Voter overrepresentation, vote misreporting, and turnout bias in postelection surveys

Peter Selb [a,b,*], Simon Munzert [a,b,1]

[a] *Department of Politics and Public Administration, University of Konstanz, P.O. Box D-85, 78457 Konstanz, Germany*
[b] *Center for Quantitative Methods and Survey Research, University of Konstanz, P.O. Box D-85, 78457 Konstanz, Germany*

ABSTRACT

Figures from postelection surveys often grossly overestimate election turnout. Two distinct phenomena are responsible for this gap: overrepresentation of actual voters and vote misreporting by actual nonvoters among survey respondents. Previous accounts of turnout bias are inconclusive in that they either focus on a single component, or fail to separate between the two. In this paper, we formally decompose turnout bias in election surveys into its constituent parts, assess their empirical prevalence and heterogeneity using an extensive collection of 49 vote validation studies from six countries, and employ Bayesian meta regression techniques to account for cross study differences. Our results indicate that both election and survey characteristics such as actual voter turnout and survey response rates differentially affect the components of turnout bias. We conclude with a discussion of the threats and potentials of our findings for survey based comparative electoral research.

1. Introduction

Rarely do survey data confront reality as bluntly as in comparing estimates of voter turnout from postelection polls to official statistics. Survey based figures commonly exceed official turnout by large margins, a phenomenon frequently dubbed 'vote overreporting' in the pertinent literature. Fig. 1 plots reported sample turnout rates from 130 postelection studies in 43 countries versus official election turnout. All but two surveys in fact overreport turnout,[2] the average difference between sample and official turnout is 13%, and the maximum observed discrepancy amounts to a staggering 42% at the 2005 national elections in Albania.

Vote overreporting certainly ranks among the big annoyances of survey based electoral research, as it threatens both the general credibility of survey data and the validity of conclusions drawn from studies of individual political behavior and attitudes (Bernstein et al., 2001; Brehm, 1993; Cassel, 2003; Hill and Hurley, 1984; Jones, 2008; Karp and Brockington, 2005; Katosh and Traugott, 1981; Presser and Traugott, 1992; Sigelman, 1982; Tittle and Hill, 1967).

While amply used in the literature, the term vote overreporting seems a misnomer, as it muddles two distinct causes of turnout bias: *over*representation of actual voters due to the disproportionate self selection of politi cally involved citizens into realized election survey samples (e.g., Brehm, 1993; Jackman, 1999; Voogt and Saris, 2003), and vote mis*reporting* by actual nonvoters among survey respondents due to pressures of social desirability and other psychological processes (e.g., Belli et al., 1999; Presser, 1990). Terminological quibbles aside, it is

---

* Corresponding author. Department of Politics and Public Administration, University of Konstanz, P.O. Box D-85, 78457 Konstanz, Germany. Tel.: +49 7531 882321; fax: +49 7531 884159.

E-mail addresses: peter.selb@uni-konstanz.de (P. Selb), simon.munzert@uni-konstanz.de (S. Munzert).

[1] Tel.: +49 7531 883087; fax: +49 7531 884159.

[2] The most conspicuous case of vote *under*reporting (the circle markedly below the diagonal line in Fig. 1) occurs with the 2008 parliamentary election in Belarus which, according to official observers, clearly fell short of OSCE commitments for democratic elections (see OSCE, 2008). It is a matter of speculation whether such massive vote underreporting may not be a consequence of polished official turnout figures.
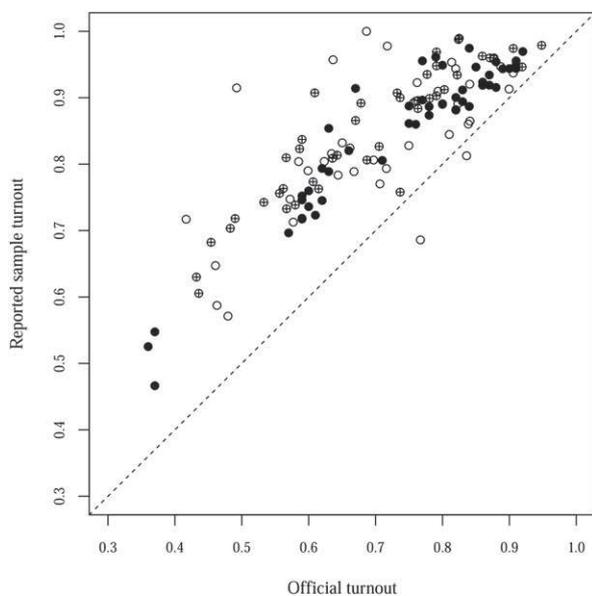
Fig. 1. Reported sample turnout rates from 130 postelection surveys versus official turnout. Data are taken from Modules 1–3 of the *Comparative Study of Electoral Systems (CSES)*, and from a collection of election surveys for which vote validation studies (VVS) are available. For more detailed information, see the Appendix to this paper.

important to keep these two concepts separate, first, since the methodological issues arising from each differ selection bias in the former (see Achen, 1986; Heckman, 1979), and misclassification bias in the latter case (see Hausman et al., 1998) and thus call for different strategies when modeling, for instance, individual turnout propensities using survey data (e.g. Brehm, 1993, 1999; Deufel and Kedar, 2010; Katz and Katz, 2010). Second, voter over representation and vote misreporting may be linked through common factors, though in inconsistent ways a possibility that renders any attempt to understand the phenomenon 'from above' essentially hopeless (Tourangeau et al., 2010).

To illustrate, consider the lively exchange between Burden (2000, 2003), McDonald (2003) and Martinez (2003) on why the turnout gap in the *American National Election Studies (ANES)* has more than doubled over the past decades, from 12% in the 1950s to 25% in the 1990s. Burden (2000) attributes rising turnout bias in ANES surveys to declining response rates that supposedly exacerbate the overrepresentation of voters in realized survey samples. To empirically support his conjecture, he regresses compound turnout gaps in 11 ANES surveys on response rates and indeed finds a significantly negative relationship. McDonald (2003) criticizes Burden's (2000) analysis on measurement as well as theoretical grounds, and suggests that declining official turnout is the chief culprit since, with decreased turnout, there is a higher share of nonvoters among survey respondents 'at risk' of misreporting their true voting behavior. Though equipped with a more appropriate estimate of official turnout (see McDonald and Popkin, 2001), McDonald (2003) still considers compound turnout gaps, as does Martinez (2003) who holds voter

overrepresentation due to panel conditioning responsible for some particularly pronounced discrepancies between sample and official turnout. As plausible as these arguments may appear, there is simply no way to judge their empirical merits unless one examines the components of turnout bias separately. For instance, it may well be that increased response rates reduce voter overrepresentation yet at the same time, also draw more embarrassed nonvoters into the sample who are more likely to misreport (Tourangeau and Yan, 2007). Likewise, higher turnout evidently reduces the population of nonvoters at risk of misreporting, but may also put up social desirability pressures, so that the individual propensity to misreport among the (fewer) nonvoters participating in election surveys increases (Karp and Brockington, 2005). Looking at compound turnout gaps alone simply will not provide any conclusive answer.

