EUROPEAN CENTRAL BANK

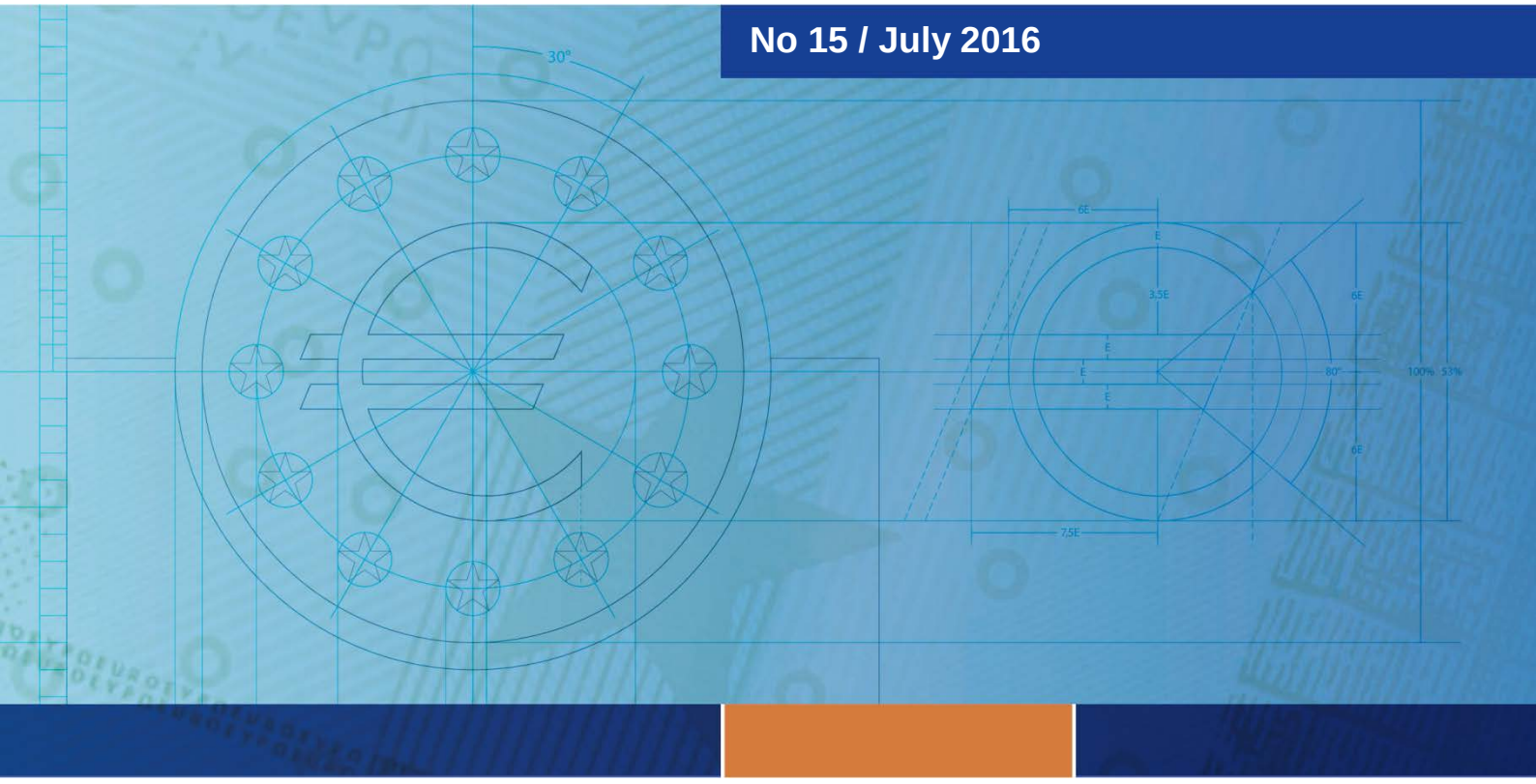
EUROSYSTEM

Statistics Paper Series

Guillaume Osier

Unit non-response in household wealth surveys

Experience from the Eurosystem's Household Finance and Consumption Survey



No 15 / July 2016

Note: This Statistics Paper should not be reported as representing the views of the European Central Bank (ECB). The views expressed are those of the author and do not necessarily reflect those of the ECB.

Contents

Abstract

The Household Finance and Consumption Survey (HFCS) is a recent initiative from the Eurosystem to collect comparable micro-data on household wealth and indebtedness in the euro area countries. The Household Finance and Consumption Network (HFCN), which comprises the European Central Bank (ECB), national central banks (NCBs), and national statistical institutes (NSIs), is in charge of the development and implementation of the HFCS. The first round of the survey was successfully conducted between 2008 and 2011, and the results were published in April 2013. The second round is now under way and will cover all the euro area countries.

This paper is a joint effort by several members of the HFCN to further investigate the issue of unit non-response in the HFCS, better describe and understand its patterns, measure its effects on the overall quality of the survey and, ultimately, propose strategies to mitigate them. The paper is divided into sections, the first section being the introduction. The second section draws up a list of the main possible sources of auxiliary information that can be relied on in order to analyse non-response patterns in the HFCS. It also presents summary indicators that can be used to quantify unit non-response. In the third section, based on the experience from the first wave of the HFCS, the report elaborates on good survey practices (e.g. interviewer training and compensation, use of incentives, persuasive contact strategies, etc.) to prevent unit non-response from occurring. The fourth section compares several reweighting strategies for coping with unit non-response a posteriori, in particular simple and generalised calibration methods. These methods are assessed with respect to their impact on the main HFCS-based estimates. Finally, based on the outcome of this empirical analysis, recommendations are made with regard to post-survey weighting adjustment in the HFCS.

Keywords: sampling weights, unit non-response, response propensity, calibration

JEL codes: C83

Non-technical summary

The Household Finance and Consumption Survey (HFCS) is a recent initiative from the Eurosystem to collect comparable micro-data on household wealth and indebtedness in the euro area countries. The Household Finance and Consumption Network (HFCN), which is composed of the European Central Bank (ECB), national central banks (NCBs), national statistical institutes (NSIs) and academics, is responsible for the development of the HFCS and its implementation in the participating countries. The first round of the survey was successfully conducted between 2008 and 2011, and the results were released in April 2013 (European Central Bank, 2013b). The second round of data collection is in progress and will cover all the euro area countries.

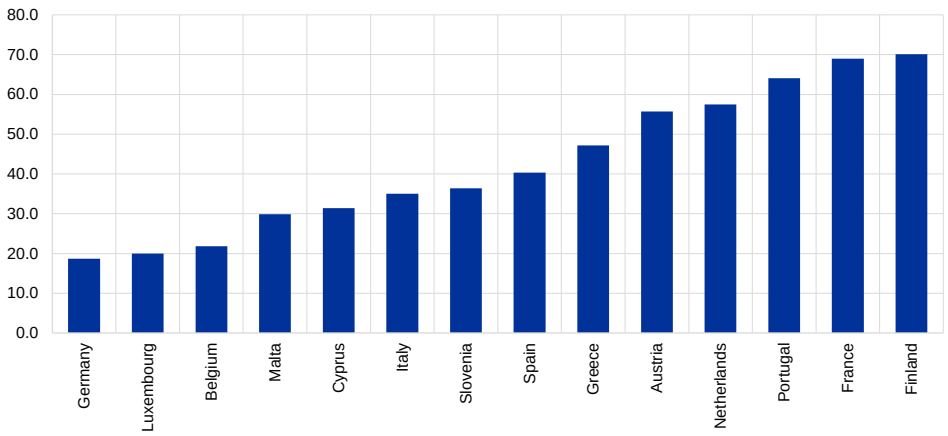
The complexity and the sensitivity of the HFCS core information on income and wealth has led to a high proportion of the selected households not participating in the survey at all. This is called unit non-response. There are also cases where a household agrees to participate in the survey, but part of the response is missing. In this case, the non-response is said to be partial. This paper only addresses the question of unit non-response in the HFCS.

Household response rates to the first wave of the HFCS range from less than 20% in Germany to around 70% in Finland. In general, all other things being equal, the HFCS achieves lower response than other European-wide household surveys (Pérez-Duarte et al., 2010), such as the European Union Statistics on Income and Living Conditions (EU-SILC) and the Household Budget Surveys (HBS).

Chart 1

Household response rates, first wave of the HFCS

(percentages)



Source: ECB, HFCS methodological report for the first wave.

Unit non-response is a key concern for survey data quality, especially when the non-respondents have a different profile from the respondents with respect to the main

survey characteristics of interest. In this case, survey estimates based on the respondents are biased, unless the estimation formulae are adjusted to reflect the underlying response mechanism. This adjustment requires auxiliary information on both the respondents and the non-respondents. In addition, unit non-response makes estimates less stable, as they rely on fewer units (the responding units). In order to compensate for the loss of sample units due to unit non-response, a larger initial sample should be drawn (on condition that additional resources can be devoted to this).

In the HFCS, wealthy households might be more reluctant to participate in the survey than other households. If so, estimates would be downwardly biased. Possible explanations could be that wealthy households may not want to reveal sensitive information on income and wealth to interviewers, or they may simply be more difficult to contact. Above all, it is essential to get the households at the top of the income and wealth distribution to participate in the survey in order to keep the sample representative of the distribution of household income and wealth in the population as a whole.

The main determinants of unit non-response in the HFCS need to be identified by gathering as much auxiliary information as possible on both the responding and the non-responding households. In order to reduce non-response bias significantly, the auxiliary variables need to be related both to the probability of responding to the HFCS and to the survey target variables on income and wealth. Possible sources of auxiliary information on the non-respondents include:

- dwelling-related information collected visually by the interviewers;
- the sampling frame itself;
- interviewer characteristics (age, education, experience, etc.);
- administrative sources, including tax sources;
- ad hoc surveys of the non-respondents;
- banking and financial sources.

It is important to consider all the available sources equally, gathering the most accurate information possible for non-response correction. The opportunity to utilise auxiliary information from other sources needs to be planned carefully in advance. When available, special attention needs to be paid to administrative sources, especially tax data and banking data, as they convey accurate information which is likely to be strongly related to the main topics of the HFCS. This information may not only help deal with unit non-response, but also serve more broadly as a benchmark against which to compare survey data.

Evidence based on the analysis of the available auxiliary sources shows that unit non-response in the HFCS is far from random, and is likely to depend not only on idiosyncratic characteristics of households, but also on external characteristics

related to the interviewers and the design of the survey itself. In addition, the response patterns in the HFCS seem to be different from one country to another.

Household response in the HFCS seems to be correlated with household income and wealth, at least in some countries. This indicates a potential risk of a causal relationship between household response and income and wealth, which may be detrimental to data quality and representativity, unless proper action is taken to both prevent and deal with unit non-response through sample reweighting.

Table 1

Weighted household response rates broken down by dwelling-related characteristics, by country

(percentages)

	Spain*	Italy	Portugal	Slovakia	Germany	Estonia	Belgium
Type of dwelling							
Individual house	67.4	61.7	62.6	50.6	21.5	76.1	47.9
Semi-detached house	66.8	61.0	50.1	66.8	19.7	70.5	41.8
Flat/apartment	50.5	50.1	83.1	54.7	14.8	64.6	32.6
Dwelling rating by interviewer							
Luxury	54.3	48.4	n.a.	29.8	23.6	58.4	43.4
Upscale	54.0	56.5	n.a.	51.0	21.4	67.4	44.4
Mid-range	.	54.3	n.a.	52.7	17.0	66.0	42.0
Modest	59.4	58.8	n.a.	56.8	18.8	77.4	39.6
Low-income	67.3	58.6	n.a.	56.2	15.9	77.0	42.8
Dwelling location							
Downtown	n.a.	56.3	n.a.	57.3	17.8	69.1	37.7
Between city centre and suburbs	n.a.	55.3	n.a.	50.5	16.8	63.5	38.6
Town outskirts	n.a.	53.4	n.a.	46.2	19.1	61.4	47.1
Isolated area/countryside	n.a.	58.6	n.a.	56.4	23.6	77.2	43.9
Dwelling – outward appearance							
Generally clean and sound	54.3	n.a.	66.2	52.6	21.4	64.6	44.5
Some peeling paint or cracks in walls	54.0	n.a.	65.1	51.1	14.6	69.3	36.8
Needs substantial painting, refilling or repair	59.4	n.a.	62.5	58.9	18.7	76.6	38.3
Comparison with the neighbourhood. Are other dwellings in the neighbourhood better or worse than this one?							
Better	60.3	n.a.	67.8	52.9	18.1	71.7	36.0
As good as this dwelling	54.1	n.a.	64.9	52.6	18.3	65.3	42.1
Worse	63.9	n.a.	61.6	53.9	26.0	73.7	49.3
No other buildings in view	59.1	n.a.	81.7	43.0	22.2	87.0	19.1
Rating of surrounding buildings by interviewer							
Luxury	57.0	54.9	n.a.	35.2	25.3	54.9	39.0
Upscale	53.4	.	n.a.	48.9	21.6	65.0	42.8
Mid-range	51.7	55.7	n.a.	52.8	16.1	67.7	42.2
Modest	66.7	.	n.a.	60.7	15.4	. (**)	41.7
Low-income	71.8	.	n.a.	53.1	10.7	75.6	41.2
Security measures							
Doorman	52.9	n.a.	72.7	74.5	13.0	100.0	38.6
Guard	58.0	n.a.	.	.	15.1	60.5	19.8
Locked lobby	.	n.a.	.	54.3	6.4	72.5	38.2
Intercom device	51.4	n.a.	63.2	46.3	17.0	60.5	31.5

* No weights available for Spain.
** Due to CAPI errors, the categories “Modest” and “Low-income” have been merged into a single “Low-income” category for Estonia.
Notes:
n.a. = not available (information not collected).
. = no observations for the category.

Popular indicators of unit non-response in sample surveys are the following outcome rates (American Association for Public Opinion Research, 2008):

- Eligibility rate, which depends on the quality of the sampling frame. It refers to the proportion of units in the gross sample that are eligible for the survey.
- Contact rate: ratio of the total number of contacted units to the total number of eligible units.
- Cooperation rate: ratio of the total number of respondents to the total number of contacted units.
- Refusal rate: ratio of the total number of units who refused to participate in the survey (or who break off the interview) to the total number of contacted units.
- Response rate: ratio of the total number of complete interviews to the number of eligible units.

These outcome rates were used in the HFCS methodological report for the first wave (European Central Bank, 2013). Although extremely popular, response rates are, however, poor predictors of unit non-response bias, as a low response rate does not automatically mean that the response bias is important and vice versa (Groves and Peytcheva, 2008). To that end, alternative indicators are required.

Key indicators of survey representativity, in particular the R-indicator, have been defined to serve as counterparts to survey response rates and are primarily directed at evaluating the non-response bias. The indicator is given by (Schouten et al., 2009):

$$R(\rho) = 1 - 2S(\rho)$$

$S(\rho)$ is the standard deviation of the response propensity ρ within the population. It measures the spread of the response propensities among the households: the higher the spread, the less "representative" is the survey response. By definition, the R-indicator lies between 0 and 1: $R(\rho) = 1$ means response is fully "representative" (i.e. uniform response propensity across the households), $R(\rho) = 0$ otherwise. Furthermore, the R-indicator also provides an upper bound of the non-response bias called the maximum potential bias:

$$|B_m| \leq \frac{S(\rho)}{\bar{\rho}} = \frac{1 - R(\rho)}{2\bar{\rho}}$$

In practice, however, there are certain objections and problems using the R-indicator and the resulting estimate for the maximum response bias. The main issue is the sensitivity of the R-indicator to model specification. The construction of such indicators relies on a stage of response modelling which relies itself on auxiliary information, the availability of which may be scarce in certain countries. Besides, as response patterns are generally country-specific and the availability of auxiliary information very heterogeneous across the countries, the R-indicator must not be used for cross-country comparisons. Nonetheless, it still remains a useful tool for

monitoring the representativeness of the sample during the fieldwork and identifying groups with relatively lower participation and for which tailored efforts are needed.

