

Web Mode as Part of Mixed-Mode Surveys of the General Population: An Approach to the Evaluation of Costs and Errors

Nejc Berzelak¹, Vasja Vehovar² and Katja Lozar Manfreda³

Abstract

Lower data collection costs make web surveys a promising alternative to conventional face-to-face and telephone surveys. A transition to the new mode has already been widely initiated in commercial research, but web surveys remains limited in academic and official research projects that typically require probability samples and high response rates. Various design approaches for coping with the problems of sampling frames, incomplete Internet use, and nonresponse in web surveys have been proposed. Mixed-mode designs and incentives are two common strategies to reach Internet non-users and increase the response rates in web surveys. However, such survey designs can substantially increase the costs, the complexity of administration and the possibility of uncontrolled measurement effects. This paper presents and demonstrates an approach to the evaluation of various survey designs with simultaneous consideration of the errors and costs. It focuses on the designs involving the web mode and discusses their potential to replace traditional modes for probability surveys of the general population. The main idea of this approach is that part of the cost savings enabled by the web mode can be allocated to incentives and complementary survey modes to compensate for the Internet non-coverage and the higher nonresponse. The described approach is demonstrated in an experimental case study that compares the performance of mixed-mode designs with the web mode and prepaid cash incentive with that of an official survey conducted using the face-to-face and telephone modes. The results show that the mixed-mode designs with the web mode and incentives can greatly increase the response

¹ Nejc Berzelak, Faculty of Social Sciences, University of Ljubljana, Kardeljeva ploščad 5, 1000 Ljubljana; nejc.berzelak@fdv.uni-lj.si

² Vasja Vehovar, Faculty of Social Sciences, University of Ljubljana, Kardeljeva ploščad 5, 1000 Ljubljana; vasja.vehovar@fdv.uni-lj.si

³ Katja Lozar Manfreda, Faculty of Social Sciences, University of Ljubljana, Kardeljeva ploščad 5, 1000 Ljubljana; katja.lozar@fdv.uni-lj.si

rate, which even surpasses that of the conventional survey modes, but still offer substantial cost savings. However, the results also show that higher response rate does not necessary translate to higher data quality, especially when the main aim is to obtain estimates that are highly comparable with those of the reference survey.

1 Introduction

Declining survey response rates and increasing survey costs force researchers to use new modes and procedures for survey data collection. Web surveys are one of the most promising approaches, especially in terms of cost savings and increased measurement quality (e.g. reduced social desirability bias due to the absence of interviewers, less mistakes in navigation or smaller item nonresponse due to the computerisation of the questionnaire, more effort from respondents who can choose to answer the questionnaire at the time and pace of their convenience, larger respondents' motivation due to the possibilities of including multimedia, etc.; Callegaro et al. 2015). Therefore, ESOMAR's reports on their prevalence in commercial research (ESOMAR, 2014) are not surprising. However, the breakthrough of web surveys in academic and official fields remains limited. This situation is mainly due to the lack of adequate sampling frames, incomplete Internet access and use in the general population and even lower response rates compared with the traditional modes (Lozar Manfreda et al., 2008; Shih and Fan, 2008). These drawbacks are particularly critical in probability samples, which are commonly required for academic and official data collection.

The potentials of web surveys to reduce research costs and increase measurement quality strongly encourage researchers to find the solutions for their application to the general population and to any other study with high data quality requirements, such as official statistic surveys. In this study, we examine a combination of two such solutions: incentives and mixed-mode survey designs. Incentives have been shown to be an effective means for increasing the overall response rates in traditional survey modes (Singer and Ye, 2013) and in web survey (Görizt, 2006). Mixed-mode designs aim at compensating for the weaknesses of each individual mode by concurrently or sequentially combining different modes within a single survey project (de Leeuw, 2005). In the case of web surveys, they can be used to reach Internet non-users and may also stimulate the response of Internet users who do not want to participate online.

This study mainly aims to present and demonstrate the methodology for benchmarking and evaluating the performance of a web survey included in a mixed-mode design with incentives in terms of 1) obtaining the results comparable to those from traditional survey modes and 2) reducing the overall survey costs in comparison with traditional modes. We put a special emphasis on the issues of Internet non-coverage and nonresponse, which are the most typical and specific problems of web surveys.

We begin with an introductory overview of the use of mixed-mode approaches and incentives to deal with the problems of incomplete Internet use and

nonresponse in web surveys. In the second part, we explain a methodological approach for the evaluation of different survey designs in terms of errors and costs. We establish the criteria for the identification of the optimal design and apply them to the empirical data from a case study. Finally, we present and discuss the results of the case study by observing the differences in sample composition, substantial data and costs.

2 Background

The problem of incomplete Internet access and use is one of the greatest threats to inference from web surveys of the general population. The differences in Internet use among countries are profound. In the European Union (EU), the proportion of Internet users ranges from 54% in Romania to 96% in Denmark (Eurostat, 2015). At the world level, according to Internet World Stats (2014), the proportion of Internet users in 2013 and 2014 was below 30% in Africa and almost 88% in North America. Only a few countries have Internet coverage above 90%, and in Europe these countries are Switzerland, the United Kingdom, Sweden, Finland, the Netherlands, Luxembourg, Norway, Denmark and Iceland (Eurostat, 2015).

The bias due to Internet non-coverage occurs if Internet non-users in the target population differ from Internet users in the characteristics measured by survey questions. Internet non-users are typically older, less educated, live in lower-income households, work in manual labour or are unemployed (Eurostat, 2015). The differences between users and non-users can contribute to the Internet non-coverage bias when a web survey is applied to the general population (Couper et al., 2007; Dever et al., 2008; Rookey et al., 2008). Although attempts to infer to the general population using post-survey adjustments have been made, their performance is often questionable and inconsistent across different variables and surveys (Lee and Valliant, 2009; Loosveldt and Sonck, 2008; Schonlau et al., 2009).

Low response rates are another prominent problem of web surveys. Whereas a persistent trend of declining response rates is observed in all survey modes (de Leeuw and de Heer, 2002), the situation is even worse in web surveys, which have been found to commonly produce lower response rates than the other modes (Lozar Manfreda et al., 2008; Shih and Fan, 2008).

Both Internet non-coverage and nonresponse problems can critically compromise data quality in web surveys by failing to reach specific parts of the target population. In consideration of the advantages of online data collection, particularly reduced costs and speed, significant efforts are unsurprisingly being devoted to the development of designs for a more efficient implementation of web surveys on the general population.

The two most plausible solutions for these problems are incentives to stimulate participation and mixed-mode survey designs that use an alternative mode to survey sampled individuals not reached by the web mode. An alternative approach

of online probability-based panels, which can also provide Internet access to households without the Internet, is clearly expensive. Consequently, only a few such panels currently exist, including LISS panel in the Netherlands, GESIS and GIP panels in Germany, ELIPSS in France, Social Science Research Institute panel in Iceland, and GfK-Knowledge Network panel in the United States.

