# Surveys are like a box of chocolate – you never know what you’re going to get? Assessing the impact of a digital chocolate voucher on the Finnish ISSP2018 response rate and data quality

Vesa Virtanen, Statistics Finland, vesa.virtanen@stat.fi

Juhani Saari, Statistics Finland, juhani.saari@stat.fi

**Abstract**

Statistics Finland has been pursuing a series of experiments since 2017 for testing new incentives to increase response rates for household surveys. The aim has been to reduce the range of incentives and expenses, to develop digital and knowledge base incentives, increase web responses and reward early bird respondents. Following an unsatisfactory trial of a digital meal voucher targeted to less educated young respondents, Statistics Finland examined the effectiveness of a digital voucher for a bar of chocolate was trialed with the Finnish ISSP 2018 sample. The survey was conducted following a web-push strategy with the incentive designed to boost both early web response as well as to improve the overall response rate. The results indicate that motivating respondents with chocolate not only seems to increase response, but also seemed to benefit the overall sample composition in terms of nonresponse bias and to reduce the selectivity of respondents based on topic interest. At the same time the effectiveness was found conditional on some respondent characteristics such as living in the rural countryside as well as certain age groups. The effects of the incentive on data quality was further assessed by comparing survey estimates on religious community membership with external statistical data. The tentative evidence was found promising and Statistics Finland will conduct one or more follow-up experiments later in 2019–2020.

**Keywords**: digital incentive, nonresponse, web-push survey, external validation, experimental culture, logistic regression, r-indicator

## Introduction

Research on response rates to mail surveys have shown that incentives are a solid tool to increase response rates. The effects of incentives on data quality remain debated however with arguments being made for both positive and negative consequences. Incentives could persuade people with lower motivation in the survey topic to participate and their response could be of poorer quality (Teisl, Roe & Vayda 2005). According to the leverage-salience theory of survey participation, noncooperation is especially problematic when the survey topic is associated with the respondent decision to participate. People who are more involved with the survey topic in question tend to respond at higher levels than others while at the same time are likely differ in the attributes on key survey variables from non-respondents that are less involved with the topic.

On the other hand, a number of other factors which are unassociated with the topic may also contribute significantly to respondent cooperation, such as the type of organization conducting the survey (government agencies and universities tend to fare better than commercial or NGOs). Thus it could be argued that incentives that are both effective in increasing response rates as well as unassociated with the survey topic would be most effective in reducing nonresponse error related to topic interest. (Groves et al. 2000 & Groves et al. 2004) Findings from previous studies clearly have indicated that bias on estimates which are most associated with the survey topic tend to be more biased than other estimates (Groves, Presser and Dipko 2004).

Statistics Finland has been conducting the Finnish data collection of the International Social Survey (ISSP) Programme data collection since 2000. The ISSP is a cross-national programme of conducting annual surveys on topics relevant to social sciences. The 2018 ISSP survey focused on religion and attitudes towards religion. Out of all the topics covered religion is among those which is most avoided and felt uncomfortable to participate (Sherkat 2007) and is thus well suited to examine the possibility to counter topic sensitivity bias via incentives. This is because it can be argued that survey participants of the 2018 ISSP round may be biased towards either the emphasis the respondent gives to religion, the lack thereof (atheists may also feel more motivated to participate a survey on religious values) or both (see Sherkat 2007). Either way there is reason to believe that this type of topic sensitivity will contribute to either the skewness or kurtosis of distributions of items assessing the overall religiousness of respondents: If topic sensitivity is present, either one or both tails should be overrepresented (as people less interested in religion are also less likely to participate).

The second key component of the leverage-salience theory also suggests that people who are less influenced by other aspects of the survey such as the reputation of the survey organization, monetary incentives or the persuasion efforts of the interviewer.