Vote validation studies (VVS) that examine official records to verify self reported votes in election surveys offer invaluable information about both sources of bias (see Traugott, 2008, for an overview). As of yet, however, such studies have almost exclusively been utilized in isolation to determine the magnitude and *individual* correlates of misreporting at a given election[3] an approach ill suited for shedding light on the puzzling observation that the gap between reported and official turnout varies so widely across elections and surveys within a single country (as observed by Burden, McDonald and Martinez in the ANES case), but also across countries. Reconsidering Fig. 1, the standard deviation of the differences between reported and official turnout rates across election studies is as high as 7.1%. Accounting for this enormous scatter is the primary aim of the present study. To this end, we first formally decompose the turnout gap in postelection surveys into its constituent parts, and then assess their empirical prevalence and heterogeneity across 49 VVS from six countries (marked with solid circles in Fig. 1). Next, we identify election and survey characteristics that may account for eventual cross study differences, and test their predictive value using Bayesian meta regression techniques that synthesize information from all the VVS. Subsequently, we will reassemble the components, and use our and alternative models' estimates to predict turnout gaps in a 'fresh' collection of postelection surveys for which the relevant election and study level features are known, but vote validation data are unavailable (marked with crossed circles in Fig. 1). The final section concludes with a discussion of the threats and potentials of our findings for survey based electoral research, with a particular emphasis on comparative approaches.

---

[3] Obviously, VVS are silent about individual correlates of voter over-representation, as such studies only allow for observing respondents', but not nonrespondents' characteristics. Other paradata generated as a by-product of the survey data collection process (e.g., contact histories, interviewer observations of should-be respondents refusing to participate in the survey) are needed to determine individual correlates of voter overrepresentation (e.g., Brehm, 1993; Stoop, 2005). However, as we demonstrate in the subsequent section, VVS data can be used to gauge at least the overall magnitude of voter overrepresentation in a study.

2. Decomposing the turnout gap

In this section, we delineate our central concepts and their linkages. In doing so, we heavily draw on Jackman's (1999) treatment of nonresponse and misreporting bias in survey estimates of proportions, and Manski's (1995, ch. 4) approach to identification in response based samples. By way of illustration, we initially set up the problem as a simple model of nested sets (see Fig. 2). The first ply divides the *original survey sample* (as opposed to the *realized sample*) into two strata: the share of individual sampling units that are actually interviewed (i.e., the respondents), S, and the share of nonrespondents, $\tilde{S} = 1 - S$, who cannot be reached during the field period or, perhaps more importantly, refuse to participate in the survey. Among both strata, there is a substratum of actual voters, T, and nonvoters, $\tilde{T} = 1 - T$. Finally, actual voters and nonvoters among survey respondents each divide into sub substrata, one of which containing those who report that they voted, R, and the other those who admittedly abstained from the voting booth, $\tilde{R} = 1 - R$. Evidently, R is the *reported sample turnout* rate, and *turnout gap*, G, can be written as

$$G = R - P, \tag{1}$$

where P is the population parameter of interest, that is, the *official turnout* rate. In the discussion to follow, we assume (1) that official turnout is appropriately measured in terms of the voting eligible population (VEP) as the target population of interest (McDonald and Popkin, 2001; Holbrook and Heidbreder, 2010), (2) that the VEP also constitutes the sampling frame of the postelection surveys under consideration (see McDonald and Popkin, 2001), and (3) that survey estimates of turnout are weighted for differential selection probabilities, so that the expectation of the survey estimate will correspond to the population parameter if (a) either the realized sample equates the original sample, or nonvoters have the same propensity to participate in election surveys as voters, and (b) respondents correctly report their true voting behavior.

It is immediately obvious from our exposition that there are two kinds of misreporting in postelection surveys, both of which can be observed from vote validation data: the fraction of nonvoters who report that they voted (which we will dub *type I misreporting*),

$$M_1 = \frac{R \cap \left(\tilde{T}|S \right)}{\tilde{T}|S} \tag{2}$$

and the fraction of actual voters who erroneously indicate that they abstained from the election (*type II misreporting*),[4]
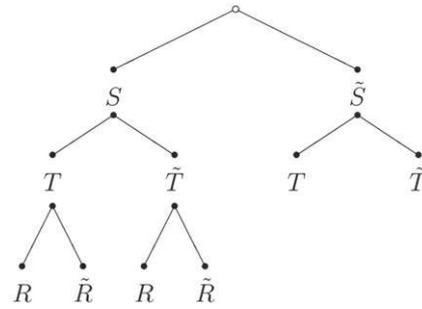
Fig. 2. Survey participation S, true voting behavior T, and reported voting behavior R: a simple model of nested strata.

$$M_2 = \frac{\tilde{R} \cap (T|S)}{T|S}. \tag{3}$$

Equipped with these definitions, we may ask how *true sample turnout*, $T|S$, and reported turnout, R, are linked. It is clear that, absent misreporting, $R = T|S$. On the other hand, with type I and type II misreporting rates included, reported sample turnout rate can be expressed as

$$R = (T|S) + M_1 \times \left(\tilde{T}|S \right) - M_2 \times (T|S) \tag{4}$$

Thus, vote misreporting among nonvoters may be partly counteracted by vote misreporting among voters in producing the reported sample turnout rate. Yet, albeit both types of misreporting may originate from common processes such as memory failure (Belli et al., 1999; Stocké and Stark, 2007), we would nevertheless expect the share of misreporting nonvoters, M_1, by far to outweigh misreporting by actual voters, M_2, primarily due to social desirability pressures (see Tourangeau et al., 2010).

Turning now to the definition of voter over representation, we may next ask on what condition the true sample turnout rate, $T|S$, would approach the population parameter, P. Simple random sampling (or other forms of random sampling that allow for correcting unequal selection probabilities) guarantees that the true turnout rate in the original sample, T, equals the population parameter, P, in the limit. However, the same will hold for $T|S$ only if $T|S$ equals $T|\tilde{S}$, since P is composed of turnout rates among respondents and nonrespondents,

$$P = (T|S) \times S + \left(T|\tilde{S} \right) \times (1 - S). \tag{5}$$

This assumption seems quite heroic, since intuition and previous findings suggest that electoral participation and survey participation are products of common underlying causes (e.g. Brehm, 1993). If so, we would expect actual voters to be overrepresented in realized election survey samples, that is, $T|S > T|\tilde{S}$. *Voter overrepresentation*, O, may thus be defined in terms of the ratio of true voter proportions among respondents and nonrespondents,

$$O = \frac{T|S}{T|\tilde{S}}, \tag{6}$$

which would assume values greater than unity if voters were indeed overrepresented among survey respondents. While both $T|S$ and $T|\tilde{S}$ are theoretical quantities that are invisible from postelection surveys alone, we may profit from VVS to directly observe $T|S$, and once again exploit the simple random sampling assumption to infer $T|S$ (see Manski, 1995, ch. 4). In particular, it follows from Equation (5) that

$$T|\tilde{S} = \frac{P - (T|S) \times S}{\tilde{S}}. \tag{7}$$

Thus, Equation (6) can be rewritten (and voter over representation be identified) as

$$O = \frac{(T|S) \times \tilde{S}}{P - (T|S) \times S}. \tag{8}$$

To illustrate, consider a scenario where official turnout $P = 0.7$, true turnout among respondents is $T|S = 0.8$, and the survey response rate is $S = 0.5$. Clearly, true turnout among nonrespondents must be $T|\tilde{S} = (0.7 - 0.8 \times 0.5)/0.5 = 0.6$ given that true turnout in the original sample approaches the population parameter P. Hence, voter overrepresentation amounts to $O = 0.8/0.6 = 1.33$ in this hypothetical case.

