**Survey quality: Response bias in retrospective questions**

Mathias Killengreen Revold, Statistics Norway, rmk@ssb.no

Mari Lande With, Statistics Norway, mlw@ssb.no

**Abstract**

*The paper examines changes in survey respondents’ answers to retrospective questions when time since the event increases. As a case, we examine self-reported electoral participation using data from two waves of the Norwegian Survey on Living Conditions. As shown in previous studies, self-reported electoral participation is much higher than actual turnout. We find little discrepancy in self-reported participation at the aggregate level when measured after approximately 6 months and after 4 years. At the individual level, there is a larger discrepancy in reported party-choice. At both time points, very few respondents admit to not remembering whether they voted and which party they voted for. Results suggest that respondents are eager to respond even though they do not remember. We find no evidence that respondents give more socially desirable answers as time passes since the event.*

**Keywords:** Survey methodology, social desirability bias, forgetting, self-reported electoral participation

**1. Background**

The paper examines changes in survey responses over time, when respondents are asked about participation in the Norwegian parliamentary elections at two time-points: shortly after it took place and after four years. Both respondents' misreporting and non-response bias may lead to inaccurate survey results. Whereas previous studies have examined changes in such biases over a relatively short period of time, our data allows us to study the effect several years removed from the event.

Self-reported electoral participation is generally higher than actual turnout, both at the individual and aggregate level (Blais, Gidengil and Nevitte 2004, Holbrook and Krosnick 2010, Selb and Munzert 2013). This is also the case in the Norwegian Survey on Living Conditions, which is the data we use. In this study, we are not primarily concerned with electoral participation but use this as a case to examine the impact of misreporting and non-response bias and thereby the quality of survey data.

**2. Theoretical perspectives and previous research**

Survey quality is affected by both over-representation of certain types of respondents and by respondents’ misreporting. Several factors may affect how respondents answer survey questions, yielding inaccurate results. Such factors include comprehension of the question asked, respondents' failure to remember events accurately (forgetting), psychological mechanisms leading people to present themselves in a socially desirable manner (social desirability bias) (Biemer and Lyberg 2003).

Generally, forgetting increases the longer it has been since the event. Responses are most accurate when events are regular and seen as important (Tourangeau, Rips and Rasinski 2000). The Norwegian parliamentary elections occur every four years and may therefore be seen an example of this type of event.

In itself, forgetting does not necessarily lead to systematic bias. However, respondents may remember events that happened outside the reference period as occurring within the reference period. Looking at health data, Kjellson, Clarke and Gerdtham (2014) finds that the number of errors increase as recall periods become longer. In the case of election surveys, this type of error could lead to non-voters who voted in previous elections erroneously reporting that they voted. For rare events, this type of over-reporting may also occur because respondents feel the event should be documented or because they try to help the interviewer (Biemer and Lyberg 2003).

Social desirability bias is typically systematic: Events or behaviours that are seen as socially acceptable are over-reported, whereas those that are perceived to be unacceptable are under-reported (Krumpal 2013). Electoral participation is a type of civic engagement that is typically regarded as socially desirable and is thus an event that is likely to be over-reported (Holbrook and Krosnick 2010). Previous studies suggest that some types of respondents are especially likely to over-report voting. People who usually vote are more likely to over-report, and over-reporters are generally similar to voters in terms of educational level and political engagement (Abelson, Loftus, and Greenwald 1992, Belli, Traugott, and Beckmann 2001).

There is also indication that respondents give more socially desirable answers if they do not remember. For example, previous studies have found that turnout over-reporting may increase as the time between the election and the interview increases (Belli, Traugott, and Beckmann 2001).

