Sampling for the European Social Survey- Round II: Principles and requirements

The objective of Work Package 3 is the "design and implementation of *workable and equivalent sampling strategies* in all participating countries". This concept stands for random (probability) samples with comparable estimates. From the statistical point of view full coverage of the population, non-response reduction and considering design effects are prerequisites for the comparability of unbiased or at least minimum biased estimates. In the following we shortly want to

- describe the theoretical background for these requirements,
- show some examples, how the requirements can be kept in the practices of the individual countries and
- explain, which information the expert panel on sampling needs from the National Co-ordinators to evaluate their proposed sampling schemes.

1. Basic principles for sampling in cross-cultural surveys

Kish (1994, p. 173) provides our starting point: "Sample designs may be chosen flexibly and there is no need for similarity of sample designs. Flexibility of choice is particularly advisable for multinational comparisons, because the sampling resources differ greatly between countries. All this flexibility assumes probability selection methods: known probabilities of selection for all population elements." Following this, an optimal sampling design for cross-cultural surveys should consist of the best random practice used in each participating country. The choice of specific design depends on the available frames, experiences and of course also the costs in the different countries. If adequate estimators are chosen the resulting values can be compared[1]. This comparability has to be the goal of the sampling strategy and its implementation for the ESS.

[1] To ensure comparability design weights have to be computed for each country. For this, the relative selection probabilities of every sample member at each stage of selection must be known and recorded. A description of the sample design data file that has to be delivered by the National teams can be found in Appendix 3.

2. *Discussion of standards set in the Technical Annex[2] / Specification for participating countries*

Only random samples provide a theoretical basis which allows us to infer from the sample to the population or sub-sets of this. As design based inference is one important goal in the project, probability samples are required. However, this is related to other requirements:

- full coverage of the target population
- high response rates (ESS: target minimum response rate: 70%)
- no substitution
- the same minimum effective sample sizes (completely responded units) in participating countries (ESS: 1,500 or 800 where population is smaller than 2 m. inhabitants).

These requirements can only be sensibly discussed in the context of random samples. They form a theoretical system that in the end ensures equivalence. The crucial point is that the practical implementation works.

Full coverage of the residential population

An important step in planning a survey is the definition of the population under study. In the case of the ESS it contains in each country persons 15 years or older who are resident within private households, regardless of nationality and citizenship or language.[3] This definition applies to all participating countries and thus every person with the defined characteristics should have a non zero chance of being selected. It follows from that, that the more completely the frame covers the persons belonging to the target population, the better the resulting sample will be. The quality of the frames – e.g. coverage, updating and access - may differ from country to country. Therefore, frames have to be evaluated carefully. The results of these evaluations have to be documented and taken into account when the data are analysed.

The following differences in frames can be expected:

a) countries with reliable lists of residents that are available for social research such as Norway, Sweden, Denmark
b) countries with reliable lists of the households/addresses that are available for social research such as Switzerland, Netherlands, U.K.
c) countries without reliable and/or available lists such as Portugal, France

Drawing a sample is more complicated if no registers (lists) are available (group c). In this instance multistage designs are usually applied, in which the selection of municipalities forms

[2] see www.europeansocialsurvey.com

[3] In countries in which any minority language is spoken as a first language by 5 % or more of the population, the questionnaire will be translated into that language.

the first stage and the selection of households within these municipalities the second stage. Because no sampling frames are available, the crucial problem is the selection of the households. There are two main ways to go about this. The first is to list all the addresses within certain districts of each selected community. The target households are then drawn from these lists. Arguably, it is possible to assess this procedure as one way of drawing a random sample, even if one which is fairly strongly clustered. Another frequently used way to find target households is the application of random route elements. The question here however, is the extent to which random routes can be judged to be "strictly random". In Lyberg´s view, e.g., these techniques do result in non-probability samples (see evaluation of the IALS DATE). At the very least, the following questions have to be answered

- How are the rules for random routes defined in the countries?
- What experience have interviewers with random walks?
- How can the whole random walk process be controlled?

in order to minimise the interviewer´s influence on the selection of respondents. An acceptable method might involve the interviewer doing the complete walk, recording the sampled addresses and notifying these to the survey office before he/she begins contacting any addresses.

