Public Opinion Quarterly

DO SURVEYS OVERESTIMATE OR UNDERESTIMATE SOCIOECONOMIC DIFFERENCES IN VOTER TURNOUT?
EVIDENCE FROM ADMINISTRATIVE REGISTERS

HANNU LAHTINEN*,
PEKKA MARTIKAINEN
MIKKO MATTILA
HANNA WASS
LAURI RAPELI

Abstract Surveys generally overestimate the overall level of voter turnout in elections due to both the misreporting of voting and nonresponse. It is sometimes argued that socioeconomic differences in turnout are exaggerated in surveys because social desirability has a more pronounced effect on eligible voters in more advantaged socioeconomic positions. However, the contribution of nonresponse bias has not been taken into consideration in these assessments. Using a register-linked survey with information on the education, occupational social class, income, and voting in the 2015 Finnish parliamentary elections of both respondents and nonrespondents, this study shows that nonresponse bias leads to not only a larger overestimation of the overall level of turnout

Hannu Lahtinen is a doctoral candidate in the Faculty of Social Sciences at the University of Helsinki, Helsinki, Finland. Pekka Martikainen is a professor of demography in the Faculty of Social Sciences at the University of Helsinki, Helsinki, Finland; visiting professor at the Karolinska Institutet at Stockholm University, Stockholm, Sweden; and research affiliate at the Max Planck Institute for Demographic Research, Rostock, Germany. Mikko Mattila is a professor of political science in the Faculty of Social Sciences at the University of Helsinki, Helsinki, Finland. Hanna Wass is an acting professor of political science in the Faculty of Social Sciences at the University of Helsinki, Helsinki, Finland. Lauri Rapeli is the research director in the Social Science Research Institute at Åbo Akademi University, Turku, Finland. The authors thank Statistics Finland, Feelback Group Ltd., the Finnish Election Study Consortium, and the European Social Survey team of the European Research Infrastructure Consortium for providing the datasets used in this study. This study was supported by the Academy of Finland, grant number 266844 to M.M. In addition, P.M. was supported by grant number 1308247, H.W. was supported by grant number 273433, and L.R. was supported by grant number 312676, all from the Academy of Finland. *Address correspondence to: Hannu Lahtinen, University of Helsinki, P.O. Box 18 (Unioninkatu 35), 00014 Helsinki, Finland; email: hannu.lahtinen@helsinki.fi.

doi:10.1093/poq/nfz022

than social desirability, but also an underestimation of educational, social class, and income-related differences in the propensity to vote. Socioeconomic differences in the probability of voting in register-based data were at least two-thirds larger than differences obtained when using standard survey techniques. This finding implies that socioeconomic inequality in electoral participation is a more pressing social problem than previous evidence might indicate.

Population-based surveys constitute the key source of our knowledge on the individual-level determinants of electoral participation. Despite their prominence, it is widely known that surveys tend to overestimate the level of overall voter turnout. This is due to two mechanisms. First, many of the actual nonvoters claim to have voted when asked, which results in overreporting due to *social desirability bias* (Karp and Brockington 2005; Selb and Munzert 2013; Sciarini and Goldberg 2016, 2017). The second mechanism involves *nonresponse bias*, which occurs when those who are less likely to vote are also less likely to respond to surveys (Selb and Munzert 2013; Sciarini and Goldberg 2016, 2017).

Although the overestimation of overall turnout in surveys is well established, we know less regarding whether turnout differences between population groups are also biased due to similar mechanisms. Studies usually have been able to measure only the effect of overreporting on socioeconomic differences in voter turnout (Sigelman 1982; Hill and Hurley 1984; Belli, Traugott, and Beckmann 2001; Karp and Brockington 2005; Ansolabehere and Hersh 2012), so little is known regarding the extent to which the nonresponse mechanism contributes to these differences.

Among the different factors that stratify individuals' voting propensity, indicators of socioeconomic position[1] are of special interest. Differences in voting based on these factors demonstrate social inequalities in parliamentary representation, which is much less directly the case for many other factors commonly used to explain turnout, such as trust, ideological strength, newspaper readership, or political interest.[2] Since fair elections are a central institution of liberal democracies, and since the equal impact of individuals is the key principle behind universal suffrage, inequalities in turnout are a crucial concern. Therefore, the importance of consistently estimating socioeconomic differences also needs to be underlined.

This study leverages a rare opportunity to simultaneously assess the contributions of both overreporting and survey nonresponse to socioeconomic

1. We use "socioeconomic position" as an umbrella term that we attribute empirically to education, occupational social class, and income—arguably the three most commonly used factors for this purpose (Wolfinger and Rosenstone 1980; Leighley and Nagler 1992; Martikainen, Martikainen, and Wass 2005; Gallego 2007; Lahtinen et al. 2017).

2. However, factors that are comparable to indicators of socioeconomic position would certainly include gender and race/ethnicity in this respect.

differences in turnout. It poses two research questions: First, to what extent do misreporting and respondent selection contribute to the overall overestimation of turnout in surveys? Second, how much do these two mechanisms contribute to the estimated turnout differences between individuals from different educational backgrounds, social class positions, and income categories? Using a survey from the 2015 Finnish parliamentary elections that was linked to administrative registers, we were able to analyze both self-reported and register-based information on voting and the socioeconomic characteristics of the survey respondents. In addition, the dataset includes validated information on the voting and socioeconomic characteristics of the nonrespondents.

We demonstrate that nonresponse leads to a larger overestimation of the overall level of turnout than does social desirability. Moreover, due to the joint effect of these two mechanisms, socioeconomic differences in turnout are underestimated when using survey data.

The Misestimation of Voting in Surveys

MISREPORTING AND RESPONDENT SELECTION AS SOURCES OF SURVEY ERROR

The total survey error paradigm disentangles the sources of errors in a survey into several components. Although these components differ slightly in different descriptions of the framework, the two main error categories are measurement and representation-related errors (Groves et al. 2009; Groves and Lyberg 2010). Here, we address errors stemming from both of these categories: Social desirability bias is a form of direct measurement error, and nonresponse bias is a representation-related error.[3] In addition, since we have information on those persons whose phone number was not obtained, we can assess some of the issues related to noncoverage. The following sections discuss these sources of error with regard to both overall turnout estimates as well as estimates on turnout differences between socioeconomic groups.

OVERESTIMATION OF OVERALL TURNOUT LEVELS

Throughout the history of surveys on political behavior, scholars have been concerned about the overestimation of voter turnout (e.g., Parry and Crossley 1950; Clausen 1968; Traugott and Katosh 1979). The pressure to provide a socially acceptable response often has been considered the main culprit behind the inflated turnout figures reported by survey respondents. The most straightforward way to address this problem has been to use validated information on turnout drawn from actual voting lists instead of self-reports (McDonald 2007; Traugott 2007). However, according to meta-analysis conducted by

3. Note that the representativeness of the respondents as a survey methodological issue should not be confused with parliamentary representation discussed in the introduction.