In addition to bias due to inaccurate responses, survey data may also be erroneous because survey participants differ from individuals that do not participate, non-response bias (Tarrant et al 1993, Selb and Munzert 2013). In our study, we use weights to adjust for systematic differences in response rates between population groups, but this is only possible if population characteristics are known. Election surveys may be particularly susceptible to non-response bias because electoral participation and survey participation are products of common underlying causes (Selb and Munzert 2013). Estimates cannot be weighted to adjust for higher participation among civic minded individuals unless differences in response likelihood correlate strongly with other variables (e.g. level of education).

*Hypotheses*

In this study we compare self-reported participation in the 2013 parliamentary election, measured in two surveys conducted in 2014 and 2017 respectively. The first survey was conducted approximately 6 months after the election, and the second three years later. Although accuracy is most likely higher when events are measured only shortly after they happened, retrospective questions with a larger time-frame are often the only available means to retrieve information.

Although non-response bias is an important source of overestimation of electoral participation, we have no reason to expect that this bias varies between the two surveys. Thus, we expect that systematic differences in self-reported electoral participation between the two surveys are caused by changes in individuals’ ability to recall the event accurately and possibly also changes in respondents’ motivation to intentionally misreport.

As time passes since the elections, we see the four following scenarios as possible effects of social desirability bias and forgetting on reported electoral participation:

H1: Respondents forget and give more socially desirable answers. Thus, reported participation rates increase over time at both the individual and the aggregate level.

H2: Respondents forget and do not feel strong social desirability pressures. Thus, the rates of do-not-know answers increase.

H3: Respondents answer more honestly because they feel less social desirability pressure years after the event. Thus, the aggregate participation rate is closer to the actual turnout. At the individual level more respondents change their answer from yes to no than vice versa.

H4: Respondents remember important events such as elections and answer truthfully at both time points. Thus, there is a small discrepancy in participation rates at the aggregate and individual level. Over-reporting is therefore due to overrepresentation of voters in the sample.

**3. Data and methodology**

The data used in this study is the Norwegian Survey on Living Conditions (SILC) from 2014 and 2017. Both years the survey included questions on participation in the last parliamentary election. Since there had been no national elections between the surveys both referred to the 2013-election. Thus, we asked about the same event 4-9 months and 3.5-4 years after it happened. Electoral participation was not the main topic of the surveys. This reduces the risk of self-selection of politically involved individuals.

The response rate in the 2014 survey was 55 per cent, and 54 per cent in 2017. Non-response weights were applied[[1]](#footnote-1). The survey has a rotating panel design and about ¼ of the 2014 sample were invited to participate again in 2017. Thus, the sample size is considerably lower in the analyses conducted on panel data.

The sample is randomly drawn among residents aged 16 or more at the time of survey. We only included respondents that had a right to vote in the 2013 parliamentary election in the analyses. Respondents who were younger than 18 in 2013 and non-citizens are therefore excluded. The old and sick are likely to have low participation rates in surveys, as well as in elections. Therefore, results for this age group should be interpreted with care.

There were 6 665 respondents in the 2014 survey and 5 353 respondents in the 2017 survey that were eligible voters in 2013 and answered the questions about electoral participation. 1094 respondents participated in both survey waves.

**4. Results from cross-sectional data**

The reported rates of voting were relatively consistent in the surveys. In the 2014 survey a total of 88.8 per cent of eligible voters reported to have voted in the 2013-election, while 88.2 reported to have voted in the same election in the 2017-survey. Thus, at the aggregate level forgetting and social desirability bias appear to have a consistent impact on the headline numbers.

**Table 1. Reported electoral participation in 2013 parliamentary election in the SILC-survey in 2014 and 2017, cross-sectional data.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Response | Voted | Did not vote | Do not want to answer | Do not know |
| **SILC 2014** | 88.8 | 10.9 | 0.3 | 0.1 |
| **SILC 2017** | 88.2 | 11.4 | 0.2 | 0.2 |

Source: Norwegian Survey on Living Conditions, 2014 and 2017

Few respondents admit to not remembering whether they voted in the election. It is just as uncommon to refuse to answer. This indicates that either people do not forget whether they voted or that respondents that do not remember try to give a response anyway.

