart 1).

Cut-off strategies. BTS lists usually contain a number of firms' characteristics that are helpful to identify them: main information included in a frame list is the demographic characteristic of the firm (i.e. address, telephone number, and contact person), its industry classification, number of employee, turnover and other (i.e., legal framework and region of activity). On the basis of these information, Institutes may choose to consider the whole frame list, or to adopt some cut-off strategy, usually based on the branches of activity or the size of the firm. Discriminating with respect to the branches of activity may be a viable option in order to better focus the survey on the sectors of interest: for instance, in the case of the EU-harmonised service sector survey, only the so-called market-service sectors are considered, excluding all the branches in which the activity of government agencies prevails (Social security, Education, Health and other public services). On the other hand, discriminating with respect to size may respond to the need of ensuring a certain stability of the sample, given the fact that, under a certain size threshold, mortality rates for the firms are quite high, as the number of newly born firms. In the current experience, some of the European institutes performing the surveys use cut-off strategies mostly excluding firms according to their size.

Chart 1 – The frame list



Sample unit. After choosing the relevant frame list, eventually discriminating with respect to sector and/or size, Institutes have to choose the unit on which to perform the sample selection

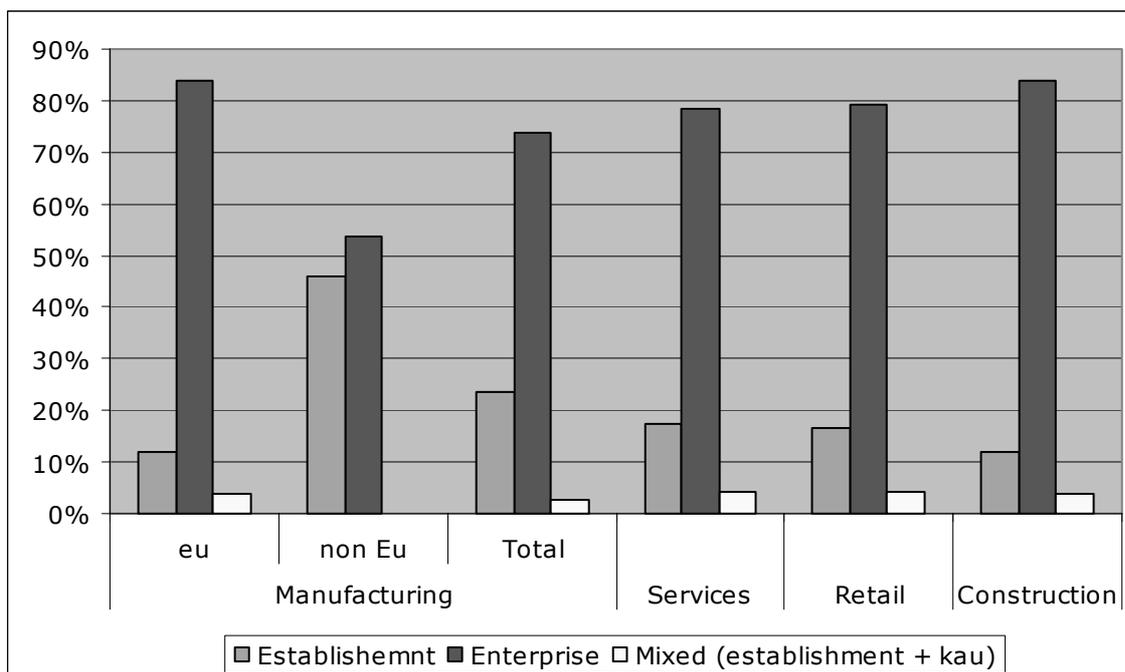
(sample unit). The main choice is between having the whole firm or some sub-section of it as the relevant sample unit: possible options in this sense include the use of establishments, local units or kind of activity units.

In particular, the choice of the establishment as the sample unit derive from the consideration that the legal entity (i.e., the firm) may not be the appropriate level to receive the most adequate information about the economic development of the firm: in other words, it is possible that different production units of a company may evolve differently, and in this sense it may be useful to provide several questionnaires to the various establishments. If it is not possible to gather information at the establishment level, local units, or kind-of-activity units (KAU) may be used. However, these options, while having the advantage of extracting a more precise information about the single establishment/KAU/local unit, may be quite difficult to implement, given the fact that the single units are not able to provide reliable disaggregated information, perhaps reporting to the firm' headquarters in order to provide the answers⁴. As a consequence, in the current experience the use of the whole enterprise as the sample unit largely prevails, even if establishments are used in some cases; only in a few occasion the Kind-of Activity Unit (KAU) is considered as the relevant sample unit, in mixed-mode surveys in which part of the sample is selected on the basis of the whole enterprise and part looking at the KAUs (see chart 2).

Reporting units and response units. If the firm is chosen as the sample unit, it is however possible to have more reporting units within the firms; in particular, response units may be the establishment/KAU/local units of the firms to which questionnaire are actually sent, after the identification of the firm as the sample unit. In this sense, the survey may have different response units, usually identified with the manager of each unit asked to respond to the questionnaire.

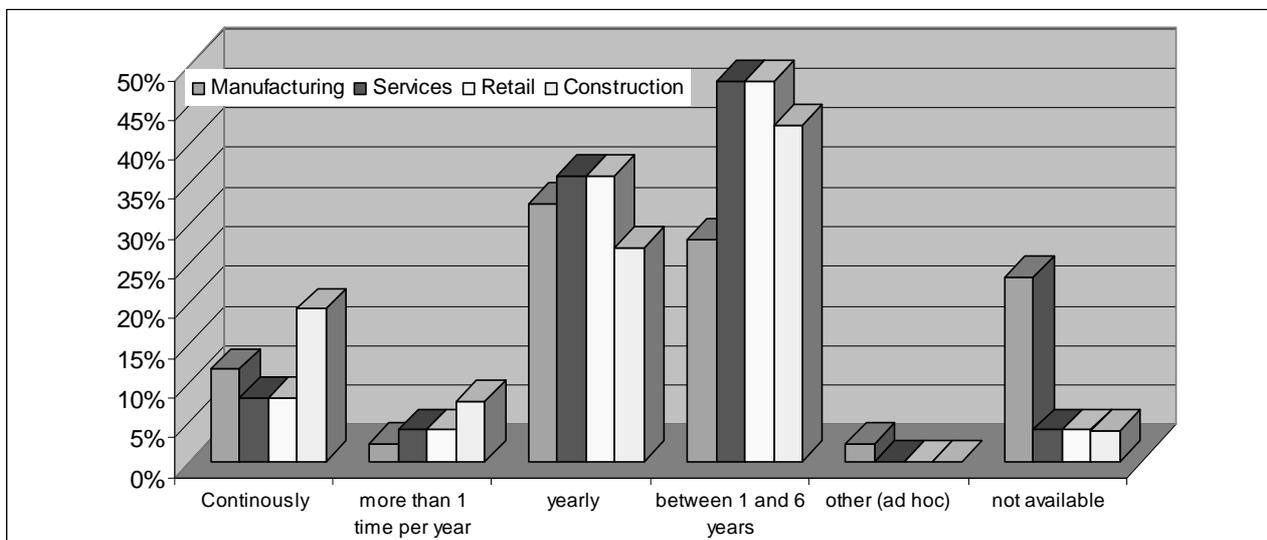
⁴ This is exactly what happened when ISAE adopted the KAU as the sample unit.

