**Can we raise response rates by using multiple modes during data collection? Evidence from two mixed-mode designs tested before the 2018 Hungarian parliamentary election**

Adam Stefkovics[[1]](#footnote-1)

**Abstract**

This paper aims to assess the potential of mixed-mode surveys in handling nonresponse in political surveys. The study examines if response rates can be improved either by offering nonrespondents additional mode options or by initially contact respondents on their preferred, more feasible mode. Mixed-mode surveys are popular because they have the potential to improve response rates, decrease other error sources or lower costs. In the case of election studies, representativeness is a core issue, hence mixed-mode designs are especially attractive.

Two mixed-mode designs have been tested in March 2018, just before the Hungarian parliamentary election. One was a sequential design, where the main data collection was face-to-face and the nonrespondents have been offered an online or phone option. The other design was a concurrent, CATI-CAWI survey, where people above 55 were sampled from an online access panel, while people over 55 answered via phone interviews. The results of MM1 suggest that by offering alternative modes for the nonrespondents response rates can only be improved to a very limited extent at least in Hungary. Besides that, unexpectedly, MM2 had the best data quality even compared to a single-mode CATI.

Keywords: nonresponse, mixed-mode survey, election studies

1. **Introduction**

Decreasing response rates of surveys keep being probably the most challenging difficulty for empirical sociology all over the world. Indeed, this phenomenon has long ago encouraged methodologists to renew traditional methods and seek for innovation. One of the most popular attempts to raise response rates is the use of multiple modes of data collection. Contacting nonrespondents by an other mode or offering alternative modes both might decrease nonresponse related errors. There is also some evidence that responding on the preferred mode have a positive impact on data quality and the speed of responding (Olson, Smyth, and Wood 2012). Mixed-mode surveys also have the potential to decrease error sources other than nonresponse, or even lower costs. The core idea is to compensate the weakness of one mode with the strength of an other mode, which results in a decrease of the total survey error (De Leeuw and Hox 2011).

This paper aims to assess the potential of mixed-mode surveys in handling nonresponse in political surveys. In election studies ensuring representativeness is a core issue. Better coverage and higher response rates provide higher reliability, on the other hand mode effects might cause additional measurement error (lower validity), especially when sensitive questions are asked (De Leeuw 2005). The paper tries to look at the whole picture by comparing other attributes of the mixed-mode designs besides nonresponse, such as data quality or costs.

1. **Data and methods**

The data is from two mixed-mode surveys conducted in March 2018, just before the Hungarian parliamentary election. In MM1 and in MM2 we used two different mixed-mode strategies in order to improve response rates and the quality of our data. In MM1 we tried to convince nonrespondents by providing additional mode options, while in MM2 we initially contacted sub-populations on the mode we suspected they would prefer.[[2]](#footnote-2) MM1 was a sequential survey (N=800), where the main data collection was face-to-face (CAPI) and the non-respondents have been offered an online or phone interview option. We used a two-stage stratified sampling procedure and random walking method in choosing the final sample members. From those who refused to participate, interviewers asked for email addresses or phone numbers. Soon after, the same questionnaires have been sent out to these sample units via email or have been contacted via telephone. MM2 was a concurrent, CATI-CAWI mixed-mode survey (N=804), where the sample was split by age. People above 55 was surveyed online, people over 55 was surveyed by phone. This meant two independent sampling procedures. The online sample was drawn from a nonprobability online access panel counting 80.000 active members, who were recruited through different channels, both online and offline. The sample was stratified according to gender, age and settlement type. The CATI sample was drawn with a geographically stratified random sampling procedure. Both data of MM1 and MM2 was weighted using four basic demographic factors (gender, age, type of settlement, educational level) through an iterative procedure. Besides the two mixed-mode survey, data was available from a benchmark single mode CATI survey (N=501).

1. **Results**

**Sample compositions, coverage**

The unweighted marginal distributions for basic demographic variables show that the MM1 survey provided the closest estimates to the census values. All surveys have quite close estimates for gender and settlement type, but not with regard to age and education. People over 60 were overrepresented in the single-mode CATI, furthermore, all surveys attracted far more respondents with higher education, especially MM2.

The available sampling frames of the three surveys suggest that MM2 had the highest coverage error. That is partly due to the design itself, as people under 55 years old who do not use internet had a zero chance of being selected into our sampling frame. That is around 18 percent of the population under 55. Compared to this, MM1 and the single-mode CATI provide lower coverage error.

**Response and participation rates**

As expected, MM1 produced the highest response rate (32%), and the single-mode CATI the lowest (20%).[[3]](#footnote-3) In MM1 by providing an online and a telephone option we were only able to achieve a 3 percent raise in response rate. These sample members answered online, we were unable to involve nonrespondents with the telephone option. Unfortunately response rate for MM2 survey is not fully comparable to the others, as part of the sample was drawn from a non-probability online access panel (Loosveldt and Sonck 2008). In these surveys nonresponse bias can occur at two different stages: the willingness to participate in the panel and the actual participation in the survey (Bosnjak et al. 2013). For MM2 the CATI sample’s response rate was 27 percent, the completion rate for the CAWI part was 51 percent. The rates for CATI suggest that higher response rates can be achieved among older people by telephone.