Fieldwork procedures have been widely recognised both in the statistical literature and survey practice as contributing to the prevention of unit non-response and the raising of survey participation. Raising response rates may result in lower unit non-response bias, provided the effort is made in such a way that the participation of all sub-groups of the population is balanced. If response rates asymmetrically increase only for particular sub-groups, there may be an adverse effect on response bias (this would happen if the response enhancement efforts simply brought in more respondents like those already in the sample). The reason is that unit non-response bias depends not only on the response rate, but also on the correlation between the response and the target variables of the survey. Thus, higher survey participation does not always mean lower non-response bias (Groves and Peytcheva, 2008), unless differences between respondents and non-respondents with respect to the key survey characteristics are reduced at the same time. In fact, if not carried out properly, retargeting mechanisms might reinforce biases.

Response rates depend on a variety of underlying factors, some of which have to do with the content of the survey and the way it is conducted while others do not. For example, response rates are generally lower in urban areas and higher in rural areas. Potential survey-specific factors include:

at the design stage of the survey,

- a questionnaire which is neither too long nor too burdensome,
- the availability of several language versions of the questionnaire,
- a sampling design which takes into account the interviewer workload (i.e. multi-stage selection to control the geographical dispersion of the sample),
- the recruitment of experienced interviewers,
- the proper training and compensation of the interviewers,
- the supporting legal basis (mandatory versus voluntary participation),
- the credibility of the different sponsors and stakeholders,
- assurances regarding data privacy and confidentiality;

during the fieldwork,

- the calling schedule of the interviews, including evenings and weekends,
- the use and type of incentives to get the households to participate,
- the use of an efficient mode of data collection (computer-assisted personal interviewing (CAPI), computer-assisted telephone interviewing (CATI), mixed, etc.),

- the use of advanced information (official letter, etc.) to introduce the survey,
- the monitoring of the fieldwork, managing interviewers, offering encouragement, assistance, identifying and responding to problems;

after the fieldwork,

- fieldwork auditing, refusal conversion strategies.

In addition to efficient fieldwork procedures, sample reweighting methods are commonly used ex post in order to compensate for unit non-response bias. Basically, there are two kinds of approach: one based on the explicit estimation of the response propensities and another based on calibration techniques. Calibration has grown today into a widely used technique in official statistics. The principle is to adjust the sampling weights using population benchmarks so the new weights reproduce exactly the population totals (in the case of quantitative variables) or the population distributions (in the case of categorical variables) for a predefined set of auxiliary variables. For instance, we may seek new sampling weights so the weighted sample distribution by age and gender characteristics is the same as the distribution in the population, as provided by the last population census.

The calibration approach offers several advantages: it is simple and flexible to implement with existing software tools. At the same time, it is able to reduce non-response bias, reduce sampling variance and deal with frame imperfections. This approach is less stringent than the traditional one based on the direct estimation of response propensities, as adjustment variables are no longer required to be available for the non-respondents. However, population benchmarks need to be known, which can still be a hurdle. Calibration is also regarded by some users as an opaque ("black box") technique. In any case, this option cannot be implemented when the probability of response depends on the topic of the survey itself (non-ignorable non-response), as might be the case with wealth surveys such as the HFCS.

The impact of both reweighting approaches (response propensity weighting and calibration) on the HFCS-based estimates seems pretty much the same. Thus, the choice between them is a matter of personal preference. Although calibration is easy, flexible and universal, some might consider it more reasonable to follow a more classical two-step approach (estimation of the response propensities and calibration to reduce sampling variance), where each step is kept under control. In any case, the choice of a reweighting scheme should be made in the light of its impact on survey estimates, especially on estimates for small sub-populations, which are usually less stable.

The main lessons we can draw from this paper are the following:

- Unit non-response is a key concern for the HFCS, as non-response rates are important in some countries and response patterns are not completely random. The factors underlying unit non-response to the HFCS are diverse and also country and survey-specific. The way the survey is designed, the legal basis

(mandatory/voluntary), the quality of interviewer training and the use of incentives may also affect response patterns.

- In order to deal with unit non-response bias, it is of utmost importance that all the auxiliary information available on the non-respondents be gathered in order to properly describe and analyse response patterns. This paper provides an extensive list of possible auxiliary sources to be considered for non-response analysis. In this respect, administrative, banking and financial data should be especially relevant as they convey information which is highly correlated to the main topics of the HFCS.

- Relevant contact and interview strategies are required to increase survey participation and reduce non-response bias. In addition, reweighting strategies need to be implemented ex post. There is no single reweighting approach which could be recommended for all cases, each having pros and cons. Basically, the choice between the traditional response propensity weighting approach and calibration techniques mainly depends on their impact on the key HFCS aggregates on income and wealth.

1 Introduction

1.1 The Household Finance and Consumption Survey (HFCS)

The Household Finance and Consumption Survey (HFCS) is a recent initiative the Eurosystem to collect comparable micro-data on household wealth and indebtedness in the euro area countries (European Central Bank, 2013). The Household Finance and Consumption Network (HFCN), which is composed of the European Central Bank (ECB), national central banks (NCBs), and national statistical institutes (NSIs), is in charge of the development of the HFCS and its implementation in all the participating countries. The first round of the survey was successfully conducted between 2008 and 2011, and the results were published in April 2013 (European Central Bank, 2013b). The HFCS data have been collected in a harmonised way in 15 euro area countries for a sample of more than 62,000 households, with varying sample sizes across the countries. The second round of HFCS data collection is in progress and will cover all the euro area countries as well as a few non-euro area countries.

The HFCS micro-data provide valuable insights into the distributional aspects of household income, wealth and indebtedness. Analysis by domain of interest is thus possible, which macroeconomic aggregates do not allow. For instance, one may assess the vulnerability of specific sub-groups of households with regards to indebtedness, which can help central banks better identify potential risks for the stability of the financial sector. Particularly since the beginning of the economic and financial downturn in 2008-2009, central bankers are increasingly using micro-data from sample surveys[1] in order to fine-tune their actions in the field of monetary policy.

In each country, household samples have been designed to make survey results representative. Nearly all HFCS national surveys follow a probabilistic sample design, which means that each household in the population has a well-defined, a priori, non-zero probability of being part of the sample. Nearly all the HFCS national sampling designs are complex in the sense that they may involve components such as stratification or geographical clustering. Stratification is a well-known technique to make sampling more precise by dividing the population into categories (strata) that are supposed to be homogeneous with regard to the target characteristics of the survey. Then, representative samples are drawn from each stratum, the selection being independent across strata. As regards geographical clustering, it consists of selecting a sample in several stages. First, a sample of “primary sampling units” (PSUs) is taken. These often represent aggregated geographical divisions such as regions or municipalities. Then, a sample of “secondary sampling units” (SSUs) is taken within each selected PSU, the selection being independent from one PSU to

[1] See, for instance, *IFC Bulletin*, No 39, Bank for International Settlements, April 2015.

another. The process can be iterated until the final sampling units, i.e. households, are reached.

In order to better capture the extreme top of the wealth distribution, many countries participating in the HFCS oversample the wealthy households (Kennickell, 2007). Oversampling wealthy households increases sampling precision, especially in the estimation of income or wealth-related indicators. However, efficient oversampling would require identifying wealthy households in the sample in advance, which is not always feasible. Spain, for instance, had access to taxable wealth data from tax sources. On the other hand, Italy, Malta, the Netherlands and Slovakia did not use oversampling for the first wave of the HFCS.

The HFCS micro-data are fully checked, edited and imputed by the individual countries. The multiple imputation technique with five implicates (Rubin, 1987) was recommended by the HFCN to impute missing values in order to better account for the variability of the imputation process. SAS routines called €MIR have been developed for this purpose and are used by many participating countries.

Table 2

Oversampling strategies, first wave of the HFCS

Country	Criteria for oversampling	Details
Belgium	Average local income	Optimal (Neyman) allocation, based on the standard deviation of income in stratum and stratum size.
Germany	Taxable income of regions	Smaller municipalities (population<100,000) and, in larger municipalities, street sections with high average income (>€92,000) are oversampled.
Greece	Regional; real estate prices	The sampling rates in Athens and Thessaloniki are proportional to the real estate prices of each cluster.
Spain	Taxable wealth of individuals	Eight wealth strata were defined, which were oversampled at progressively higher rates.
France	Wealth	For the wealthy sample, four strata were defined: wealthy city dwellers, equity-based wealth, real estate-based wealth, lower wealth. For each primary unit and each stratum, an allocation proportional to main residences was computed. Then a systematic selection was made within each primary unit.
Italy	No oversampling	-
Cyprus	Electricity consumption	61% of the gross sample was drawn from households within the top 10% by electricity consumption.
Luxembourg	Personal income subject to social contributions	20% of the gross sample was drawn from the top income decile according to the social security register and the self-employed headed fiscal household sub-population.
Malta	No oversampling	-
Netherlands	No oversampling	-
Austria	Regional	Some oversampling in Vienna because of higher expected non-response.
Portugal	Region	Metropolitan areas of Lisbon and Porto oversampled, 50% of gross sample drawn from these areas.
Slovenia	Region	The municipalities of Ljubljana and Maribor were oversampled, as higher non-response was to be expected.
Slovakia	No oversampling	To help interviewers fulfil the prescribed income quotas, a list of streets with higher incidence of wealthy households in each stratum was provided.
Finland	Individual income and socio-economic status from population register	High-income employees, self-employed and farmers were oversampled, as well as “others” and “no tax”.

Source: ECB, HFCS methodological report for the first wave.

1.2 The issue of unit non-response in the HFCS

Unit non-response is a long-standing issue in sample surveys. In particular, since the launch of the HFCS, the complexity and the sensitivity of the core survey information on income and wealth have led to a high proportion of the selected households not participating in the survey at all. This is called unit non-response. Unit non-response generally stems from non-contact (the household cannot be reached), refusal (the household is reached but refuses to participate in the survey) or incapacity (e.g. health problems or language barriers). There are also cases where a household agrees to participate in the survey, yet part of the response is missing (e.g. the interviewee may refuse to answer certain questions, which are considered too intrusive, or may not know the answer). In this case, non-response is said to be partial (also called item non-response). In this paper, only the question of unit non-response in the HFCS is addressed.

Household response rates for the first wave of the HFCS appear to be quite heterogeneous across countries, ranging from 18.7% in Germany, 20% in Luxembourg and 21.8% in Belgium to 69% in France and around 70% in Finland. Even though all sample surveys are liable to unit non-response, response rates tend to be higher in wealth surveys than in other household surveys, all other things being equal. For example, a comparative study of the HFCS, the European Union Statistics on Income and Living Conditions (EU-SILC) and the Household Budget Survey (HBS) showed that the HFCS generally achieves lower response than the two other surveys (Pérez-Duarte et al., 2010).

Unit non-response in the HFCS is a key concern for survey data quality, especially when the non-respondents have a different profile from the respondents with respect to the survey characteristics of interest. In this case, there is a relationship between survey response and the target characteristics of the survey. As a result, survey estimates based on the respondents are biased, unless the estimation formulae are adjusted in order to reflect the underlying response mechanism. The adjustment requires auxiliary information which is available both for the respondents and the non-respondents in order to describe the response mechanism. In order to reduce non-response bias effectively, the auxiliary information used needs to be related both to the survey response and the target characteristics of the survey. The availability of such auxiliary information is actually a major practical consideration, which often limits the scope for non-response correction.

In addition, efficient contact strategies and interviewing protocols need to be put in place in order to convince as many sample households as possible to participate in the survey, while keeping the overall sample representative of the distribution of household income and wealth. Unit non-response also makes estimates less stable (i.e. higher variance), as they rely on fewer units (the responding units). In order to compensate for the loss of sample units due to unit non-response, a larger initial sample should be drawn (provided that additional resources can be devoted to the survey).

Unit non-response in sample surveys is driven by a lot of different factors, some of them being country-specific (e.g. the social and cultural environment, the degree of

political participation, the level of trust in institutions), while others are survey-specific (e.g. the overall survey design, whether or not the survey is mandatory for the respondents, the degree of sensitivity and complexity of the main survey topic) or even unit-specific. For example, in the HFCS, wealthy households might be more reluctant to participate in the survey than other households. If so, estimates would be downwardly biased. Possible explanations could be that wealthy households may not want to reveal sensitive information on income and wealth to interviewers, or they may simply be more difficult to contact. Above all, it is essential to get the households at the top of the income and wealth distribution to participate in the survey in order to keep the sample representative of the distribution of household income and wealth in the population as a whole. This is precisely why many countries chose to oversample the wealthy households for the HFCS.

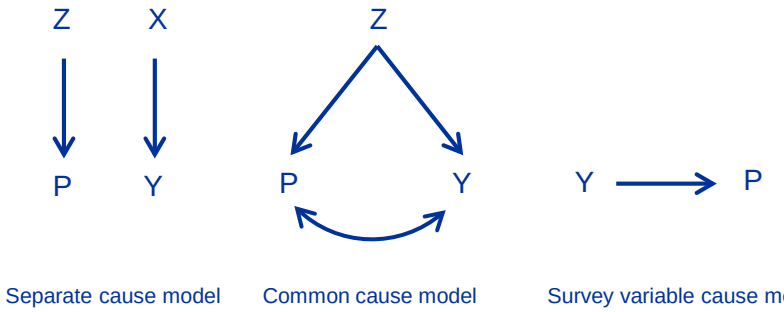
In this paper, the issue of unit non-response bias in the HFCS is further investigated, with the aim of better describing and understanding its patterns, measuring its effects on the overall quality of the survey and, ultimately, proposing strategies to reduce its impact. The paper is divided into several sections, the first section being the introduction. The second section draws up a list of the main possible sources of auxiliary information on the non-respondents to the HFCS. Furthermore, unit non-response indicators are presented. In the third section, on the basis of country experiences from the first wave of the survey, the report elaborates on good practices (e.g. interviewer training and compensation, use of incentives, persuasive contact strategies, etc.) to increase response rate. Finally, in the fourth section, post-survey weighting adjustments are presented. These methods are aimed at dealing with unit non-response bias a posteriori.

As already mentioned, unit non-response bias is caused by the relationship between survey response and the target variables of the survey. First of all, proper non-response analysis requires making assumptions about the underlying response mechanisms. There are basically three main mechanisms (Groves and Peytcheva, 2008):

- Missing completely at random (MCAR), where the probability of response is not related to the target variables of the survey. Non-response is MCAR when the probability of response is uniform for all units.
- Missing at random (MAR), where the probability of response is related to the target variables via other characteristics.
- Not missing at random (NMAR), where there is a causal relationship between the probability of response and the target variables of the survey.

This terminology follows that introduced by Rubin (1987) and Little and Rubin (2002). MCAR and MAR mechanisms are said to be ignorable, while NMAR is non-ignorable. These mechanisms are illustrated in the figure below (Groves and Peytcheva, 2008).

Figure
Examples of response mechanisms



Source: Groves and Peytcheva (2008).

- On the left, the probability of response (P) depends on characteristics (Z) and the survey variables of interest (Y) depend on characteristics (X) which are not related to (Z). Thus, (P) and (Y) are not related: this is a missing completely at random (MCAR) mechanism.
- In the middle, the probability of response (P) is related to the survey variables of interest (Y) via a set of confounding factors (Z): this is a non-causal relationship and the mechanism is called missing at random (MAR).
- On the right, a not missing at random (NMAR) mechanism is shown, where there is a causal relationship between the probability of response (P) and the target variables of the survey (Y).

When dealing with unit non-response, we generally assume that the underlying mechanism is ignorable (MAR). Thus, non-response bias could be taken care of using auxiliary variables that are available both on the respondents and the non-respondents. Otherwise, non-ignorable non-response is more challenging to deal with, as the probability of response depends on variables (the target variables of the survey) which are by definition available only for the respondents.

1.3 Short historical review

Unit non-response in sample surveys has been a long-established problem for survey statisticians, who have been developing adapted survey designs to deal with unit non-response and limit its impact on estimates (e.g. Hansen and Hurwitz, 1946). The approach proposed by Hansen and Hurwitz consists of calling non-respondents (or a sub-sample thereof) back using a more efficient mode of data collection in order to get their survey data. Such an approach is costly, however, and there is no guarantee that all the non-respondents that had been re-approached will agree to cooperate the second time. A less expensive option is to ask the non-respondents only a few basic questions (Kersten and Bethlehem, 1984). This basic question approach (BQA) allows a few basic items to be collected help understand non-response patterns.

