DACSEIS
IST-2000-26057

Workpackage 1
Data Quality in Complex Surveys

Deliverable 1.1

2003

List of contributors:

Wolf Bihler, Statistisches Bundesamt
Jan Bjørnstad, Li-Chun Zhang, Statistics Norway
Anthony Davison, Sylvain Sardy, EPFL
Alois Haslinger, Statistics Austria
Paul Knottnerus, Statistics Netherlands
Seppo Laaksonen, Statistics Finland
Ralf Münnich, Dominik Ohly, Josef Schürle, Rolf Wiegert, University of Tübingen
Ueli Oetliker, Jean-Pierre Renfer, OFS-Office fédéral de la Statistique Suisse
Andreas Quatember, JKU/IFAS Linz
Chris Skinner, Yves Berger, University of Southampton

Main responsibility:

Ralf Münnich, University of Tübingen

IST–2000–26057–DACSEIS

The DACSEIS research project is financially supported within the IST programme of the European Commission. Research activities take place in close collaboration with Eurostat.

http://europa.eu.int/comm/eurostat/research/

http://www.cordis.lu/ist/

http://www.dacseis.de/

Preface

One major aim of the DACSEIS project is to produce a recommended practice manual on the measurement of accuracy on estimates from complex surveys. However, this is strongly related to the investigation of variance estimation methods in the context of complex surveys partially considering non-response with different non-response quotes.

Deliverable D1.1 is aimed at giving an overview to the methodology to be considered, the problems in practice - under strong consideration of the DACSEIS orientation - with respect to household- and individual surveys, as well as to the basic set-ups for the simulation study that will be fully carried out in deliverable D1.2 in connection with the methodological workpackages WP5 - WP11.

Since almost all parties are involved in this deliverable, the responsibility for the different parts of the document are split as follows:

1.	Anthony Davison, Chris Skinner
2.	Ralf Münnich
3.	Ralf Münnich, Rolf Wiegert
4.	Ralf Münnich, Rolf Wiegert
5.	Seppo Laaksonen
5.1	Andreas Quatember, Alois Haslinger
5.2	Paul Knottnerus
5.3	Seppo Laaksonen
5.4	Wolf Bihler
5.5	Ueli Oetliker, Jean-Pierre Renfer
5.6	Jan Bjørnstad
A	Ralf Münnich
B	Ralf Münnich, Rolf Wiegert

I would like to personally thank all colleagues of the project for highly valuable comments and the wonderful cooperation to fulfill the aims of the project. Further thanks go to Jana Meliskova, UN/ECE, for the support regarding the implementation of metadata, as well as the members of the *Task Force Variance Estimation*, Eurostat, for helpful discussions on user needs.

Ralf Münnich Eberhard–Karls–University Tübingen.

Contents

List of Figures

List of Tables

Chapter 1

Classification of Variance Estimation Methods

1.1 Direct methods

A method of variance estimation depends fundamentally upon:

- the point estimator; and
- the distribution with respect to which the variance is defined.

In the classical design-based approach to survey sampling, the distribution is defined with respect to the randomisation mechanism generating the sample. Thus the variance is interpreted with respect to the distribution of the point estimator over repeated samples generated by the given sampling design.

In this section, we restrict attention to this classical case, but in general it may be appropriate to consider augmenting or replacing this sampling distribution by other distributions, for example:

non-response — the distribution induced by an assumed stochastic non-response model, often in combination with the randomisation mechanism generating the sample;

imputation — the distribution induced by a stochastic mechanism generating imputed values, often in combination with the mechanisms generating non-response and the sample;

conditional variance — the distribution conditional on ancillary statistics, such the sample sizes within poststrata; or

model variance — the distribution with respect to an assumed model; this might be conditional upon the outcomes of the sampling and non-response mechanisms or the distribution might reflect the combined outcomes of the model and these mechanisms.

Returning to the classical case, it is natural to classify variance estimation methods by both the point estimator and the sampling design. For any given estimator and sampling design, classical design-based sampling theory may be used to determine the variance of the estimator and to obtain an expression for an approximately unbiased estimator of this variance. This leads to what is called here the *direct method* of variance estimation. This approach is primarily applied to linear point estimators and we consider variance estimation for non-linear estimators in Sections 1.2 and 1.3. A *linear point estimator* takes the form

$$\widehat{\theta} = \sum_{s} w_i y_i, \tag{1.1}$$

where s denotes the sample, y_i the value of the survey variable and w_i the weight for the i-th unit. An important example is the Horvitz–Thompson estimator of the population total, where $w_i = \pi_i^{-1}$, the reciprocal of the first order inclusion probability. A number of direct methods have been proposed for general sampling designs, for example the Sen–Yates–Grundy estimator of the variance of the Horvitz–Thompson estimator $\widehat{\theta}_{HT}$:

$$\widehat{V}\left(\widehat{\theta}_{HT}\right) = \sum\sum_{i<j \in s} \frac{\pi_i \pi_j - \pi_{ij}}{\pi_{ij}} \left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j}\right)^2. \tag{1.2}$$

Direct methods of variance estimation are most simply obtained for linear estimators where the weights w_i are fixed, that is, do not depend upon the sample s. It is possible, however, to obtain general variance estimation procedures for the wider class of linear estimators, where the weights w_i may depend upon s but not upon the survey variable. This more general class of linear estimators includes, for example, the ratio estimator, often considered as non-linear. It is possible to express such variance estimators, as in (1.2), as quadratic forms in the y_i values, where the coefficients of the squares and products of the y_i depend only upon the design and the weights and not on the survey variable RAO (1988).

The application of such general results to specific estimators and sampling designs can be complex in practice, in particular because of difficulties in determining second order inclusion probabilities. The latter issue is addressed in workpackage 6, where some simplified variance estimators, dependent only upon first-order inclusion probabilities, have been researched. It remains useful to consider variance estimators tailored to different specific sampling schemes.

Stratified sampling is usually relatively straightforward to handle with linear estimators if variances can be estimated satisfactorily within strata and then summed across strata.

Multistage designs can represent a greater challenge. If, as is common, sampling takes place without replacement at each stage, then expressions for exactly unbiased variance estimators will typically involve components for each stage. A recursive approach to variance estimation can be constructed, based upon variance estimation approaches appropriate at each stage of sampling (cf. DURBIN, 1953, RAJ, 1966, and DEVILLE, 1999). Simplified procedures are available under certain special multistage designs (e.g. DURBIN, 1967) and a variety of approximations are available. A widely used approach is to approximate sampling at the first stage by with replacement sampling leading to a simple variance estimator RAO (1988). The resulting overestimation is often considered negligible.

Special attention has also been given to the development of variance estimators for two-phase and multi-phase sampling (Fuller, 2003).

1.2 Linearization methods

Direct methods of variance estimation apply to classes of linear statistics, but many point estimators of interest in official statistics fall outside these classes. Linearization methods enable direct methods to be extended to non-linear estimators. Below we first describe linearization using Taylor series, and then outline a generalization most useful for complex non-linear estimators.

As discussed in more detail by Wolter (1985), Ch. 6, we consider a non-linear estimator, which may be expressed as a differentiable function, $\widehat{\theta} = g(T)$, of a vector of linear estimators T, where a direct method may be used to obtain a satisfactory estimator, $\widehat{V}(T)$, of the variance-covariance matrix of T. Letting $\nabla_g(\cdot)$ be the (vector) derivative of the function $g(\cdot)$, the basic linearization variance estimator takes the form

$$V_L(\widehat{\theta}) = \nabla_g(T)^{\mathrm{T}}\,\widehat{V}(T)\nabla_g(T)\,. \tag{1.3}$$

An alternative approach, which in practice will usually lead to the same variance estimator, is to approximate the variance of the non-linear estimator by the variance of a linear statistic and then use a direct method for that statistic. If T may be expressed as $\sum w_i y_i$, where y_i is now a vector, then the approximating linear statistic for $\widehat{\theta} = g(T)$ will be of the form $\sum w_i z_i$ where $z_i = \nabla_g(T)\ y_i$ is treated as a fixed variable Woodruff (1971). Sometimes it is not possible to express the function $g(\cdot)$ in closed form, for example when the point estimator is defined as the solution of an estimating equation or the limit of an iterative estimation method. Rao (1988), Binder (1996) and Deville (1999) discuss the use of linearization methods in such cases of 'implicit parameters'.

A similar but more general approach to construction of linearization variance estimators for a compactly differentiable statistic $\widehat{\theta}$ is through what are commonly called the influence function components Hampel *et al.* (1986), which may be interpreted as derivatives of $\widehat{\theta}$ with respect to changes in the weights placed on individual observations; for this reason the technique is sometimes known as the infinitesimal jackknife. Here the term influence function is best avoided, because the influence of an observation on an estimator depends also on the weight attached to the observation. The extra generality stems from use of a von Mises rather than Taylor series expansion of the statistic of interest, enabling theoretical variance formulae to be obtained for estimators such as the sample median and other quantiles. In the case of a simple random sample of size n, the empirical version of the resulting jackknife linearization variance formula is

$$n^{-2}\sum_{j=1}^{n} l_j^2, \tag{1.4}$$

where l_j, the so-called empirical influence value, is the empirical version of the derivative corresponding to the jth observation. When the statistic considered has form (1.1), for example, we find that $l_i = nw_iy_i - \sum_s w_jy_j$, and (1.4) reduces to the familiar

$$(1-f)\frac{1}{(n-1)n}\sum_{i=1}^{n}\left(nw_iy_i - \sum_s w_jy_j\right)^2, \tag{1.5}$$

apart from replacement of n^{-2} by a quantity that corrects the downward bias of (1.4) and includes a finite population correction. Apart from this change, this approach yields many familiar variance estimators, but by putting the linearization idea in a more general context makes extensions to robust and other implicitly-defined estimators more direct and transparent. DAVISON and HINKLEY (1997), Chapter 2, give a brief overview and illustrate the computation of empirical influence values, while theoretical justification of jackknife linearization using notions of functional differentiation may be found in CAMPBELL (1980); see also FERNHOLTZ (1983) or VAN der VAART (1998), for example.

One important way in which non-linear estimators arise in official statistics is through the use of calibration methods, such as regression estimation. The simplest such estimator has form (1.1), but as the weights depend on the sampling scheme, though not generally on the survey variable, they must be treated as random variables. Such estimators arise for various reasons in addition to calibration, another example being imputation to account for case non-response. DEVILLE (1999) shows how the application of Woodruff's method to regression estimation corresponds to the calculation of regression residuals.

As comparison of expressions (1.4) and (1.5) suggests, linearization does not define a unique variance estimator, although all alternatives should be equivalent asymptotically within the asymptotic framework with which this method is derived. Some explicit general approaches to the choice of method, avoiding this indeterminacy, are given by BINDER (1996) and DEVILLE (1999); a standard approach is to aim to achieve an approximately unbiased variance estimator, as in (1.5).

Deliverable D5.2 contains formulae and code for implementation of jackknife linearization with calibration and regression imputation in stratified surveys, which should be implemented for comparison with other approaches discussed below.

1.3 Resampling methods

In some ways the simplest approach to resampling is the jackknife, for which the estimator $\widehat{\theta}$ is recomputed based on samples from which each observation has been deleted in turn. This corresponds to numerical differentation, and in principle should produce variances similar to those obtained by jackknife linearization, or linearization in the case of statistics based on averages. In practice the variance estimates thus obtained tend to be somewhat larger than those found using linearization, partly because linearization variances involve exact derivatives of the functional. The computational costs are quite different, as the basic leave-one-out jackknife requires as many recalculations as there are units in the sample. This is clearly infeasible for surveys of any appreciable size, for which we suggest using the grouped jackknife with replacement (SHAO and TU, 1995, Section 5.2.2), with a total number of recalculations R taken to give a similar computational burden as that for the bootstrap, discussed below. The group sizes should be kept as equal as possible,

d, say, so we should have $Rd = n$, where n is the total sample size. The implementation details for these jackknives will depend on the particular survey under simulation, and in particular the sizes of their strata. It may be necessary to generalize the delete-d jackknife variance estimator to deal with varying group size d, in cases where the number of strata exceeds R. The grouped jackknife without replacement underestimated the variances appreciably in a pilot study, but it is worth incorporating it into at least some of the simulations, to see if the underestimation is quite systematic. It seems straightforward to allow for deterministic imputation of missing values, and for calibration, and the approach to doing so is discussed in deliverable D5.2.

Variant jackknifes can be defined for more complex sampling plans, for example by jackknifing the primary sampling units in cluster sampling. The key idea is to seek to drop quantities which are statistically independent, or approximately so, and the choice of these will depend upon the sampling plan.

Balanced half–sampling (McCarthy, 1969) is the simplest form of balanced repeated replication. When a sample consists of two observations in each stratum, then a half–sample is formed by taking one observation from each of the H strata. The statistic can be recalculated 2^H times, and the results combined to estimate the variance of the original statistic. When H is large, balanced half–sampling is computationally intensive, but ideas from experimental design allow fewer recalculations, say $L < 2^H$, resulting in the same result for linear statistics as would computation of all 2^H replicates. The simplest approach when each stratum has more than two observations is to split the stratum randomly into two groups of nearly equal size, and we recommend that this be applied for the simulation studies.

Shao *et al.* (1998) adjust balanced repeated replication to the presence of non-response, by taking into account a deterministic or random imputation mechanism, and their proposal should be implemented for comparison with other approaches.

The bootstrap is the most simply implemented of the resampling approaches. We propose that for data sets in which the sampling fraction in each stratum is lower than 1/20, the usual nonparametric bootstrap should be used, with stratification corresponding to the sampling strata, and with about $R = 200$ replications. Further stratification will be needed for variances of estimators for change, to allow for three artifical strata within each sampling stratum: one stratum for individuals present on both occasions of sampling, one for those present on the first occasion, and one for those present on the second occasion. It seems quite feasible to use the control variable idea from Davison *et al.* (1986) (see also Davison and Hinkley, 1997, Section 9.3) to reduce simulation error, and thereby to reduce substantially the amount of bootstrap simulation required.

Chapter 2

Criteria for evaluating estimation and variance estimation methods

The aim of this chapter is to give an overview to accurately measuring estimators and variance estimators. Special emphasis will be placed on the applicability and use of the measures within the simulation study. However, three different cases will be distinguished:

1. Theoretical measurement of the accuracy of estimators;
2. Theoretical measurement of the accuracy of variance estimators;
3. Empirical measurement of the accuracy of estimators.

Additional efforts have to be undertaken with respect to non-response and correction for non-response.

2.1 Theoretical measures of accuracy

The main characteristics of the measurement of accuracy in the context of *recommended practice* is the comparison of different estimators and in DACSEIS also of variance estimators. Therefore, it is important to adequately compare distributions of estimators in practical use. This can be done either by the investigation of special measures or by applying direct comparisons between estimators. Since these measures will be applied to the simulation study, the measures will be presented with respect to the R runs of the simulation study.

One major measure of interest is the mean square error MSE of an estimator $\widehat{\theta}$ for the population parameter θ which is defined as:

$$\mathrm{MSE}\,(\widehat{\theta}) = \mathrm{E}(\widehat{\theta} - \theta)^2 \quad . \tag{2.1}$$

This can be split into

$$\mathrm{MSE}\,(\widehat{\theta}) = \mathrm{var}(\widehat{\theta}) + (\mathrm{E}(\widehat{\theta}) - \theta)^2 \quad , \tag{2.2}$$

where $\mathrm{E}(\widehat{\theta}) - \theta$ denotes the bias of the estimator $\widehat{\theta}$.

According to SÄRNDAL *et al.* (1992), p. 41, the MSE will appropriately compare between several different estimators $\widehat{\theta}_1, \widehat{\theta}_2, \ldots$ for one and the same parameter θ. From representation (2.2) follows that large biases or variances will considerably influence the MSE. However, if the MSE of an estimator $\widehat{\theta}$ is small, the estimated value is likely to be close to the true value.

The $\mathrm{Bias}\,\widehat{\theta}$ and $\mathrm{var}\widehat{\theta}$ as parts of the MSE in (2.2) can themselves be applied as measures of accuracy, taking only one important aspect into consideration. However, in practice it is likely to have no information or adequate estimate of the bias which then may lead to use the variance only.

In a simulation study, when the universe is fixed, the true values can easily be determined. Within the simulation study, the $\mathrm{MSE}\,(\widehat{\theta}\,)$ will then be approximated by

$$\mathrm{MSE}\,(\widehat{\theta}\,) \doteq \frac{1}{R}\sum_{r=1}^{R}\left(\widehat{\theta}_r - \frac{1}{R}\sum_{i=1}^{R}\widehat{\theta}_i\right)^2 + \left(\frac{1}{R}\sum_{r=1}^{R}\widehat{\theta}_r - \theta\right)^2 \quad , \tag{2.3}$$

where $\widehat{\theta}_r$ is the r-th estimate from R runs in the simulation.

As a quadratic form, the MSE surely is an appealing measure with brilliant mathematical properties with regards to optimisation (cf. GABLER and STENGER, 2000, pp. 305 for optimal sampling plans). Nevertheless, from statistical decision theory, apart from equal weighting and quadratic terms other functions on standard deviation and bias could be considered which may lead to

$$\Psi(\widehat{\theta}\,) = g \cdot \left(\sqrt{\mathrm{var}\,\widehat{\theta}}\right)^{\alpha} + (1-g)\cdot(\mathrm{Bias}\,\widehat{\theta}\,)^{\alpha} \quad , \tag{2.4}$$

with $g \in (0;1)$, $\alpha > 0$, and Ψ is a general decision functional. However, the practical impact on these general measures are not clear. For $\alpha = 1$ and suitable g, one may refer to the idea of interval estimation which will be investigated later in the text. Further investigation of the general functional Ψ seems sensible since a fine distinction between standard deviation and bias of an estimator can be easily achieved.

All of the above measures can be transformed principally to relative measures with respect to the true value θ or its expected estimate. However, we differentiate between the following cases:

1. As a generalisation of the coefficient of variation to estimators, we use

$$\mathrm{CV} = \frac{\sqrt{\mathrm{var}\,\widehat{\theta}}}{\mathrm{E}\widehat{\theta}}$$

2. and the equivalent for biased estimators

$$\mathrm{CV}^* := \frac{\sqrt{\mathrm{MSE}\,\widehat{\theta}}}{\mathrm{E}\widehat{\theta}} = \frac{\mathrm{RMSE}\,\widehat{\theta}}{\mathrm{E}\widehat{\theta}}$$

where RMSE denotes the root mean square error.

Both relative measures could be referenced to the true value θ which yield the other two relative measures. However, in a comparison the denominator would cancel out. Further, the remaining two cases are equivalent for unbiased estimators.

In a simulation, we apply

$$\mathrm{CV}^* = \frac{\sqrt{\frac{1}{R}\sum_{r=1}^{R}\left(\widehat{\theta}_r - \frac{1}{R}\sum_{r=1}^{R}\widehat{\theta}_r\right)^2 + \left(\frac{1}{R}\sum_{r=1}^{R}\widehat{\theta}_r - \theta\right)^2}}{\frac{1}{R}\sum_{r=1}^{R}\widehat{\theta}_r} \quad . \tag{2.5}$$

When turning to interval estimation, one should consider the coverage rate of confidence intervals. Assuming that an estimator is at least approximately unbiased and normal, we could apply for the confidence interval

$$\left[\widehat{\theta} - z_{1-\alpha/2}\cdot\sqrt{\mathrm{var}\widehat{\theta}}\,;\,\widehat{\theta} + z_{1-\alpha/2}\cdot\sqrt{\mathrm{var}\widehat{\theta}}\right] \quad . \tag{2.6}$$

In practice, however, the variance has to be estimated from the sample which gives

$$\left[\widehat{\theta} - z_{1-\alpha/2}\cdot\sqrt{\widehat{\mathrm{var}}\widehat{\theta}}\,;\,\widehat{\theta} + z_{1-\alpha/2}\cdot\sqrt{\widehat{\mathrm{var}}\widehat{\theta}}\right] \quad . \tag{2.7}$$

Applying biased estimators, the derivation of a confidence interval is more complex. Following SÄRNDAL *et al.* (1992), pp. 164, under the assumption of at least approximately normally distributed estimates, the standard deviation of the estimator can be substituted by the root mean square error which yields as an equivalent to equation (2.6)

$$\left[\widehat{\theta} - z_{1-\alpha/2}\cdot\sqrt{\mathrm{MSE}\,\widehat{\theta}}\,;\,\widehat{\theta} + z_{1-\alpha/2}\cdot\sqrt{\mathrm{MSE}\,\widehat{\theta}}\right] \quad . \tag{2.8}$$

However, this can deliver only approximations to the confidence intervals which work the better the less biased an estimator is. With rising bias relative to the standard deviation of the estimator, the relative coverage of confidence intervals which should be $(1-\alpha)\cdot 100\%$ decreases from the correct value. Applying the bias-ratio which is defined as

$$\mathrm{BR}\,(\widehat{\theta}\,) := \frac{\mathrm{Bias}\,\widehat{\theta}}{\sqrt{\mathrm{var}\,\widehat{\theta}}} \quad , \tag{2.9}$$

one can obtain in the case of a standard normal distribution the relation according to figure 2.1.

