Meta-Analysis of the Effects of Voting Advice Applications*

Simon Munzert (Hertie School)†
Sebastian Ramirez Ruiz (Hertie School)

This article has been accepted for publication in Political Communication

Abstract. We review the influence of voting advice applications (VAAs) on three core outcomes: turnout, vote choice, and issue knowledge. In a meta-analysis of 55 effects reported in 22 studies, comprising 73,673 participants in 9 countries, we find strong evidence for positive effects of VAA usage on reported turnout (OR = 1.87; 95% CI = [1.50, 2.33]) and vote choice (OR = 1.44; 95% CI = [1.16, 1.78]) as well as modest evidence on knowledge increase (partial correlation = 0.09; 95% CI = [-0.01, 0.18]). At the same time, we observe large heterogeneity in effect sizes, for which study design turns out to be a key driver. Effects are substantively weaker in causally more rigorous experimental studies. We highlight the need for more well-powered experimental research as well as studies focusing on the acquisition of knowledge in VAA usage.

Keywords. Meta-analysis, VAA, Voting Advice Application, Effect, Impact, Turnout, Vote Choice, Issue Knowledge, Heterogeneity

*This research was generously supported by a grant from the Volkswagen Foundation Computational Social Science Initiative, reference 92 143. We are grateful to Bernhard Clemm von Hohenberg, Francesca Minetto, Henrik Domansky, Konstantinos Gemenis, Laurens Klein Kranenburg, and Valérie-Anne Mahéo-Le Luel for their valuable feedback on the manuscript. Supplemental data for this article can be accessed on this Dataverse repository.

†Corresponding author. Address: Friedrichstrasse 180, 10117 Berlin, Germany. Email: munzert@hertie-school.org, Web: simonmunzert.github.io.

INTRODUCTION

The vision of the internet as a liberating tool for global citizenship has been severely battered. Once praised as a savior for deliberative democracy (Coleman and Blumler, 2009; Gimmler, 2001), it is now seen by many as one of its biggest challenges (Settle, 2018; Sunstein, 2007). However, in times of rampant misinformation and powerful partisan media online, voting advice applications (VAAs) testify to the empowering capabilities of digital tools. Aside from the informational benefits of these voter guides, many studies have suggested sizeable effects on political participation and vote choice.

The body of evidence about the effects of VAAs on political behavior has been growing quickly. In this note, we present the first quantitative review of VAA effects on individual turnout, vote choice, and accumulation of issue knowledge. Summarizing 55 effects from 22 studies covering over 73,673 participants and 25 elections in 9 countries, our analysis substantively extends the body of evidence from previous qualitative reviews of the VAA effects literature (Garzia, 2010; Garzia and Marschall, 2012). Using cross-classified random-effects meta-analyses, we find strong evidence for positive effects of VAA usage on reported turnout (odds ratio (OR) = 1.87; 95% confidence interval (CI) = [1.50, 2.33]) and vote choice (OR = 1.44; 95% CI = [1.16, 1.78]) as well as modest evidence on knowledge increase (partial correlation = 0.09; 95% CI = [-0.01, 0.18]).

While some optimism might be warranted, the results have to be read under the caveat that many of the early impact evaluations have been limited by a reliance on convenience samples, lack of random assignment of VAA usage, or both. This raises questions about self-selection bias. In fact, moderator analyses show that the identified effect sizes are much lower when focusing on more recently deployed experimental designs (effects on turnout: OR = 1.04; 95% CI = [0.90, 1.72], effects on vote choice: OR = 1.24; 95% CI = [0.87, 1.25]). We call for more well-powered experimental research as well as studies focusing on the acquisition of knowledge in VAA usage.

THE NATURE AND GROWING POPULARITY OF VAAS

A Voting Advice Application (VAA) is an online tool, either a website or a mobile app, that guides voters' choices. After completing a series of items to indicate their agreement with policy statements on salient issues, users are shown how their positions on issues correspond to those of each party competing in an election (see, e.g., Garzia and Marschall, 2012).

Over the past decade, VAAs have become extremely popular, particularly in European countries with multi-party systems such as the Netherlands (the popular tool *Stemwijzer* was used 6.8 million times at the 2017 national election with an electorate of 12.9 million eligible voters), Germany (15.7m user sessions of *Wahl-O-Mat* at the 2017 federal election with 61.7m eligible voters), Switzerland (1.3m user sessions of *smartvote* at the 2015 federal election with 5.3m eligible voters), and also during second-order elections, such as the elections to the European Parliament. Beyond user session counts, which can contain repeated usage, the trend of VAA proliferation is also present in representative election surveys that suggest a reach of more than 20% of the electorate in recent Swiss, German, Finnish and Danish elections (Germann and Gemenis, 2019). Considering the significant numbers of users, even small effects can have a potentially substantial impact on the aggregate, which makes evaluating them a noteworthy object of research.

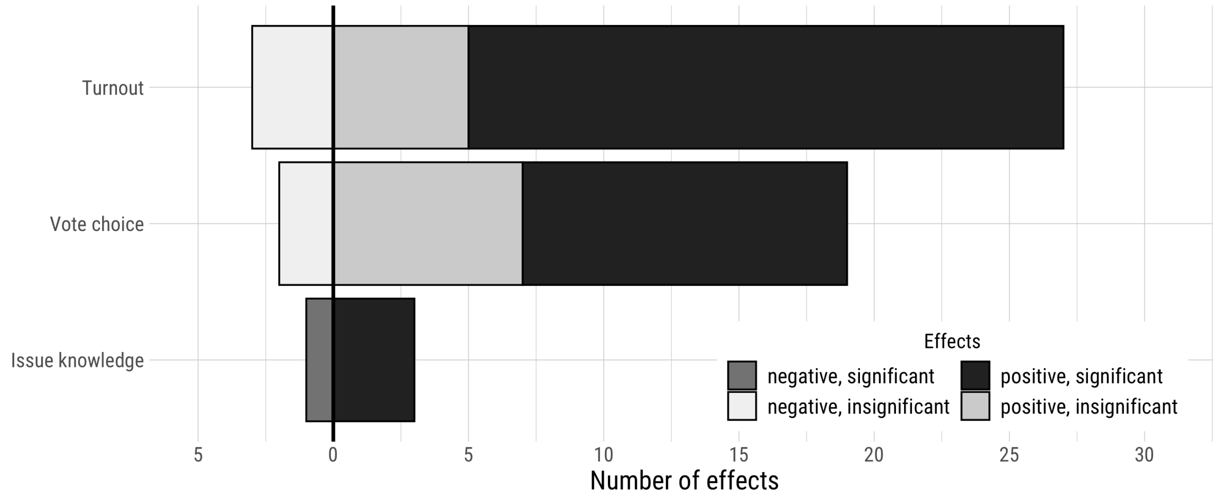
BEHAVIORAL AND COGNITIVE EFFECTS OF VAAS

We group studies examining the potential effects of VAAs on voters' political behavior into three categories according to their outcome of interest: The decision to turn out to vote, whom to vote for, and the accumulation of knowledge about policy stances of political platforms (Garzia and Marschall, 2019). We restrict our discussion to the effects of VAA usage (perceived as binary treatment) as the explanatory variable, which is also a key selection criterion in the meta-analysis. Furthermore, we refrain from discussing heterogeneity of effects across subpopulations (e.g., in different educational or knowledge strata) because such effects are neither frequently nor consistently reported in the literature. We follow our discussion about the behavioral and cognitive effects of VAAs with an overview of some of the proposed mechanisms and measurement issues specific to each outcome of interest. Forestalling the quantitative synthesis, Figure 1 summarizes our compilation of VAA effects research. The retrieval process for the studies summarized in this figure is described in detail in the method section.

EFFECTS ON TURNOUT

Mechanisms. As can be seen in Figure 1, the question of whether VAAs help boost turnout is at the core of the VAA effects research program and has received the widest attention. This is not without reason: Exposure to online information has been shown to be positively associated with higher levels of political engagement (Boulianne, 2009), and VAAs as interactive tools to communicate political information seem particularly suited to empower and engage citizens. By efficiently teaching users about their congruence with parties' positions on various political issues, they can both stimulate interest and reduce the costs of information (Garzia et al., 2017).

Figure 1: Summary of VAA effects reported in 22 studies, by outcome, direction, and significance at $p < 0.05$.



Measurement issues. An important challenge for synthesizing findings in the literature is the heterogeneity of outcome measures. The most commonly studied turnout outcome is a binary indicator of participation in the election under analysis. A few studies focus on related outcomes, such as mobilization (switch from non-voting to voting from one election to the other or changing intention to vote during the campaign; see, e.g., Vassil, 2011) or intention to vote (Marschall and Schultze, 2012). While such differences in outcomes and measurement can induce meaningful effect heterogeneity, the amount of research looking at alternative outcomes is low, so we collapsed all studies that focus on any of these outcomes into the turnout category.

As visible in the figure, the vast majority of effects in the literature are positive, and most report a statistically significant boost in turnout among VAA users compared to non-users.

EFFECTS ON VOTE CHOICE

Mechanisms. The second major outcome is (change in) vote choice. The rationale is straightforward: By matching voters to parties based on voters' issue preferences, they support proximitybased issue voting (Wagner and Ruusuvirta, 2012). As long as issue congruence plays a role in voting behavior, VAA exposure can potentially influence party choice. seems to be a powerful relationship: In the aggregate (see again Figure 1), the majority of studies identify positive and significant effects of VAA usage on vote choice.

Measurement issues. There is no gold standard of measuring VAA effects on vote choice. Most studies focus on changes in vote choice, identified either by actual vote switching between elections or by within-campaign changes in reported vote intention (or choice) after VAA usage. Inferences are then usually based on the somewhat heroic assumption that every observed change in preferences by VAA users can be attributed to VAA exposure, and no change implies the absence of an effect. However, switching vote choice is a rather high benchmark for VAA effects on voting preferences. VAAs cannot only convert but also reinforce attitudes by providing confirming advice to previously decided voters (Klein Kranenburg and Rosema, 2019). At a lower level, VAA usage could also induce changes in vote certainty, the likelihood to vote for a particular party, or simply change a user's sympathy towards particular parties. Just measuring reported changes in vote intentions thus might lead to underestimating attitudinal effects. Still, the by far most common way employed in the literature to operationalize attitudinal changes is to focus on change in vote choice. In our study sample, vote switching from the previous to the current election (used 16 times) and change of vote intention in the ongoing campaign (i.e., pre-VAA usage intention vs. post-VAA usage intention; used 5 times) are the most common operationalizations of the vote choice outcome.

Although the literature has addressed the need for more fine-grained operationalizations to capture the mechanisms connecting the user's initial intention to vote, affinity for a party, and the advice given by the VAA (see Garzia and Marschall, 2019), only a few studies offer this level of nuance. The lack of alternative metrics and incomparability of some outcome measures results in a category populated by measures that pick up vote switching or changes in vote intention.

EFFECTS ON ISSUE KNOWLEDGE

Mechanisms. A more recent line of research has turned to the question of whether VAAs help establish knowledge about these positions. After all, VAAs are designed to provide information about parties' issue stances by comparing them to the users' issue preferences. This mechanism is fundamental for downstream effects on actual changes in voting behavior.

Measurement issues. While studies in this domain often refer to "political knowledge" more generally, knowledge is usually operationalized as VAA users' ability to identify party positions on selected issues, which is consistent with what most VAAs are designed to convey. Therefore, we refer to "issue knowledge" in the following sections.

As indicated in Figure 1, research on the consequences for issue knowledge is still limited to very few studies. We nevertheless consider these studies here to provide an early synthesis of research on this outcome.

METHOD

STUDY RETRIEVAL

To identify relevant studies, we searched the databases *Web of Science*, *Dimensions*, *Google Scholar*, and *ProQuest dissertations and theses* using the Boolean expression "('voting advice application' OR 'online voter guide' OR 'voting aid') AND ('impact' OR 'effect')". Our database recovery strategy employed broad terms deliberately to ensure high sensitivity in the retrieval of relevant literature. To identify additional (and potentially unpublished) research, we screened the work listed on the pages of the ECPR Research Network on Voting Advice Applications, which provides a curated bibliography on research related to VAAs in general (see http://vaa-research.net/?page_id=18), searched recent programs of leading conferences in political science (APSA, MPSA, EPSA, IPSA, and ECPR), and directly contacted scholars of VAA effects research. Furthermore, we screened the reference lists of the identified relevant studies. Finally, we screened studies citing our selected articles through Google Scholar's "cited by" feature. The search was conducted over a period of several weeks and was completed on September 23, 2019.