In addition, Groves and Couper (1998) developed a conceptual framework for survey participation. Their framework includes factors which cannot be controlled, such as the social environment, and factors which can be controlled, such as the design of the survey or the profile of the interviewers. Today's research is largely focused on responsive designs (Groves and Herringa, 2006; Laflamme and Karaganis, 2010), where the fieldwork is continuously monitored and adjusted in order to ensure good representativity of the sample, especially with respect to the population categories that are usually hard to reach. Fieldwork monitoring relies on indicators available prior to sampling and also on information gathered during the data collection process (paradata) (Blom, 2009).

In parallel, more efficient sample weighting methods have been developed to deal with unit non-response a posteriori. They range from the traditional response propensity weighting approach to calibration-based techniques (Deville and Särndal, 1992). Response propensity weighting was defined long ago (Oh and Scheuren, 1983) and consists of explicitly estimating the response propensities from auxiliary information observed both for the respondents and the non-respondents and adjusting the sample weights by the inverse of the estimated response propensities. Therefore, in order to be effective, such methods would require a great deal of auxiliary information on both the respondents and the non-respondents. Calibration techniques have been developed more recently (Deville and Särndal, 1992). They are often regarded as more flexible, as no information is needed on the non-respondents. However, population benchmarks are required in order to calibrate the weights. In any case, sample reweighting is intrinsically related to the availability of strong auxiliary sources.

With the declining trend in household survey participation (Meyer et al., 2015), the issue of unit non-response has grown today into a very important challenge to survey data quality and representativity. Efficient sample designs and fieldwork procedures and powerful reweighting strategies are of vital importance in order to limit its impact on survey results as far as possible.

2 Describing unit non-response in the HFCS

2.1 Sources of auxiliary information

First of all, the main determinants of unit non-response in the HFCS need to be identified by gathering as much auxiliary information as possible both for the responding and the non-responding households. In this section, we present possible sources of auxiliary information on the non-respondents to the HFCS.

As already mentioned, in order to reduce non-response bias significantly, the auxiliary variables need to be related both to the probability of response to the HFCS and to the survey target variables on income and wealth (e.g. Särndal and Lundström, 2005). The relevance of any source of auxiliary information in treating unit non-response needs to be considered in the light of these two requirements. As regards the former, i.e. the relationship with the probability of response to the HFCS, summary response indicators (basically response rates broken down by auxiliary variables) could be produced to examine the relationship between the response indicators and the auxiliary information. Fitting regression-based models (e.g. logistic models) to the data is another option. As regards the latter requirement, as, by definition, no survey information is available on the non-respondents, we cannot determine whether there is a statistical relationship between the auxiliary variables and the target variables on income and wealth. At this stage, only evidence based on other sources is available.

2.1.1 Auxiliary information already available in the survey datasets

The HFCS datasets which are transmitted by individual countries to the ECB already contain dwelling-related information observed for both the responding and the non-responding households. This information might be useful for describing and analysing unit non-response. In particular, we could assume that this information is related to the survey target variables on income and wealth. For instance, the higher the dwelling rating is, the higher household income and wealth are likely to be.

- Type of dwelling

1 – Individual house, 2 – Semi-detached house, 3 – Flat/apartment, 4 – Other kind of dwelling

- Dwelling rating

1 – Luxury, 2 – Upscale, 3 – Mid-range, 4 – Modest, 5 – Low-income

- Dwelling location

1 – Downtown, 2 – Area between city centre and suburbs, 3 – Town outskirts, 4 – Isolated area, countryside

- Dwelling – outward appearance

1 – Generally clean and sound, 2 – Some peeling paint or cracks in walls, 3 – Needs substantial painting, refilling or repair, 4 – Dilapidated

- Dwelling – comparison with the neighbourhood (Are other dwellings in the neighbourhood better or worse than this one?)

1 – Better, 2 – As good as this dwelling, 3 – Worse, 4 – There are no other buildings in view

- Dwelling – rating of surrounding buildings

1 – Luxury, 2 – Upscale, 3 – Mid-range, 4 – Modest, 5 – Low-income, 6 – Very low income

- Dwelling – security measures

1 – Doorman, 2 – Guard, 3 – Locked lobby, 4 – Intercom device, 5 – Other

These variables are meant to be collected visually by the interviewers both on the responding and the non-responding households. However, only very few countries actually collected and provided all this information for the first wave of the HFCS (see next table). The situation will improve for the second wave of the survey, as more countries have committed themselves to collecting and transmitting the information to the ECB. Nonetheless, there are still several countries which will either provide no variables at all or provide them only for the interviewed households. Although relevant, these variables also have quality problems: the subjectivity of interviewer ratings when interviewers cannot enter the dwelling, the difficulty of distinguishing between the proposed response categories (e.g. between “luxury” and “upscale” dwellings), the difficulty of distinguishing between the centre, suburbs and outskirts of small towns, and the difficulty of collecting data on all the security measures inside and outside a dwelling. Furthermore, the collection of these variables raises methodological issues, particularly with regard to comparability from one country to another and, within a country, from one interviewer to another. Even though strict instructions were given to interviewers as to how to collect these variables, the assessment still relies strongly on the interviewer’s own perception. Notwithstanding all this, as we will see later on, these variables are related to some degree to household response.

In order to improve the collection and the quality of these variables, the dwelling location data could be collected only for households in densely populated areas, following, for instance, the DEGURBA methodology,[2] which was recently introduced by the European Commission to classify municipalities on the basis of their degree of urbanisation. As regards the dwelling rating variables, in order to make these

[2] http://ec.europa.eu/eurostat/web/degree-of-urbanisation/overview

variables more comparable across the interviewers, sample pictures could be shown to the interviewers during training sessions. Internet tools such as Google Street View might also be helpful if the assessment turns out to be difficult for certain dwellings.

Table 3

Availability of dwelling-related information in the HFCS datasets

	AT	BE	CY	DE	EE	ES	FI	FR	GR	HU	IE	IT	LU	LV	MT	NL	PT	SI	SK
First wave of the HFCS																			
Type of dwelling	Yes	No	Part	Part		Yes	Part	Part	Part			Yes	Part		Part	No	Yes	Yes	Part
Dwelling rating	Yes	No	Part	No		Yes	No	No	No			Yes	Part		Part	No	No	Yes	Part
Dwelling location	Yes	No	No	No		No	No	No	No			Yes	Yes		Part	No	No	Yes	Part
Dwelling – outward appearance	Yes	No	Part	No		Yes	No	No	No			No	Part		Part	No	Yes	Yes	Part
Comparison to the neighbourhood (are other dwellings in the neighbourhood better or worse and this one?)	Yes	No	No	No		Yes	No	No	No			No	Part		Part	No	Yes	Yes	Part
Rating of surrounding buildings	Yes	No	No	No		Yes	No	No	No			Yes	Part		Part	No	No	Yes	Part
Security measures	Yes	No	No	No		Yes	No	No	Part			No	Part		Part	No	Yes	Yes	Part
Second wave of the HFCS																			
Type of dwelling	Yes	Yes	Part	Part	Yes	Yes	Part	Yes	Part	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes	Yes
Dwelling rating	Yes	Yes	Part	No	Yes	Yes	No	Yes	No	Yes	No	Yes	Yes	No	Yes	No	No	Yes	Yes
Dwelling location	Yes	Yes	No	No	Yes	No	No	No	No	Yes	Yes	Yes	Yes	No	Yes	No	No	Yes	Yes
Dwelling – outward appearance	Yes	Yes	Part	No	Yes	Yes	No	No	No	Yes	No	No	Yes	No	Yes	No	Yes	Yes	Yes
Comparison to the neighbourhood (are other dwellings in the neighbourhood better or worse and this one?)	Yes	Yes	Part	No	Yes	Yes	No	No	No	Yes	No	No	Yes	No	Yes	No	Yes	Yes	Yes
Rating of surrounding buildings	Yes	Yes	Part	No	Yes	Yes	No	No	No	Yes	No	Yes	Yes	No	Yes	No	No	Yes	Yes
Security measures	Yes	Yes	Part	No	Yes	Yes	No	No	Part	Yes	No	No	Yes	No	Yes	No	Yes	Yes	Yes

Sources: ECB (for the first wave) and ad hoc consultation of the countries (for the second wave).
Notes:
- "Yes" indicates that information on the respondents and the non-respondents is available in the HFCS datafiles.
- "Part" indicates that only information on the respondents is available in the HFCS datafiles.
- "No" indicates that no information at all is available in the HFCS datafiles (not collected, or collected but not provided).
- In Luxembourg (LU), the variables might be not delivered if any concerns about household privacy arose in the anonymisation process.
- In Hungary (HU), the variables are collected only for the part of the sample which chooses to be interviewed face-to-face.
- No data for Estonia (EE), Hungary (HU), Ireland (IE) and Latvia (LV) for the first wave of the survey.

2.1.2 The sampling frame

Another source of auxiliary information about the non-respondents is the sampling frame itself. However, the information therein is often limited to a few basic descriptors such as the NUTS2 region or the municipality. Nonetheless, there are cases where the sampling frame can be linked (based on a unique identification code) to a wider range of auxiliary information, such as tax sources.

Table 4

Sampling frames used by the individual countries, first wave of the HFCS

Country	Sampling frame(s)
Belgium	Telephone register and street register
Germany	Clusters of addresses from municipalities (NSIs); list of street sections, population registers of municipalities
Greece	List of municipalities (census); random routing for secondary sampling units
Spain	Municipal census (list of addresses) supplemented by tax office information; list of addresses
France	List of geographical units (based on census); list of dwellings
Italy	List of municipalities; resident lists from municipalities
Cyprus	Customer register of the Electricity Authority of Cyprus
Luxembourg	Addresses of fiscal households from social security register
Malta	Dwelling register of the NSI
Netherlands	Postal addresses
Austria	List of enumeration districts; register of postal addresses
Portugal	List of geographical areas; list of private dwellings, from census
Slovenia	List of districts from census; list of persons aged 16+ from population register
Slovakia	List of municipalities, households chosen by random walk
Finland	Central population register using master sample of 50,000 persons aged 16+ and members living in the same household-dwelling unit

Source: ECB, HFCS methodological report for the first wave.

2.1.3 Interviewer characteristics

This information is easily accessible and may prove to be strongly related to household response. For example, the Banca d'Italia is used to collect a great deal of auxiliary information on interviewers' socio-demographic characteristics as part of its national Survey on Household Income and Wealth (SHIW), which is the reference source for the HFCS data in Italy. The following tables are based on the SHIW data for 2012. They show that, for wealth surveys, interviewers with a higher level of education and experience tend to achieve better response rates, all else being equal.

Table 5

Weighted household response rates by socio-demographic characteristics of the interviewers, SHIW, 2012

(percentages)

		Response rate
Age	<30	33.3
	30-49	54.4
	50-64	55.0
	>64	48.2
Gender	Male	58.5
	Female	52.3
Level of education	Lower secondary	51.2
	Upper secondary	53.2
	Tertiary	59.0
Currently working on other surveys?	Yes	54.4
	No	52.3
Number of years of experience as interviewer	0-2	49.9
	3-4	42.6
	5-6	48.9
	7-8	54.7
	9-10	50.6
	10+	55.3
Number of years of experience working on the HFCS	0-2	46.4
	3-4	55.0
	5-6	54.4
	7-8	67.6
	9-10	53.0
	10+	58.9
Motivation for working as interviewer	This is my main source of income	56.1
	To have an additional source of income	51.2
	I love this job	51.4
	Other	60.3
Job status	Employee	50.3
	Self-employed	50.5
	Unemployed	54.5
	Other non-working condition	54.9

Source: Banca d'Italia, Survey on Household Income and Wealth (SHIW).

In the context of the SHIW, the Banca d'Italia also collects detailed information about interviewers' attitudes toward wealth. This information does not seem to affect interviewer performance in a significant way. On the other hand, response rates tend to be higher if the interviewers consistently resort to certain "good practices" when they try to contact households and introduce the survey (all other things being equal). For instance, the average household response rate is 55% when the interviewer always mentions the social utility of the survey to the household, but falls to 47% when the interviewer often mentions this. Similarly, the interviewers who always ask neighbours for information when nobody is home tend to achieve higher response rates. Thus, unit non-response appears to be correlated with interviewer-

related characteristics. However, the correlation of those characteristics with the survey target variables on income and wealth is more questionable. In that case, interviewer-related characteristics may not be a key to reducing unit non-response bias.

Table 6

Weighted household response rates by interviewers' attitudes toward wealth, SHIW, 2012

(percentages)

	Totally disagree	2	3	4	5	6	7	8	9	Totally agree
The concentration of too much wealth in a few hands creates problems in society	.	62.1	74.5	25.0	39.0	63.5	49.8	56.3	50.4	55.3
The rich enjoy unjustified advantages in life	51.8	56.0	62.4	55.2	48.5	54.7	56.1	52.1	51.2	54.9
No one should have a better chance in life just because of having inherited considerable wealth	50.6	49.0	59.5	50.9	50.8	59.9	49.8	53.7	52.9	54.7
Everyone has a chance to become rich by his/her own effort	52.3	49.8	47.3	50.6	50.0	52.6	57.1	52.5	56.2	55.4
Rich people can contribute (through the charitable sector) to make our society fairer	47.5	59.5	72.5	61.5	47.4	53.5	53.7	55.1	45.8	59.7
You become rich by your own efforts	55.5	56.5	56.9	56.8	49.6	55.7	51.1	56.4	42.8	41.3
You become rich by inheritance	31.1	65.9	45.2	68.5	57.4	55.4	53.1	52.7	55.5	50.0
You become rich if you have good ideas	53.0	39.1	56.9	57.1	56.1	55.6	51.5	54.1	49.6	55.1

Source: Banca d'Italia, Survey on Household Income and Wealth (SHIW).
Note: . = no observations for the category.

Table 7

Weighted household response rates by interviewers' introduction strategies, SHIW, 2012

(percentages)

	Always	Often	Rarely	Never
"I introduce myself and I immediately show my identification card"	52.7	55.3	.	.
"I immediately say that I am not selling anything"	51.0	55.8	.	.
"I show immediately a copy of the advance letter"	54.2	51.8	.	.
"I use a standard introduction for all respondents"	54.9	53.1	.	.
"I tell them something about the study and about the Banca d'Italia"	54.2	51.7	.	.
"I mention the survey and its social utility"	55.2	47.1	.	.
"I immediately show all the material about the survey"	55.7	53.4	37.6	.
"Before every new study I rehearse the introduction, so I can say my things smoothly without hesitation"	54.6	50.5	55.0	.
"I vary my introduction depending on the situation"	53.7	53.2	56.2	35.7
"I adjust my language and the words I use to suit the people I interview"	52.2	54.9	54.4	.
"If asked about the length, I try to be vague"	47.7	54.1	54.9	51.4
"If nobody is at home, I leave a message (card, letter)"	53.0	56.2	51.5	45.0
"If nobody is at home, I ask the neighbours for information"	58.4	51.3	48.2	.

Source: Banca d'Italia, Survey on Household Income and Wealth (SHIW).
Note: . = no observations for the category.

Table 8

Weighted household response rates by interviewers' contact strategies, SHIW, 2012

(percentages)

	Strongly agree	Agree	Neither agree, nor disagree	Disagree	Strongly disagree
During the initial contact, it is more important to gain interest than to seek a quick decision to participate	55.3	52.6	30.9	45.2	.
Reluctant respondents should always be persuaded to participate	53.4	53.7	53.3	53.4	51.8
With enough effort, even the most reluctant respondent can be persuaded to participate	53.7	55.6	47.9	50.6	55.2
An interviewer should respect the privacy of the respondent	49.1	54.9	51.1	54.8	.
If a respondent is reluctant, a refusal should be accepted	57.8	52.8	46.6	52.9	66.0
One should always emphasise the voluntary nature of participation	54.3	53.0	53.4	54.9	45.1
Most respondents can be approached in the same way, in the same manner	52.6	55.2	46.0	64.5	.
Every respondent needs a unique approach	54.6	53.9	50.7	47.7	51.8
Give everyone the feeling that there are a lot of other households that have already agreed to participate	52.7	57.4	42.3	46.6	.
An interviewer should project a positive image of himself/herself	50.2	51.2	53.5	60.0	51.1
If a respondent appears likely to refuse, it is better to withdraw and try again at a later moment	53.6	53.5	.	.	.
Make clear that YOU believe in the study	52.9	62.1	34.4	.	.
Interviewers must convey to the respondents that they can be trusted	52.7	54.8	75.1	63.5	.