According to SÄRNDAL *et al.* (1992), p. 165, the loss of confidence interval coverage rate can be neglected for $|\mathrm{BR}(\widehat{\theta}\,)| \leq 0{,}1$. Even for somewhat higher values of the bias-ratio, the above negative effect is relatively small.

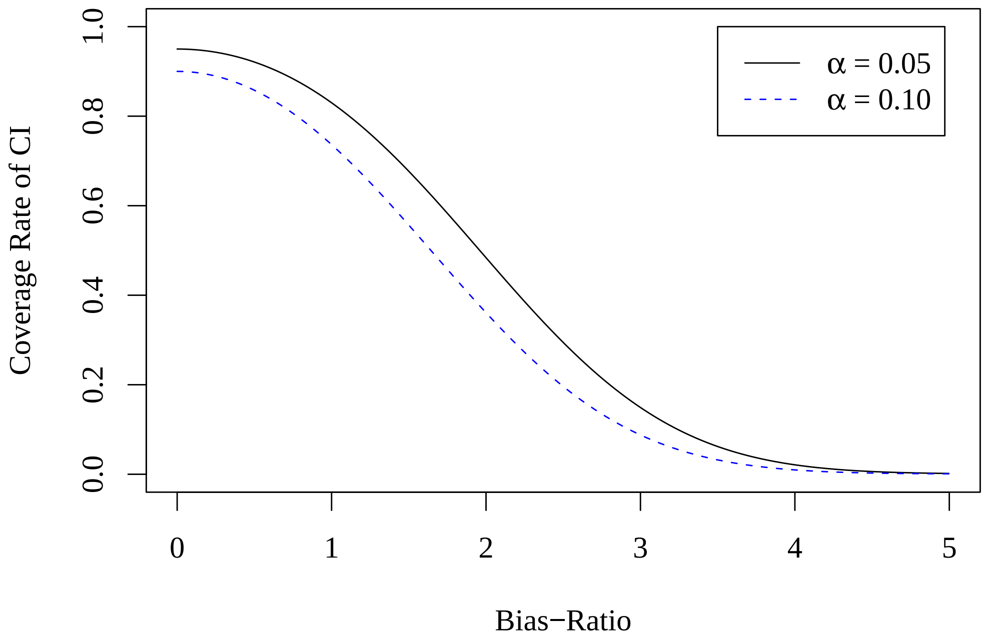


Figure 2.1: Relation of confidence interval coverage rate with respect to the bias-ratio

The coverage rate of confidence intervals is then

$$\mathrm{CR}_\alpha(\widehat{\theta}\,) \doteq \mathrm{P}\left(\widehat{\theta} - z_{1-\alpha/2} \cdot \sqrt{\mathrm{var}\widehat{\theta}} < \theta < \widehat{\theta} + z_{1-\alpha/2} \cdot \sqrt{\mathrm{var}\widehat{\theta}}\right) \quad . \tag{2.10}$$

Within the simulation study one should use

$$\widehat{\mathrm{CR}}_\alpha(\widehat{\theta}\,) = 1 - \frac{1}{R}\sum_{r=1}^{R} \delta\left\{\left(\widehat{\theta} - z_{1-\alpha/2} \cdot \sqrt{\mathrm{var}\widehat{\theta}} > \theta\right) \vee \left(\widehat{\theta} + z_{1-\alpha/2} \cdot \sqrt{\mathrm{var}\widehat{\theta}} < \theta\right)\right\} . \tag{2.11}$$

Remarks:

1. The non-coverage of confidence intervals may be asymmetric due to a correlation between point estimator and variance estimator. The confidence interval boundaries can then be improved by applying an adequate transformation which stabilises the variance estimate (cf. DAVISON and HINKLEY, 1997, p. 195).

2. Applying the MSE in equation (2.8) causes an unnecessarily large confidence interval with an asymmetric non-coverage in the direction of the bias. This possible disadvantage could be avoided by applying a bias corrected estimate. In practice, however, this may not be straightforward due to the difficulty in adequately estimating biases. Further, it is not clear how estimated bias influences the loss of coverage rate.

Additionally to the above measures, higher moments could be applied such as the skewness and curtosis which seem more adequate as measures of normality. Further, as more robust measures, one could also apply the median, the mean absolute deviation, the inter-quartile deviation, or more generally, coefficients of dispersion and their corresponding relative measures.

2.2 Measurement of accuracy of variance estimates

The same methodology can now be applied to variance estimators. However, it seems difficult to estimate the variance of the variance. Further, it is not transparent how the accuracy of estimates is influenced by the properties of the variance estimates when considering higher moments of the variance estimate. Assuming an estimator is biased, it is more appropriate to apply the measures to the mean square error estimates if they are available.

However, in order to compare the accuracy of complex designs with simple random sampling, on may apply the design effect of a design with a given estimator.

2.3 Empirical measurement of the accuracy

In practice, several of the above measures can not or can only be approximately determined like the bias and the mean square error due to the unknown true value. In these cases, estimates of the measures have to be gained from the sample.

Within the simulation study, the aim is to investigate in addition to the quality of estimators and variance estimators, the accuracy of the estimates of the measures.

A comparison of methods under *realistic* conditions will be enabled by the simulation study since the true values will be available as well as the models used for non-response. Following the general simulation scheme (cf. final report of workpackage 3), all parameters and measures of interest can be drawn from the sample and the universe to be able to evaluate the methodology in detail.

2.4 Methods for use under non-response

In principle, the same measures will be applied to either imputed samples or with methods that use weighting to compensate for non-response. However, additionally to the standard programme of measures, the neff may be considered in selected cases (cf. Chapter 3 of deliverable D11.1).

Chapter 3

Overview of individual and household surveys in European Official Statistics

3.1 The DACSEIS questionnaire

3.1.1 Introduction

One of the main goals of DACSEIS is to generate a recommended practice manual for variance estimation methods with special emphasis on individual and household surveys. In order to achieve the most valuable information for this manual, the correct settings for the supporting simulation study had to be investigated. In addition to the set-ups of the available surveys that are described in workpackages 2 and 3, further information of interest had to be obtained. The main interest in terms of the simulation study are the designs of European individual and household surveys, their corresponding non-response rates as well as the current usage of variance estimation methods in these surveys. To obtain this information, a questionnaire was developed and distributed via the working group of *assessment of quality in statistics* at Eurostat. The questionnaire is attached in Appendix B.1.

Since the simulation study will mainly be performed on labour force surveys (LFS) including the microcensuses and household and budget surveys (HBS), the output of the questionnaire will be split with respect to LFS, HBS, and others. It was expected to gain further information as input for the simulation study in terms of *close to reality* set-ups, as well as to gain an overview of the software in use which is intended as input for workpackage 4, the *evaluation of statistical software*.

The questionnaire itself consists of two main parts A and B. Part A contains questions for the National Surveys (12 items) whereas part B is devoted to variance-estimation-software in use (5 items). The full evaluation of the questionnaire can be drawn from appendix B.2.

3.1.2 Overview of the questionnaire

The heading column in Table 2 (Appendix B.2) shows the names of all responding countries in alphabetical order. One can see that all countries of the EU, newly associated candidates and others are in this list. Columns of the Table 2 (Appendix B.2) show the responses to the 12 and 5 questions (parts A, B) in short abbreviations. Part A contains questions for the type of survey, for the characteristic: mandatory or voluntary, for the amount of non-response, for periods and to other most relevant features of surveys for individual and household data. The heading line of Table 2 shows in abbreviated form the core of responses of the questions in parts A, B. In the cells of Table 2 (Appendix B.2) one can find all results in a condensed and abbreviated form. The other Tables 1 - 4 (Appendix B.3) with results of selected evaluations is divided in two parts according to HBS and LFS. It shows the results of classifying the responses of all countries according to the properties of HBS and LFS in a detailed form.

About 25 countries in Europe as well as Israel received the questionnaire. Responses from 22 countries came back to the coordinator of DACSEIS, 3 countries did not respond, NR-Quota 12%. In some of the evaluated questionnaires there exists item non-response to a greater or smaller extent. Table 2 (Appendix B.2) gives an overview to item non-response in the surveys.

3.1.3 Overview of some main results

The answers show that about 80% of the participating countries apply household surveys, labour-force surveys and others that integrate these aspects like the German and Austrian Microcensus (MZ). Most of these surveys gain household as well as individual data for variables like income or employment. Around these central surveys in the several countries a variety of complementary surveys for special purposes is gathered.

The designs of the surveys are very different. The developed countries only show some common properties in the designs of their National Surveys. For example, two stage stratified sampling, rotational schemes, households or individuals as units etc.

Mandatory are just 50% of the LFS with voluntary parts in it. The other half of the requested surveys are strictly voluntary. Regarding the HBS's and the other surveys only 20% up to 30-% of the surveys are mandatory.
The non-response in the case of mandatory surveys varies between 5% and 28%. In all other cases non-response quotes up to 90% are possible. Imputation as a tool to reduce remarkable non-response is applied by the National Offices.

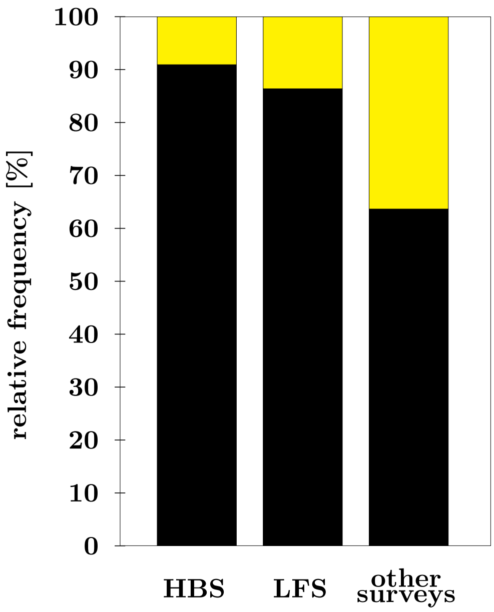
A variety of software is used for variance estimation. Own software of the National Institutes, SAS, PL/I, EXCEL, SPSS, S-PLUS, SUDAAN etc. The whole variety is demonstrated as a graph in figure 3.14 For the future in some countries it is planned to implement POULPE, updates of SAS and new versions of SUDAAN.

3.2 Evaluation of the questionnaire

In the following section the responses to all the questions are gathered and presented in a condensed form restricted to the most relevant surveys HBS and LFS. Other types of surveys which are part of responses to question 1, vary over a wide spectrum of different surveys. A gathering regarding those ones is not meaningful. For the responses of the questions over all countries there is a table and a corresponding graph or diagram to show the results in percentages of the questioning in the European countries as well in some outside of the EU. The diagrams show ratios of certain characteristics of the surveys relative to those in the whole set.

3.2.1 Surveys of interest

In Figure 3.1 corresponding to question 1 it is shown that 91% of the responding countries carry out a HBS and 86% a LFS. This result supports the restriction to carry out special detailed evaluations for HBS and LFS.



Legend for surveys:

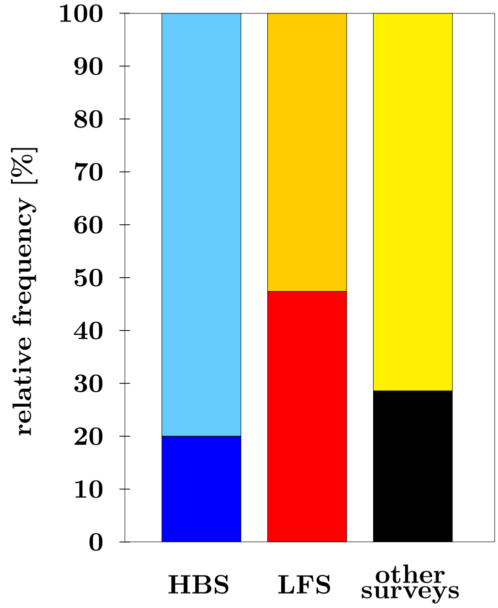
- survey not mentioned
- survey mentioned in questionnaire

Data table:

	HBS	LFS	other surveys
not mentioned	9.1%	13.6%	36.4%
mentioned	90.9%	86.4%	63.6%

Figure 3.1: Quotes for surveys mentioned by countries in the questionnaire

A corresponding Figure 3.2 (questions 1 and 2) shows the proportions between mandatory and voluntary in the surveys of the different countries. For HBS this proportion is 4 : 1 and for LFS 1.1 : 1.



Legend for surveys:

survey voluntary

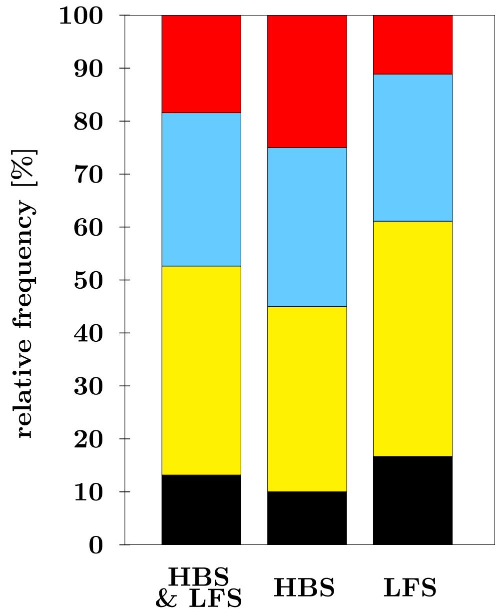
survey mandatory

Data table:

	HBS	LFS	other surveys
voluntary	80.0%	52.6%	71.4%
mandatory	20.0%	47.4%	28.6%

Figure 3.2: Distribution of surveys with respect to mandate classes

Figure 3.3 deals with question 3: periodicity of HBS, LFS. For the categories monthly, quarterly, yearly and others the diagram shows approximately balanced proportions. This result leads to the conclusion that a preference for a distinct periodicity does not exist for the surveys dealt with in the countries that filled in the questionnaire. The maximum value for HBS & LFS is about 40% for the category quarterly.



Legend for time intervals:

period: other

period: yearly

period: quarterly

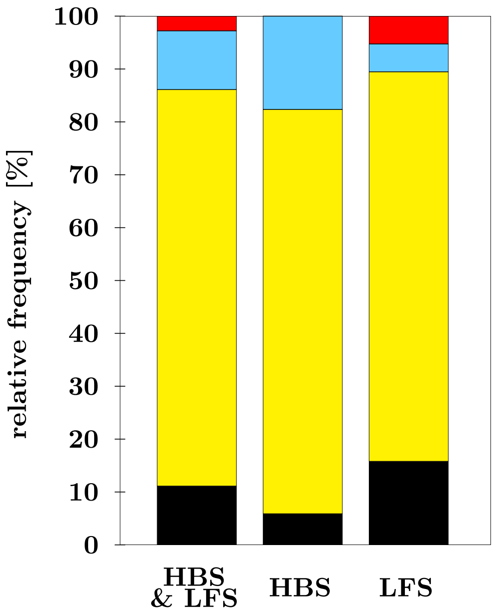
period: monthly

Data table:

	HBS & LFS	HBS	LFS
other	18.4%	25.0%	11.1%
yearly	28.9%	30.0%	27.8%
quarterly	39.5%	35.0%	44.4%
monthly	13.2%	10.0%	16.7%

Figure 3.3: Distribution of surveys with respect to time intervals

The Figure 3.4 is related with question 4: type of design classes, e.g. cross-sectional etc. for HBS as well as for LFS. Only 75% of this surveys in the countries belong to the category: cross-sectional with rotation. Types like panel, cross-sectional without rotation and others are less than 12% for HBS & LFS. Conclusion: This characteristic shows a homogeneity of the countries with respect to the design classes mentioned above.



Legend for design classes:

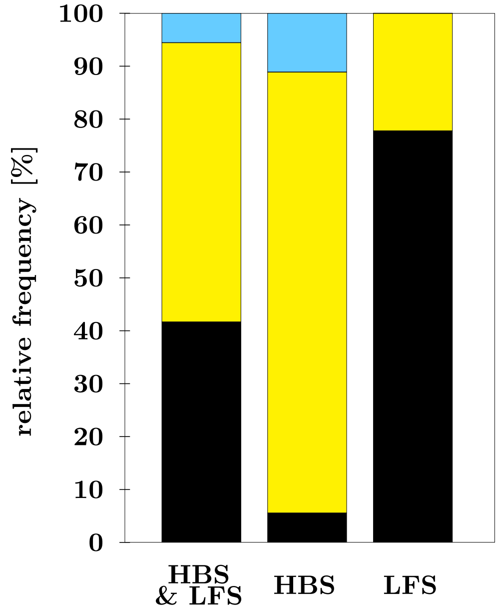
- other
- cross- sectional without rotation
- cross- sectional with rotation
- panel

Data table:

	HBS & LFS	HBS	LFS
other	2.8%	0.0%	5.3%
cr- sect. wo. rot.	11.1%	17.7%	5.3%
cr- sect. w. rot.	75.0%	76.5%	73.7%
panel	11.1%	5.9%	15.8%

Figure 3.4: Distribution of surveys with respect to design classes

The responses to question 5 are the basis for Figure 3.5. Three categories of non-response rates show in detail how the size of the ranges applied are selected. For HBS & LFS the non-response quote that varies in the range [0%;15%] has a relative frequency of about 42%. Analogously, only 53% are in the range (15%;50%] and a low 6% in the range (50%;100%]. The HBS and the LFS differ significantly in the percentages of non-response for the categorial level of [0%;15%].



Legend for non-response classes:

- non-response quote: (50%; 100%]
- non-response quote: (15%; 50%]
- non-response quote: [0%; 15%]

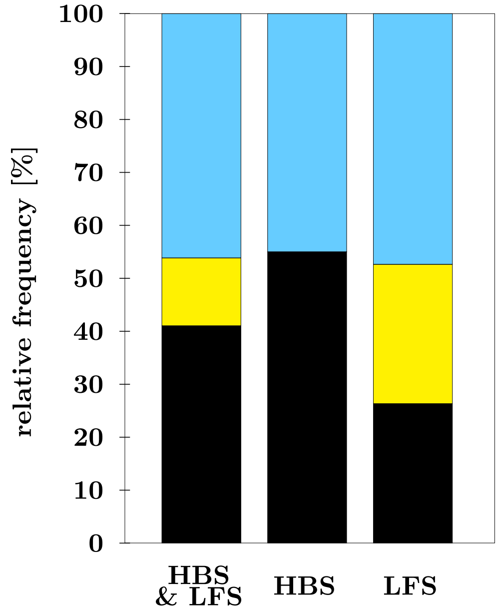
Data table:

	HBS & LFS	HBS	LFS
(50%; 100%]	5.6%	11.1%	0.0%
(15%; 50%]	52.8%	83.3%	22.2%
[0%; 15%]	41.7%	5.6%	77.8%

Figure 3.5: Distribution of surveys with respect to non-response classes

Further details on the distribution can be drawn from the Table 4 (Appendix B.3).

Figure 3.6 is related to questions 6 and 7 and shows the distribution of the most relevant types of units in HBS, LFS and their proportions in percent. The column for HBS & LFS shows that category 3 (households and individuals) have just the same relative share (46%). However, HBS and LFS differ clearly in categories 1 and 2 (households only, individuals only) which only have a relatively small percentage.