Even in countries where reliable frames exist, we have to expect pitfalls. For example, it will be difficult to fully cover people with illegal status. Such systematic losses because of undercoverage cannot be ruled out in practice. However, they must be documented carefully.

Response rates

Non-response is the next problem for the representativeness of the target population in the sample. A carefully drawn gross sample from a perfect frame can be worthless if non-contacts and refusals lead to systematic biases. Therefore, it is of essential importance to plan and implement a sufficient number of contacts as well as appropriate field work strategies for the persuasion of the target persons to participate in the survey. However, the fixed goal of 70% is particularly challenging for some countries where response rates of 50 percent or even lower are common (see Technical report of round I). Nevertheless, all efforts should be done to avoid non-response because it includes the danger of biased samples, and cell weighting is not such a global means of "repairing" samples, as it is sometimes argued (Häder/Gabler 1997).

To sum it up, the transition process from the gross sample to the net sample is of high importance for the quality of the collected data. Comparability of estimates can be achieved

only if the net samples are not seriously biased. That is more likely if the response rates are fairly high and appropriate auxiliary data is collected to aid weighting.

Much attention should be paid to providing as much and as good auxiliary data as possible. What data may be available, depends on a country; basically these are of 3 types: (i) external information derived from registers and other administrative records and her surveys updated to survey period if possible (micro and aggregated), (ii) internal information about attempts to contact and successfulness of these (incl. reasons for refusals, partial information on non-respondents), (iii) sub-sample of non-respondents with key questions of the usual survey (necessary especially, if serious bias because of high non-response has to be expected) . This information may be used for describing non-response and adjustments for non-response.

Note that both unit and item non-response have to be registered.

Design Effects/ Effective Sample Size

As indicated, a variety of complex sample designs such as multistage stratified and clustered sampling was used in the first round of the ESS and can also be expected to be used in round II. For determining the sample sizes for each country the respective design effects have to be considered to ensure the comparability of estimates with respect to their confidence intervals. The design effect is defined as ratio of the actual variance of a variable under the actual sampling design to the variance computed under the assumption of simple random sampling. The problem is that design effects do not only vary from survey to survey because of their different designs but differ within one survey from question to question. “In general, for a well designed study, the design effect usually ranges from 1 to 3.” (Shackman 2001) It is essential that National Co-ordinators and the fielding organizations analyse the data from round I to calculate appropriate intraclass correlation coefficients for the sample designs used in their countries. The cluster size of the selection units also influences the design effect. It should be chosen as small as possible because: The larger the average cluster sizes are, the lower the effective sample size is and the more interviews have to be conducted to reach the minimum size of 1.500. In that sense a large number of selection units should be the goal with only a few interviews in each. Another important effect is that of departures from equal probability selection methods, that requires design weighting as correction of different inclusion probabilities. In particular, in countries where the only available frames are of households or addresses, design effects will be larger than in countries where frames of persons are available. This fact has also to be considered when computing the sample sizes.

3. Summary

Comparability of sampling means that the national surveys must provide estimates that are subject to minimal bias of parameters of the equivalent populations. The basic requirement to use probability samples together with the additional requirements discussed in this paper leads *theoretically* to comparable estimates. However, in the end the quality of the data depends also on the implementation process, e.g. the *practical* applications[4]. Therefore, this process has to be monitored carefully.

4. Handling of the Work-package

In round I we installed an expert panel on sampling. This panel will continue it´s work.
Members are the following sampling specialists:

- Sabine Häder (Centre for Survey Research and Methodology, Germany)
- Siegfried Gabler (Centre for Survey Research and Methodology, Germany)
- Seppo Laaksonen (University of Helsinki, Finland)
- Peter Lynn (University of Essex, U.K.)

Each of the experts will be assigned about six countries to liase and support. However, the decision to 'sign off' a design will be made together by the whole team.
As a starting point of the assessment of the sampling designs we need the information available from the tenders. The National Co-ordinators should ensure that the questions listed in paragraph 5 can be answered with the help of the tenders. That means that the survey organisations have to be informed by the NCs about these requirements in advance of handing in the tenders. Additionally, we ask the NCs to give their comments to the proposed designs, e.g. to evaluate them with the help of their experience. At least the following points should be treated:

- Is the proposed design good or best practice in the country concerned? If the country is a participant of round 1: Is it the same design or are there improvements implemented?
- Does the survey organisation have experiences with the proposed design (e.g. from round 1 of the ESS)?
- Is the proposed response rate realistic?