Having the three components of turnout bias written out that way, we may solve Equation (8) for $T|S$ and substitute this into Equation (4), so that the reported sample turnout, R, can be re expressed as

$$R = \frac{O \times P}{(O-1) \times S + 1} + M_1 \times \left(1 - \frac{O \times P}{(O-1) \times S + 1}\right) - M_2$$
$$\times \frac{O \times P}{(O-1) \times S + 1}$$
$$= \frac{(M_1 \times O - M_1) \times S + (1 - M_1 - M_2) \times O \times P + M_1}{(O-1) \times S + 1}. \tag{9}$$

Finally, by substituting R in Equation (1) by the right hand side of Equation (9), turnout gap, G, can be decom posed as

$$G = \frac{(M_1 \times O - M_1) \times S + (1 - M_1 - M_2) \times O \times P + M_1}{(O-1) \times S + 1} - P. \tag{10}$$

Note that this model of turnout bias naturally accounts for some logical constraints which are, perhaps, of minor empirical interest but nevertheless need to be taken into consideration when investigating the phenomenon. For instance, if official turnout, P, approaches 1, the share of nonvoters at risk of misreporting their voting behavior, M_1, converges toward 0 (see McDonald, 2003), and the true voter ratio among respondents versus nonrespondents draws near 1. Compound turnout bias, G, in turn, can never exceed $1 - P$, as 1 constitutes the upper bound for reported sample turnout, R. That way, our formulation allows us to focus on the empirically relevant relationships between election and study characteristics on the one hand, and the components of turnout bias on the other, while accounting for the logical restrictions.

3. Vote validation studies: a systematic review

In the previous section, we have presented definitions of turnout bias components that lend themselves to oper ationalization using data from vote validation studies (VVS). Though indispensable for our purposes, VVS are relatively rarely conducted due to data privacy and economic constraints. We were nevertheless able to gather data from 49 such studies carried out in six countries: Ireland (2 studies), Norway (12), New Zealand (5), Sweden (16), the United Kingdom (6) and the United States (8).[5] Most of these studies pertain to national elections (parlia mentary or presidential), though four Norwegian studies were implemented subsequent to local elections. See the Appendix for more detailed information about the studies.

In this section, we will use Bayesian methods to combine study estimates of vote misreporting and voter overrepresentation, and to assess the degree of heteroge neity in these estimates across VVS. Bayesian methods offer a unified modeling framework for synthesizing informa tion from different studies, to adjust for covariates through regression models (see Section 4), and to make out of sample predictions for other studies (see Section 5; Warn et al., 2002).

In notational terms, we index studies with $j = 1, 2,\ldots J$, where $J = 49$ is the number of election studies attended with a VVS component. Furthermore, we denote realized sample sizes N_j, with N_{1j} indicating the numbers of self reported nonvoters, and N_{2j} the numbers of self reported voters in study j. We designate $\tilde{N}_j$ to represent the number of nonrespondents per study (i.e., the difference between original and realized sample sizes). As opposed to our previous treatment of misreporting and over representation, we now also consider the fact that sample proportions are subject to sampling error, and denote the underlying true misreporting parameters, in line with Jackman (1999), η_{1j} and η_{2j}, and the overrepresentation parameter ω_j. All the parameters are modeled on the log scale to rid their bounds and to make them approximately normally distributed. Accordingly, the log proportion of nonvoters in study j who misreport their true voting behavior, M_{1j}, is assumed to follow a normal distribution with mean $\log(\eta_{1j})$ and known variances representing study specific sampling error, $\hat{\sigma}^2_{M_{1j}}$,[6]

$$\log(M_{1j}) \sim \text{Normal}\left(\log(\eta_{1j}), \hat{\sigma}^2_{M_{1j}}\right). \tag{11}$$

---

[5] Two other U.S. election studies, namely the 2006 and 2008 *Cooperative Congressional Election Studies*, provide vote validation information but have been excluded from our analysis. The reason is that these studies use nonprobability samples of U.S. citizens (for details, see http://projects.iq.harvard.edu/cces/). Though highly sophisticated, it seems questionable to what extent such study designs meet the assumptions we have made to identify the magnitude of voter overrepresentation from VVS data.

[6] In doing so, we assume that the vote validation process is free of error, an admittedly heroic assumption (see Presser et al., 1990; Cassel, 2004).

The Taylor approximation may be used to estimate the variance of $\log(M_{1j})$ from the sample data (see Cornell and Mulrow, 1999),[7]

$$\hat{\sigma}^2_{M_{1j}} \quad \frac{1 \quad M_{1j}}{M_{1j}} \times \frac{1}{N_{1j}}. \tag{12}$$

Having accounted for the within study variability of $\log(M_{1j})$ this way, $\log(\eta_{1j})$ is modeled as a function of a constant term (so to speak the grand propensity of nonvoters to misreport their voting behavior across studies), β_0, and a study specific deviation from that grand propensity, u_j, for which we assume a normal distribution with mean zero and variance σ_u^2,

$$\begin{aligned} \log(\eta_{1j}) \quad & \beta_0 + u_j \\ u_j &\sim \text{Normal}(0, \sigma_u^2). \end{aligned} \tag{13}$$

An analogous model of type II misreporting is set up by replacing observed and estimated shares of misreporting nonvoters, M_{1j} and η_{1j}, by the respective proportions of misreporting voters, M_{2j} and η_{2j}, and substituting N_{1j} with the number of self reported voters, N_{2j}. To model voter overrepresentation, we replace M_{1j} and η_{1j} by O_j and ω_j, and use the Taylor approximation for the variance of the log ratio of proportions instead of the formula given in Equation (12) (see Cornell and Mulrow, 1999):

$$\hat{\sigma}^2_{O_j} \quad \frac{1}{N_j \times (T_j|S_j) \times (\tilde{T}_j|S_j)} + \frac{1}{\tilde{N}_j \times (T_j|\tilde{S}_j) \times (\tilde{T}_j|\tilde{S}_j)}, \tag{14}$$

where N_j is the number of respondents in the realized sample and N_j is the number of nonrespondents.