Chart 2 – The sample Unit



Updating of the frame list. Finally, the list should be updated quite frequently in order to keep track of the changes in the structure of the reference Universe and avoid the aforementioned problems of possible undercoverage (in the case of new firms entering the market), ineligibility (old firms exiting the market) and duplicate entries (some firms may be included twice due to a recent merger or as separate establishment of the same firm). For the manufacturing survey, many Institutes (35%, see chart 3) update their frame list yearly, with a 30% doing it less frequently (between one and six years) and some of them using a continuous or semi-continuous mode of updating the list. For the other surveys, the updating is less frequent (the modality “between 1 and 6 years” prevails), even if also in the case of the retail, services and especially construction surveys some of the institutes make a continuous updating of their lists.

Chart 3 – Updating of the frame list.



A.1.3 Sampling methods

As reported in OECD (2003)⁵, the collection of data in order to make inference on the Universe' behaviour may be based alternatively on the analysis of the whole Universe itself, or on the collection of information on a sample of units selected in a representative way. Usually, BTS are based on information stemming from a sample considered to be representative of the entire Universe. In order to build up a representative sample, choices have to be made relatively to:

- The sampling method
- The sampling size

The problem of the treatment of missing data is separately discussed below in section A.1.4.

Sampling methods. As mentioned before, BTS are generally based on a fixed panel of reporting units, i.e. the same set of units is surveyed in each period, in order to reduce the sample variance. A fixed sample structure may however cause representitiveness problems, because the panel may loose its power of correctly representing the reference Universe if it is not updated regularly. For this reason, usually Institutes adopt a rotating panel method, in which a fixed percentage of response units are replaced at regular intervals. More precisely, in the ideal rotating panel the largest and most important firms should always be included and never rotated

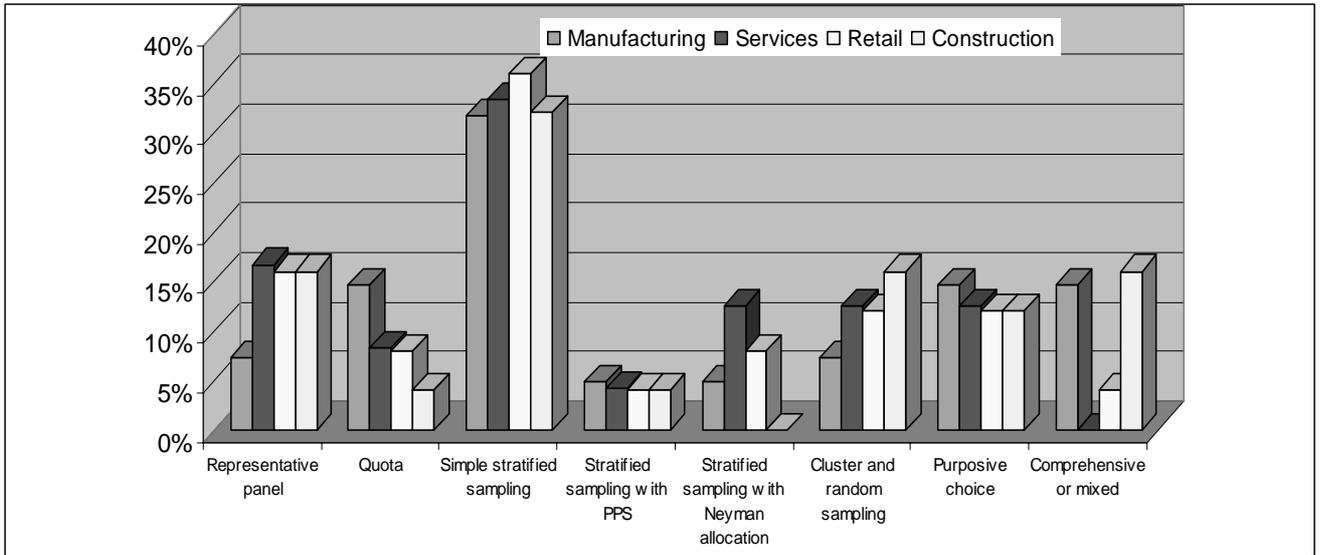
⁵ See OECD, Business Tendency Surveys: an Handbook

out; the smallest firms should be selected with a much smaller probability and be rotated out on a regular basis⁶.

More in detail concerning the method used to extract the sample, in the current experience comprehensive surveys with cut off are used in a fairly limited number of cases (see chart 4). They are generally limited to a particular category of firms (typically, those exceeding a predetermined size), the other being generally extracted with random sampling techniques. In the case of the manufacturing and construction survey, more than 20% of the Institutes use purposive or quota sampling methods, not based on any probabilistic or statistical assumption to extract the firms from the relevant frame list; this percentage drops to about the 10% of the Institutes for the survey on the service and retail sector. In the remaining cases, samples are extracted using statistically founded methods in order to ensure the sample representativeness with respect to the whole Universe. In particular, the most frequently used method of sample selection is the simple stratified sample, in order to keep track of the heterogeneity of sample population; in some cases, stratified sampling techniques using PPS or optimal allocation methods are used: in the former, the stratum size is defined proportionally to the size of the universe (i.e. the sampling fraction is constant in each stratum); with the latter, the number of firms selected in each stratum is chosen taking into account not only the corresponding Universe size, but also the variability of the firms belonging to that stratum. In this case, a greater heterogeneity of the firms corresponds to a larger stratum size; on the other hand, in strata where firms are relatively similar, a minor number of firms are needed to get the desired precision of the results. In the case of stratified sample (simple stratified sample, PPS and OAS according to Nyman), stratification variables usually include the official NACE industry classification, size (as expressed according to turnover or employee of the firms) and the region of residence of the firms.