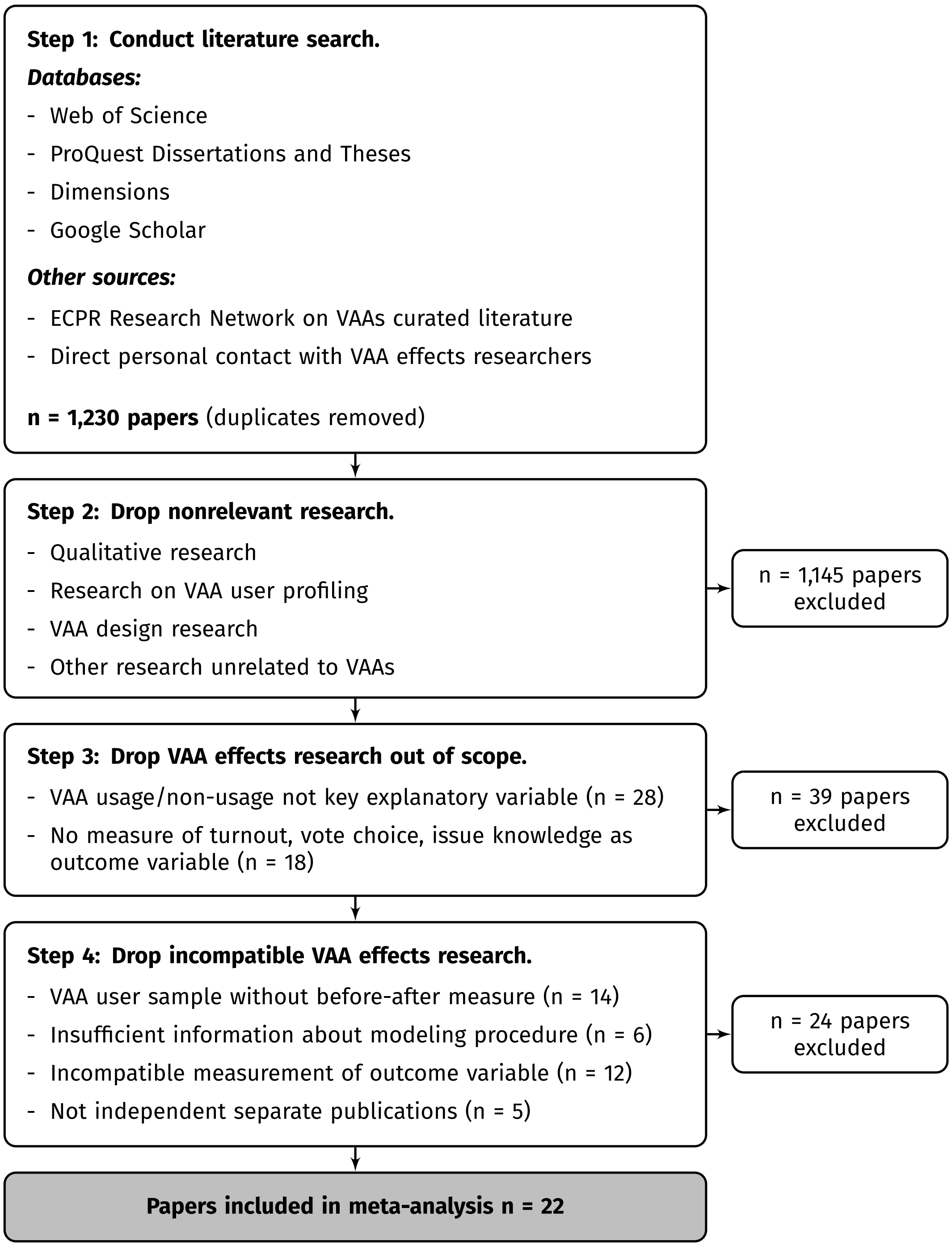
INCLUSION CRITERIA

To be included in our analysis, study effects had to conform with the following criteria: (1) They had to focus on a variant of turnout, change in vote choice, or issue knowledge as the outcome variable. (2) They had to report a quantitative estimate for the main effect of a binary VAA exposure variable on one of the relevant outcomes (see below for effect size transformations). (3) They had to report a comparison with a meaningful baseline (e.g. before and after measures, or treatment and control groups). (4) They had to provide sufficiently detailed information on the modeling procedure to allow us to infer the effect size.

To apply these criteria, the study selection process was divided in four steps, as illustrated in Figure 2. In Step 1, the titles and abstracts of the articles retrieved through the initial search were screened. We identified 1,230 unique papers in total.

In Step 2, after the full-text screening, which we took care of ourselves, we selected those studies that were designed to identify attitudinal and behavioral effects of VAA usage. We excluded

FIGURE 2: Illustration of study collection and selection procedure



studies that focused on VAA design issues or user profiling, followed an exclusively qualitative approach, and overall research unrelated to the effects of VAAs. As a result, 1,145 papers (93% of the originally identified papers) were discarded.

In Step 3, we excluded studies that focused on VAA effects research, but were outside of the scope of our analysis for at least one of the following reasons: (1) They did not focus on a variant of turnout, change in vote choice, or issue knowledge as the outcome variable. This included studies about the effects on information seeking (e.g., Manavopoulos et al., 2018) or the perceived utility of VAAs (Alvarez et al., 2014). While these outcomes are per se worth studying, the amount of evidence did not justify a quantitative assessment. (2) They did not report a quantitative estimate for the effect of a binary VAA exposure variable on one of the relevant outcomes. For meta-analytic effects to be meaningful, it is important to maintain consistency in the nature of the treatment. The by far most common VAA effect assessed in the literature is a binary operationalization of exposure, that is whether a person used the tool or not. A few studies consider alternative implications of VAA usage, such as irritation by the VAA recommendation (Israel et al., 2017) or the information value of VAA results (Gemenis, 2018). Again, these studies add important nuance to the analysis of VAA effects. However, since our selection strategy was guided by rendering a meaningful synthesis of comparable effects that is also largely representative of the literature, they had to be excluded. A total of 39 studies were dropped at Stage 3. Figure 2 provides counts of the studies that applied to each of the exclusion criteria for steps 3 and 4. In many cases, multiple reasons within the same step applied, hence the counts do not add up to the total number of studies excluded.

Finally, in Step 4 we excluded studies that focused on the effects of VAAs on our outcomes of interest but were incompatible with our analytic strategy for at least one of the following reasons: (1) They did not report a comparison with a meaningful baseline by, e.g., contrasting outcomes of VAA users with non-users or of VAA users before and after actual usage. This is a criterion that is not satisfied by many of the earlier studies that focused exclusively on VAA user samples without baseline outcome measures (e.g., Fivaz and Nadig, 2010; Marschall and Schmidt, 2008). As all units in these sample were exposed to the VAA, it is not possible to assess the outcome under the control condition. Self-reported outcomes, such as agreement to the statement "I did not want to vote, but the voting advice application has motivated me to cast my vote" are clearly unsatisfactory and insufficient to provide a meaningful effect measure. (2) They did not provide sufficiently detailed information on the modeling procedure to allow us to infer the effect size (e.g., Ruusuvirta and Rosema, 2009; Westle et al., 2015). The minimum information needed was an effect size together with an uncertainty estimate, or information that could be used to reconstruct these quantities (see the following section). (3) They measured the outcome on a scale that was incompatible with our modeling strategy. Again, we prioritized consistency in the outcome measure to be able to arrive at informative effect estimates. Therefore, studies such as the one by Garry et al. (2018), who measure the change in the propensity to vote for a party (on a 0-10 support scale) contingent on the advice received by the respondent, or some of the effects reported by Pianzola et al. (2019), who study changes on propensity to vote scales, had to be excluded. (4) The same samples and almost identical modeling strategies were utilized in different publications. In the instances where separate publications could not be treated as independent but the reported effects were different, we chose the study that conveyed the more conservative effect size. As a result of the final step, another 24 studies were excluded, leaving us with 22 unique studies as a basis for our analyses. Table A1 in the Online Appendix provides an overview of selected effects that were excluded from the meta-analysis along with the reasons for exclusion.

The remaining studies were then systematically classified according to sample and election type, the existence of a before-after measure, the general design type, and the outcome under study.

The studies identified and used in the meta-analysis are listed in Tables A2 to A4 in the Online Appendix.

EFFECT SIZE CALCULATION

The vast majority of VAA effects studies with a focus on turnout and/or vote choice model binary outcomes, typically using logit or probit regressions. Therefore, we use log-odds as effect size measures, which we believe are more informative and can be interpreted more intuitively than more generic effect size measures (such as Hedges' *g*). Effect sizes from studies that do not report log-odds are transformed if possible. Probit coefficients and standard errors are multiplied by factor 1.6 to approximate log-odds (Amemiya, 1981). Unstandardized regression coefficients from linear probability models are multiplied by factor 4 (Gelman and Hill, 2006). Finally, in cases where mere percentage point differences between VAA non-users and users are reported, we directly compute log-odds from the reconstructed 2x2 table. In presenting the results, the estimated log-odds are exponentiated to odds ratios to facilitate interpretation. Issue knowledge is measured on different scales. To render effect sizes comparable, we convert them into partial correlations (Aloe and Becker, 2012).

MODELING STRATEGY

We summarize evidence separately by outcome and implement cross-classified random-effects models (CCREM; see Fernández-Castilla et al., 2019), assuming that the studies included in the meta-analysis represent random samples from a larger population of studies. Several studies report multiple effect sizes reflecting alternative modeling strategies. Furthermore, some studies present evidence from different countries and elections. We only consider effect sizes based on the model presented by the authors as the most robust design to guard against self-selection bias. Still, we cannot assume the independence of effect sizes within studies even when they are based on independent data because they are subject to the same data-pre-processing decisions and modeling strategies. Also, some samples are re-used in different studies and thus represented more than once in our analysis. The cross-classified multilevel structure explicitly accounts for these forms of dependence (Cheung, 2014; Fernández-Castilla et al., 2019). Effect sizes (level 1) are nested within both studies and elections (cross-classified levels 2 and 3). The weight with which an effect contributes to the overall CCREM estimate is a function of the effect estimate's precision and the amount of residual model heterogeneity, reflected in the model-implied variance-covariance matrix of the outcomes. In each model, all weights sum up to 100%. Meta-analyses were implemented using the metafor package (Viechtbauer, 2010), the estimations were carried out via REML (Viechtbauer, 2005).

We also offer an alternative strategy to account for dependency in the data by excluding quasi-duplicate studies that re-analyze study data which had been used in predecessor studies. The overall estimates are marginally lower for turnout and virtually identical for vote choice (see Figure B8 in the Online Appendix).

RESULTS

The search resulted in a final overall (subgroup) sample of $k = 55$ effects ($k_{\text{turnout}} = 30$; $k_{\text{vote choice}} = 21$; $k_{\text{knowledge}} = 4$) from $n = 22$ studies ($n_{\text{turnout}} = 13$; $n_{\text{vote choice}} = 10$; $n_{\text{knowledge}} = 4$), with a total (subgroup) sample size of $s = 73,673$ ($s_{\text{turnout}} = 52,573$; $s_{\text{vote choice}} = 32,980$; $s_{\text{knowledge}} = 9,272$).

Online Appendix A provides a full list of papers which are used in the meta-analysis as well as additional descriptive statistics.

OVERALL EFFECTS ANALYSES

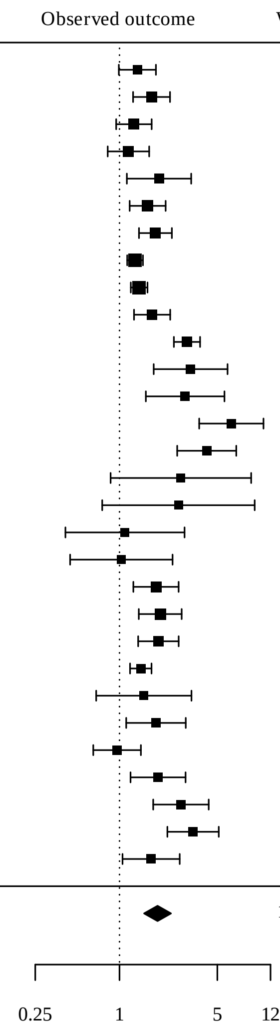
For the effects of VAA usage on turnout, the forest plot shown in Figure 3 displays the effect sizes (odds ratios, OR) reported in each study as well as the election (country and year) in which the VAA was implemented and the study sample size together with the number of VAA users within each sample. The CCREM estimate of the average observed effect of VAA usage on turnout is 1.87 (95% CI = [1.50, 2.33]). This is massive and implies that the odds of voting are expected to be on average about 90 percent higher for VAA users than non-users. At the same time, there is a large heterogeneity of effect sizes, ranging from no effect (Gemenis, 2018; Munzert et al., Forthcoming) to an OR of over 6 (for the 2012 Dutch general election in Garzia et al., 2017) (test for heterogeneity: $Q(df = 29) = 138.46, p-val < .0001, I^2 = 83.13\%$; see Higgins and Thompson, 2002). Also visible from the plot is the dominance of the multi-election studies by Garzia and colleagues which, taken together, account for 64 percent of the weight going into the average.

Figure 4 reports the effects of VAA usage on change in vote choice. The CCREM estimate of the average observed effect is 1.44 (95% CI = [1.16, 1.78]). The heterogeneity between effect sizes is still substantive but somewhat lower than for turnout ($Q(df = 20) = 46, p = .0009, I^2 = 78\%$), and relatively more studies report null effects. Again, two multi-election studies dominate the synthesis (Andreadis and Wall, 2014; Klein Kranenburg, 2015), accounting for 54 percent and 35 percent of the overall weight, respectively.