[4] Of course, there are also other influences on the data quality besides those because of sampling – such as features of the interviewers, question wording, translation problems and so on. However, we cannot deal with them in this paper.

If the information contained in the tenders and the additional comments of the National Co-ordinators is sufficient the expert panel is enabled to 'sign off' the proposals without delay. If the information is not sufficient, the respective expert will start a dialogue with the National Co-ordinator (and possibly the survey organisation involved) in order to clarify details or propose amendments. If necessary, other sampling specialists in the country concerned will be joined in the discussion, so that their knowledge of local practices, arrangements and vocabulary can be drawn on. Similarly where necessary, the panellist will visit the country to give help and support. These consultations will be conducted as efficiently as possible to give maximum time for the design to be implemented in good time according to the specification.

5. Information need to be contained in the tenders

Answers to the following questions concerning sampling should be given in the tenders from the survey organisations.

Description of the target population

- Are the ESS specifications of the sampling 'universe' adhered to (i.e. all residents aged 15+, regardless of nationality or citizenship, excluding only the homeless and the institutional population)?

Description of the sampling frame

- Is the quality of the proposed sampling frame suited to its proposed purpose (in terms of (coverage, updating, access, etc)?

Detailed (!) description of the sample design

- If lists are to be used, how, if at all, are they to be stratified?
- Is the design single- or multi-stage?
- Which stages are defined?
- How much clustering is proposed?

Sample size

- How has the effective sample size been calculated, including estimates of response rates and design effects due to clustering or necessary weighting?[5]
- Will any population subgroups be over-sampled ?
- What steps will be taken to achieve the target response rate?

The National Co-ordinators are responsible for inquiring the survey organisations about these points. As a result, the assigned sampling expert shall be enabled to fill in the following form (as an example see the Irish form of round I):

[5] For the computation see Appendix 1 of the Specification for participating countries of the ESS.

Country:	Ireland
NC:	Richard Sinnott
Survey Institute:	Economic and Social Research Institute (ESRI)
Expert:	Sabine Häder
Reference study:	Irish Social and Political Attitudes Survey 2002 and others

Target Population, Population coverage	Persons aged 15 years and over who are resident in private households in the Republic of Ireland
Problems	15 year olds will only be interviewed with parental consent

Sampling frame	For stratification and selection of addresses: Computer-based National Electoral Register. This register is updated annually. The version of 2002/03 will be used.

Sampling design	Stratified three-stage probability sampling PSUs are aggregates of District Electoral Divisions (DEDs). There is a total of 3,440 DEDs in Ireland. PSUs are ordered geographically (north-south, east-west). This makes it possible to take a systematic sample giving an implicit stratification. Stage 1: Selection of 220 PSUs proportional to size of population. Stage 2: Systematic sample of 19 to 20 addresses within each PSU. Addresses are sorted by geographical propinquity. Stage 3: Random selection of one individual within each household (Next-Birthday-Method)
Problems	- Given the nature of the sampling frame, which is the only one available in Ireland, one is constrained to use the population aged 18 and over in selecting the PSUs. - The sample from the electoral register is a sample of addresses. In effect this is very close to a sample of households in Ireland as there is a very low level of multi-household occupancy at addresses in Ireland. If there is more than one household behind an address the household to be interviewed is selected via Kish grid (given the nature of the electoral register the chances of more than one household at any address are very small indeed).

Design effects	$DEFF_c = 1 + (11 - 1) * 0.02 = 1.2$ $DEFF_p = 1.33$ $DEFF = 1.6$
Remark	Differential household sizes between PSUs are already included in the calculation of the design effect $DEFF_p$.

Target response rate	About 63%
Problems	The target response rate (70%) will not been reached.

Sample size	A net sample of size $n_{net} = n_{eff} * DEFF = 1,500 * 1.6 = 2,400$ interviews will be conducted. With 10% of ineligibles and a response rate of about 63% the gross sample size must be $n_{gross} = 2,400 / (0.63 * 0.9) = 4,233$. For each cluster 19 to 20 addresses have to be drawn. That results in an average of 11 interviews per cluster.