⁶ There may be some industry in some country where there are no dominant firms; in this case, no firm should be included as a permanent member. When there is such a dominant firm, it should always be included.

Chart 4 – Sampling Methods



Sample size. Once the method for sample selection has been chosen, the Institutes have to decide about the size of the sample they want to use. In fact, the measurement error incurred in the survey is strictly dependent on – together with the method used for sample selection - the size of the sample. If random sample techniques are used, well known algorithms allow evaluating the precision of the estimates, given the method used for sample selection and the size of the sample⁷. Formally, the qualitative estimate of the percentage of answers p for a generic question i can differ (in absolute value) from the true value P (relative to the whole universe) for a quantity larger than d , with a confidence level $1-\alpha$ (with α for instance equal 5%).

We have:

$$\Pr\{|p - P| \geq d\} = \alpha \tag{1}$$

Assuming that p is normally distributed, it follows:

$$d = z_{\alpha/2} s_p \tag{2}$$

⁷ See on this W. G. Cochran, “Sampling Techniques”, 3rd edition, Wiley, New York, 1977.

where $z_{\alpha/2}$ is the value of the normal deviate for a predetermined value of α (e.g. for $\alpha=5\%$, $z_{\alpha/2}=1,96$) and sp is the standard error of the qualitative variable p .

As an example, in the case of OAS methods, the standard error may be derived as follows:

$$s_p = \left(\frac{(\sum_k W_k \sqrt{p_k q_k})^2}{n} - \sum_{hk} W_k p_k q_k \right)^{1/2} \quad (3)$$

being:

k the generic stratum;

N = the Universe size; N_k = the number of universe firms belonging to the generic stratum k

$W_k = N_k / N$ the weight of the stratum k

$p_k q_k$ the estimated qualitative variance of p in the stratum k .

A.1.3 Treatment of Missing data

Sample sizes reported above do not account for the actual number of answers received each month by the Institutes: indeed, they only represent the magnitude of the theoretical sample, the actual number of answers depending on the rate of response to the survey. More specifically, it is possible to distinguish between “unit” non response and “item” non response: the latter represents missing data to questions on otherwise completed interviews, the former occurs in the case of the absence of the entire interview – or “unit”. If unit or item non response involves a dominant firm, missing data problems could eventually be judged severe; if repeated attempts fail to produce an interview, imputation of missing data may be needed. In contrast, if there is missing data among the smallest firms, the problem is not critical and the data can be easily imputed based on the responses of similar sized firms (on this, see also the section on the recommended best practices at the end of the paper).

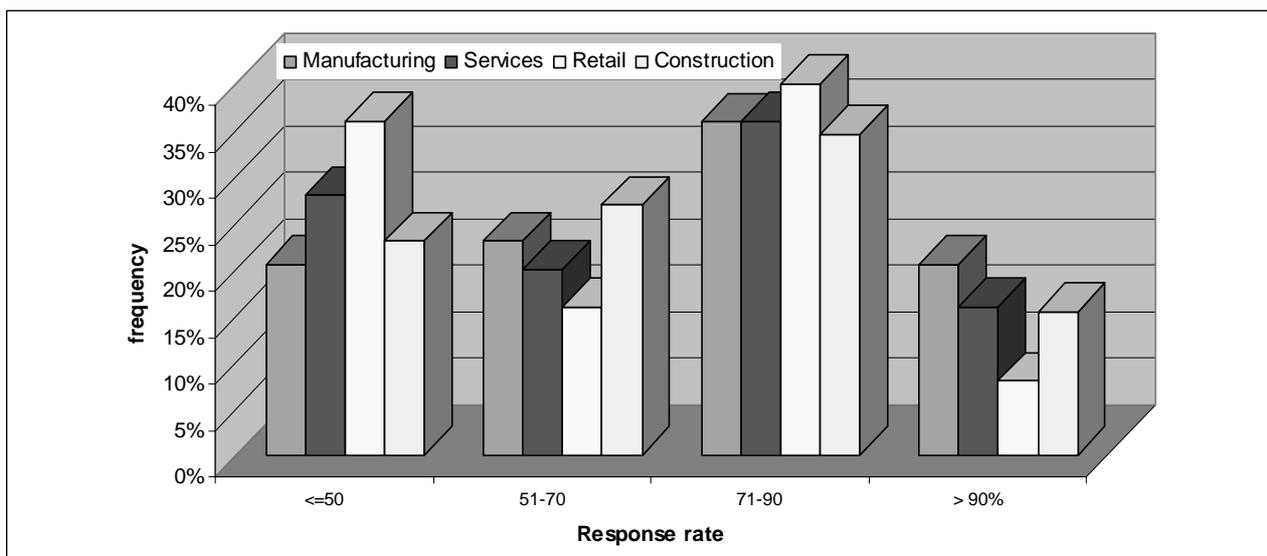
With respect to unit non-response⁸, chart 5 reports the Institutes simple response-rate⁹ distribution for all the 4 surveys considered here: generally speaking, response rates appear to be relatively higher for the manufacturing survey, with over 56% of the institutes reporting

⁸ Measures of item non response are not currently available for BTS.

⁹ The method for calculating response rates is not stated explicitly in the EU metadata; the simplest measure of calculating response rates is the ratio between the number of the firms answering in a given month and that of the whole sample, and it is probably the one that has been used here. For alternative, more appropriate methods for calculating response (and non-response) rates see OECD (2003).

response rates equal or higher than 70% and more than the 20% of Institutes with response rate even higher than 90% of the sample. Moreover, for the manufacturing survey the percentage of institutes reporting response rate lower than 70% is only 43%, raising to about 50% or more for the other surveys.

