20 November 2019

Economic Commission for Europe

Conference of European Statisticians

Expert meeting on measuring poverty and inequality: SDGs 1 and 10

Geneva, Switzerland
5–6 December 2019
Item C of the provisional agenda
Improving response rate and sampling precision in surveys

Non-response Analysis in the Household Budget Surveys in Bosnia and Herzegovina

Note by the Agency for Statistics of Bosnia and Herzegovina*

Summary

Non-response is one of the most important non-sampling errors and a pervasive problem for every manager of statistical surveys. It occurs in censuses as well as in sample surveys. Non-response increases the total survey error and it has two effects on the statistical results: it contributes to an increase in the total variance of survey estimates and it can introduce a bias in estimates. Therefore, the analysis of survey non-response and the implementation of adequate methods for non-response reduction are essential elements in optimizing the overall survey design.

The Household Budget Survey is one of the most complex household surveys with quite big sample size in the Agency for Statistics of Bosnia and Herzegovina. Starting in 2004, the survey is conducted as a cross-sectional survey with irregular periodicity (every three or four years) and on the basis of a sample of about 7,500 households. At the time of its introduction in the practice of the Agency for Statistics of Bosnia and Herzegovina, the household budget survey was methodologically harmonized in line with agreements and recommendations of Eurostat. Since then, significant improvements in the HBS methodology were made and the survey was conducted four times so far. These improvements had the consequences in the extension of survey instruments and the increase of the average time of the interviews, which, among other factors, influenced response rates.

The aim of this paper is to analyse non-response rates and main reasons for non-responses within household budget surveys in Bosnia and Herzegovina in the period 2004-2015. The main goal of this work is to indicate trend in development of this kind of non-sampling error and to propose new methods

* Prepared by Edin Šabanović.

for non-response reduction. For this purpose, methods of descriptive and inferential statistics will be used. In order to describe non-responses within each survey year and to monitor changes over time, basic measures of central tendency will be calculated and most important non-response indicators will be presented by statistical graphs. Statistical significances of the changes in total non-response rates will be tested using t-tests for independent samples and one-way ANOVA tests for independent samples. The results of this analysis will be disaggregated by survey strata in order to monitor territorial differences. The paper shows the significant increase of total non-response rates within household budget surveys in Bosnia and Herzegovina. Territorial differences in response rates are also significant, which will require different approaches in solving this problem. Additionally, main causes for increasing non-response rates will be presented.

In the second part, the paper will critically present methods for decreasing non-response, used to date. It will be done by comparing current practices with all possible factors which can influence coverage and response rates for any statistical survey (survey design, sampling frame quality, data collection methods, communication strategy, response burden imposed, etc.). The paper will discuss new approaches for non-response reduction and how to harmonize them with new data collection method within the household budget survey, which is in preparation. Namely, the Agency for Statistics of Bosnia and Herzegovina plans to start with mixed data collection mode within the household budget survey in 2020, where diaries will be collected with PAPI technique, while final interview will be performed with CAPI application. The paper ends with realistic recommendations and further steps for non-response reduction in the household budget survey in Bosnia and Herzegovina.

1. Introduction

Modern statistical production is increasingly based on sampling surveys rather than on statistical censuses. There are two main reason for that: budget limitation and high sampling survey accuracy, which is a result of sophisticated statistical techniques and computer development.

However, there are errors in every statistical survey, no matter is it based on census or sample survey. Total survey error is a cumulative value of all posible errors in statistical survey and it defines survey accuracy, which is measured as a deviation of a survey estimate from its underlying true parameter value (Biemer, 2010, p. 817). Non-response in statistical surveys is one of possible generators of survey errors.

2. Non-response in sampling surveys

There are two main types of non-response in statistical surveys: unit non-response and item non-response. The first type occurs in the case where sampled unit did not participate in the survey. A special case of unit non response is a wave non-response, which is caused by attrition in longitudinal statistical surveys. From the other hand, item non-response considers missing answers to specific items (survey variables).