Finally, Figure 5 displays the summary of four studies exploring the effect of VAA usage on issue knowledge. Note that the effect size metric is now shifted to partial correlations. The average effect

Figure 3: Effect sizes (odds ratios) of VAA usage on turnout.

		Election		Sample		Estimates	
Author(s) and Year	Type	Country	Year	Size	VAA users	Weight	OR [95% CI]
Enyedi, 2016	EXP	Hungary	2010	1727	1164	2%	1.34 [0.99, 1.82]
Garzia & De Angelis, 2014	O−NP	Finland	2007	1062	313	5%	1.69 [1.25, 2.30]
Garzia & De Angelis, 2014	O−NP	Finland	2011	1030	443	5%	1.27 [0.95, 1.70]
Garzia & De Angelis, 2014	O−NP	Germany	2009	1187	116	5%	1.16 [0.82, 1.63]
Garzia & De Angelis, 2014	O−NP	Netherlands	2003	1016	325	3%	1.92 [1.13, 3.25]
Garzia & De Angelis, 2014	O−NP	Netherlands	2006	1793	687	6%	1.59 [1.18, 2.14]
Garzia & De Angelis, 2014	O−NP	Netherlands	2010	1693	706	5%	1.81 [1.38, 2.37]
Garzia & De Angelis, 2014	O−NP	Switzerland	2007	3127	256	10%	1.29 [1.14, 1.47]
Garzia & De Angelis, 2014	O−NP	Switzerland	2011	3254	368	10%	1.38 [1.20, 1.58]
Garzia et al., 2017	EXP	Italy	2013	888	454	4%	1.71 [1.27, 2.30]
Garzia et al., 2017	O−NP	EU	2009	19592	921	4%	3.03 [2.45, 3.76]
Garzia et al., 2017	O−NP	Finland	2003	1152	255	2%	3.22 [1.75, 5.92]
Garzia et al., 2017	O−NP	Germany	2013	1820	260	2%	2.94 [1.54, 5.62]
Garzia et al., 2017	O−NP	Netherlands	2012	1574	631	2%	6.30 [3.71, 10.69]
Gemenis & Rosema, 2014	O−P	Netherlands	2006	2356	903	2%	4.20 [2.58, 6.83]
Gemenis, 2018	O−SM	Greece	5/2012	890	?	1%	2.75 [0.86, 8.73]
Gemenis, 2018	O−SM	Greece	6/2012	896	?	1%	2.64 [0.75, 9.25]
Gemenis, 2018	O−SM	Greece	1/2015	1013	?	1%	1.09 [0.41, 2.92]
Gemenis, 2018	O−SM	Greece	9/2015	806	?	1%	1.03 [0.44, 2.39]
Germann & Gemenis, 2019	O−SM	Switzerland	2007	3578	322	5%	1.82 [1.26, 2.64]
Germann & Gemenis, 2019	O−SM	Switzerland	2011	3657	366	6%	1.95 [1.37, 2.78]
Germann & Gemenis, 2019	O−SM	Switzerland	2015	2931	410	4%	1.90 [1.36, 2.65]
Heinsohn et al., 2016	O−P	Germany − EU	2014	449	127	2%	1.42 [1.19, 1.69]
Mahéo, 2017	EXP	Canada	2014	277	138	1%	1.49 [0.68, 3.27]
Marschall & Schultze, 2012	O−NP	Germany	2009	1153	421	2%	1.82 [1.12, 2.97]
Munzert et al., 2020	EXP	Germany	2017	979	499	2%	0.96 [0.65, 1.42]
Mykkänen & Moring, 2006	O−NP	Finland − EU	2004	1362	?	2%	1.89 [1.20, 2.96]
Mykkänen & Moring, 2006	O−NP	Finland	2003	1511	?	2%	2.75 [1.74, 4.34]
Ruusuvirta et al., 2009	O−P	Netherlands	2006	2356	892	2%	3.35 [2.20, 5.12]
Vassil, 2012	EXP	Estonia − EU	2009	279	97	2%	1.68 [1.05, 2.69]
Overall estimate (CCREM with study and election REs)						100%	1.87 [1.50, 2.33]

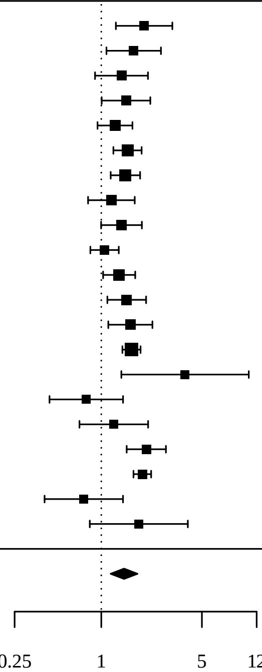


Note: Estimates are represented by black boxes sized proportionally to their weight. The estimated log-odds are exponentiated to odds ratios to facilitate interpretation. Study design types are: *Observational - panel* (O-P), *Observational - no panel* (O-NP), *Observational - selection and matching models* (O-SM), *Experimental* (EXP). See Appendix A for full study details.

is moderately positive with p.corr. = 0.09 (95% CI = [-0.01, 0.18]). Three out of four studies report a positive effect, while the effect reported by van de Pol (2016), the study with the largest sample size in the meta-analysis, is negative. Overall, the cumulative evidence on this relationship is the weakest of all in our study and dominated by research on Germany.

Figure 4: Effect sizes (odds ratios) of VAA usage on change in vote choice.

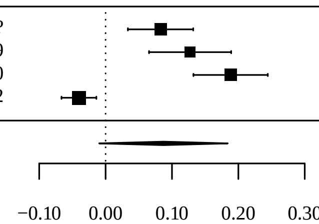
Author(s) and Year	Type	Election: Country	Election: Year	Sample: Size	Sample: VAA users	Weight	OR [95% CI]
Andreadis & Wall, 2014	O−P	Finland	2003	578	153	3%	1.98 [1.26, 3.12]
Andreadis & Wall, 2014	O−P	Finland	2007	617	188	3%	1.68 [1.09, 2.60]
Andreadis & Wall, 2014	O−P	Finland	2011	526	237	4%	1.38 [0.90, 2.11]
Andreadis & Wall, 2014	O−P	Germany	2009	1175	147	4%	1.48 [1.01, 2.19]
Andreadis & Wall, 2014	O−P	Netherlands	2003	1170	387	8%	1.25 [0.94, 1.65]
Andreadis & Wall, 2014	O−P	Netherlands	2006	1721	685	11%	1.52 [1.21, 1.91]
Andreadis & Wall, 2014	O−P	Netherlands	2010	1537	661	10%	1.47 [1.16, 1.86]
Andreadis & Wall, 2014	O−P	Switzerland	2007	1787	166	5%	1.18 [0.81, 1.71]
Andreadis & Wall, 2014	O−P	Switzerland	2011	1645	206	6%	1.38 [1.00, 1.91]
Enyedi, 2016	EXP	Hungary	2010	1727	1124	2%	1.05 [0.84, 1.32]
Klein Kranenburg, 2015	O−P	Netherlands	2006	1628	624	9%	1.33 [1.03, 1.72]
Klein Kranenburg, 2015	O−P	Netherlands	2010	1485	619	6%	1.50 [1.10, 2.05]
Klein Kranenburg, 2015	O−P	Netherlands	2012	1026	411	5%	1.59 [1.12, 2.27]
Klein Kranenburg, 2015	O−P	Netherlands	2015	5211	1516	15%	1.62 [1.40, 1.87]
Kleinnijenhuis et al., 2019	O−P	Netherlands	2010	1159	426	1%	3.82 [1.38, 10.58]
Mahéo, 2016	EXP	Canada	2014	211	211	1%	0.79 [0.44, 1.42]
Munzert et al., 2020	EXP	Germany	2017	923	483	1%	1.22 [0.71, 2.11]
Pianzola, 2014a	O−SM	Switzerland	2007	4067	?	2%	2.05 [1.50, 2.81]
Pianzola, 2014b	O−SM	Switzerland	2011	9163	?	2%	1.93 [1.67, 2.22]
Pianzola et al., 2019	EXP	Switzerland	2011	1775	1349	1%	0.76 [0.40, 1.42]
Vassil, 2012	EXP	Estonia − EU	2009	394	186	1%	1.82 [0.83, 3.99]
Overall estimate (CCREM with study and election REs)						100%	1.44 [1.16, 1.78]



Note: Estimates are represented by black boxes sized proportionally to their weight. The estimated log-odds are exponentiated to odds ratios to facilitate interpretation. Study design types are: *Observational - panel* (O-P), *Observational - no panel* (O-NP), *Observational - selection and matching models* (O-SM), *Experimental* (EXP). See Appendix A for full study details.

Figure 5: Effect sizes (partial correlations) of VAA usage on issue knowledge.

Author(s) and Year	Type	Election: Country	Election: Year	Sample: Size	Sample: VAA users	Weight	P. corr. [95% CI]
Heinsohn et al., 2016	O−P	Germany − EU	2014	1569	?	25%	0.08 [0.03, 0.13]
Munzert et al., 2020	EXP	Germany	2017	979	499	24%	0.13 [0.07, 0.19]
Schultze, 2014	O−NP	Germany	2009	1145	440	25%	0.19 [0.13, 0.24]
Van de Pol, 2016	O−SM	Netherlands	2014	5571	3522	26%	−0.04 [−0.07, −0.01]
Random−effects model						100%	0.09 [−0.01, 0.18]



Note: Estimates are represented by black boxes sized proportionally to their weight. Study design types are: *Observational - panel* (O-P), *Observational - no panel* (O-NP), *Observational - selection and matching models* (O-SM), *Experimental* (EXP). See Appendix A for full study details.

Altogether, the effect estimates are very stable across different model specifications, weighting schemes, and study subsets (see Figures B8 and B9 in the Online Appendix).

MODERATOR ANALYSES

The meta-analysis has revealed a substantive heterogeneity of effect sizes. We now turn to various study- and context-level characteristics as potential moderators to explore why VAA usage effects, while large on average, vary so much. We focus on turnout and vote choice; the sample of issue knowledge studies is too small to make any meaningful inference on sources of heterogeneity.

A particular challenge to the detection of the effects is the issue of self-selection into VAA usage, which has been recognized previously as a source of endogeneity bias (Garzia et al., 2017; Gemenis and Rosema, 2014; Pianzola, 2014)—those who tend to consult VAAs are more likely to engage politically in the first place. As different research designs guard against this bias to a different extent, we think of study designs as a prime suspect explaining the observed heterogeneity of effects.

Building on the typology of generations of VAA studies proposed by Germann and Gemenis (2019), we distinguish between four study design types according to their capability of tackling the selection bias issue. The scheme entails three categories for different observational study setups and an experimental category. Accordingly, we categorized studies as (1) *Observational - no panel* when VAA usage is observed (usually indirectly via survey-based reports) and the outcome is measured in a singular observation (post VAA usage) without a pre-VAA baseline, (2) *Observational - panel* when the outcome is measured before and after VAA use allowing to identify changes, (3) *Selection and matching models* when authors estimate the effect through a two-equation structure or a matching approach to account for non-random selection into VAA use, and (4) *Experimental* when VAA usage is randomly assigned or encouraged (see Figure A2 in the Online Appendix for a publication timeline of the studies by study design type).

Figure 6: Estimated effects of study design on turnout and vote choice. *Note:* Estimates along with 95% confidence intervals are based on mixed-effects CCREMs, with k equalling the number of effects available by group.

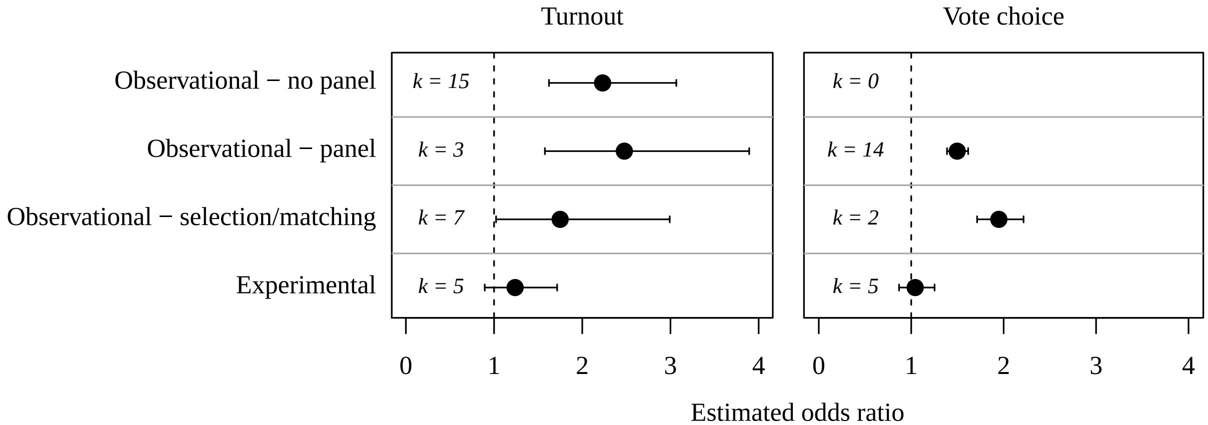


Figure 6 summarizes the estimated effects of study design on turnout and vote choice using mixed-effects models with the same random-effects structure as in the empty models. It reveals major differences between designs, supporting the idea that experimental designs, which are theoretically better suited to isolate selection effects, tend to find much weaker effects of VAA usage. For both turnout (OR = 1.24; 95% CI = [0.90, 1.72]) and vote choice (OR = 1.04; 95% CI = [0.87, 1.25]), the estimated effects for this subgroup fail statistical significance at conventional levels. It is also worth noting that these estimates come with fairly high precision. In contrast, the effects of the observational panel and non-panel studies are significantly larger. Estimates from selection models are somewhat in between (turnout) or even larger (vote choice).