Smets and van Ham (2013), only 11 percent of turnout studies have been able to use such information. Validated information on electoral participation is often laborious and expensive, or even impossible to obtain, and it possibly has its own errors as well (Berent, Krosnick, and Lupia 2016). Thus, a number of other means for overcoming this problem have been proposed, such as question-and-response wordings or honesty pledges (e.g., Abelson, Loftus, and Greenwald 1992; Belli, Moore, and VanHoewyk 2006; Duff et al. 2007; Hanmer, Banks, and White 2014; McDonald, Scott, and Hanmer 2017; Morin-Chassé et al. 2017), or developing statistical models to correct for biases (Katz and Katz 2010).

Although less studied than overreporting, potentially an even larger source of inflated turnout bias arises from the selection of survey respondents (Berent, Krosnick, and Lupia 2016; Sciarini and Goldberg 2017). At least part of the (self-)selection effect in inflated turnout estimates can be interpreted in light of the leverage-saliency theory of survey participation (Groves, Singer, and Corning 2000). According to this theory, individuals weight issues they consider salient (i.e., issues that are important for them) when they make the decision as to whether or not to participate in a survey. Leverage, in turn, is the direction of the effect of a specific salient issue, that is, whether it functions as a motivating (positive leverage) or demotivating (negative leverage) factor. If the salient issues of positive leverage outweigh those of negative leverage, then the decision to respond to a survey is made. Some strong predictors of electoral participation, such as political interest or a sense of civic duty (Smets and van Ham 2013), arguably are salient positive-leverage issues that affect whether or not one participates in a political survey.[4] This means that the propensity to participate in elections correlates with the propensity to participate in surveys, which results in an overestimation of the overall turnout rate.

To address selection bias, efforts have been made to increase survey response rates (Burden 2000; Keeter et al. 2000; Schmeets 2010). However, the relationship between response rates and response bias is far from straightforward, especially when the determinants of survey participation are highly correlated with the variables of interest (Groves and Peytcheva 2008), as in this case. An increasing response rate can even accentuate the biases (Lin and Schaeffer 1995; Groves 2006; Selb and Munzert 2013).

BIAS IN SOCIOECONOMIC DIFFERENCES IN TURNOUT

Studies have demonstrated that individuals in more advantaged socioeconomic positions are more likely to overreport voting, which implies that surveys may overestimate the socioeconomic gaps in turnout (Sigelman 1982; Hill and Hurley 1984; Silver, Anderson, and Abramson 1986; Belli, Traugott, and Beckmann 2001; Bernstein, Chadha, and Montjoy 2001; Karp and

4. Salient factors with negative leverage (such as a lack of trust in political and/or research-conducting institutions) are also potentially important, and they tend to correlate negatively with the propensity to vote (Smets and van Ham 2013).

Brockington 2005; Ansolabehere and Hersh 2012). The explanation for this pattern has been that the social pressure to vote is more pronounced among those in more advantaged positions, which also leads to a stronger desirability bias among them (Bernstein, Chadha, and Montjoy 2001). Ansolabehere and Hersh (2012, p. 458) have made perhaps the boldest conclusion in this respect, ending up questioning the overall validity of conventional (resource-based) theories of electoral participation:

> The dramatic effect of misreporting on models of participation demands a renewed effort at theory-building. Sociodemographic and political resources do not explain all that much about why certain people vote and others do not. These variables . . . simply perform the dubious function of identifying survey respondents who think of themselves as voters.

However, this conclusion may have been premature, as studies have not been able to assess the extent of nonresponse bias with respect to socioeconomic differences in turnout. There are grounds to hypothesize that the respondent self-selection effect could bias the estimates of socioeconomic differences in turnout in the opposite direction; that is, it underestimates socioeconomic differentials. Drawing again from the leverage-saliency theory of survey participation (Groves, Singer, and Corning 2000), factors that have positive leverage in shaping people's decisions to participate in a survey, such as a high level of political interest and a feeling of civic duty, correlate with not only participating in a survey and elections, but also an advantaged socioeconomic position (Jackson 1995; Hillygus 2005). This means that survey nonresponse bias may be larger for those in low socioeconomic positions. For example, the differences in political interest or in a sense of civic duty can be smaller between middle-class respondents and nonrespondents than between working-class respondents and nonrespondents.

Data and Methods

DATASET

The dataset used was the Health and Political Engagement Survey (n = 2,001, cooperation rate 18.9 percent),[5] collected between January 15 and February 18, 2016, via telephone interviews by a private research company, Feelback Group Ltd. The original survey sample (including nonrespondents) has been linked to information from administrative registers for several common sociodemographic background indicators,

5. This is the minimum cooperation rate according to the AAPOR (2016) Standard Definitions guide. That is, cooperation rate = (respondents)/(respondents + called, refusals + called, no answer + called, number was no longer in use). Corresponding numbers were 2,001/(2,001 + 4,31 7 + 3,935 + 325) = 0.1888.

including education, social class, and income. In addition, the survey contains register-based information on voter turnout in those electoral wards that used electronic voting registers, administered by the Ministry of Justice, in the 2015 parliamentary elections (research permit from Statistics Finland number TK-53-1532-15).

Wards using these electronic registers included 24.2 percent of the individual-level electorate. Individuals in the survey sample were matched by Statistics Finland using personal identity codes, and the anonymized dataset was provided to the research team. Every citizen of Finland has a unique personal identity code, which is used in practically all administration settings, including voting registers and other administrative registers. Also, our baseline sample, obtained directly from the Finnish Population Register Centre, contained these codes for each individual. Although the personal identity codes could not be directly accessed by the research team for privacy reasons, they allowed officials in Statistics Finland to conduct exact matching in linking different data sources. The linkage can be expected to be virtually error free, and therefore our data are of considerably higher quality than what has been used in many previous validation studies (for more discussion, see the methodological considerations section at the end of this article).

Telephone numbers were obtained from the national database of Finnish phone numbers, which is collaboratively maintained by the Finnish telephone operators. This database includes all the phone numbers maintained by the Finnish tele-operators, excluding prepaid plans or the numbers that users have declared ex-directory.

Our sampling frame for the survey was a random sample of 25,000 individuals originally drawn from the Finnish population register. For the analysis in this study, we excluded individuals who were aged 24 years or younger since their socioeconomic positions are not likely to have been well established yet, as well as 2,321 individuals for whom we were unable to identify all three indicators regarding socioeconomic position. After making these exclusions, 19,997 individuals were included. Register-based information on voting was available for 4,754 of the individuals who lived in the electoral wards that employed the electronic voting registers. This group constitutes the baseline sample in our analysis, as presented in the topmost box of figure 1.