Almost every non-response can affect the quality of the survey, since it increases the total survey error. Non-response has two effects on survey estimates (Cornish, 2002, p. 1):

(i) it contributes to an increase of sampling variance of estimates, as the sample size is reduced; and

(ii) it contributes to bias of estimates, when respondents differ from respondents in the survey characteristics.

The findings of several statisticians (Curtin et al., 2000, p. 414; Keeter et al., 2000, p. 126) show that changes in non-response rates do not necessarily alter the results of the survey and that the bias is not a simple function of nonresponse level. The survey bias is rather a multiplicative function of the nonresponse level and the differences in the survey question between respondents and non-respondents. The recommendation to survey practicioners is that the best solution for non-response is not to have any. Since, it is not possible today, the best solution for this problem is the effort to minimize non-response rates. There is no common opinion which is a minimum response rate in order to garantee survey accuracy and representativeness. There are examples of policy stating that no survey should be approved that anticipates less than a 50% response rate (according Rubin D.B., 1978, p. 20), or opinions that response rate of 60% is good, while 70% is very good (Babbie E., 2007, p. 262). Other authors stated that minimum of 85% of response is an adequate and that below 70% of responses there is a serious possibility for survey bias (Singleton R., Bruce S., 2005, p. 145). In line with findings of Curtin and Keeter, Groves (2006, p. 650) found that there is no minimum response rate below which survey estimates are necessarily subject to bias.

Because of all these reasons, it is very important to minimise survey non-responses. In order to do it, it is needed to monitor, measure and report non-responses in statistical surveys and to apply appropriate methods for decreasing non-response.

3. Methods for decreasing survey non-response

In order to be able to reduce survey non-response, this phenomenon must first be monitored and reported. For a good strategy of non-response reduction, it is very important to identify and measure different components of non-response (e.g. refusals, non-contacts, vacant dwelling, persons temporarily absent, etc.). Knowing the reasons of non-response, helps us reducing it and investigating whether the bias is introduced or not. The most common indicators used for reporting non-response in household surveys are following:

- overall non-response rate,
- non-contact rate,
- refusal rate and
- percentage of vacant dwelling.

The non-response in statistical surveys is a multi-factor phenomenon. Therefore, good survey management and strategy for non-response reduction have to be adjusted to this multidimensional nature of the non-response. The selection of methods for non-response reduction is related to all possible factors influencing survey response. The most common methods, which can reduce non-response in household surveys are as follows:

(i) Improving the quality of the sampling frame;
(ii) Data collection method;
(iii) Questionnaire design and respondent burden;
(iv) Length of data collection period and time in year for data collection;
(v) Legal obligation and protection of individual data confidentiality;
(vi) Communication strategy;
(vii) Trained survey staff;
(viii) Total survey budget;
(ix) Use of respondent incentives;
(x) Language of the questionnaire;
(xi) Postsurvey adjustments.

Up-to-date sampling frame is one of the first factors influencing the quality of the household survey. Such frame significantly decreases coverage errors and non-contacts during the data collection. Sampling frames of a good quality are usually created on the basis of population censuses or population registers and their quality depends of the quality of these bases and of the frequency of their updating.

An appropriate data collection method is one of the most important factors, which directly affects response burden and co-operation of respondents in the survey. The time of face-to-face surveys passes and new methods of data collection appear and they are usually more attractive and cheaper. Computer Assisted Paper Interview (CAPI), Computer Assisted Telephone Interview (CATI) and Computer Assisted Web Interview (CAWI) become a regular practice of many national statistical institutes. Although at the beginning of their use, new methods of collecting data confirmed the decrease of overal non-response rates and number of non-sampling errors, their implementation must be adjusted to survey specifics and to specific groups of target population. In last years, reduced telephone number listings and people's increasing resistance to unwanted phone calls, alternatives to telephone surveys, such as internet-based approaches, should be investigated (O'Toole J. et al., 2008). One of practical solutions to this problem is the implementation of mixed mode of data collection.