Chart 5 – Survey response rates



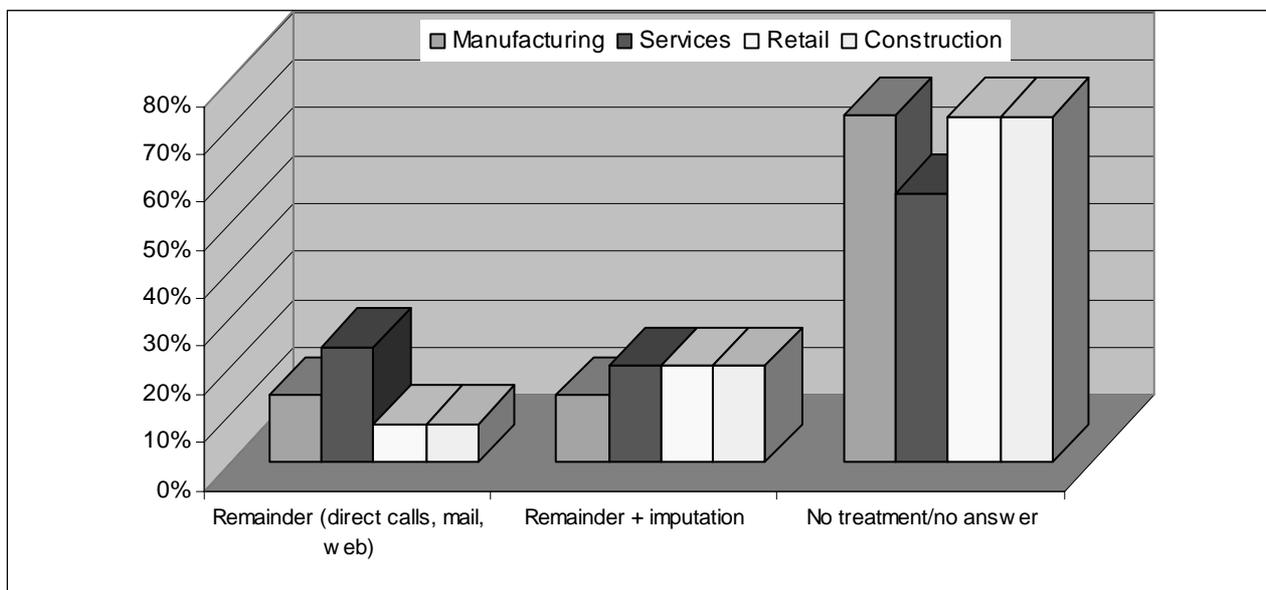
However, non responses have a not negligible weight in the surveys considered, especially in those on the retail, services and construction sectors. A non-constant answering behaviour – both because of unit or item non response (see again Donzé et alii, 2004) - from the companies participating the survey may indeed seriously damage the overall quality of survey data, causing bias and excessive variance of survey estimates; moreover, missing data problem may arise from an incorrect identification of the response unit, which has to be as much stable as possible in order to minimize possibility of non response or incorrect answers.

For these reason, the use of appropriate methods for dealing with missing data problems is advisable; however, in the current practices (chart 6), only a quite small fraction of Institutes do indeed report about the use of such methodologies. In particular, almost in all surveys some 70% of the Institutes do not report (i.e., they don't use them, or do not report to use them) the use of methods for the treatment of missing data¹⁰; remaining Institutes use either a simple system of

¹⁰ This result is in contrast with those provided by Donzé et alii (2004), according to which almost all the Institutes explicitly declared to use follow up methods to increase the response rates; however, our data are taken from the official EU metadata, where the Institutes do not report such follow up activities.

reminder (by telephone, fax, mail and e-mail) or more complex system of data imputation and re-weighting of the data.

Chart 6 – Treatment of missing data



A.2 Key Issues in weighting methodologies

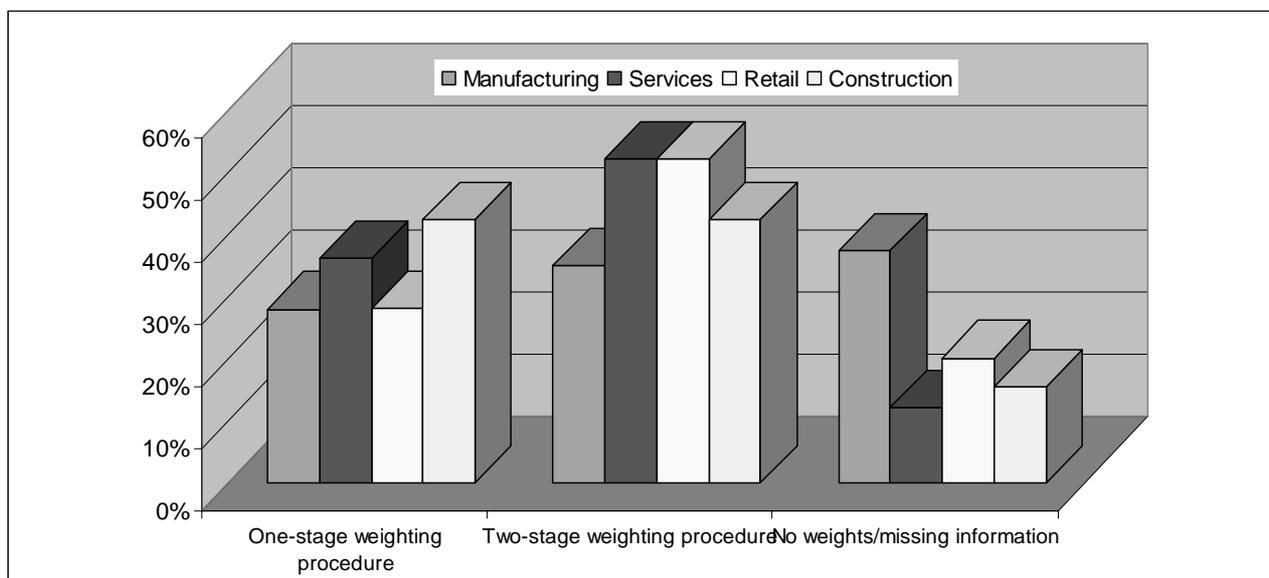
Weighting is meant to transform the realised sample into estimates of the reference population, improving the precision of the estimates. Weights are usually based on auxiliary sources - coming from outside the firm or from the survey itself – providing information on the size or output of the firm or of the sector in which the firm operates. In the current experience, almost 40% of Institutes do not use a weighting method in the case of the manufacturing survey¹¹, the percentage dropping to circa 20% for the other surveys (see chart 7); this means that in these cases the surveys represent all firms equally regardless to size. However, Institutes outside the EU are mostly non using weights (or, non reporting about the use of them) in the manufacturing survey, while those participating to the EU harmonised programme use weights in almost 80% of the cases, as for the other surveys considered.