Additionally, we explore the moderating effects of additional context and design features, specifically the country of VAA deployment, election year, election type (first vs. second-order), sample size, sample type, and operationalization of vote choice outcome variable. Most of them did not substantively help explain variation (see Figures B1 to B4 in the Online Appendix). Effect sizes in turnout models tend to be smaller in studies with smaller sample sizes and more recent elections

(2015–2017). These are also typical characteristics of studies with an experimental setup. Given the limited number of studies, it is difficult to tease out the causal source of heterogeneity. Figure B7 in the Online Appendix reports results from mixed-effects models with multiple moderators. Standard errors are inflated (likely due to multicollinearity issues), but the larger estimated effects of observational studies persist.

DISCUSSION AND CONCLUSION

In their still short coming-of-age period, voting advice applications have drawn considerable academic attention. The number of effect evaluations has reached a critical mass worthy to be synthesized. Overall, VAAs have received significant acclaim for their power to mobilize and change people's minds. Are VAAs' empowering effects one of the biggest success stories of the internet that nobody outside of the academic sphere is talking about?

The meta-analysis indicates that the existing evidence to answer this question is ambiguous—on second sight at least: Averaging over the existing literature indeed reveals an emerging consensus about the power of VAAs to boost turnout and make people re-consider their vote choice. However, moderator analyses provide an important qualification: The search for causal VAA effects has been plagued by self-selection issues, potentially inflating effect estimates. The literature has been aware of the problem for quite some time (Pianzola, 2014) and suggested several strategies to tackle the issue (Germann and Gemenis, 2019). More recently, the experimental encouragement design has been established as the new gold standard (Enyedi, 2016; Mahéo, 2016; Munzert et al., Forthcoming; Pianzola et al., 2019; Vassil, 2011). Studies following the experimental paradigm tend to fail in showing a substantive positive impact of VAA usage. Furthermore, publication bias as another rival explanation of the dominance of positive and significant effects in the literature is unlikely to play a major role (see analysis in Online Appendix C).

A look into the larger literature on mobilization and persuasion effects during election campaigns helps put the findings into perspective. Experimental evidence on door-to-door canvassing points to a complier average causal effect on turnout of about 0.8 (Europe-focused studies; see Bhatti et al., 2019) to 2.5 percentage points (US-focused studies; see Green et al., 2013). The projected effects of VAA usage on the overall turnout of the electorate range from 0.7 (2009 German parliamentary election Garzia et al., 2017) over 1.2 (2007 Swiss federal election Germann and Gemenis, 2019) to 4.4(6.8) percentage points (2006(2012) Dutch parliamentary elections Garzia et al., 2017; Gemenis and Rosema, 2014). In the context of our findings and considering the fact that voters typically interact with VAAs for no longer than a couple of minutes, not all these reported effects seem equally plausible .

Notwithstanding the above, our analysis has its limitations. First, individual studies are necessarily able to offer more nuance than a quantitative synthesis of the literature. For instance, we analyzed different measures of turnout jointly rather than separately and did not consider individual-level moderators or more complex mechanisms. Second, the number of studies for some outcomes is still modest. It will require more research in the future, in particular such that employs methodologically robust designs that allow for the identification of the effects of interest, to enrich our knowledge on the size of these effects. Third, although being a very young literature, the studies that were considered in this meta-analysis covered a wide range of empirical strategies. With samples of varying quality and identification strategies of varying strengths, some of the individual effects provide rather poorly identified average treatment effects while others offer well-identified encouragement-specific local average treatment effects. Fundamentally, however, we do not see meta-analysis as a tool to precisely identify an ATE for a pre-defined population, but to quantify the tendencies and heterogeneity in effects reported in published and unpublished VAA effects research.

The empirical study of VAA effects on political behavior has become more mature. However, the case is far from closed. Our findings point to future directions. First, more evidence is needed on the probably most obvious direct consequence of VAAs usage: A boost of citizens' knowledge about parties' issue positions. This might be particularly consequential in low-information settings, such as local or second-order elections, which were also underrepresented in our sample. Second, with regards to downstream effects on voting behavior, we hope to see more studies that innovate in measurement of outcomes and treatments by considering, e.g., actual VAA results (see, e.g., Gemenis, 2018), validated voting behavior, or changes in latent preferences. Third, with the dominance of research on countries such as Germany and the Netherlands, more evidence is needed from other (also non-Western) contexts (see Figure A1 in the Online Appendix for the geographic distribution of VAA deployment). Finally, experimental designs are no panacea to the study of VAA effects. They tend to be costly to implement and therefore carried out in smaller samples. Therefore, more well-powered experimental studies are needed. Additionally, noncompliance limits the generalizability of these findings, as compliers might be substantively different from the overall population of VAA users. While much remains to be explored, we hope that this meta-analysis can serve as a useful reference for future research on VAA effects.

REFERENCES

Aloe, Ariel M and Betsy Jane Becker. 2012. "An effect size for regression predictors in meta-analysis." *Journal of Educational and Behavioral Statistics* 37(2):278–297.

Alvarez, R Michael, Ines Levin, Alexander H Trechsel and Kristjan Vassil. 2014. "Voting advice applications: How useful and for whom?" *Journal of Information Technology & Politics* 11(1):82–101.

Amemiya, Takeshi. 1981. "Qualitative response models: A survey." *Journal of Economic Literature* 19(4):1483–1536.

Andreadis, Ioannis and Matthew Wall. 2014. "The impact of voting advice applications on vote choice." *Matching Voters with Parties and Candidates. Voting Advice Applications in Comparative Perspective* pp. 115–128.

Bhatti, Yosef, Jens Olav Dahlgaard, Jonas Hedegaard Hansen and Kasper M. Hansen. 2019. "Is Door-to-Door Canvassing Effective in Europe? Evidence from a Meta-study across Six European Countries." *British Journal of Political Science* 49(1):279–290.

Boulianne, Shelley. 2009. "Does Internet Use Affect Engagement? A Meta-Analysis of Research." *Political Communication* 26(2):193–211.

Cheung, Mike W-L. 2014. "Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach." *Psychological Methods* 19(2):211–229.

Coleman, Stephen and Jay G Blumler. 2009. *The Internet and democratic citizenship: Theory, practice and policy*. Cambridge University Press.

Enyedi, Zsolt. 2016. "The Influence of Voting Advice Applications on Preferences, Loyalties and Turnout: An Experimental Study." *Political Studies* 64(4):1000–1015.

Fernández-Castilla, Belén, Marlies Maes, Lies Declercq, Laleh Jamshidi, S. Natasha Beretvas, Patrick Onghena and Wim Van den Noortgate. 2019. "A demonstration and evaluation of the use of cross-classified random-effects models for meta-analysis." *Behavior Research Methods* 51(3):1286–1304.

Fivaz, Jan and Giorgio Nadig. 2010. “Impact of voting advice applications (VAAs) on voter turnout and their potential use for civic education.” *Policy & Internet* 2(4):167–200.

Garry, John, James Tilley, Neil Matthews, Fernando Mendez and Jonathan Wheatley. 2018. “Does receiving advice from Voter Advice Applications (VAAs) affect public opinion in deeply divided societies? Evidence from a field experiment in Northern Ireland.” *Party Politics* 0(0):1–8.

Garzia, Diego. 2010. *The Effects of VAAs on Users' Voting Behaviour: An Overview*. Vol. Voting Advice Applications in Europe: The State of the Art ScriptaWeb pp. 13–47.

Garzia, Diego, Alexander H Trechsel and Andrea De Angelis. 2017. “Voting advice applications and electoral participation: a multi-method study.” *Political Communication* 34(3):424–443.

Garzia, Diego and Stefan Marschall. 2012. “Voting Advice Applications under review: the state of research.” *International Journal of Electronic Governance* 5(3-4):203–222.

Garzia, Diego and Stefan Marschall. 2019. Voting Advice Applications. In *Oxford Research Encyclopedia of Politics*. Oxford University Press.

Gelman, Andrew and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Gemenis, Kostas. 2018. “The Impact of Voting Advice Applications on Electoral Turnout: Evidence from Greece.” *Statistics, Politics and Policy* 9(2):161–179.

Gemenis, Kostas and Martin Rosema. 2014. “Voting advice applications and electoral turnout.” *Electoral studies* 36:281–289.

Germann, Micha and Kostas Gemenis. 2019. “Getting Out the Vote With Voting Advice Applications.” *Political Communication* 36(1):149–170.

Gimmler, Antje. 2001. “Deliberative democracy, the public sphere and the internet.” *Philosophy & Social Criticism* 27(4):21–39.

Green, Donald P, Mary C McGrath and Peter M Aronow. 2013. “Field experiments and the study of voter turnout.” *Journal of Elections, Public Opinion and Parties* 23(1):27–48.

Higgins, Julian PT and Simon G Thompson. 2002. "Quantifying heterogeneity in a meta-analysis." *Statistics in medicine* 21(11):1539–1558.

Israel, Jonas, Stefan Marschall and Martin Schultze. 2017. "Cognitive dissonance and the effects of Voting Advice Applications on voting behaviour: evidence from the European Elections 2014." *Journal of Elections, Public Opinion and Parties* 27(1):56–74.

Klein Kranenburg, Laurens. 2015. The impact of voting advice applications on party choice in Dutch national and provincial elections. Master's thesis University of Twente.

Klein Kranenburg, Laurens and Martin Rosema. 2019. "Voting Advice Applications and Electoral Choice: A Typology of Effects." Paper prepared for presentation at the ECPR Conference, Wroclaw, Poland, 5–8 September 2019.

Mahéo, Valérie-Anne. 2016. "The Impact of Voting Advice Applications on Electoral Preferences: A Field Experiment in the 2014 Quebec Election." *Policy & Internet* 8(4):391–411.

Manavopoulos, Vasilis, Vasiliki Triga, Stefan Marschall and Lucas Constantin Wurthmann. 2018. "The Impact of VAAs on (non-Voting) Aspects of Political Participation: Insights from Panel Data Collected During the 2017 German Federal Elections Campaign." *Statistics, Politics and Policy* 9(2):105–134.

Marschall, Stefan and Christian K Schmidt. 2008. Preaching to the converted or making a difference? Mobilizing effects of an internet application at the German general election 2005. In *Non-party actors in electoral politics*. Nomos Verlagsgesellschaft mbH & Co. KG pp. 259–278.

Marschall, Stefan and Martin Schultze. 2012. "Voting Advice Applications and their effect on voter turnout: the case of the German Wahl–O–Mat." *International Journal of Electronic Governance* 5(3-4):349–366.

Munzert, Simon, Pablo Barberá, Andrew M. Guess and JungHwan Yang. Forthcoming. "Do Online Voter Guides Empower Citizens? Evidence from a Field Experiment with Digital Trace Data." *Public Opinion Quarterly* .

Pianzola, Joëlle. 2014. “Selection biases in voting advice application research.” *Electoral Studies* 36:272–280.

Pianzola, Joëlle, Alexander H Trechsel, Kristjan Vassil, Guido Schwerdt and R Michael Alvarez. 2019. “The impact of personalized information on vote intention: Evidence from a randomized field experiment.” *The Journal of Politics* 81(3):833–847.

Ruusuvirta, Outi and Martin Rosema. 2009. Do online vote selectors influence electoral participation and the direction of the vote? In *ECPR General Conference 2009, Potsdam*. pp. 1–22.

Settle, Jaime E. 2018. *Frenemies: How social media polarizes America*. Cambridge University Press.

Sunstein, Cass R. 2007. *Republic.com 2.0*. Princeton, NJ: Princeton University Press.

van de Pol, Jasper. 2016. *Voting wiser: The effect of Voting Advice Applications on political understanding*. University of Amsterdam: PhD Thesis.

Vassil, Kristjan. 2011. *Voting Smarter? The impact of voting advice applications on political behavior*. European University Institute: PhD Thesis.

Viechtbauer, Wolfgang. 2005. “Bias and efficiency of meta-analytic variance estimators in the random-effects model.” *Journal of Educational and Behavioral Statistics* 30(3):261–293.