The design of survey questionnaire depends of the nature of survey, as well as of data collection mode. It has direct influence on respondent burden and, consequently, to the response rate. Therefore, the design of the questionnaire must be adjusted to the method and the procedures of the data collection. Language of the questionnaire must be adjusted and local interviewers must be used, if there are specific ethnic groups in the survey target population.

The length of the interview, the length of the whole data collection and the period in year for data collection are factors, which are correlated each other and with previous two factors/methods. Provided the other factors remain unchanged, the longer data collection, the higher non-response rate.

Statistical surveys, which are by law mandatory, can have higher response rate in developed countries, where the implementation of laws shows high standards. In the same time, it increases the level od statistical data protection and confidentiality in survey and national statistical office.

Good communication strategy and well trained survey staff are factors reducing non-response. They affect the success of the first contacts with respondents as well as measurement errors, which are very important for the final result of the interview and the quality of collected data.

The total survey budget is a factor which is highly corrrelated to all others. In every country, even those developed, survey budgets are limited. Therefore, the conduction of the survey must be based on cost-benefit approach in defining every specific budget line.

The use of respondent incentives direct depends of the total survey budget. Their implementation increases survey response and it has been evident that prepaid monetary incentives work better in comparison to postpaid (i.e. conditional) and non-monetary incentives (Church, 1993; Singer et al., 1999).

In addition to previous factors/methods for non-response reduction, the postsurvey adjustments are used in order to reduce the non-response and to improve the overall data quality. These techniques are applied after the data collection and they usually include data editing, missing data imputation, weighting adjustments, and poststatification techniques. The results of their implementation depend of the availability and the quality of auxiliary variables and applied adjustment methods.

4. Non-response in Household Budget Surveys in Bosnia and Herzegovina

The analysis of non-response within the household budget surveys in Bosnia and Herzegovina was performed on data from all four survey years: 2004, 2007, 2011 and 2015. Table 1 presents overall non-response rates for the country and its administrative regions (two entities and district) in that period:

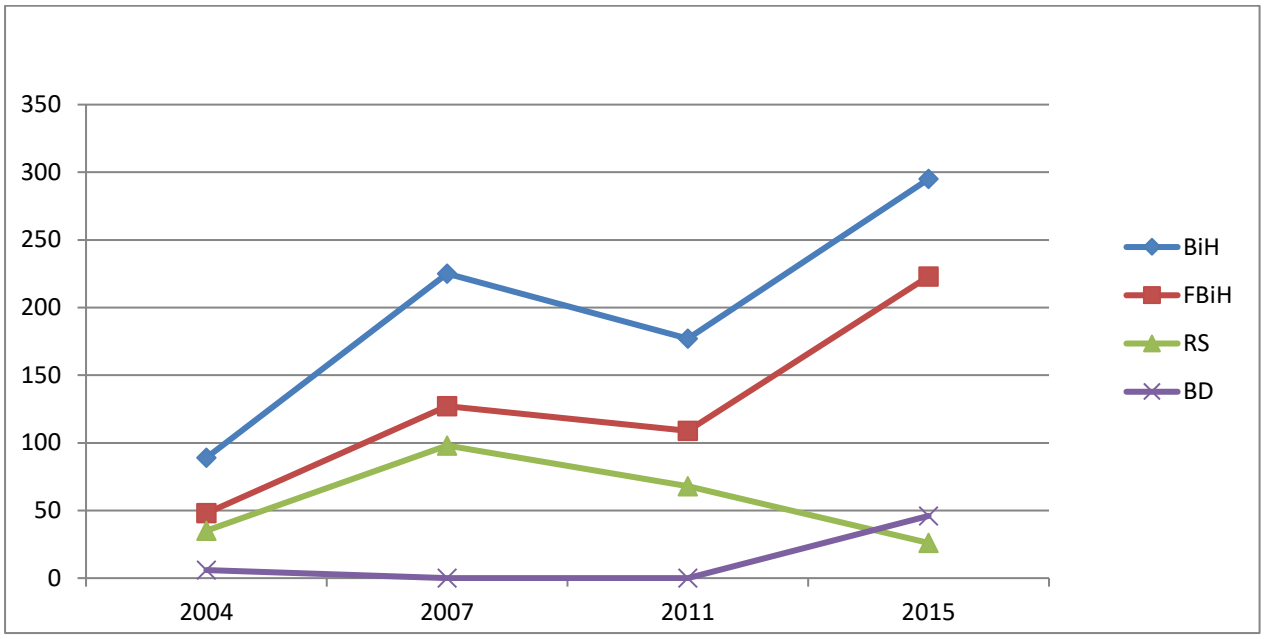
Table 1. Overal non-response rates, national and by entities, BiH HBS, HBS 2004, 2007, 2011 and 2015