After forming this baseline sample, we further decomposed it based on their survey answer status or the reason for not being interviewed. First, there were 54 individuals who did not speak Finnish or Swedish, as well as 1,668 individuals whose phone number was not available, meaning they were not approached for an interview. We subsequently included in this latter group 52 individuals who were called but whose phone number was no longer in use with the latter group, bringing that group size to 1,720 observations (see figure 1). After these

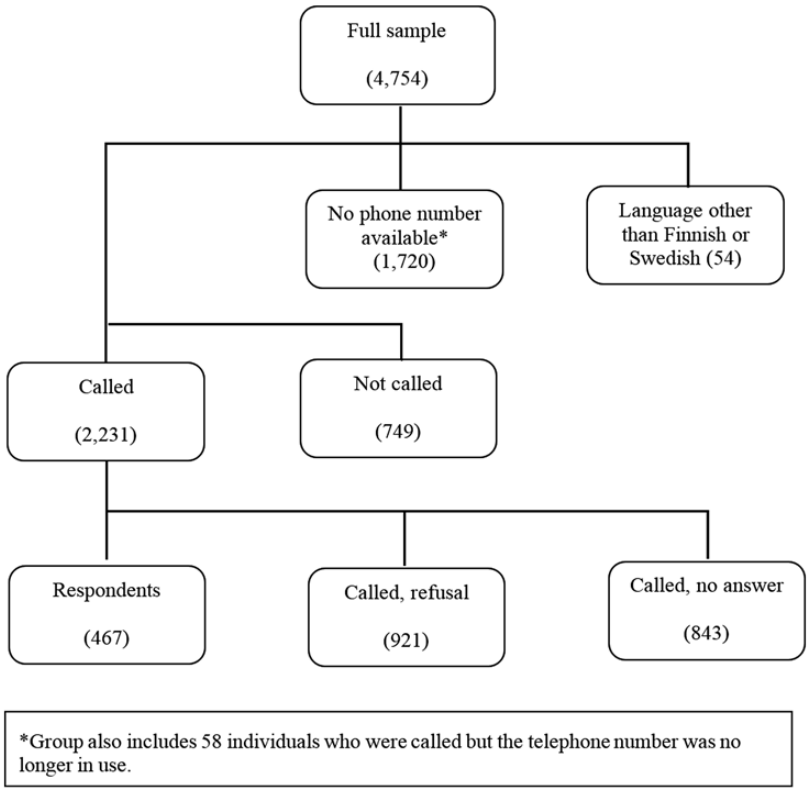

Figure 1. Sample description. Within brackets: number of individuals with all three measures of socioeconomic position and available validated voting data.

exclusions, 2,980 individuals were available for analyses. Of them, 2,231 were called and 749 were not called. The reason why certain individuals were not called was that their calling quota was filled during the fieldwork period. In the interview phase, quotas based on native language, gender, age, and the region of living were applied to this baseline random sample (for more information on calling quotas, see table A1 in the online appendix).

Those persons who were approached were further divided into three groups: those who did not answer (843 individuals); those who refused to participate (921 individuals), and survey respondents (467 individuals).

VARIABLES

The outcomes presented in this study are two dichotomous variables indicating whether an individual voted in the 2015 Finnish parliamentary elections, with the first outcome being based on self-reports and the second

on administrative registers.[6] To enable comparisons between the self-reported data and the register-based data and facilitate interpretation of the overall results, each of the indicators of socioeconomic position have been recoded into three categories. Education indicates the highest degree an individual has obtained (basic/secondary/higher). Social class was measured using the socioeconomic classification provided by Statistics Finland and categorized as follows: manual/intermediate/upper nonmanual. The intermediate class includes lower nonmanual and self-employed individuals. There is a relatively wide consensus that such a three-class breakdown can be considered hierarchal in terms of status or prestige (e.g., Erikson and Goldthorpe 1992, p. 45). Income was measured as tertiles based on household income after taxes. We primarily used register-based information to define the socioeconomic variables. In a few cases (n = 37), register-based information on income or social class was unavailable and self-reported position was used.

Control variables were kept at a minimum because we were primarily interested in the differences between groups rather than the effects that are independent of third factors. However, we made an exception with regard to age, which was controlled for as cubic splines with four knots (Durrleman and Simon 1989). Thus, structural changes, especially educational expansion, were taken into account. An increasing number of studies suggest that the effect of education on various outcomes should be considered in relative rather than absolute terms (for reviews, see Persson 2015; Bills 2016). For example, having secondary-level qualifications as a result of 12 years of schooling would imply a relatively high status for somebody who was born in 1930, whereas for somebody born in 1980 this would imply an average status. Likewise, income varies strongly by age, especially between working-aged and pensioned populations, but the rank order between individuals remains quite similar throughout this transition.

MODELING STRATEGY AND METHODS

The first analysis assessed how nonresponse and overreporting contributed to overall turnout. The second analysis estimated the nonresponse effect by comparing the educational, social class, and income-related turnout estimates provided by the different survey answer groups described in figure 1. Third, we assessed the overall misestimation of socioeconomic differences in turnout caused by both misreporting and nonresponse. This was done by comparing the estimates based on self-reported voting to survey respondents with validated voting and the estimates on both respondents and nonrespondents.

6. The formulation of the question in the survey was: "Sometimes individuals abstain from voting. Did you vote in the last parliamentary elections held in April 2015?" Answer categories were "yes," "no," and "did not have the right to vote." One respondent fell into the third category, and four respondents answered "don't know" spontaneously; all were excluded from the analysis. None of them had register-based information on voting available.

Finally, five series of robustness checks and replications were run (see the online appendix). First, we compared the estimates of socioeconomic differences in turnout from the baseline sample to the voting registers in their entirety, consisting of a quarter of the entire electorate. Second, we compared the distributions of those with validated voting information available to all survey respondents and the full baseline sample. Third, we replicated our main analysis of misestimation of overall socioeconomic differences in a multivariate setting. Fourth, we evaluated the extent to which the results differed depending on whether the socioeconomic indicators were based on registers or self-reports. And finally, we replicated the results on the aggregate level using two different surveys.

The first analysis (and the additional analysis presented in table A3 in the online appendix) was based on cross-tabulations. The other results are presented as predicted probabilities post-estimated from the logistic regression model, holding age as observed (Hanmer and Kalkan 2013).

Results

The first column of table 1 shows that turnout rates are overestimated when relying on self-reported survey estimates. Our validated turnout from the full sample was 74.3 percent.[7] The self-reported turnout of the survey participants was 15.4 percentage points higher, which reflects both the overreporting and selection of the respondents. However, the selection bias was three times the size of the misreporting bias (11.8 percent points vs. 3.6 percent points). This shows that nonresponse was a far more important source of error in measuring turnout than was social desirability or motivated memory failure-related bias. The gap between the reported turnout of all respondents and those for whom only validated turnout information was available was only 0.7 percentage points and statistically nonsignificant.