Legend for unit classes:

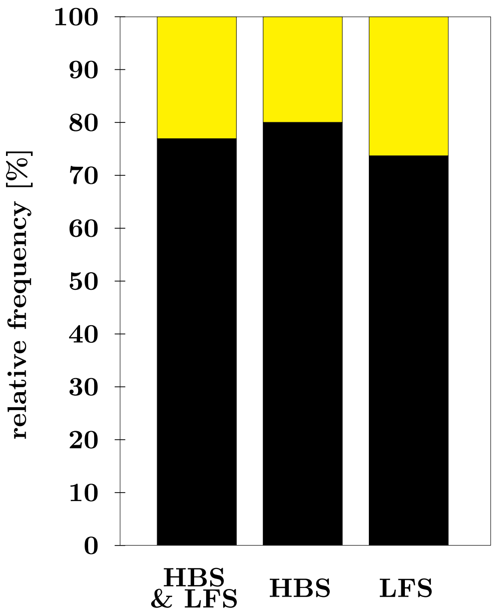
- units: households and individuals
- units: individuals
- units: households

Data table:

	HBS & LFS	HBS	LFS
both	46.2%	45.0%	47.4%
individuals	12.8%	0.0%	26.3%
households	41.0%	55.0%	26.3%

Figure 3.6: Distribution of surveys with respect to definition of survey units

Figure 3.7 shows an overview for sample selection schemes. 2- or 3- stage sampling is dominant for both surveys and ranges between 73% - 80%. The proportion of all other types ranges around a mean of 23%.



Legend for surveys:

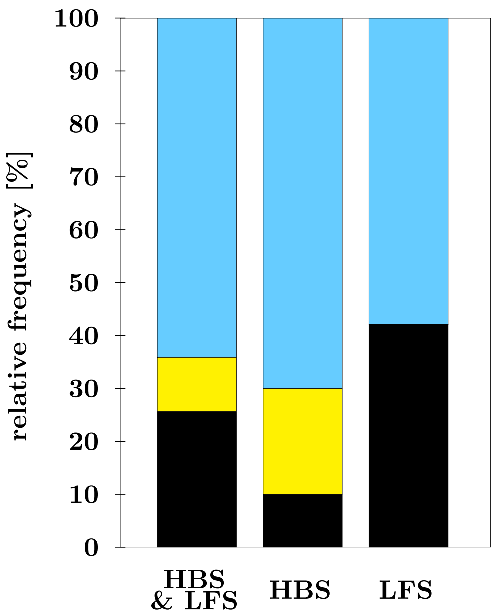
- other sampling design
- 2- or 3- stage sample

Data table:

	HBS & LFS	HBS	LFS
other	23.1%	20.0%	26.3%
2/3- stage	76.9%	80.0%	73.7%

Figure 3.7: Overview of sample selection schemes

Figure 3.8 shows results for HBS and LFS with respect to the scaling of the variables. From the graph one can see that both types: metric and categorial have a proportion of about 66% over all countries. This gives an impression that a sufficient homogeneity regarding this feature for 2/3 of the countries is valid.



Legend for variable classes:

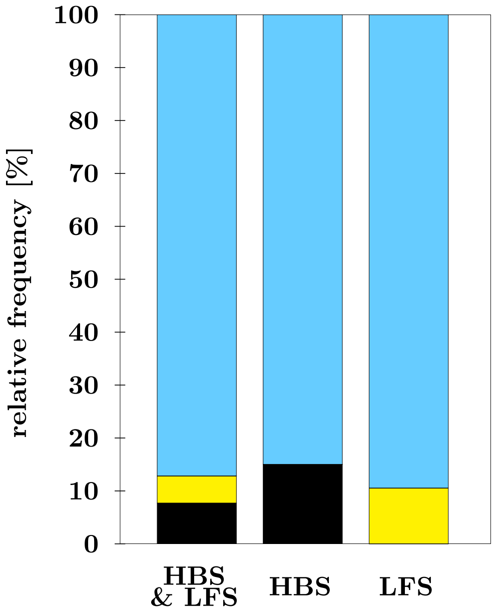
- variable: categoric and metric
- variable: metric
- variable: categoric

Data table:

	HBS & LFS	HBS	LFS
both	64.1%	70.0%	57.9%
metric	10.3%	20.0%	0.0%
categoric	25.6%	10.0%	42.1%

Figure 3.8: Distribution of surveys with respect to scaling of variables

In some respects, Figure 3.9 corresponds with Figure 3.6. Question 10 asks especially for the possibility to get from HBS or LFS individual as well as household data simultaneously. In about 88% (mean) both types of data are surveyed. Combined with the results of figure 3.1 these proportions show that approximately 90% of all surveys in the countries obtain household and individual data within HBS and LFS.



Legend for data classes:

- data: households and individuals
- data: individuals
- data: households

Data table:

	HBS & LFS	HBS	LFS
both	87.2%	85.0%	89.5%
individuals	5.1%	0.0%	10.5%
households	7.7%	15.0%	0.0%

Figure 3.9: Overview surveying household / individual data

As mentioned above in Section 3.2, it seems to be reasonable to gather the surveys into three categories HBS (1), LFS (2), and other surveys (3). The surveys of category 3 show a wide spectrum of quite different surveys. The responses of these surveys differ greatly in completeness and accuracy. Therefore, the evaluation was restricted to those interesting features which could be figured out reliably with the responses. 14 of the 22 countries gave information about complementary surveys. Most of these additional surveys are of type voluntary (71%). Due to this fact the non-response is higher. The non-response quotes are shown in figure 3.11. 62% of the surveys carry out variance estimation. This result is slightly higher than results for HBS and LFS. With regards to the survey category: other, 50% gain individual data.

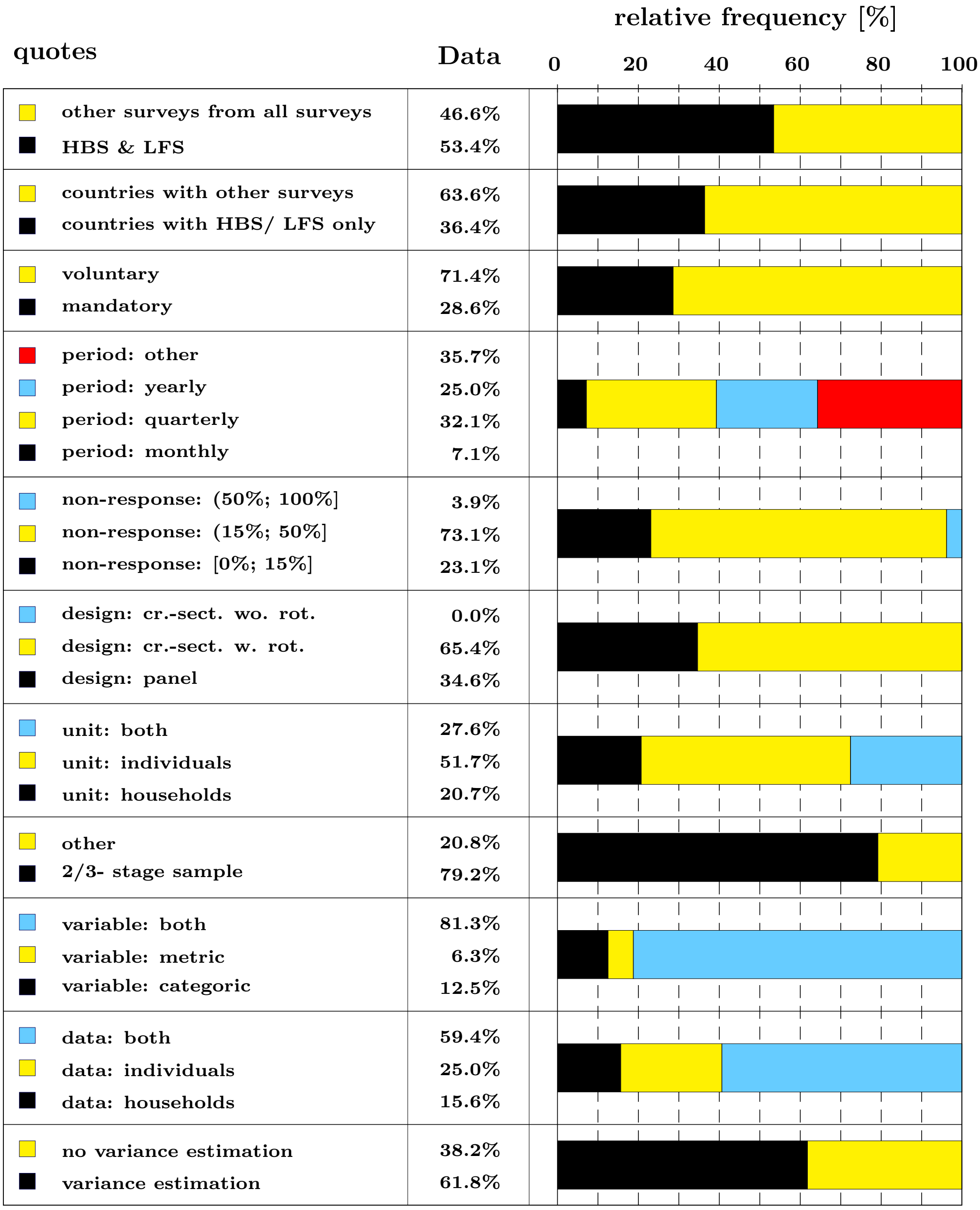


Figure 3.10: Evaluation of additional surveys

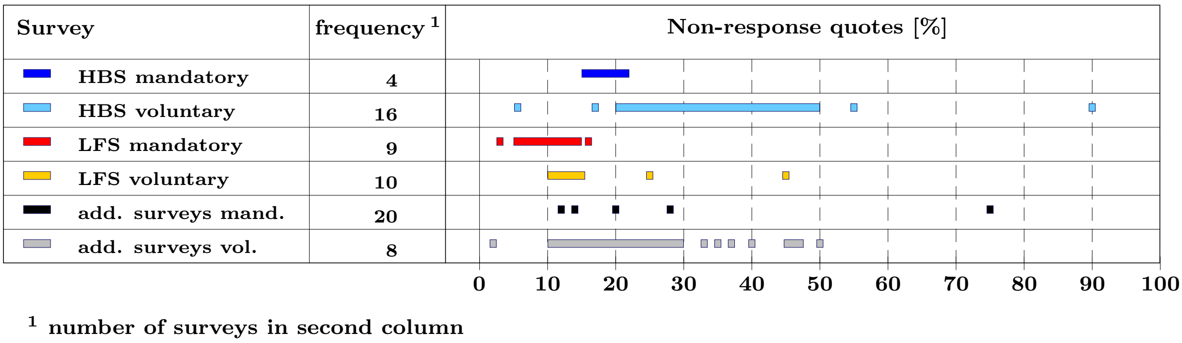
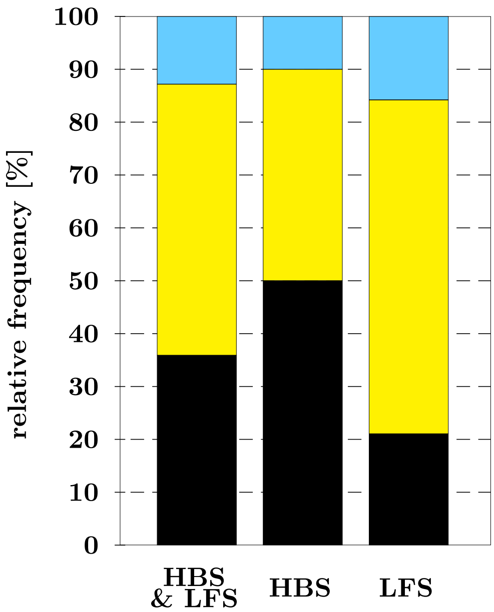


Figure 3.11: Comparison of non-response quotes: HBS, LFS and additional surveys

3.2.2 Variance estimation methods in use

Figure 3.12 gives an overview of variance estimation methods that are currently in use in the questioned countries. The responses to question 11 to a certain extent are not reliable and not meaningful with regard to details. Due to this the classification uses only three categories for HBS and LFS. An interesting result is that on average just 64% is using standard and other methods to estimate variances. About 36% do not calculate any variances to control the quality of surveyed data. This seems to be an unsatisfying proportion.



Legend for variance estimation:

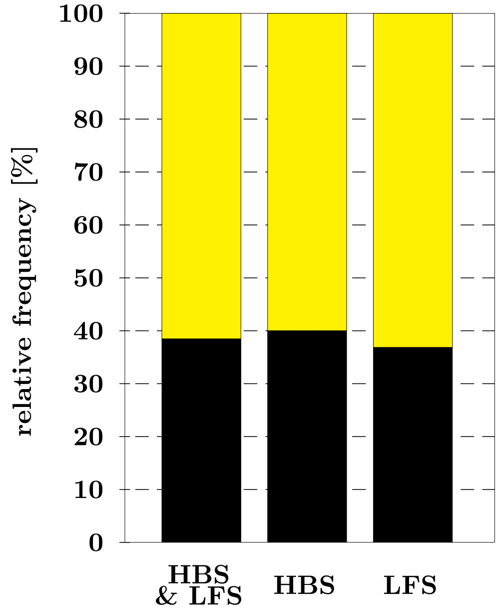
- re-sampling, approximative & other
- standard methods
- no variance estimation methods

Data table:

	HBS & LFS	HBS	LFS
other	12.8%	15.8%	10.0%
standard m.	51.3%	63.2%	40.0%
no var. est.	35.9%	21.1%	50.0%

Figure 3.12: Overview of variance estimation procedures applied to surveys

Figure 3.13 deals with question 12 which asks for changes and modifications of applied survey designs and further modifications planned for the future. The result shows total homogeneity in the three categories applied. 40% of the countries do not intend to change anything and the complemental 60% will look for efficient changes.



Legend for surveys:

- no change of survey intended
- change of survey design intended

Data table:

	HBS & LFS	HBS	LFS
no change	61.5%	60.0%	63.2%
change of design	38.5%	40.0%	36.8%

Figure 3.13: Distribution of responses for change in survey design

3.2.3 Software in use

Figures 3.14 and 3.15 show overviews for software packages in use, their preferences across the countries and the intention to implement new software in the future. The results contain responses of questions B.1 - B.5 in a condensed form. SAS seems to be the most used software package. 8 times it was mentioned, SUDAAN only 4 times and the other packages show lower numbers like 3, 2, 1. 60% of the 17 respondent countries want to implement new or additional software, 22% plan no software changes and just 18% will implement modified software which will improve the treatment of non-response.

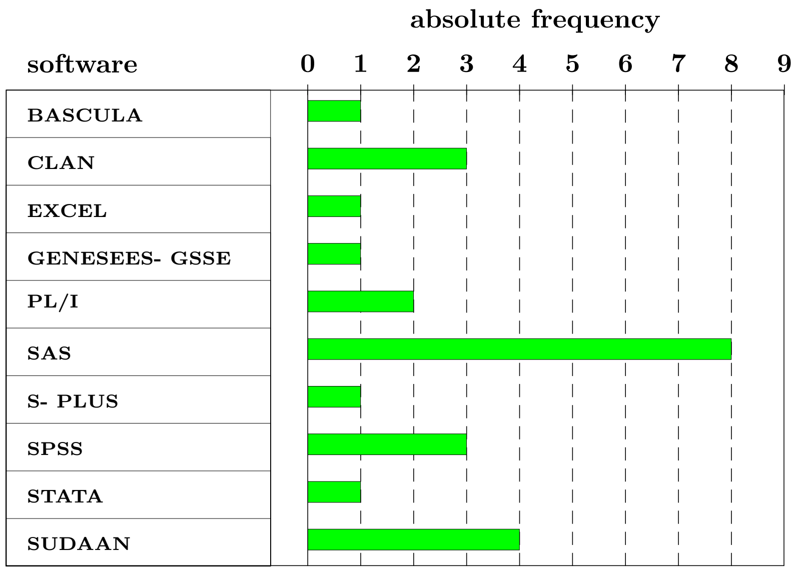


Figure 3.14: Overview of software packages in use

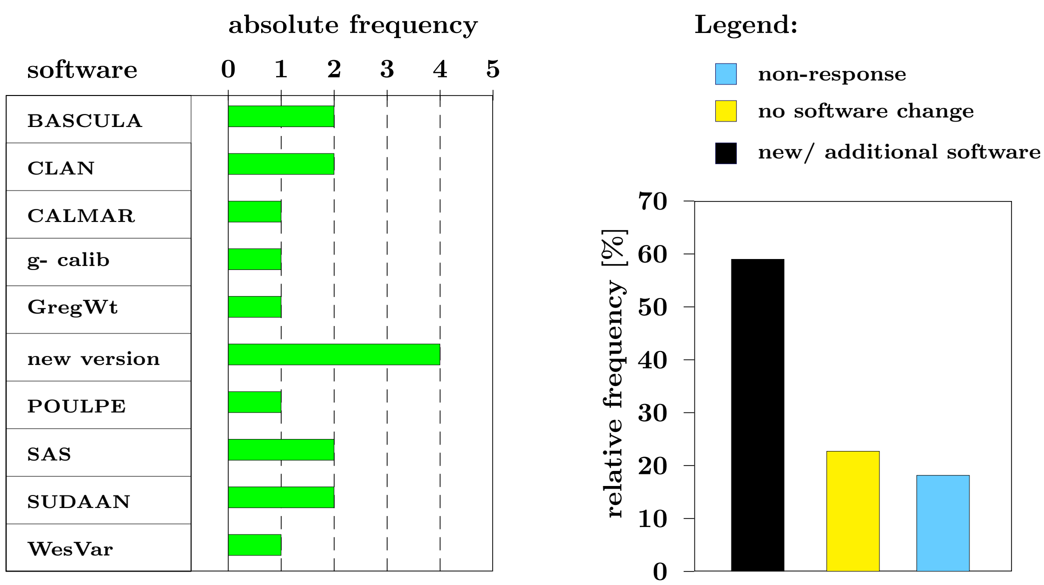


Figure 3.15: Overview to new or additional software under consideration

3.3 Summary

Evaluating the 22 questionnaires it soon became evident that a descriptive treatment of responses seems to be the best way to get clear as well as reasonable results. For three categories of surveys: HBS, LFS, and category: other surveys, a set of figures and related graphs were carried out. They show the results of the questioning and of the evaluation in detail. Furthermore, one can recognise the following main results:

- In figure 3.1 corresponding with question 1 it is shown that 91% of the responding 22 countries carry out a HBS and 86% a LFS (or equivalent types of surveys).

- The set of other surveys contains many different surveys. For these elements the same modus operandi as in the case of HBS and LFS is used. We recognised remarkable differences in the mandate-type, in the non-response quotes, as well as in the category: period.

- The graphs for HBS and LFS show a convincing homogeneity of characteristics of surveys across the responding countries. For example, non-response quota of surveys in countries where it is mandatory belong generally to the same category of non-response. If the survey is voluntary no homogeneity regarding non-response classes can be recognised. The graphs for HBS and LFS show convincing homogeneity of characteristics for the surveys of the responding countries. For example, non-response quotes of surveys in countries with category mandatory are part of the same non-response category. If the survey is voluntary no homogeneity regarding non-response classes can be recognised.

- Regarding table and graph of software in use, Figure 3.14, one can easily see that SAS with 8 users is the dominant software.

- About 62% of the countries (average) carry out variance estimation for HBS as well as for LFS 62% for category: others.

For more details of main results of the evaluation see Subsection 3.1.3. One can see that there exists generally homogeneity separately for countries with HBS and for those with LFS. For the category: others, this result fails. According to the gathering applied and to the figures evaluated the main conclusion regarding household and individual data is that in the relevant EU-countries equal standards are approximately valid.

Full details of the evaluation of the questionnaire can be drawn from the appendix B.2.

Chapter 4

Metadata for the dissemination of DACSEIS results

One major aim of the DACSEIS project is to generate a recommended practice manual on the accuracy of estimators. For an optimal use as a manual and as an internet based database, it is convenient to formulate adequate metadata on the items of the recommended practice manual. These metadata should focus on the simulation output to most appropriate support the recommendations on the use of estimators and variance estimators in a practical environment. Therefore, the metadata should cover all aspects from the survey to simulation output covering the dissemination of information. The final implementation will be available in the recommended practice manual and, hence, in the report for deliverable D1.2.