Year	BiH	FBiH	RS	BD
2004	17,2	18,1	15,9	17,0
2007	18,5	20,1	16,4	10,7
2011	24,0	24,8	21,1	30,7
2015	33,2	35,3	23,6	52,1

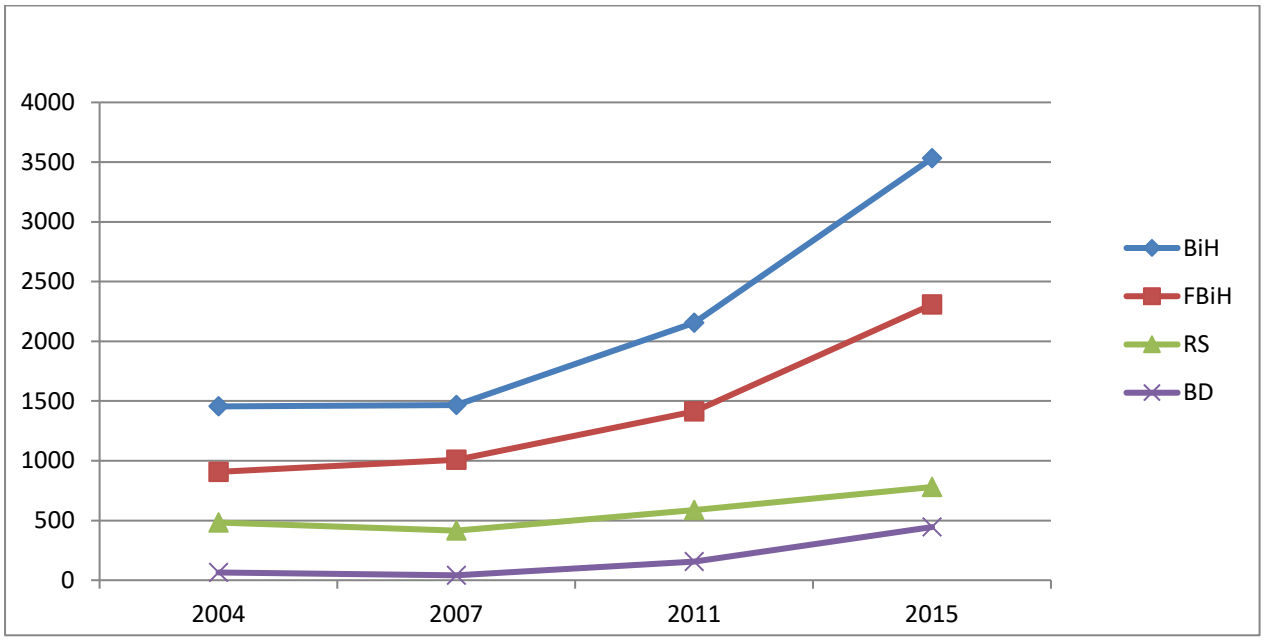
There is a huge increase in the non-response in the analyzed period. According the results of the one-way ANOVA test, there are significant differences in mean non-response rates between survey years ($F=320,752$; $p=0,000<0,05$). The same pattern has been observed at all three subnational levels with stronger increase of non-response in Brcko district BiH in the last two surveys (FBiH: $F=206,386$; $p=0,000$; RS: $F=30,560$; $p=0,000$; BD: $F=135,092$; $p=0,000$). The Tukey post hoc tests shown that there are significant differences in non-response rates between every pair of survey years in Brcko district BiH, while at the state and entity levels these differences are not significant only between 2004 and (BiH: $p=0,285$; FBiH: $p=0,062$; RS: $p=0,941$) and in Republika Srpska between 2011 and 2015 (RS: $p=0,063$)..