## Administering the survey

The survey was administered as a web-push survey in Finnish and Swedish (5,2 % percent of the target population being native Swedish speakers) with an initial postal request to respond over the Internet to which the chocolate incentive was conditioned. A paper questionnaire was sent prat the 3rd stage of mail contacts, after which no mention of the chocolate incentive was included. Thus only early bird web respondents in experiment group 1 and 2 got a bar of chocolate if they answered before given day (during 3 weeks). Web respondents received a text message code to their telephone and so they were able to pick up their bar of chocolate from the local kiosk service that practically covers all major municipalities. This does mean however that the trouble of picking up the reward varies especially in the rural areas, and we shall therefore also include in interaction term in the model to test, whether the incentive is less effective in non-urban areas.

The web-push survey which is followed by paper questionnaire results in a bipartition of the sample into PAPI and online respondents. Because the incentive was designed in order to boost online response, it may have also given rise to a particular mode effect with people otherwise less inclined to respond online doing so. This selectivity of the mode and its effects on the quality of survey estimates is however difficult to entangle.

In order to motivate response across different data collections, Statistics Finland follows a standard policy of including a lottery of fifty pieces digital gift voucher certificates worth 100 € each in all surveys, which is mentioned in every contact attempt with the wording “among all those responding to Statistics Finland surveys”, including sample units in the chocolate experiment group. What this means is that the experimental group is subject to one conditional lottery based plus the redeemable chocolate bar whereas the control group only receives the mention of the standard gift certificate lottery.

## Research design

The analysis proceeds in three steps. First, we take a look at the effect of the incentive on overall response rate in the experimental and control strata. Next in order to assess the effects of the incentive on cooperation and response mode, we estimate logistic regressions, using auxiliary data and level of contacting effort, predicting respondent cooperation with the survey request as well as web-response conditional on cooperation. The third step of the analyses examines data quality and the incentives effect on countering bias resulting from topic interest. We assess the quality of data in the two subgroups using benchmark external data on religious community membership as well as reported turnout in the 2015 parliamentary elections.

A common procedure is to compare the distributions of certain key survey variables to known statistics that are available for the target population. Two such sources of religion data are available which the ISSP is also attempting to estimate: belonging to a religious congregation (SVT 2018) as well as voter turnout in the 2015 national elections.

In order to increase the statistical power of the experiment design, the ISSP primary sample of 2500 was supplemented by an additional sample of 700 the survey in order to increase the statistical power of the research design. This additional sample was to have been administered as a web-only survey with two letters (one invitational and second a reminder) with the primary sample receiving up to four letters.

## Results

### Effects on the response rate and online response

Table 1 illustrates the overall increase in the response rate as well as the increase in willingness to respond online in the initial phase of data collection. In both perspectives the overall effect of the incentive seems to achieved both the intended effects: The overall response rate was around 4,6 percent better in the experiment group while the proportion of early bird online response was also significantly higher.

In terms of data collection economics, the incentive and its administration had both direct and indirect effects on the overall costs of the survey. We first compared the costs related to mailing of paper questionnaires, return envelopes as well as manual coding of returned questionnaires. The increase in early stage web-response was able to yield savings on all of these: In the incentive sub-sample the average expenses per sample unit were around 0,30 € lower for postage and 0,15 € lower for manual coding due to the relative increase in web-response. This saving was somewhat offset by the costs related to general administration of the incentive as well as the chocolate bars themselves increased the cost per sample unit by 1,54 €. Thus, the overall costs were around 1,10 € higher.

However, since the overall response rate was higher (increasing the relative effectiveness of the initial survey request), it would make more sense to compare the average administration cost per net response. Put this way on average, a single net respondent took around 0,92 € more financial effort with the incentive. Or put another way, a relative increase of 1 percentage point in the net response rate was achieved with a little more than 0,20 euros increase in the per respondent expenditure. In addition, as we have shown this increase in motivation effort has also contributed to the overall quality of the net sample, the benefit of which cannot be directly quantified in financial terms.

Table 1. Response rate in the experiment vs. control group as well as the costs related to administering the incentive per sample unit and as well as net response

|  |  |  |
| --- | --- | --- |
|  | **Control group** | **Incentive group** |
| N | 1500 | 1700 |
| Response rate\*(% of which online) | 35,0(67,7) | 39,6(76,0) |
| ***Expenses*** |  |  |
| Per respondent | - | +0,92 € |
| Per sample unit | - | +1,10 € |

Note: Increase in response rate found significant in the response propensity model presented in Table 2.