Another interesting detail from table 1 is that turnout among the full sample (74.3 percent) was close to the number of those who were approached but refused to participate in the interview (72.3 percent) or who did not answer the phone (73.8 percent). This was probably because those who can be approached are, on average, already in some ways in advantaged positions in society, namely being native speakers of either Finnish or Swedish rather than belonging to a language minority and owning a non-prepaid (cell or landline) phone plan. The group with the second-largest turnout (after the actual respondents) was "not called," who consisted of those whose sample quota had

7. The difference between the turnout from the full sample and the official turnout rate for citizens residing in Finland (70.1 percent) is mainly due to restricting the age of the study population to those at least 25 years of age. Therefore, the youngest voters, who have relatively low turnout rates (Bhatti, Hansen, and Wass 2012), are excluded.

Table 1. Size of the overall turnout bias (respondent selection and overreporting) and distribution of categories of socioeconomic position by answer status

Validated turnout by answer status				The distribution of socioeconomic variables by answer status								
				Educational groups			Social classes			Income tertiles		
	%	s.e.	*N*	Basic	Secondary	Higher	Manual	Intermediate	Upper non-manual	Low	Middle	High
Full sample	74.3	0.62	4,754	21.0	44.8	34.2	34.5	48.3	17.2	32.9	33.6	33.4
Respondents	86.1	1.57	467	22.7	41.5	35.8	30.8	53.3	15.8	33.4	38.1	28.5
Called, refusal	72.3	1.49	921	18.6	51.6	29.9	39.4	45.7	14.9	29.2	37.7	33.1
Called, no answer	73.8	1.54	843	17.9	47.8	34.3	34.2	49.5	16.4	28.0	38.6	33.5
Not called	83.7	1.38	749	34.4	36.3	29.2	32.6	52.1	15.4	45.0	31.2	23.8
No number available	69.4	1.10	1,720	17.5	44.5	38.0	33.9	46.3	19.8	31.4	29.0	39.7
Other language	37.0	6.70	54	22.2	35.2	42.6	29.6	46.3	24.1	51.9	31.5	16.7
Self-reported turnout												
All respondents[a]	89.0	0.72	1,839									
If validated turnout available	89.7	1.40	467									
Overall bias[b]	15.4	1.43										
Misreporting bias[c]	3.6	1.04										
Nonresponse bias[d]	11.8	1.56										

NOTE.—Standard errors (s.e.) obtained by bootstrapping (2,000 replications).

[a]Self-reported turnout of survey respondents, regardless of the availability of validated voting data.

[b]Self-reported turnout of respondents minus validated turnout of full sample.

[c]Self-reported turnout of respondents minus validated turnout of respondents.

[d]Validated turnout of respondents minus validated turnout of full sample (= overall bias minus misreporting bias).

already been filled. This means they had "easy to survey" demographics in terms of their age, gender, native language, and region of living. Therefore, the high turnout of this group is not surprising either.

An additional analysis (not shown) indicated that among the 467 individuals for whom both self-reported and validated turnout information was available, the voting information of 25 individuals (5.4 percent) differed between the two sources. Only four persons underreported their voting (all of those had voted in advance), whereas 21 overreported it. This indicates that the error in self-reported voting is mostly not random, but evident of systematic overestimation. Among the 65 validated nonvoters, 21 (32 percent) claimed to have voted. Among validated nonvoters, those with the highest educational qualifications, intermediate social class, or highest income tertile were most likely to overreport their voting. However, due to the small sample size, our results on socioeconomic differences in overreporting are only tentative and statistically insignificant.

Table 2 presents the differences in validated turnout between socioeconomic groups stratified based on their survey answer status. The gap between individuals with the highest and lowest socioeconomic positions was smaller among the respondents than among the overall population, regardless of the measure of socioeconomic position. In terms of education, the difference was small (27.2 vs. 26.0 points), whereas the differences in social class (23.2 vs 19.3) and income (18.0 vs. 8.9 points) were more substantial. We also observed relatively small socioeconomic differences among those who were not called. Another group with only minor socioeconomic differences was those with a mother tongue other than Finnish or Swedish. However, the number of observations in this group was too small to make strong conclusions. Large socioeconomic differences in turnout were observed among those who refused to participate or whose phone number could not be obtained from the operators' lists.

For the analysis shown in figure 2, we compared estimates of socioeconomic turnout differences on the basis of three strategies. "Respondents" (n = 1,832, circle-shaped symbols in figure 2) represent the conventional survey estimate of respondents and self-reported turnout. We compared that to the respondents using the validated turnout measure (n = 467, diamond-shaped symbol), and finally, to the estimates obtained from the full baseline sample (n = 4,722, square symbol). As in table 2, estimates of socioeconomic differences in validated turnout were larger than among survey respondents. In addition, when comparing socioeconomic differences based on survey respondents and self-reported turnout to those from validated turnout estimates of the full sample, the differences are even more pronounced than when using validated turnout information obtained from respondents. For example, conventional survey estimates of the differences between the most- and least-advantaged groups are 16.2 percent (= 95.1 – 78.9) for education, 13.9 percent (= 95.1 – 81.2) for social class, and 6.6 percent (= 92.1 – 85.5)

Table 2. Socioeconomic differences in validated turnout by survey answer status

	Education							
	Basic		Secondary		Higher		Difference: Higher–Basic	
	%	s.e.	%	s.e.	%	s.e.		*N*
Full sample	59.6	1.7	70.2	1.0	86.8	0.8	27.2	4,754
Respondents	69.5	5.3	83.2	2.6	95.5	1.5	26.0	467
Called, refusal	55.7	4.0	69.5	2.0	85.8	2.0	30.1	921
Called, no answer	59.8	4.6	71.3	2.2	83.1	2.1	23.3	843
Not called	77.2	2.8	82.4	2.3	92.8	1.7	15.6	749
No number available	52.3	3.0	61.9	1.7	85.2	1.4	32.9	1,720
Other language	27.5	12.8	32.8	10.5	45.4	10.0	17.9	54
	Social class							
	Manual		Intermediate		Upper non-manual		Difference: Upper non-manual–Manual	
	%	s.e.	%	s.e.	%	s.e.		*N*
Full sample	63.8	1.2	77.1	0.9	87.0	1.2	23.2	4,754
Respondents	75.7	3.6	89.2	1.9	94.9	2.5	19.2	467
Called, refusal	64.6	2.5	74.3	2.1	86.8	2.9	22.2	921
Called, no answer	65.3	2.8	76.6	2.0	82.8	3.2	17.5	843
Not called	77.6	2.7	84.7	1.8	93.1	2.4	15.5	749
No number available	54.3	2.0	73.3	1.5	85.8	1.9	31.5	1,720
Other language	39.8	12.1	29.6	9.0	47.5	13.7	7.7	54

Continued

Table 2. ***Continued***

	Income							
	Low		Middle		High		Difference:	
	%	s.e.	%	s.e.	%	s.e.	High–Low	*N*
Full sample	64.6	1.3	74.1	1.1	82.7	0.9	18.1	4,754
Respondents	80.8	3.4	87.4	2.4	89.7	2.6	8.9	467
Called, refusal	63.9	3.1	73.9	2.3	77.0	2.3	13.1	921
Called, no answer	67.0	3.2	71.7	2.4	81.4	2.3	14.4	843
Not called	75.3	2.5	89.4	2.0	91.0	2.1	15.7	749
No number available	57.8	2.2	64.7	2.1	81.2	1.5	23.4	1,720
Other language	31.0	8.5	41.4	11.5	47.1	16.0	16.1	54

NOTE.—Predicted probabilities based on binary logistic models (each row is based on a different model), adjusted using three cubic splines for age. Standard errors (s.e.) obtained using the delta method.