Source: Banca d'Italia, Survey on Household Income and Wealth (SHIW).
Note: . = no observations for the category.

The relevance of using interviewer-related information for non-response correction was also emphasised by other HFCS national experiences, for instance, in Austria and Belgium.

2.1.4 Administrative sources

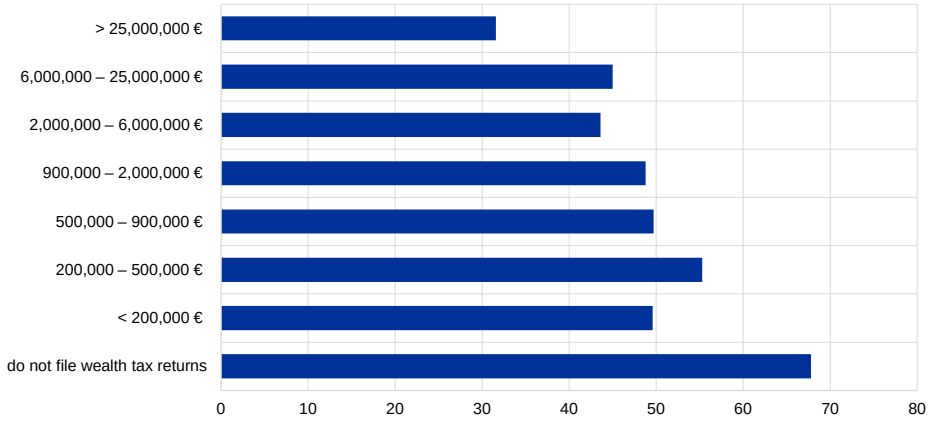
Administrative sources may prove to be powerful in reducing non-response bias. For example, France and Spain had access to tax sources on income and wealth. Tax data were used at the sample selection stage of the HFCS in order to oversample the wealthy households, thereby making their samples more representative of the overall distributions of income and wealth in the population.

In addition, this auxiliary information, which might be linked to the survey data through a unique identification code, is expected to be strongly correlated both with the characteristics of interest to the survey and the propensity for households to respond. For example, the chart and table below, which are based on the Survey of Household Finances (EFF) which is conducted by the Banco de España, show that wealthy households tend to participate less in the HFCS than other households.

Chart 2

Cooperation rate in the Spanish EFF by wealth category, 2008[3]

(percentages)



Source: Banco de España, Survey of Household Finances (EFF).

Conversely, using auxiliary data from social security files, the Banque centrale du Luxembourg (BCL) has observed that higher income classes tend to have higher response rates than the others (Mathä et al., 2012).

Table 9

Cooperation rate in the LU-HFCS, first wave

(percentages)

Response rate by category		
By nationality	National households	19.7
	Non-national households	20.3
By income class	< €7,000	18.4
	> €7,000	24.2
Total		20.0

Source: Banque Centrale du Luxembourg, Household Finance and Consumption Survey, first wave.

Other national experiences with administrative sources include Statistics Finland using income registers and Eesti Pank using income-related information from the Estonian Tax and Customs Board.[4]

However, the use of administrative sources in the HFCS, like in any other surveys, poses some extra difficulties, such as conceptual gaps between administrative and

3 See the following reference: http://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosOcasionales/11/Fich/do1103e.pdf

4 This income-related information refers to:
- domestic gross employee income,
- gross income from self-employment,
- domestic gross income from public pensions,
- foreign gross income from public pensions,
- gross income from private and occupational pension plans,
- gross income from unemployment benefits.

survey data, access restrictions due to confidentiality reasons (especially with tax sources), problems of timeliness, or technical difficulties in relation to database management.

In any case, there is a growing interest among members of the European Statistical System (ESS) in using administrative sources instead of collecting the information through interviewing (Jäntti, Törmälehto and Marlier, 2013). Using administrative sources would offer a couple of significant advantages: first, administrative data are generally high-quality data, and, second, this would reduce the burden of the survey questionnaire, as the number of questions would be reduced. Surveys could then focus on key information which cannot be retrieved from administrative databases.

2.1.5 Ad hoc survey of the non-respondents

The collection of relevant auxiliary information on non-respondents through an ad hoc survey is a long-established practice. Several strategies for re-approaching non-respondents have been defined in the literature (e.g. Stoop, 2005). One possible approach is to call back the non-respondents (or a sub-sample of them) using more experienced interviewers and more targeted modes of data collection, trying to convince them to reverse their decision and to participate in the survey. The objective is to increase the overall response rate. However, refusal conversion strategies remain costly (in terms of time and resources) and risky, as the desired outcome (refusal conversion) is by no means guaranteed.[5]

One alternative and cheaper approach, called the basic question approach (BQA), consists of asking a very limited number of questions to those households who refuse to cooperate (Kersten and Bethlehem, 1984). Actually, experience has shown that households who refuse to participate in a survey may agree to reply to a very light version of the questionnaire (one page maximum). Obviously, the challenge is to ensure that this light questionnaire captures as much of the core HFCS information as possible (e.g. housing tenure status, value of household main residence, household income – in bands – and employment status). Furthermore, the basic questionnaire is itself affected by unit non-response (caused by “hard” refusals), making the collected data useless unless unit non-response is “ignored”, i.e. treated as uniform. In addition, data is usually harder to obtain from the households which have not been contacted for the main questionnaire. In fact, this reduced questionnaire is often restricted to the refusals.

2.1.6 Banking and financial sources

For Estonia’s national HFCS, Eesti Pank has had access to a great deal of auxiliary information from commercial banks, the Estonian Central Register of Securities, financial intermediaries and insurance funds. This information refers to the amount of

[5] However, interesting refusal conversion experiences have been reported in the context of the European Social Survey.

wealth assets (e.g. sight accounts, time deposits or shares) owned by households and is likely to be well-correlated with the HFCS target variables on household wealth. Assuming this information is also correlated with the probability to respond, it should be quite powerful in dealing with unit non-response bias. On the other hand, as with administrative sources, access to banking and financial data is generally constrained by legal restrictions and practical considerations.

Table 10

Examples of banking and financial sources (used by Eesti Pank)

Auxiliary information	Source
Outstanding balance on all the loans (sum of real estate collateralised and consumer loans)	Commercial banks
Monthly payments on the loans (sum of real estate collateralised and consumer loans)	
Amount of the lease payments per month	
Outstanding balance on leasing	
Amount in euro in sight accounts	
Amount in euro in time deposits	
Current market value of shares in publicly traded companies	Estonian Central Register of Securities; financial intermediaries; insurance funds
The value of social security plans with an account balance	
The value of voluntary pension schemes or whole life insurance contracts	

Source: Eesti Pank.

2.2 Unit non-response indicators

2.2.1 Outcome rates

Popular indicators of unit non-response in sample surveys are the following outcome rates (American Association for Public Opinion Research, 2008; National Research Council, 2013):

- Eligibility rate, which measures the quality of the sampling frame. It refers to the proportion of units in the gross sample that are eligible for the survey.
- Contact rate: ratio of the total number of contacted units to the total number of eligible units.
- Cooperation rate: ratio of the total number of respondents to the total number of contacted units.
- Refusal rate: ratio of the total number of units who refused to participate in the survey (or break off the interview) to the total number of contacted units.
- Response rate: ratio of the total number of complete interviews to the total number of eligible units.

These outcome rates were used in the HFCS methodological report for the first wave (European Central Bank, 2013).

Table 11

Response behaviour indicators, first wave of the HFCS

Country	Gross sample size (number of households)	Net sample size (number of households)	Response rate (%)	Weighted response rate (%)	Refusal rate (%)	Cooperation rate (%)	Contact rate (%)	Eligibility rate (%)
Belgium	11,376	2,364	21.8	n.a.	57.6	27.2	80.1	95.4
Germany	20,501	3,565	18.7	n.a.	69.7	21.1	94.2	92.9
Greece	6,354	2,971	47.2	48.7	46.4	47.8	98.7	99.1
Spain	11,782	6,197	56.7 *	n.a.	34.8	58.4	97.2	92.6
France	24,289	15,006	69.0	69.6	30.0	69.0	100.0	89.8
Italy	15,592	7,951	52.1 *	53.2	37.8	57.8	90.2	97.8
Cyprus	3,938	1,237	31.4	32.4	56.6	35.7	88.0	100.0
Luxembourg	5,000	950	20.0	19.3	63.7	21.0	95.5	94.9
Malta	3,000	843	29.9	30.4	34.1	44.3	67.5	94.0
Netherlands	2,263	1,301	57.5 *	n.a.	42.5	57.5	100.0	100.0
Austria	4,436	2,380	55.7	56.4	39.6	56.7	98.1	96.3
Portugal	8,000	4,404	64.1	59.0	10.3	80.2	79.9	85.9
Slovenia	965	343	36.4	35.6	45.9	41.6	87.5	97.8
Slovakia	2,000	2,057	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Finland	13,525	10,989	82.2 *	85.0	11.1	86.2	95.4	98.7

Source: ECB, HFCS methodological report for the first wave.
Notes:
n.a. = not available.
* Response rates for the whole sample. More comparable response rates are the response rates for households interviewed for the first time, which are 40.3% in Spain, 35.0% in Italy and 70.1% in Finland. This figure is not available for the Netherlands.

Cases with unknown eligibility require special attention. It is recommended that a certain proportion of those cases be treated as eligible and taken into account in the calculation of response indicators. When estimating this proportion, one must be guided by the best available objective information. The share of eligible cases among the cases with unknown eligibility may be estimated by the share of eligible cases calculated among those with known eligibility:

$$eE = EL / (EL+NE) \quad (1)$$

where EL is the total number of known eligible cases and NE is the total number of known ineligible cases. Hence, the total number of eligible units, that is the denominator of both the contact and the response rates is EL+eE*UE, where UE is the total number of units with unknown eligibility. In any case, every effort should be made during the fieldwork to reduce as much as possible the number of cases with unknown eligibility.

In the case of sampling designs where the units are sampled with unequal probabilities (such as the HFCS), it is possible to calculate weighted response rates. The weighting coefficient to be used in the HFCS is the household case design weight, which basically refers to the inverse of the household selection probabilities. Weighted response rates are better indicators of survey output quality, while non-weighted rates are more appropriate for measuring fieldwork process quality. In the HFCS, weighted response rates would be more comparable across the countries, as

the weights would reflect the complex sampling features used at national level (e.g. oversampling the wealthy, multi-stage selection or weighting adjustments).

Household response rates broken down by auxiliary information available both for the respondents and the non-respondents provide valuable insight into the main potential determinants of unit non-response. Measures of association (such as Cramer's V coefficient) should be derived from those cross-tabulations in order to quantify the strength as well as the statistical significance of those associations.

2.2.2 Response representativity indicators

The response rate is often regarded as key descriptor of the quality of a survey, yet it is in fact a poor predictor of non-response bias (Groves and Peytcheva, 2008). In order to circumvent this issue, we produce response representativity indicators based on estimated response propensities. Key indicators of survey representativity were defined in the framework of the "RISQ" project[6] (Representativity Indicators for Survey Quality). Such quality indicators, called the R-indicators, may serve as counterparts to traditional survey response rates and, contrary to the traditional outcome rates, are primarily directed at evaluating the non-response bias (Schouten et al., 2009 and 2011). The R-indicator is given by:

$$R(\rho) = 1 - 2 \times S(\rho) \quad (2)$$

$S(\rho)$ is the standard deviation of the response propensities in the population, which measures the spread of the response propensities among the households: the higher the spread, the less "representative" is the survey response. By definition, the R-indicator lies between 0 and 1: $R(\rho) = 1$ means the response is fully "representative" (i.e. uniform response propensity across all households), $R(\rho) = 0$ means the complete opposite. Furthermore, the R-indicator also provides an upper bound of the non-response bias B_m called the maximum potential bias:

$$|B_m| \leq \frac{S(\rho)}{\bar{\rho}} = \frac{1-R(\rho)}{2\bar{\rho}} \quad (3)$$

In practice the value of the indicator has to be estimated:

$$\hat{R}(\hat{\rho}) = 1 - 2 \times \hat{S}(\hat{\rho}) \quad (4)$$

where $\hat{S}^2(\hat{\rho}) = \sum_i w_i \times (\hat{\rho}_i - \bar{\rho}_w)^2 / [\sum_i w_i - 1]$, the $\hat{\rho}_i$ are the estimated response propensities (based on a response model), and $\bar{\rho}_w = \sum_i w_i \times \hat{\rho}_i / \sum_i w_i$ is the mean estimated propensity based on the household weights w_i.

The decomposition of the variance $S^2(\rho)$ of the response propensities into "between" and "within" components of the response propensities for the sample sub-groups is the foundation of the partial R-indicators at variable level. The unconditional partial R-indicator corresponds to the between sub-group variance, while the within variances are the basis for the conditional partial indicators. Those indicators can be

[6] http://www.risq-project.eu/

further decomposed into the category level R-indicators showing the contributions to the variation of the respective categories (Heij et al., 2010; Bańkowska, Osiewicz and Pérez-Duarte, 2015).

Table 12

Unconditional and conditional partial R-indicators

	Unconditional	Conditional
$S^2(\rho) =$	$S^2_{between}(\rho)$	$S^2_{within}(\rho)$
Variable level	$P_U(X_k) = \sqrt{\frac{1}{N}\sum_{h=1}^{H} n_h(\bar{\rho}_h - \bar{\rho})^2}$	$P_C(X_k) = \sqrt{\frac{1}{N}\sum_{l=1}^{L}\sum_{i \in U_l} d_i(\rho_i - \bar{\rho})^2}$
Category level	$P_U(X_k, h) = \sqrt{\frac{n_h}{N}}(\bar{\rho}_h - \bar{\rho})$	$P_C(X_k, h) = \sqrt{\frac{1}{N}\sum_{l=1}^{L}\sum_{i \in U_l} d_i\Delta_{h,i}(\rho_i - \bar{\rho}_l)^2}$
Notes	X_k is a categorical variable with *H* categories and is a component of the vector *X* $n_h = \sum_{i=1}^{n} d_i\Delta_{h,i}$ is the weighted sample size in the category *h*, where $\Delta_{h,i}$ is a 0-1 dummy variable for sample unit i being a member of stratum *h* U_l is a cell in the cross-classification of all model variables except X_k	

Source: Bańkowska, Osiewicz and Pérez-Duarte (2015).

2.2.3 Numerical exercise

For the specific purpose of this exercise, computer codes were developed in order to:

- calculate household response rates broken down by auxiliary information,
- estimate the household response propensities based on the auxiliary information available,
- calculate response representativity indicators (R-indicator and maximum potential bias) from the estimated response propensities.

The auxiliary information used for this exercise includes both the dwelling-related information already provided in the HFCS datasets and, when possible, other nationally available information. Data from seven countries were used: Belgium, Estonia, Germany, Italy, Portugal, Slovakia and Spain.

Table 13

Auxiliary information used in each country

Country	Reference period	Description
Belgium (BE)	HFCS wave 2	HFCS dwelling-related information +location of residence + municipality-based information (dwelling space, taxable income, number of inhabitants, etc.) and interviewer characteristics
Estonia (EE)	HFCS wave 2	HFCS dwelling-related information + age, gender (household's reference person), income, pension assets, outstanding loans and deposits
Germany (DE)	HFCS wave 1	HFCS dwelling-related information + age (year of birth), gender, citizenship of household's reference person + dwelling location (four categories: urban – wealthy area, urban – other area, non-urban – wealthy area, and non-urban – other area)
Italy (IT)	2012 wave	HFCS dwelling-related information and interviewer characteristics
Portugal (PT)	HFCS wave 1	HFCS dwelling-related information + NUTS2 region
Slovakia (SK)	HFCS wave 2	HFCS dwelling-related information + region, municipality size and income class (three categories: high, medium, low)
Spain (ES)	HFCS wave 1	HFCS dwelling-related information

The next table presents the household response rates broken down by the dwelling-related variables that are available in the HFCS datasets. Actually, these variables appear to be correlated with response rates, although the trends are also different from one country to another. In Portugal, for example, households living in detached or semi-detached houses have a much lower response rate than those living in apartments, although the opposite is observed in the other countries. In Germany, households that are declared "luxury" by interviewers report higher response rates, while other countries exhibit the opposite trend.