There are two types of household non-responses: incomplete interviews and no interviews. In the analyzed period, the number of households in both types has been increased, as shown on following graphs:

Graph 1. Number of households with incomplete interviews by geographical areas, BiH HBS 2004, 2007, 2011 and 2015



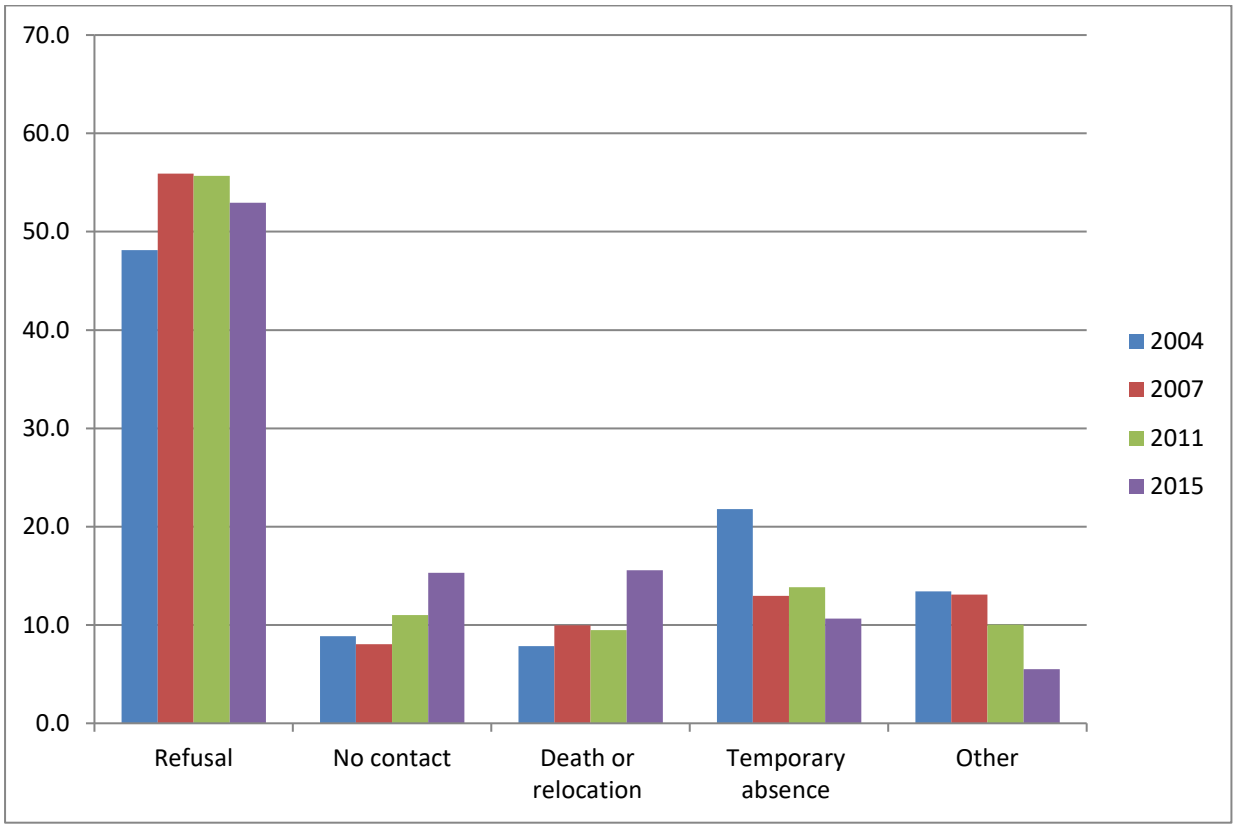
Graph 2. Number of households with no interviews by geographical areas, BiH HBS 2004, 2007, 2011 and 2015