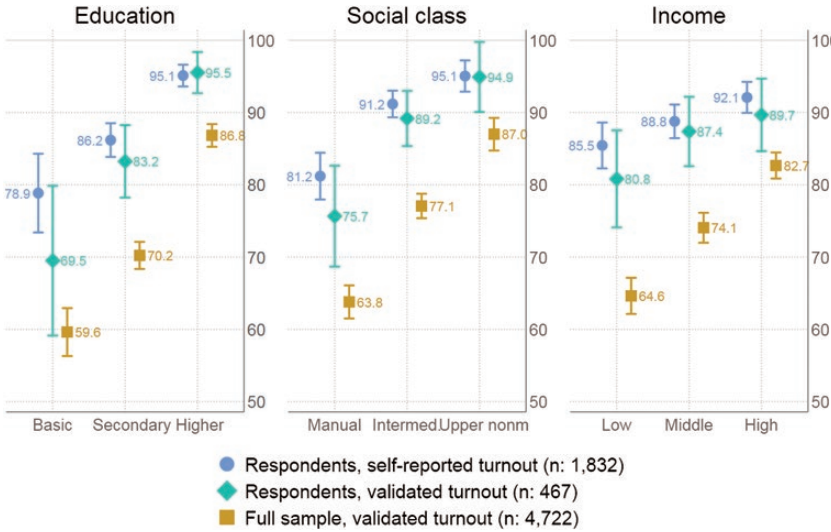

Figure 2. Voter turnout (%) by education, social class, and income using different measurement strategies. Predictions from logit models, adjusted using three cubic splines for age; 95 percent confidence intervals were calculated using the delta method.

for income. Corresponding register-based estimates from the full sample were much larger: 27.2 percent (= 86.8 – 59.6) for education, 23.2 percent (= 87.0 – 63.8) for social class, and 18.1 percent (= 82.7 – 64.6) for income. This means that register-based differences are roughly two-thirds larger for education and social class and 170 percent larger for income compared to self-reported survey estimates.

Although probability differences are interesting as such, the next obvious question is how much this underestimation matters. To assess the substantial importance of it, the risk ratios in turnout between various socioeconomic groups may be the most interesting measure, since they directly capture the over/underrepresentation of different groups in parliamentary decision-making. This is especially true since turnout gaps have more relevance when the overall turnout is lower, when an absolute gap of the same size in turnout between two population groups means a larger difference in political voice (see Verba, Schlozman, and Brady 1995) between them.[8] When using self-reported voting information obtained from the survey respondents, the corresponding

8. The logic behind our approach is quite similar to that of "representation ratios/scales" used by Leighley and Nagler (2013), Rosenstone and Hansen (2003), and Verba, Schlozman, and Brady (1995), although in our case we had the possibility to calculate such representativeness in a more straightforward manner.

risk ratios were 1.21 (= 95.1/78.9) for education, 1.17 (= 95.1/81.2) for social class, and 1.08 (= 92.1/85.5) for income. When using register-based information on the full sample, the risk ratio of turnout between the highest and lowest groups was 1.46 (= 86.8/59.6) for education, 1.36 (= 87.0/63.8) for social class, and 1.28 (= 82.7/64.6) for income. This means that the register-based measure of bias in parliamentary representation is roughly twofold for education and social class, and for income more than threefold, compared to what was obtained from the survey.[9]

SENSITIVITY/REPLICATION ANALYSES

This section assesses the external validity of our study, namely whether the main results are valid beyond this specific survey, using three different analyses. First, we compared the register-based information of our base sample to information based on full voting registers (24.2 percent of the entire individual-level electorate). Table A2 in the online appendix shows that the estimates were almost exactly the same between datasets, implying that our baseline sample represents Finnish voting patterns well.

Second, we compared the distributions of socioeconomic factors and some other socio-demographic factors—namely age, marital status, family size, and main type of economic activity—between those for whom validated voting was available and our baseline random sample, and those for whom information on voting was available (i.e., living in the wards covered by registers and eligible to vote). The analysis does not indicate serious concerns that those for whom there is background information would be a biased population with respect to any of these variables. Comparing the distributions of those for whom register-based voting is available to those without such information in the baseline sample (columns 3 and 4 in table A3, online appendix) shows that all categories pertaining to each of the variables are within 2.5 percentage points of each other, with the exception of education. A 4.4 points larger share of those with secondary educational qualifications was observed, and correspondingly a 3.4 points smaller share of those with basic qualifications. In addition, the overall turnout among all those with validated information, 69.7 percent (not applying the age restriction as in the main analysis), was very

9. When inspecting odds ratios, in turn, the lower overall turnout (while still over 50 percent) results in a lower odds ratio for an equal probability difference between groups than in a high-turnout context. When using the same register-based adjusted probabilities presented in figure 2, on self-reported turnout probabilities obtained from survey respondents, the odds ratio of turnout between the highest and lowest groups would be 5.2 (= 19.41/3.74) for education, 4.5 (= 19.41/4.32) for social class, and 2.0 (= 11.66/5.90) for income. Register-based probabilities from the full sample give corresponding odds ratios of 4.5 (= 6.58/1.48) for education, 3.8 (= 6.69/1.76) for social class, and 2.6 (= 4.78/1.82) for income. However, unlike risk ratios, odds ratios give no straightforward interpretation of representativeness arising from these differences.

close to the official turnout rate for the 2015 parliamentary elections for those residing in Finland, 70.1 percent.

Third, although not the main focus of this study, the question of how our results apply in a multivariate setting is nevertheless interesting. This is because multivariate models are typically used when studying the determinants of turnout with survey data. Figure A4 replicates the results shown in figure 2 in a multivariate setting, where socioeconomic factors were adjusted for each other as well as other variables with similar categorization as presented in table A3 in the online appendix (age, marital status, family size, and main type of economic activity). The underestimation of socioeconomic turnout differences among survey respondents persists also in the multivariate model. In this multivariate model, conventional survey-estimated differences between the highest and lowest categories are 12.7 percent (= 93.9 – 81.2) for education, 6.2 percent (= 91.8 – 85.6) for social class, and 1.5 percent (= 89.4 – 87.9) for income. Corresponding register-based estimates are 19.1 percent (= 83.4 – 64.3) for education, 9.5 percent (= 79.5 – 70.0) for social class, and 8.4 percent (= 78.4 – 70.0) for income.