### Nonresponse model

In order to disentangle the independent as well as possible conditional effects of our incentive on both overall response and willingness to respond online, we estimate two multivariate models which include both contacting effort variables as well as respondent characteristics. The dependent variable in the first Model 1 signifies overall cooperation, whether the sampled case completed the survey in either web or postal mode. The second model predicts completing the survey online conditioned on cooperation. The third and final model predicts the outcome of having redeemed the chocolate bar for those who were included in the experimental group and responded online during the first two months of data collection.

The effort variable equals the number of questionnaires that were administered in each of the strata and varies between 2 and 4. The major difference between primary and supplementary sample strata was, that the initial letters sent to those allocated in the supplementary ones was that there was no indication of a further questionnaire being sent via conventional mail. However, during the final stage of data collection, a decision was made to send one additional questionnaire via post to Experiment group 2 because it was deemed possible to increase the overall response rate and to enhance the sample size of the primary sample by merging this supplementary strata with the primary sample. Because of this we also run all the statistical models again in such a way that we ignore the small number of postal response in Group 4 which were received after this additional effort, but no significant changes in our parameter estimates was observed.

The first two models include three measures of effort, including whether or not the respondent was included in the incentive or control group. These include the number and type of mail questionnaires as well as the initial wordings of the advance letter (whether or not the respondent could be expecting a paper questionnaire following the initial web-push invitation). The predictor “supplementary sample” thus signifies that only one paper questionnaire was sent at a later stage of data collection.

In addition to the conventional postal reminders, we were also able to send out one additional reminder targeting respondents from whom a valid e-mail address was available (this was the case for a total of 41.9 percent of our sample). The availability of an e-mail address was strongly associated with respondent age, education, gender as well as type of municipality. Sample units residing in urban areas are more likely to have a valid e-mail.

In addition to the 4 rounds of postal reminders, an additional e-mail reminder was administered to every sample unit from which we were able to obtain a valid e-mail address from Statistics Finland population census database. The availability of a valid e-mail address is somewhat associated with respondent age, type of municipality as well as education (more educated younger people in cities had better coverage than did less educated older people). The incentive was mentioned in the e-mail request which was sent a week after the initial survey request letter in the incentive group whereas the standard lottery of gift certificates was mentioned in the control group e-mail.

Table 2 presents the results of three response propensity models — an overall response model, a web response model conditional on response and an incentive model conditional on valid web-response. These models include frame variables which were used in weighting the data. These include ISCED 2011 education, gender, age, language and type of municipality. The use of the response incentive was found to predict both response as well as choosing web over paper questionnaire. The availability of a valid e-mail address yielded similar results, although was found not to be significantly associated with redeeming the chocolate bar even though clearly contributing to the overall response rate online.