Viechtbauer, Wolfgang. 2010. “Conducting meta-analysis in R with the metafor package.” *Journal of Statistical Software* 36(3):1–48.

Wagner, Markus and Outi Ruusuvirta. 2012. “Matching voters to parties: Voting advice applications and models of party choice.” *Acta Politica* 47(4):400–422.

Westle, Bettina, Christian Begemann and Astrid Rütter. 2015. “The “Wahl-O-Mat” in the course of the German Federal Election 2013–Effects of a German VAA on users’ election-relevant political knowledge.” *Zeitschrift für Politikwissenschaft* 24(4):389–426.

Meta-Analysis of the Effects of Voting Advice Applications

Online Appendix

CONTENTS

APPENDIX A OVERVIEW OF COLLECTED STUDIES

Figure A1: Geographic distribution of VAAs. The purple filling indicates countries covered by studies that were included in the meta-analysis. *Source:* 2016 Global Census of the ECPR Research Network on Voting Advice Applications (http://vaa-research.net/?page_id=146).

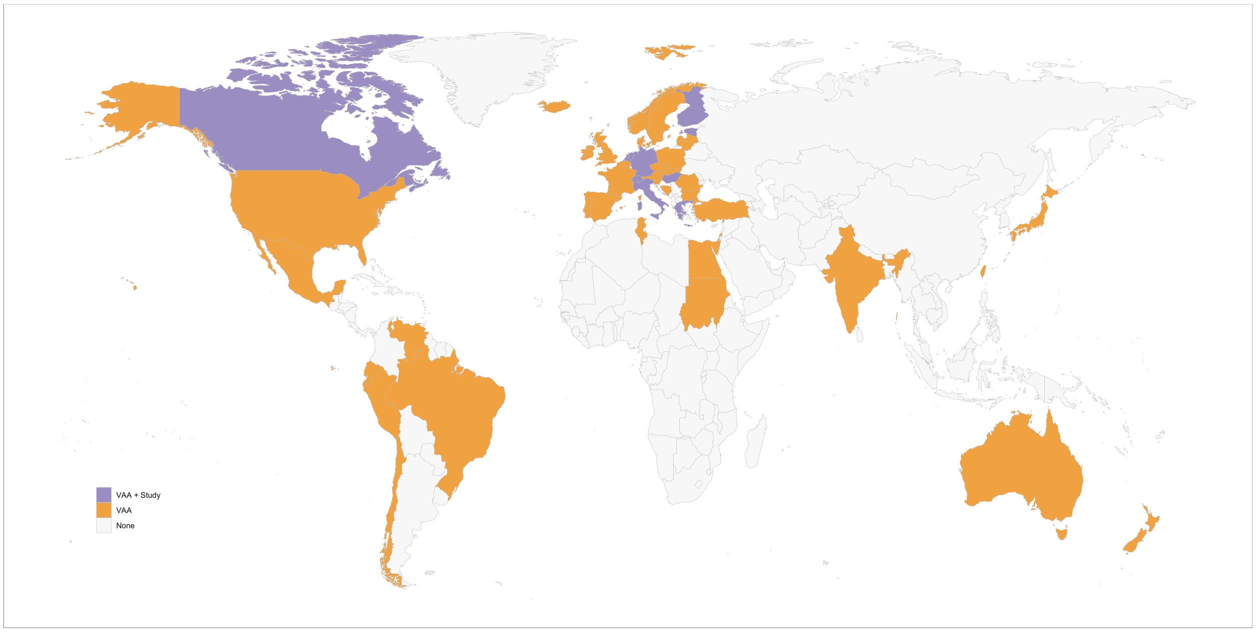


Figure A2: Publication timeline of 22 studies included in the meta-analysis, by study type.

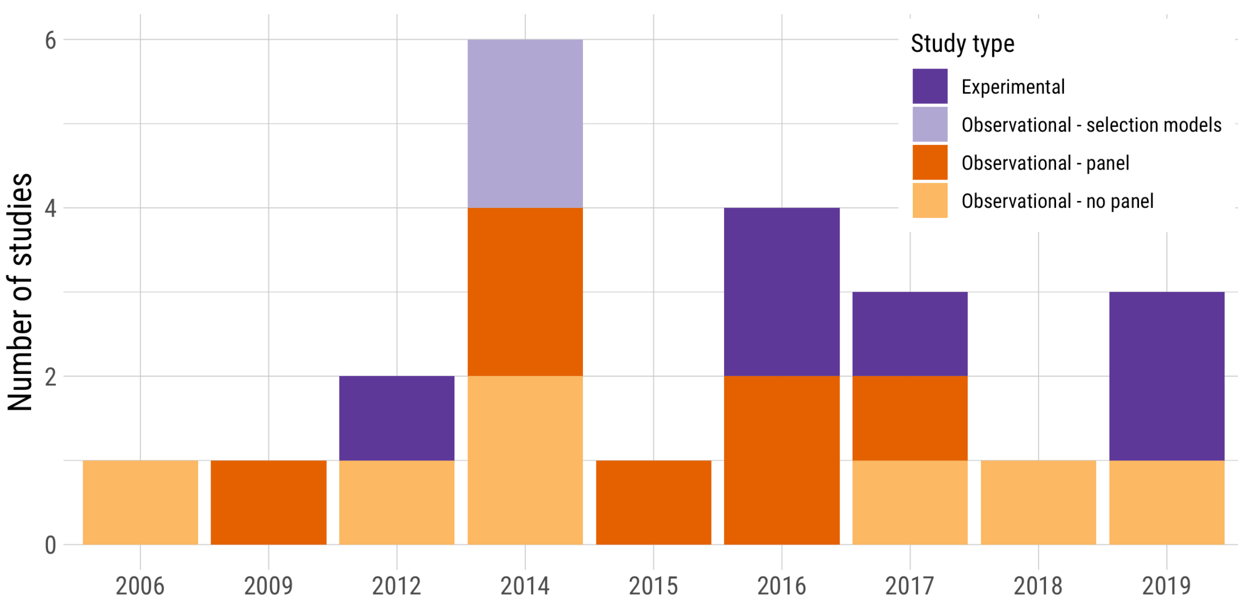


Table A1: Overview of selected within-study effects excluded from the meta-analysis.

Study	Election	Outcome of interest	Exclusion reason
Alvarez, Levin, Trechsel and Vassil (2014)	EU 2009	Perceived utility	IVAR; NOV
Alvarez, Levin, Mair and Trechsel (2014)	EU 2009	Vote choice	IVAR
Boogers (2006)	Netherlands 2006	Self-assessed impact on information seeking	IVAR; NOV
Boogers (2006)	Netherlands 2006	Turnout	IVAR
Boogers (2006)	Netherlands 2006	Vote choice	IVAR
De Rosa (2010)	Italy - EU 2009	Motivation to seek information	IVAR; NOV
De Rosa (2010)	Italy - EU 2009	Issue knowledge	IVAR
De Rosa (2010)	Italy - EU 2009	Turnout	IVAR
Dinas et al. (2014)	EU 2009	Turnout	IVAR
Enyedi (2016)	Hungary 2010	Vote choice	IVAR
Fivaz and Nadig (2010)	Switzerland 2007	Turnout	NMB; IMOV
Fivaz and Nadig (2010)	Switzerland 2007	Vote choice	NMB; IMOV
Garry et al. (2018)	N. Ireland 2016	Party support (propensity to vote for)	IMOV
Garzia et al. (2017)	Finland 2007	Turnout	NISP - Duplicate of Garzia and Angelis (2014)
Garzia et al. (2017)	Finland 2011	Turnout	NISP - Duplicate of Garzia and Angelis (2014)
Garzia et al. (2017)	Germany 2009	Turnout	NISP - Duplicate of Garzia and Angelis (2014)
Garzia et al. (2017)	Netherlands 2003	Turnout	NISP - Duplicate of Garzia and Angelis (2014)
Garzia et al. (2017)	Netherlands 2006	Turnout	NISP - Duplicate of Garzia and Angelis (2014)
Garzia et al. (2017)	Netherlands 2010	Turnout	NISP - Duplicate of Garzia and Angelis (2014)
Garzia et al. (2017)	Switzerland 2007	Turnout	NISP - Duplicate of Garzia and Angelis (2014)
Garzia et al. (2017)	Switzerland 2011	Turnout	NISP - Duplicate of Garzia and Angelis (2014)
Gemenis (2018)	Greece 2015	Turnout	IVAR
Israel et al. (2016)	Germany - EU 2014	Vote choice	IVAR
Israel et al. (2017)	Germany - EU 2014	Vote choice	IVAR
Israel et al. (2017)	Germany - EU 2014	Vote choice	IVAR
Kamoen et al. (2015)	Netherlands 2012	Issue knowledge	IVAR
Kamoen et al. (2015)	Netherlands 2012	Vote choice	IVAR
Ladner and Pianzola (2010)	Switzerland 2007	Turnout	NMB
Ladner et al. (2012)	Switzerland 2007	Vote choice	IVAR
Mahéo (2017)	Canada 2014	Attention to campaign	NOV
Mahéo (2017)	Canada 2014	Information seeking	NOV
Manavopoulos et al. (2018)	Germany 2017	Information seeking	IVAR; NOV
Marschall and Schmidt (2008)	Germany 2005	Information seeking	NOV
Marschall and Schmidt (2008)	Germany 2005	Turnout	NMB
Marschall and Schmidt (2010)	Germany - EU 2009	Turnout	NMB
Marschall and Schultze (2012*b*)	Germany 2009	Turnout	NISP - Duplicate of Marschall and Schultze (2012*a*)
Nuytemans et al. (2010)	Belgium 2009	Vote choice	NSI
Pianzola et al. (2019)	Switzerland 2011	Propensity to vote for most preferred party	IMOV
Ramos et al. (2019)	EU 2014	Turnout	NMB
Ruusuvirta and Rosema (2009)	Netherlands 2006	Vote choice	NSI
Schultze (2013)	Germany 2009	Issue knowledge	NISP - Duplicate of Schultze (2014)
Walgrave et al. (2008)	Belgium 2004	Vote choice	NSI
Wall et al. (2014)	Netherlands 2010	Vote choice	NMB; IMOV
Wang (2016)	Taiwan 2012	Turnout	IVAR
Westle et al. (2015)	Germany 2013	Issue knowledge	NSI

Notes: IVAR: VAA usage not independent variable; NOV: No measure of turnout, vote choice, issue knowledge as outcome; NMB: No meaningful baselines; NSI: Insufficient information about modeling; IMOV: Incompatible measurement of outcome variable; NISP: Not independent separate publication

TABLE A2: Overview of VAA turnout effects studies used.

Study	Election	Sample size	VAA users	Model	Effect (orig)	SE (orig)	Effect (conv)	SE (conv)	Study Design Type
Enyedi (2016)	Hungary 2010	1727	1164	mean differences	3.10		0.29	0.16	Experimental
Garzia and Angelis (2014)	Finland 2007	1062	313	logit	0.53	0.15	0.53	0.15	Obs — no panel
Garzia and Angelis (2014)	Finland 2011	1030	443	logit	0.24	0.15	0.24	0.15	Obs — no panel
Garzia and Angelis (2014)	Germany 2009	1187	116	logit	0.15	0.17	0.15	0.17	Obs — no panel
Garzia and Angelis (2014)	Netherlands 2003	1016	325	logit	0.65	0.27	0.65	0.27	Obs — no panel
Garzia and Angelis (2014)	Netherlands 2006	1793	687	logit	0.46	0.15	0.46	0.15	Obs — no panel
Garzia and Angelis (2014)	Netherlands 2010	1693	706	logit	0.59	0.14	0.59	0.14	Obs — no panel
Garzia and Angelis (2014)	Switzerland 2007	3127	256	logit	0.26	0.07	0.26	0.07	Obs — no panel
Garzia and Angelis (2014)	Switzerland 2011	3254	368	logit	0.32	0.07	0.32	0.07	Obs — no panel
Garzia et al. (2017)	Italy 2013	888	454	mean differences	10.70		0.54	0.15	Experimental
Garzia et al. (2017)	EU 2009	19592	921	logit	1.11	0.11	1.11	0.11	Obs — no panel
Garzia et al. (2017)	Finland 2003	1152	255	logit	1.17	0.31	1.17	0.31	Obs — no panel
Garzia et al. (2017)	Germany 2013	1820	260	logit	1.08	0.33	1.08	0.33	Obs — no panel
Garzia et al. (2017)	Netherlands 2012	1574	631	logit	1.84	0.27	1.84	0.27	Obs — no panel
Gemenis and Rosema (2014)	Netherlands 2006	2356	903	logit	1.44	0.25	1.44	0.25	Obs — panel
Gemenis (2018)	Greece 5/2012	890	NA	entropy bal + LR	1.01	0.59	1.01	0.59	Obs — selection/matching
Gemenis (2018)	Greece 6/2012	896	NA	entropy bal + LR	0.97	0.64	0.97	0.64	Obs — selection/matching
Gemenis (2018)	Greece 1/2015	1013	NA	entropy bal + LR	0.09	0.50	0.09	0.50	Obs — selection/matching
Gemenis (2018)	Greece 9/2015	806	NA	entropy bal + LR	0.03	0.43	0.03	0.43	Obs — selection/matching
Germann and Gemenis (2019)	Switzerland 2007	3578	322	entropy bal + LR	0.60	0.19	0.60	0.19	Obs — selection/matching
Germann and Gemenis (2019)	Switzerland 2011	3657	366	entropy bal + LR	0.67	0.18	0.67	0.18	Obs — selection/matching
Germann and Gemenis (2019)	Switzerland 2015	2931	410	entropy bal + LR	0.64	0.17	0.64	0.17	Obs — selection/matching
Heinsohn et al. (2016)	EU - Germany - EU 2014	449	127	log panel reg with FEs	0.35	0.09	0.35	0.09	Obs — panel
Mahéo (2017)	Canada 2014	277	138	logit	0.40	0.40	0.40	0.40	Experimental
Marschall and Schultze (2012*c*)	Germany 2009	1153	421	logit	0.60	0.25	0.60	0.25	Obs — no panel
Munzert et al. (Forthcoming)	Germany 2017	979	499	IV model	-0.01	0.05	-0.04	0.20	Experimental
Mykkänen and Moring (2006)	EU - Finland - EU 2004	1362	NA	logit	0.63	0.23	0.63	0.23	Obs — no panel
Mykkänen and Moring (2006)	Finland 2003	1511	NA	logit	1.01	0.23	1.01	0.23	Obs — no panel
Ruusuvirta and Rosema (2009)	Netherlands 2006	2356	892	frequency tables	6.00		1.21	0.22	Obs — panel
Vassil (2011)	EU - Estonia - EU 2009	279	97	IV model	0.13	0.06	0.52	0.24	Experimental

Table A3: Overview of VAA vote choice effects studies used.