Special Features of the design	Given the sample design it is necessary to re-weight the data according to household size due to unequal selection probabilities. This is implemented as a standard procedure in our samples and is discussed in the written documentation previously provided by the ESRI.

6. References

ESS-Sampling Expert Panel: The Sample. In . ESS 2002/2003 Technical Report Edition 1 (September 2003). http://naticent02.uuhost.uk.uu.net/archive/tech_report.htm

Kish, L. (1994). Multipopulation Survey Designs: Five Types with Seven Shared Aspects. International Statistical Review 62, 167-186.

Lyberg, L. (2000): Review of IALS – a commentary on the technical report. In: Measuring Adult Literacy. The International Adult Literacy Survey in the European context. London: Office for National Statistics.

Häder, S. and Gabler, S. (1997). Deviations from the population and optimal weights. In: W. Saris and M. Kaase (1997). Eurobarometer - Measurements for opinions in Europe. ZUMA-Nachrichten Spezial No.2.

Shackman, G. (2001). Sample size and design effect. http://www.albany.edu/~areilly/albany_asa/confweb01/agenda01.htm

Appendix 1: Sampling issues in the "Specifications for participating countries"; Round 2 of the ESS

5.1 *Population coverage*

The survey will be representative of all persons aged 15 and over (no upper age limit) resident within private households in each country, regardless of their nationality, citizenship or language[6]. Potential under-coverage of certain groups, say because of language problems or sampling frame deficiencies, or for any other reason, must be discussed with the sampling panel prior to deciding on the final sampling method, so that the problem can if at all possible be remedied.

5.2 *The sample*

The sample is to be selected by strict random probability methods at every stage and respondents are to be interviewed face-to-face. Where a sample frame of individuals is not available, countries may use a sample frame of households or of addresses. In these cases, procedures for selecting a household from a multi-household address (where appropriate), and an individual within a household will be specified and agreed in advance with the sampling panel. In any event, the relative selection probabilities of every sample member must be known and recorded (see Appendix 3), as should any remaining systematic non-coverage problems. Quota sampling is not permitted at any stage, nor is substitution of non-responding households or individuals (whether 'refusals' or 'non-contacts'). Over-sampling of certain subgroups must be discussed and agreed in advance with the sampling panel.

5.3 *Effective sample size*

The **minimum 'effective achieved sample size' should be 1,500**, after discounting for design effects, or 800 in countries with populations of less than 2 million. Thus, with the help of the sampling panel, each country should determine the appropriate size of its initial issued sample by taking into account the realistic estimated impact of clustering, eligibility rates (where appropriate), over-sampling and response rate on the effective sample size. The sampling panel will help to calculate the actual gross achieved sample size required in order to achieve an effective sample size of 1,500 interviews.

[6] Please note that questionnaires are to be available in all languages spoken as a first language by 5 per cent or more of the population and interviewers must be available to administer them (see 5.12). For speakers of certain minority languages (spoken by fewer than 5 per cent of the population), however, it may be possible to adapt the questionnaire produced by another participating country. If National Co-ordinators wish to offer translated questionnaires to these smaller minority language groups, they should refer to the CCT for advice. Countries are not, however, required to interview language minorities under the 5% cut-off.

5.4 *Documentation of sampling procedures*

The precise sampling procedures to be employed in each country, and their implications for representativeness, must be documented in full and submitted in advance to the expert panel for 'signing off' and subsequently to the CCT for reference. This precaution is to ensure that all countries within the ESS have defensible (and equivalent) national probability samples of their resident (aged 15 and over) populations. The following details will be required before the sampling panel can 'sign off' a country's sample design:

- a description of the sampling frame and of the units it comprises (including information on units that might be used either to stratify the sample or to vary probabilities of selection for certain subgroups, and estimates of any likely under-coverage, duplication and ineligibles)
- for those using multi-stage samples, a description of how the units at each stage will be selected to result in a random sample of individuals, plus the inclusion probabilities of units at each stage of selection
- details of whether and how the survey is to be clustered geographically, and how the initial clusters are to be selected
- full details of any stratification to be employed
- the calculations on which the predicted effective sample size has been based.