Table 2. Response propensity models for Cooperation and Online response

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Model 1:** Response | **Model 2**: Online response, conditional on response | **Model 3:** Redeemed the chocolate, conditional on web response |
|   | B | S.E. | Exp(B) | B | S.E. | Exp(B) | B | S.E. | Exp(B) |
| **Data collection** |  |  |  |  |  |  |  |  |  |
| E-mail available | **0.271\*\*\*** | **0.085** | **1.312** | **0.552\*\*\*** | **0.158** | **1.737** | 0.141 | 0.242 | 1.152 |
| Incentive vs no incentive | **0.202\*\*\*** | **0.075** | **1.224** | **0.391\*\*\*** | **0.140** | **1.479** |  |  |  |
| Supplementary sample | **-0.289\*\*\*** | **0.092** | **0.749** | **2.153\*\*\*** | **0.299** | **8.613** | 0.013 | 0.256 | 1.013 |
|  |  |  |  |  |  |  |  |  |  |
| Sex (=male) | **-0.234\*\*\*** | **0.075** | **0.791** | 0.031 | 0.141 | 1.032 | -0.178 | 0.221 | 0.837 |
| Language other than finnish/swedish(=yes) | **-0.466\*\*\*** | **0.125** | **0.628** | -0.108 | 0.253 | 0.897 | **-0.579†** | **0.341** | **0.561** |
| **Age group (vs. <45)** |  |  |  |  |  |  |  |  |  |
| 45-64 | 0.110 | 0.089 | 1.116 | -0.992 | 0.173 | 0.371 | -0.151 | 0.256 | 0.860 |
| >65 | **0.511\*\*\*** | **0.117** | **1.667** | **-1.211\*\*\*** | **0.199** | **0.298** | **-0.572†** | **0.345** | **0.564** |
| **Education (vs. ISCED 1-2 or unknown education)** |  |  |  |  |  |  |  |  |  |
| ISCED 3-5 | -0.095 | 0.100 | 0.910 | 0.198 | 0.192 | 1.219 | 0.026 | 0.329 | 1.026 |
| ISCED 6+ | **0.500\*\*\*** | **0.106** | **1.648** | **0.549\*\*\*** | **0.197** | **1.731** | -0.069 | 0.332 | 0.933 |
| **Degree of urbanisation** |  |  |  |  |  |  |  |  |  |
| Semi-urban | -0.166 | 0.108 | 0.847 | -0.173 | 0.198 | 0.841 | -0.003 | 0.332 | 0.997 |
| Rural | -0.113 | 0.117 | 0.893 | -0.074 | 0.217 | 0.929 | **-0.659\*** | **0.315** | **0.517** |
|  |  |  |  |  |  |  |  |  |  |
| Constant | **-0.915\*\*\*** | **0.160** | **0.400** | 0.250 | 0.282 | 1.284 | 1.602 | 0.318 | 4.962 |
|  |  |  |  |  |  |  |  |  |  |
| Nagelkerke | 0.059 |  |  | 0.218 |  |  | 0.038 |  |  |
| Cox & Snell | 0.044 |  |  | 0.151 |  |  | 0.025 |  |  |
| Likelihood Ratio Chi² | **142.8\*\*\*** |  |  | **197.5\*\*\*** |  |  | 12.8 |  |  |
| Valid N | 3200 |  |  | 1204 |  |  | 511 |  |  |

+*p* < .10. \**p* < .05. \*\**p* < .01. \*\*\**p* < .001.

As expected both age and education were associated with cooperation as well as online response. Older people showed better overall cooperation although they clearly preferred to give their response via paper questionnaire. What is somewhat surprising was that the older respondents were also more reluctant in redeeming the chocolate bar they would have been entitled to. Respondents living in rural municipalities showed a similar result which can be explained by the lack of a local R-kiosk, where the code could have been redeemed. Non-Finnish speakers were also found less likely to cooperate and redeemed the incentive less often than other respondents. The lower cooperation of this group can be at least partially explained by the fact that the survey was administered only in Finnish and Swedish, though this does not explain their unwillingness to redeem the incentive (since they already responded to the questionnaire!). Men were also found less likely to participate whereas their willingness to use the online questionnaire did not differ from women.

### Non-response bias

We also ran two distinct response propensity models with frame variables (gender, age, type of municipality, language, education and e-mail address) for both groups separately in order to assess the overall bias of both sub-samples. Following the suggestion of Betlehem, Cobben & Schouten (2011, 138) we can assess the net effect of each frame variable on the response in different sub-samples.

First and foremost it is worth noticing that the models do not fit well in neither of the sub-samples, which is good news since a good fit would indicate a strongly biased response. Respondent education and age group are most strongly associated with response in both models, though the observed change between sub-groups seem puzzling. The effect of education (both level and field of study) is found greatly reduced with the introduction of the incentive indicating, that the incentive may have encouraged less educated to respond. At the same time however we also observe a slight increase in the contribution of age and language groups as well as the availability of an e-mail address. Women and especially younger people may have been more receptive to the incentive while the latter finding can be explained by the fact that we were able to motivate the respondent better by mentioning the chocolate incentive in the e-mail reminder, which was perhaps seen less motivating among the control group respondents. At any case these results on overall representativeness are ambiguous at best.