4.1 General remarks on the DACSEIS metadata

The metadata should include all necessary and several additional keys including possible values for the different items of the recommended practice platform. The items obviously cover estimators and variance estimators. However, for the simulation study further items are needed including the surveys and concrete formulation of tasks of interest as parts of surveys with a subset of variables. Additionally, to best cover the aim for practice recommendations, several changes to the original survey, such as design and non-response, have to be included in the design of metadata for the recommended practice manual. Therefore, an adequate set of metadata cannot be the same as for a single survey in practice.

A general scheme for the metadata can be drawn from figure 4.1.

The following metadata focus on the relevant aspects of the DACSEIS recommended practice. Therefore, the keys and values are strongly related to the needs of the DACSEIS simulation study and concentrate on the relevant aspects for recommended practice. However, the main conception tries to respect for generality and easy enhancement of the following metadata to more general needs. As a basis for these metadata, the *Guidelines for the Modelling of Statistical Data and Metadata* in United Nations (1995) were adopted to the needs of the DACSEIS project.

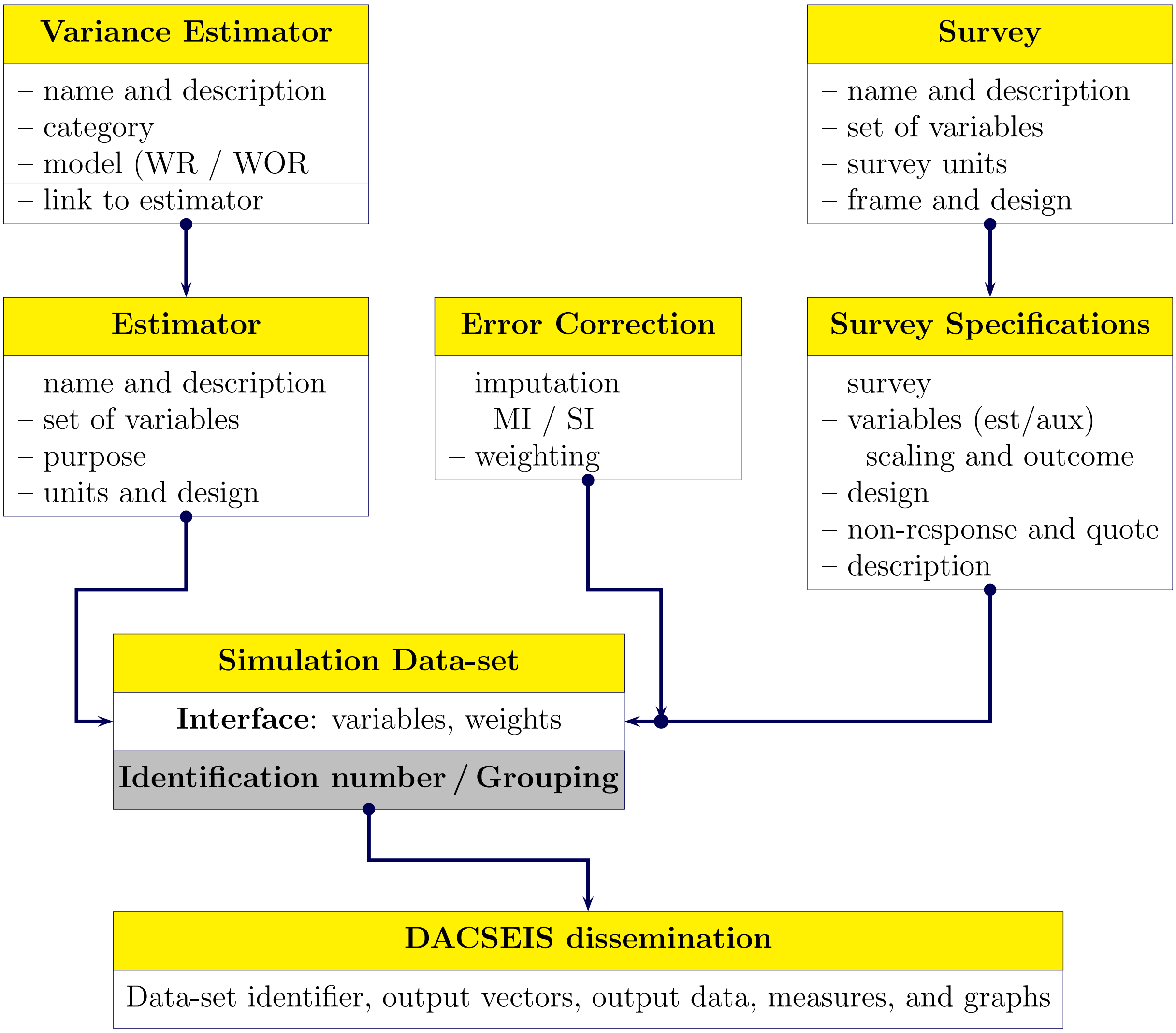


Figure 4.1: Structure of the DACSEIS metadata

4.2 Survey data

According to the general remarks on metadata and their use for the simulation study and recommended practice, it seems convenient to split the survey metadata in two groups, the general survey data as the *ideal* universes from the six DACSEIS surveys and the specialised survey data for use in the simulation studies. The latter contain a subset of variables of the surveys for special simulation tasks and additionally concrete information on non-response in the data.

4.2.1 Surveys

The metadata for the six DACSEIS universes include the following information:

Name of survey Dutch LFS, Finnish LFS, Austrian Microcensus, German Microcensus, German EVS, Swiss HBS;

Description This item should redirect to the WP2 description as a PDF subdocument including some graphs (reference to additional sources);

Survey units Households, individuals (as for the simulation);

Variable list List of variables that are available including scaling, list of values, and voluntary or mandatory variables;

Survey plan Original design, further designs for the simulation study (with descriptions).

4.2.2 Survey specifications in the simulation study

The *survey specifications* will be defined to allow for concrete simulation tasks regarding the choice of variables, sampling units, design, and peculiarities of survey data, e.g. non-response in data. The model cases will have to be designed with the aim to standardise an interface to further procedures that have to be defined such as imputation methods.

Survey specification name The name should cover the survey acronym and a unique identifier;

Target variable(s) One or two target variables have to be identified from the set of variables in the survey section; the target values should also be identified to describe precisely the aim of the simulation to be performed;

Auxiliary variables As before. The auxiliary variables should cover only sample information in this item.

Auxiliary information Any further information, e.g. register based information or aggregate universe-based information, should be identified in this section. As an additional document, the description on the use of the information should be added.

Non-response information This information should be categorised as indicated below. This may differ from general metadata on non-response models since it has only to do with the concrete non-response simulation information.

Description Short description of the task, including aim and features.

Full information data-set This information must contain the *true* values. The aim is to use the information with efficiency measures.

4.2.3 Non-response in surveys

The non-response metadata are strongly connected towards the simulation study. Therefore, the information on non-response modelling is related to the survey and a certain selection of variables. It will be fully integrated in the section on survey specifications for the simulation study.

Type of non-response Unit / item non-response;

Model MCAR, MAR, and NMAR;

Variables of non-response The variables that have to be *masked* as nonrespondants have to be indicated. Estimation variable, auxiliary variable;

Non-response quote 0%, 5%, 10%, 25%, 40%, *true* non-response quote;

Description of model Reference to the non-response modelling in the surveys;

Description Some remarks on the non-response mechanism;

Code Pseudo-Code, R, C++, or SAS code.

4.3 Correction for non-response

According to the section on model cases, where the inclusion of non-response was introduced, this section will add the methods for the correction of non-response.

4.3.1 Imputation methods

Type SI or MI;

Method E.g. regression imputation;

NR model Reference to the accompanying non-response model;

Description Some remarks on the method;

References 1 – 3 references including the original source. This should be given as BiBTEX-reference from the RPM database;

Formula Formula or algorithm for estimation method (in LATEX);

Code Pseudo-Code, as well as R, C++, and possibly SAS code.

Within the simulation, the imputation methods will be applied to survey specifications that already include certain non-response mechanisms.

4.3.2 Weighting

Weighting techniques may be applied in different areas and for different reasons within the estimation process, e.g. design weighting and weighting for non-response. Therefore, generating specialised metadata for weighting procedures generally leads to a selection of specific estimators that may incorporate explicitly weights. Hence, no further specifications of metadata within the DACSEIS simulations are needed. The unique identification of weighting for non-response within simulation tasks can be drawn from the the task name in which no imputation method is specified and the fact that the point and variance estimator allow for weighting.

4.4 Estimation and variance estimation methods

This section contains the information on the point and variance estimators. Since variance estimators have to be applied to specific point estimators within the estimation or simulation process, a direct link between the estimators will be added to the variance estimator.

4.4.1 Estimation methods

Name of method If available.

Purpose of method total, mean, ratio, proportion, median, change rate, correlation, quantiles.

Survey design simple random sampling, cluster sampling, stratified random sampling, unequal probability sampling, multi-stage sampling, multi-phase sampling.

Replacement with / without / both (*both* is default).

Variables of interest estimation and auxiliary variables with number of variables (generally, 1, 2, or n), scale (binary, nominal, ordinal, categorical, metric, mixed for 2 or n), and aggregation level (if not sampling unit, e. g. for small area methods).

Description Some remarks on the method

References 1 – 3 references including the original source. This should be given as BiBTEX-reference from the RPM database.

Formula Formula or algorithm for estimation method (in LATEX).

Code Pseudo-Code, as well as R, C++, and possibly SAS code.

Software packages The standard software packages should listed as a table to see if method is available (yes, no, partially with footnote). Software: BASCULA, CLAN, POULPE, SAS, STATA, SUDAAN, WESVAR

Additional remarks Link to simulation results.

4.4.2 Variance estimation methods

Link to estimator Generally, variance estimation methods are designed for a special estimation method. Purpose, design, and variables of interest will be copied from estimator.

Name If available.

Category of variance estimation method
1 DI Direct,
2 TL Taylor Linearisation,
3 RG Random Group,
4 BRR Balanced Repeated Replication,
5 JK Jackknife,
6 BM Bootstrap Methods.

Replacement with / without / both (*with* is default).

Description cf. estimator.

References 1 – 3 references including the original source. This should be given as BiBTEX-reference from the RPM database.

Formula Formula or algorithm for estimation method (in LATEX).

Code Pseudo-Code, as well as R, C++, and possibly SAS code.

Software packages The standard software packages should listed as a table to see if method is available (yes, no, partially with footnote, cf. estimator).

Additional remarks Link to simulation results.

4.5 Simulation study

The metadata for the simulation study will aim for the dissemination of the DACSEIS *recommended practices*. However, this will consist of the simulation set-ups that each indicate one single or a grouped simulation run on a identified task, with a given imputation or weighting rule, if needed, and a given point estimator and variance estimator respecting for a common interface. The interface will be checked within the module for the simulation data-sets and conclude with the results in the simulation output.

4.5.1 Simulation data-sets

Task Selection of task;

Non-response correction Selection of either weighting or imputation method;

Estimator Selection of the point estimator respecting for the two before mentioned items;

Variance estimator Selection of the variance estimator for the use with the specified point estimator.

The data may be grouped to allow for a comparison aiming at practice recommendations.

4.5.2 Simulation output

The simulation output will be standardised according to the before mentioned task including the information on the accompanying universe including its survey specification (with and without non-response, if applicable) and will contain the vector of point and variance estimates. A standard set of measures will be given. The aim is to easily update the measures if needed, e.g. during run-time on the RPM server.

Task Indicate ID of the task;

Simulation data-set This will precisely indicate the simulation;

Vector of estimates As plain data-file;

Vector of variance estimates As plain data-file;

Standard set of measures The measures will be given according to the recommendations in the measures section.

Chapter 5

Non-response modelling for the simulation study

The artificial universes have been formed for the five so-called DACSEIS surveys, that is, for the Austrian and German microcensuses, for the Dutch and Finnish labour force surveys and for the Swiss household budget survey. Initially the universes are complete, without any missingness or other deficiencies. Consequently, these universes can be used for examining variance estimation problems in such an ideal situation, although still there are complexities in estimating figures for small domains and areas, for example. The DACSEIS project aims at examining more difficult and realistic cases too, so that non-response is included in simulations. This requires to create missingness in the samples that are used within simulations by a non-response mechanism.

Our strategy in the creation of non-response is roughly as follows:

We assume that each unit in the universe has a certain propensity to respond or not to respond to a survey question, either to all questions in the survey, or to some questions only. Consequently, we will include this kind of propensities or probabilities in each universe. How to do it? We have tried to find a training data set as close to a real situation as possible, and next we have estimated such probabilities. Our strategy has not been to estimate only one set of probabilities but some such ones that can be believed to be possible in real life. For these selections we have used standard tools, and assumptions behind these, starting from a quite simple case and going on to more complex ones. This strategy means that we have first assumed that the missingness mechanism is a simple MAR (missing at random) or unconfounded, next a more complex MAR, and in some cases, we have based our mechanism to NMAR (missing not at random).

The estimation of response propensities is usually based on logistic regression but also the direct statistical probabilities (frequencies) are used. In this chapter we call these functions as 'non-response models.' The estimation results of these models can be easily implemented for the DACSEIS simulation setup. The chapter shows how the models have been constructed for each country data set. The key estimation results are also included. Moreover, some formulae for calculating variance estimates have been given.

5.1 Non-response for the Austrian Microcensus

5.1.1 Introduction

The sampling frame of the Austrian Microcensus (=AMC) consists of all dwellings (inhabited or not) of the Austrian Housing Census. These dwellings are the sampling units for the AMC even if for some variables the universe of households or persons is the universe of interest. The reason for this is, that there is no central register of persons in Austria.

For the purpose of the selection of dwellings for the AMC the universe of dwellings is splitted into two parts. One part (=part A) consists of dwellings situated mainly in large (more urban) municipalities and the other (=part B) consists of dwellings situated in smaller (mainly rural) municipalities (see: workpackage 2: "Report on the Structures of the Universes and the Sample Deisgns of the National Surveys", p.10-12).

Within a selected dwelling we may get either

- all information on the interesting dwelling-, households- or persons-characteristics or
- the information, that there is actually nobody living in this dwelling, because the dwelling is a second home, simply not occupied or even not existing
- the information on some dwelling characteristics, but absolutely no data about households or persons, who are actually living in this dwelling, because these persons are nonrespondents in the sense of nobody was found at home or they were refusing to answer (although in principle the answering to the AMC is mandatory)

So the non-response of people in the AMC occurs in such a way, that we get absolutely no information about the nonresponding persons with the exception of: the part of the dwelling universe, the federal state and for part A-dwellings the stratum and for part B-dwellings the PSU, to which the dwelling belongs. No other information about the nonresponding persons of such a dwelling is gathered in the caes of the AMC.

5.1.2 Variables used for the modeling

The Austrian pseudo universe includes the following variables on person level:

PART Part of the dwelling universe, to which the dwelling, in which the person is living, is situated (1=part A, 2=part B)

FED_STATE Federal State (1=Burgenland, 2=Carinthia, 3=Lower Austria, 4=Upper Austria, 5=Salzburg, 6=Styria, 7=Tyrol, 8=Vorarlberg, 9=Vienna)

STRATUM Stratum within part and federal state

PSU For part B-dwellings: PSU to which the dwelling belongs

DWELL_NR Number of dwelling within stratum (part A) or PSU (part B)

HH_NR Number of household within dwelling (0-3)

PERS_NR Number of person within household (0-10)

AGE Age in years

GENDER (0=male, 1=female)

EDUCATION Highest educational level (0=not completed compulsory education, 1=completed compulsory education, 2=completed apprenticeship, 3=medium secondary level, 4=secondary academic school, 5=upper secondary level school, 6=post secondary school, 7=tertiary level school, 8=university, 9=school-aged child)

NATION Nationality (0=Austria, 1=former Yugoslavia, 2=Turkey, 3=other countries)

EMPLOYMENT Working at least one hour a week (0=yes, 1=no, 2=unknown)

Only the information for the dwelling characteristics is available, when people to be interviewed are not responding. We can therefore model the non-response using only the variables PART, FED_STATE and STRATUM or PSU.

5.1.3 The Non-response Model

It is possible to develop a response homogenity group (RHG) model (or a missing completely at random (MCAR) model) where each group has different response probabilities conditional to the values of these variables. In fact for part A of the dwelling universe, we get a total amount of $1,189$ three-dimensional strata and different response-rates and in part B there is a total of $1,710$ PSUs with variing response-rates.

Table 5.1: The partition of federal states of the AMC pseudo universe into strata (part A) and PSUs (part B)

Federal State	Part A-Strata	Part B-PSUs
Burgenland	110	78
Carinthia	132	94
Lower Austria	134	482
Upper Austria	134	369
Salzburg	131	71
Styria	123	410
Tyrol	115	206
Vorarlberg	146	0
Vienna	164	0
Austria	1,189	1,710

As you can see, the federal states of Vienna and Vorarlberg consist only of dwellings, that are situated in part A of the dwelling universe.

For all of this total amount of 2,899 disjunctive parts of the universe of dwellings different response rates can be calculated, which leads to a close to reality non-response model, that follows the complicated structure of the real Austrian dwelling universe. It is important to mention once again, that persons within the same dwelling do not have special response rates in the AMC, because the non-response occurs only on dwelling level.

In the following Table 5.2 (and the figures below) the response-rates conditional only to the two dwelling variables Part of the universe and federal state are presented to give an overview:

Table 5.2: MCAR conditional to the variables PART and FED_STATE

Part of the universe	Federal State	Stratum (A)/ PSU (B)	Response-Rate (%)
A	Burgenland	.	94.1
A	Carinthia	.	86.5
A	Lower Austria	.	89.3
A	Upper Austria	.	90.8
A	Salzburg	.	90.1
A	Styria	.	89.5
A	Tyrol	.	89.5
A	Vorarlberg	.	93.8
A	Vienna	.	87.8
A	.	.	90.0
B	Burgenland	.	95.6
B	Carinthia	.	90.2
B	Lower Austria	.	90.3
B	Upper Austria	.	94.9
B	Salzburg	.	96.5
B	Styria	.	96.0
B	Tyrol	.	92.5
B	.	.	93.5
.	.	.	91.1

In Table 5.2 also the total response-rates for the two parts of the dwellings universe (A: 90.0%, B: 93.5%) and the overall response-rate for the Austrian Microcensus, which is 91.1%, is presented.