Table 14

Weighted household response rates broken down by dwelling-related characteristics, by country

(percentages)

	ES*	IT	PT	SK	DE	EE	BE
Type of dwelling							
Individual house	67.4	61.7	62.6	50.6	21.5	76.1	47.9
Semi-detached house	66.8	61.0	50.1	66.8	19.7	70.5	41.8
Flat/apartment	50.5	50.1	83.1	54.7	14.8	64.6	32.6
Dwelling rating by interviewer							
Luxury	54.3	48.4	n.a.	29.8	23.6	58.4	43.4
Upscale	54.0	56.5	n.a.	51.0	21.4	67.4	44.4
Mid-range	.	54.3	n.a.	52.7	17.0	66.0	42.0
Modest	59.4	58.8	n.a.	56.8	18.8	77.4	39.6
Low-income	67.3	58.6	n.a.	56.2	15.9	77.0	42.8
Dwelling location							
Downtown	n.a.	56.3	n.a.	57.3	17.8	69.1	37.7
Between city centre and suburbs	n.a.	55.3	n.a.	50.5	16.8	63.5	38.6
Town outskirts	n.a.	53.4	n.a.	46.2	19.1	61.4	47.1
Isolated area/countryside	n.a.	58.6	n.a.	56.4	23.6	77.2	43.9
Dwelling – outward appearance							
Generally clean and sound	54.3	n.a.	66.2	52.6	21.4	64.6	44.5
Some peeling paint or cracks in walls	54.0	n.a.	65.1	51.1	14.6	69.3	36.8
Needs substantial painting, refilling or repair	59.4	n.a.	62.5	58.9	18.7	76.6	38.3
Comparison with the neighbourhood. Are other dwellings in the neighbourhood better or worse than this one?							
Better	60.3	n.a.	67.8	52.9	18.1	71.7	36.0
As good as this dwelling	54.1	n.a.	64.9	52.6	18.3	65.3	42.1
Worse	63.9	n.a.	61.6	53.9	26.0	73.7	49.3
No other buildings in view	59.1	n.a.	81.7	43.0	22.2	87.0	19.1
Rating of surrounding buildings by interviewer							
Luxury	57.0	54.9	n.a.	35.2	25.3	54.9	39.0
Upscale	53.4	.	n.a.	48.9	21.6	65.0	42.8
Mid-range	51.7	55.7	n.a.	52.8	16.1	67.7	42.2
Modest	66.7	.	n.a.	60.7	15.4	. (**)	41.7
Low-income	71.8	.	n.a.	53.1	10.7	75.6	41.2
Security measures							
Doorman	52.9	n.a.	72.7	74.5	13.0	100.0	38.6
Guard	58.0	n.a.	.	.	15.1	60.5	19.8
Locked lobby	.	n.a.	.	54.3	6.4	72.5	38.2
Intercom device	51.4	n.a.	63.2	46.3	17.0	60.5	31.5

* No weights available for Spain.
** Due to CAPI errors, the categories “Modest” and “Low-income” have been merged into a single “Low-income” category for Estonia.
Notes:
n.a. = not available (information not collected).
. = no observations for the category.

Further analysis based on nationally available auxiliary information shows that non-response tends to be higher in urban areas (probably due to higher non-contact), a little higher among the female population (in Estonia and Germany) and higher among foreigners (in Germany). Based on HFCS data for Italy and Belgium, household response is also found to be correlated with interviewer-related characteristics.

Table 15a

Weighted household response rates by nationally available auxiliary variables (Slovakia)

(percentages)

Region	Bratislava (capital city)	43.5
	Trnava	59.1
	Trencin	51.7
	Nitra	51.3
	Zilina	52.6
	B. Bystrica	55.5
	Presov	57.3
	Kosice	55.9
Municipality size	Less than 2,000 inhabitants	57.2
	2,000-4,999	57.9
	5,000-9,999	49.9
	10,000-19,999	53.8
	20,000-49,999	50.3
	More than 50,000 inhabitants	48.9
Income class	High	49.0
	Middle	53.5
	Low	50.5

Source: Countries' own calculations.

Table 15b

Weighted household response rates by nationally available auxiliary variables (Portugal)

(percentages)

NUTS2 Region	North	59.5
	Algarve	66.7
	Centre	62.7
	Lisbon	49.3
	Alentejo	75.5
	Madeira	81.1
	Açores	67.2

Source: Countries' own calculations.

Table 15c

Weighted household response rates by nationally available auxiliary variables (Estonia)

(percentages)

Age group	First quintile	61.5
	Second quintile	58.7
	Third quintile	66.0
	Fourth quintile	71.3
	Fifth quintile	76.0
Gender	Male	67.2
	Female	68.2
Deposits	Lowest 20% (bottom quintile)	66.6
	20-40%	65.3
	40-60%	67.5
	60-80%	69.7
	Highest 20% (top quintile)	69.8
Income	Lowest 20% (bottom quintile)	60.4
	20-40%	70.1
	40-60%	73.3
	60-80%	66.7
	Highest 20% (top quintile)	65.8

Source: Countries' own calculations.

Table 15d

Weighted household response rates by nationally available auxiliary variables (Germany)

(percentages)

Year of birth	1945 or earlier	20.1
	1946-1955	23.3
	1956-1965	18.6
	1966-1975	15.9
	1976-1985	14.5
	1986-1992	16.0
Gender	Male	19.6
	Female	16.9
Citizenship	National	19.1
	Foreigner	11.8
Dwelling location	Urban area – wealthy	17.5
	Urban area – other	14.2
	Non-urban area – wealthy	17.5
	Non-urban area - other	19.7

Source: Countries' own calculations.

Table 15e

Weighted household response rates by nationally available auxiliary variables (Belgium)

(percentages)

Average taxable income by household	First quintile	32.0
	Second quintile	36.1
	Third quintile	45.9
	Fourth quintile	45.8
	Fifth quintile	45.7
Average dwelling space according to the kind of dwelling visited (detached, semi-detached, apartment)	First quintile	27.8
	Second quintile	38.2
	Third quintile	43.9
	Fourth quintile	44.0
	Fifth quintile	47.5
Municipality size (number of inhabitants)	First quintile	42.1
	Second quintile	46.9
	Third quintile	45.5
	Fourth quintile	36.9
	Fifth quintile	35.2
Average value of real estate according to the kind of visited dwelling (detached, semi-detached, apartment)	First quintile	36.5
	Second quintile	40.7
	Third quintile	44.2
	Fourth quintile	47.1
	Fifth quintile	39.6
Median price of dwelling according to the kind of dwelling visited (detached, semi-detached, apartment)	First quintile	35.2
	Second quintile	40.1
	Third quintile	42.4
	Fourth quintile	46.4
	Fifth quintile	48.0
Percentage of mortgage contracts in default in the municipality	First quintile	50.0
	Second quintile	46.2
	Third quintile	40.7
	Fourth quintile	32.8
	Fifth quintile	33.7
Wealth index of the municipality	First quintile	30.9
	Second quintile	36.0
	Third quintile	46.3
	Fourth quintile	43.5
	Fifth quintile	47.4

Source: Countries' own calculations.

The next step consists of producing response representativity indicators based on the estimated response probabilities. In order to estimate the response probabilities, the following two-stage approach was implemented:

A logistic model is defined, where the 0/1 household response is the dependent response variable and the dwelling-related variables that are in the HFCS datasets (+ depending on the country, any other auxiliary variables available at national level)

are the explanatory variables. This regression model is weighted by the household design weights.

In order to smooth the distribution of the estimated response propensities, thereby preventing extreme propensity values, response homogeneous groups (RHG) are created following the methodology proposed by Beaumont and Haziza (2007). Basically, RHGs are created using the K-means classification algorithm, where the "optimal" number of classes is determined so the R-square coefficient from the ordinary linear regression of the estimated propensities (based on the weighted logistic regression at step 1) on the smoothed values (calculated as the average propensity value in each RHG) is higher than a pre-defined threshold (e.g. 0.95).

On the basis of these estimated probabilities, the R-indicator and the maximum potential non-response bias were calculated for the seven countries being studied.

Table 16

R-indicator and maximum potential non-response bias, by country

(percentages)

Belgium	R-indicator	64.6
	Maximum potential non-response bias	42.1
Estonia	R-indicator	77.1
	Maximum potential non-response bias	16.9
Germany	R-indicator	87.5
	Maximum potential non-response bias	32.3
Italy	R-indicator	64.7
	Maximum potential non-response bias	32.0
Portugal	R-indicator	67.0
	Maximum potential non-response bias	25.6
Slovakia	R-indicator	80.3
	Maximum potential non-response bias	18.3
Spain	R-indicator	80.8
	Maximum potential non-response bias	17.4

Source: Countries' own calculations.

In practice, however, there are certain objections to using the R-indicator and problems with the resulting estimate for the maximum non-response bias. The main issue is the sensitivity of the R-indicator to model specification.

For example, sensitivity analysis was conducted on the HFCS data for Estonia and Belgium. Using the Estonian HFCS data for wave 2, the R-indicator based only on the HFCS dwelling-related characteristics was compared to the indicator based both on the HFCS dwelling-related characteristics and the auxiliary information from banking and financial sources available at national level (see Table 10 in section 2.1.6). The addition of nationally available auxiliary information to the S-file variables had an impact on the R-indicator, which decreased from 81% to 77%. This shows how sensitive this indicator is to model specification.

Table 17

R-indicator and maximum potential non-response bias, Estonia, second wave of the HFCS

(percentages)

HFCS dwelling-related variables + banking and financial information available at national level	R-indicator	77.1	[76.6 , 77.6]
	Maximum potential non-response bias	16.9	[16.6 , 17.3]
HFCS dwelling-related variables only	R-indicator	81.1	[80.6 , 81.5]
	Maximum potential non-response bias	14.0	[13.7 , 14.3]

Source: Eesti Pank, Household Finance and Consumption Survey, second wave.
Note: The values between brackets are the lower and upper bounds of confidence intervals (95% level).

Similarly, Belgium has adopted a rotational structure for the second wave of the HFCS. In this case, the value of the R-indicator changes based on whether the calculation relies on the whole sample, the sub-sample of the new entries at the second wave, or the sub-sample of the panel units.

Table 18

R-indicator, Belgium, second wave of the HFCS

(percentages)

Whole sample	64.6	[64.0 , 65.1]
New entries	72.6	[72.0 , 73.2]
Panel units only	57.5	[56.3 , 58.7]

Source: Nationale Bank van België/Banque Nationale de Belgique, Household Finance and Consumption Survey, second wave.
Note: The values between brackets are the lower and upper bounds of confidence intervals (95% level).

Thus, R-indicators cannot be used for cross-country comparisons, as the response models rely on different sets of auxiliary variables in each country and the indicator is heavily dependent on the model specification. However, R-indicators can be implemented at country level as a tool for monitoring the representativeness during the data collection (Bańkowska, Osiewicz and Pérez-Duarte, 2015). They can be computed for different amounts of effort, e.g. number of attempts, level of interviewer's experience. Furthermore, R-indicators (including partial R-indicators) are key indicators in the design of responsive surveys, i.e. survey designs that are tailored to optimise the response rate and reduce response selectivity (Schouten et al., 2011).

2.2.4 Conclusions

There are many possible sources of auxiliary information for non-response correction in the HFCS: conventional information from the sampling frame, administrative sources, interviewer characteristics, etc. It is important to consider all the available sources equally, gathering the most accurate information possible for non-response correction. The opportunity to utilise auxiliary information from other sources needs to be planned carefully in advance. When available, special attention needs to be paid to administrative sources, especially tax data and banking data, as they convey accurate information which is likely to be strongly related to the main topics of the

HFCS. This information may not only help deal with unit non-response, but also serve more broadly as a benchmark against which to compare survey data.

Evidence based on the analysis of the available auxiliary sources shows that unit non-response in the HFCS is far from random, and is likely to depend not only on idiosyncratic household characteristics, but also on external characteristics related to the interviewers and the design of the survey itself. In addition, the response patterns in the HFCS seem to be different from one country to another.

Household response in the HFCS seems to be correlated with household income and wealth, at least in some countries. This indicates a potential risk of a causal relationship between household response and income and wealth, which may be detrimental to data quality and representativity, unless proper action is taken to both prevent and deal with unit non-response bias.

The main survey outcome rates (i.e. eligibility rate, contact rate, cooperation rate, refusal rate and overall response rate) should be calculated in order to provide valuable measures of the prevalence of unit non-response in a survey. In order to identify the main drivers of unit non-response, these rates should be broken down using auxiliary information available both for the respondents and the non-respondents, and measures of association (such as Cramer's V coefficient) should be derived from those cross-tabulations in order to quantify both the strength and the statistical significance of the different correlations. Multivariate techniques, such as logistic regression, could also be used to test the significance of the different response explanatory variables.

However, the response rate is not a good measure of non-response bias, as a low response rate does not automatically mean that the response bias is important and vice versa. To that end, response representativity indicators (R-indicators) based on the distribution of the response propensities may provide more relevant measures of the level of response bias. The construction of such indicators relies on a stage of response modelling which relies itself on auxiliary information, the availability of which may be scarce in certain countries. Besides, as response patterns are generally country-specific and the availability of auxiliary information very heterogeneous across the countries, the R-indicator must not be used for cross-country comparisons. Nonetheless, it still remains a useful tool for monitoring the representativeness of the sample during the fieldwork and identifying groups with relatively lower participation and for which tailored efforts are needed.

3 Other factors influencing the response rate

3.1 Introduction

In this section, we summarise field procedures that have been widely recognised both in the statistical literature and survey practice as contributing to the prevention of unit non-response and the raising of survey participation. Detailed tables based on country experiences from the first wave of the HFCS are provided in the annex. Raising response rates may result in lower unit non-response bias, provided the effort is made in such a way that the participation of all sub-groups of the population is balanced. If response rates asymmetrically increase only for particular sub-groups, there may be an adverse effect on response bias (this would happen if the response enhancement efforts simply brought in more respondents like those already in the sample). The reason is that unit non-response bias depends not only on the response rate, but also on the correlation between the response and the target variables of the survey. Thus, higher survey participation does not always mean lower non-response bias (Groves and Peytcheva, 2008), unless differences between respondents and non-respondents with respect to the key survey characteristics are reduced at the same time. In fact, if not carried out properly, retargeting mechanisms might reinforce biases.

Response rates depend on a variety of underlying factors, some of which have to do with the content of the survey and the way it is conducted while others do not. For example, response rates are generally lower in urban areas and higher in rural areas. Potential survey-specific factors include:

- the data collection organisation and the body sponsoring the survey,
- the use and type of incentives,
- the topic of the survey,
- the sample design,
- the mode of the survey (face-to-face, telephone, web, etc.),
- the interviewer's experience, training, compensation, workload,
- the call scheme of the interviewers (i.e. the times/days of the week when interviewers attempt to establish contact with respondents),
- the monitoring of the fieldwork, managing interviewers, offering encouragement, assistance, identifying and responding to problems,

- the length and the design of the questionnaire/interview, which must not put too high a burden on the interviewees, the availability of several language versions (especially in countries with several official languages),
- the fieldwork period (length, time of the year),
- the credibility of promises to protect the confidentiality of the data,
- the perceived usefulness of the data for the intended purposes,
- the design and persuasiveness of informational materials intended for respondents,
- the supporting legal basis (mandatory versus voluntary participation).