Therefore, we limit our discussion to the more generally feasible mixed-mode survey designs with incentives in this study. Although past studies have already devoted attention to these topics also in the context of web surveys, the elaboration of cost-related factors is severely lacking. Thus, we first present the main general points related to mixed-mode designs involving web surveys and incentives, and then elaborate and demonstrate an approach for the evaluation of costs and errors in such designs.

2.1 Mixed-mode survey designs

Although mixed-mode survey designs existed well before the emergence of web surveys, the specifics of the new mode strongly facilitated their development and use. The main rationale for using mixed-mode designs involving the web mode is to exploit the cost advantages of web surveys and at the same time overcome their main problems, particularly non-coverage and nonresponse.

Modes can be combined in various ways at all stages of the survey project: at the initial contact stage, during data collection (response) stage or during the follow-up of non-respondents (de Leeuw, 2005). Many mode combinations are possible because the contact and data collection stages of web surveys are explicitly separated (Vehovar and Lozar Manfreda, 2008). The three most common combinations are as follows:

1. Telephone, mail or even **SMS invitation to a web survey** to overcome the problem of lacking contact information in the sampling frames and to stimulate the response rates by increasing the legitimacy of the request (e.g. Bosnjak et al., 2008; Messer and Dillman, 2011; Porter and Whitcomb, 2007).
2. **Concurrent use of web and other data collection modes** for different sample members at the same stage of the data collection to overcome the Internet non-coverage problem and to increase response. The most appropriate mode for each respondent can be selected in advance by the researcher (Blyth 2008; Rookey et al., 2008) or offered as a choice to the respondent (e.g. Medway and Fulton 2012; Parackal, 2003; Vehovar et al., 2001). However, although the latter option may seem respondent-friendly and advantageous, there is a strong evidence that offering a choice often decreases rather than increases the response rate (Dillman et al., 2014; Medway and Fulton 2012). This approach is therefore generally not recommended.

3. **Sequential use of different data collection modes** to increase the response rates and to reach Internet non-users. The most common approach is to begin with the web mode and then follow up the web non-respondents using one of the traditional survey modes (e.g. Dillman et al., 2014; Greene et al., 2008; McMorris et al., 2009). Alternatively, the web can be used to follow up non-respondents that used the other modes (e.g. Greene et al., 2008; McMorris et al., 2009).

Mixed-mode designs have specific problems. First, their administration is usually significantly more complex than that of single-mode surveys. A well-established sample monitoring system is essential to assure a proper assignment and transition of sampled individuals to different modes. The additional workload and required resources also have a direct effect on the costs of a survey project.

Second, the different modes may affect the comparability of the data obtained. Each mode has specific characteristics that may influence the answers provided by respondents. For example, compared with telephone and face-to-face surveys, the web mode is self-administered rather than interviewer-administered, the questions are presented visually rather than aurally, and the responses are provided by electronic input rather than orally. To what degree these factors cause mode effects and the related between-mode differences remains to be thoroughly investigated, as the findings of methodological studies are currently largely inconsistent. Some of these studies have found no differences among the modes (Knapp and Kirk 2003; Revilla and Saris 2010), whereas others have reported mode effects in different data quality indicators. Examples of mode effects include between-mode differences in the response order effects (e.g. Galesic et al., 2008; Malhotra 2008), response length to open-ended questions (Kwak & Radler, 2002), non-substantive answers (Bech and Kristensen, 2009), item nonresponse (Heerwegh and Loosveldt, 2008; Smyth et al., 2008) and non-differentiation (Tourangeau, 2004; Fricker et al., 2005). The most consistently observed effects are related to social desirability, in which the web mode almost universally produces less socially desirable answers than the interviewer-administered modes (e.g. Lozar Manfreda and Vehovar, 2002; Kreuter et al., 2008; Tourangeau et al., 2013).

Empirically investigating the mode effects is a demanding task that requires a researcher to disassemble error sources into individual components (sampling, non-coverage, nonresponse and measurement errors) and to identify which of these components are affected by the mode. This process is usually only possible with complex and well-controlled experimental designs. However, when the main objective is to compare the overall performance of different survey designs, all error sources of each individual mode can be taken into account without the separate identification of each component. We use the latter approach in this paper, as our goal is not to analyse different error sources but to observe the overall comparability of the obtained estimates among survey designs.

2.2 Incentives

Based on the evidence from research on traditional survey modes (Church, 1993; Singer and Ye, 2013), several studies have expected and confirmed that incentives can also improve the response rates in web surveys (Görizt, 2006; Parsons and Manierre, 2014). However, the efficiency of the incentives strongly depends on their type (monetary or non-monetary, pre-paid, post-paid, or lottery), value, survey sponsor (commercial or non-commercial), sample type (panel or non-panel studies) and several other factors (see e.g. Bosnjak and Tuten, 2003; Downes-Le Guin et al., 2002; Görizt, 2006, 2008; Görizt et al., 2008).

Pre-paid incentives have proved to be more effective than post-paid (promised) incentives for different survey modes (e.g. Church, 1993; James and Bolstein, 1992). Their benefits are even higher in web surveys because they increase the legitimacy of the survey request that is often limited online. Furthermore, immediate delivery dispels respondents' doubts about whether the incentive will actually be delivered or not. A small-value pre-paid incentive can already significantly improve the response rate, with higher values bringing relatively small additional benefits (e.g. Downes-Le Guin et al., 2002; Villar, 2013). However, the use of pre-paid incentives in web surveys is mostly limited to list-based surveys, in which people are personally addressable prior to their participation in the survey.

Although post-paid incentives are usually less effective, Görizt (2006) suggested that they seem to work better in online rather than in offline studies. A possible reason is that such incentives are common online and that Internet users may have gotten used to them (Bosnjak and Tuten, 2003).

Despite the potential benefits to the response rate, incentives may increase the nonresponse bias by attracting specific sample members (e.g. people with lower income, with more free time, who are younger, who are more computer literate, etc.). Görizt (2008) showed that incentives had a greater effect on low-income members of an online panel. Altered sample compositions have also been reported also by some other authors (e.g. Parsons and Manierre, 2014; Sajti and Enander, 1999). Furthermore, incentives may stimulate some respondents to complete the questionnaire more than once. This scenario is most likely to happen if a web survey is not based on a list of sampled persons where the possibilities for effective access control are limited (Comley, 2000).

Incentives are usually expected to increase the commitment of respondents to the task of answering questions because they feel compensated for their effort. However, with post-paid incentives, some people may use non-optimal response strategies just to reach the end of the questionnaire and become eligible for the incentive. Little empirical evidence about the effects of incentives on data quality exists. Görizt et al. (2008) found generally high data quality without systematic differences in the item nonresponse, length of answers to open-ended questions, discrepancy of answers, completion time and response styles between respondents receiving and those not receiving the incentive.