Study	Election	Sample size	VAA users	Model	Effect (orig)	SE (orig)	Effect (conv)	SE (conv)	Study Design Type
Andreadis and Wall (2014)	Finland 2003	578	153	probit	0.43	0.14	0.68	0.23	Obs — panel
Andreadis and Wall (2014)	Finland 2007	617	188	probit	0.32	0.14	0.52	0.22	Obs — panel
Andreadis and Wall (2014)	Finland 2011	526	237	probit	0.20	0.14	0.32	0.22	Obs — panel
Andreadis and Wall (2014)	Germany 2009	1175	147	probit	0.25	0.12	0.40	0.20	Obs — panel
Andreadis and Wall (2014)	Netherlands 2003	1170	387	probit	0.14	0.09	0.22	0.14	Obs — panel
Andreadis and Wall (2014)	Netherlands 2006	1721	685	probit	0.26	0.07	0.42	0.12	Obs — panel
Andreadis and Wall (2014)	Netherlands 2010	1537	661	probit	0.24	0.07	0.39	0.12	Obs — panel
Andreadis and Wall (2014)	Switzerland 2007	1787	166	probit	0.10	0.12	0.16	0.19	Obs — panel
Andreadis and Wall (2014)	Switzerland 2011	1645	206	probit	0.20	0.10	0.32	0.17	Obs — panel
Enyedi (2016)	Hungary 2010	1727	1124	mean differences			0.05	0.12	Experimental
Klein Kranenburg (2015)	Netherlands 2006	1628	624	logit	0.29	0.13	0.29	0.13	Obs — panel
Klein Kranenburg (2015)	Netherlands 2010	1485	619	logit	0.41	0.16	0.41	0.16	Obs — panel
Klein Kranenburg (2015)	Netherlands 2012	1026	411	logit	0.47	0.18	0.47	0.18	Obs — panel
Klein Kranenburg (2015)	Netherlands 2015	5211	1516	logit	0.48	0.07	0.48	0.07	Obs — panel
Kleinnijenhuis et al. (2019)	Netherlands 2010	1159	426	logit with REs	1.34	0.52	1.34	0.52	Obs — panel
Mahéo (2016)	Canada 2014	211	211	linear regression	-0.24	0.30	-0.24	0.30	Experimental
Munzert et al. (Forthcoming)	Germany 2017	923	483	IV model	0.05	0.07	0.20	0.28	Experimental
Pianzola (2014*a*)	Switzerland 2007	4067	NA	IV model	0.18	0.04	0.72	0.16	Obs — selection/matching
Pianzola (2014*b*)	Switzerland 2011	9163	NA	Heckman + PS matching	0.16	0.02	0.66	0.07	Obs — selection/matching
Pianzola et al. (2019)	Switzerland 2011	1775	1349	IV model	-0.07	0.08	-0.28	0.32	Experimental
Vassil (2011)	EU - Estonia - EU 2009	394	186	IV model	0.15	0.10	0.60	0.40	Experimental

Table A4: Overview of VAA issue knowledge effects studies used.

Study	Election	Sample size	VAA users	Model	Effect (orig)	SE (orig)	Effect (conv)	SE (conv)	Study Design Type
Heinsohn et al. (2016)	EU - Germany - EU 2014	1569	NA	panel regression with FEs	0.23	0.07			Obs — panel
Munzert et al. (Forthcoming)	Germany 2017	979	499	IV model	0.04	0.01	0.16	0.04	Experimental
Schultze (2014)	Germany 2009	1145	440	SEM probit	0.34	0.05	0.55	0.08	Obs — no panel
van de Pol (2016)	Netherlands 2014	5571	3522	entropy bal + LR	-0.06	0.02	-0.06	0.02	Obs — selection/matching

APPENDIX B SUPPORTING TABLES AND FIGURES

Table B1: Turnout study design results of mixed effects moderator analysis

	Estimate	k	Std. Error	Z. Value	Conf. Low	Conf. High	P. Value
Experimental	0.21	5	0.17	1.29	−0.11	0.54	0.2
Observational - no panel	0.8***	15	0.16	4.94	0.48	1.12	< 0.001
Observational - panel	0.91***	3	0.23	3.93	0.46	1.36	< 0.001
Observational - selection/matching	0.56*	7	0.27	2.04	0.02	1.1	0.04

Notes: The estimates are presented in log-odds. The tests for residual heterogeneity and of moderators are significant at the 95%: QE(df = 26) = 133.0338, p-val < .0001, QM(df = 4) = 44.9021, p-val < .0001

Table B2: Vote choice study design results of mixed effects moderator analysis

	Estimate	k	Std. Error	Z. Value	Conf. Low	Conf. High	P. Value
Experimental	0.04	5	0.09	0.45	−0.14	0.23	0.65
Observational - panel	0.4***	14	0.04	10.37	0.33	0.48	< 0.001
Observational - selection/matching	0.67***	2	0.07	10.16	0.54	0.8	< 0.001

Notes: The estimates are presented in log-odds. The test for residual heterogeneity is not significant at the 95%: QE(df = 18) = 15.0816, p-val = 0.6564. The test of moderators is significant: QM(df = 3) = 210.7735, p-val < .0001

Table B3: Turnout country results of mixed effects moderator analysis

	Estimate	k	Std. Error	Z. Value	Conf. Low	Conf. High	P. Value
Finland	0.81***	4	0.16	5.16	0.5	1.12	< 0.001
Germany	0.48***	4	0.17	2.82	0.15	0.82	< 0.001
Netherlands	1.07***	6	0.15	7.23	0.78	1.36	< 0.001
Other	0.43***	11	0.12	3.49	0.19	0.67	< 0.001
Switzerland	0.71***	5	0.15	4.79	0.42	1.01	< 0.001

Notes: The estimates are presented in log-odds. The tests for residual heterogeneity and of moderators are significant at the 95%: QE(df = 25) = 102.3942, p-val < .0001, QM(df = 5) = 58.2662, p-val < .0001

Table B4: Vote choice country results of mixed effects moderator analysis

	Estimate	k	Std. Error	Z. Value	Conf. Low	Conf. High	P. Value
Finland	0.63***	3	0.18	3.57	0.29	0.98	< 0.001
Germany	0.45*	2	0.2	2.29	0.06	0.84	0.02
Netherlands	0.5***	8	0.13	3.83	0.24	0.75	< 0.001
Other	0.07	3	0.19	0.37	−0.31	0.45	0.71
Switzerland	0.43***	5	0.13	3.41	0.18	0.68	< 0.001

Notes: The estimates are presented in log-odds. The tests for residual heterogeneity and of moderators are significant at the 95%: QE(df = 16) = 27.5137, p-val = 0.0361, QM(df = 5) = 19.6308, p-val = 0.0015

Table B5: Turnout election year results of mixed effects moderator analysis

	Estimate	k	Std. Error	Z. Value	Conf. Low	Conf. High	P. Value
2003-2005	0.86***	4	0.23	3.73	0.41	1.32	< 0.001
2006-2008	0.75***	6	0.19	4.02	0.39	1.12	< 0.001
2009-2011	0.63***	9	0.14	4.51	0.36	0.91	< 0.001
2012-2014	0.68***	7	0.18	3.68	0.32	1.04	< 0.001
2015-2017	0.27	4	0.22	1.23	−0.16	0.7	0.22

Notes: The estimates are presented in log-odds. The tests for residual heterogeneity and of moderators are significant at the 95%: QE(df = 25) = 127.1568, p-val < .0001, QM(df = 5) = 41.8469, p-val < .0001

Table B6: Vote choice election year results of mixed effects moderator analysis

	Estimate	k	Std. Error	Z. Value	Conf. Low	Conf. High	P. Value
2003-2005	0.33	2	0.18	1.86	−0.02	0.67	0.06
2006-2008	0.34**	5	0.13	2.55	0.08	0.61	0.01
2009-2011	0.35***	10	0.12	2.88	0.11	0.59	< 0.001
2012-2014	0.33	2	0.2	1.64	−0.06	0.72	0.1
2015-2017	0.44***	2	0.15	2.98	0.15	0.73	< 0.001

Notes: The estimates are presented in log-odds. The tests for residual heterogeneity is significant at the 95%: QE(df = 16) = 44.0260, p-val = 0.0002. The test of moderators is not significant: QM(df = 5) = 11.0442, p-val = 0.0505

Table B7: Turnout election type results of mixed effects moderator analysis

	Estimate	k	Std. Error	Z. Value	Conf. Low	Conf. High	P. Value
First order	0.65***	25	0.12	5.33	0.41	0.89	< 0.001
Second order	0.57***	5	0.18	3.26	0.23	0.91	< 0.001

Notes: The estimates are presented in log-odds. The tests for residual heterogeneity and of moderators are significant at the 95%: QE(df = 28) = 131.7178, p-val < .0001, QM(df = 2) = 32.4878, p-val < .0001

Table B8: Vote choice election type results of mixed effects moderator analysis

	Estimate	k	Std. Error	Z. Value	Conf. Low	Conf. High	P. Value
First order	0.35***	18	0.12	2.98	0.12	0.57	< 0.001
Second order	0.43***	3	0.14	3	0.15	0.7	< 0.001

Notes: The estimates are presented in log-odds. The tests for residual heterogeneity and of moderators are significant at the 95%: QE(df = 19) = 45.5668, p-val = 0.0006, QM(df = 2) = 10.7327, p-val = 0.0047

Table B9: Turnout sample size results of mixed effects moderator analysis

	Estimate	k	Std. Error	Z. Value	Conf. Low	Conf. High	P. Value
Large (>3000)	0.75***	8	0.12	6.09	0.51	0.99	< 0.001
Medium (2000-3000)	0.84***	14	0.13	6.75	0.6	1.09	< 0.001
Small (<1000)	0.27	8	0.14	1.91	−0.01	0.55	0.06

Notes: The estimates are presented in log-odds. The tests for residual heterogeneity and of moderators are significant at the 95%: QE(df = 27) = 133.6281, p-val < .0001, QM(df = 3) = 48.1064, p-val < .0001

Table B10: Vote choice sample size results of mixed effects moderator analysis

	Estimate	k	Std. Error	Z. Value	Conf. Low	Conf. High	P. Value
Large (>3000)	0.52***	3	0.09	5.87	0.35	0.7	< 0.001
Medium (2000-3000)	0.28***	12	0.08	3.4	0.12	0.44	< 0.001
Small (<1000)	0.35**	6	0.12	2.79	0.1	0.59	0.01

Notes: The estimates are presented in log-odds. The test for residual heterogeneity is not significant at the 95%: QE(df = 18) = 26.6059, p-val = 0.0867. The test of moderators is significant: QM(df = 3) = 38.3041, p-val < .0001

Table B11: Turnout sampling type results of mixed effects moderator analysis

	Estimate	k	Std. Error	Z. Value	Conf. Low	Conf. High	P. Value
Non-probability	0.56***	11	0.14	4.08	0.29	0.83	< 0.001
Probability	0.73***	16	0.14	5.2	0.46	1.01	< 0.001

Notes: The estimates are presented in log-odds. The tests for residual heterogeneity and of moderators are significant at the 95%: QE(df = 25) = 133.5, p-val < .0001, QM(df = 2) = 27.99, p-val < .0001