In order to be efficient, non-response prevention requires identification of the most important obstacles, which may well vary by country. Some of these obstacles cannot be controlled, e.g. the topic of the survey is inherent to the survey, although one can attempt to make the questionnaire interesting, well-designed, less intrusive, relevant to the respondent, and natural in its flow. The length of the survey is also an important aspect, if the interviewer is honest about this when introducing the survey. A long interview will put off respondents, although too short an interview may have adverse effects too, as respondents may feel that the survey is not rigorous enough. The legal basis cannot be controlled either, at least on the short run. Most of the remaining factors listed above refer to fieldwork procedures which, to a large extent, can be controlled with a positive influence on response rates.

The effectiveness of non-response prevention measures may vary across countries, as each has different norms and attitudes. Some countries also have long-established experience of wealth surveys (e.g. France, Italy and Spain), while others have very little experience. Experienced countries should undoubtedly achieve better response rates than the inexperienced ones. In any case, the following recommendations have proven to work, although some countries may need to make more effort than others to achieve the desired effects.

3.2 Interviewer experience and training

There is strong empirical evidence based on HFCS data that experienced interviewers tend to achieve higher response rates. There is also evidence that interviewers who are confident about their ability to elicit co-operation tend to achieve higher response rates. This confidence can increase with training.

In addition to training on how to conduct the actual interview, interviewers should be trained in ways of making contact and establishing rapport with the respondent, of persuading them of the importance of the survey, and of combating various kinds of refusal. It is therefore also essential that they are convinced about the topic of the survey and make themselves familiar with it so they would be able to answer any potential questions from the interviewees. Interviewers should also be briefed on the administrative procedures, filling in contact forms, reporting to the office, etc. Material

should be prepared and made available to them so that they can consult it whenever they need to.

The involvement of staff from the NCBs in the training of interviewers for the HFCS is essential, as they will be able to familiarise interviewers with the topic and, moreover, show them the importance of the survey to the NCB and hence also make them feel their own role is an important and responsible one. Moreover, because NCB staff bring a high level of subject-matter expertise, they are able to monitor and correct deviations from the intended interpretation of the survey materials and procedures.

3.3 Interviewer workload and compensation

Response rates can also be affected by the total amount of work allocated to each interviewer. An interviewer with many cases to deal with will naturally begin with the "easiest" ones (from their point of view), e.g. the households which can be located and accessed most easily or those for which a telephone number is available. Those cases form a sub-sample which may well be non-random. If interviewers work in parallel on several surveys (or even have several jobs), they may prioritise one survey over another (based, for example, on the payment scheme). Furthermore, less attention paid to a survey may lead the interviewers to be less persuasive when they attempt to get households to participate in that particular survey.

The level and the type of payment also affect the interviewers' incentive to work on the survey and thus the response rate. The most commonly used payment scheme in the HFCS (see next table) is payment per completed interview. This kind of payment may offer a strong incentive for the interviewers to make contact and convince the contacted households to participate in the survey. On the other hand, it may encourage the interviewers to focus more on the "easiest" cases in order to optimise their income and to give up, perhaps too early, on the households which are more reluctant to cooperate, thereby influencing any non-response bias. Payment based on the interviewer's time does not have this disadvantage, but the interviewers may not feel the need to work so hard to achieve cooperation. "Bonus" payments above a certain response rate may have a positive effect. However, the bonus system may be regarded by interviewers as unfair, since different areas and different groups of respondents may pose specific difficulties in achieving the desired response rate.

3.4 Efficiently organising and monitoring the fieldwork

Fieldwork should be efficiently organised and the progress of interviewers closely monitored. Regular reports (fortnightly or, preferably, weekly) about contacted households, refusals and cooperating households, if possible for sub-groups, should be requested. Additional information about the non-responding households may further help in identifying particular problems so that compensating action can be taken.

3.5 Contact strategy and interview protocol

Interviewers who routinely make calls at different times of a day (including evenings) and on different days of the week (including weekends), and who make several attempts over an extended period of time, tend to achieve better contact rates. Such a contact regime should be a part of any survey protocol. To allow supervisors to monitor interviewers' behaviour and to facilitate more detailed analysis of response behaviour by survey methodologists, interviewers should complete a contact record after each attempted visit or other action in each case, including the date, time, outcome in a broad sense, and some indication of obstacles to progress.[7] The contact protocol should be flexible enough to respond to unforeseen difficulties in the course of the fieldwork, such as low contact rates, which might require more frequent calls, calls at different times of day, or more calls at the weekend).

The length of the interviewing period should also be sufficient to allow enough time to reach households that are difficult to contact and to allow more conversion attempts for households that are initially reluctant to participate. Some periods (e.g. August) may be less productive, because many households are absent and, therefore, harder to contact. However, the length of the interview period must not be detrimental to the timeliness of the survey (especially when we consider the complexity of the data editing and imputation process in the HFCS).

The initial contact in the HFCS is made either by telephone or face-to-face. The only exception is the Netherlands, which uses web interviewing. In several countries, the mode is chosen by the interviewer (telephone or visit). Once cooperation has been achieved it is possible to switch to another mode to complete the interview (e.g. telephone or self-completed paper questionnaire).

3.6 Use of advance letters and incentives

A letter sent in advance of the interviewer's call has proved to have beneficial effects on response rates. The advance letter is mainly aimed at introducing the survey, its purpose, its sponsors and the different institutions involved. It can emphasise the importance of the survey, the commitment of the different stakeholders to guaranteeing data privacy and confidentiality and also provide contact information (both at the survey agency and the NCB sponsoring the survey) in case of further questions. The letter usually informs the household about the forthcoming visit of an interviewer, asking the household to get prepared to reply to the survey (e.g. by collecting all the necessary documentation). Advance letters are also important for interviewers, as they make them feel more confident when they approach households, since the latter are already informed about their visit.

In the countries where the HFCS is conducted by the NCB, the common practice is to send an advance letter signed by the NCB governor or another person in

[7] Such a contact form has been developed for the European Social Survey.

authority. The letter is sometimes accompanied by a brochure presenting the main objectives of the survey and a link to a dedicated page on the NCB's website.

Because many people discard unknown envelopes without opening them, the envelopes should be designed to achieve a high degree of saliency, without appearing to be "junk mail". In order for the advance letter to produce its desired effect, it should be sent very shortly before the anticipated time of the interviewer's first contact with the household, otherwise it may be forgotten by the time the interviewer calls or visits. It would therefore be advisable to co-ordinate the sending of the advance letters with the interviewers' work plans, rather than sending all the letters at the start of fieldwork, and it may be useful to include a brief but informative brochure with the letter. Interviewers should also be given copies of the fieldwork material, which they can use either to remind the respondents of what they have seen already or to show them what they may have thrown away.

The use of incentives to foster survey participation has proved to increase response rates, although their effects vary across countries and across the different types of incentive. According to the literature, incentives that are promised up front to all the contacted households tend to perform best, possibly because they are regarded by the households as a kind of moral obligation for them to respond. In some instances, monetary incentives seem to work better than non-monetary incentives. However, it is important to note that different types of incentive attract different population groups. For example, pure monetary incentives tend to attract lower income groups, which may generate a certain amount of undesired bias. Non-monetary incentives are found to work better than money for some of the HFCS surveys carried out by the NCBs (especially if they involve items which cannot be acquired in other ways, e.g. attractive gifts only produced by the NCB). However, the appeal of non-monetary incentives also varies across countries. For example, lottery tickets may work well in Germany, but are considered totally unacceptable in Spain. Charity donations are very well received by Irish households, but they are not used in other countries.

3.7 Refusal conversion

Interviewers must be trained in how to avoid getting a refusal on the first attempt. This can be partly achieved if the interviewer allows respondents to take their time regarding their participation in the survey and does not pressure them to make a quick decision. Experience has shown that respondents often refuse initially because they do not fully understand the survey and do not realise that their information will not be made available in an identifiable form. Where an initial refusal entirely eliminates the legal possibility of further contacts (as in Germany), every effort must be made to keep the initial contact low key and focused on imparting information, except in cases where the respondent is clearly willing to participate.

Often an initial refusal may be considered a "soft" refusal, or an "initial reluctance" to participate. Many survey organisations employ specialised staff ("refusal convertors") who concentrate on re-approaching such households in order to seek their

cooperation. This practice has worked pretty well in some surveys, although it requires strong commitment from the survey organisation. Needless to say, all the efforts to increase survey participation (including refusal conversions) must be conducted in accordance with accepted ethical and legal standards. Above all, it is completely unproductive to “harass” a household to get it to reply to the survey (there is no certainty that the information collected that way would be of good quality, and the organisation(s) involved in the data collection might run into legal issues, which might hamper their credibility).

In the European Social Survey, most countries have attempted refusal conversions, but to varying degrees and with different conversion rates – varying from 22% to 70%. Of course, the success of refusal conversion strategies depends on how easily the interviewer accepted initial refusals. Also, respondents may be more inclined to refuse initially if they expect a second attempt anyway (in some countries, a low initial response rate is very common, suggesting the “culture” in that country might be such that people expect a second attempt). These factors need to be taken into consideration when explaining response rates.

However, refusal conversions are recommended, not only to enhance the overall response rate, but mainly because they are thought to contribute efficiently to the reduction of non-response bias. This will be accomplished if the converted refusals present characteristics similar to those of the non-respondents as a whole. However, the empirical evidence supporting this is rather mixed. In the European Social Survey, however, significant differences have been found between initially cooperative respondents and initially reluctant respondents for a variety of socio-demographic and attitude variables. This finding seems to confirm the usefulness of refusal conversions.

3.8 Conclusions

It is extremely important to increase survey participation through the implementation of adequate prevention measures at each stage of the survey. In order to be efficient, the most important obstacles need to be identified in advance, which may well vary from country to country.

3.8.1 Actions before the fieldwork

- A questionnaire which is neither too long nor too burdensome
- The availability of several language versions of the questionnaire
- A sampling design which takes into account the interviewer workload (i.e. multi-stage selection to control the geographical dispersion of the sample)
- The recruitment of experienced interviewers
- The proper training and compensation of the interviewers

- The supporting legal basis (mandatory versus voluntary participation)
- The credibility of the different sponsors and stakeholders
- Assurances regarding data privacy and confidentiality

3.8.2 Actions during the fieldwork

- The calling schedule of the interviews, including evenings and weekends
- The use and type of incentives to get the households to participate
- The use of an efficient mode of data collection (CAPI, CATI, mixed etc.)
- The use of advanced information (official letter, etc.) to introduce the survey
- The monitoring of the fieldwork, managing interviewers, offering encouragement, assistance, identifying and responding to problems

3.8.3 Actions after the fieldwork

- Fieldwork auditing, refusal conversion strategies

4 Re-weighting for unit non-response in the HFCS

In spite of all the efforts survey managers and operations staff put into preventive actions, unit non-response can never be fully eradicated. That is why post-survey weighting adjustment methods are also needed to compensate for non-response bias, provided powerful auxiliary information correlated both to household response and the HFCS target variables on income and wealth is available. A review of the main sources of auxiliary information can be found in the second section of this paper. In the following, the main unit non-response correction methods are described. These methods are in line with the common guidelines for the construction of survey weights which were agreed on by the HFCN and are followed by all participating countries.

4.1 Main unit non-response correction methods

4.1.1 Response propensity weighting

Traditional non-response correction methods involve adjusting the sampling weights for the respondents by using the inverse of the estimated response propensities as adjustment factors. Response propensities are predicted on the basis of auxiliary covariates available both for respondents and non-respondents using statistical techniques such as logistic modelling or response homogeneous groups.

Let p_k denote the response probability for household k (we assume $p_k > 0$ for all k) and $x_k = (x_k^1, x_k^2 \ldots x_k^J)$ a vector of J auxiliary variables which are observed both for the respondents and the non-respondents. The x_k are supposed to have the desired properties of "good" response predictors, meaning they are correlated both to household response and to the HFCS target information on income and wealth. A logistic model then assumes the following relationship between the response probability p_k and the auxiliary variables x_k:

$$p_k = exp\left(\sum_{j=1}^{J} a_j x_k^j\right) / \left[1 + exp\left(\sum_{j=1}^{J} a_j x_k^j\right)\right] = exp(a \cdot x_k^T) / [1 + exp(a \cdot x_k^T)] \quad (5)$$

where $a = (a_1, a_2 \ldots a_J)$ is the vector of the model parameters. The vector a is often estimated by maximum likelihood. The logistic model also provides statistical indicators (p-values) to help determine the most significant non-response predictors among all the potential candidates. Furthermore, in order to control for extreme propensity values, it is common practice to smooth the distribution by creating response homogeneous groups, where the response propensity is assumed to be equal for all units. The methodology proposed in Beaumont and Haziza (2007) could be used to construct the groups. The availability of auxiliary information on both the respondents and the non-respondents is essential, and a lack of such information

would limit the scope for non-response correction. In particular, in the case of non-ignorable non-response – i.e. where non-response depends on the survey characteristics of interest – traditional approaches cannot be implemented as no information is available on the non-respondents. Consequently, alternative methods are needed, and this is where calibration techniques come in.

4.1.2 Simple calibration

Originally developed by Deville and Särndal (1992), calibration has grown today into a widely used technique in official statistics. The principle is to adjust the sampling weights using population benchmarks so that the new weights reproduce exactly the population totals (in the case of numerical variables) or the population distributions (in the case of categorical variables) for a predefined set of auxiliary variables. Thus, calibration makes survey estimates consistent with auxiliary information taken from external sources, such as population censuses or administrative registers, which are usually regarded as accurate. As shown by Deville and Särndal (1992), the sampling variance can be expected to be reduced when the calibration variables are correlated with the target variables of the survey.

The calibration approach is also easy and flexible to implement using existing software, such as the SAS macro CALMAR, developed by France's National Institute of Statistics and Economic Studies (INSEE), the SPSS program G-Calib, developed by Statistics Belgium, the SAS program CLAN, developed by Statistics Sweden, or the ad hoc functions available in the R package "Sampling" (Tillé and Matei, 2012) or in Stata (function calibrate).

In addition, the technique can be used to reduce non-response bias (Särndal and Lundström, 2005), assuming that the calibration variables are correlated with both the probability of response to the survey and the variables of interest. This approach is less stringent than the traditional one based on the explicit estimation of response propensities, as the adjustment variables are no longer required to be available for the non-respondents and no explicit response modelling is needed. In particular, this allows information collected within the survey to be used. By definition, this information is updated and, therefore, expected to be more relevant for coping with non-response bias. However, population benchmarks still need to be known, which can be a hurdle. Survey estimates based on large samples (e.g. the Labour Force Survey) may provide external benchmarks for calibration, provided that both the actual and the reference surveys follow the same concepts and definitions. In any case, this solution is still not applicable when the probability of response depends on the topic of the survey itself (i.e. non-ignorable non-response), as no benchmarks are available.

Särndal and Lundström (2005) propose the use of calibration as an integrated approach for dealing with unit non-response, reducing sampling variance and for dealing with frame imperfections. Auxiliary information available for both the respondents and the non-respondents, but for which population totals are not available from external sources (e.g. paradata, interviewer-related information), can

be used as calibration variables. In that case, the calibration totals are the unbiased estimates based on the household base weights. Särndal and Lundström (2005) emphasise that such a unified approach is flexible and easy to implement with existing software tools, and is not computationally demanding.

Table 19

Calibration variables and sources used by the individual countries, first wave of the HFCS

Country	Age	Gender	Household size	Region	Other	Source
Belgium	x	x	x			NSI
Germany	x		x	x	municipality size, homeownership, size of main residence (for homeowners); education, labour status and nationality (of main income earner)	Micro census
Greece	x		x	x	homeownership, education	LFS, EU-SILC
Spain	x	x	x		municipality size	Census 2008
France	x	x		x	municipality size, education, type of household, job qualifications, labour and wealth income	NSI
Italy	x	x		x	municipality size	NSI
Cyprus	x	x	x	x	-	Census
Luxembourg	x	x	x		nationality, labour status, stratum	Social security register
Malta	x	x	x	x	employment status	LFS
Netherlands	x	x			household population, homeownership, education	EU-SILC, CBS Statline
Austria*			x		municipality size	Micro census
Portugal	x	x	x	x	total number of households	Population statistics, LFS
Slovenia	x	x	x		-	Population register
Slovakia	x	x	x	x	Homeownership	NSI
Finland	x	x	x	x	selected income variables	Population and tax register

Source: ECB, HFCS methodological report for the first wave.
Notes:
LFS – Labour Force Survey; NSI – national statistical institute; EU-SILC – European Union Statistics on Income and Living Conditions; CBS – Statistics Netherlands.
* Cell-based poststratification.