5.1.4 Outlook

So the proposal for the modeling of the non-response of the AMC is to give each dwelling of the pseudo universe for a MCAR-Model conditional to dwelling variables a value of

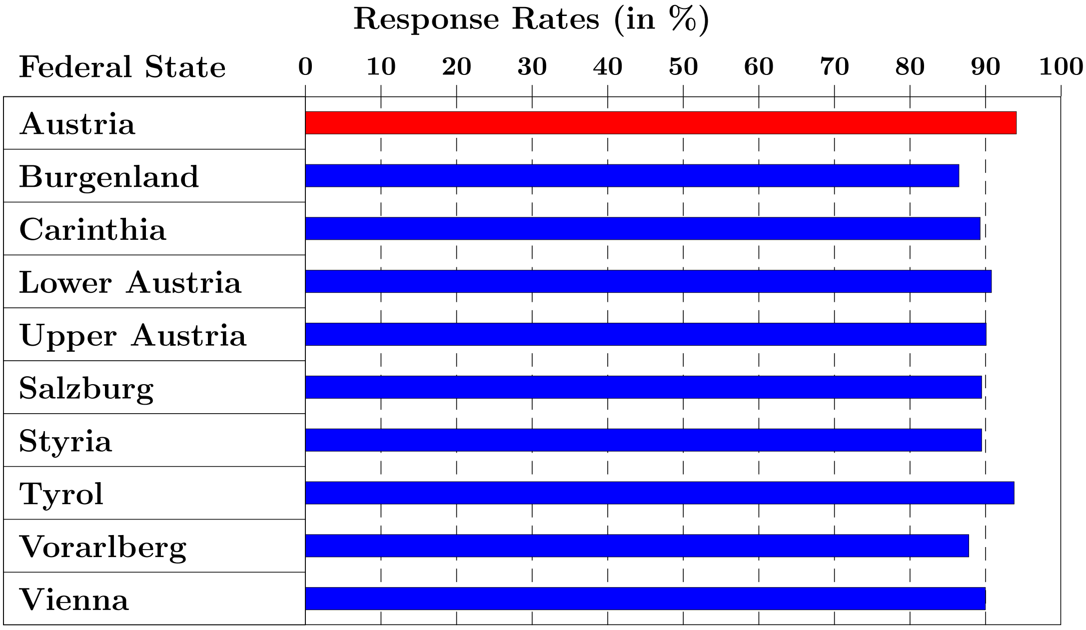


Figure 5.1: The Response-Rates in Part A of the AMC 2001/1

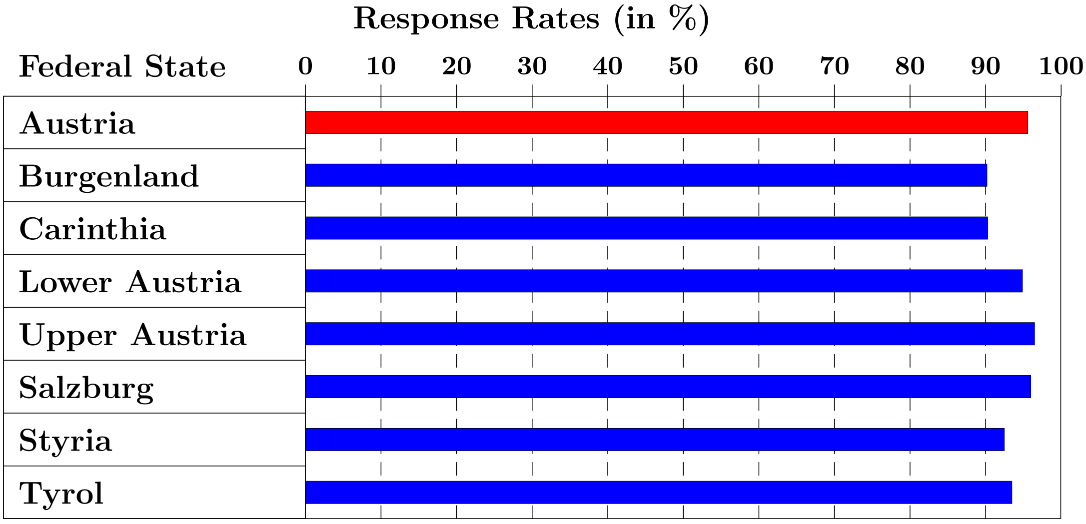


Figure 5.2: The Response-Rates in Part B of the AMC 2001/1

a **response/non-response-variable** (1/0) depending on the value of the response-probablities within the stratum resp. the PSU, in which the dwelling is situated. If a ”nonresponding” dwelling is selected for the sample from the pseudo universe, the person-related data then are ignored like in practice. Therefore we are able to compare the accuracy of estimators calculated from samples, where non-response does not, with the accuracy of estimators, where non-response does occur.

For the results of these simulation using this nonresponse mechanism see Section 2.2 of deliverable D1.2.

To have a more realistic nonresponse model, than the above mentioned available from the survey, for the main simulations, which were done in R, it was decided to adopt

Table 5.3: Adopted Response Probabilities for the Austrian Microcensus

Region	**NAT**	**HS**	**SEX**	**AGE**	**Response Probability**
Vienna	0	1	0	0	0.805
	0	1	0	1	0.850
	0	1	1	0	0.851
	0	1	1	1	0.847
	0	2			0.856
	0	3			0.859
	1	1			0.840
	1	2			0.849
	1	3			0.849
Rest of Austria	0	1	0	0	0.871
	0	1	0	1	0.810
	0	1	1	0	0.893
	0	1	1	1	0.860
	0	2			0.893
	0	3			0.894
	1	1			0.849
	1	2			0.885
	1	3			0.894

the German nonresponse mechanism to the Austrian circumstances. For this purpose it was needed, that for the persons, that were included in the sample, the values of some auxiliary variables were known (which means, that in practice at least these values should be observable from each person in the survey). According to the German nonresponse mechanism (see section 6.4) these variables are: nationality NAT (0: Austrian, 1: non-Austrian), household size HS (1/2/3: 3+), sex SEX (0: male, 1: female) and age AGE (0: <60, 1: 60+).

The decision, whether all persons of a dwelling responsed to some interesting question or not, was made according to the region of Austria in which the dwelling is situated and to the values of the auxiliary variables of a reference person within a dwelling. With these variables like in Germany for a MAR model nine nonresponse (resp. response) classes were built in two regions of Austria, that were Vienna and the rest of Austria. For the fixing of this total of 18 different nonresponse rates these rates of Hamburg (for Vienna) and Bayern (for the rest of Austria) were adopted to the original total nonresponse rate of dwellings in the first Austrian Microcensus of 2001, which was 11.7 %. This gave the nonresponse rates, that are shown in Table 5.3.

To be able to simulate nonresponse also with other total nonresponse rates of - let's say p % (for p=5, 25 and 40) -, the 18 *original* nonresponse rates of Table 5.3 were recalculated by the multiplication with the ratio of p and 11.7.

5.2 Non-response for the Dutch LFS

In this section we present some item non-response (NR) models including the specification of the parameters for the Dutch LFS. These model specifications are to be used in the simulation study. Before describing these models some remarks are in order.

1. In general, it is difficult to estimate the relationship between the dichotomous non-response variable and the target variable y, i.e., the labour status variable, because it is unknown whether a nonrespondent is employed or unemployed.
2. The parameters mentioned in this section are based on some experiences but need not be realistic or accurate, in general. The parameters are such that the mean non-response rate is approximately 40%. The main aim of the simulation study is to get insight into the question to what extent it is possible to correct a survey that is affected by non-response. Therefore, in first instance, it is of interest to examine some worst case and best case models/scenarios apart from the question of which model is most realistic. For simplicity's sake and without loss of generality we only consider models with a binary labour status variable: unemployed or not unemployed. This is of interest for examining the so-called NMAR situation, i.e., the non-response behaviour depends on the labour status.
3. In the context of the Dutch LFS a record has one target variable y (labour status) and a number of register variables such as sex, age, region, marital status, ethnicity. Hence, non-response for the Dutch LFS amounts to item non-response for the variable y.
4. The Dutch LFS is a household survey while the universe used in the simulations is based on persons. In order to obtain model specifications that can be used for the non-response behaviour of households, we propose to look at the eldest person of the household and to apply the model specifications to this person.

In the next subsection we present some NR models. In the last subsection we pay some attention to the difference between the use of a random response model and the use of a fixed response model in the context of large sample surveys.

1. *Model specifications*

First, we introduce some notation

$$NR_i = \begin{cases} 1 & \text{if the } i\text{th person does } \textit{not} \text{ respond when included in the sample} \\ 0 & \text{otherwise} \quad (i = 1, \ldots, N) \end{cases}$$

$$p_i \equiv \mathrm{P}\left(NR_i = 1\right)$$

$$URB_i = \begin{cases} 1 & \text{if the } i\text{th person lives in Overijssel, Utrecht, Noord-Holland,} \\ & \text{Zuid-Holland, or Noord-Brabant} \\ 0 & \text{otherwise} \end{cases}$$

$$ETN_i = \begin{cases} 1 & \text{if the } i\text{th person belongs to a noneuropean ethnic group} \\ 0 & \text{otherwise} \end{cases}$$

$$UPL_i = \begin{cases} 1 & \text{if the } i\text{th person belongs to the unemployed labour force} \\ 0 & \text{otherwise} \end{cases}$$

In the remainder of this section we assume that the relationship between non-response behaviour and the underlying variables can be described by a so-called logit model

$$\ln \frac{p_i}{1-p_i} = \alpha_0 + f(UPL_i, URB_i, ETN_i)$$

or, equivalently,

$$p_i = \frac{\exp\{\alpha_0 + f(UPL_i, URB_i, ETN_i)\}}{1+\exp\{\alpha_0 + f(UPL_i, URB_i, ETN_i)\}}.$$

Now the MCAR (missing completely at random) model becomes

$$\ln\left(\frac{p_i}{1-p_i}\right) = -0,405.$$

When applied to persons this simple response model would result in a non-response (mean) rate of 40%. It is interesting to see what the non-response rate among the households (approximately) becomes when this model is applied to the eldest person of the household as described above. This gives a first global idea of the difference between logit models when they are applied to persons or to households. It is expected that the difference is close to zero.

As MAR (missing at random) model we propose

$$\ln\left(\frac{p_i}{1-p_i}\right) = -0.754 + 0.442URB_i + 0.582ETN_i.$$

A pure NMAR is

$$\ln\left(\frac{p_i}{1-p_i}\right) = -0.427 + 0.832UPL_i.$$

A somewhat more complicated NMAR model is

$$\ln\left(\frac{p_i}{1-p_i}\right) = -0.590 + 0.221URB_i + 0.291ETN_i + 0.416UPL_i.$$

In order to evaluate the sensitivity of the outcomes of the simulations, one may change the parameter values proposed so far by, for instance, +1% which leads to a change of the non-response rate yielding the elasticities of the non-response rate with respect to the corresponding parameters. The same remark can be made with respect to the mean squared error of the target variable [MSE(UPL)].

1. *Fixed or random response model*

In this subsection it is argued that the difference between fixed and random response models can be ignored, provided that the sample size is relatively small ($n \ll N$). Define the dichotomous random response variable R_i by

$$R_i \equiv 1 - NR_i.$$

Let the R_i be independent drawings from a Bernoulli distribution with parameter ρ and define the random vector r by $r = (R_1, \ldots, R_N)^T$. Let y be the target variable and let $\overline{y}_{s,r}$ denote the sample mean of y from the net sample of respondents.

Note that drawing a sample of size n followed by n independent drawings from a Bernoulli distribution is equivalent to N drawings from a Bernoulli distribution in a first step followed by a sample of size n from the whole population which now consists of two strata: respondents and nonrespondents. Making use of the familiar formula for conditional variances, the variance of the *random* response model can be decomposed according to

$$\mathrm{Var}(\overline{y}_{s,r}) = \mathrm{E}\{\mathrm{Var}(\overline{y}_{s,r}\,|r, n_r)\} + \mathrm{Var}\{\mathrm{E}(\overline{y}_{s,r}\,|r, n_r)\}, \tag{5.1}$$

where n_r is the size of the net sample (without nonrespondents). The first term is the so-called within variance. This term can be seen as a lower bound for the variance obtained by applying the *fixed* response model. Note that n_r is fixed here whereas, in fact, in the fixed response model n_r is random. In other words, the variance of the fixed response model is greater than the first term

$$\mathrm{Var}(\overline{y}_{s,r}\,|r) \geq \mathrm{E}\{\mathrm{Var}(\overline{y}_{s,r}\,|r, n_r)\}$$

and, consequently, the second term in (5.1) which can be seen as the between variance component is an overestimation of the difference between the variance from the random response model and the variance from the fixed response model. Define N_r as the (random) size of the stratum of respondents and $\overline{Y}_r$ as the (random) mean of y for this stratum. Let σ_y^2 denote the (adjusted) population variance and $\sigma_{y,r}^2$ the stratum variance. Making use of the classical sampling formulas we get for the first term

$$\begin{aligned}
\mathrm{E}\{\mathrm{Var}(\overline{y}_{s,r}\,|r, n_r)\} &= \mathrm{E}\left\{\left(\frac{1}{n_r} - \frac{1}{N_r}\right)\sigma_{y,r}^2\right\} \\
&= \mathrm{E}\left\{\mathrm{E}\left(\frac{1}{n_r} - \frac{1}{N_r}\right)\sigma_{y,r}^2 \middle| N_r, n_r\right\} \\
&= \mathrm{E}\left\{\left(\frac{1}{n_r} - \frac{1}{N_r}\right)\sigma_y^2\right\} \\
&\approx \left(1 - \frac{n}{N}\right)\frac{\sigma_y^2}{\rho n}.
\end{aligned}$$

Likewise, for the second term we have

$$\begin{aligned}
\mathrm{Var}\{\mathrm{E}(\overline{y}_{s,r}\,|r, n_r)\} &= \mathrm{Var}(\overline{Y}_r) \\
&= \mathrm{E}\{\mathrm{Var}(\overline{Y}_r\,|N_r)\} + \mathrm{Var}\{\mathrm{E}(\overline{Y}_r\,|N_r)\} \\
&= \mathrm{E}\left(\frac{1}{N_r} - \frac{1}{N}\right)\sigma_y^2 + \mathrm{Var}(\mu_y) \qquad (\mu_y \equiv \tau_y/N) \\
&\approx \left(\frac{1}{\rho N} - \frac{1}{N}\right)\sigma_y^2 \;+\; 0 = (1-\rho)\frac{\sigma_y^2}{\rho N} \\
&= (1-\rho)\frac{\sigma_y^2}{\rho n} \times \frac{n}{N}.
\end{aligned}$$

Hence, the second term is much smaller than the first term when $n << N$ and, consequently, the variance of the fixed response model is approximately equal to that of the random response model.

5.3 Non-response for the Finnish LFS

5.3.1 Variables used for modelling

Finnish LFS pseudo universe includes 5 variables and their categories:

region (NUTS 2) (0=Uusimaa, 1= Southern Finland, 2= Eastern Finland, 3=Mid-Finland, 4=Northern Finland, 5=Aaland)

age (5-year age group) (0 = 15-19, 1 = 20-24, 2 = 25-29, 3 = 30-34, 4 = 35-39, 5 = 40-44, 6 = 45-49, 7 = 50-54, 8 = 55-59, 9 = 60-64, 10 = 65-69, 11 = 70-74)

gender (0 = male, 1 = female)

lfstat (labour force status) (0=Employed, 1=Unemployed, 2=Conscripts, 3=Students, 4=Disabled, 5=Pensioners, 6=Persons performing domestic work, 7=Others)

education (highest level of education) (0=Premier or less, 1=Upper secondary, 2=Post secondary non-tertiary, 3=First stage of tertiary education, 5A Programmes 4=First stage of tertiary education, 5B Programmes, 5= Second stage of tertiary education)

For response modelling we have used some transformed variables:

- 10-year age group
  age10=int((age+2)/2)
- Group of regions

region2	region
0	0
1	1+5
2	2
3	3
4	4

- Labour force status
  lfstat2= (0=Employed, 1=Unemployed, 2= Inactives)
  Note that this is available only for the universe

- Education

isced2	isced
0	0
1	1+2
3	3
4	4+5

5.3.2 Response Models

MAR conditional to pre-stratification by gender and age10

Gender	Age10	Probability for responding for the universe, %
0	1	89.6
1	1	90.5
0	2	83.3
1	2	87.7
0	3	83.2
1	3	87.1
0	4	83.9
1	4	87.4
0	5	85.2
1	5	86.2
0	6	86.5
1	6	87.8
.	.	86.4

More complex MAR model using logistic regression

Response indicator (resp=1 if responded, resp=0 otherwise) has been modelled using logistic regression. Our estimation based on real LFS data gave the following results:

$$\begin{aligned}\operatorname{logit}(resp) = 2.6349 &- 0.2114 * (gender = 0) + 0.2715 * (age10 = 1)\\ &- 0.3406 * (age10 = 2) - 0.3827 * (age10 = 3) - 0.2664 * (age10 = 4)\\ &- 0.1789 * (age10 = 5) - 0.5342 * (region2 = 0) - 0.1062 * (region2 = 1)\\ &+ 0.1633 * (region2 = 2) + 0.1606 * (region4 = 3) - 0.6218 * (isced2 = 0)\\ &- 0.2429 * (isced2 = 1) + 0.1037 * (isced2 = 3)\end{aligned}$$

These logits may be transformed to response probabilities for each individual of the universe using the formula:

$$p = \exp\left(\frac{\operatorname{logit}(resp)}{1 + \exp(\operatorname{logit}(resp))}\right) \tag{5.2}$$

NMAR model

Now we add variable lfstat2 to the list of explanatory (auxiliary) variables. The values of lfstat2 for nonrespondents are based on 'good guesses' since true data are not available. Now we obtain - analogously to case B - the following estimated model:

$$\begin{aligned}\text{logit}(resp) = 2.5706 &- 0.2263 * (gender = 0) + 0.2955 * (age10 = 1)\\ &- 0.4070 * (age10 = 2) - 0.4771 * (age10 = 3) - 0.3179 * (age10 = 4)\\ &- 0.2215 * (age10 = 5) - 0.5831 * (region2 = 0) - 0.1321 * (region2 = 1)\\ &+ 0.1843 * (region2 = 2) + 0.1408 * (region4 = 3) - 0.5266 * (isced2 = 0)\\ &- 0.1754 * (isced2 = 1) + 0.1284 * (isced2 = 3) + 0.2861 * (lfstat2 = 0)\\ &- 0.7063 * (lfstat2 = 1).\end{aligned}$$

The response probabilities are to be calculated as in case B.

Some summary points

In these 3 examples the response probabilities are to be fixed approximately for a real-life level in Finland. If in simulations, it is wanted to use different response levels but using the same response structures, it will be easily done so that a certain new coefficient implying that level will be included in all models.

5.4 Non-response for the German Microcensus

For the unit-non-response households, there isn‘t any register information available. But the interviewers can observe in a part of the nonresponding households some variables which can be used as auxiliary information. This are the variables household size (1/2/3+), sex, age (<60/60+), nationality (German/not German) and state of residence (main/secondary); these variables are based on the reference person of the household. There are combinations of these X-variables which are never observed: sex is only observed in households of size one, age is only observed if sex is known and the state of residence is only observed in German households. Unfortunately, even in the planned combinations the auxiliary variables are not always known (in almost 40% of the nonresponding households nothing is known apart from the region).

In Order to get a full table, we estimated the missings in a contingency table (non-response households by auxiliary variables) for each federal state: for known sex in household size 1, the missings of age are allocated according to the proportion of the units with known age, separately for each sex. The missings of sex in household size 1 are allocated according to the common distribution of known sex and (if necessary estimated) age. The missings of household size are allocated according to the distribution of the known houshold size and (if necessary estimated) sex and age. The missings of nationality are allocated proportional to German/not German according to the proportion of the units with known nationality, given sex, age, household size and federal state. For the resulting contingency table see Table 5.4.

For a MAR model, the response probabilities of each cell are estimated by the number of the responding households (Table 5.5), divided by the number of responding households plus the number of nonresponding households (flats which are definitively empty were not included). The age-sex cells for the non-German households are collapsed because the numbers are very small. For remaining cells with no observed non-response, the overall response rate of 96.9% is taken instead of 100%. Table 5.6 shows the final response probabilities. The state of residence is not included because this variable is not in the DACSEIS universe.

We don't consider other models (like a logit model) because of the missings in the X-variables. There is also no information for a NMAR model. Because of the very low non-response rates (the MZ is a mandatory survey) it seems sufficient to consider only this simple MAR model.