Finally, the effect of incentives may depend on the whole context of a specific web survey. A meta-content analysis of a larger number of web surveys (Lozar Manfreda and Vehovar, 2002) showed that incentives decreased the drop-out rates but did not increase the proportion of respondents who began the survey participation. The latter was more influenced by other design features, such as pre-notification and content of the questions. The meta-analysis by Cook et al. (2000) found even lower response rates with incentives, although the authors attributed this outcome to the confounding between survey length and use of incentives, where disproportionately long or tedious surveys more often offered incentives.

2.3 Consideration of costs and errors

Most of the existing studies that compare web surveys with other survey modes have focused on the response rates (Dolničar et al., 2009; Lozar Manfreda et al., 2008; Shih and Fan, 2008). These comparisons are insufficient, as the response rate is the only one indicator of data quality used. Existing comparisons also rarely consider the costs of data collection, which can be crucial for an overall assessment of web survey performance. That is, the cost savings earned by using a web survey instead of a more expensive survey mode (e.g. face-to-face or telephone) can be invested into additional recruitment measures (e.g. incentives) to increase the response rates. Such measures can significantly improve the performance of the web mode.

The costs–error evaluation of different alternative survey designs requires the separate consideration of costs at all contact and data collection stages for each mode. Several means can be used to identify the optimum approach among competing survey designs by taking into account the errors and costs. Some strategies include finding the design with the lowest costs at the fixed error level, the design with the smallest error at fixed costs (available budget), or the design with the smallest product of costs and a selected error measure (Vehovar et al., 2010).

These calculations rely on certain assumptions regarding the errors that are expected to occur in the actual survey implementation. Informed decision making about these assumptions can be made on the basis of experience, similar surveys or a pilot study.

3 Case study comparing errors and costs: methodology

To explore the possibilities of surveying the general population using the web mode, we performed an experimental study as part of a survey on information communication technologies (ICT) in Slovenia in 2008. According to Eurostat, 77% of Slovene households had Internet access in 2014, although this percentage was around 60% at the time of the survey. This penetration rate is too low to make

a web survey feasible as a standalone mode for surveying the general population. To overcome the Internet non-coverage problem, we used a mixed-mode design. We also used incentives to address the problem of nonresponse, which was expected to be high.

3.1 Experimental design

The Survey on ICT Usage by Individuals and in Households is a Eurostat survey conducted in all EU member states on an annual basis. In Slovenia, this survey is conducted by the Statistical Office of the Republic of Slovenia (SORS). In April and May 2008, the SORS fielded the survey using the mixed-mode design with face-to-face and telephone modes, depending on the availability of telephone numbers (further referred to as the *SORS survey*). The sample of 2,504 Slovenian citizens, together with their home addresses, was obtained from the Central Register of Population (CRP). Parallel to the official SORS survey, we implemented several experimental mixed-mode designs based on the same questionnaire (Berzelak et al., 2008).

In this study, we focused on the experimental mixed-mode designs with a mail follow-up of web survey non-respondents and two incentive conditions. In June 2008, mail invitations to take part in the web survey were sent to 305 randomly sampled individuals from the CRP (further on “*web > mail*” design). The letters included questionnaire access instructions, a unique access code for each sampled person and the statement that non-respondents would receive a paper questionnaire in a follow-up letter. The follow-up was carried out 10 days later. Two weeks later, the second follow-up mail letter was sent to the remaining non-respondents. Access to the web questionnaire was available throughout the whole data collection period.

The sampled individuals were randomly assigned to two incentive treatments: 100 sampled members received a EUR 5 banknote as a pre-paid incentive in the first mail letter and the remaining 205 received no incentive.

The survey questionnaire consisted of 44 questions (approximately 125 items) covering the topics like household access to various ICTs, individual’s frequency of computer and Internet use, devices used to access the Internet, frequency of using various online services, experiences with online shopping, and background information about the target person.

3.2 Research questions

On the basis of the experimental data, we explored the following research questions:

Q1. Response rate: Can the “web > mail” mixed-mode design with pre-paid cash incentives produce a response rate comparable with that in the reference SORS survey?

Q2. Sample composition: How do the final sample compositions of the “web > mail” designs with and without incentives differ from the reference SORS data?

Q3. Difference in substantive results: What is the difference between the survey estimates of the “web > mail” designs with and without incentives compared to the reference SORS data?

Q4. Does **weighting** based on demographic variables eliminate any differences in estimates between the “web > mail” designs and the SORS survey?

Q5. Costs: What is the cost of the compared designs for equal target sample sizes?

Q6. Cost and error comparison: Which “web > mail” design (with or without incentives) performs better when considering the costs and errors simultaneously?

To investigate these research questions, we performed several comparisons between the experimental survey designs and the official SORS survey.

3.3 Estimation of the bias

Although the **response rates** are straightforward to calculate, they have been shown to be insufficient predictors of the nonresponse bias (Groves & Peytcheva, 2008). Thus, we also compared **sample composition** between the experimental designs and the SORS survey, considering the final sample composition and the sample composition gained prior to the mail follow-up, which includes only the Internet users.

To **compare substantive results**, we selected key survey variables associated with ICT use. We compared the unweighted and weighted estimates from the experimental designs with the weighted reference SORS survey. The weights for the experimental survey designs were calculated using the raking method by sex, age and educational structure of respondents, as it is usually also done by the SORS. As the differences in estimates can substantially depend on the content of the variables, we also calculated the average difference in substantive results across several variables.

To calculate the difference in estimates for each variable, we followed the definition of the bias as the difference between the expected value over all respondents and the true value of the estimated parameter (Groves et al., 2004):

$$Bias(\bar{y}) = E(\bar{y}) - \bar{Y} \quad (3.1)$$

The above definition of the bias considers all the error sources that may affect the differences between the compared modes, including Internet non-coverage, nonresponse and measurement errors. Although it does not enable the explicit separation of error sources, it suffices for the evaluation of the overall differences among the compared survey designs.

Clearly, the true value is rarely known in practice and is usually estimated using a reference (“gold standard”) survey. In such cases, the above equation can be extended to account for the variance in the reference survey estimates (Biemer, 2010). However, because our main goal is to assess to what degree the experimental survey designs can provide results comparable with the official survey, we regard the estimates from the SORS survey as true values. Therefore, we refer to the difference between the experimental designs and the SORS survey as the bias, and it is calculated using Equation (3.1).