The final sample design will be fully documented by each national team in the national technical report of the survey. This documentation will be translated into one or more variables within the national data file to indicate the relative selection probabilities of cases and to enable appropriate weighting strategies to be calculated.

5.5 *Target response rates*

Outcomes of all approaches to addresses, households and individuals in the sample will be defined and recorded according to a pre-specified set of categories that distinguish non-eligibility, non-contacts and refusals (see section 5.8). Model 'contact forms' will be produced by the CCT, for translation and use by national teams. Countries may use their own contact forms if they wish, ensuring that these collect data on all of the variables specified by the CCT. Example contact forms can be seen at www.europeansocialsurvey.org/fieldwork/contact_form_address.doc.

The proportion of non-contacts should not exceed 3 per cent of all sampled units, and the **minimum target response rate** - after discounting ineligibles (and other 'deadwood', as defined by the CCT - see section 5.7) - **should be 70%**. As seen in Round 1, this figure is likely to be exceeded in certain countries. Countries that participated in Round 1 and achieved lower response rates will nevertheless be expected to aim for the same 70% target in Round 2. Survey organisations should thus cost their surveys with this response rate in mind and consider what steps may be required to achieve it.

Appendix 2: Rules for estimating design effects

Effective Sample Size

The effective sample size (neff) is the size of a simple random sample which would produce the same precision (standard errors) as the design actually used. Typically, neff is less than the actual number of achieved interviews, m, as certain aspects of survey design - for example, clustering or the use of differing selection probabilities - tend to reduce the precision of estimates. The reduction of precision is known as the design effect (DEFF):

DEFF = Actual sampling variance / Sampling variance with simple random samples of same size;

DEFF = m/neff, so neff = m/DEFF

We therefore need to be able to predict the value of DEFF for a proposed sample design, in order to determine how many interviews should be achieved so as to produce a particular value of neff. We suggest that two components of *DEFF* should be taken into account at the design stage - the design effect arising from differing selection probabilities ($DEFF_p$) and the design effect arising from clustering ($DEFF_c$) . Then $DEFF = DEFF_p$ x $DEFF_c$. We then also need to predict the survey response rate (and the proportion of ineligibles on the sampling frame, if relevant) in order to determine the size of the initial sample (n) required in order to achieve approximately m interviews.

Design Effects due to Differing Selection Probabilities

In some countries which have accessible population registers, it will be possible to select an equal-probability sample from the survey population. In other countries, it will be necessary to select the sample in stages, with the penultimate stage being residential addresses. In this case, each person's selection probability will depend on their household size. Another reason why differing selection probabilities might be used is if important minority groups were to be over-sampled.

If differing selection probabilities are to be used - for whatever reason - the associated design effect should be predicted. This can be done very simply, using the following formula

$$DEFF_p = \frac{m(\sum_i m_i w_i^2)}{(\sum_i m_i w_i)^2}$$

where there are m_i respondents in the i^{th} selection probability class, each receiving a weight of w_i, where α means 'proportional to'

$$w_i \alpha \frac{N_i}{m_i}$$

(This formula assumes that the population variance of survey variables will not vary over selection probability classes - a reasonable assumption in most situations)

Design Effects Due to Clustering

It is anticipated that in most countries it will be efficient to select a multi-stage, clustered, sample. In such situations there will also be a design effect due to clustering:

$DEFF_c = 1 + (b\text{-}1)\ \rho$

where b is the mean number of respondents per cluster and ρ is the intra-cluster correlation (or "rate of homogeneity") - a measure of the extent to which persons within a clustering unit are more homogeneous than persons within the population as a whole (see Kish, 1994, Survey Sampling, pp. 161-164 (New York: Wiley and Sons, Inc.)). This design effect can be estimated, at least crudely, from knowledge of other surveys and/or the nature of the clustering units.

In practice, all elements of the overall design effect, including that due to differing selection probabilities and that due to clustering, will take different values for different survey estimates. For sample design purposes, an average value should be used.

Example: How to determine the size of issued sample

We have prescribed neff > 1500.