Fourth, we analyzed the extent to which the socioeconomic turnout differences among the survey respondents varied in terms of whether their position was measured by relying on self-reported or register-based information. The two left panels of table A5 in the online appendix show that there was little difference between turnout estimates for education and income in this respect. However, the difference between manual and upper nonmanual classes was somewhat more pronounced when using register-based measurements.

Fifth, table A5 compares our results from the Health and Political Engagement (H&E) Survey to two other surveys, namely the 2015 Finnish National Election Study (FNES), which was conducted just after the 2015 parliamentary elections (April 24–July 7, 2015) by face-to-face interview, and the Finnish subset of round 8 of the European Social Survey (ESS), which was collected between September 15, 2016, and March 8, 2017. Table A5 shows that, overall, the measures used for the H&E and FNES surveys were in good agreement. There was a slightly larger gap in turnout between the income tertiles in the H&E survey and a slightly larger gap between social classes in the FNES survey. On the other hand, the ESS yielded larger estimates of the turnout differences between socioeconomic groups than did the other two surveys. However, the gaps in education and social class with respect to turnout were also smaller in the ESS dataset compared to the register-based information that was presented in table 2 in the main analysis (as well as figure 2 and online appendix table A2). In terms of income, the differences were of roughly equal size between the ESS and register-based data. This can partly be explained by the fact that the ESS dataset did not use a continuous income measure; rather, income was measured across 10 categories. Dividing these findings into tertiles as precisely as the data will allow results in the middle "tertile" comprising 39 percent of the respondents, which leads to the

conclusion that the lowest and highest income tertiles are somewhat more polarized groups than in the other datasets.

Discussion

This study has contributed to existing knowledge on the misestimation of voter turnout in surveys by using an exceptional dataset that includes information on both respondents' and nonrespondents' socioeconomic positions and voter turnout. We demonstrated that survey-based estimates of turnout are inflated due to both overreporting of voting and the fact that survey respondents are more likely to vote than nonrespondents. Approximately 30 percent of the validated nonvoters reported that they had voted, which is close to estimates obtained in previous studies (Bernstein, Chadha, and Montjoy 2001; Karp and Brockington 2005; Katz and Katz 2010; Ansolabehere and Hersh 2012). However, selection of the respondents was a far more serious concern, overestimating turnout by 12 percentage points relative to four points contributed by misreporting. This observation is in line with previous studies (Berent, Krosnick, and Lupia 2016; Sciarini and Goldberg 2016, 2017).

Bias related to respondent selection leads to not only an underestimation of the overall turnout level, but also the tendency to underestimate the turnout gaps between socioeconomic groups. One possible explanation for this is that the difference in terms of political interest and a sense of civic duty may be greater among respondents and nonrespondents from disadvantaged socioeconomic positions than among those respondents and nonrespondents from more advantaged socioeconomic positions. Unfortunately, since a register-based indicator of political interest or civic duty does not exist, this hypothesis could not be directly tested here. However, the observation made in the previous literature (and also tentatively with our data), that a social desirability bias in voting is stronger among those in more advantaged positions, indirectly speaks to the different levels of social pressure and expectations regarding civic duty.

Among the different (non)response groups, the overall turnout was second lowest and socioeconomic turnout differences second largest among those that were contacted and refused to answer, and lowest turnout and largest socioeconomic differences were found among those whose phone number could not be obtained. Thus, important phases of the respondent selection process occur both at the moment of contact between the survey interviewer and interviewee and before it. The latter problem is related to the coverage error of survey samples (e.g., Groves et al. 2009, ch. 3).

We observed somewhat larger socioeconomic differences in turnout when using validated voting relative to self-reported voting among respondents, despite the fact that nonvoters in more advanced socioeconomic positions were generally more likely to overreport voting. This, perhaps counterintuitive,

observation can be explained by an "opportunity effect," namely that the level of turnout of those in the most-advantaged groups is too high for overreporting to have much absolute effect (cf. Deufel and Kedar 2010; Sciarini and Goldberg 2016).

In addition to systematic error due to social desirability, the self-reported voting information potentially includes more random error, which tends to bias regression coefficients toward zero (Madger and Hughes 1997). However, only four underreporters of voting observed in our data suggest that the extent of the random error is likely to be modest.

In order to minimize the misestimation of turnout, it would be tempting to specifically target groups with the largest socioeconomic differences in voting for more answers. Unfortunately, such groups are also arguably the most difficult to include in a survey, as they include persons who were contacted but refused to participate and those for whom a phone number was not available. The selection bias caused in particular by the noncoverage of the phone number information is very difficult to address, at least in telephone surveys. In addition, efforts at minimizing overreporting among survey respondents are of course warranted. We recommend using validated turnout information, if available. If not, the approaches developed in previous studies to decrease overreporting through, for example, the way questions are worded and the inclusion of introductory comments or response alternatives (Duff et al. 2007; Hanmer, Banks, and White 2014; Morin-Chassé et al. 2017) can be used.

METHODOLOGICAL CONSIDERATIONS

This study has its limitations. A relatively small number of respondents with register-based information on voting leads to reduced statistical power in some analyses. However, our general conclusion that register-based estimates of socioeconomic differences in turnout are larger than survey-based estimates is less likely to be affected by this limitation, since self-reported information on voting was available for all except five respondents. In addition, such a finding was further validated in the sensitivity analysis section by comparing the estimates from different surveys with estimates from a full voting register.

We have used the register-based information of the baseline sample as a rare "gold standard" measure (see Biemer 2010) of socioeconomic voting patterns. That is, we assume that it contains accurate information that is generalizable to the entire Finnish electorate. In the following, we shall address the validity of this assumption through three points: 1) overall sample coverage, 2) coverage of the electronic voting registers, and 3) the accuracy of the validated information. First, we were able to observe the coverage error arising from missing information on phone numbers or native language, and assessing these sources of errors was an explicit part of our analysis. However, we could not directly observe whether our baseline sample represents the overall Finnish electorate

residing in Finland. Although no source of information is perfect, Nordic-style population registers, from which the original sample was drawn, are usually considered to have exceptionally reliable coverage of the total population and used as a benchmark of high-quality censuses (Official Statistics of Finland 2015; Skinner 2018). Therefore, we can be reasonably confident that the sample coverage is not a major problem.