The two main weighting procedures used by the Institutes are based either on a one-stage or a two-stage scheme. In the first case, a weight is associated to each reporting unit, in order to take into account its relative importance inside the sample; in the second case, a firm-specific

¹¹ However, most of EU countries actually use weighting methods, see again Donzé et alii (2004).

weight is used to calculate strata-level results, further aggregating the data with some external sources in order to obtain the industry aggregates. In the current practices, two-stage weighting procedures slightly prevail over the single-stage option in all the surveys considered; in the second stage, the variables used as weights usually come from an external source. The variables that are more often currently considered as weights in the first stage are employee or turnover provided by the firm itself within the survey; on the other hand, official data on industry-level employee, turnover or value added are generally used for the second level of aggregation. In most cases, institutes use the same system of weights for all the questions of the surveys, even if sometimes separated systems of weights are used for particular questions (for instance, external orders).

Chart 7 – The use of weighting methodologies



B. Consumers Surveys

As recently stated by Curtin (2004)¹², Consumer Sentiment surveys are now regularly performed worldwide in at least forty-five countries, and their outcomes are both widely used in the business and financial press and analysed by economists and policy makers. They measure consumer opinions and expectations using qualitative questions; as for BTS, also in the case of CS an accurate sample design, together with the use of appropriate methods for weighting raw data and build up the sentiment index, are crucial to ensure a good quality of the survey. Differently from

¹² See Curtin, R. (2004), “Consumer Sentiment Surveys: Worldwide Review and Assessment”, paper presented at the 27th CIRET Conference, Warsaw, September 2004.

BTS, where a panel structure of the sample often prevails, for CS each month an independent cross-section of households is usually drawn, with possible re-interviewing of the units after a fixed interval. For CS, the main choices that have to be made are the identification of the sample frame, the method to be used for sample selection and that concerning the aggregation of individual results.

B.1 Key Issues in Efficient Sampling Design

As in the case of BTS, also consumers surveys are usually based on information stemming from a sample considered to be representative of the entire Universe. In order to build up a representative sample, choices have to be made relatively to:

- The sampling frame
- The sampling methods

B.1.1 Sample frame

In the case of the consumer survey, the construction of the sample frame implies the following steps:

- Identification of the appropriate frame list
- Eventual Adoption of a cut off-strategy
- Identification of the sample unit, reporting unit and response unit
- Updating of the frame list

Frame list. it is crucial that the right population is being sampled, and that all the members of the population are correctly located, so all of them may have a chance of being sampled¹³: in this sense the adopted list should be up-to-date and complete, possibly with some cut-off strategy linked to the age or some other characteristic of the population of interest. Looking at the actual experience in OECD countries¹⁴, the frame list is usually build up from either an official population register - supposed to include every adult member of the country population – or from telephone registers (half of the 27 countries considered use the first option, half the second); in the latter case, a risk of possible exclusion of a part of the relevant population from the list do exists, because in some countries individuals may choose not to be comprised in the telephone

¹³ See on this “Standard and Best Practices”, American Association for Public Opinion Research, available on line at www.aapor.org

¹⁴ Information available refer to the EU (25 countries), US and Japan.

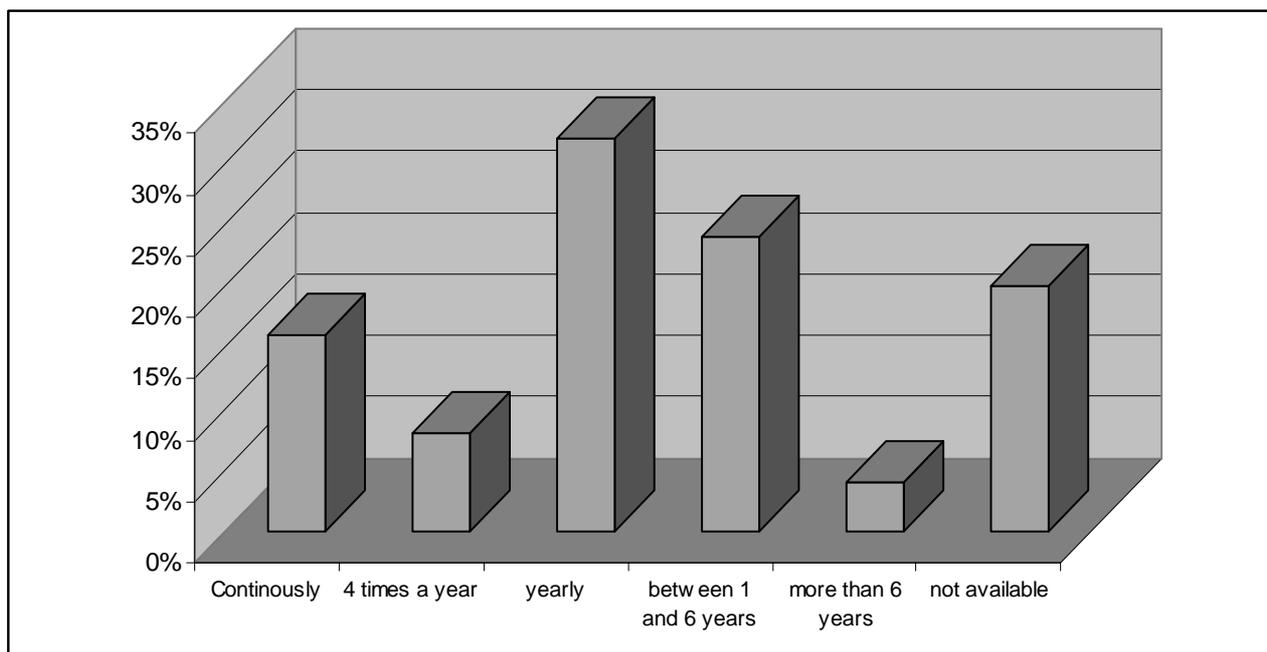
directories (so called “opposition list” or “red list”). In these cases, sample procedures may be adapted accordingly (adopting random extraction of telephone numbers, or using different sample techniques for people excluded from the telephone directory), or appropriate system of weights should be used in order to take into account the possible bias arising from the exclusion of possibly relevant sample units from the list (see below).