Table B12: Vote choice sample type results of mixed effects moderator analysis

	Estimate	k	Std. Error	Z. Value	Conf. Low	Conf. High	P. Value
Convenience	−0.26	2	0.28	−0.94	−0.8	0.28	0.35
Non-probability	0.5***	9	0.11	4.45	0.28	0.72	< 0.001
Probability	0.42***	10	0.11	3.77	0.2	0.64	< 0.001

Notes: The estimates are presented in log-odds. The tests for residual heterogeneity and of moderators are significant at the 95%: QE(df = 18) = 34.62, p-val = 0.011, QM(df = 3) = 21.14, p-val < .0001

Table B13: Operationalization of vote choice results of mixed effects moderator analysis

	Estimate	k	Std. Error	Z. Value	Conf. Low	Conf. High	P. Value
Between elections	0.46***	16	0.11	4.24	0.25	0.67	< 0.001
In-campaign	0.19	5	0.16	1.23	−0.11	0.5	0.22

Notes: The estimates are presented in log-odds. The tests for residual heterogeneity and of moderators are significant at the 95%: QE(df = 19) = 34.43, p-val = 0.016, QM(df = 3) = 19.45, p-val < .0001

Figure B1: Mixed effects moderator analysis on country of election

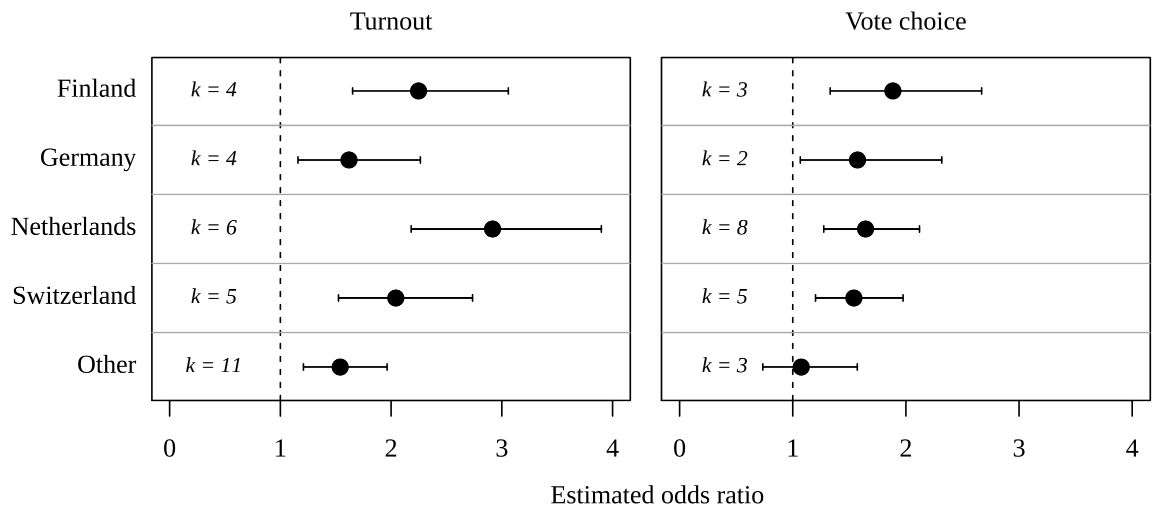


Figure B2: Mixed effects moderator analysis on election year

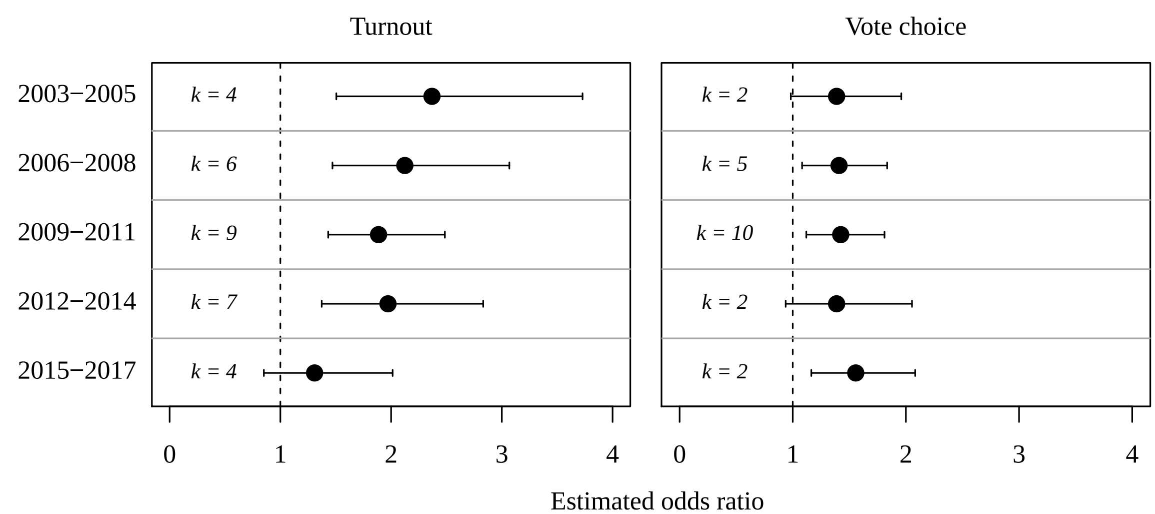


Figure B3: Mixed effects moderator analysis on election type

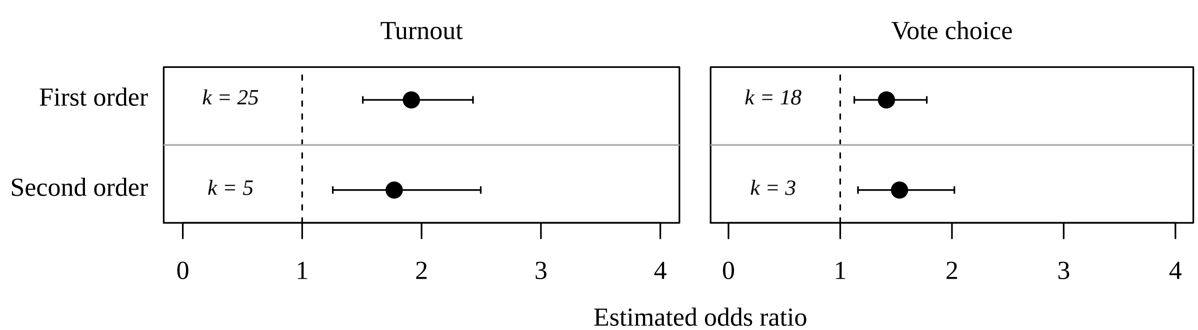


FIGURE B4: Mixed effects moderator analysis on sample size

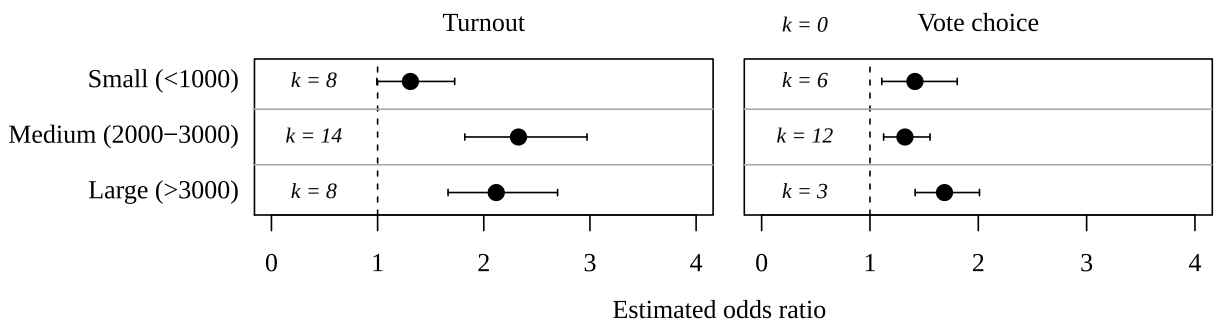


FIGURE B5: Mixed effects moderator analysis on sample type

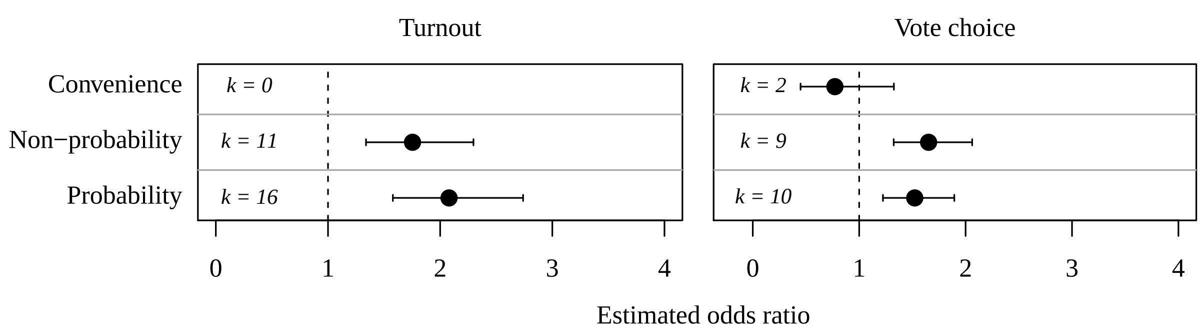


FIGURE B6: Mixed effects moderator analysis on operationalization of vote choice

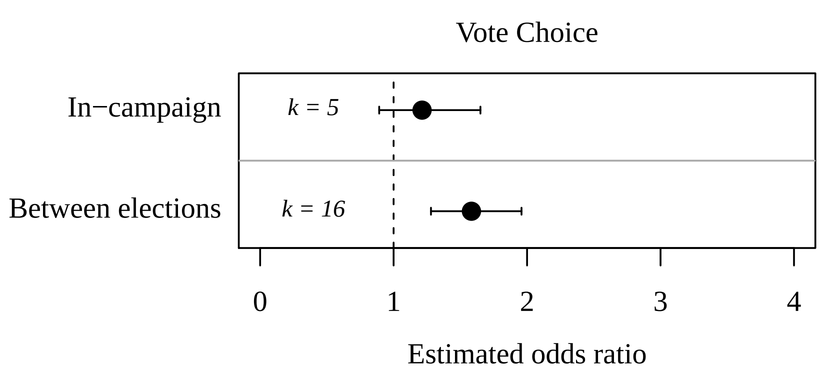


Figure B7: Mixed effects moderator analysis, multiple moderators

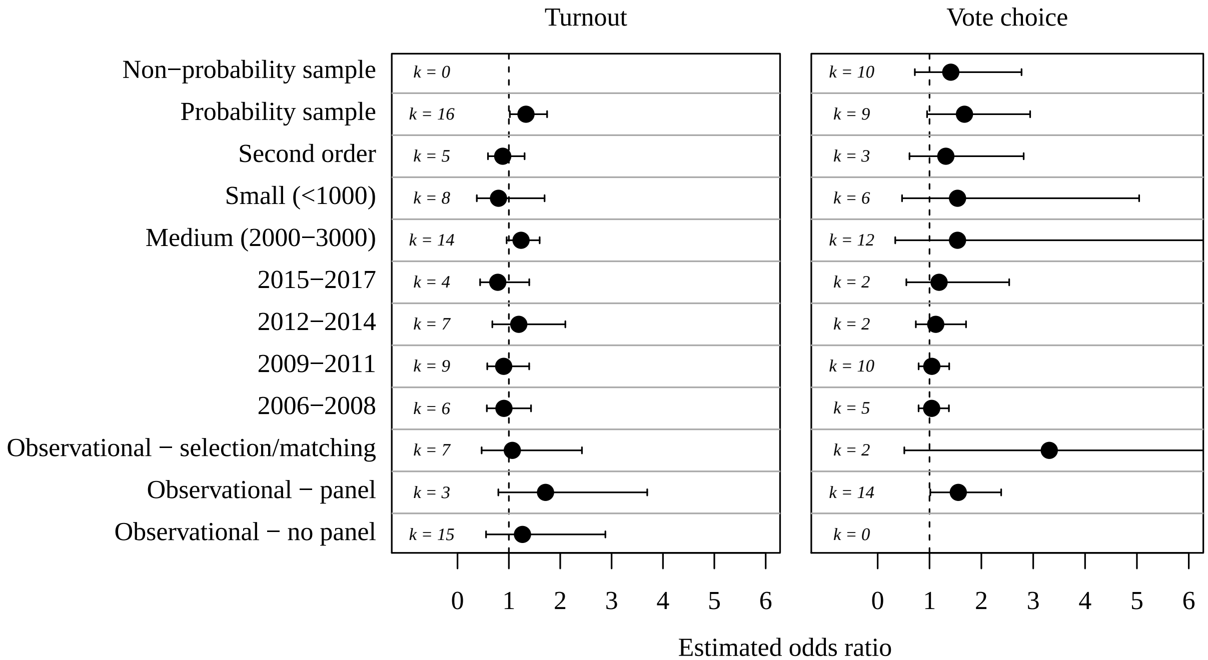
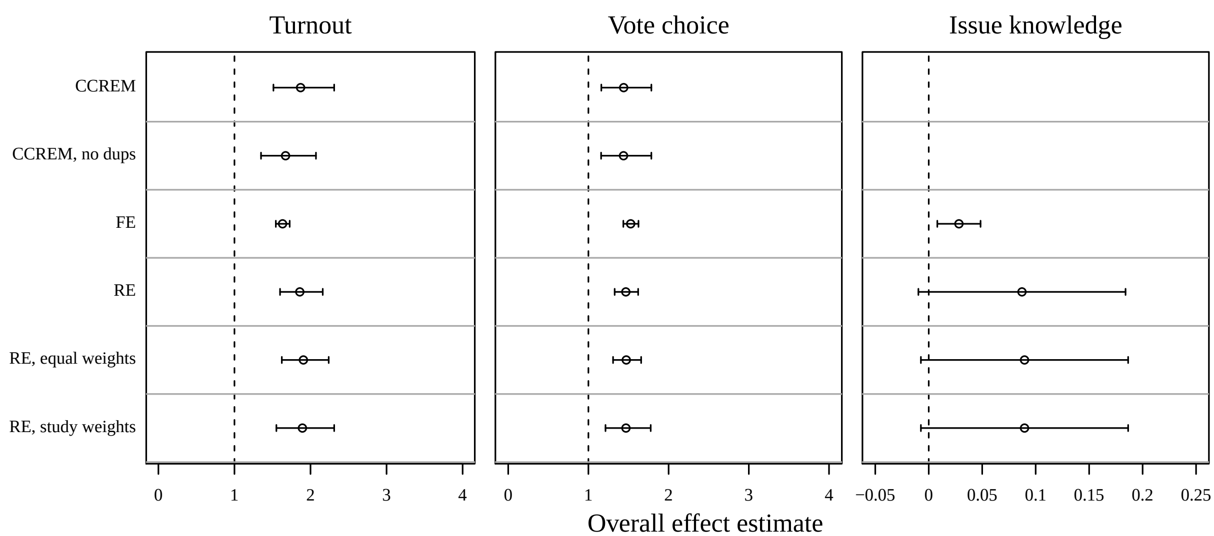