To determine m, we must first estimate $DEFF = DEFF_p$ x $DEFF_c$

1. Suppose the proposed clustering units are administrative areas of around 5,000 households on average and that based on data from other surveys, we expect that for these areas, ρ will take values of around 0.02 for many variables. Then, if we are proposing a design with a mean of 15 interviews per cluster:

$DEFF_c = 1 + (15 - 1)$ x $0.02 = 1.28$.

[Note: If there is no available empirical evidence at all upon which to base an estimate of Δ , then we suggest that a value of 0.02 should be used.]

2. Suppose that the only available sampling frame is a list of addresses and that these must be selected with equal probabilities. The proposed design is then randomly to select one person to interview at each address. This is the only aspect of the proposed design that involves differing selection probabilities. Then, we can use population statistics on the distribution of household size to estimate the number of respondents in each selection probability class, thus:

No. of persons aged 18+ in household i	**Proportion of households in population H_i/H**	**No. of achieved interviews m_i**	**Relative weight**		
			w_i	**m_iw_i**	**$m_iw_i^2$**
1	0.35	0.35m	1	0.35m	0.35m
2	0.45	0.45m	2	0.90m	1.80m
3	0.12	0.12m	3	0.36m	1.08m
4	0.06	0,06m	4	0.24m	0.96m
5	0.02	0.02m	5	0.10m	0.50m
				1.95m	4.69m

The population distribution of household size appears in the first two columns. From this, we can predict that the sample distribution will be as shown in the third column. We can thus predict $DEFF_p$:

$$DEFF_p = m \times 4.69m/(1.95m)^2 \quad = \quad 4.69/1.95^{\,2} \quad = 1.23$$

3. Thus, we predict DEFF = 1.28 x 1.23 = 1.57. Consequently, to achieve neff > 1,500 with this design, we would need m > 1,500 x 1.57 = 2,355.

4. The final stage is to calculate the sample size to select initially in order to be likely to achieve around 2,355 interviews. Suppose we anticipate a response rate of 80% and that 5% of the sampling frame units will be ineligible (e.g. addresses which do not contain a resident household), then:

n = (m / 0.80) / 0.95 = 3 ,098

So we would select a sample of at least 3,100 addresses.

Appendix 3: European Social Survey Round II – Sample Design Data File

This is the description of the data file needed regarding the sample design.

The file should contain one record for each selected address, in the case of address-based samples, or for each selected individual, in the case of individual-based (population register) samples. (In the rare event of a design in which where it is possible to select more than one household at an address, extra records will be needed for any extra households thus selected. At the time of writing, the only design of this sort is the Netherlands.)

Column 1: Name of the variable
Column 2: Short description of the variable - including logical relationship with other variables.
Column 3: Source of the variable.
S = initial office-based sample selection process
F = field-based sampling process (interviewers)
Column 4 : Reason for variable
W = For the construction of design weights
SE = For estimation of design effects and standard errors

The annex contains examples of how the variables should be interpreted for different types of sample design. It is suggested that you look closely at these examples. National co-ordinators may seek advice from their nominated member of the sampling expert panel if necessary.

Note on missing values: Use code 9 for items that are missing through interviewer error. This should apply only to certain items (F in column 3) in the case of address-based sampling. Use 'blank' for items that are not applicable (e.g. 3rd-stage probabilities, in the case of a 2-stage design).
We assume that SPSS or SAS system files will be delivered.

NAME of variable	*MEANING OF VARIABLE (label) and codes*	*I*	*Info*
IDNO	Address number *(same as in main datafile)* Code = number	X	SE
IDEXTRA	Extra code to distinguish between households/persons selected at the same address (*applies only to the Netherlands*) Code = 1, 2, … (within IDNO)	X	SE
PROB1	Selection probability at first stage of sampling. Must be numeric > 0, with 12 decimal places (or as many as necessary to provide at least 4 significant figures for all cases). For multi-stage samples, all records with the same value of PSU should have the same value of PROB1.	S	W
PROB2	Only to be used if there are at least 2 stages of selection. Otherwise, PROB2 should be blank. Selection probability at second stage of sampling (conditional upon first stage). Must be numeric > 0, with 12 decimal places (or as many as necessary to provide at least 4 significant figures for all cases). All records with the same combination of PSU/SSU should have the same value of PROB2.	S or F	W