Inference proceeds via Markov chain Monte Carlo (MCMC) methods to compute a joint posterior density for all the model parameters. We use WinBUGS to run the required computations (Lunn et al., 2000). Since the posterior distribution is produced by simulation, we can make direct inferences for the parameters of interest (i.e., η_{1j}, η_{2j}, and ω_j), instead of their log transforms. Bayesian methods require the specification of prior distributions on the model parameters. We use uninformative priors to let the data determine the parameters. In particular, we use normal priors with mean 0 and large variance for β_0, and vague uniform priors within a range of $(0, 10)$ for the cross study standard deviation, σ_u. To monitor convergence, we set up three chains with randomly chosen starting values for the parameters, each with 20,000 iterations, of which

we discard the first 10,000 before summarizing the parameters' posterior probabilities.

Fig. 3 presents medians of the posterior densities for η_{1j}, η_{2j}, and ω_j plus their 90% credible intervals. Considering first type I misreporting, we can see from the last line in column 1 that, on average,[8] a remarkable proportion of about 0.27 of the nonvoters wrongly indicate that they voted. This proportion is highly variable across VVS, ranging from approximately 0.10 in the 1970 Swedish national election study to an astonishing 0.68 in the 1985 Norwegian Storting election survey. The cross study standard deviation of these proportions is estimated 0.14. Expectedly, the results pertaining to type II misreporting are far less alerting. On average, less than 1% of the voters misreported their actual voting behavior. After all, a maximum estimated percentage of about 3.4 occurs in the ANES midterm study 1990, perhaps making it worth while to model the determinants of cross study differences in η_{2j} as well. Finally, the voter overrepresentation parameter is, on average, 1.4, that is, there are about 1.4 voters in the realized sample per voter among nonrespondents. With the possible exception of the election survey, there is no study in our review where voters did not disproportionately self select into survey samples, i.e., all the other ω_j estimates significantly exceed 1. And again, over representation ratios are highly scattered across studies, with an estimated standard deviation of 0.34. Voter over representation is strongest in the 1999 Norwegian local election study, with three times as many voters among respondents than among nonrespondents.

In sum, almost all the studies considered are affected by vote misreporting among nonvoters (but not necessarily among voters) and voter overrepresentation, although to varying degrees.

4. Contextual determinants of overrepresentation and misreporting: meta regression analyses

Which election and study features may be responsible for the considerable cross study variation in misreporting rates and overrepresentation ratios previously observed? Some authors suggest that the actual turnout rate, P, is the prime suspect. Actual voter turnout may affect the extent of the turnout gap in several ways. First, the share of nonvoters at risk of misreporting their true behavior converges to 0 as official turnout rates approach 1 (see McDonald, 2003), a logical constraint that our model formulation in Equation (10) already takes care of. In a similar vein, the true voter ratio among respondents and nonrespondents converges to 1. Under full turnout, there are simply no nonvoters to misreport their behavior, or to be underrepresented in the realized sample. The relatively small turnout gaps to be observed in Fig. 1 for surveys from high turnout elections are indicative of these limiting conditions. More interestingly, actual turnout and components of turnout bias may be linked through psychological processes. Karp and Brockington (2005) demonstrate in

---

[7] Alternatively, we could model the binomial outcomes underlying M_{1j} and the other summary measures directly in a fully Bayesian setting, thereby evading the assumption that $\hat{\sigma}^2_{M_{1j}}$ is known, although in fact it has to be estimated from the data (see Warn et al., 2002). However, given our relatively large sample sizes, differences in results should be (and, as we found in preliminary analyses, actually are) minuscule (see Thompson et al., 2001). Therefore we maintain the summary statistic models which allow us to easily include design weights to account for unequal sampling probabilities. Additional results are available on request. Modeling the binomial outcomes directly would also allow us to consider individual-level characteristics potentially affecting survey participation and misreporting. However, VVS usually do not include sampling frame information on nonrespondents, which further renders this strategy redundant.

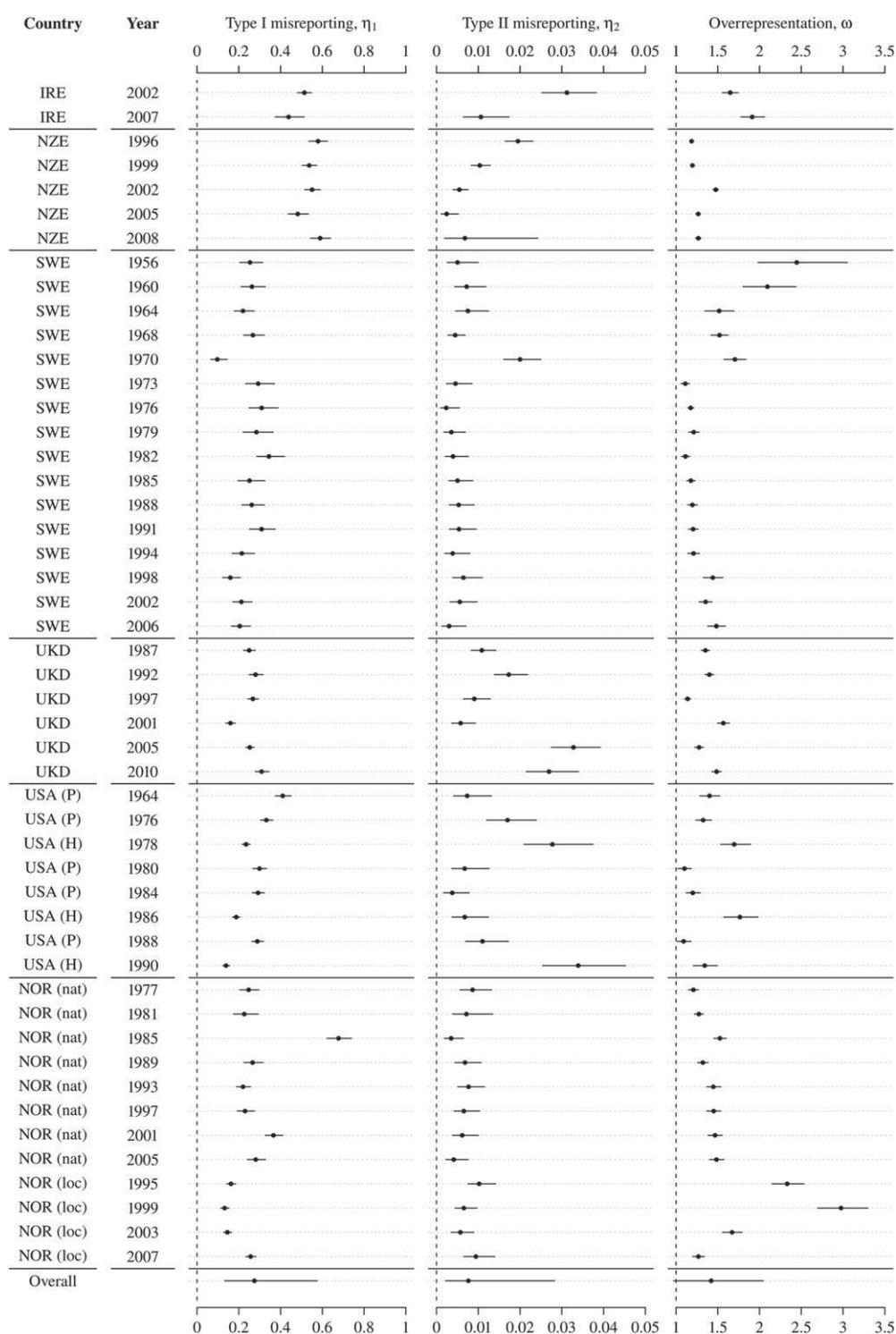[8] Overall predictions are based on the posterior density of β_0 with cross-study variance σ_u^2.