The electoral participation rates in both surveys are more than ten percentage points higher than the actual turnout: While 89 per cent reported to have voted in the both surveys[[2]](#footnote-2), the actual turnout was 78 per cent. This is in line with the average over-reporting in post-election studies (Selb and Munzert 2013). Electoral participation was over-reported at approximately equal levels for men and women.

The reported rate of voting was close to actual turnout for one age group: respondents who were 18 or 19 in 2013. This group voted for the first time in 2013. At the time they were interviewed this was the only parliamentary election they could have participated in. This limits the risk of misremembering which election the question refers to. Electoral participation was overreported by at least five percentage points in the other age groups. The rates are most over-reported among respondents in the oldest age group (80 years or older). However, this could be due to the oldest and sickest in this age-group both having lower probability of voting and participating in the surveys.

The Norwegian SILC underestimates the differences in participation rates between education groups. While only 65 per cent of persons with education lower than the upper secondary level voted, approximately 80 per cent with this education level reported to have voted in the surveys. The actual participation rate for individuals with tertiary education was 90 per cent, whereas 95 per cent of the respondents in this group reported to have voted.

**Table 2. Reported and actual participation rate in 2013 parliamentary election in various groups**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Actual turnout** | **SILC 2014** | **N 2014** | **SILC 2017** | **N 2017** |
| **All** | **78.2** | **89.1** | 6 665 | **88.6** | 5 353 |
| Men | 76.8 | 87.9 | 3 456 | 88.4 | 2 739 |
| Women | 79.9 | 90.3 | 3 209 | 88.8 | 2 614 |
| **Age in 2013** |  |  |  |  |  |
| 18-19 years | 70.4 | 68.6 | 204 | 70.7 | 178 |
| 20-24 years | 63.0 | 75.0 | 470 | 75.1 | 415 |
| 25-44 years | 75.6 | 87.8 | 2 056 | 86.5 | 1 765 |
| 45-66 years | 82.8 | 93.2 | 2 772 | 92.4 | 2 214 |
| 67-79 years | 87.6 | 92.8 | 929 | 95.2 | 660 |
| 80 years or older | 70.7 | 90.5 | 234 | 93.2 | 121 |
| **Education** |  |  |  |  |  |
| <Upper secondary | 65.1 | 80.0 | 1 140 | 78.1 | 814 |
| Upper secondary | 78.5 | 89.1 | 2 964 | 88.4 | 2 234 |
| Tertiary | 90.4 | 95.2 | 2 542 | 92.8 | 2 288 |

Source: Election statistics and the Norwegian Survey on Living Conditions, 2014 and 2017

There were only minor changes in the aggregate reported participation rates between the surveys for the population groups we look at. Again, this indicates that the mechanisms leading to misreporting affect the results similarly when measured several years after the event as when measured after approximately 6 months.

**5. Results from panel data**

Of the 1094 respondents who participated in both survey waves, 43 changed their response from “voter” to “non-voter” from the first to the second survey, whereas 28 respondents changed their response from “non-voter” to “voter”[[3]](#footnote-3). As observed in the cross-section analysis, very few individuals refused to answer the question or said they did not know. In total, approximately 7 per cent of panel participants changed their voter status from the first to the second time they were asked.

**Table 3. Number of respondents by voter status in 2014 and 2017. Panel participants.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 2017: Voted in the 2013 parliamentary election | | | |  |
|  | Yes | No | Do not want to answer | Don't know | Total |
| 2014: Voted in the 2013 parliamentary election |  |  |  |  |  |
| Yes | 971 | 43 | 1 | 2 | 1017 |
| No | 28 | 47 | 0 | 0 | 75 |
| Do not want to answer | 0 | 1 | 1 | 0 | 2 |
| Total | 999 | 91 | 2 | 2 | 1094 |

Source: Norwegian Survey on Living Conditions, 2014 and 2017

Overall, a slightly larger percentage changed their answer from yes to no, than vice versa: 4 per cent versus 2.5 per cent. However, this pattern is not consistent across groups. Thus, there is no evidence that respondents’ answers are more affected by social desirability bias as time passes, as one would expect if this bias increases as the risk of misremembering increases. Neither is there clear evidence that answers are less affected by desirability bias over time as respondents can distance themselves from the event several years later.