Since the achieved sample sizes were similar in all years (about 7500 HHS), the hugest increase of non respondent households was in Federation BiH.

The most difficult cases of the non-response are those related to not interviewed households. Five reasons for no interviews have been monitored: refusal, no contact, death or relocation of all household members, temporary absence of all household members and other reasons. We analyzed the average rates of such cases in survey years. Graph 3 shows that the dominant reason for no interview was initial refusal for participation in the household budget surveys.

Graph 3. Structure of households with no interview by reasons of the non-response, BiH HBS 2004, 2007, 2011 and 2015



The rates of initial refusals are statistically significant between survey years (F=8,295; p=0,000) and the Tukey post hoc test indicated significant increases in proportion of refusal households in last three years compared with the first survey run. Differences between last three years are not statistically significant. The similar pattern has been observed at entity/district levels.

The indicators of non-response within the household budget surveys in Bosnia and Herzegovina in the analyzed period have clearly shown the significant increase of this non-sampling error, which can contribute to the increase the total survey error and damage the total quality of the surveys. Since 2004, when the survey was first time launched, no major improvements were made in order to decrease non-response. Surveys were only repeated in more or less the same scope and the same quality of the survey processes. Additionally, survey instruments were in 2011 and 2015 extended by ad hoc modules on social inclusion, which increased the respondent`s burden. It is one of factors than can explain the increase of number of refusals, although it is not visible in the graph.

5. Methods for improving response in Household Budget Surveys in Bosnia and Herzegovina

Although the non-responses in Bosnian household budget surveys are still smaller than in the majority of European countries (Eurostat, p. 14), there are needs for the implementation of various methods aimed to increase the response. Some of the possible methods are considered in the following paragraphs.

The first method, which can significantly improve response rates in all household surveys in Bosnia and Herzegovina is the creation of new sampling frame. The Agency for Statistics of

Bosnia and Herzegovina is still using the old master sample from 2009, which is significantly exhausted and out-of-date. It contains data of about 81,000 households for the selection of HBS sample and it is a huge source of non-response, especially of non-contacts. The creation of new master sample on the basis of the 2013 Population Census is still very slow and it is a conditio sine qua non for launching future household surveys.

Within the HBS, data is still collected in a face-to-face interview (PAPI), which is a source of response burden and non-sampling errors (interviewer`s errors, measurement errors, etc.) and a cause of high data collection costs. The introduction of computer assisted data collection (CAPI) is planned for the final interview of HBS. New data collection method should contribute to decrease of non-sampling errors and to decrease of the length of data collection period.

Although there is a legal basis for launching household budget surveys and for the protection of individual data, there is still room for improvements. Efforts must go towards the improvements of the overall image of the Agency for Statistics of Bosnia and Herzegovina. They must be realized through significantly better communication strategy, which should be focused on presenting statistical results in public and anouncing the launching of the HBS just before its field-work. The improved quality of survey staff and an adequate survey budget are pre-conditions for the realization of these activities.