Fig. 3. Vote misreporting by actual nonvoters (η_1) and voters (η_2), and voter overrepresentation (ω) in 49 vote validation studies: medians of the bias estimates' posterior densities (points) and their 90% credible intervals (lines).

their inventive analysis of 23 VVS that higher effective election turnout increases the propensity to misreporting among actual nonvoters participating in election surveys. Accordingly, actual turnout levels set the 'descriptive social norm' that prescribes how socially desirable the act of voting is (see Gerber and Rogers, 2009).

On the other hand, study characteristics may account for differences in voter overrepresentation and misreporting.

Burden (2000) hypothesizes that voter overrepresentation decreases with higher response rates, *RR*. Implicit in this conjecture is the notion of a 'continuum of resistance' (see Groves, 2006). Accordingly, nonvoters are generally more hesitant to participate in election surveys. Increased response rates, in turn, primarily result from the survey institutions' gradual success to overcome the nonvoters' inhibitions to agree to be interviewed. This view has not gone uncontested. As Lin and Schaeffer (1995) argue, increasing response rates may, unless they reach 100%, even accentuate differences between respondents and nonrespondents. This would be the case if additional effort to increase response rates disproportionately drew more voters into the realized sample, thereby reducing the fraction of voters among nonrespondents, and thus boost ing the voter overrepresentation ratio. In any case, lower overrepresentation might turn out to be a dearly purchased success when higher response rates draw more embar rassed nonvoters into the realized sample who are more likely then to misreport (Tourangeau and Yan, 2007). Response rates vary from 0.40 to 0.95 in our collection of VVS, with a mean of 0.71.

A second study specific factor possibly affecting mis reporting behavior is the interview mode. Different inter view modes offer varying degrees of privacy to the respondent, and the absence of interviewers is often assumed to reduce social desirability pressures (for an extensive overview of studies examining mode effects for sensitive topics, see Tourangeau and Smith, 1996). Evidence on this conjecture is mixed. While Locander et al. (1976) do not find substantive differences in misreporting rates across interviewer aided and self administered survey modes (e.g., postal or online), Stocké (2007) detects significantly higher reported turnout rates among nonvoters for the former. We include a variable, *SAI*, that indicates the proportion of self administered interviews within a study. In our collection of VVS, 32 studies were conducted in an interviewer aided manner, while the remaining 17 studies at least partly employed self administered interview modes.

Some authors argue that careful question wordings may help to reduce social desirability pressures, and thus, vote misreporting (e.g., Abelson et al., 1992; Belli et al., 2006; Box Steffensmeier et al., 2000; Duff et al., 2007). Such *extended question formats, EQF*, go beyond the traditional dichotomy ('I voted in the election', 'I didn't vote in the election'), and provide extended introductory statements like 'Talking to people about the general election, we have found that a lot of people didn't manage to vote', or addi tional response categories such as 'I thought about voting but I didn't' and 'I usually vote but didn't vote this time', to make it easier for the respondents to be truthful. 29 of the 49 studies we consider employed such extended question formats.[9]

We also code a variable that covers the proportion of post election interviews that were preceded by a pre election *Panel* wave. As Martinez (2003) argues, the mere

[9] For a documentation of the question wordings used in the studies under consideration, see Table 2 of the online Appendix to this paper.

fact of being interviewed prior to the election may remind respondents of reasons why they might want to vote, and thus motivate some respondents to turn out in the election who may have abstained otherwise (see Granberg and Holmberg, 1991; Smith et al., 2003). Therefore, panel effects may be expected to increase voter over representation by nonvoter conversion. 35 of our 49 VVS include panel components, but only in three studies all the respondents were interviewed prior to the election.

To model type I misreporting (and, analogously, type II misreporting and voter overrepresentation) as a function of covariates, we simply add these variables to Equation (13). At this stage, we also include a country level random effect, $v_{k(j)}$, to account for unobserved cultural or institutional differences (or temporally stable study and election char acteristics) that potentially affect the turnout bias components,

$$
\begin{aligned}
\log(\eta_{1j}) \quad & \beta_0 + \beta_1 P_j + \beta_2 RR_j + \beta_3 SAI_j + \beta_4 EQF_j \\
& + \beta_5 Panel_j + u_j + v_{k(j)}, \\
u_j \sim & \, \mathrm{N}(0, \sigma_u^2), \\
v_{k(j)} \sim & \, \mathrm{N}(0, \sigma_v^2).
\end{aligned}
\tag{15}
$$

Fig. 4 reports the corresponding results for the three models. As expected, higher actual turnout has an ambiv alent effect on the bias components that any study of compound turnout bias simply sweeps under the carpet. While increased turnout alleviates voter over representation (see McDonald, 2003), it tends to exacer bate the problem of misreporting among nonvoters, arguably by putting up social desirability pressures (see Karp and Brockington, 2005). Interestingly, higher turnout also tends to reduce the portion of actual voters who mistakenly report that they abstained from the election. An ad hoc explanation of this finding may be that, with increased turnout, actual voters unable to recall their true voting behavior more and more consider voting as the default behavior and, in case of doubt, opt for the perceived standard. Our results also indicate that response rate is positively related to voter overrepresentation, and nega tively (though not significantly at conventional levels) to misreporting among nonvoters. Apparently, Burden's (2000) and others' speculations of relevant differences between respondents and nonrespondents automatically vanishing with increasing response rates seem overly optimistic. Just the opposite, increased response rates seem to primarily result from the interviewers' success to draw more voters into the sample (also see Groves, 2006; Lin and Schaeffer, 1995). In line with Martinez (2003), we find that voter overrepresentation increases with higher shares of respondents being interviewed in a pre election panel wave, while misreporting rates remain unaffected. Finally, interview mode and question format have negligible effects, with one possible exception: long question formats designed to relieve respondents from social desirability pressures apparently increase the rate of misreporting among actual voters.

Hardly surprising, the factors included in our models cannot fully account for the observed cross study vari ability in voter overrepresentation and vote misreporting.
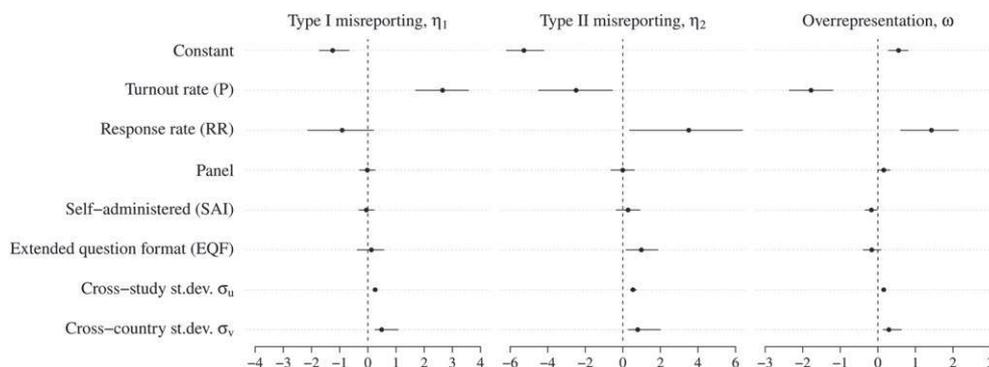
Type I misreporting, η_1 Type II misreporting, η_2 Overrepresentation, ω

Fig. 4. Bayesian meta regression results: medians of the parameter estimates' posterior densities (points) and their 90% credible intervals (lines).