Cut-off. List participants are often cut-off on the basis of their age; in EU, the cut-off age may vary across countries (with a minimum age of 14-18 years and a maximum age in some cases); in some countries, cut-offs on a geographical base are also applied.

Sample Unit and response unit Also in the case of consumer surveys, the response unit may differ from the sample unit. Typically, samples are devised to be representative of all households, with the selected respondent reporting on the economic status of the household. In fact, while it is true that individuals are the respondents, it is the economic status of the household that is the primary interest of the survey.

Updating of the list To ensure an high quality of the survey, the frame list should be updated frequently in order to monitor as close as possible the evolution of the relevant population and ensure a good coverage, minimising missing and duplicate entries. According to the available information, in OECD countries more than half of the institutes update the frame at least yearly, even if in almost 30% of cases the lists are updated every two or even 5-6 years and in one case the list is not updated since 1991 (see chart 8).

Chart 8: Updating of the list for consumers surveys



B.1.2 Sampling methods

Once the relevant frame list has been chosen, possibly adopting some cut-off strategy, sample units should be extracted with an appropriate method in order to ensure the highest possible representativeness of the resulting sample. Key issues in sampling methodology involve the choice of the appropriate sampling method, together with that of the optimal sample size.

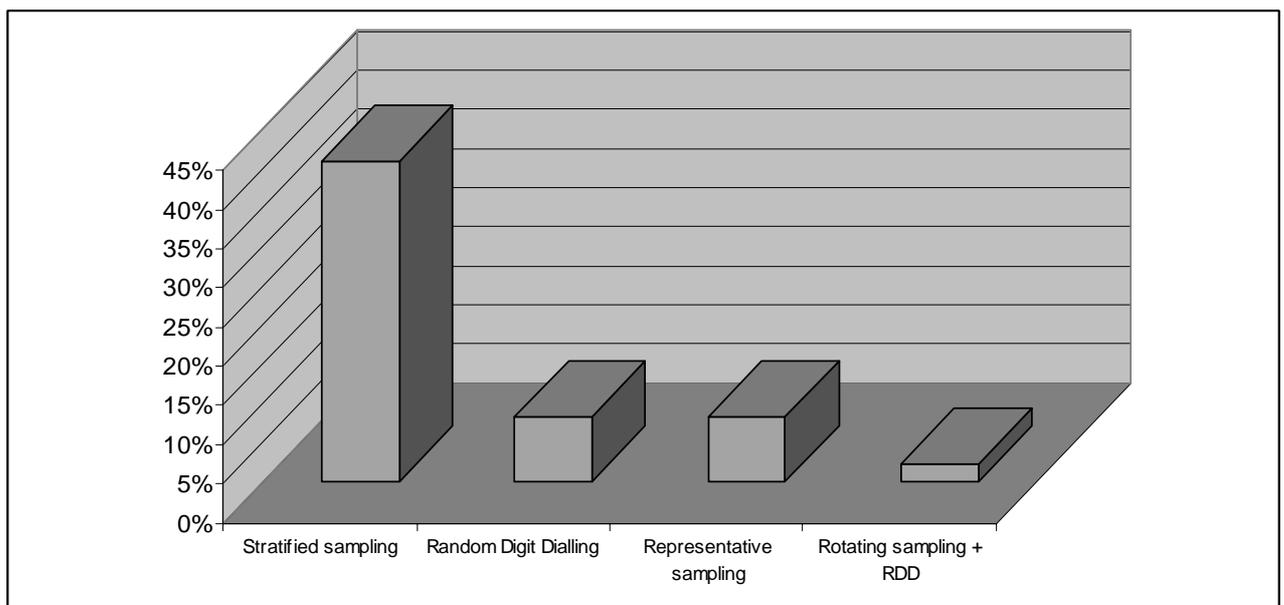
Sampling Methods. All the surveys considered in this study adopt some form of random or probabilistic sampling based on statistical or probabilistic method. Usually, each month an independent cross-section of households is drawn; however, in the EU a general strategy of simple random sampling is used, while in the US a rotating sampling design is applied, in which the respondent chosen in each drawing is re-interviewed six months later, in order to provide a regular assessment of change in consumers' attitudes and behaviour (see chart 9). The most widely used sampling methodology is that of stratified sampling, in which units are extracted from the relevant frame list, with a single or multiple stage¹⁵ strategy, stratifying on one or more variables; in some cases, the Random Digit Dialling Method is used, in which the institutes

¹⁵ In a first stage the sample may be stratified according to some characteristic of the population (region of residence; type of occupation etc); in the second stage the sample units are chosen according to their per quota (i.e., with respect to gender) distribution in the population.

extract directly (with various methods, see for instance Curtin, 2000¹⁶) the telephone numbers of the respondents from the chosen telephone registers. As for the variables considered for stratification, they usually include information on respondent socio-demographic characteristic, such as gender and age, type of occupation, region/town of residence, size of the household; in some cases, information about the respondent social class and the type of settlement of the households is also used.