Table 5.4: Number of Nonresponding Households in the German Microcensus

Federal State	German						Not German		
	Household Size						Household Size		
	1				2	3 +	1	2	3 +
	Male		Female						
	Age<60	Age 60+	Age<60	Age 60+					
Baden-Württemberg	169	34	53	44	83	66	80	14	24
Bayern	191	170	48	289	168	150	60	12	17
Berlin	96	48	30	20	34	21	11	0	1
Brandenburg	30	13	11	46	18	14	2	0	1
Bremen	21	7	12	10	8	10	2	1	0
Hamburg	96	6	21	26	30	11	8	3	6
Hessen	206	28	97	67	198	170	85	44	88
Mecklenburg-Vorpommern	56	7	11	40	51	32	0	0	0
Niedersachsen	162	223	56	434	227	150	27	6	3
Nordrhein-Westfalen	1128	272	492	568	953	748	111	70	126
Rheinland-Pfalz	26	26	0	34	29	15	6	3	3
Saarland	2	3	4	0	4	2	1	2	0
Sachsen	42	10	8	16	20	13	0	0	0
Sachsen-Anhalt	48	8	21	11	39	22	3	0	0
Schleswig-Holstein	76	30	49	62	98	71	20	11	4
Thüringen	89	18	24	32	63	90	0	2	0

Table 5.5: Number of Responding Households in the German Microcensus

Federal State	German						Not German		
	Household Size						Household Size		
	1				2	3 +	1	2	3 +
	Male		Female						
	Age<60	Age 60+	Age<60	Age 60+					
Baden-Württemberg	4307	1299	3348	5100	12216	11638	894	778	1805
Bayern	5159	1477	3992	5737	14969	15122	963	739	1340
Berlin	2568	500	1885	2065	4705	2897	364	228	432
Brandenburg	960	245	567	1217	3560	3655	21	19	29
Bremen	387	117	325	437	896	537	69	54	81
Hamburg	1254	297	1189	1171	2444	1410	255	174	323
Hessen	2463	719	1946	2802	8005	6624	496	499	1061
Mecklenburg-Vorpommern	658	182	381	745	2105	2196	10	6	9
Niedersachsen	3016	974	2315	3754	10212	8886	294	230	579
Nordrhein-Westfalen	6706	2213	5352	9109	23138	18310	998	1022	2360
Rheinland-Pfalz	1424	512	1097	2011	5284	5057	234	209	445
Saarland	450	160	340	685	1419	1196	53	40	85
Sachsen	1747	515	1011	2779	6498	5723	33	40	61
Sachsen-Anhalt	916	309	582	1507	3778	3440	28	15	37
Schleswig-Holstein	1202	357	902	1494	4199	3179	76	68	142
Thüringen	779	273	493	1309	3140	3408	11	12	27

Table 5.6: Response Probabilities in the German Microcensus (in %)

Federal State	German						Not German		
	Household Size						Household Size		
	1				2	3 +	1	2	3 +
	Male		Female						
	Age<60	Age 60+	Age<60	Age 60+					
Baden-Württemberg	96,2	97,4	98,4	99,1	99,3	99,4	92,0	98,0	99,0
Bayern	96,4	89,7	98,8	95,2	98,9	99,0	94,0	98,0	99,0
Berlin	96,4	91,3	98,4	99,0	99,3	99,3	97,0	96,9	96,9
Brandenburg	97,0	94,8	98,2	96,4	99,5	99,6	91,0	96,9	97,0
Bremen	95,0	94,5	96,5	97,8	99,1	98,2	97,0	98,0	96,9
Hamburg	92,9	98,1	98,2	97,8	98,8	99,2	97,0	98,0	98,0
Hessen	92,3	96,3	95,3	97,7	97,6	97,5	85,0	92,0	92,0
Mecklenburg-Vorpommern	92,1	96,0	97,2	94,9	97,6	98,6	96,9	96,9	96,9
Niedersachsen	94,9	81,4	97,6	89,6	97,8	98,3	92,0	97,0	99,0
Nordrhein-Westfalen	85,6	89,0	91,6	94,1	96,0	96,1	90,0	94,0	95,0
Rheinland-Pfalz	98,2	95,2	96,9	98,3	99,5	99,7	98,0	99,0	99,0
Saarland	99,6	98,0	99,0	96,9	99,7	99,8	98,0	95,0	96,9
Sachsen	97,6	98,2	99,2	99,4	99,7	99,8	96,9	96,9	96,9
Sachsen-Anhalt	95,0	97,6	96,6	99,3	99,0	99,4	90,0	96,9	96,9
Schleswig-Holstein	94,0	92,1	94,9	96,0	97,7	97,8	79,0	86,0	97,0
Thüringen	89,7	93,9	95,3	97,6	98,0	97,4	96,9	86,0	96,9

5.5 Non-response for the Swiss HBS

The Swiss household budget survey 1998 (HBS98) was conducted in 12 monthly waves of a stratified simple random sample of private households in Switzerland with 9,295 fully participating households. Each participating household had to report during a full month at a very detailed level all its expenditures and income. Since this was quite a demanding task, participation had to be strongly solicited and thus unit non-response is not negligible in HBS98. Besides, households that did not deliver a completed report at the end of the month were classified as non-participating and were joined to the unit non-response. Households that returned all reporting notebooks and participated fully to all telephone interviews were considered as responding completely.

Therefore item non-response is not very frequent in the HBS98 data. It concerns mainly the income data. Missing data were detected based on supplementary information like the economic activity status given by the participating persons themselves (implying a certain type of income) or simply based on a *don't want to give*-declaration of the persons. For most households only part of the income is missing while part of it is given (see Figure 5.4). This missing part was imputed by a regression model based on the total income per household explained by the total expenditures per household.

For the regression models, the income and expenditures data had to be transformed logarithmically for obtaining more symmetric distributions. The regression was done separately for each socio-economic group in a two-step process. First, a L1 (or LAD) regression was applied in order to identify outliers. After excluding those outliers, a classic L2 (or LS) regression was applied as a second step to the remaining points in order to obtain the coefficients for imputing. The missing income was obtained by subtracting the given income from the imputed total income. In cases where this difference was negative, the value was not imputed (see left part of the distributions in Figure 5.4).

In the following, we suggest simulating item non-response in the HBS98 universes by considering the imputed values as the true missing values. Based on the distribution and on the amounts of these *true missing* values, we can estimate two parameters for item non-response. On the one hand, we can calculate the probability of item non-response per household (meaning that the household has some income missing). On the other hand, we can also calculate the probability of how much of the income is missing (and thus reduce the given income by this percentage). Based on these ideas, we propose three models of different levels of complexity in Table 5.7.

The first **model** (**E, simple**) takes into account only the general percentage of missing income compared to the total income of all households (NR = 1.0827%). It suggests selecting completely at random (MCAR, no auxiliary variables used) some households with the given probability and to put their complete income to missing. The imputation model would have to be simplified for this model (→ without condition) since there is no partially given income.

In the **intermediate model (F)**, non-responding households would also be selected completely at random (NR = 5.4330 %). However, the income would not be put completely to missing but only a fraction of it. This fraction is calculated using the following formula (there is again no auxiliary variable used):

$$\text{part of missing income} = \mathbf{RND}^{[1/p]-1}$$

where ***RND*** is a uniformly distributed random number between 0 and 1 and p is the average part of missing income amongst households with missing income: 17.41 % for all households together.

The function $f(x) = x^{[1/p]-1}$ was chosen because it allows to convert the Uniform distribution into a distribution where the missing part of income is distributed as such that its average corresponds to the observed average missing part (parameter p) of all households as is shown in Figure 5.3.

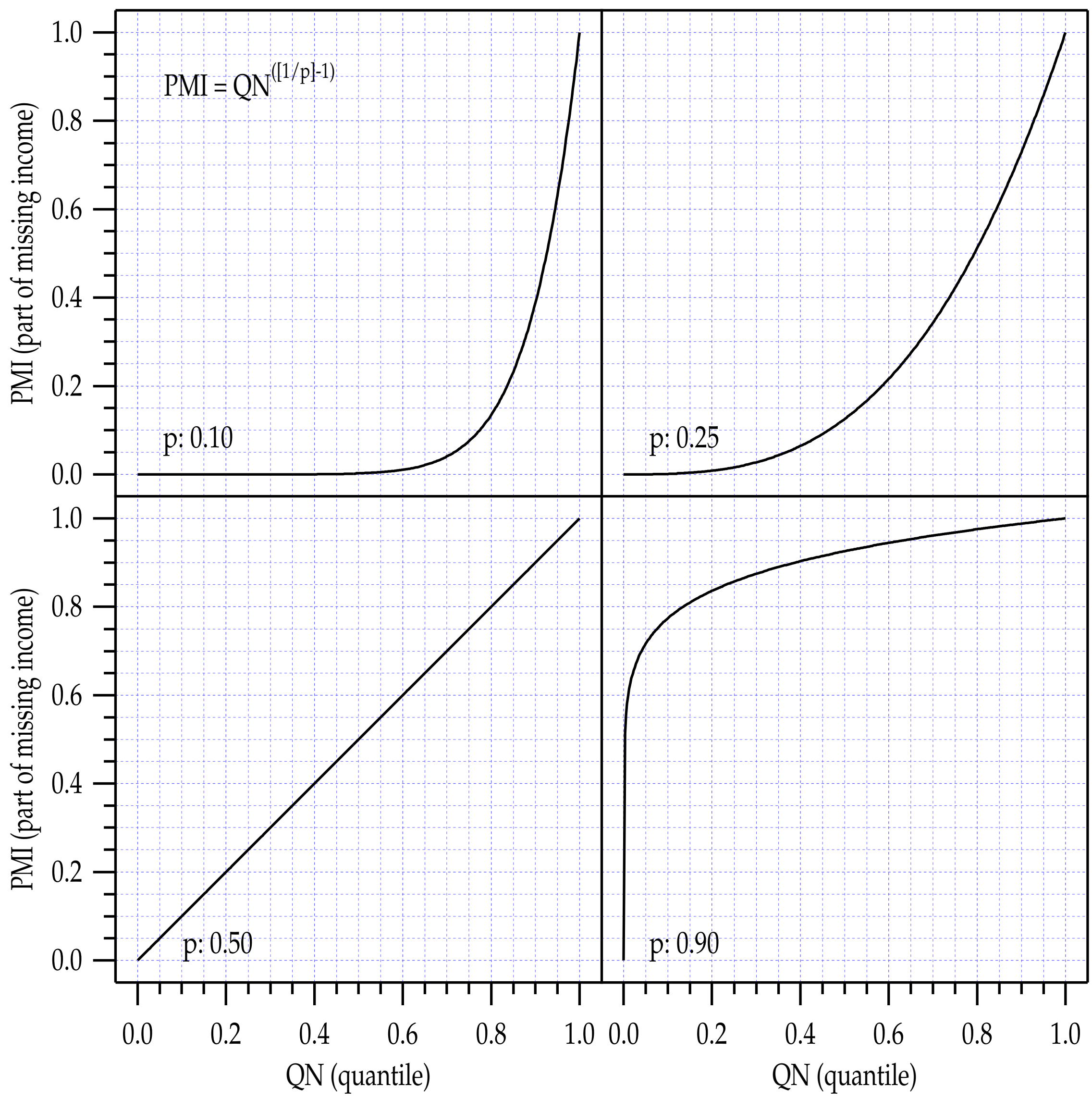


Figure 5.3: Four examples of the function $f(x) = x^{[1/p]-1}$ with $p = 0.1, 0.25, 0.5$ and 0.9.

In the **complex model (G)**, the same idea as in model (F) is used. However for this model, both non-response parameters (NR_i and p_i) depend on an auxiliary variable (STASOCIO). This model is therefore rather of the MAR type (missing at random).

Again, not the complete income is put to missing but only a fraction that is calculated using the formula:

$$\text{part of missing income} = \mathbf{RND}^{[1/p_i]-1}$$

Table 5.7: The suggested item non-response models.

	Simple Model (E)	Intermediate Model (F)	Complex Model (G)
Item non-response: INCOME	Complete INCOME missing for non-responding households	Partial INCOME missing for non-responding households	Partial INCOME missing for non-responding households
Aux. variable: STASOCIO	(not used)	(not used)	Item non-response probability for each class
Non-response probability per household	1.0827 % all households	5.4330 % all households	[-1] Others: 6.0914% [1] Employees: 5.3528% [2] Self-employed: 9.0028% [3] Farmers: 22.4138% [4] Unemployed: 6.3492% [5] Pensioners: 1.9333%
Missing part of INCOME	100 %	$RND^{(1/0.1741)-1}$ where $RND \sim U[0;1]$	$RND^{(1/p_i)-1}$ where $p([-1]) = 0.5324$ $p([1]) = 0.1557$ $p([2]) = 0.2201$ $p([3]) = 0.2347$ $p([4]) = 0.2620$ $p([5]) = 0.1578$
Number of parameters	1	2	$2 \cdot 6$
Imputation model	Always complete imputed value	Imputation only if imputed value is greater than non-missing part of INCOME	

For both models with resulting partial missing income (F and G), the imputation must be implemented as it was done with the real data: impute the total income only under the condition to increase the income compared to the given non-missing part of the income.

5.6 Effect of simulating non-response at the universe stage

For the sake of clarifying the main effect, we consider the simplest possible case of simple random sample (SRS) of size n from a universe (population) of size N and random non-response, i.e., missing completely at random, i.e.. We consider the problem of estimating the population mean $\overline{Y}$ of some variable y of interest, $\overline{Y} = \sum_{i=1}^{N} y_i/N$. The estimator

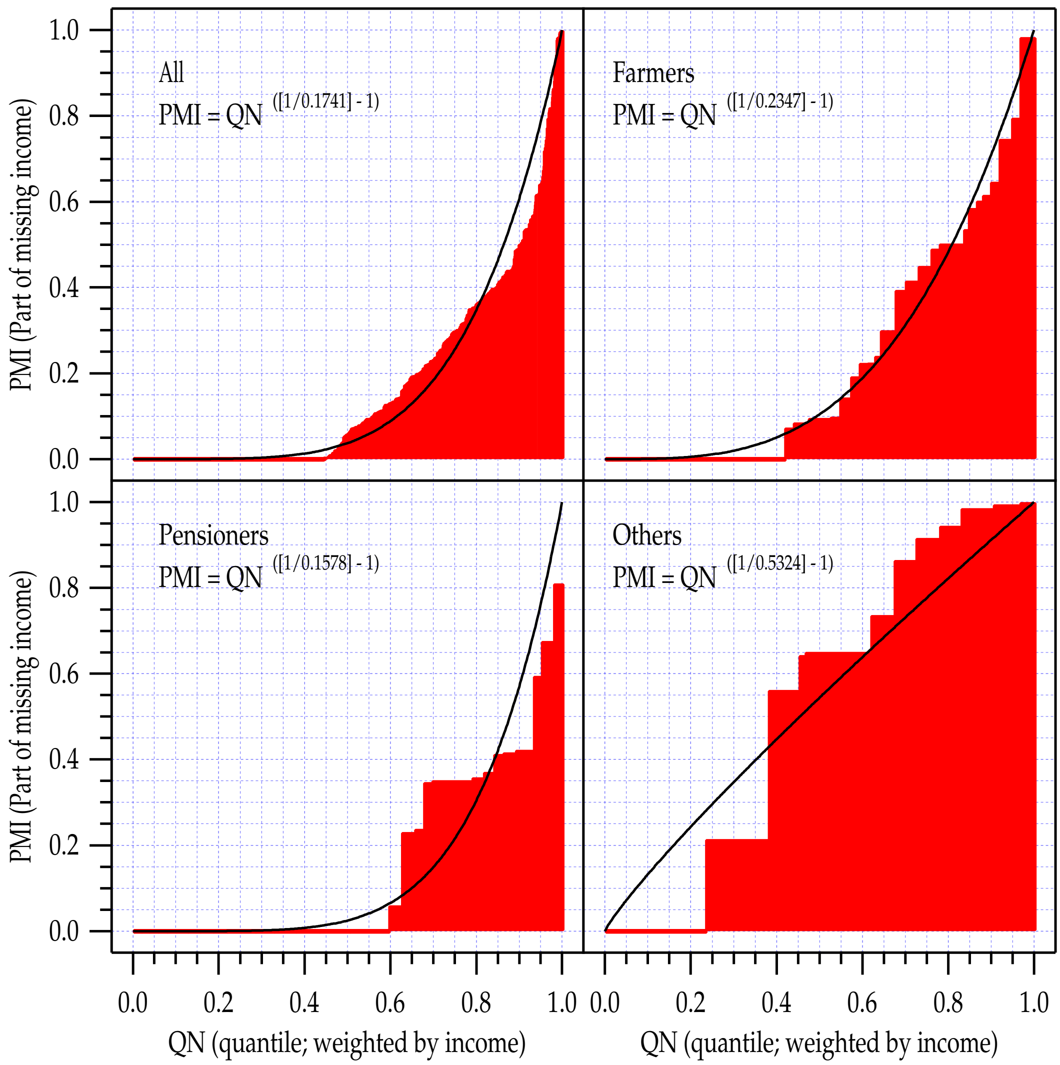


Figure 5.4: The observed distributions of the part of the imputed (= missing) income of the population with detected missing income (5.43% of all households) for model F (→ All) or G (three examples). The solid lines indicate the distributions of the missing part of the income according to the suggested model.

is the observed mean $\overline{y}_r$ from the response sample. Let $\mathbf{R} = (R_1, R_2, \ldots, R_N)$ be the response indicators, where $R_i = 1$ if unit i in the population respond and 0 otherwise.

Random non-response means that the R_i's are independent with the same response probability,

$$\mathrm{P}\left(R_i = 1\right) \ = p.$$

If response is generated for the universe in the simulation study, it means that $\mathbf{R} = \mathbf{r}$ is given and one simulation study gives us estimates of $\mathrm{Var}(\overline{y}_r|\mathrm{r})$ and $\mathrm{E}(\overline{y}_r|\mathrm{r})$. The main purpose is to estimate the unconditional variance, $\mathrm{Var}(\overline{y}_r)$. To assess the effect of generating non-response at the universe stage, we shall study the expected variance ratio of the conditional and unconditional variance,

$$\frac{\mathrm{EVar}(\overline{y}_r|\mathrm{R})}{\mathrm{Var}(\overline{y}_r)}.$$

Let $f = n/N$ be the sampling variance and $\sigma^2 = \dfrac{1}{N-1}\sum_{i=1}^{N}(y_i - \overline{Y})^2$ the population variance. The two main results state:

$$\frac{\mathrm{EVar}(\overline{y}_r|\mathrm{R})}{\mathrm{Var}(\overline{y}_r)} \approx \frac{1-f}{1-pf} \quad \in (1-f, 1) \tag{5.3}$$

and

$$\mathrm{Var}(\overline{y}_r) - \mathrm{EVar}(\overline{y}|\mathrm{R}) \approx \frac{\sigma^2}{pn} f(1-p) \tag{5.4}$$

as a sort of first-order approximation, to the order of $1/n$ which typically is very accurate for sample sizes of $n \geqslant 1000$.

A second-order approximation to the order of $1/n^2$ is given by

$$\frac{\mathrm{EVar}(\overline{y}_r|\mathrm{R})}{\mathrm{Var}(\overline{y}_r)} \approx \frac{1-f}{1-pf} \cdot (1 + \delta_n/n) \tag{5.5}$$

and

$$\mathrm{Var}(\overline{y}_r) - \mathrm{EVar}(\overline{y}|\mathrm{R}) \approx \frac{\sigma^2}{pn} f(1-p) + \frac{\sigma^2}{(pn)^2} f^2(1-p) \tag{5.6}$$

where

$$\delta_n = \frac{1-p}{p} \cdot \frac{\left(1 + f - \frac{1}{1-pf}\right)}{1 + \frac{1}{np} \cdot \frac{1-p}{1-pf}}.$$

We note that $\delta_n \geqslant 0 \Leftrightarrow f \leqslant (1-p)/p \Leftrightarrow p \leqslant 1/(1+f)$which typically is always the case for household and person surveys where f is very small. In this case,

$$\delta_n \leqslant \lim_{n\to\infty} \delta_n = \frac{(1-p)f}{p}\left(1 - \frac{p}{1-pf}\right) = \delta.$$

We note that this upper bound is less than $1-f$ if $f \leqslant p$, which of course is typically the case in household/person surveys.

We have of course that $\delta_n < 0 \Leftrightarrow f > (1-p)/p \Leftrightarrow p > 1/(1+f)(\Rightarrow p > 0.5)$ which may occur for business surveys. In this case

$$\delta_n > -\frac{(1-p)f}{p}\left(1 - \frac{p}{1-pf}\right) = -\delta$$

Some illustrations

Case 1) $\underline{f = 0.001,\ p = 0.6}$

$$\frac{\mathrm{EVar}(\overline{y}_r|\mathrm{R})}{\mathrm{Var}(\overline{y}_r)} \approx \frac{1-f}{1-pf} = .9996$$

2nd order approximation gives the same result to 5 decimal places for $n \geqslant 100$.

Case 2) *f = 0.5, p=0.6*

$$\frac{\mathrm{EVar}(\overline{y}_r|\mathrm{R})}{\mathrm{Var}(\overline{y}_r)} \approx \frac{1-f}{1-pf} = .7143$$

2nd order approximation gives the same result to 4 decimal places for $n \geqslant 1000$ and for $n = 100$, the variance ratio equals 0.7146.

So we see that when the sampling fraction is ignorable one gets approximately the correct variance when conditioning on **r**, i.e., one can simulate then non-response at the universe stage in the simulation study. However, when f is not ignorable the conditional variance one obtains from generating non-response at the universe stage severely understates the actual variance of the estimator.