Both study and country specific random effects, u_j and $v_{k(j)}$, still vary considerably across studies, as indicated by the estimates of σ_u and σ_v. This suggests that other, as yet unmodeled, study, election and country characteristics affect the components of turnout bias. It is indeed a general problem of meta analytical approaches to completely identify possible patterns of predictor variables, since (a) the number of studies that inform such analyses is usually limited (and so are the degrees of freedom), and (b) the potential for measuring relevant covariates heavily depends on the least common set of information the underlying studies document. For the time being, a hierar chical model specification including both study and country random effects therefore seems appropriate.

5. Compound turnout bias: out-of-sample predictions

How far does our model get us in understanding gaps between reported and official turnout in general, that is, in studies other than the 49 VVS closely investigated in the preceding section? To answer this question, we draw on a collection of 42 'fresh' postelection studies from 19 countries included in Modules 1 3 of the *Comparative Study of Electoral Systems* (CSES) for which the election and study level covariates previously used to predict bias components are known, but for which VVS data is unavailable.[10] The rationale behind this approach to external validation is that if our models captured the essential contextual dependencies of voter over representation and vote misreporting, then it should also be possible to accurately predict known turnout gaps based on contextual information alone. To do so, we combine covariate information from the selected CSES surveys with the posterior densities of our parameter estimates (including v_k in the case of the U.S., which is the only country for which both VVS and independent CSES data are

available) to compute study specific estimates of η_{1j}, η_{2j} and ω_j, which are then reassembled following Equation (10) to yield out of sample predictions of turnout gaps to be compared to their observed values.

As benchmarks, we also fit the compound turnout bias models according to Burden (2000), McDonald (2003) and Martinez (2003) to the 49 VVS, and use these parameter estimates to predict turnout gaps in the CSES surveys.[11]

In validation terms, one important criterion will be the distance between observed turnout bias values and their estimates expressed as the empirical root mean squared error (RMSE) of (the median of) the estimates' posterior probabilities. Clearly, smaller RMSEs indicate better point estimates. As to these estimates' uncertainty, we calculate 90% Bayesian credible intervals from the highest posterior density regions that can be immediately interpreted in terms of the probability that the true value of the estimated parameter is inside a given interval. A second validation criterion is then the *coverage probability* of the credible intervals, that is, the proportion of the time that the intervals actually contain the observed value of interest. The actual coverage probability should approximate the nominal level of 90% as closely as possible. A final validation criterion bears on the logical constraints of turnout bias discussed in Section 2. In particular, turnout gap, G_j, cannot exceed 1 P_j, since 1 constitutes the upper bound for re ported sample turnout, R (see Equation (1)). Thus, we report the relative number of times that $\hat{G}_j > 1$ P_j, which should, of course, be zero with theoretically sound estimators.

---

[10] In order to avoid irregularities as those encountered in the intro ductory section (see Footnote [1]), we focus our analysis on countries with a flawless democratic record, as indicated by Polity IV values of 10 (see http://www.systemicpeace.org/polity/polity4.htm). Out-of-sample predictions for a wider collection of countries including those with lower Polity values came out less precisely, although the relative fit of alternative models (see below) was reconfirmed. Detailed results are available on request.

[11] The Burden model is a simple bivariate linear regression of turnout gaps on survey response rates, which yields parameter estimates [stan dard errors] of $\hat{B}_j$ 0.334[0.036]$-$0.316[0.050]RR_j, with a root mean squared error (RMSE) of 0.038. The McDonald model includes both response rates and official turnout, and yields $\hat{B}_j$ 0.411[0.031]$-$0.241[0.040]$RR_j - 0.177[0.031]P_j$, RMSE 0.029, with our VVS data. Finally, Martinez' model specification includes response rates, a second-order election dummy and proportions of respondents having been interviewed in a pre-election panel wave, $\hat{B}_j$ 0.331[0.036]$-$0.313[0.048]$RR_j - 0.010[0.019]Panel_j + 0.034[0.015]$ SOE_j, RMSE 0.036. Some of these estimates are substantially different from those found by Burden, McDonald and Martinez on a sample of 11 *ANES* surveys. However, for comparative purposes, we consider it more appropriate to refit their models using the VVS data from which our own model's estimates derive.

Fig. 5 pits predicted against observed turnout gaps in the 42 selected CSES studies. Predictions are based on the four alternative models: the compound turnout bias models according to Burden (2000), McDonald (2003) and Martinez (2003) as well as our own. Eyeballing Fig. 5 already suggests that our model and McDonald's model based on official turnout and response rates fit the observed turnout gap data reasonably well, while Burden's response rate based model and Martinez' model including response rates and the second order election and panel variable perform considerably worse. On closer examina tion, our model outperforms the McDonald and the other models along all the validation criteria (see Table 1). The RMSE for our model is just 0.048 and by construction it always yields estimates within the logically possible limits. In contrast, the other models yield out of bound predic tions in 3 9 out of 42 cases, and produce RMSEs ranging from 0.053 for the McDonald model to 0.074 for Martinez'. The coverage probability is quite close to its nominal level of 90% for our model, and relatively far off the mark for the others, indicating some overconfidence. Much more important though, our model is the only one to give disaggregate insight into how election and study charac teristics relate to the components of turnout bias.

Table 1
Empirical root mean squared errors (RMSE), coverage probabilities and out-of-bounds predictions for four alternative models of turnout bias.

	Our model	Burden (2000)	McDonald (2003)	Martinez (2003)
RMSE	0.048	0.073	0.053	0.074
Coverage probability	0.857	0.643	0.690	0.643
Out of bounds	0.000	0.214	0.071	0.191

6. Discussion

In this study, we have demonstrated that voter over representation and vote misreporting vary tremendously across studies and elections, and that part of this hetero geneity emanates in a predictable manner from actual turnout and survey response rates. So why should we, as students of comparative political behavior, bother?