The self-reported voters were also asked about party choice. Table 4 shows changes in the answers about party choice for panel participants with “voter status” in both waves. 73 per cent reported to have voted for the same party both survey years, whereas 16 per cent changed their answer to a different party. Nearly 8 per cent refused to disclose party choice in at least one of the surveys. A relatively small percentage, 2 per cent, changed their answer from a party to “don’t know”.

**Table 4. Respondents by combinations of answers about party choice in 2014 and 2017. Panel participants. Per cent.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Same party** | **Changed party** | **Change from "don't know"** | **Changed to "don't know"** | **"Don't know" both surveys** | **"Do not want to answer"** | **N** |
| Per cent | 73.1 | 16.1 | 0.8 | 2.1 | 0.0 | 7.9 | 971 |

Source: Norwegian Survey on Living Conditions, 2014 and 2017

If individuals who answered “don’t know” or “do not want to answer” either year are excluded from the analysis, the percentage of voters who changed party is slightly higher, at 18 per cent. Discrepancy in responses about party choice is clearly higher than for voting, even though only those who answered consistently as “voters” are included. This suggests that the impact of time on response accuracy varies between the type of phenomena measured. As there are more alternatives, party choice is more difficult to remember than voting, and it is less likely to give the same response by chance.

Furthermore, more than one party may be seen as a socially acceptable answer. This also reduces the risk that those who do not remember, consistently choose the same answer. Interestingly, relatively many respondents change their answer about party choice, while very few admit that they do not remember. This suggests that respondents try to give an answer even when they have forgotten which party they actually voted for. Between the surveys there were local elections in 2015. Possibly, party choice in the parliamentary elections is confused with that of the local elections.

**6. Changes in answers in selected demographic groups**

Table 5 shows the percentages who changed their answers between the surveys for selected demographic groups. The percentage who changed their answer is highest among persons with education lower than upper secondary level, and also decreases with age. This is the case both for the question about electoral participation and party choice.

**Table 5. Percentage who changed their answer between survey waves in selected groups. Panel participants. Per cent.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Changed answer, electoral participation** | **N (panel participants)** | **Changed answer, party choice** | **N (voters both waves)** |
| **All** | **6.7** | **1094** | **18.0** | **868** |
| Men | 7.0 | 576 | 17.8 | 464 |
| Women | 6.4 | 518 | 18.3 | 404 |
| **Age in 2013** |  |  |  |  |
| 18-24 years | 11.4 | 106 | 22.5 | 67 |
| 25-44 years | 8.1 | 323 | 15.4 | 258 |
| 45-66 years | 5.8 | 483 | 20.6 | 404 |
| 67 years or older | 3.8 | 182 | 14.2 | 139 |
| **Education** |  |  |  |  |
| <Upper secondary | 14.7 | 149 | 21.7 | 102 |
| Upper secondary | 6.5 | 453 | 18.1 | 352 |
| Tertiary | 4.4 | 492 | 17.1 | 414 |

Source: Norwegian Survey on Living Conditions, 2014 and 2017

**7. Discussion**

We started the paper by formulating four hypotheses about how time affects changes in survey responses. According to H1, the probability that respondents forget and give more socially acceptable answers increases with time, therefore reported electoral participation would increase. However, our results do not support this, as there is little change in the reported participation rate between the surveys. Neither the cross-sectional nor the panel analyses indicate that respondents give more socially desirable answers as time passes.