The second possible limitation arises from the fact that those persons for whom the register-based information on voting is available did not constitute a random sample of the Finnish electorate. Municipalities could choose whether or not to employ these registers and in which electoral wards they were used. However, there is no individual-level self-selection measure at play here, and consequently, major biases in the relationship between socioeconomic factors and voting should not be expected. Our robustness checks further indicated no major biases in the background variable distributions of those for whom register-based voting was available and those for whom it was not available.

Third, the errors in the voting registry or its linkage, although they can never be entirely ruled out, are unlikely to be of concern. Here, we argue that our voting data is of considerably better quality than in many traditional validations, where an exact linkage using personal identity codes was not available. As a pragmatic argument, electronic voting registers must work in real-life situations, since their primary function is the actual administration of polling places. The personal identity codes have existed in their current form in Finland for almost 50 years, and they are used for almost all administrative information collection purposes. There is also a long tradition of research cooperation and data-sharing among government agencies and with Finnish academia. As an example, the correct linkage of the individuals has been around 99.5 percent in the Finnish Hospital Discharge Register during the most recent decades (Sund 2012).

GENERALIZABILITY

Although directly comparable results with regard to our second research question have apparently not been published, we believe our findings hold also in other countries. First, our estimates of the overall bias in turnout are similar to those observed, for example, in Switzerland (Sciarini and Goldberg 2016, 2017) and elsewhere (Selb and Munzert 2013). Second, the nonresponse rates in Finnish surveys are of a similar magnitude as in other countries (e.g., Stoop 2012), and the relationship between response rates and response bias is far from straightforward (Lin and Schaeffer 1995; Groves 2006; Groves and Peytcheva 2008; Selb and Munzert 2013). Third, the overall turnout rate can be an issue. If we assume that the share of overreporters among all nonvoters is constant, the importance of overreporting relative to respondent selection should be higher in lower-turnout contexts. However, previous empirical evidence is not in line with such an assumption. Instead, overreporting bias tends to be stronger in higher-turnout contexts, possibly due to stronger desirability norms (Karp and

Brockington 2005; Selb and Munzert 2013), whereas selection bias is lower in high-turnout contexts (Selb and Munzert 2013). Fourth, as speculated by Karp and Brockington (2005), local cultural norms may partly explain differing levels of overreporting. At least stereotypically, Finnish society has been seen as placing a high value on honesty (Isotalus 2009, p. 18). However, in our results, the share of overreporters relative to all nonvoters was roughly in line with the numbers obtained in many previous studies from other countries (Karp and Brockington 2005; Katz and Katz 2010; Ansolabehere and Hersh 2012).

Overall, we have no specific reason to believe that our main results cannot be generalized to other contexts. Nevertheless, the generalizability should be directly tested if an opportunity for it arises in some other national context.

Conclusion

In our register-based data, socioeconomic differences in the probability of voting were at least two-thirds larger than those obtained when using standard survey techniques. These results challenge the claim that the literature on political participation generally overestimates socioeconomic turnout biases (Bernstein, Chadha, and Montjoy 2001; Ansolabehere and Hersh 2012). In contrast, our results imply that socioeconomic inequality in electoral participation—and hence, in political voice—is a more pressing social problem than conventionally thought.

Supplementary Data

Supplementary data are freely available at *Public Opinion Quarterly* online.

References

Abelson, Robert P., Elizabeth F. Loftus, and Anthony G. Greenwald. 1992. "Attempts to Improve the Accuracy of Self-Reports of Voting." In *Questions about Questions*, edited by Judith M. Tanur, 138–53. New York: Russell Sage Foundation.

American Association for Public Opinion Research. 2016. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys,* 9th ed. American Association for Public Opinion Research. Available at https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf.

Ansolabehere, Stephen, and Eitan Hersh. 2012. "Validation: What Big Data Reveal About Survey Misreporting and the Real Electorate." *Political Analysis* 20:437–59.

Belli, Robert F., Sean E. Moore, and John VanHoewyk. 2006. "An Experimental Comparison of Question Forms Used to Reduce Vote Overreporting." *Electoral Studies* 25:751–59.

Belli, Robert F., Michael W. Traugott, and Matthew N. Beckmann. 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* 17:479–98.

Berent, Matthew K., Jon A. Krosnick, and Arthur Lupia. 2016. "Measuring Voter Registration and Turnout in Surveys: Do Official Government Records Yield More Accurate Assessments?" *Public Opinion Quarterly* 80:597–621.

Bernstein, Robert, Anita Chadha, and Robert Montjoy. 2001. "Overreporting Voting: Why It Happens and Why It Matters." *Public Opinion Quarterly* 65:22–44.

Bhatti, Yosef, Kasper M. Hansen, and Hanna Wass. 2012. "The Relationship Between Age and Turnout: A Roller-Coaster Ride." *Electoral Studies* 31:588–93.

Biemer, Paul P. 2010. "Total Survey Error: Design, Implementation, and Evaluation." *Public Opinion Quarterly* 74:817–48.

Bills, David B. 2016. "Congested Credentials: The Material and Positional Economies of Schooling." *Research in Social Stratification and Mobility* 43:65–70.

Burden, Barry C. 2000. "Voter Turnout and the National Election Studies." *Political Analysis* 8:389–98.

Clausen, Aage R. 1968. "Response Validity: Vote Report." *Public Opinion Quarterly* 32:588–606.

Deufel, Benjamin J., and Orit Kedar. 2010. "Race and Turnout in U.S. Elections Exposing Hidden Effects." *Public Opinion Quarterly* 74:286–318.

Duff, Brian, Michael J. Hanmer, Won-Ho Park, and Ismail K. White. 2007. "Good Excuses: Understanding Who Votes with an Improved Turnout Question." *Public Opinion Quarterly* 71:67–90.

Durrleman, Sylvain, and Richard Simon. 1989. "Flexible Regression Models with Cubic Splines." *Statistics in Medicine* 8:551–61.

Erikson, Robert, and John Goldthorpe. 1992. *The Constant Flux: A Study of Class Mobility in Industrial Societies*. London: Clarendon Press.

Gallego, Aina. 2007. "Unequal Political Participation in Europe." *International Journal of Sociology* 37:10–25.

Groves, Robert M. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *Public Opinion Quarterly* 70:646–75.

Groves, Robert M., Floyd J. Fowler, Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau. 2009. *Survey Methodology,* 2nd ed. Hoboken, NJ: Wiley & Sons.

Groves, Robert M., and Lars Lyberg. 2010. "Total Survey Error: Past, Present, and Future." *Public Opinion Quarterly* 74:849–79.

Groves, Robert M., and Emilia Peytcheva. 2008. "The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis." *Public Opinion Quarterly* 72:167–89.

Groves, Robert M., Eleanor Singer, and Amy Corning. 2000. "Leverage-Saliency Theory of Survey Participation: Description and an Illustration." *Public Opinion Quarterly* 64:299–308.