It is well documented for the single study case that voter overrepresentation and vote misreporting can considerably bias survey based analyses of turnout. On the one hand, survey participation has been demonstrated to be positively associated with electoral participation. That
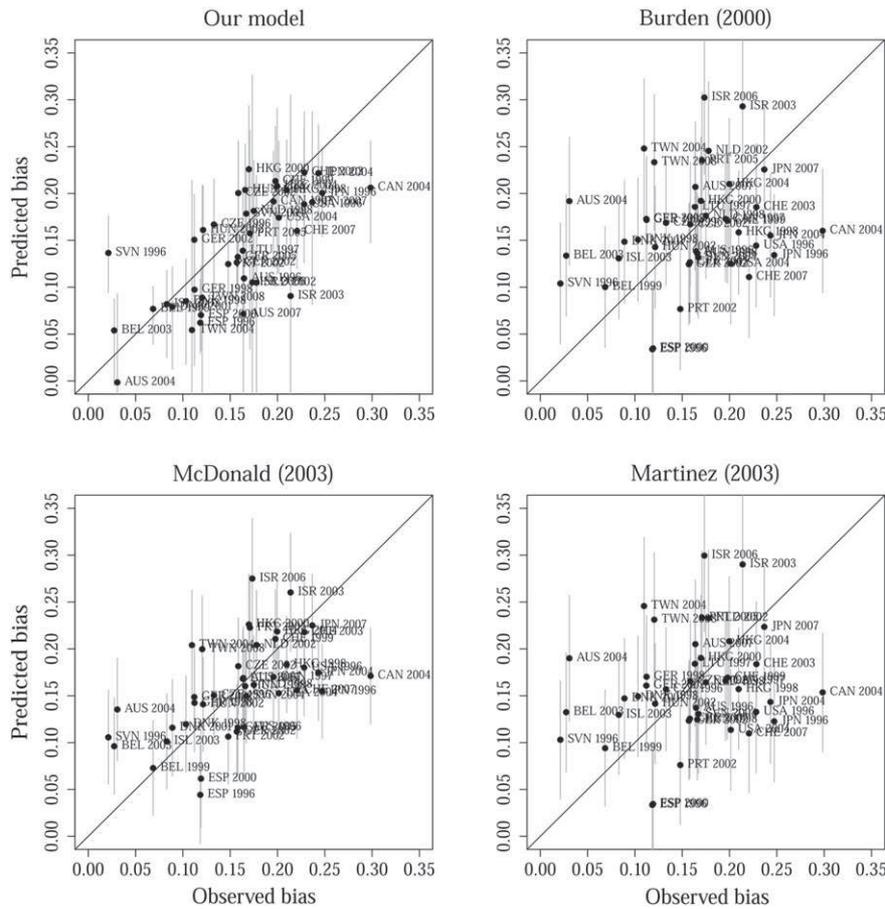
Fig. 5. Observed versus predicted turnout bias: out-of-sample predictions from our model, and from model specifications according to Burden (2000), McDonald (2003) and Martinez (2003). Lines indicate 90% credible intervals.

way, nonvoters that are successfully being interviewed in election surveys are not typical in that they still represent the politically more involved faction of nonvoters, so that survey based figures tend to blur actual differences between the voters' and nonvoters' characteristics (see Brehm, 1993; Stoop, 2005). On the other hand, those respondents most likely to vote are often just those respondents most likely to misreport if they, for one reason or other, failed to vote. Regarding comparisons of voters' and nonvoters' characteristics based on self reported voting behavior from surveys, increased misreporting among nonvoters that otherwise share many characteristics of voters therefore probably results in an over estimation of the differences between actual voters and nonvoters (see Bernstein et al., 2001; Silver et al., 1986).

These problems are far less frequently appreciated in survey based comparative work (for some notable exceptions, see Brockington, 2004, 2009; Chen, 2011; Karp and Brockington, 2005), although our results suggest that their distorting effects may propagate to the contextual level. If, for example, certain institutional arrangements (such as proportional representation) foster voter turnout, they may at the same time increase misreporting rates among nonvoters, so that survey based estimates of turnout differences between electoral systems may be overstated. This is exactly what Karp and Brockington (2005) observe in their sensitive cross national analysis of both reported and validated turnout!

Previous strategies to deal with these issues include ex ante efforts to reduce vote misreporting by improved survey instruments and data collection procedures (e.g., Abelson et al., 1992; Belli et al., 1999, 2006; Box Steffensmeier et al., 2000; Duff et al., 2007; Gerber and Rogers, 2009; Holbrook and Krosnick, 2010a, 2010b; Presser, 1990; Silver et al., 1986), or to reduce voter over representation via increased survey response rates (e.g., Arzheimer and Klein, 1999; Brehm, 1994; Keeter et al., 2000; Schmeets, 2010). However, our results indicate that increasing response rates may, in part, even aggravate the problems, and that changes in question formats have negligible effects. More promising estimation solutions to retroactively adjust for misclassification and sample selection exist (e.g., Brehm, 1993, 1999; Deufel and Kedar, 2010; Katz and Katz, 2010), but require information about the bias components. By making the bias components amenable for empirical scrutiny, assessing their prevalence and heterogeneity using vote validation data, and deter mining potential contextual correlates, our study is a first step to provide such information, possibly even for those surveys that do not include vote validation studies.

Acknowledgment

References

Abelson, R.P., Loftus, E.F., Greenwald, A.G., 1992. Attempts to improve the accuracy of self-reports of voting. In: Tanur, J.M. (Ed.), Questions About Questions: Inquiries into the Cognitive Bases of Surveys. Russell Sage Foundation, New York, pp. 138–153.

Achen, C.H., 1986. Statistical Analysis of Quasi-Experiments. University of California Press.

Arzheimer, K., Klein, M., 1999. The effect of material incentives on return rate, panel attrition and sample composition of a mail panel survey. International Journal of Public Opinion Research 11, 368–377.

Belli, R., Traugott, M., Young, M., McGonagle, K., 1999. Reducing vote overreporting in surveys – social desirability, memory failure, and source monitoring. Public Opinion Quarterly 63, 90–108.

Belli, R., Moore, S., VanHoewyk, J., 2006. An experimental comparison of question forms used to reduce vote overreporting. Electoral Studies 25, 751–759.

Bernstein, R., Chadha, A., Montjoy, R., 2001. Overreporting voting: why it happens and why it matters. Public Opinion Quarterly 65, 22–44.

Box-Steffensmeier, J., Jacobson, G., Grant, J., 2000. Question wording and the house vote choice – some experimental evidence. Public Opinion Quarterly 64, 257–270.

Brehm, J., 1993. The Phantom Respondents: Opinion Surveys and Political Representation. University of Michigan Press, Ann Arbor, MI.

Brehm, J., 1994. Stubbing our toes for a foot in the door? Prior contact, incentives and survey response. International Journal of Public Opinion Research 6, 45–63.

Brehm, J., 1999. Alternative corrections for sample truncation: applications to the 1988, 1990, and 1992 senate election studies. Political Analysis 8, 183–199.

Brockington, D., 2004. The paradox of proportional representation: the effect of party systems and coalitions on individuals' electoral participation. Political Studies 52, 469–490.

Brockington, D., 2009. It's about the benefits. Choice environments, ideological proximity and individual participation in 28 democracies. Party Politics 15, 435–454.

Burden, B., 2000. Voter turnout and the national election studies. Political Analysis 8, 389–398.

Burden, B., 2003. Internal and external effects on the accuracy of NES turnout: reply. Political Analysis 11, 193–195.