**Data quality**

***Item-nonresponse***

Interestingly MM1 had the highest item-nonresponse rate altogether (mean=17,4%), and the CATI single-mode the lowest (mean=10,9%). In MM2 the CAWI part resulted in higher item-nonresponse rates, but only for the substantive (non-demographic) questions. No significant differences have been found regarding the length of the open-ended answers.

***Accuracy***

MM2 had the closest estimate (72,9%) to the voter turnout rate of the Hungarian general elections held in April 2018 (69,5%). MM1 underestimated (50,4%), the single-mode CATI survey overestimated (82,3%) the turnout rate.[[4]](#footnote-4) With regard to party preferences, MM2 was also the proved to be the most efficient in predicting the election results. MM2 had the lowest mean squared error (23,3), and estimated all of the listed parties’ vote share within sampling error.

***Satisficing, extreme answers, non-differentiation***

Acquiescence (the tendency to agree with questions regardless of its content) was used as an indicator for satisficing. Signs of acquiescence have been found in MM1 in a higher rate than in MM2.[[5]](#footnote-5) In MM2, online respondents were much less likely to satisfice. Similarly, non-differentiation was more likely among respondents of the MM1 survey, although no significant differences have been found between the surveys in choosing extreme response options.

**Length of data collection, costs**

With the programming of the questionnaires included MM1 took 25 days, MM2 took half as long (12 days). The single-mode CATI would have been around 15 days, if same sample sizes had been applied. The CAWI part of MM2 took 5 days only, and more than half of the respondents filled out the questionnaire on the day it was sent out. Seemingly surveying online is far faster than face-to-face or by phone. Nevertheless a “too fast” data collection like ours might also cause measurement error, as early respondents possibly differ from later respondents in their characteristics (Dillman 2014).

With regard to fieldwork costs, the cost of MM1 was 60 percent more than MM2’s and 39 percent more than the single mode CATI’s.[[6]](#footnote-6) Adding that using multiple modes might increase labour costs, which largely depends on the survey firm’s experience and quality.

1. **Discussion**

The purpose of this study was to examine the efficiency of using multiple modes in order to improve on response rates and other attributes of a survey in the case of measuring political attitudes in Hungary. Two designs have been tested: a sequential design with CAPI as the main mode and a concurrent CAWI-CATI survey. The results of MM1 suggest that by offering alternative modes for the nonrespondents response rates can only be improved to a very limited extent. On the other hand in MM2, which was a non-traditional design using nonprobability based channels, data quality proved to the best. It seems that the high coverage error of MM2 did not translate into significant coverage bias, nor did lower response rate cause high nonresponse bias. The most likely cause of this might be that basic demographic variables (e.g. age) has quite limited impact on political attitudes in Hungary as previous research suggest. The data of MM2 also outlines that experimenting with non-probability online access panels should not be ignored by scientists. MM2 offers a significant save in costs as well, but the “gap” between face-to-face and online surveying is not that huge as it is in most of the Western-European countries (Villar, A. and Fitzgerald, R. 2017). Although we were not able to achieve a breakthrough in raising response rates by using multiple modes, still the data shows that a well-designed mixed-mode survey can potentially decrease costs with keeping the survey’s representativeness and data quality on a high level.

1. **Questions for discussion**
2. Which are the more efficient sequential mixed-mode design?
3. Are there feasible CAWI-first designs? How can we assess nonresponse bias in such designs?
4. What are your experiences with non-probability online access panels in terms of nonresponse?
5. **References**

Bosnjak, M., Haas, I., Galesic, M., Kaczmirek, L., Bandilla, W., & Couper, M. P. (2013). Sample composition discrepancies in different stages of a probability-based online panel. Field Methods, 25(4), 339-360.

De Leeuw, E. D. (2005). To mix or not to mix data collection modes in surveys. Journal of official statistics, 21(5), 233-255.

De Leeuw, E. D., & Hox, J. J. (2011). Internet surveys as part of a mixed-mode design. Social and behavioral research and the Internet, 45-76.

Loosveldt, G., & Sonck, N. (2008, June). An evaluation of the weighting procedures for an online access panel survey. In Survey Research Methods (Vol. 2, No. 2, pp. 93-105).

Olson, K., Smyth, J. D., & Wood, H. M. (2012). Does giving people their preferred survey mode actually increase survey participation rates? An experimental examination. Public Opinion Quarterly, 76(4), 611-635.

Villar, A., & Fitzgerald, R. (2017). Chapter 16 Using mixed modes in survey research: Evidence from six experiments in the ESS. Values and Identities in Europe: Evidence from the European Social Survey, 259-293.

1. PhD student, ELTE Faculty of Social Sciences, Sociology PhD Programme, Hungary [↑](#footnote-ref-1)
2. Previous research suggest that younger respondents are more likely to prefer online answering, while older respondents prefer phone interviews. [↑](#footnote-ref-2)
3. For all response rates I used AAPOR Response Rate 1. [↑](#footnote-ref-3)
4. If we look at the rates of those who said they will definitely vote in the elections. [↑](#footnote-ref-4)
5. No data from the single-mode CATI. [↑](#footnote-ref-5)
6. Counting with same sample sizes. [↑](#footnote-ref-6)