H2 suggested that with time, respondents forget, but do not feel strong social desirability pressures, so that the rates of “do-not-know” answers increase. As very few respondents change their answer to “don’t know”, there is no support for this hypothesis. The relatively high percentage of respondents changing their answer about party choice, as well as the very low percentage who answers “don’t know” to both questions, clearly indicates that respondents try to give an answer even if they are unsure.

According to H3, respondents would answer more honestly in the second survey because they feel less social desirability pressure years after the event. Thus, the aggregate participation rate should be closer to the actual turnout and at the individual level more respondents would change their answer from yes to no than vice versa. However, the aggregate participation rate hardly changes, and only a slightly larger number of respondents change their answer from yes to no. Thus, there is little evidence that social desirability bias decreases with time.

H4 stated that respondents would both remember and answer truthfully at both time points, and that over-reporting would only be caused by non-response bias. Although the small discrepancy in participation rates at the aggregate level is in line with this, the observed changes in individuals’ responses suggest that many respondents forget yet are eager to give a response anyway. Nevertheless, it is difficult from our results to estimate the relative impact of non-response bias and recall and social desirability bias.

At the individual level, answers to the question about electoral participation are more consistent than answers to the question about party choice, even though only those who answered consistently (yes) to the first question were able to answer the second question twice. One explanation for this could be that it is easier to remember whether you voted than which party you voted for. Consequently, over-reporting of electoral participation could be caused solely by non-response bias, whereas answers to questions that are harder to remember are more influenced by respondents’ misreporting. Non-response bias could also explain the particularly high voting rates reported by respondents with a low level of education, as this group may be represented by its’ most civic minded members.

Although we find no evidence that social desirability bias is more prevalent years after the election than shortly after the event, it is still possible that respondents give equally biased responses at both time-points. The relatively large change in responses about reported party choice, combined with the low percentage “do-not-know”-answers both years, indicate that respondents do forget and still try to give an answer. Thus, it may be that respondents that are unsure whether they voted give the socially desirable response, and that this contributes to the high reported voting rates.

**8. References**

Abelson, R., Loftus, E., and Greenwald, A. (1992), Attempts to improve the accuracy of self-reports of voting, Questions about questions, p.p. 138-153.

Belli, R., Traugott, M., & Beckmann, M. (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, 479-498.

Biemer P. and Lyberg L. (2003), Introduction to Survey Quality, J.Wiley, N.Y.

Blais, A., Gidengil, E., and Nevitte, N. (2004), Where does turnout decline come from?, European journal of political research, 43, pp. 221-236.

Holbrook, A. and Krosnick, J. (2010), Social desirability bias in voter turnout reports: Tests using the item count technique, Public Opinion Quarterly, 74, pp. 37-67.

Kjellson, G., Clarke. P., and Gerdtham U. (2014), Forgetting to remember or remembering to forget: A study of the recall period length in health care survey questions, Journal of Health Economics, 35, pp. 34-46.

Krumpal, I. (2013), Determinants of social desirability bias in sensitive surveys: a literature review, Quality & Quantity, 47, pp. 2025-2047.

Selb, P., and Munzert, S. (2013), Voter overrepresentation, vote misreporting, and turnout bias in postelection surveys, Electoral Studies, 32, pp. 186-196.

Tarrant, M., Manfredo, M., Bayley, P., and Hess, R. (1993), Effects of recall bias and nonresponse bias on self-report estimates of angling participation, North American Journal of Fisheries Management, 13, p.p. 217-222.

Tourangeau, R., Rips, L., and Rasinski, K. (2000). The psychology of survey response. Cambridge University Press.

1. Sex, age-group, family size and education level are used in the calculation of the weights. [↑](#footnote-ref-1)
2. Do not know and do not want to answer are excluded. [↑](#footnote-ref-2)
3. In both survey years the self-reported electoral participation was slightly higher among the panel-respondents than for the cross-sectional samples. This means that over-reporting is even higher for the panel sample. [↑](#footnote-ref-3)