The question of respondent incentives was analyzed many times within household surveys in Bosnia and Herzegovina (LSMS, HBS, LFS, etc.). In the previous implementations of some of these surveys, incentives were used in order to improve response rates. They were every time non-monetary and prepaid incentives in a form of small gifts for households (pocket computer, chemical pen or copybook). But, a serious analysis of whether incentives increase response has never been made. It can be recommended for testing the use of incentives within HBS, since it is the most complex survey for respondents.

The only method of the postsurvey adjustments within household budget surveys were weighting and very complex imputations. Weighting was adjusted to unit non-responses in order to get population totals, while nereast neighbor imputations decreased item non-responses. Further improvements must be directed to the poststratification using auxiliary variables, such as sex, age and education. The pre-condition for use of the poststratification is to get demographic estimates on the basis of Census results and vital statistics, which is still in process.

6. Conclusion

The significant increase of total non-response within household budget surveys in Bosnia and Herzegovina was noted. The non-response is caused by various reasons, the most important of which are initial refusals to participation in the survey, caused by huge response burden. Additionally, imperfections of the existing sampling frame contribute to the increase of non-contacts. From the first launching of the household budget survey in Bosnia and Herzegovina, significant methodological improvements were made. Survey questionnaires were extended and improved, which, from the other side, increased response burden. But, regarding the field work, surveys were conducted in the same scope and without any improvements of the communication strategy with respondents. Consequently, there is a lot of room for improvements in the nereast future. Several methods aimed to decrease unit and item non-responses in HBS were proposed. The most important ones are related to the creation of new master sample, to the introduction of new data collection methods and to improvement of the image of statistical institutions. All these methods will have a cumulative positive effect to decrease of non-response and other non-sampling errors, which will lead to better precision of survey estimates and to reduction of potential bias.

References

Babbie E. (2007). *The Practice of Social Research*. 11th ed. Belmont, CA: Wadsworth

Biemer, P. (2010). *Total survey error. Design, implementation, and evaluation*. Public Opinion Quarterly, 74(5): 817–848.

Church A. (1993). *Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis*. Public Opinion Quarterly, March 1993, DOI: 10.1086/269355

Cornish J. (2002). *RESPONSE PROBLEMS IN SURVEYS: Improving response & minimising the load*. UNSD Regional Seminar on 'Good Practices in the Organisation and Management of Statistical Systems' for ASEAN countries, Yangon Myanmar, 11-13 December 2002

Curtin R., Presser S., Singer E. (2000). *The Effects of Response Rate Changes on the Index of Consumer Sentiment*. Public Opinion Quarterly; Winter 2000; 64, 4; ABI/INFORM Global

EUROSTAT. (2015). *Household Budget Survey-Wawe 2010-Quality Report*, DOC HBS/2015/01/EN

Keeter S., Miller C., Kohut A., Grover R. M., Presser S. (2000). *Consequences of Reducing Nonresponse in a National Telephone Survey*. Public Opinion Quarterly 64:125–148. by the American Association for Public Opinion Research

Groves R.M. (2006). *Nonresponse Rates and Nonresponse Bias in Household Surveys.* Public Opinion Quarterly, Vol. 70, No. 5, Special Issue 2006

O'Toole J., Sinclair M., Leder K. (2008). *Maximising response rates in household telephone surveys*. BMC Medical Research Methodology 2008, **8**:71

Rubin D.B. (1978). *Multiple Imputations in Sample Surveys-A Phenomenological Bayesian Approach to Nonresponse*. Educational Testing Service.

Singer, E., Van Hoewyk, J., Gebler, N., Raghunathan, T., and McGonagle, K. (1999). *The effect of incentives in interviewer-mediated surveys*. Journal of Official Statistics, 15(2): 217–230

Singleton R., Bruce S. (2005). *Approaches to Social Research*. 4th ed. New York: Oxford University Press

Toepoel V. & Schonlau M. (2017). *Dealing with nonresponse: Strategies to increase participation and methods for postsurvey adjustments*. Mathematical Population Studies, 24:2, 79-83, DOI: 10.1080/08898480.2017.1299988