As we want to estimate the bias across several survey items and compare them among the experimental designs, we calculated the standardised bias for each item. The standardised bias is defined as the absolute bias divided by the standard deviation of the item in the reference SORS survey:

$$St.bias(\bar{y}) = \frac{E(\bar{y}) - \bar{Y}}{SD(\bar{Y})} . \quad (3.2)$$

For the items measured on the nominal scale, the statistic of interest is the proportion, and the standardised bias for each item is calculated as follows:

$$St.bias(p) = \frac{p - \pi}{\pi(1 - \pi)} . \quad (3.3)$$

We calculated the average standardised bias across items in each experimental design for each data collection stage (before and after the mail follow-up) and for the unweighted and weighted data. This approach enables the evaluation of both the incentives and the mixed-mode follow-up. Note that we did not test the statistical significance of the observed differences (biases) among the modes, as the objective was not to make inference about the target population but to investigate the presence of the bias in the specific survey implementation compared with the reference survey.

3.4 Identification of the optimal design

The identification of the most optimal survey design requires a simultaneous evaluation of costs and errors. To estimate the overall costs of each design, we considered the fixed and variable costs at each contact and data collection stage. We then included them in the cost model outlined by Vehovar et al. (2010) to simulate the costs at various target sample sizes.

When looking for the most optimal survey design in terms of errors and costs, it is beneficial to consider both, the bias of the estimate and also the corresponding sampling variance $Var(\bar{y})$. Namely, the variance decreases with the sample size n . If the lower costs of a specific survey design enable the increase in the sample size, the sampling variance of an item of interest will be reduced. Following the *total survey error* principles, each item can be evaluated using the standard approach of the root mean squared error (RMSE) (Kish, 1965):

$$RMSE(\bar{y}) = \sqrt{Var(\bar{y}) + Bias(\bar{y})^2} . \quad (3.4)$$

The total survey error does not depend only on the difference between the true value and the expected value from the survey but also on the precision of the point estimates. Therefore, the potential higher bias of a cheaper survey mode may be outweighed by the lower sampling variance enabled by the larger sample size in some circumstances.

Similar to the bias, the RMSE can be aggregated across several survey items. The relative RMSE is obtained for each item by dividing the RMSE by the value obtained from the reference SORS survey, against which we compare the estimated parameters:

$$rRMSE(\bar{y}) = \frac{\sqrt{Var(\bar{y}) + Bias(\bar{y})^2}}{\bar{Y}} . \quad (3.5)$$

Thus, we were able to calculate the average rRMSE across the selected key survey items. To evaluate the costs–error performance of each design, we observed which of the two experimental designs would produce the smallest average rRMSE at the given fixed budget. This circumstance resembles the practical situation survey researchers often face when deciding on how to best spend the available budget to reach the highest possible data quality, which can be defined in terms of accuracy of estimates, comparability with another survey (as in our case) or many other quality-related criteria. To simulate the amount of the total error in the two “web > mail” designs, we assumed that the difference between the designs and the official SORS survey would remain unchanged, but the sampling variance would change because of the different sample sizes reachable by the specified budget.

4 Results

4.1 Response rates

Comparing the response rates in the experimental mixed-mode designs (“web > mail” designs) with those in the official SORS survey enables us to explore whether such designs can produce response rates comparable with those in the traditional modes. We expect follow-up stages in the mixed-mode design to increase the response rates by reaching the sampled Internet non-users and converting those who do not want to participate online for some reason. We expect additional increases in the response rates in the experimental group with the pre-paid incentive.

The unit response rates for the compared survey designs by individual stages of the survey project are presented in Figure 1. The response rates are calculated as the percentage of the survey respondents out of all the eligible sampled individuals.

The **mixed-mode design without incentives** produced a final response rate of 33%. At the first stage, when the respondents were only able to complete the online questionnaire, only 11% of the sampled individuals participated in the survey. The follow-up reminder with paper (mail) questionnaire increased the response rate by 13 percentage points, and the final reminder increased it by another 9 percentage points. It is possible that some of the respondents would have completed the web questionnaire later even without follow-up. However, this proportion is unlikely to be high, considering that only 21% of those who participated after the second and the third reminders opted for the web mode. These findings confirm the result of previous research that offering a follow-up in a different mode can help increase the response rates.

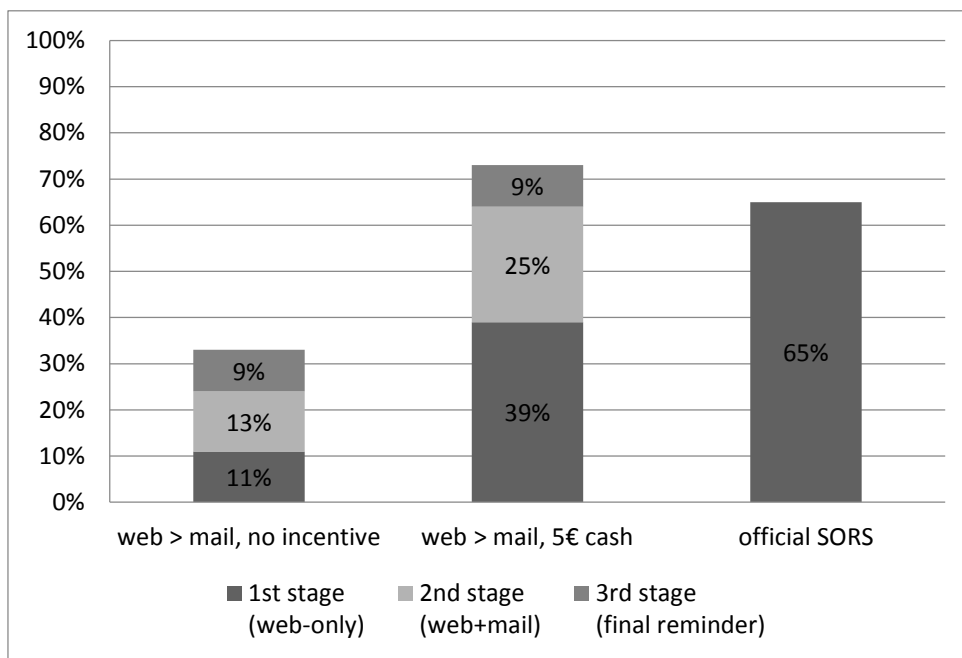


Figure 1: Unit response rates produced by the evaluated survey designs.

The observed **effect of incentives** on the response rates in the mixed-mode survey design is profound. Figure 1 shows that the EUR 5 pre-paid cash incentives already increased the response rate from 11% to 39% at the first stage, in which the respondents were able to complete the online questionnaire only. Although the effect somewhat diminished at the follow-up stages, it still produced the final response rate of 73%, which surpassed the response rate achieved by the official SORS survey. This finding confirms the large positive effect of pre-paid cash incentives on increasing the willingness of participation in both web and mail modes.