Hanmer, Michael J., Antoine J. Banks, and Ismail K. White. 2014. "Experiments to Reduce the Over-Reporting of Voting: A Pipeline to the Truth." *Political Analysis* 22:130–41.

Hanmer, Michael J., and Kerem Ozan Kalkan. 2013. "Behind the Curve: Clarifying the Best Approach to Calculating Predicted Probabilities and Marginal Effects from Limited Dependent Variable Models." *American Journal of Political Science* 57:263–77.

Hill, Kim Quaile, and Patricia A. Hurley. 1984. "Nonvoters in Voters' Clothing: The Impact of Voting Behavior Misreporting on Voting Behavior Research." *Social Science Quarterly* 65:199–206.

Hillygus, D. Sunshine. 2005. "The Missing Link: Exploring the Relationship between Higher Education and Political Engagement." *Political Behavior* 27:25–47.

Isotalus, Eija. 2009. "White Lies or Relative Truth? Cultural Considerations of the Finnish Concept of Honesty and Intercultural Learning." In *Speech Culture in Finland*, edited by Richard Wilkins and Pekka Isotalus, 17–41. Lanham, MD: University Press of America.

Jackson, Robert A. 1995. "Clarifying the Relationship between Education and Turnout." *American Politics Quarterly* 23:279–99.

Karp, Jeffrey A., and David Brockington. 2005. "Social Desirability and Response Validity: A Comparative Analysis of Overreporting Voter Turnout in Five Countries." *Journal of Politics* 67:825–40.

Katz, Jonathan N., and Gabriel Katz. 2010. "Correcting for Survey Misreports Using Auxiliary Information with an Application to Estimating Turnout." *American Journal of Political Science* 54:815–35.

Keeter, Scott, Carolyn Miller, Andrew Kohut, Robert M. Groves, and Stanley Presser. 2000. "Consequences of Reducing Nonresponse in a National Telephone Survey." *Public Opinion Quarterly* 64:125–48.

Lahtinen, Hannu, Mikko Mattila, Hanna Wass, and Pekka Martikainen. 2017. "Explaining Social Class Inequality in Turnout: The Contribution of Money and Health." *Scandinavian Political Studies* 40:388–410.

Leighley, Jan E., and Jonathan Nagler. 1992. "Socioeconomic Class Bias in Turnout, 1964–1988: The Voters Remain the Same." *American Political Science Review* 86:725–36.

———. 2013. *Who Votes Now? Demographics, Issues, Inequality, and Turnout in the United States*. Princeton, NJ: Princeton University Press.

Lin, I.-Fen, and Nora Cate Schaeffer. 1995. "Using Survey Participants to Estimate the Impact of Nonparticipation." *Public Opinion Quarterly* 59:236–58.

Magder, Laurence S., and James P. Hughes. 1997. "Logistic Regression When the Outcome Is Measured with Uncertainty." *American Journal of Epidemiology* 146:195–203.

Martikainen, Pekka, Tuomo Martikainen, and Hanna Wass. 2005. "The Effect of Socioeconomic Factors on Voter Turnout in Finland: A Register-Based Study of 2.9 Million Voters." *European Journal of Political Research* 44:645–69.

McDonald, Michael P. 2007. "The True Electorate: A Cross-Validation of Voter Registration Files and Election Survey Demographics." *Public Opinion Quarterly* 71:588–602.

McDonald, Jared A., Zachary A. Scott, and Michael J. Hanmer. 2017. "Using Self-Prophecy to Combat Vote Overreporting on Public Opinion Surveys." *Electoral Studies* 50:137–41.

Morin-Chassé, Alexandre, Damien Bol, Laura B. Stephenson, and Simon Labbé St-Vincent. 2017. "How to Survey About Electoral Turnout? The Efficacy of the Face-Saving Response Items in 19 Different Contexts." *Political Science Research and Methods* 5:575–84.

Official Statistics of Finland: Population Structure. 2015. *Quality Description: Population Structure 2015*. Helsinki: Statistics Finland. Available at http://www.stat.fi/til/vaerak/2015/01/vaerak_2015_01_2016-09-23_laa_001_en.html, accessed August 30, 2018.

Parry, Hugh J., and Helen M. Crossley. 1950. "Validity of Responses to Survey Questions." *Public Opinion Quarterly* 14:61–80.

Persson, Mikael. 2015. "Education and Political Participation." *British Journal of Political Science* 45:689–703.

Rosenstone, Steven J., and John M. Hansen. 2003. *Mobilization, Participation, and Democracy in America.* New York: Longman.

Schmeets, Hans. 2010. "Increasing Response Rates and the Consequences in the Dutch Parliamentary Election Study 2006." *Field Methods* 22:391–412.

Sciarini, Pascal, and Andreas C. Goldberg. 2016. "Turnout Bias in Postelection Surveys: Political Involvement, Survey Participation, and Vote Overreporting." *Journal of Survey Statistics and Methodology* 4:110–37.

———. 2017. "Lost on the Way: Nonresponse and Its Influence on Turnout Bias in Postelection Surveys." *International Journal of Public Opinion Research* 29:291–315.

Selb, Peter, and Simon Munzert. 2013. "Voter Overrepresentation, Vote Misreporting, and Turnout Bias in Postelection Surveys." *Electoral Studies* 32:186–96.

Sigelman, Lee. 1982. "The Nonvoting Voter in Voting Research." *American Journal of Political Science* 26:47–56.

Silver, Brian D., Barbara A. Anderson, and Paul R. Abramson. 1986. "Who Overreports Voting?" *American Political Science Review* 80:613–24.

Skinner, Chris. 2018. "Issues and Challenges in Census Taking." *Annual Review of Statistics and Its Application* 5:49–63.

Smets, Kaat, and Carolien van Ham. 2013. "The Embarrassment of Riches? A Meta-Analysis of Individual-Level Research on Voter Turnout." *Electoral Studies* 32:344–59.

Stoop, Inke. 2012. "Unit Non-Response Due to Refusal." In *Handbook of Survey Methodology for the Social Sciences*, edited by Lior Gideon, 121–47. New York: Springer.

Sund, Reijo. 2012. "Quality of the Finnish Hospital Discharge Register: A Systematic Review." *Scandinavian Journal of Public Health* 40:505–15.

Traugott, Michael W. 2007. "Validation Studies." In *The SAGE Handbook of Public Opinion Research*, edited by Wolfgang Donsbach and Michael W. Traugott, 408–16. London: Sage.

Traugott, Michael W., and John P. Katosh. 1979. "Response Validity in Surveys of Voting Behavior." *Public Opinion Quarterly* 43:359–77.

Verba, Sidney, Kay Lehman Schlozman, and Henry Brady. 1995. *Voice and Equality. Civic Voluntarism in American Politics.* Cambridge, MA: Harvard University Press.

Wolfinger, Raymond E., and Steven J. Rosenstone. 1980. *Who Votes?* New Haven, CT: Yale University Press.