Cassel, C., 2003. Overreporting and electoral participation research. American Politics Research 31, 81–92.

Cassel, C., 2004. Voting records and validated voting studies. Public Opinion Quarterly 68, 102–108.

Chen, T., 2011. Uncovering the micro-foundations of turnout and electoral systems. Electoral Studies 30, 295–308.

Cornell, J.E., Mulrow, C.D., 1999. Meta-analysis. In: Adèr, H.J., Mellenbergh, G.J. (Eds.), Research Methodology in the Social, Behavioral and Life Sciences. Sage, London, pp. 285–323.

Deufel, B.J., Kedar, O., 2010. Race and turnout in U.S. elections exposing hidden effects. Public Opinion Quarterly 74, 286–318.

Duff, B., Hanmer, M.J., Park, W., White, I.K., 2007. Good excuses: understanding who votes with an improved turnout question. Public Opinion Quarterly 71, 67–90.

Gerber, A.S., Rogers, T., 2009. Descriptive social norms and motivation to vote: everybody's voting and so should you. Journal of Politics 71, 178–191.

Granberg, D., Holmberg, S., 1991. Self-reported turnout and voter validation. American Journal of Political Science 35, 448–459.

Groves, R., 2006. Nonresponse rates and nonresponse bias in household surveys. Public Opinion Quarterly 70, 646–675.

Hausman, J.A., Abrevaya, J., Scott-Morton, F.M., 1998. Misclassification of the dependent variable in a discrete-response setting. Journal of Econometrics 87, 239–269.

Heckman, J., 1979. Sample selection bias as a specification error. Econometrica 47, 153–161.

Hill, K.Q., Hurley, P.A., 1984. Nonvoters in voters' clothing: the impact of voting behavior misreporting on voting behavior research. Social Science Quarterly 65, 199–206.

Holbrook, A.L., Krosnick, J.A., 2010a. Measuring voter turnout by using the randomized response technique. Evidence calling into question the method's validity. Public Opinion Quarterly 74, 328–343.

Holbrook, A.L., Krosnick, J.A., 2010b. Social desirability bias in voter turnout reports. Tests using the item count technique. Public Opinion Quarterly 74, 37–67.

Holbrook, T., Heidbreder, B., 2010. Does measurement matter? The case of VAP and VEP in models of voter turnout in the United States. State Politics and Policy Quarterly 10, 157–179.

Jackman, S., 1999. Correcting surveys for non-response and measurement error using auxiliary information. Electoral Studies 18, 7–27.

Jones, E., 2008. Vote overreporting: the statistical and policy implications. Policy Perspectives 15, 83–97.

Karp, J., Brockington, D., 2005. Social desirability and response validity: a comparative analysis of overreporting voter turnout in five countries. The Journal of Politics 67, 825–840.

Katosh, J., Traugott, M., 1981. The consequences of validated and self-reported voting measures. Public Opinion Quarterly 45, 519–535.

Katz, J., Katz, G., 2010. Correcting for survey misreports using auxiliary information with an application to estimating turnout. American Journal of Political Science 54, 815–835.

Keeter, S., Miller, C., Kohut, A., Groves, R.M., Presser, S., 2000. Consequences of reducing nonresponse in a large national telephone survey. Public Opinion Quarterly 64, 125–148.

Lin, I., Schaeffer, N., 1995. Using survey participants to estimate the impact of nonparticipation. Public Opinion Quarterly 59, 236–258.

Locander, W., Sudman, S., Bradburn, N., 1976. An investigation of interview method, threat and response distortion. Journal of the American Statistical Association 71, 269–275.

Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing 10, 325–337.

Manski, C.F., 1995. Identification Problems in the Social Sciences. Harvard University Press, Cambridge.

Martinez, M., 2003. Comment on "Voter turnout and the national election studies". Political Analysis 11, 187–192.

McDonald, M., 2003. On the overreport bias of the National Election Study turnout rate. Political Analysis 11, 180–186.

McDonald, M.P., Popkin, S.L., 2001. The myth of the vanishing voter. American Political Science Review 95 (4), 963–974.

OSCE, 2008. OSCE Election Observation Mission Belarus – Parliamentary Elections, 28 September 2008. Technical Report. http://www.osce.org/odihr/elections/belarus/33564.

Presser, S., 1990. Can changes in context reduce vote overreporting in surveys? Public Opinion Quarterly 54, 586–593.

Presser, S., Traugott, M.W., 1992. Little white lies and social science models: correlated response errors in a panel study of voting. Public Opinion Quarterly 56, 77–86.

Presser, S., Traugott, M.W., Traugott, S., 1990. Vote "Over" Reporting in Surveys: the Records or the Respondents. American National Election Studies Technical Report Series, No. 010157.

Schmeets, H., 2010. Increasing response rates and the consequences in the Dutch Parliamentary Election Study 2006. Field Methods 22, 391–412.

Sigelman, L., 1982. The non-voting voter in voting research. American Journal of Political Science 26, 47–56.

Silver, B., Anderson, B., Abramson, P., 1986. Who overreports voting? The American Political Science Review 80, 613–624.

Smith, J.K., Gerber, A.S., Orlich, A., 2003. Self-prophecy effects and voter turnout: an experimental replication. Political Psychology 24, 593–604.

Stocké, V., 2007. Response privacy and elapsed time since Election Day as determinants for vote overreporting. International Journal of Public Opinion Research 19, 237–246.

Stocké, V., Stark, T., 2007. Political involvement and memory failure as interdependent determinants of vote overreporting. Applied Cognitive Psychology 21, 239–257.

Stoop, I., 2005. The Hunt for the Last Respondent. Social and Cultural Planning Office, The Hague.

Thompson, S.G., Turner, R.M., Warn, D.E., 2001. Multilevel models for meta-analysis, and their application. Statistical Methods in Medical Research 10, 375–392.

Tittle, C.R., Hill, R.J., 1967. The accuracy of self-reported data and prediction of political activity. Public Opinion Quarterly 31, 103–106.

Tourangeau, R., Groves, R.M., Redline, C.D., 2010. Sensitive topics and reluctant respondents. Demonstrating a link between nonresponse bias and measurement error. Public Opinion Quarterly 74, 413–432.

Tourangeau, R., Yan, T., 2007. Sensitive questions in surveys. Psychological Bulletin 133, 859–883.

Tourangeau, R., Smith, T.W., 1996. Asking sensitive questions. The impact of data collection mode, question format, and question context. Public Opinion Quarterly 60, 275–304.

Traugott, M.W., 2008. Validation studies. In: Donsbach, W., Traugott, M.W. (Eds.), The SAGE Handbook of Public Opinion Research. Sage, London, pp. 408–416.

Voogt, R., Saris, W., 2003. To participate or not to participate: the link between survey participation, electoral participation, and political interest. Political Analysis 11, 164–179.

Warn, D.E., Thompson, S.G., Spiegelhalter, D.J., 2002. Bayesian random effects meta-analysis of trials with binary outcomes: methods for the absolute risk difference and relative risk scales. Statistics in Medicine 21, 1601–1623.