Workshop on Labour Force Survey Methodology
Budapest, 16–17 May 2019
Workshop topic: Field work: Methods for decreasing non-response

Non-response Analysis in the Labour Force Surveys in Bosnia and Herzegovina

Prepared by Edin Šabanović[1]

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 in 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. non-contact, vacant dwelling, person temporarily absent, refusal, 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.

---

[1] Agency for Statistics of Bosnia and Herzegovina

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, which is the most important for the final result of the interview.

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 Labour Force Surveys in Bosnia and Herzegovina

The analysis of non-response within the labour force surveys in Bosnia and Herzegovina was performed on data for the period 2009-2018. 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 LFS, 2009-2018

Year	BiH	FBiH	RS	BD
2009	11,30	9,48	14,46	11,50
2010	11,14	9,34	13,16	14,47
2011	12,48	11,63	13,30	14,57
2012	13,55	12,50	14,61	16,09
2013	14,36	13,18	16,39	14,34
2014	16,28	16,90	15,76	14,42
2015	19,62	20,88	18,98	14,34
2016	23,09	23,48	19,38	32,95
2017	25,44	23,11	26,57	35,48
2018	27,47	23,62	31,64	34,93

There is a huge increase in non-response in the analyzed period. The difference in the average non-response rates in 2018 and 2009 is statisticaly significant ($z=29.72$, $p=0.000$). The same trend has been observed in all three administrative entities with atypical increase of non-response in Brcko district BiH in the last three years of the survey.

Four main reasons for the non-response have been monitored: vacant dwelling, temporarily absent people, refusal and other. We analyzed the number and average rates of such cases in the period 2009-2018. Graph 1 shows the huge increase in number of non-responded households by reasons of the non-response:

**Graph 1. Number of non-responded households by reasons of the non-response, BiH LFS 2009-2018**
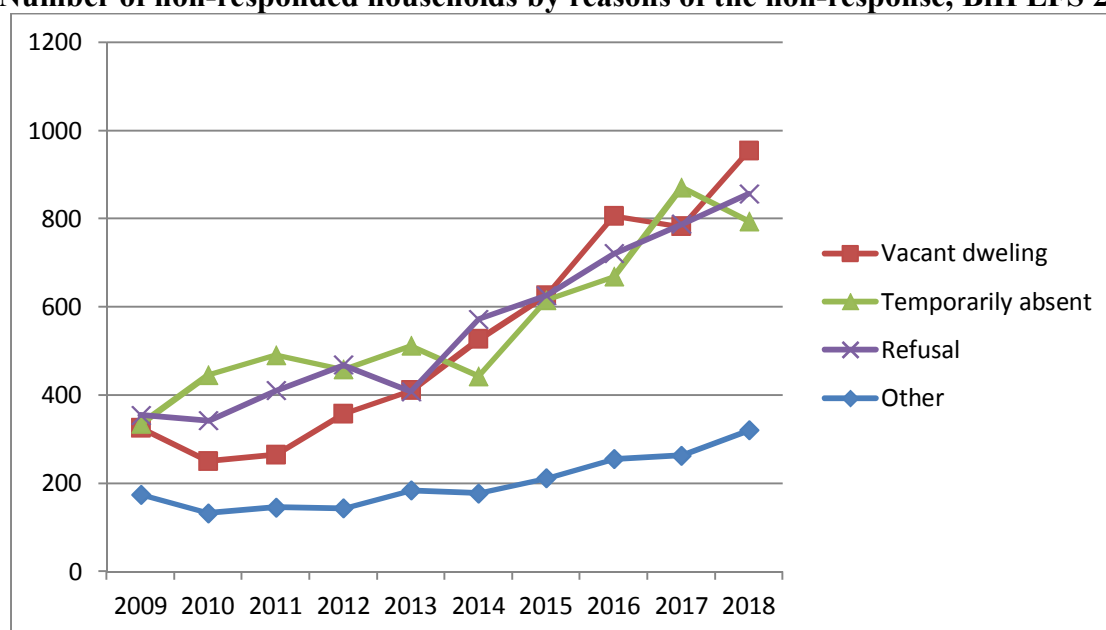
According the results of the one-way ANOVA test, there are significant differences in mean rates of vacant dwellings ($F=15.439$; $p=0.000$), mean rates of temporarily absent people ($F=13.940$; $p=0.000$) and mean rates of refusals ($F=2.717$; $p=0.004$). The trends of these mean rates are presented on the graph 2. Results of the Tukey post-hoc test shown that mean rates of refusals are more homogeneous in comaprison to mean rates of vacant dwellings and temporarily absent people. They are classified into two homogeneous subsets, compared with four and five sets, respectively.