Considering only response rates, the results indicate that the combination of web and mail modes with appropriate incentives may be used to replace the traditional combination of face-to-face and telephone modes to survey the general population. However, as noted earlier, the response rate should not be regarded as

the only indicator of data quality because it is not necessarily a good predictor of the nonresponse bias.

4.2 Composition of the samples

We explored the biases in the sample composition of the experimental “web > mail” designs compared with the official SORS survey. The weighted official data were used as reference, as weighting can be expected to bring the compared socio-demographic estimates very close to the population characteristics.

Note that various error sources, including sampling frame non-coverage, Internet non-coverage, nonresponse and measurement errors, can contribute to a biased sample composition. As all evaluated survey designs use the same highly accurate sampling frame of the CRP, the errors due to sampling frame non-coverage are likely to be low. The same is true for the between-mode differences in measurement errors because the analysed questions are simple, factual, and not explicitly sensitive. Finally, the problem of Internet non-coverage is avoided in the “web > mail” designs by enabling Internet non-users to complete the paper questionnaire in the follow-up stages. Therefore, we expect that any bias in the final sample composition of the “web > mail” design compared with that in the official SORS survey is mostly due to the differences in the nonresponse error. However, the biases in the first (web-only) stage of the experimental designs are likely to be due to both nonresponse and Internet non-coverage.

The reduction of Internet non-coverage problem can be observed by comparing the estimates across the stages of the “web > mail” designs. At the first stage (web-only) of the design without incentives, the percentage of daily Internet users was 91%. After the mail option was offered (2nd stage) and final reminders sent (3rd stage), this proportion dropped to the final 62%. Although the proportion of Internet users did not reach the official estimate of 56% in the SORS survey, it substantially reduced the differences compared to the first (web-only) stage. While the web mode at the first stage of the design with incentives was more successful in reaching less frequent Internet users compared to the design without incentives (82% reported using the Internet almost every day), the decrease after the mail follow-up was much less pronounced and produced the final estimate of 73% of daily Internet users

Consistent with previous research, Table 1 shows that the **sample composition produced at the first stage** of the “web > mail” design without incentives is substantially affected by the typical characteristics of Internet users. Compared to the official survey, the sample over-represents men, younger and those with higher income. However, the education of respondents is similar to the official data. With the incentives, the respondents of the first stage of the “web > mail” design are similar to those of the SORS survey in terms of gender and age, but the bias for education is larger. This finding suggests that incentives especially attract female, older and higher-educated Internet users.

The **final sample composition** after the mail follow-ups in the “web > mail” design without incentives shows the increased proportion of women, older respondents and respondents with lower income. All of these characteristics are similar to those of the official SORS survey, although the average age remains substantially lower. This finding indicates that more respondents with characteristics consistent with Internet non-users were reached by the follow-ups, thus resulting in decreased non-coverage error. However, the combination of mail follow-up and web survey with pre-paid incentives produced mixed results. The mail follow-up made the sample more similar to that of the SORS sample in terms of age, income and education, but it increased the bias in the gender structure.

Table 1: Composition of samples produced by the evaluated survey designs.

	Web > mail (unweighted)				Official SORS (n=1052)
	No incentive		€5 cash		
	Stage 1: Web-only (n=22)	Final: Web+mail (n=67)	Stage 1: Web-only (n=39)	Final: Web+mail (n=73)	
% of men	64%	52%	49%	45%	52%
Mean age	24.7	29.0	27.9	28.3	32.5
Mean household size	3.8	3.8	3.8	3.8	3.8
Median household month. income (EUR)	2,000	1,500	2,000	1,700	1,500
% with more than secondary education	20%	21%	28%	24%	19%

The findings suggest that although the web survey with pre-paid cash incentives combined with the mail questionnaire significantly increased the response rates, it did not produce the sample composition similar to that in the SORS survey. The biases were especially large for gender, age and education. Interestingly, despite the profoundly lower response rates, the “web > mail” design without incentives produced the basic characteristics of respondents more similar to the official data.

4.3 Biases in substantive items

The socio-demographic composition of a sample can be largely corrected by weighting if the relevant information is available from reliable official sources, as in our case. To gain further insights into the performance of the compared experimental designs, we analysed the biases in substantive variables related to the survey topic using unweighted and weighted data. Again, bias is defined as the difference in estimates between the experimental designs and the official data, as

the comparability of estimates is considered to be the most important data quality criterion in this case.

In the unweighted data, the differences between the “web > mail” designs and the SORS data are higher in the “web > mail” design with incentives than in that without incentives for six out of the seven selected key survey variables (Table 2). The opposite situation is observed only for the question about online shopping. In both experimental designs, the biases are most profound in the estimates related to computer and Internet use. The situation does not improve much after weighting by sex, age and education. Although the differences between the “web > mail” designs and the official survey are reduced for most variables, the improvement is relatively limited. For most variables the biases remain higher in the design with incentives.

Table 2: Biases as the differences in estimates between the experimental designs and the official data for selected variables before and after weighting.

	Web > mail (final)				Official SORS (n=1052)
	No incentive		€5 cash		
	Unweighted	Weighted	Unweighted	Weighted	
Mean number of ICT devices in household ^a)	4.13 (-0.08)	4.15 (-0.06)	4.47 (+0.26)	4.50 (+0.29)	4.21
% living in household with Internet access	86% (+4 pp)	87% (+5 pp)	93% (+11 pp)	91% (+9 pp)	82%
% living in household with broadband Internet access	70% (-1 pp)	70% (-1 pp)	73% (+2 pp)	69% (-2 pp)	71%
% using computer every day or almost every day	76% (+15 pp)	72% (+11 pp)	83% (+22 pp)	80% (+19 pp)	61%
% using internet every day or almost every day	62% (+6 pp)	56% (+0 pp)	76% (+20 pp)	73% (+17 pp)	56%
% using mobile telephone	100% (+3 pp)	100% (+3 pp)	100% (+3 pp)	100% (+3 pp)	97%
% shopping over the Internet in the last 3 months	28% (+13 pp)	26% (+11 pp)	24% (+9 pp)	22% (+7 pp)	15%

^a The possession of the following devices was counted: TV, landline telephone, mobile telephone, desktop computer, laptop computer, personal digital assistant and gaming console.

Next, we calculated the average standardised bias in the estimates of the proportions in the “web > mail” designs against the official SORS survey for all 117 categorical variables in the questionnaire (Table 3). We did not include the comparison of mean estimates because only six items were measured at the scale level and even those were related more to background information than substantive

topics. The included variables thus take into account all key aspects of ICT use covered by the questionnaire (frequency of computer and Internet use, devices, and online services).

The unweighted data after the first web-only stage show that the average difference between the “web > mail” designs and the official survey is lower in the design with incentives (1.36) than in the design without incentives (1.51). However, the opposite is observed after the mail follow-up stages. Although offering the mail option decreased the average difference in both “web > mail” designs, the improvement was higher in the design without incentives (0.95 vs. 1.04). The weights further slightly decreased the average differences, but the design with incentives again performed worse. Note that the weights were only calculated for the final data. This is primarily due to small sample sizes after the first stage, which limited the possibilities to apply weighting procedures comparable to those used by the statistical office.