Table 3. Multivariate analysis of response behavior in the incentive and control group.

|  |  |
| --- | --- |
|   | **Wald X²** |
| **Control group** | **Incentive group** |
| Sex | **3,81**  | **4,28** |
| Age group | **19,03** | **23,32** |
| Degree of urbanisation | 1,94 | 1,80 |
| Language: Other | **7,86** | **10,47** |
| ISCED 2011 | **29,19** | **19,08** |
| ISCED Field 2013 | **16,70** | **14,38** |
| Availability of E-mail | **3,27** | **9,88** |
|   |
| Valid N | 1500 | 1700 |
| Nagelkerke R² | 0,082 | 0,087 |

**\* Bolded Wald test values indicated significance at 0.05-level**

### Data quality using external criteria

Our research interest concerning the comparison of estimates of voter turnout to register based turnout rate is twofold: First we seek to test the hypothesis that the incentive was able to attract people who are less by comparing the estimates of self-reported voting in the last parliamentary election of 2015 to actual turnout. Second, comparing this estimate to serves the purpose of external validation or an indicator of data quality. The less biased the estimate, the better is the quality of our data. Since overreporting of turnout is a commonly identified inconsistency in found in most election surveys (Belli, Traugott & Beckmann 2001), we are not surprised to see the overall estimate to be above the official turnout of 70,1 percent.

Respondents in the incentive subsample reported a turnout rate of 83,4 percent whereas the control stratum estimate was somewhat higher (86,9 percent) both of which clearly indicate the overall survey to be overrepresent voters. Though this difference is so small that it remains outside conventional levels of significance (Chi²~2.65; df=1; p~0.1), we conclude this as tentative evidence that the incentive was likely able to contribute data quality by convincing people less active in societal matters to participate.

### Topic interest

According to the leverage-salience theory of survey participation, interest in religion (whether as an extention of their faith or secularism) may be associated with willingness to respond to a survey request with a religious topic. Accordingly the reputation of the survey organization may be significant for others whereas some people may only be responsive to direct rewards awarded to respondents (Groves et al. 2004). Therefor we hypothesize that an increase in the reward of participation should adjust the distribution of those survey variables which are most strongly associated with topic interest. This is because there is reason to believe that with minor rewards, the non-response bias in terms of topic interest will be greater.

Thus we expect that the increase in response which can be accounted to the chocolate incentive should have attracted people with less emphasis on religious topics. This was examined by comparing the distribution of a key survey item measuring overall religiousness between the experiment and control samples. The survey item measuring overall religiousness “Would you describe yourself as--- 1 Extremely religious – 7 Extremely non-religious” which was recoded into three categories: Strongly non-religious (1-2), indifferent (3-5) and strongly religious respondents (6-7). Consistent with our expectations we found the experimental group reported somewhat less often extreme values on both ends of the scale, though the difference observed lacked statistical significance. Though the number of religiously indifferent respondents was found higher in the incentive group

Table 4. Self-reported religiousness in the incentive vs. control group sub-samples

|  |  |
| --- | --- |
|   | Sub-sample |
| Control | Incentive |
| Strongly religious (6-7) | 7,4% | 7,8% |
| Indifferent (3-5) | 58,6% | 61,5% |
| Strongly non-religious (1-2) | 34,0% | 30,6% |

Chi²~1.529; df=2; p~0.466

Finally we were able to assess the quality of data by comparing the estimated belonging to various religious communities to benchmark register based data obt­ained from Statistics Finland population statistics. Table 5 presents the register based distribution which is compared with the estimated distributions of both sub-samples.

For comparison of data quality we calculate an index of dissimilarity as a summary measure for the consistency of the ISSP subsample and register based distributions. In this comparison, a value of 0 indicates that there is no dissimilarity between the religious community membership estimate and benchmark data. This value can be interpreted as the percentage of respondents that should change categories to produce the same distribution as in the benchmark data. For the distribution in the control group, we measure a dissimilarity of 8.2 whereas the observed dissimilarity is somewhat higher in the incentive group (12.9). This finding seems to indicate that the latter distribution is more biased than the control group. Thus somewhat contrary to our expectations we found the while the total sample seemed to significantly overestimate the proportion of non-members, this bias was found somewhat more prominent among the incentive group. The difference in the bias observed between experiment and control group remain small however. Though misreporting of religious community membership could be a possible explanation for observed difference, we nevertheless find the observed differences between sub-samples too small for inference.

