

Estimation of response propensities using the previous survey

Miika Honkala

Statistics Finland, e-mail: miika.honkala@stat.fi

Abstract

This paper studies how response propensities, estimated using the dataset of the previous survey, predict actual response rates. In this study, two consecutive datasets of same survey were available. Response propensities were estimated to the older dataset using logistic regression model. Then the propensities were imputed to the newer dataset. The result was that the imputed response propensities predicted response behavior quite well.

Keywords: response propensity, response rate

1 Introduction

Many surveys are carried out annually. The implementation of the surveys and response behavior remain quite similar in consecutive years. If a survey has conducted a number of times previously, it may be possible to determine different optimal designs for different subgroups on the basis of the past experience (Tourangeau et al. 2017). However, it is good to check several rounds if available and look forward how regular response rates are (Laaksonen 2016). Schouten et al. (2017) present adaptive survey design which offers several methods for data-driven tailoring of data collection.

One possibility to utilize a previous survey is to estimate response propensities before the data collection of the survey, using the dataset of the previous year. For the data collection, it may be beneficial to know estimated response propensities in advance. If response propensities are known, the data collection can be designed in a new way. The response propensities may be utilized to tailor the data collection.

This study was carried out using Statistics Finland's data. Two datasets were available: datasets of European Social Survey from rounds 7 (2014) and round 8 (2016). In this paper, ESS7 means the dataset of round 7 and ESS8 means the dataset of round 8. Response propensities were estimated to the ESS7 using logistic regression model. After this, the propensities were imputed from the ESS7 to the ESS8. An interesting question was how the imputed propensities predict actual response rates. Were individuals with low response propensities often non-respondents? On the other hand, were individuals with high response propensities often respondents?

2 European Social Survey

ESS is a cross-national survey that has been conducted across Europe since 2001. Its target population consists of all residents 15 years or older who are residents of the country within private households. The ESS is conducted every two years using face-to-face interviews. In Finland, the sample size of ESS was 3 400 in the rounds 7 and 8. Response rate was 62.7% in the round 7 and 57.7% in the round 8. ESS's websites include more information about the ESS.

3 Methods and results

The sizes of the datasets were 3400, including respondents and non-respondents. Both datasets contained a binary response indicator (1 = respondent, 0 = non-respondent) and a lot of register variables. A response propensity model was fitted to the ESS7. The model was a logistic regression model where the dependent variable was the response indicator. Explanatory variables of the model were selected from the register variables. In the final model, the explanatory variables were municipality group, gender and interaction age group x education. These variables had a statistically significant effect on response. All the explanatory variables were classified variables. The model was made using SAS program.

After modeling, the response propensities were imputed from the ESS7 to the ESS8. The imputation method was a donor-recipient method based on the explanatory variables of the model. A donor in the ESS7 and a recipient in the ESS8 had same characteristics (the same values in the variables municipality group, gender and interaction age group x education). For example, suppose that a donor person in the ESS7 had response propensity 0.53 and following characteristics: municipality group = 2 (semi-urban municipalities), gender = 1 (male) and age group x education = 5 (30-44 years, no final examination). If the ESS8 included people who had exactly same values in these variables, these people got imputed response propensity 0.53.

The ESS8 were divided into groups according to imputed response propensities. Actual response rates in the six response propensity groups in the ESS8 are shown in Figure 1.

Figure 1. Response rates in the ESS8.

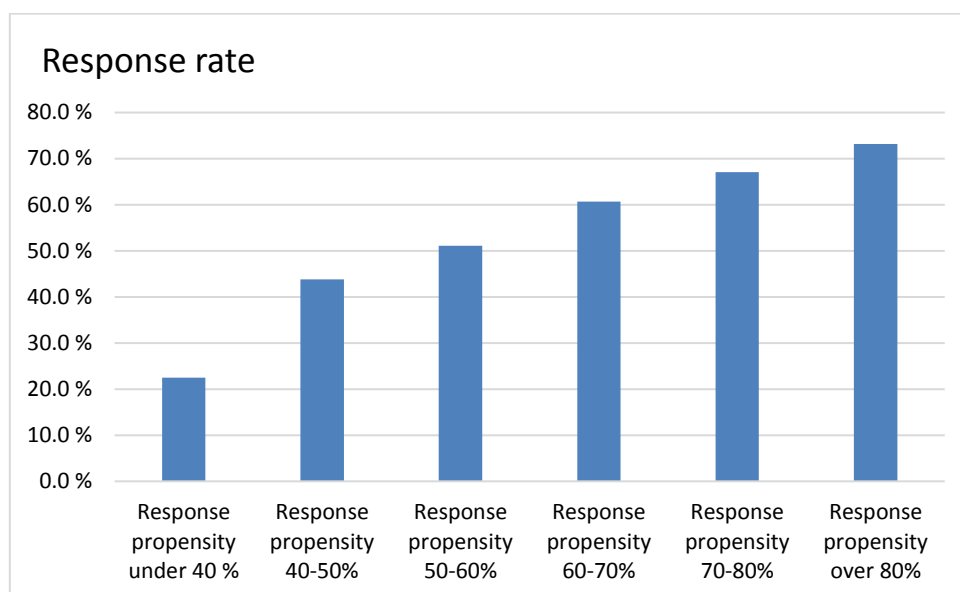


Figure 1 shows that imputed response propensities predict actual response rates quite well. The curve is rising, so there is a clear link between the imputed response propensities and the actual response rates. In a group where response propensities are under 40%, response rate is slightly more than 20%. In a group where response propensities are over 80%, response rate is more than 70%.

4 Conclusion

Using a good response propensity model, response behavior can be predicted before the data collection of the survey has begun, if the dataset of the previous survey is available. Predicting response rates is useful in surveys that are carried out annually, and where the sampling design and the implementation of the survey generally remains the same in consecutive years. When possible challenging respondents are known in advance, it is possible to plan the data collection in a new way and consider motivation letters or incentives for some respondents, for example. In telephone interview surveys, response propensities can determine the number of calls. The lower response propensity, the more contact attempts. It could be interesting to test this kind of responsive data collection which is based on the previous survey.

Utilizing a previous survey could also be beneficial when planning a new sampling design. Laaksonen (2016) presents a stratified sampling method which is based on the response rates of the previous survey. In that method, the strata are formed according to auxiliary variables. Sample sizes are bigger in groups where response rates have been low, and smaller in groups where response rates have been high in the previous survey. This kind of sampling method may lead to more representative set of respondents.

References

ESS's websites,

<http://www.europeansocialsurvey.org/> (referenced 25.5.2018)

Laaksonen, S. (2016). Anticipation of unit nonresponse or not in the sampling designing From the point of view of the European Social Survey (ESS). International Workshop on Household Survey Nonresponse, 2016, Oslo.

Schouten, B., Peytchev, A. and Wagner, J. (2017). *Adaptive survey design*. Chapman and Hall, New York.

Tourangeau, R., Brick, M., Lohr, S. and Li, J. (2017). Adaptive and Responsive Survey Designs: a Review and Assessment. *Journal of the Royal Statistical Society A* **180**, 203-223.