4.1.3 Generalised calibration

Generalised calibration (Särndal and Lundström, 2005; Särndal, 2007) can produce non-response adjusted weights which reflect characteristics that are observed only for the respondents and for which no benchmark totals are available. This is a major improvement compared to simple calibration. In particular, survey variables of interest can themselves be used for non-response correction.

Let us assume the response propensity p_k of household k depends on a set of J variables $z_k = (z_k^1, z_k^2 \ldots z_k^J)$. The response variables z_k are measured within the survey, meaning they are observed in the sub-sample r of the responding units. The z_k may be continuous characteristics or dummies for population categories (e.g. by occupation, education, citizenship or gender). Contrary to simple calibration, no population benchmarks are required for the z_k. Thus, survey variables of interest can

themselves be used. For example, in the HFCS, we may have J = 1 and z_k = total net wealth for household k. We seek to modify the sampling weights d_k using an adjustment factor which is a function of the response variables z_k:

$$\omega_k = d_k \times F(z_k\lambda) \quad (6)$$

where $\lambda = \left(\lambda^1, \lambda^2 \cdots \lambda^J\right)^T$ is a column vector of dimension J and $F(z_k\lambda)$ is regarded as an adjustment factor for unit non-response, assuming the response propensity p_k can be written as the inverse of $F(z_k\lambda)$. The function F is assumed to be monotonic and twice differentiable, for instance $F(x) = 1 + x$ (linear weighting) or $F(x) = \exp(x)$ (« raking ratio » weighting).

Furthermore, let $x_k = (x_k^1, x_k^2 \ldots x_k^J)$ be a set of J calibration variables for which the population totals $X = (X^1, X^2 \ldots X^J) = (\sum_k x_k^1, \sum_k x_k^2 \ldots \sum_k x_k^J)$ are supposed to be known from reliable auxiliary sources, such as population censuses, administrative registers or sample surveys based on large sample sizes. As with the response variables z_k, the x_k may be continuous characteristics (e.g. household total income) or dummy variables pertaining to population categories. The idea of the generalised calibration approach is to estimate the unknown parameter λ by solving the system of the J calibration equations based on the x_k:

$$\sum_{k\in r} d_k \times F(z_k\lambda) \times x_k = X \quad (7)$$

If the calibration variables x_k are correlated with the target variables of the survey, then survey estimates are expected to be more stable (i.e. lower variance). This system of J equations can be solved by using iterative methods (e.g. Newton-Raphson algorithm).

For instance, in the case of linear calibration $(F(x) = 1 + x)$, we obtain the following expression for the final weights:

$$\omega_k \;=\; d_k \times (1 + z_k\lambda) \;=\; d_k \times \left[1 + z_k\left(\sum_{(k\in r)} d_k z_k^T x_k\right)^{-1}\left(X - \hat{X}\right)^T\right] \quad (8)$$

This is the generalised regression estimator (GREG) using the calibration variables x_k as regression variables and the response variables z_k as instrumental variables (Särndal, 2007). In order to circumvent identifiability problems, the number of response variables needs to be the same as the number of calibration variables.

The generalised calibration approach is implemented in the new version of the SAS macro CALMAR (CALMAR2) (Sautory, 1993 and 2003) and in the R package "Sampling" (function gencalib). Simulation studies have shown that the use of this approach can lead to a significant reduction in non-response bias when the response model is well specified (i.e. when the response propensities $p_k \propto [F(z_k\lambda)]^{-1}$). The price to pay for bias reduction is an increase in sampling variance when the calibration variables are poorly correlated with the response variables (Osier, 2013). Powerful calibration variables are needed in order to keep control over the increase in variance. There is also a potential risk of bias amplification when the instruments z_k are too weak, meaning poorly correlated with the calibration variables x_k (Haziza and Lesage, 2013). That is why this technique, although interesting, must be implemented with great caution.

4.2 Numerical example

This numerical example is aimed at illustrating the impact of the different correction methods on the HFCS estimates, namely mean and median household wealth. It is based on the Austrian HFCS data from the first wave, and its purpose is to check the impact on the main HFCS estimates of several weighting schemes. These schemes do not intend to mimic the ones used at national level (there is not enough information in the HFCS datasets to do that), but rather to allow further comparisons between the different non-response correction methods:

- Response propensity weighting: estimation of the response propensities from a logistic model based on the seven dwelling-related variables (SC variables) available in the HFCS datasets (S-file) (see section 2.1.1):
    - Type of dwelling
    - Dwelling rating
    - Dwelling location
    - Dwelling – outward appearance
    - Dwelling – comparison with the neighbourhood (Are other dwellings in the neighbourhood better or worse than this one?)
    - Dwelling – rating of surrounding buildings
    - Dwelling – security measures
- Two-step approach: (i) estimation of the response propensities from a logistic model based on the seven SC variables available in the S-file, (ii) calibration to population counts by age and gender.
- One-step approach: simple calibration to the population counts by age and gender categories and the estimated population totals (based on the household design weights) for each of the response categories of the SC variables.

For comparison purposes, the results are expressed as a percentage of those obtained using the household design weights, i.e. when unit non-response is "ignored". Basically, it appears that for a given indicator the impact of the different weighting schemes is pretty much the same, although there are also differences observed from one indicator to another.

Table 20

Mean and median wealth for the main HFCS wealth components based on different weighting schemes – percentage of the estimates based on the household design weights (HFCS variable SD0300)

HFCS code	Label	Indicator	Estimation based on the final weights	Response propensity weighting	Two-step approach	One-step approach
DA1000	Total real assets	Mean	103	97	99	98
		Median	99	96	93	92
DA1110	Household's main residence	Mean	102	98	101	101
		Median	100	95	98	98
DA1120	Other real estate property	Mean	103	97	102	101
		Median	107	104	95	95
DA1130	Household's vehicles	Mean	101	99	100	98
		Median	100	100	100	100
DA1131	Valuables	Mean	108	101	101	99
		Median	114	100	99	99
DA1140	Self-employment businesses	Mean	101	100	104	101
		Median	100	96	100	96
DA2100	Total financial assets	Mean	105	99	100	99
		Median	104	98	97	96
DA2101	Deposits	Mean	105	96	95	96
		Median	101	95	95	95
DA2102	Mutual funds	Mean	97	107	109	107
		Median	107	100	98	88
DA2103	Bonds	Mean	97	98	99	99
		Median	93	106	106	100
DA2104	Non self-employment private businesses	Mean	117	166	155	138
		Median	100	114	114	114
DA2105	Shares, publicly traded	Mean	95	105	100	94
		Median	100	95	100	100
DA2106	Managed accounts	Mean	100	100	100	100
		Median	100	100	100	100
DA2107	Money owed to households	Mean	98	95	97	95
		Median	100	93	93	93
DA2108	Other assets	Mean	112	93	89	93
		Median	100	93	90	93
DA2109	Voluntary pension/whole life insurance	Mean	98	102	101	97
		Median	102	100	99	99
DA3001	Total assets	Mean	103	97	98	97
		Median	96	88	83	79
DN3001	Net wealth	Mean	103	97	98	97
		Median	99	91	85	82

Source: ECB, HFCS, first wave.

4.3 Recommendations

Reweighting methods need to be used ex post in order to compensate for unit non-response bias. Basically, there are two kinds of approach: one based on the explicit estimation of the response propensities and another based on calibration techniques. The calibration approach offers several advantages: it is simple and flexible to implement with existing software tools. At the same time, it is able to reduce non-response bias, reduce sampling variance and deal with frame imperfections. On the other hand, calibration is regarded by some users as an opaque technique, unlike the more traditional response propensity weighting approach.

The impact of both kinds of approach on the HFCS-based estimates seems pretty much the same. Thus, the choice between them is a matter of personal preference. Although calibration is easy, flexible and universal, some might consider it more reasonable to follow a more classical two-step approach (estimation of the response propensities and calibration to reduce sampling variance), where each step is kept under control. In any case, the choice of a reweighting scheme should be made in the light of its impact on survey estimates, especially on estimates for small sub-populations.

The generalised calibration approach may be well suited to cope with non-ignorable non-response, i.e. mechanisms which depend on variables which are observed only for the respondents. However, the approach might lead to unstable and even biased results if the calibration variables are not well chosen. Therefore, this approach must be used with great caution.

5 General conclusion – main lessons

There are several important lessons that can be drawn from this paper:

- Unit non-response is a key concern for the HFCS, as non-response rates are important in some countries and response patterns are not completely random. The factors underlying unit non-response to the HFCS are diverse and also country and survey-specific. The way the survey is designed, the legal basis (mandatory/voluntary), the quality of interviewer training, the use of incentives, etc. may also affect response patterns.

- To deal with unit non-response bias, it is of utmost importance that all the auxiliary information available on the non-respondents be gathered in order to properly describe and analyse response patterns. This paper provides an extensive list of possible auxiliary sources to be considered for non-response analysis. In this respect, administrative, banking and financial data should be especially relevant as they convey information which is highly correlated to the main topics of the HFCS.

- Relevant contact and interview strategies are required to increase survey participation and reduce non-response bias. In addition, reweighting strategies need to be implemented ex post. There is no single reweighting approach which could be recommended for all cases, each having pros and cons. Basically, the choice between the traditional response propensity weighting approach and calibration techniques mainly depends on their impact on the key HFCS aggregates on income and wealth.

Annex
Countries' fieldwork practices from the first wave of the HFCS

Table 1

Interviewer training, first wave of the HFCS

Country	Description
Belgium	6 hours, 40 participants in the presence of NCB staff. Issues covered: content of the survey (history, goals, architecture of the questionnaire, the nature of the different questions, navigation through the CAPI questionnaire, etc.); financial terminology and products, contact procedure.
Germany	11 hours, 230 participants. Issues covered: plenary sessions on the background of the survey, sampling design, interviewer conduct, interviewer material and special features of the CAPI program. Practical training sessions in smaller groups of about 20-30 interviewers: screening and household matrix, contact protocol, interview with financially knowledgeable person, interview with a household member. Additional training: one or two pre-test interviews in the week following the interviewer training. Interviewers that didn't complete at least one pre-test interview had to take part in additional face-to-face training.
Greece	8 hours, 20 participants. Issues covered: sample selection procedures and recruitment techniques; presentation of the study: general info; reading through the questionnaire question by question; explaining each question and its structure; explaining the follow-up to each item; simulation: completion of questionnaires with a supervisor; role playing and simulation of a real interview.
Spain	28 hours, 80 participants. Issues covered: one centralised training session just prior to the start of data collection in a hotel around 100 km outside Madrid to try and ensure full-time commitment to this task. First, the 14 field managers were given a two and a half days briefing. Following this, three and a half days of training. During these sessions the questionnaire was analysed in detail by going through hypothetical cases and getting familiar with this particular CAPI application. Various representatives of the NCB participated in these sessions to explain the importance and difficulty of the project and to clarify any matters arising during the explanation of the questionnaire. Arguments to reduce non-cooperation were also discussed, as well as appropriate ways of approaching households. Prior to the training, all interviewers were sent material to familiarise themselves with the study. At the end of the training, all interviewers had to conduct a mock interview with a predefined script and an exercise in gaining cooperation.
France	INSEE collects survey data through its network of interviewers who have often been working for INSEE for many years. Thus they have a good experience of interviewing and are aware about interviewing instructions required by INSEE to ensure high quality data collection. Interviewer training was conducted during the preparation of the survey. The team which designed the survey and the manager responsible for the data collection process gave face-to-face instructions to the interviewers. Each session involved between 20 and 30 interviewers and lasted two days. Then there was an in-home session dedicated to exercises so that interviewers could experience the questionnaire and some typical issues that could arise during the interviews. This in-home session was supposed to last six hours.
Italy	8 hours, 192 participants. During the fieldwork, there is constant communication between the NCB, the survey agency and the interviewers in order to exchange suggestions, experiences and so on.
Cyprus	5 hours, 20 participants. Three training sessions of five hours each were organised and the interviewers were split in three groups. Issues covered: explanation of the definitions and questionnaire, actual practice of the CAPI questionnaire and practices for persuading respondents to participate.
Luxembourg	6 hours, 41 participants. The interviewers were divided in six groups. During the first part of the training, the survey was presented and the questionnaire was explained in detail. During the second part of the training, they learned about the questionnaire on CAPI. Due to the small size of the groups (maximum eight persons), the interaction between interviewers and trainer(s) was easier. Written handbook provided for interviewers with information on how to use the CAPI, documents related to the survey (documents received by the households, list of addresses, pre-selection interview to identify the financially knowledgeable person, list to document the number of contacts, paradata section), explanation of the survey, key items, and the loops within the survey.
Malta	9 hours, 30 participants. Issued covered: introducing yourself, how to handle refusals, rude persons, awkward situations; the interviewer manual, making the first contact, identifying the financially knowledgeable person, making the appointment and the paper material distributed (non-contact form, the payments diaries, the introductory letter), familiarisation with the CAPI program and questionnaire, hands-on training on the CAPI application software, the questionnaire.
Netherlands	Not applicable (web-based data collection).
Austria	7 hours, 85 participants.

	Issues covered: introduction to the survey with special emphasis on the goal of the survey and the importance of the work of interviewers, administrative provisions including contact strategy for households and contact details at the Oesterreichische Nationalbank for further questions, interviewer behaviour, structure and content of the questionnaire, including the exact definition of the term household within the meaning of the HFCS, special features of the questionnaire (loops and Euro loops, paradata), test interview with a complicated household (i.e. a household owning a business and having loans), clarification of technical terms, interview documents (Interviewer Handbook, Glossary, Information Folder, etc.).
Portugal	16 hours, 163 participants. Issues covered: motivation, efforts to minimise non-response, questionnaire and related concepts, CAPI, “practical” session with simulated interviews.
Slovenia	7 hours, 22 participants. Issues covered: instructions regarding contacting the household, entering the household, technical issues with CAPI, content of the survey, answering the interviewers’ questions, handing over the materials.
Slovakia	4 hours, 50 participants. Issues covered: description of HFCS, data protection issues, CAPI simulation, training in refusal conversion.
Finland	40 hours (including general NSI training), all new interviewers participated. Issues covered: all interviewers are trained on a one-week face-to-face basic training course when they begin work, including training in refusal conversion. New interviewers attend a special EU-SILC training session. All interviewers also have SILC-specific remote training (getting to know the instructions and the questionnaire).

Source: Metadata provided by the countries.