The main results can be derived as follows:

I) $\mathrm{Var}(\overline{y}_r)$

The response sample is denoted s_r with size n_r. Then conditional on n_r, s_r is a SRS from the universe U. This implies that

$$\mathrm{E}(\overline{y}_r|n_r) = \overline{Y} \text{ and } \mathrm{Var}(\overline{y}_r|n_r) = \frac{\sigma^2}{n_r} \cdot \frac{N-n_r}{N}$$

This implies that

$$\mathrm{Var}(\overline{y}_r) = \mathrm{EVar}(\overline{y}_r|n_r) = \sigma^2(\mathrm{E}(1/n_r) - 1/N)$$

Now, n_r is binomial (n,p) and a 2^{nd} order Taylor approximation around np gives us:

$$\mathrm{E}\left(\frac{1}{n_r}\right) = \frac{1}{np} + \frac{1-p}{n^2p^2}.$$

It follows that

$$\mathrm{Var}(\overline{y}_r) \approx \frac{\sigma^2}{pn}(1-pf) + \frac{\sigma^2(1-p)}{(np)^2} \tag{5.7}$$

II) $\mathrm{Var}(\overline{y}_r|\mathrm{r})$ and $\mathrm{EVar}(\overline{y}_r|\mathrm{R})$

Given **r**, let $\overline{Y}_r$ and σ_r^2 be the population mean and variance in the response universe U_r.

Now conditional on **r** and n_r, s_r is a SRS from U_r. This implies that

$$\mathrm{E}(\overline{y}_r|\mathrm{r}, n_r) = \overline{Y}_r \text{ and } \mathrm{Var}(\overline{y}_r 1\mathrm{r}, n_r) = \frac{\sigma_r^2}{n_r} \cdot \frac{N_r - n_r}{N_r}$$

Hence,

$$\mathrm{Var}(\overline{y}_r|\mathrm{r}) = \mathrm{EVar}(\overline{y}_r|\mathrm{r}, n_r) = \sigma_r^2\left(\mathrm{E}\left(\frac{1}{n_r}\middle| r\right) - \frac{1}{N_r}\right)$$

Given $\mathbf{r}$, n_r is hypergeometric with expected value nq_r where $q_r = N_r/N$. Again, using 2^{nd} order Taylor approximation for $\mathrm{E}(1/n_r|\mathbf{r})$ we find that

$$\mathrm{Var}(\overline{y}_r|\mathrm{r}) \approx \frac{\sigma_r^2}{nq_r}(1-f) + \frac{\sigma_r^2(1-q_r)}{(nq_r)^2}(1-f)$$

Using only the first term

$$\mathrm{EVar}(\overline{y}_r|\mathrm{R}) = \mathrm{E}(\sigma_r^2/q_r)(1-f)$$

Conditional on N_r, U_r is a SRS from U. Hence, $\mathrm{E}(\sigma_r^2|N_r) = \sigma^2$. Since N_r is binomial $(N{,}p)$ we find

$$\mathrm{EVar}(\overline{y}_r|\mathrm{R}) \approx \frac{\sigma^2}{pn}(1-f) + \frac{\sigma^2(1-f^2)(1-p)}{(np)^2} \tag{5.8}$$

Using the first terms in (5.7) and (5.8) gives us (5.6) and (5.4), while (5.5) and (5.6) follow from the 2^{nd} order expressions in (5.7) and (5.8).

Appendix A

Miscellaneous

A.1 Dissemination of computer code

The pseudo-code will consist of a head and a body. The head will include input and output variables, additional data, NR and imputation scheme applied, and a description of further variables needed for the pseudo-code. The body will consist of the general algorithm. In a first step, the dissemination should consist of plain ASCII file and the algorithm R-code, where specific R-code or syntax should be avoided or at least be described in detail.

A.2 Example for metadata and code dissemination

The Generalised Regression Estimator

$$\widehat{\tau} = \sum_{i \in \mathcal{S}} \frac{1}{\pi_i} \cdot y_i + \left(\sum_{i \in \mathcal{U}} X_i + \sum_{i \in \mathcal{S}} \frac{1}{\pi_i} \cdot X_i \right) \cdot \widehat{B} \quad ,$$

where $\widehat{B}$ is the coefficient vector of the (weighted) least squares estimate.

Design: Simple random sampling

Model: With and without replacement

Purpose: Total

Estimation variable: 1, metric (individual level, counts in cluster possible)

Auxiliary variable: n, metric (c.f. estimation variable)

References: Särndal *et al.* (1992), Lohr (1999)

Software Packages: S-Plus / R

Additional remarks: No.

```
##########################################################################
##   GREG total estimator                                               ##
##########################################################################
##########################################################################
##  Variables:
##
##  y        values of variable of interest            n
##  x        matrix of aux variables                   n x #covar
##  w        weights (eg for compensating NR)          n
##  iip      inverse inclusion porbabilites (1st order) n
##  Xtot     matrix of true total values               #covar
##  N        number of units in universe
##  W        flag for weighted regression
##  V        flag for including variance estimate (F -> NA)
##
##  Variables in programme:
##
##  n        number of units in sample
##  ll       number of regression coefficients (#covar)
##  B        regression coeff. from Y on
##               incl. weights w                       ll = #covar
##
##  Return values:
##
##  GREG.strat.total       : GREG estimate
##  GREG.strat.total.var   : approx. variance estimate for GREG
##
##########################################################################

GREG.total <- function(y, x, w = NULL, iip = NULL, W=F, Xtot, N, V=F) {

    x <- as.matrix(x)
    n <- length(y)
    ll <- dim(x)[2]
    if (length(w) == 0) w <- rep(1,n)
    if (length(iip) == 0) iip <- rep(N/n,n)

    if (W) {
        lreg <- lm(y ~ x, weights = w)
    } else {
        lreg <- lm(y~x)
    }

    B <- as.matrix(coef(lreg)[2:(ll+1)])
    GREG <- as.numeric(sum(iip * y) + (as.matrix(Xtot) - as.matrix(apply(x
        ,2,function(Z) sum(iip*Z)))) %*% B)

    GREGvar <- ifelse(V, sum(N^2 * (sum(lreg$residuals)^2 / (n-1)) / n *
        (1-n/N)), NA)

    list(GREG.strat.total = GREG, GREG.strat.total.var = GREGvar)

}
```

Appendix B

Summary tables of the questionnaire

B.1 The questionnaire

A. National Surveys on household and individual data

1. Which regular sample surveys on household and individual data do you/does your institution apply in your country?
2. Are those surveys voluntary or mandatory ?
3. Are the samples drawn in one or several periods? Please indicate number and time of periods.
4. Are the samples organised as panels, cross sectional or rotational samples?
5. What are the non-response rates?
6. How are the universes defined?
7. Please indicate the survey units.
8. Please give a rough description of the sample selection schemes.
9. Are the variables of interest categorical, metric or both?
10. Do your surveys allow for surveying household as well as individual data? If only individual data are available, is it possible to identify the underlying households?
11. Are you using variance estimation methods for estimates? If yes, could you give a short description?
12. Do you intend to modify your existing surveys or to implement new surveys in the future?

B. Variance estimation software in use

1. Do you use central organised software (e.g. mainframe)?

2. Which statistical software and which version do you/does your organisation use for variance estimation?

3. Is this software also used for variance estimation within the complex surveys mentioned in Part A? If more than one software package is applied, please relate them to the respective surveys.

4. Did you implement additional software routines in your software packages concerning variance estimation?

5. Do you intend to apply new software packages or new versions?

B.2 Overview to Evaluation of the questionnaire

Table 1: Abbreviations

Abbreviation	Meaning
add.	additional
approx.	approximation
dw	dwellings
est.	estimation
GREG	general regression estimator
HH	households
instit.	institutional
npq	variance of n independent Bernoulli trials
municip.	municipalities
pers.	person
pop.	population
PPS/πPS	proportional probability sampling
PSU	primary sampling unit
resid. econ. act.	resident economically active
rot.	rotating
s.a.	see above
STRWR	stratified random sampling with replacement
StS	stage sample
strat.	stratified
suppl.	supplementary
sw	software
syst.	systematically
var.	variance

Table 2: Evaluation of the Questionnaire (14 pages)

(0) country	(1) survey	(2) mandate	(3) period	(4) sample organisation	(5) Non-Response	(6) universe	(7) unit	(8) design
AT Austria	Austrian Microcensus	core: mandatory specials: voluntary	annual	one eight of sample is quarterly rotated	10% - 15%	non-institutional population	persons, families, households, dwellings	2-StS; 1. municipalities 2. dwellings 1-StS; with syst. selection for urban domain
BE Belgium	HBS	voluntary	communities: annual households: quarterly	continuous; monthly, weekly	90%	all private households	households and individuals	3-StS; 1. sections (NUTS 1) 2. households 3. individuals
	LFS	mandatory	HH quarterly; geographical sections: all 3 years	continuous; weekly sub-sample	16%	all private households	households and individuals	3-StS; 1. sections (NUTS 2) 2. households 3. individuals
	Travel Survey	voluntary	no	panel, 4 times	about 50%	all private households	households and individuals	187 strata defined by NUTS 2 regions, HH selcted by simple random sampling
BG Bulgaria	sample survey (HBS)	voluntary	monthly	cross-sectional	5.6%	all HH from population census	HH; 6000 random sampled	2-StS; 1. clusters 2. households
CZ Czech Republic	HBS	voluntary	continuous survey	quota sample, panel-like rotation		set of private households	private households	quota sampling stratified by: social group, net income, children, HH- members
	Microcensus	voluntary	3 - 5 years	cross-sectional	about 25%	set of permanent occupied dwellings	persons in dw inclusive long term absentees	2-StS; 1. census districts-clusters 2. dwellings
DK Denmark	HBS	voluntary	every 3 months	cross-sectional	about 35%	all private households	households	probability sample within geographical clusters
EE Estonia	HBS	voluntary	once a year	rotational	about 40%	households	households, HH- members	systematic stratified sample; stratification variable: regional indicator
	LFS	voluntary	once a quarter	rotational	10% - 13%	persons in the age of 15 - 74	HH- members in the age of 15 - 74	systematic stratified sample; stratification variable: regional indicator

(0) country	(1) survey	(2) mandate	(3) period	(4) sample organisation	(5) Non-Response	(6) universe	(7) unit	(8) design
FI Finland	Consum Barometer	voluntary	quarterly	3 waves, lag 6 months	30%	resid. econ. act. pop., age 15 - 74	individuals	systematic sub-selection on LFS sample
	ECHP	voluntary	yearly	2 wave rotating panel	25% - 30%	non-institutional households	households and individuals	estimation systematic π PS sampling
	HBS	voluntary	yearly		35%	non-institutional households	HH and individuals	regional stratified SRSWOR, est. syst. π PS sampling
	Income Distribution	voluntary	yearly	2 wave rotating panel design	15% - 30%	non-institutional households	households and individuals	stratified SRSWOR on key persons; network smaple: est. syst. π PS sampling
	LFS	voluntary	monthly	5 wave rot. panel design	14%	resid. econ. act. pop., age 15 - 74	individuals	systematic selection
	National Border Survey	voluntary	continuous		10% - 20%	foreigners visiting Finland	vehicle/ person	selection of persons from the flow leaving Finland
	Time Use Survey	voluntary	about 10 years		40%	resid. non-instit. pers., age 15 - 74	HH and individuals	regional stratified SRSWOR est. syst. π PS sampling
FR France	Business Cycle and Living Costs	mandatory	monthly expect August	rotating	about 28%	telephone survey, set of dwellings	individuals in dwellings	
	Household Survey	voluntary	3 times a year	rotating	about 22%	individuals elder 14	individuals elder 14 in dwellings	master sample with stratification
	LFS	mandatory	weekly	rotating	about 12%	individuals elder 14		cluster sampling of areas
	Rents and Add. Costs	mandatory	Jan., April, July, October	rotating	about 14%	dw: urban areas and tenant HH	dwellings	master sample with stratification
DE Germany	Survey of Income and Expenditure	voluntary	every 5 years	partly split in 4 parts for quarters		private HH, net income under 35000 DM	households and individuals	quota samples
	HBS	voluntary	once a year	rotating panel		private HH, net income under 35000 DM	households and individuals	quota samples
	German Microcensus	mandatory	annual	rotating sample	about 3%	all private HH, individuals living permanently in Germany	households and individuals	stratified random sample of sample areas, fraction 1%

(0) country	(1) survey	(2) mandate	(3) period	(4) sample organisation	(5) Non-Response	(6) universe	(7) unit	(8) design
EL Greece	LFS	mandatory	quarterly	rotational	5% - 10%	private households	households	2-StS; 1. random sample of small areas 2. syst. sample of HH
IE Ireland	HBS	voluntary	conducted every 7 years	rotational	55%	private households	households	2-StS; with rotation in second stage
	Quarterly National HH Survey	voluntary	continuous quarterly	rotating sample	about 75%	private households	persons	2-StS; 80% quarterly overlap in stage 2
IL Israel	Family Expenditure Survey	mandatory in practice: voluntary	once a year in addition: samples for new dwellings and immigrants	cross-sectional	15%	permanant population aged over 14 exclusive Bedouin tribes and institutional persons	dwellings; households and persons elder 15	2-StS; 1. localities over 5000 HH by PPS 2. dwellings systematically chosen
	Income Survey			appended to LFS	22%			fourth investigation of LFS
	LFS			cross- sectional with rotating panel design	9%	s.a., in addition: survey for Bedouin tribes and institutional pers.		2-StS; 1. localities over 5000 HH by PPS 2. dwellings systematically chosen
	Social Survey			cross-sectional	yet started	see Family Expenditure Survey	dw; HH and pers. elder 20	strat. sample drawn from Nat. Population Register
IT Italy	Consumer Expenditure Survey	mandatory	annual in November	cross-sectional	18%	households	households	2-StS; 1. municipalities 2. households and persons in HH
	ECHP	mandatory	annual in November	panel survey		Italian residents in households older 14 years	households	2-StS; 1. municipalities 2. households and persons in HH
	LFS	mandatory	annual in January	rotational	10%	Italian residents in households	households and individuals	2-StS; 1. municipalities 2. households and persons in HH
	Multi Purpose HH Survey	mandatory	annual in November	cross-sectional	20%	households and individuals	households	2-StS; 1. municipalities 2. households and persons in HH

(0) country	(1) survey	(2) mandate	(3) period	(4) sample organisation	(5) Non-Response	(6) universe	(7) unit	(8) design
LT Lithuania	HBS	voluntary	once a year	rotational	20% - 30%	private HH	households	2-StS; sample with probability proportional to size of cluster and households
	LFS	voluntary	twice a year	rotational	10% - 15%	residents of Lithuania elder 14	persons	simple random sample of 3000 members of population register
NL Netherlands	Consumer Survey	voluntary	monthly	cross-sectional	35%	private HH	households	2-StS; 1. sample of municip. 2. sample of addresses
	Income Panel Survey	voluntary	yearly	panel	2%	frame: population register	individuals	simple random survey of 0.61% of population
	Integrated Continuous Survey on Living Conditions	voluntary	continuous monthly	cross-sectional	40%	persons from population register of municipalities	individuals of all ages	2-StS; 1. sample of municip. 2. sample of persons
	LFS	voluntary	continuous monthly	rotational	45%	non-instit. pop. in Netherlands	households	2-StS; 1. sample of municip. 2. sample of addresses
	Housing Demand Survey	voluntary	once in 4 years	cross-sectional	1998- 2000: 44,7- 47,6%	drawn from popul. register of municipalities	HH and persons aged 18 and over	random sample among all persons aged 18 and over
	Survey on Daily Recreational Activities	voluntary	every 5 years	cross-sectional	50%	non-institutional population in Netherlands	individuals of all ages	2-StS; 1. sample of municip. 2. sample of persons within these municipalities
	Socio- Economic Panel Survey	voluntary	yearly additions since 1984	panel	12%; add sample: 70%	private HH	HH and persons	2-StS; 1. sample of municip. 2. sample of addresses

(0) country	(1) survey	(2) mandate	(3) period	(4) sample organisation	(5) Non-Response	(6) universe	(7) unit	(8) design
NO Norway	Household Consumption Survey	register sampling		based on register	30% - 40%		households and individuals	PPS
	LFS	register sampling		cluster sample base: register	10% - 15%		individuals	stratified systematic cluster sample based on register
	Survey of Living Conditions	register sampling		based on register	20% - 23%		individuals	2-StS; with one PSU per stratum, systematic sampling within PSU
	Survey of Income and Assets	register sampling		union of SLC and HCS				union of HH Consumption Survey and Survey of Living Conditions
PT Portugal	Demand Tourism Survey	mandatory	quarterly	rotational	12%	persons elder 14 in private households	sample: dwellings; measurement: households/ persons	Sub-sample of Master sample 1989, Master sample: cluster sample with 2- stage design
	ECHP	mandatory	only one time	panel		set of households		
	HBS	mandatory	every 5 years	cross-sectional	20%	set of households		
	LFS	mandatory	each quarter	rotational	10%	set of persons in private HH		
ES Spain	HBS	mandatory	quarterly	rotational	about 16%	private households	private households	2-StS; 1. census areas (PPS) 2. private HH (PPS)
	LFS	mandatory	quarterly	rotational	about 8%	population in private HH	private households	

(0) country	(1) survey	(2) mandate	(3) period	(4) sample organisation	(5) Non-Response	(6) universe	(7) unit	(8) design
SI Slovenia	HBS	voluntary	quarterly	cross-sectional	17%	private HH and all members	households and individuals	stratified samples; for small settlements: 2-StS; for lager settlements: syst. sample
	LFS	voluntary	quarterly	panel rota-. ting sample	12%	private HH and all members	HH and individuals	stratified systematic sample by regions and size of settlement
	Consumer Attitude Survey	voluntary	drawn in one period	cross-sectional	37%	persons elder 15	individuals	stratified systematic sample by regions and size of settlement
	Crime Victimisation Survey	voluntary		cross-sectional	33%	persons elder 15	individuals	
	Quarterly Survey on Travels of Domestic Population	voluntary		cross-sectional	25%	persons elder 14	individuals	
SE Sweden	LFS	voluntary	quarterly	rotational	15%	individuals of age 18 - 64	using register of total population; when survey unit is HH, then individuals are included, else only individuals	stratification and simple random sampling without replacement in each stratum; sometimes: systematic sampling
	Activities after Education fs†	voluntary	quarterly or yearly	normally cross-sectional	normally 30% - 35%	normally individuals of age 18 - 64, some survey: under limit 0 years, upper limit 85 years		

† Labor Market Entry And Continuity; EU and EMU Sympathies; General Elections; Income and Income Distribution; Living Conditions; Occupational Injuries; Party Preferences; Pupil Follow-up; Pupil Panels for Longitudinal Studies; Staff Training Statistics; Time Use Survey; Work Environment

(0) country	(1) survey	(2) mandate	(3) period	(4) sample organisation	(5) Non-Response	(6) universe	(7) unit	(8) design
CH Switzerland	Dwelling Survey					telephone register for most mentioned surveys	households or individuals	stratified random samples, NUTS- regions as stratification variable
	Family Survey							
	HBS	voluntary	quarterly	cross-sectional	level 1: 50%, level 2: 30%			
	LFS	voluntary		rotational				
	National Transport Survey		quarterly	cross-sectional				
	Rental-fee Survey							
	Swiss Survey on Health	voluntary	quarterly					
UK United Kindom	Family Expenditure Survey	voluntary	quarterly	cross-sectional	40%	private households in UK	households	3-StS; 1. areas 2. addresses 3. HH from post-code address file
	General Household Survey	voluntary	quarterly	cross-sectional	33%	private households in UK	households	
	International Passenger Survey	voluntary	quarterly	cross-sectional	19%	passengers at all ports in Great Britain	individuals	2-StS; 1. time periods 2. passengers stratified by port, season, day of week
	LFS	voluntary	quarterly	5 waves rotating panel	25%	private households in UK	adults	single stage systematic cluster sample of addresses
	Omnibus Survey	voluntary	quarterly	cross-sectional	35%	individuals in private households in Great Britain	adults	3-StS; 1. areas 2. addresses 3. HH from post-code address file