Table 5. Estimated distributions of belonging into religious denominations compared with benchmark data

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Religious denomination in population statistics** | **Control group** | **Incentive group** |
| Evangelical Lutheran | 3 912 187 | **71,0 %** | **62,8 %** | **59,2 %** |
| Russian Orthodox | 61 441 | 1,1 % | 2,2 % | 1,7 % |
| Catholic | 14 169 | 0,3 % | 1,1 % | 0,3 % |
| Other christian | 55 087 | 1,0 % | 1,5 % | 3,5 % |
| Jewish | 1 093 | 0,0 % | 0,2 % | 1,0 % |
| Muslim | 15 359 | 0,3 % | 1,7 % | 0,2 % |
| Buddist | 1 814 | 0,0 % | -  | -  |
| Hindu | 353 | 0,0 % | -  | -  |
| Other religion | 889 | 0,0 % | 0,7 % | 0,3 % |
| **None** | **1 450 738** | **26,3 %** | **29,8 %** | **33,8 %** |
| Total | 5 513 130 | 100 | 100 | 100 |
| **Index of dissimilarity** |  |  | **8,2** | **12,9** |

## Discussion

All in all the results of the experiment can be seen as successful even though some of the results concerning data quality using both internal and external criteria yielded somewhat ambiguous results. The incentive was successful in terms of its primary intended effect of increasing response rate. Secondary objective of increasing the representativeness of the sample was met only somewhat with mixed results. Although we found some tentative evidence that the chocolate motivation was able to convert some people usually underrepresented in surveys as respondents among the less educated, this effect was somewhat offset by the relative increase in bias in terms of sample age distribution.

Comparing the estimated distributions of electoral participation as well as religious denominations represented in the sample we also found mixed results: The incentive sub sample was able to reduce the bias related to political efficacy by attracting non-voters to the respondent pool. This we argue is a very important finding since there is reason to believe, that the underlying mechanisms of survey and electoral participation may be strongly associated, though beyond the scope of this paper. Though small in terms of statistical inference, we also remain puzzled by the observed increase in response bias related to religious community membership, though this could be explained by bias related to self-reporting.

All in all, the relatively small increase in survey effort seemed to yield a number of interesting effects on data quality, which Statistics Finland will investigate further in the future. Statistics Finland is hoping that respondents could be motivated by different statistical information- and web-based incentives which will be conducted along with future relevant surveys. Statistics Finland will conduct one or more follow-up experiments later in 2019–2020.

## References

Belli, R. F., Traugott, M. W., & Beckmann, M. N. (2001). What leads to voting overreports? Contrasts of overreporters to validated voters and admitted nonvoters in the American National Election Studies. *Journal of Official Statistics*, *17*(4), 479.

Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation: description and an illustration. *The Public Opinion Quarterly*, *64*(3), 299-308.

Groves, R. M., Presser, S., & Dipko, S. (2004). The role of topic interest in survey participation decisions. *Public Opinion Quarterly*, *68*(1), 2-31.

Mykkänen, J. (2011). Uskonto ja äänestäminen. *Muutosvaalit*, 292-309.

Sherkat Darren (2007). Religion and Survey Non-Response Bias. Toward Explaining the Moral Voter Gap between Surveys and Voting. Sociology of Religion 2007, 68:1, 83-95).

Official Statistics of Finland (OSF): Population structure [e-publication].
ISSN=1797-5395. 2018, Appendix table 1. Population according to language 1980–2018 . Helsinki: Statistics Finland [referred: 27.6.2019].
Access method: <http://www.stat.fi/til/vaerak/2018/vaerak_2018_2019-03-29_tau_001_en.html>

Teisl, M. F., Roe, B., & Vayda, M. E. (2005). Incentive effects on response rates, data quality, and survey administration costs. *International Journal of Public Opinion Research*, *18*(3), 364-373.