Table 2

Fieldwork organisation, first wave of the HFCS

Country	Description
Belgium	Supervision organised by the survey company. Each interviewer is evaluated after his/her first three interviews and feedback is delivered to him/her. Both local and travelling interviewers.
Germany	Local and travelling interviewers. Each interviewer is assigned to a regional fieldwork manager. In addition, the fieldwork is overseen by fieldwork managers at the survey agency headquarters. Usually interviews are conducted by local teams. So called "full-timers" travel and conduct interviews with households and persons who have moved to an area where no other (trained) interviewer lives. In the second part of the fieldwork phase the "full-timers" also travelled to big cities to support local interviewers with regular interviews. One interviewer per municipality. In large cities one interviewer per sample point.
Greece	Local and travelling interviewers. During fieldwork there is a continuous contact between the fieldwork manager, fieldwork supervisors and interviewers, both in person and by telephone, to provide feedback. The fieldwork supervisors are responsible for checking at least 20% of each interviewer's work, filling in a specific form for these cases.
Spain	Local and travelling interviewers.
France	INSEE has a subsidiary in each region which allows local supervision of the interviewer teams. Supervisors are the first step in the dialogue with interviewers. Then every regional subsidiary reports response rates and data collection incidents to the manager in charge of supervising national collection progress who collates the information centrally.
Italy	Local interviewers, ten local supervisors who are responsible for monitoring and dealing with different teams of interviewers, three national supervisors who are responsible for monitoring the entire fieldwork, two IT experts for the exchange of data transmission between interviewers and the agency and one person responsible for the whole project with a direct contact with the Banca d'Italia. A team for the editing of raw data.
Cyprus	Local and travelling interviewers. The company undertaking the fieldwork organisation used techniques that ensured the proper flow of the fieldwork when using CAPI (e.g. the time that the interview began and the time that the interview was completed appeared on the server where the data were automatically saved, and the identity of each interviewer was also available). When the paper and pencil interview (PAPI) version of the questionnaire was used, there was no opportunity to check this through server, but only through the paradata section completed by the interviewer by hand after the end of the interview.
Luxembourg	Local interviewers. Each interviewer is attached to a unique supervisor. At the launch phase of the survey, interviewers received about 20 addresses and were asked to return to the survey agency after having successfully completed two questionnaires in order to discuss any problems encountered and to provide supervisors with early feedback from the fieldwork. After the launch phase, interviewers went regularly to the survey agency to transfer the completed questionnaires and to receive a new set of addresses. In addition to regular face-to-face meetings, supervisors also had phone contacts with the interviewers. Fieldwork supervisors were in charge of the management of the interviewers and had an important role in assuring the quality of data collection. A supervisor managed a team of interviewers, distributed new addresses to them, transferred the data collected from the laptop, handled problems encountered during the fieldwork, responded to phone calls from households, controlled the interviewers' work, supported and encouraged the interviewers to continue, even if the fieldwork was difficult, and provided feedback to the survey supervisor and the IT team.
Malta	Local interviewers, three supervisors who followed the progress made by each interviewer. The interviewers can contact their supervisor either by phone, by going directly to the survey agency or by e-mail for any difficulty. Every fortnight the interviewers have to report their progress.
Netherlands	Not applicable.
Austria	Local interviewers. The NCB developed special interviewer sheets with various performance indicators for every single interviewer that were very closely monitored together with the recorded data. The head of field staff in the survey agency reported to the experts in the NCB team.
Portugal	Local interviewers, national supervision team that ensures large-scale analysis and supports regional teams. Regional organisation with hierarchical structure –project manager, supervisors, interviewers.
Slovenia	Local interviewers. The manager of the operative centre was responsible for the start of the project, daily supervision, selection of interviewers, leading training seminars (together with field supervisors) etc. He reported any problem to the project manager. Each day field supervisors checked the previous day's work. The daily report and comments from the manager of the operative centre warned the interviewers about any mistakes. He was in contact with them at least two times per week. The field supervisor was responsible for ongoing control of telephone and field contact with the households and reported any problem to the manager of the operative centre and the project manager. After the interviewing was finished, they checked the remaining materials, tokens, etc. and calculated the payment for each interviewer.
Slovakia	Local and travelling interviewers. Connection with interviewers during the whole fieldwork (face-to-face, by telephone or by e-mail), giving the instructions according to the regular data monitoring, controlling the fulfilment of the prescribed quotas.
Finland	Local interviewers. Centrally organised supervision in a specialised unit in Statistics Finland. Help desk operates by e-mail, telephone or online during the fieldwork period. Controlled data return system from the interviewer to the central organisation with automatic checks, interviewer-specific follow-up of non-response and interruption rates and other qualities of the returned units. Extra work for non-response prevention in local interviewer groups is supported.

Source: Metadata provided by the countries.

Table 3

Contact strategies, first wave of the HFCS

Country	First contact	Minimum required contact attempts	Possibility to re-contact households	Number of households re-contacted
Belgium	Telephone or visit, depending on sampling frame.	4 for households contacted face-to-face, 10 for households contacted by telephone.	Yes	n.a.
Germany	Telephone or visit	5	No	-
Greece	Visit	4	Yes	n.a.
Spain	Visit	5	Yes	800
France	Visit		No	-
Italy	Visit	3	Yes	n.a.
Cyprus	Telephone	4	No	
Luxembourg	Telephone or visit. However, addresses distributed to the interviewers did not contain any phone numbers.	3	No	-
Malta	Telephone, if the number is available. Otherwise visit.	3	Yes	n.a.
Netherlands	Web		Yes	n.a.
Austria	Telephone or visit, chosen by the interviewer.	5	Yes	400
Portugal	Visit	3, but in urban areas the number was higher.	Yes	n.a.
Slovenia	Visit	4	Yes	13
Slovakia	Visit	1	Only for verification.	n.a.
Finland	Telephone	3 for personal interviews, no limit for telephone interviews.	Yes	n.a.

Source: Metadata provided by the countries.

Table 4

Advanced information provided to households, first wave of the HFCS

Country	Type of advanced information	Advanced letter signed by
Belgium	Letter and brochure	NCB Governor
Germany	Letter, brochure and data collection web page	NCB President
Greece	Letter	Bank of Greece logo and contact details, no signature.
Spain	Letter, brochure and data collection web page.	NCB Governor
France	Letter	INSEE (Regional Director)
Italy	Letter, brochure, press release and data collection web page.	NCB Governor
Cyprus	Letter	NCB Governor
Luxembourg	Letter, brochure and data collection web page.	NCB Governor
Malta	Letter	On behalf of NCB Governor, no signature
Netherlands	Letter, brochure and data collection web page.	Survey agency director
Austria	Letter and data collection web page, brochure about the survey	NCB Governor
Portugal	Letter and brochure	Head of Demographic and Social Statistics (NSI) and Head of Economic and Research Department (NCB)
Slovenia	Letter and brochure	NCB Governor
Slovakia	Letter, contact details of the persons responsible for the project both at the NCB and the survey agency.	NCB Governor
Finland	Letter, brochure and data collection web page.	Director of Social Statistics (NSI)

Source: Metadata provided by the countries.

Table 5

Incentives given to respondents, first wave of the HFCS

Country	Value and description
Belgium	Monetary (€10) with option to donate the amount to charity.
Germany	€10 per household; €2.50 for each household to be used at interviewer's discretion. Additional monetary incentive (€10 per person) for low income households in the second part of the field phase.
Greece	Pocket calculator.
Spain	Token gift.
France	None.
Italy	A print representing a collection of images of Lire banknotes issued in 1915.
Cyprus	A token gift from the Central Bank of Cyprus and entry into a prize draw with the chance to win a €1,000 holiday voucher.
Luxembourg	None.
Malta	Gift voucher worth €40.
Netherlands	Points given by the survey agency to panel members for completed interviews. On average, the value of points received per interview was 25 cents, but the value increased to up to two times the original value for persons who participated longer on the panel. Points are paid out either in cash or in the form of a donation to a charity or entry into a state lottery, as the household prefers. In addition, three or four times a year, the panellists receive a magazine with background information and an overview of the main results of research based on the panel data.
Austria	Silver coin denominated at €5 (actual value of the silver coin around €15 at the time of the field phase). Additional prize draw to win a gift voucher for a travel agency (one voucher worth €1,000 and five vouchers worth €200 each).
Portugal	Token gift.
Slovenia	A special edition €3 coin.
Slovakia	Small gift, value €3.70.
Finland	None.

Source: Metadata provided by the countries.

References

American Association for Public Opinion Research (2008), Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 5th edition, Lenexa, Kansas: AAPOR.

Bańkowska, K., Osiewicz M. and Pérez-Duarte, S. (2015), "Measuring nonresponse bias in a cross-country enterprise survey", *Statistics Paper Series*, No 12, ECB, Frankfurt am Main, December.

Blom, A. (2009), "Nonresponse Bias Adjustments: What Can Process Data Contribute?", *ISER Working Paper Series*, No 2009-21, Institute for Social and Economic Research (ISER), University of Essex, United Kingdom.

Deville, J.C. and Särndal, C.E. (1992), "Calibration Estimators in Survey Sampling", *Journal of the American Statistical Association*, Vol. 87, No 418, pp. 376-382.

European Central Bank (2013), "The Eurosystem Household Finance and Consumption Survey: methodological report for the first wave", *Statistics Paper Series*, No 1, European Central Bank, April.

European Central Bank (2013b), "The Eurosystem Household Finance and Consumption Survey: results from the first wave", *Statistics Paper Series*, No 2, European Central Bank, April.

Jäntti, M., Törmälehto, V.M. and Marlier, E. (eds.) (2013), "The use of registers in the context of EU-SILC: challenges and opportunities", *Statistical working papers*, Eurostat.

Groves, R.M. and Couper, M.P. (1998), *Nonresponse in Household Interview Surveys*, Wiley, New York.

Groves, R.M. and Heeringa, S.G. (2006), "Responsive design for household surveys: tools for actively controlling survey errors and costs", *Journal of the Royal Statistical Society*, Series A (Statistics in Society), Vol. 169, Issue 3, pp. 439-457.

Groves, R.M. and Peytcheva, E. (2008), "The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis", *Public Opinion Quarterly*, Vol. 72, Issue 2, pp. 167-189.

Hansen, M.H. and Hurwitz, W.N. (1946), "The Problem of Nonresponse in Sample Surveys", *Journal of the American Statistical Association*, Vol. 41, No 236, pp. 517-529.

Haziza, D. and Beaumont, J.F. (2007), "On the construction of imputation classes in surveys", *International Statistical Review*, Vol. 75, pp. 25-43.

Haziza, D. and Lesage, E. (2013), "On the problem of bias amplification of the instrumental calibration estimator with missing survey data", 2013 Graybill Conference, Fort Collins, Colorado.

Heij, V. de, Schouten, B. and Shlomo, N. (2010), "RISQ manual, Tools in SAS and R for the computation of R-indicators and partial R-indicators".

Kennickell, A.B. (2007), "The Role of Over-sampling of the Wealthy in the Survey of Consumer Finances", *Survey of Consumer Finances Working Papers*, Federal Reserve Board, July 2007.

Kersten, H.M.P. and Bethlehem, J.G. (1984), "Exploring and reducing the nonresponse bias by asking the basic question", *Statistical Journal of the United Nations Economic Commission for Europe*, Vol. 2, No 4, pp. 369-380

Laflamme, F. and Karaganis, M. (2010), "Implementation of Responsive Collection Design for CATI Surveys at Statistics Canada", paper presented at the Symposium on Recent Advances in the Use of Paradata (Process Data) in Social Survey Research, London, December.

Little, R.J.A. and Rubin, D.B. (2002), *Statistical Analysis with Missing Data*, 2nd edn., Wiley, New York.

Mathä, T., Porpiglia, A. and Ziegelmeyer, M. (2012), "The Luxembourg Household Finance and Consumption Survey (LU-HFCS): Introduction and Results", *Working Paper Series*, No 73, Banque Centrale du Luxembourg.

Meyer, B.D., Mok, W.K.C. and Sullivan, J.X. (2015), "Household Surveys in Crisis", *Journal of Economic Perspectives*, Vol. 29, No 4, pp. 199-226.

National Research Council (2013), *Nonresponse in Social Science Surveys: A Research Agenda*, The National Academies Press, Washington, DC.

Oh, H.L. and Scheuren, F.S. (1983), "Weighting adjustments for unit nonresponse", in Madow, W.G., Olkin, I. and Rubin, D.B. (eds.), *Incomplete Data in Sample Surveys*, Vol. 2,, Academic Press, New York, pp. 143-184.

Osier, G. (2013), "Dealing with Non-ignorable Non-response using Generalised Calibration: A Simulation Study based on the Luxemburgish Household Budget Survey", *Economie et Statistiques, Working papers du STATEC*, No 65, STATEC, Luxembourg.

Pérez-Duarte, S., Sánchez-Muñoz, C. and Törmälehto, V.M. (2010), "Re-weighting to reduce unit non-response bias in household wealth surveys: a cross-country comparative perspective illustrated by a case study", paper presented at the Q2010 European Conference on Quality in Official Statistics, Helsinki.

Rubin, D.B. (1987), *Multiple Imputation for Nonresponse in Surveys*, Wiley, New York.

Särndal, C.E. and Lundström, S. (2005), *Estimation in Surveys with Nonresponse*, Wiley Series in Survey Methodology.

Särndal, C.E. (2007), "The calibration approach in survey theory and practice", *Survey Methodology*, Vol. 33, No 2, Statistics Canada, pp. 99-119.

Sautory, O. (1993), "La macro CALMAR: Redressement d'un échantillon par calage sur marges", *INSEE Working Paper*, No F 9310, INSEE.

Sautory, O. (2003), "Calmar 2: a new version of the Calmar calibration adjustment program", in *Proceedings of Statistics Canada's Symposium 2003*.

Schouten, B., Calinescu, M. and Luiten, A. (2009), "Optimizing quality of response through adaptive survey designs", *Statistics Netherlands Discussion Paper*, No 201118, Statistics Netherlands.

Schouten, B., Cobben, F. and Bethlehem, J. (2009), "Indicators for the representativeness of survey response", *Survey Methodology*, Vol. 35, No 1, pp. 101-113.

Schouten, B., Shlomo, N. and Skinner, C.J. (2011), "Indicators for Monitoring and Improving Representativeness of Response", *Journal of Official Statistics*, Vol. 27, No 2, pp. 231-253.

Stoop, I. (2005), *The hunt for the last respondent*, The Netherlands Institute for Social Research.

Tillé, Y. and Matei, A. (2012), The R package "Sampling".

Household Finance and Consumption Network

This paper contains research conducted within the Household Finance and Consumption Network (HFCN). The HFCN consists of survey specialists, statisticians and economists from the ECB, the national central banks of the Eurosystem and a number of national statistical institutes.

The HFCN is chaired by Oreste Tristani (ECB) and Carlos Sánchez Muñoz (ECB). Michael Haliassos (Goethe University Frankfurt), Tullio Jappelli (University of Naples Federico II), Arthur Kennickell (Federal Reserve Board) and Peter Tufano (University of Oxford) act as external consultants, and Sébastien Pérez-Duarte (ECB) and Jiri Slacalek (ECB) as Secretaries. The HFCN collects household-level data on households' finances and consumption in the euro area through a harmonised survey. The HFCN is aimed at studying in depth the micro-level structural information on euro area households' assets and liabilities. The objectives of the network are:

1) understanding economic behaviour of individual households, developments in aggregate variables and the interactions between the two;
2) evaluating the impact of shocks, policies and institutional changes on household portfolios and other variables;
3) understanding the implications of heterogeneity for aggregate variables;
4) estimating choices of different households and their reaction to economic shocks;
5) building and calibrating realistic economic models incorporating heterogeneous agents;
6) gaining insights into issues such as monetary policy transmission and financial stability.

The paper is released in order to make the results of HFCN research generally available, in preliminary form, to encourage comments and suggestions prior to final publication. The views expressed in the paper are the author's own and do not necessarily reflect those of the European System of Central Banks (ESCB).

Acknowledgements

I would like to thank my colleagues from the Household Finance and Consumption Network (HFCN) for their support in this work and their useful comments to improve the quality of this paper. I am grateful to the members of the Task Team on Unit Non-response, and in particular to:

• Ms Maria Luisa Alcoforado Farinha (Banco de Portugal)
• Ms Jaanika Meriküll (Eesti Pank)
• Mr Andrea Neri (Banca d'Italia)
• Ms Panagiota Tzamourani (Deutsche Bundesbank)
• Mr Laurent Van Belle (Nationale Bank van België/Banque Nationale de Belgique)
• Mr Tibor Zavadil (Národná banka Slovenska)

for producing and sharing non-response indicators based on their national data.

Guillaume Osier

At the time of writing, Guillaume Osier was at the European Central Bank, Frankfurt, Germany; now at STATEC, Luxembourg;
email: guillaume.osier@statec.etat.lu

© European Central Bank, 2016

Postal address	60640 Frankfurt am Main, Germany
Telephone	+49 69 1344 0
Website	www.ecb.europa.eu

All rights reserved. Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the ECB or the authors.

This paper can be downloaded without charge from www.ecb.europa.eu and from RePEc: Research Papers in Economics. Information on all of the papers published in the ECB Statistics Paper Series can be found on the ECB's website.

ISSN	2314-9248 (online)
ISBN	978-92-899-2345-3
DOI	10.2866/84445
EU catalogue No	QB-BF-16-003-EN-N