PROB3	Only to be used if there are at least 3 stages of selection. Otherwise, PROB3 should be blank. Selection probability at third stage of sampling (conditional upon first two stages). Must be numeric > 0, with 12 decimal places (or as many as necessary to provide at least 4 significant figures for all cases).	F	W
PROB4	Only to be used if there are 4 stages of selection. Otherwise, PROB4 should be blank. Selection probability at fourth stage of sampling (conditional upon first 3 stages). Must be numeric > 0, with 12 decimal places (or as many as necessary to provide at least 4 significant figures for all cases).	F	W
PSU	To be used if there is more than 1 stage of selection. For 1-stage designs, PSU should be blank. Indicator of the first stage unit (primary sampling unit) to which the sample address/person belongs. Integral. Note that a key to this variable is NOT necessary. In other words, the PSUs can be completely anonymous.	S	SE
SSU	Only to be used if there are at least 3 stages and more than one 3rd-stage unit is selected within each 2nd-stage unit (only Netherlands?). Otherwise, SSU should be blank. Indicator of the second stage unit (secondary sampling unit) to which the sample person belongs. Integral. Need only be unique within PSU.	S	SE
STRATEX1	Only to be used if explicit stratification was used to select the PSUs (independent selections from each stratum). Otherwise, STRATEX1 should be blank. Indicator of stage 1 explicit stratum to which the sample address/person's PSU belongs. All records with the same value of PSU should have the same value of STRATEX1.	S	SE
STRATIM1	Only to be used if implicit stratification was used to select the PSUs (systematic selections - possibly PPS - from an ordered list). Otherwise, STRATIM1 should be blank. Indicator of order on the list (order of selection) of the sample address/person's PSU. All records with the same value of PSU should have the same value of STRATIM1. All records with the same value of STRATIM1should have the same value of PSU.	S	SE
STRATIM2	Only to be used if there were at least 3 stages and if implicit stratification was used to select the SSUs (systematic selections from an ordered list). Otherwise, STRATIM2 should be blank. Indicator of order on the list (order of selection) of the sample address/person's SSU within the PSU. All records with the same combination of PSU/SSU (if any) should have the same value of STRATIM2.	S	SE
STRATVAL1	Only to be used if implicit stratification was used to select the PSUs. Otherwise, STRATVAL1 should be blank. The numeric value of the (first) variable used to stratify PSUs.	S	SE
STRATVAL2	Only to be used if implicit stratification was used to select the PSUs, based upon a combination of at least two variables. Otherwise, STRATVAL2 should be blank. The numeric value of the second variable used to stratify PSUs.	S	SE

STRATVAL3	Only to be used if implicit stratification was used to select the PSUs, based upon a combination of at least three variables. Otherwise, STRATVAL3 should be blank. The numeric value of the third variable used to stratify PSUs.	S	SE

Annex: Examples

Example 1: Single-stage unclustered sample of persons selected from population register, stratified by region.

PROB1 = n/N for all records (n is number of persons selected, N is total eligible number on the register);
STRATEX1 = region indicator;
PROB2, PROB3, PROB4, PSU, TSU, STRATIM1, STRATIM2, STRATVAL1, STRATVAL2, STRATVAL3 blank for all records.

Example 2: Clustered sample of addresses within PSUs, with field selection of one person per address. PSUs selected PPS from an ordered list (order of population size within regions). Equal number of addresses selected from each PSU, systematically from a geographically-ordered list.

Notation: *m* PSUs selected out of *M* in population;
N_i addresses in i^{th} PSU, $\sum_{i=1}^{M} N_i = N$;
total of *n* addresses selected, i.e. *n*/*m* in each of *m* PSUs;
total number of eligible persons at kth address is μ_k.

PROB1 = $\frac{m \times N_i}{N}$

PROB2 = $\frac{n}{m \times N_i}$

PROB3 = $\frac{1}{\mu_k}$

PSU = 1, 2, 3, ….. m;
STRATIM1 = 1, 2, 3, ….. m in order PSUs were selected;
STRATIM2 = 1, 2, 3, ….. n/m in order addresses were selected within PSU;
STRATVAL1 = region indicator;
STRATVAL2 = population size of the PSU;
PROB4, SSU, STRATEX1, STRATEX2, STRATVAL3 blank for all records.