Table 3: Average standardised biases in the estimates of proportions as the differences between the experimental designs and the official data.

	Web > mail			
	No incentive		€5 cash	
	Stage 1: Web-only (n=22)	Final: Web+mail (n=67)	Stage 1: Web-only (n=39)	Final: Web+mail (n=73)
Average standardised difference for proportions (unweighted data)	1.51	0.95	1.36	1.04
Average standardised difference for proportions (weighted data)	n/a	0.87	n/a	0.99

We can conclude that monetary pre-paid incentives can significantly increase participation in a mixed-mode design involving the web mode. The response rates in our case even surpassed those using the traditional telephone and face-to-face combination without incentives. However, the incentives strongly influenced the self-selection of specific respondents and increased the bias, calculated as the difference in estimates against the official data. This finding suggests that the mail follow-up is more effective in reducing the difference between the experimental survey designs and the official survey than incentives.

4.4 Evaluation of costs and errors

As we stressed above, the consideration of research costs is crucial to understand the performance of a specific survey design. This is especially true for web surveys, in which cost savings may enable larger sample sizes and implementation of additional measures to increase the data quality.

To illustrate the cost advantages and to perform further comparisons, we calculated the costs of obtaining a hypothetical sample size of 1,000 respondents for each evaluated survey design, taking into account the response rates obtained in the presented case study (Figure 1). Table 4 shows that the two “web > mail” designs are substantially cheaper than the combined face-to-face and telephone survey. This is mainly due to the elimination of the interviewer-related costs. Comparing both “web > mail” designs, the costs are clearly substantially higher in the design with the pre-paid incentives.

Table 4: Costs of data collection for a simulated final sample size of n=1.000 in the compared survey designs.

	Web > mail						Official SORS
	No incentive			€5 cash			
	Stage 1	Stages 1 & 2	Stages 1, 2 & 3	Stage 1	Stage 1 & 2	Stages 1, 2 & 3	
Cost for the final sample size n = 1000	€1,835	€5,418	€6,857	€8,206	€9,919	€10,374	€25,885

To demonstrate the approach for identifying the more optimal design, we searched for the design with the lowest average relative RMSE for a given fixed budget. Note that the RMSE calculation does not apply to the official SORS survey, used as reference with which the experimental survey designs are compared. As stated above, our study explores the most optimal alternative design to provide the results as comparable to the existing SORS survey as possible, but with potentially lower costs. The key data quality indicator is in this case comparability rather than accuracy. Correspondingly, the estimates from the SORS survey are considered to be “true values” against which the alternative designs are compared, making the RMSE calculation for the SORS survey inapplicable.

Table 5 presents the sample sizes that can be reached at a fixed budget of EUR 10,000 for the three compared survey designs. Assuming that the response rates are the same as in our empirical case, larger initial and final sample sizes can be achieved by the “web > mail” survey designs than by the face-to-face/telephone survey.

As noted earlier, the simulation of the total error in the two “web > mail” designs at the given fixed budget of EUR 10,000 assumes that the biases remain equal as in the empirical case, but considers the changes in the sampling variance because of the different sample sizes. Table 5 shows that the average relative RMSE across the seven key variables at a fixed budget is lower for the “web > mail” design without incentives than that with incentives. In consideration of our previous findings, this result is not surprising as the design with incentives produced greater differences with respect to the reference survey (Tables 1–3) and is substantially more expensive to implement (Table 4).

Table 5: Comparison of the mean RMSE for the selected seven key variables at a fixed budget of EUR 10,000.

Survey design	Initial sample size	Final response rate	Final sample size	Mean rRMSE	rRMSE × costs
Web > mail, no incentive	4,610	33%	1507	0.8272	8272.74
Web > mail, €5 cash	1,318	74%	962	0.9077	9078.47
Official SORS	580	65 %	372	n/a	n/a

This demonstration adds another perspective to the above findings that a higher response rate influenced by incentives does not provide more comparable data to the official face-to-face/telephone survey. Again, this finding emphasises the inconsistent relation between response rates and the nonresponse bias. That is, higher response rates do not necessarily translate to lower nonresponse bias.

5 Summary and conclusions

The answer to the question of whether web surveys can be a viable alternative to traditional modes for surveys of the general population is not simple and straightforward. It requires a consideration of several issues and a careful planning of measures, including the evaluation of survey design elements to deal with Internet non-coverage and nonresponse problems. Furthermore, selecting and applying appropriate criteria is necessary to compare the costs and errors among different survey designs to find the optimal solution for specific research needs. This is particularly relevant to web surveys because of their substantial cost-saving potential. Of course, the magnitude of errors, costs of survey design implementation, and definition of the optimal ratio between costs and errors strongly depend on a specific survey project. This study mainly aimed to discuss an approach for the simultaneous evaluation of costs and errors to assess various survey design possibilities.

We discussed mixed-mode designs involving the web mode and incentives as two common measures to treat the Internet non-coverage and nonresponse problems of web surveys. In the survey practice, a web survey is usually considered as a complimentary or a replacement mode for an existing face-to-face or telephone survey. Comparability of data between the two modes and sufficient cost savings to justify the change in survey design are usually the most important evaluation criteria in such situations. As demonstrated, the simultaneous observation of response rates, biases (in our case defined in terms of the difference against the official survey), RMSE across several key variables and costs can produce a much higher informative value about the design's performance than response rates as the sole indicator of the data quality.

In our case study, the provision of EUR 5 cash incentives had a tremendous effect on the response rate in the design with a mail follow-up to the web survey. It increased the survey participation by 40 percentage points compared to the design without incentives. The response rate even surpassed that of the official survey that combines the face-to-face and telephone modes.

However, we identified several differences in the substantive results between the mixed-mode design with the web mode and the official face-to-face/telephone survey. Although weighting generally decreased the differences between the modes, they remained substantial in some variables. Furthermore, despite the substantially higher response rate, the design with incentives led to higher differences in estimates. This finding again confirmed that response rates are not necessarily good predictors of the nonresponse bias (Groves & Peytcheva, 2008).

We also demonstrated that the web mode could substantially decrease survey costs unlike a combined face-to-face and telephone survey. This finding remains true even when the web mode is included as part of the mixed-mode design rather than as a standalone mode and when pre-paid incentives are provided to all sampled persons. However, the reduced sampling variance, which resulted from the increased sample size obtained due to lower costs, was not sufficient to substantially reduce the total survey error, which was defined in terms of the comparability of estimates with those of the official survey. We showed that the relative bias was greater in the case of incentives, which strongly influenced the self-selection of specific respondents. Thus, the total survey error even increased despite the higher response rate. Incentives should therefore not be regarded as a universal tool for data quality improvements. It is important to critically evaluate their appropriateness in the context of a specific survey project and assess their performance as part of the survey pretesting.