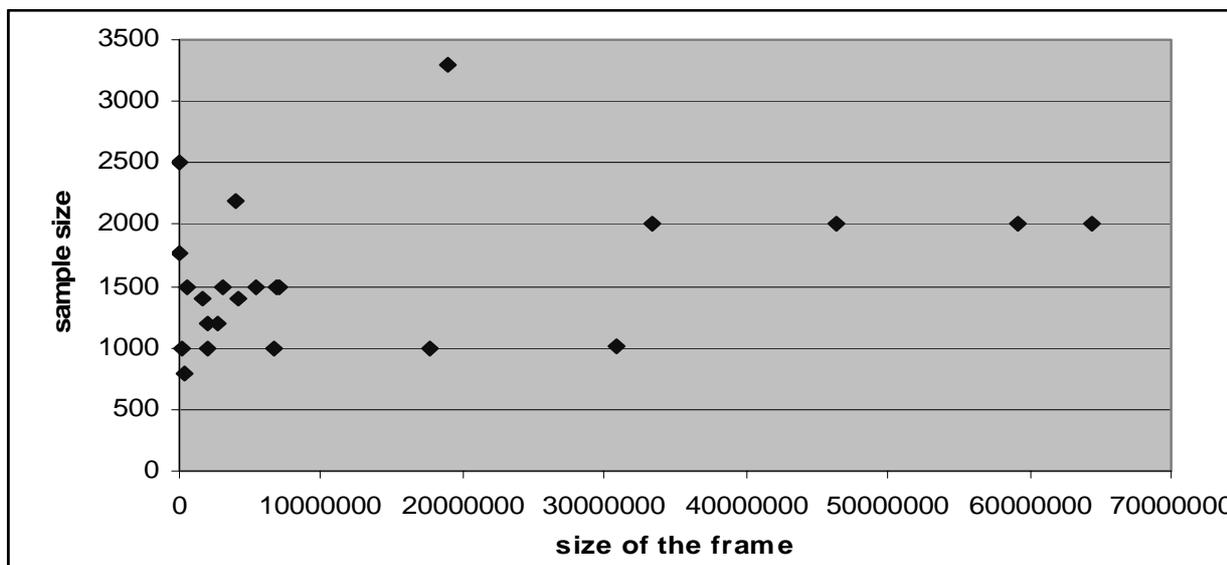
Sample size In the literature, there is no consensus about what an appropriate sample size for representative surveys should be. As emerges quite clearly from chart 10, in the current experience, sample size generally converges to about 2000; this size provides standard errors (or confidence intervals) that are presumably acceptable for this type of survey.

Chart 9: Sampling methods



¹⁶ See R. Curtin, "Surveys of Consumers", The University of Michigan, 2000, available at <http://www.sca.isr.umich.edu/documents.php?c=i>

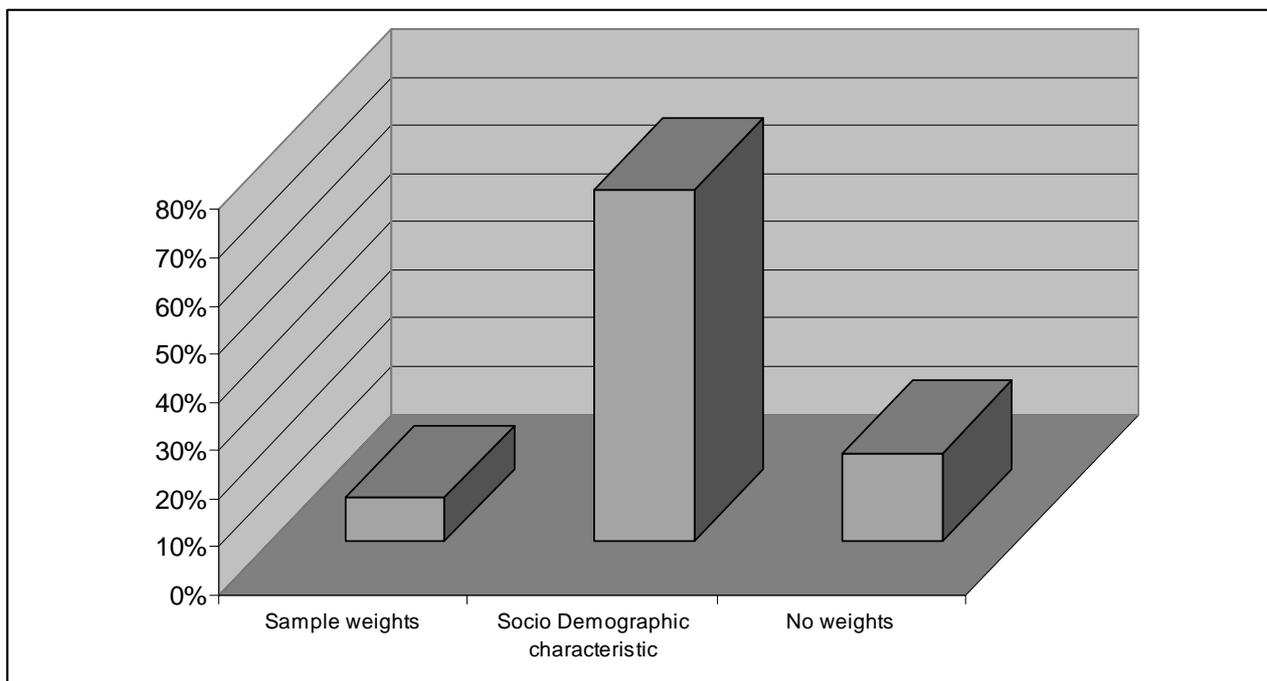
Chart 10 – Sample size and survey population by country



B.2 Key Issues in Weighting Methodologies

Information gathered from survey' respondent may be appropriately weighted to derive aggregate information on consumers' opinion and expectations. Weighting may improve the precision of the estimates; weighting methodologies may be based on auxiliary information (socio economic weights) or on inverses selection probabilities of the sample units (sample weights). In particular, most of the institutes considered use socio demographic characteristics of the households to weight individual responses and derive aggregated results (see chart 11): among the demographic characteristic considered the most common are the gender and age of the respondent, the region of residence and the size of the township; the socio-economics characteristic often considered by the institutes are the economic occupation of the households, her level of education, type of the area/municipality where she lives and her housing conditions. However, a number of institutes do not use weighting; this procedure is appropriate only when every household has an equal chance of selection and when there is no differential no-response; in other words, only when these two criteria are met there is no need for selection probability weights or socio-economic weights.

Chart 11 – Weighting Methods for consumers' surveys.



C. Minimum requirements and recommendations

The identification of the key issues and the analysis of the current practices allow us to derive a first draft assessment of the minimum requirements needed for the realisation of both BT and consumers surveys in the field of sample design and weighting methodologies. Discussion of minimum requirements, and consequently of the draft recommendation to be addressed to the Institutes, are presented separately for Business Tendency (section C.1) and Consumers surveys (section C.2).

C.1 Business tendency surveys

C.1.1 Sample design

- The frame list should include an as exhaustive as possible account of the active firms relevant to the survey of interest in order to ensure an high coverage of the reference population. As a consequence, the use of official or statistical registers of active firms is recommended over that of - more partial – business or membership register, including

only a subset of the statistical Universe of reference¹⁷. Institutes are also advised to use cut-off strategies in order to stabilize the panel of the respondents (size cut-off) and for a better identification of the survey' objectives (branch cut-off).