The indicators of non-response within the labour force surveys in Bosnia and Herzegovina in the last ten years have clearly shown the significant increase of this non-sampling error, which contributes to the

increase the total survey error and damages the total quality of the surveys. Since 2006, when the survey was first time launched, no major improvements were made in order to decrease non-response. Surveys were only repeated each year in more or less the same scope and the same quality of the survey processes. Although the non-response in Bosnian labour force survey are still better that in the majority of European countries (Eurostat, p. 21-22), 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

Graph 2. Average rates of non-responded households by reasons of the non-response, BiH LFS 2009-2018
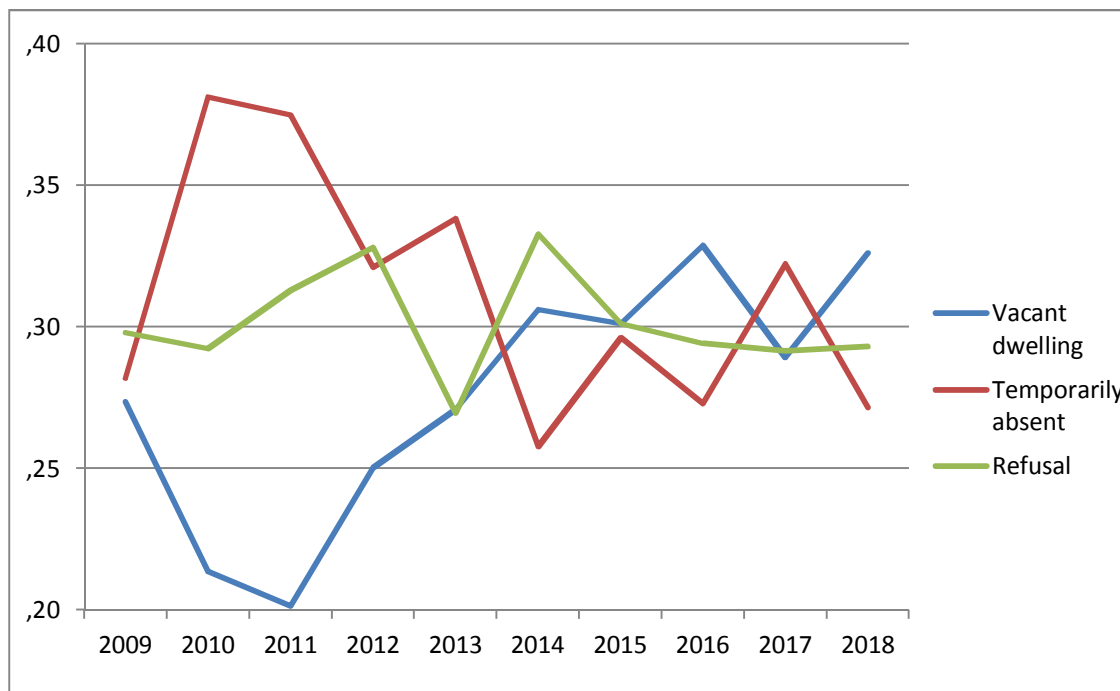
data of about 68,000 households for the selection of LFS sample and it is a huge source of non-response, especially of non-contacts. The introduction of the continuous LFS is conditioned by the creation of new master sample on the basis of the 2013 Population Census, whose realization is still very slow.

Within the LFS, data is still collected in a face-to-face interview (PAPI), which is a source of response burden and non-sampling errors (interviewers, measurement errors, etc.) and a cause of high data collection costs. The introduction of computer assisted data collection is planned for a first continuous labour force survey. New data collection will start with CAPI method and over time will swich to CATI method. 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 labour force 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 LFS 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 within a pilot continuous labour force survey.

The only method of the postsurvey adjustments within labour force surveys were weighting and very simple imputations. Weighting was adjusted to unit non-responses in order to get population totals, while mean or median imputations decreased item non-responses in a limited extent. 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.

5. Conclusion

The significant increase of total non-response within labour force surveys in Bosnia and Herzegovina was noted. The non-response is caused by various reasons, the most important of which are vacant dwellings, temporarily absent people and refusals. The imperfection of sampling frame contributes to increase of non-responses as one of most important non-sampling errors. From the first launching of the labour force survey in Bosnia and Herzegovina, no significant methodological improvements were made. Surveys were conducted in the same scope and the only adjustments were made in designing specific ad-hoc modules according the recommendations of Eurostat. Consequently, there is a lot of room for improvements in the nereast future. Several methods aimed to decrease unit non-responses in LFS were proposed. The most important ones are related to the creation of new master sample and for new data collection methods. The new master sample is conditio sine qua non for the implementation of such methods. It will allow the introduction od the continuous labour force survey on the basis of monthly data collection. Its introduction will allow the implementation of other methods for impriving survey responses.

6. References

1. Babbie E. (2007). *The Practice of Social Research*. 11th ed. Belmont, CA: Wadsworth
2. Biemer, P. (2010). *Total survey error. Design, implementation, and evaluation*. Public Opinion Quarterly, 74(5): 817–848.
3. 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
4. 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
5. 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
6. EUROSTAT. (2019). *Quality report of the European Union Labour Force Survey 2017*, Statistical reports, 2019 edition
7. 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
8. Groves R.M. (2006). *Nonresponse Rates and Nonresponse Bias in Household Surveys*. Public Opinion Quarterly, Vol. 70, No. 5, Special Issue 2006
9. O'Toole J., Sinclair M., Leder K. (2008). *Maximising response rates in household telephone surveys*. BMC Medical Research Methodology 2008, **8**:71
10. Rubin D.B. (1978). *Multiple Imputations in Sample Surveys-A Phenomenological Bayesian Approach to Nonresponse*. Educational Testing Service.
11. 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
12. Singleton R., Bruce S. (2005). *Approaches to Social Research*. 4th ed. New York: Oxford University Press
13. 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