(0) country	(9) variable	(10) data: HH/ individuals	(11) variance estimation	(12) change of design intended	B (1) cental software	B (2) software for var. est.	B (3) software for surveys	B (4) suppl. sw-routines	B (5) new software
AT Austria	mainly categorical, some metric	households and individuals	similar to German MZ	transition to continuous survey	yes	SAS and PL/I-programs	SAS for MZ, PL/I for economic surveys	no	no
BE Belgium	majority metric	households and individuals	no	no	yes	occasionally with SPSS, Excel	SPSS 9/ 10		compare g- CALIB 1.0 with CALMAR; study POULPE
	majority categorical	households and individuals	simple estimation for categorical variables, npq- like formula	no					
	majority categorical	HH and individuals	no	no					
BG Bulgaria	categorical and metric	HH and individuals	no	no					
CZ Czech Republic	categorical and metric	households	no	possibly new survey on living conditions	no	SAS 6.01	no	no	no
	categorical and metric	households and individuals	no	no					
DK Denmark	categorical and metric	HH and individuals	no	no					
EE Estonia	categorical and metric	households and individuals	totals: Horvitz- Thompson; means: ratio estimator		no	SUDAAN 7.1, CLAN 97, taylored programs	SUDAAN	no	new versions of SUDAAN
	categorical and metric	households and individuals							

(0) country	(9) variable	(10) data: HH/ individuals	(11) variance estimation	(12) change of design intended	B (1) cental software	B (2) software for var. est.	B (3) software for surveys	B (4) suppl. sw-routines	B (5) new software
FI Finland	categorical	individuals			no				
	categorical and metric	HH and individuals	not yet						
	categorical and metric	HH and individuals	design and calibration; STRWR				SUDAAN		new version
	categorical and metric	HH and individuals	design and calibration; STRWR				SUDAAN		
	categorical and metric	households and individuals	design and calibration; GREG				CLAN		
	categorical and metric	individuals	design and calibration; GREG				CLAN		
	categorical and metric	HH and individuals	design and calibration; STRWR				SUDAAN		
FR France	categorical	individuals	not commonly used; ad- hoc software: POULPE		no				investigate/ improve existing software
	mainly categorical	HH and individuals		Europ. survey on living conditions					
	mainly categorical	mainly individuals							
	mainly metric	households							
DE Germany	mainly metric	households and individuals	regression estimator for stratified random sample as approx. to calibration technique	harmonisation, integration of Income and Expediture Survey and HBS	no	self- made standard software	no	no	SAS, CLAN
	mainly metric	HH and individuals	no						
	categorical	households and individuals	classical Horvitz-Thompson Estimator for strat. random sampling as approximation	converting to continuous survey					

(0) country	(9) variable	(10) data: HH/ individuals	(11) variance estimation	(12) change of design intended	B (1) cental software	B (2) software for var. est.	B (3) software for surveys	B (4) suppl. sw-routines	B (5) new software
EL Greece	mainly categorical	households and individuals	yes	no	no official software				
IE Ireland	categorical and metric	households and individuals	simple replication method		mainframe and PC network	SAS			investigate CLAN and GREGWT for variance estimation
	categorical and metric	HH and individuals	no						
IL Israel	categorical and metric	households and individuals	surveys calculated with estimated sampling errors; Post-stratification, Jackknife, random groups method (Wolter)	no	own software with SAS	own SAS-routines for totals and ratios	SAS tools according to estimate procedures	several in- house procedures	on demand new software will be developed
	categorical and metric	HH and individuals							
	categorical and metric	households and individuals							
	categorical and metric	individuals							
IT Italy	categorical and metric	households and individuals	linearisation method and generalised regression estimator (Taylor expansion)	no	own software; **GENESEES-GSSE**	new version of **GENESEES-GSSE**		yes	subroutine for variance estimation; Jackknife linearisation
	categorical and metric	HH and individuals		no					
	categorical and metric	HH and individuals		redesign of LFS					
	categorical and metric	HH and individuals		new survey to substitute MPHS					

(0) country	(9) variable	(10) data: HH/ individuals	(11) variance estimation	(12) change of design intended	B (1) cental software	B (2) software for var. est.	B (3) software for surveys	B (4) suppl. sw-routines	B (5) new software
LT Lithuania	metric	households and individuals	experiments are done	use data of population census to compensate Non-Response	no	SAS callable SUDAAN, CLAN 97, own programs	SUDAAN may be used	yes	no
	categorical and metric	households and individuals	Post-stratification; estimation according to sampling strategy	calibrated estimation will be used			CLAN, SAS, self written programs		
NL Netherlands	categorical	households	only incidentally to determine optimal regional sample size	modification in 2002	BASCULA for weighting, own software	BASCULA 3, SPSS 8.0	variance estimation module developed for complex designs	additional routines for variance estimation	BASCULA may be updated
	categorical and metric	HH and individuals		no					
	categorical	households and individuals		no					
	categorical	HH and individuals		no					
	categorical and metric	HH and individuals		no					
	categorical and metric	individuals		may be included in a wide-ranging survey					
	categorical and metric	HH and individuals		yes, modification					

(0) country	(9) variable	(10) data: HH/ individuals	(11) variance estimation	(12) change of design intended	B (1) cental software	B (2) software for var. est.	B (3) software for surveys	B (4) suppl. sw-routines	B (5) new software
NO Norway	categorical and metric	households and individuals	no variance estimation; PSU per stratum design call for collapsed stratified design						
	categorical	households and individuals	no var. est.; design and est. method is not well adapted to analytic est. of sample variance						
			no variance estimation						
			no variance estimation						
PT Portugal	categorical and metric	households and individuals	Jackknife techniques	methodological review	no	SAS 6.12 for UNIX	SAS marco CALJACK	no	yes; testing BASCULA
	categorical and metric	HH and individuals		will be substituted by EU- SILC					
	categorical and metric	HH and individuals							
	categorical and metric	HH and individuals							
ES Spain	metric	households	ultimate cluster	modifications with data from last census	mainframe	SAS 6.12, PL/I	own application for these software	no	no
	categorical	individuals and dw	replicated half samples						

(0) country	(9) variable	(10) data: HH/ individuals	(11) variance estimation	(12) change of design intended	B (1) cental software	B (2) software for var. est.	B (3) software for surveys	B (4) suppl. sw-routines	B (5) new software
SI Slovenia	categorical and metric	households and individuals	Taylor linearisation; design: with replacement		no	SUDAAN 7.0	yes	no	yes
	categorical and metric	HH and individuals							
	categorical and metric	individuals							
	categorical and metric	individuals							
	categorical and metric	individuals							
SE Sweden	categorical and metric	when survey unit is HH, then individuals are included, else only individuals	standard formulas, Taylor linearisation	probably	yes	CLAN 97 developed by Statistics Sweden as complement to SAS	yes	CLAN 97 routines	no
	categorical and metric								

(0) country	(9) variable	(10) data: HH/ individuals	(11) variance estimation	(12) change of de-sign intended	B (1) cental software	B (2) software for var. est.	B (3) software for surveys	B (4) suppl. sw-routines	B (5) new software
CH Switzerland	categorical and metric	households and individuals	closed formula, calibrations only approximately	in some cases	yes	SAS, SPSS, S-PLus	usually SAS / SAS macros	SAS macros	new versions
	categorical and metric	HH and individuals	closed formula, with replacement						
	categorical and metric	HH and individuals							
	categorical and metric	households and individuals	summary of methods used: Davison and Canty, 1998; 5 steps Post-stratification						
	categorical and metric	HH and individuals	Jackknife						
	categorical and metric	HH and individuals	Post-stratification						
	categorical and metric	HH and individuals	Post-stratification or Jackknife						
UK United Kindom	categorical and metric	households	linearised variances; calibration weighting or Post-stratification	ongoing reviews take place for continuous survey	no	STATA 7	yes	STATA macros	SUDAAN, WESVAR, SAS used on occasion
	categorical and metric	households							
	categorical and metric	individuals							
	categorical and metric	individuals							
	categorical and metric	individuals							

B.3 Short tables

country	AT	BE	BG	CZ	DK	EE	FI	FR	DE	EL	IE	IL	IT	LT	NL	NO	PT	ES	SI	SE	CH	UK
HBS [1]	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
survey mandatory	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0
period: monthly	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
period: quarterly	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	1	1
period: yearly	0	0	0	0	0	1	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0
period: other	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	1	1	0	0	1	0	0
design: panel [2]	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
design: cr-sect. w. rot.	0	0	1	0	1	1	0	1	1	0	1	1	1	1	1	0	0	1	0	1	1	0
design: cr-sect. wo. rot.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1
nonresp. $\leq$ 15% [3]	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15% < nr $\leq$ 50%	0	0	0	0	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1
nonresp. > 50%	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
unit: households	0	0	1	1	1	1	0	1	0	0	1	0	1	1	1	0	0	1	0	0	0	1
unit: individuals	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
unit: both [4]	0	1	0	0	0	0	1	0	1	0	0	1	0	0	0	1	1	0	1	1	1	0
2/ 3-stage sample	0	1	1	0	0	1	1	1	0	0	1	1	1	1	1	1	1	1	1	0	1	1
variable: categoric	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
variable: metric	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0
variable: both	0	0	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	0	1	1	1	1
data: households	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0
data: individuals	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
data: both [5]	0	1	1	0	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	1
variance estimation	0	0	0	0	0	1	1	0	0	0	1	1	1	0	0	0	1	1	1	1	1	1
change of design [6]	0	0	0	1	0	0	0	1	1	0	0	0	0	1	1	0	0	1	0	0	1	1

[1] HBS substituted: FR: Household Survey; IL: Family Expenditure Survey and Income Survey; IT: Consumer Expenditure Survey; NL: Consumer Survey; NO: Household Consumption Survey; SE: Income and Income Distribution Survey; UK: Family Expenditure Survey

[2] design: panel survey, cross- sectional survey with rotation or cross- sectional survey without rotation

[3] three intervals for Non-Response quotes: [0%; 15%], (15%; 50%] and (50%; 100%]

[4] unit: households and individuals

[5] data: households and individuals

[6] Question: Is change of sample design intended?

Table 1: Evaluation HBS with respect to countries

Table 2: Evaluation LFS with respect to countries

country	AT	BE	BG	CZ	DK	EE	FI	FR	DE	EL	IE	IL	IT	LT	NL	NO	PT	ES	SI	SE	CH	UK
LFS	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
survey mandatory	1	1	0	0	0	0	0	1	1	1	0	1	1	0	0	0	1	1	0	0	0	0
period: monthly	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
period: quarterly	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	1	1	1	0	1
period: yearly	1	0	0	0	0	0	0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0
period: other	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
design: panel [1]	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
design: cr-sect. w. rot.	1	1	0	0	0	1	0	1	1	1	0	1	1	1	1	0	1	1	0	1	1	0
design: cr-sect. wo. rot.	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
design: other	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
nonresp. $\leq$ 15% [2]	1	0	0	0	0	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	0	0
15% < nr $\leq$ 50%	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
nonresp. > 50%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
unit: households	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	1	1	0	0	0	0
unit: individuals	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	1	0	0	0	0	0	1
unit: both [3]	1	1	0	1	0	1	0	0	1	0	0	1	0	0	0	0	0	0	1	1	1	0
2/ 3-stage sample	1	1	0	1	0	1	0	1	1	1	0	1	1	0	1	0	1	1	1	0	1	0
variable: categoric	1	1	0	0	0	0	0	1	1	1	0	0	0	0	1	1	0	1	0	0	0	0
variable: metric	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
variable: both	0	0	0	1	0	1	1	0	0	0	0	1	1	1	0	0	1	0	1	1	1	1
data: households	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
data: individuals	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
data: both [4]	1	1	0	1	0	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	0
variance estimation	1	1	0	0	0	1	1	0	1	1	0	1	1	1	0	0	1	1	1	1	1	1
change of design [5]	1	0	0	0	0	0	0	0	1	0	0	0	1	1	0	0	0	1	0	0	1	1

[1] design: panel survey, cross- sectional survey with rotation, cross- sectional survey without rotation or other
[2] three intervals for Non-Response quotes: [0%; 15%], (15%; 50%] and (50%; 100%]
[3] unit: households and individuals
[4] data: households and individuals
[5] Question: Is change of sample design intended?

Table 3: Summary of evaluation HBS, LFS

evaluation HBS	number of reponses [1]	response results = yes absolute	relative	response results = no absolute	relative
HBS	22	20	0,91	2	0,09
survey mandatory	20	4	0,20	16	0,80
period: monthly	20	2	0,10	18	0,90
period: quarterly		7	0,35	13	0,65
period: yearly		6	0,30	14	0,70
period: other		5	0,25	15	0,75
design: panel [2]	17	1	0,06	16	0,94
cross-sect. w. rot.		13	0,76	4	0,24
cross-sect. wo. rot.		3	0,18	14	0,82
nonresp. ≤ 15% [3]	18	1	0,06	17	0,94
15% < nr ≤ 50%		15	0,83	3	0,17
nonresp. > 50%		2	0,11	16	0,99
unit: households	20	11	0,55	9	0,45
unit: individuals		0	0,00	20	1,00
unit: both [4]		9	0,45	11	0,55
2/ 3-stage sample	20	16	0,80	4	0,20
variable: categoric	20	2	0,10	18	0,90
variable: metric		4	0,20	16	0,80
variable: both		14	0,70	6	0,30
data: households	20	3	0,15	17	0,85
data: individuals		0	0,00	20	1,00
data: both [5]		17	0,85	3	0,15
variance est. [6]	20	10	0,50	10	0,50
change of design [7]	20	8	0,40	12	0,60

evaluation LFS	number of reponses [1]	response results = yes absolute	relative	response results = no absolute	relative
LFS	22	19	0,86	3	0,14
survey mandatory	19	9	0,47	10	0,53
period: monthly	18	3	0,17	15	0,83
period: quarterly		8	0,44	10	0,56
period: yearly		5	0,28	13	0,72
period: other		2	0,11	16	0,89
design: panel [2]	19	3	0,16	16	0,84
cross-sect. w. rot.		14	0,74	5	0,26
cross-sect. wo. rot.		1	0,05	18	0,95
design: other		1	0,05	18	0,95
nonresp. ≤ 15% [3]	18	14	0,78	4	0,22
15% < nr ≤ 50%		4	0,22	14	0,78
nonresp. > 50%		0	0,00	18	1,00
unit: households	19	5	0,26	14	0,74
unit: individuals		5	0,26	14	0,74
unit: both [4]		9	0,47	10	0,53
2/ 3-stage sample	19	14	0,74	5	0,26
variable: categoric	19	8	0,42	11	0,58
variable: metric		0	0,00	19	1,00
variable: both		11	0,58	8	0,42
data: households	19	0	0,00	19	1,00
data: individuals		2	0,11	17	0,89
data: both [5]		17	0,89	2	0,11
variance est. [6]	19	15	0,79	4	0,21
change of design [7]	19	7	0,37	12	0,63

[1] number of "1" - responses counted for each question
[2] design: panel survey, cross- sectional survey with rotation
cross- sectional survey without rotation (LFS: or other)
[3] 3 intervals for Non-Response quotes: [0%; 15%], (15%; 50%], (50%; 100%]
[4] unit: households and individuals
[5] data: households and individuals
[6] variance estimation for survey
[7] Question: Is change of sample design intended?

Table 4: Non-Response- quotes for HBS and LFS by countries

country	survey		Non-Response- quotes in percent	m/ v[1]
BE	HBS	90%		vol.
BG	HBS	5.6%		vol.
CH	HBS	30% - 50%		vol.
CZ	HBS			vol.
DE	HBS			vol.
DK	HBS	35%		vol.
EE	HBS	40%		vol.
ES	HBS	16%		mand.
FI	HBS	35%		vol.
FR	HBS	22%		vol.
IE	HBS	55%		vol.
IL	HBS	15% - 22%		mand.
IT	HBS	18%		mand.
LT	HBS	20% - 30%		vol.
NL	HBS	35%		vol.
NO	HBS	30% - 40%		vol.
PT	HBS	20%		mand.
SE	HBS	30% - 35%		vol.
SI	HBS	17%		vol.
UK	HBS	40%		vol.

Summary[2]	HBS mandatory		4
Summary[2]	HBS voluntary		16
Summary[2]	LFS mandatory		9
Summary[2]	LFS voluntary		10

AT	LFS	10% - 15%		mand.
BE	LFS	16%		mand.
CH	LFS			vol.
CZ	LFS	25%		vol.
DE	LFS	3%		mand.
EE	LFS	10% - 13%		vol.
EL	LFS	5% - 10%		mand.
ES	LFS	8%		mand.
FI	LFS	14%		vol.
FR	LFS	12%		mand.
IL	LFS	9%		mand.
IT	LFS	10%		mand.
LT	LFS	10% - 15%		vol.
NL	LFS	45%		vol.
NO	LFS	10% - 15%		vol.
PT	LFS	10%		mand.
SE	LFS	15%		vol.
SI	LFS	12%		vol.
UK	LFS	25%		vol.
			0 10 20 30 40 50 60 70 80 90 100	

Color Definition: HBS mandatory, HBS voluntary, LFS mandatory, LFS voluntary

[1] v/ m: mandatory/ voluntary [2] summary: number of surveys in last column

References

Binder, D. A. (1996): Linearization methods for single phase and two-phase samples: A cookbook approach. *Survey Methodology* **22**, 17–22.

Campbell, C. (1980): A different view of finite population estimation. In *Proceedings of the Section on Survey Research Methods*, pp. 319–324. Alexandria, Virginia: American Statistical Association.

Davison, A. C. and Hinkley, D. V. (1997): *Bootstrap Methods and Their Application.* Cambridge: Cambridge University Press.

Davison, A. C., Hinkley, D. V. and Schechtman, E. (1986): Efficient bootstrap simulation. *Biometrika* **73**, 555–566.

Deville, J. C. (1999): Variance estimation for complex statistics and estimators: Linearization and residual techniques. *Survey Methodology* **25**, 193–203.

Durbin, J. (1953): Some results in sampling theory when units are selected with unequal probabilities. *Journal of the Royal Statistical Society series B* **15**, 262–269.

Durbin, J. (1967): Design of multistage surveys for the estimation of sampling errors. *Applied Statistics* **16**, 152–164.

Fernholtz, L. T. (1983): *Von Mises Calculus for Statistical Functionals.* Volume 19 of *Lecture Notes in Statistics.* New York: Springer.

Fuller, W. A. (2003): Estimation for multiple phase samples. In *Analysis of Survey Data*, eds R. L. Chambers and C. J. Skinner, pp. 307–322. Chichester: Wiley.

Gabler, S. and Stenger, H. (2000): Minimax strategies in survey sampling. *Journal of Statistical Planning and Inference* **90**, 305–321.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986): *Robust Statistics: The Approach Based on Influence Functions.* New York: Wiley.

Lohr, S. (1999): *Sampling: Design and analysis.* Pacific Grove, CA: Duxbury Press.

McCarthy, P. J. (1969): Pseudo-replication: Half samples. *Review of the Interational Statistics Institute* **37**, 239–264.

Raj, D. (1966): Some remarks on a simple procedure of sampling without replacement. *Journal of the American Statistical Association* **61**, 391–396.

Rao, J. N. K. (1988): Variance estimation in sample surveys. In *Handbook of Statistics, Volume 6*, ed. P. R. K. amd C. R. Rao, pp. 427–447. Amsterdam: Elsevier Science.

Särndal, C. E., Swensson, B. and Wretman, J. (1992): *Model Assisted Survey Sampling.* New York: Springer-Verlag.

Shao, J., Chen, Y. and Chen, Y. (1998): Balanced repeated replication for stratified multistage survey data under imputation. *Journal of the American Statistical Association* **93**, 819–831.

Shao, J. and Tu, D. (1995): *The Jackknife and Bootstrap.* New York: Springer-Verlag.

United Nations (1995): Guidelines for the modeling of statistical data and metadata. Technical report, United Nations.

Van der Vaart, A. W. (1998): *Asymptotic Statistics.* Cambridge University Press.

Wolter, K. M. (1985): *Introduction to Variance Estimation.* New York: Springer-Verlag.

Woodruff, R. S. (1971): A simple method for approximating the variance of a complicated estimate. *Journal of the American Statistical Association* **66**, 411–414.