In what follows, we present the limitations of the presented case study and the application of the proposed methodology for the evaluation of costs and errors. The sample sizes in the two experimental mixed-mode designs were small, potentially leading to unstable estimates. Furthermore, our calculations of the biases assumed that the official survey provided the true values of the estimated parameters. This issue is not necessarily a limitation by itself, as the comparability between modes is sometimes even a more important aspect of data quality than accuracy. This condition is especially true when the aim is to prevent breaks in time series, as is the case with many longitudinal surveys. Nevertheless, determining which design provides the estimates closer to the true population value would be certainly beneficial. Moreover, the evaluated designs presented only a small subset of the possible mixed-mode implementations. Alternative designs, such as simultaneous mode options or other mode combinations, could result in different relations between errors and costs.

It is also important to take into account that the study covers only one and in this context relatively specific survey topic, i.e. ICT usage. The survey topic is directly correlated with the Internet use as the key prerequisite to participate in a web survey, which was offered in the first stage of the investigated survey designs.

Furthermore, the survey questionnaire contained predominantly factual rather than attitudinal questions. Repeating the study with different survey topics would be informative as other topics may result in different biases caused by self-selection and mode effects.

Despite these limitations, the discussed methodology offers survey researchers a valuable tool for making informed decisions on the feasibility of various survey design possibilities. The proposed approach can provide important insights when a switch to alternative data collection modes is considered for an existing survey. The presented case study also confirms the efficiency of mixed-mode designs and incentives to improve the response rates in web surveys, but at the same time it cautions against the over-reliance on response rates in data quality assessment.

References

- [1] Bech, M. and Kristensen, M. B. (2009): Differential response rates in postal and web-based surveys among older respondents. *Survey Research methods*, **3**, 1-6.
- [2] Berzelak, N., Vehovar, V., and Lozar Manfreda, K. (2008): *Nonresponse in web surveys within the context of survey errors and survey costs*. Paper presented at the 2nd MESS Workshop, Zeist, The Netherlands.
- [3] Biemer, P. P. (2010): Total survey error: Design, implementation, and evaluation. *Public Opinion Quarterly*, **74**, 817-848.
- [4] Blyth, B. (2008): Mixed-mode: The only "fitness" regime? *International Journal of Market Research*, **50**, 241-266.
- [5] Bosnjak, M. and Tuten, T. L. (2003): Prepaid and promised incentives in web surveys: An experiment. *Social Science Computer Review*, **21**, 208-217.
- [6] Bosnjak, M., Neubarth, W., Couper, M. P., Bandilla, W., and Kaczmirek, L. (2008): Prenotification in web-based access panel surveys: The influence of mobile text messaging versus e-mail on response rates and sample composition. *Social Science Computer Review*, **26**, 213-223.
- [7] Callegaro, M., Lozar Manfreda, K., and Vehovar, V. (2015): *Web survey methodology*. Los Angeles [etc.]: Sage.
- [8] Church, A. H. (1993): Estimating the effect of incentives on mail survey response rates: A meta-analysis. *Public Opinion Quarterly*, **57**, 62-79.
- [9] Comley, P. (2000): *Pop-up surveys. What works, what doesn't work and what will work in the future*. Paper presented at the ESOMAR Worldwide Internet Conference Net Effects 3, Dublin, Ireland.
- [10] Converse, P. D., Wolfe, E. W., Xiaoting, H., and Oswald, F. L. (2008): Response rates for mixed-mode surveys using mail and e-mail/web. *American Journal of Evaluation*, **29**, 99-107.

- [11] Cook, C., Heath, F., and Thompson, R. (2000): A meta-analysis of response rates in web- or Internet-based surveys. *Educational & Psychological Measurement*, **60**, 821-836.
- [12] Couper, M. P., Kapteyn, A., Schonlau, M., and Winter, J. (2007): Noncoverage and nonresponse in an Internet survey. *Social Science Research*, **36**, 131-148.
- [13] de Leeuw, E. D. (2005): To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, **21**, 233–255.
- [14] de Leeuw, E. D. and de Heer, W. (2002): Trends in household survey nonresponse: A longitudinal and international comparison. In R. M. Groves, Dillman, D. A., Eltinge, J. L. and Little, R. J. A. (Eds.): *Survey Nonresponse*, 41-54. New York, NY: John Wiley & Sons.
- [15] Dever, J. A., Rafferty, A., and Valliant, R. (2008): Internet surveys: Can statistical adjustments eliminate coverage bias? *Survey Research Methods*, **2**, 47-62.
- [16] Dillman, D. A., Smyth, J. D., and Christian, L. M. (2014): *Internet, phone, mail and mixed-mode surveys: The tailored design method*. Hoboken, NJ: Wiley.
- [17] Dolničar, S., Laesser, C., and Matus, K. (2009): Online versus paper: Format effects in tourism surveys. *Journal of Travel Research*, **47**, 295-316.
- [18] Downes-Le Guin, T., P. Janowitz, R. Stone, and S. Khorram. (2002): Use of pre-incentives in an Internet survey. *Journal of Online Research*, **25**, 1-7.
- [19] ESOMAR. (2014): *Global market research 2014*. Amsterdam: ESOMAR.
- [20] Eurostat. (2015): *Individuals – internet use in the last three months* (Database). Luxembourg, LUX: Eurostat.
- [21] Fricker, S. S., Galesic, M., Tourangeau, R., and Yan, T. (2005): An experimental comparison of Web and telephone surveys. *Public Opinion Quarterly*, **69**, 370-92.
- [22] Galesic, M., Tourangeau, R., Couper, M. P., and Conrad, F. G. (2008): Eye-tracking data. New insights on response order effects and other cognitive shortcuts in survey responding. *Public Opinion Quarterly*, **72**, 892-913.
- [23] Göritz, A. (2006): Incentives in web studies: Methodological issues and a review. *International Journal of Internet Science*, **1**, 58-70.
- [24] Göritz, A. (2008): The long-term effect of material incentives on participation in online panels. *Field Methods*, **20**, 211-225.
- [25] Göritz, A., Wolff, H.-G., and Goldstein, D. G. (2008): Individual payments as a longer-term incentive in online panels. *Behavior Research Methods*, **40**, 1144-1149.
- [26] Greene, J., Speizer, H., and Wiitala, W. (2008): Telephone and web: Mixed-mode challenge. *Health Services Research*, **43**, 230-248.