- As a minimum requirement, frame lists should be updated as soon as a new census of active firms is made available by official statistical offices. This means that the lists should be updated yearly when new data are available, or at the soonest after the release of the new information, in order to take into account the evolution in the structure of the sectors on which the surveys are performed.
- As for the sample unit, the choice of the establishment may be considered ideal in order to gather more precise information on both industrial composition and regional distribution of the activity of the enterprise. However, gathering information at the establishment level is in practice very difficult; a viable option may be that of using Kind-of-activity units, if the Institutes are particularly interested in correct information about the industrial structure. The use of local units is on the other hand advisable to provide data correct with respect to regional distribution. Firm level data may be misleading both with respect to the regional and industrial structure of activity, but are often the only possible option for the Institutes conducting the surveys because of the difficulty in having the firms responding at a more disaggregated level.
- Even if the firm is identified as the sample unit, it is advisable – whenever is possible - to have different reporting units and response units within the firm. Both in the case when the firm or the establishment/KAU/local unit is chosen as the sample/reporting/response unit, it is strongly recommended for the quality of the survey that the Institutes ensure that the same person (the response unit) answer the questionnaire every period.
- As a minimum requirement for sampling methods, a fixed panel should be used, established on a statistical basis, using a rotating pattern of updating, with a fixed percentage of participants being replaced at regular intervals.
- As a minimum requirement, sampling methods should be based on sound probabilistic consideration. The use of exhaustive sampling (i.e., taking into consideration the whole population of firms) may be a viable option for small countries or for a sub-set of the sample (e.g., very large firms operating in a given sector). Avoiding of purposive or ad hoc sampling methods is strongly recommended. Different probabilistic methods of

¹⁷ Access to official statistical registers is not always easy for all the institutes; the recommendation should be intended in the sense that official census registers should be used as soon as they are updated and are made available to the Institutes performing the surveys.

sample selection are currently used: as a general consideration, the more heterogeneous is the population of interest (in terms of industrial structure, size, and kind of activity), the more is advisable the use of stratification-based sampling methods. In this sense, stratified sampling with optimal allocation to strata gives the best results in terms of reduction of the sampling error.

- It is critical for all survey organisations to define what procedures are used to adjust for item non response or missing data. As a minimum requirement, all the institutes are advised to closely monitor the impact of missing data (especially if they concern large dominant firms), and to use follow up techniques in order to obtain the cooperation of the respondents and reduce the impact of missing data on the quality of the survey; these techniques may include reminder notices of all types (telephone, fax, e-mail, etc). Ensuring that the same person answer the survey in any given period may strongly improve response rates. Follow up techniques may be based on the postal system or, better, on CATI techniques or on the newly developed web-based survey modes. The use of imputation methods for the treatment of remaining missing data should be considered with care, in order to avoid possible distortions; in this sense, the use of re-weighting techniques -taking into account the different composition of the panel in adjacent surveys – may be advisable to reduce possible bias.

C.1.2 Weighting methods

The use of weights is strongly recommended in order to improve the precision of the estimates: in this sense, auxiliary information stemming from external sources (mainly official statistics) may be of great help. The following three main issues should be taken into consideration when forming survey weights¹⁸:

1. Sampling probability. It is important that the inverse of the sampling probability of the business be a component of the business weight.
2. Business size. A measure of business size relevant to the subject of interest should always be used to weight responses to qualitative questions. The most readily available relevant measures of size are turnover and employment. These would normally be collected in the survey and be multiplied by the inverse of the sampling probability to form the overall business weight.

¹⁸ Formulae related to forming survey weights can be found at: <http://www.oecd.org/dataoecd/11/51/33659920.pdf>

3. The above variables should be used to generate branch or cell level estimates (e.g. detailed industry, possibly split by business size). Branch level estimates should then be weighted together to form aggregate estimates (e.g. total manufacturing) based on population statistics of value added or the most closely related variable available.

If probability sampling is not used¹⁹ and thus probabilities of selection are not available, a minimum requirement if business weights (e.g. employment / turnover) are to be used is that estimation cells are split by business size group (e.g. small, medium and large businesses). This would require approximate population level estimates from external sources for these estimation cells to use as weights in order to form aggregates as outlined in point 3. That is, businesses with large differences in size (e.g. 50 employees compared to 500) should only be included in the same estimation cell if the inverse of sampling probability is part of the business weight. If this is not done then the survey estimates will be biased.

C.2 Consumers Surveys

C.2.1 Sample design

- The frame list should include an as exhaustive as possible account of the adult population on which to perform the survey. As a consequence, the use of official census or statistical registers of active firms is to be preferred to that of - more partial – telephone register, from which part of the population may be excluded, on a voluntary base. When telephone registers are considered as frame list, the use of appropriate methods is recommended in order to correct for possible bias. Cut off strategies with respect to age are advisable; in this sense, further harmonisation is needed, even within the EU, in order to adopt a comparable strategy of exclusion under a certain age threshold. On the other hand, Institutes are recommended not to use cut-off strategies based on region of residence, which present a high risk of inducing bias in the final estimates.
- As a minimum requirement, frame lists should be updated yearly, in order to monitor as close as possible the evolution of the relevant population.
- As for sampling methods, a major difference emerge between the EU and the US experience: in Europe, cross section samples (with independent drawing of the sample being made every month) is generally in use, while in the US a rotating panel design

¹⁹ For example, in section A1.3 we saw that a number of institutes used non probabilistic sampling methods such as: Representative panel, Quota sampling and Purposive sampling.

prevails, in which the same cross-section of individual is re-interviewed after six months, in order to assess changes in attitudes and behaviour of individuals. The possibility of adopting a rotating panel design may possibly be carefully considered by the EU institutes and the Commission, given the possible advantages in terms of research option made available, and comparability of results, in presence of possibly negligible costs.

- As a minimum requirement, random sampling techniques have to be used in order to ensure the survey representitiveness. In case of heterogeneous population (large countries, with relevant differences in town size, regional population, regional development, etc), the use of stratified sampling methods – including Random Digit Dialling (RDD) sampling methods - may be preferred to simple random sampling.

C.2.2 Weighting

- Use of weighting methods is advisable to ensure a better representitiveness of the results. Demographic characteristics of the households should be used, considering among them the gender and age of the respondent, the region of residence and the size of the township; socio-economics – such as economic occupation of the household, her level of education, type of the area/municipality where she lives and her housing conditions characteristic - may be also considered as weights.