- [27] Groves, R. M., Fowler, F. J. Jr., Couper, M. P., Lepkowski, J. M., Singer, E., and Tourangeau, R. (2004): *Survey methodology*. Hoboken, NJ: John Wiley & Sons.
- [28] Groves, R. M. and Peytcheva, E. (2008): The impact of nonresponse rates on nonresponse bias. *Public Opinion Quarterly*, **72**, 167-189.
- [29] Heerwegh, D. and Loosveldt, G. (2008): Face-to-face versus web surveying in a high-Internet-coverage population: Differences in response quality. *Public Opinion Quarterly*, **72**, 836-846.
- [30] Internet World Stats (2014): *Internet users in the world*. Retrieved 30 June 2014, from <http://www.internetworldstats.com/stats.htm>
- [31] James, J. M. and Bolstein, R. (1992): Large monetary incentives and their effect on mail survey response rates. *Public Opinion Quarterly*, **56**, 442-453.
- [32] Kish, L. (1965): *Survey sampling*. New York, NY: John Wiley & Sons.
- [33] Knapp, H. and Kirk, S. A. (2003). Using pencil and paper, Internet and touch-tone phones for self-administered surveys: does methodology matter? *Computers in Human Behavior*, *19*, 117–134. [http://doi.org/10.1016/S0747-5632\(02\)00008-0](http://doi.org/10.1016/S0747-5632(02)00008-0)
- [34] Kreuter, F., Presser, S., and Tourangeau, R. (2008): Social desirability bias in CATI, IVR, and web surveys. The effects of mode and question sensitivity. *Public Opinion Quarterly*, **72**, 847-865.
- [35] Kwak, N. and Radler, B. (2002). A comparison between mail and Web surveys: Response pattern, respondent profile, and data quality. *Journal of Official Statistics*, *18*, 257–73.
- [36] Lee, S. and Valliant, R. (2009): Estimation for volunteer panel web surveys using propensity score adjustment and calibration adjustment. *Sociological Methods & Research*, **37**, 319-343.
- [37] Loosveldt, G. and Sonck, N. (2008): An evaluation of the weighting procedures for an online access panel survey. *Survey Research Methods*, **2**, 93-105.
- [38] Lozar Manfreda, K. and Vehovar, V. (2002): Survey design features influencing response rates in web surveys. Paper presented at the ICIS 2002 The International Conference on Improving Surveys, Copenhagen, Denmark.
- [39] Lozar Manfreda, K., Bosnjak, M., Berzelak, J., Haas, I., and Vehovar, V. (2008): Web surveys versus other survey modes: A meta-analysis comparing response rates. *International Journal of Market Research*, **50**, 79-104.
- [40] Malhotra, N. (2008): Completion time and response order effects in web surveys. *Public Opinion Quarterly*, **72**, 914-934.
- [41] McMorris, B. J., Petrie, R. S., Catalano, R. F., Fleming, C. B., Haggerty, K. P., and Abbott, R. D. (2009): Use of web and in-person survey modes to gather data from young adults on sex and drug use. *Evaluation Review*, **33**, 138-158.

- [42] Medway, R. L. and Fulton, J. (2012): When more gets you less: A meta-analysis of the effect of concurrent web options on mail survey response rates. *Public Opinion Quarterly*, **76**, 733-746.
- [43] Messer, B. L. and Dillman, D. A. (2011): Surveying the General Public over the Internet Using Address-Based Sampling and Mail Contact Procedures. *Public Opinion Quarterly*, **75**, 429-457.
- [44] Parackal, M. (2003): Internet-based & mail survey: A hybrid probabilistic survey approach. Paper presented at the AusWeb 2003: The 9th Australian World Wide Web Conference, Goald Coast, Australia.
- [45] Parsons, N. L. and Manierre, M. J. (2014): Investigating the relationship among prepaid token incentives, response rates, and nonresponse bias in a web survey. *Field Methods*, **26**, 191-204.
- [46] Porter, S. R. and Whitcomb, M. E. (2007): Mixed-mode contacts in web surveys: Paper is not necessarily better. *Public Opinion Quarterly*, **71**, 635-648.
- [47] Rookey, B. D., Hanway, S., and Dillman, D. A. (2008): Does a probability-based household panel benefit from assignment to postal response as an alternative to Internet-only? *Public Opinion Quarterly*, **72**, 962-984.
- [48] Sajti, A. and Enander, J. (1999): Online survey of online customers, value-added market research through data collection on the Internet. *Proceedings of the ESOMAR World-Wide Internet Conference Net Effects*, 35-51, Amsterdam: ESOMAR.
- [49] Schonlau, M., van Soest, A., Kapteyn, A., and Couper, M. (2009): Selection bias in web surveys and the use of propensity scores. *Sociological Methods & Research*, **37**, 291-318.
- [50] Shih, T.-H. and Fan, X. (2008): Comparing response rates from web and mail surveys: A meta-analysis. *Field Methods*, **20**, 249-271.
- [51] Singer, E. and Ye, C. (2013): The use and effects of incentives in surveys. *The ANNALS of the American Academy of Political and Social Sciences*, 645, 112-141.
- [52] Smyth, J. D., Christian, L. M., and Dillman, D. A. (2008): Does “Yes or No” on the telephone mean the same as “Check-All-That-Apply” on the Web? *Public Opinion Quarterly*, **72**, 103-13.
- [53] Tourangeau, R. (2004): Experimental design considerations for testing and evaluating questionnaires. In S. Presser, J. M. Rothgeb, M. P. Couper, J. T. Lessler, E. Martin, J. Martin, & E. Singer (Eds.): *Methods for Testing and Evaluating Survey Questionnaires*, 209-224. Hoboken, New Jersey: John Wiley & Sons, Inc.
- [54] Tourangeau, R., Couper, M. P., and Conrad, F. C. (2013): “Up means good”: The effect of screen position on evaluative ratings in web surveys. *Public Opinion Quarterly*, **77**, 69-88.

- [55] Vehovar, V., and Lozar Manfreda, K. (2008): Overview: Online surveys. In N. G. Fielding, R. M. Lee & G. Blank (Eds.): *The handbook of online research methods*, 177-194. Thousand Oaks, CA: SAGE Publications.
- [56] Vehovar, V., Berzelak, N., and Lozar Manfreda, K. (2010): Mobile Phones in an Environment of Competing Survey Modes: Applying Metric for Evaluation of Costs and Errors. *Social Science Computer Review*, **28**, 303-318.
- [57] Vehovar, V., Lozar Manfreda, K., and Batagelj, Z. (2001): Sensitivity of electronic commerce measurement to the survey instrument. *International Journal of Electronic Commerce*, **6**, 31-52.
- [58] Villar, A. (2013): Feasibility of using web to survey at a sample of addresses: a UK ESS experiment. Paper presented at the Genpoweb, London, UK. Retrieved from <http://www.natcenweb.co.uk/genpopweb/documents/ESS-experiment-Ana-Villar.ppt>