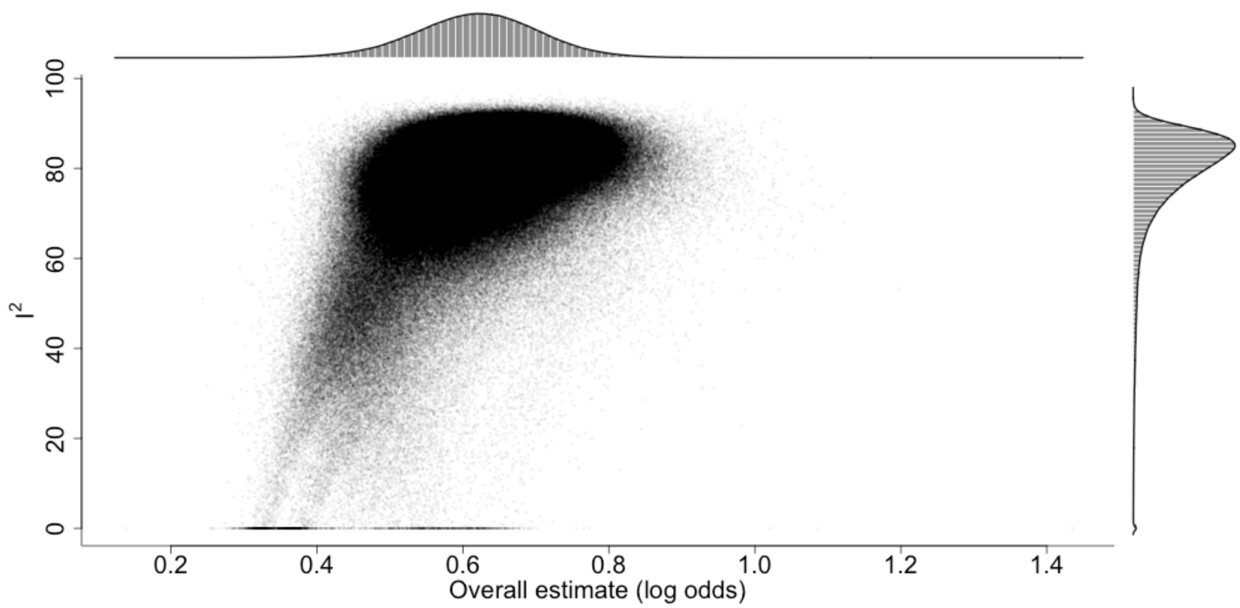
Figure B8: Sensitivity of overall effect estimates across several model specifications, by outcome



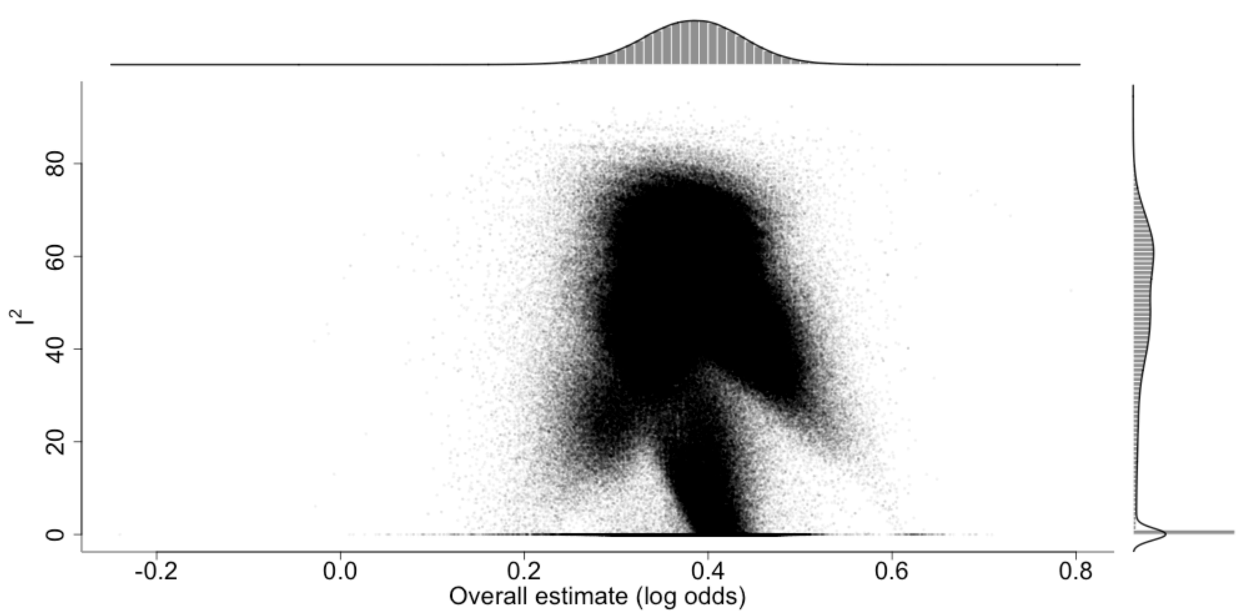
Note: The figure reports overall effect estimates together with 95% CIs across different model specifications by outcome type. The specifications used are (a) cross-classified random effects models (CCREM, as reported in the main text), (b) cross-classified random effects models excluding quasi duplicate studies that re-analyze study data which had been used in predecessor studies (CCREM, no dups), (c) fixed effects models (FE), standard random effects models (RE), random effects models assigning equal weights for effects (RE, equal weights), and random effects models assigning equal weights for studies (RE, study weights).

FIGURE B9: Graphical display of heterogeneity (GOSH) plot based on fixed-effects models in all possible subsets of effects.

(A) Turnout



(B) Vote choice



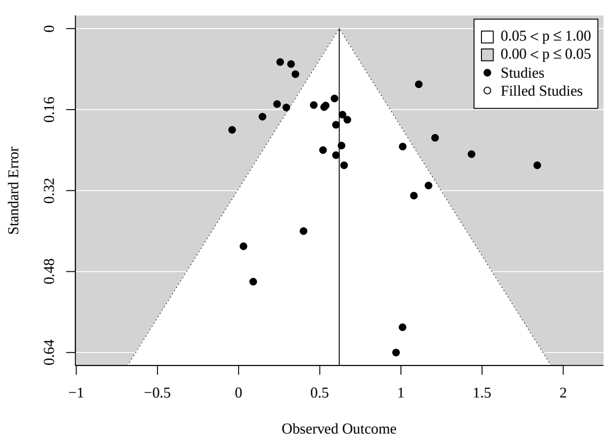
Note: The GOSH plot (Olkin et al., 2012) is based on fixed-effects models computed with all 2^{k-1} possible subsets of effects. The overall estimate is plotted against between-study heterogeneity I^2.

APPENDIX C PUBLICATION BIAS ANALYSIS

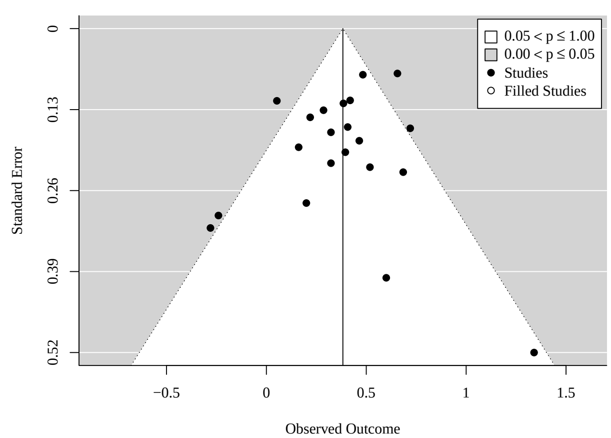
To assess publication bias, we have utilized a trim-and-fill method, which takes a two-step approach to identifying and adjusting for publication bias based on the funnel plot (see Duval and Tweedie, 2000). First, the model trims out small-N studies to obtain a more symmetrical funnel plot and estimates a new summary effect based on the larger-N studies. Second, the model restores the trimmed studies and adds the imputed "missing" counterparts of the effects around the new summary effect estimate. There are no "missing" studies imputed in either case (see Figure C1). The results do not suggest support for potential publication bias for the reported effects. Given that the funnel plot-derived trim-and-fill method assumes publication bias as the only reason for asymmetry, we note that there are various alternative explanations in this case for the observed asymmetry, such as methodological heterogeneity.

FIGURE C1: Trim-and-fill funnel plots, random-effects models

(A) Turnout



(B) Vote choice



APPENDIX D SOFTWARE STATEMENT

The entire analysis was run under OS X 10.15.4 using R version 3.6.2 (R Core Team, 2019). In the empirical analysis, we made use of the following R software packages:

dmetar (Harrer et al., 2019),
dplyr (Wickham and Francois, 2015),
ggplot2 (Wickham, 2016),
haven (Wickham and Miller, 2018),
janitor (Firke, 2018),
magrittr (Bache and Wickham, 2014),
metafor (Viechtbauer, 2010),
readxl (Wickham and Bryan, 2018),
stringr (Wickham, 2015),
writexl (Ooms, 2018), and
xtable (Dahl, 2016).

REFERENCES

Alvarez, R Michael, Ines Levin, Alexander H Trechsel and Kristjan Vassil. 2014. "Voting advice applications: How useful and for whom?" *Journal of Information Technology & Politics* 11(1):82–101.

Alvarez, R Michael, Ines Levin, Peter Mair and Alexander Trechsel. 2014. "Party preferences in the digital age: The impact of voting advice applications." *Party Politics* 20(2):227–236.

Andreadis, Ioannis and Matthew Wall. 2014. "The impact of voting advice applications on vote choice." *Matching Voters with Parties and Candidates. Voting Advice Applications in Comparative Perspective* pp. 115–128.

Bache, Stefan Milton and Hadley Wickham. 2014. *magrittr: A Forward-Pipe Operator for R*. R package version 1.5.

Boogers, Marcel. 2006. "Enquête bezoekers Stemwijzer." *Universiteit van Tilburg* .

Dahl, David B. 2016. *xtable: Export Tables to LaTeX or HTML*. R package version 1.8-2.

De Rosa, Roberto. 2010. "cabina-elettorale. it (Provides advice to Italian voters since 2009)." *Voting Advice Applications in Europe: The State of the Art* pp. 247–257.

Dinas, Elias, Alexander H Trechsel and Kristjan Vassil. 2014. "A look into the mirror: Preferences, representation and electoral participation." *Electoral Studies* 36:290–297.

Duval, Sue and Richard Tweedie. 2000. "Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis." *Biometrics* 56(2):455–463.

Enyedi, Zsolt. 2016. "The Influence of Voting Advice Applications on Preferences, Loyalties and Turnout: An Experimental Study." *Political Studies* 64(4):1000–1015.

Firke, Sam. 2018. *janitor: Simple Tools for Examining and Cleaning Dirty Data*. R package version 1.0.0.

Fivaz, Jan and Giorgio Nadig. 2010. "Impact of voting advice applications (VAAs) on voter turnout and their potential use for civic education." *Policy & Internet* 2(4):167–200.

Garry, John, James Tilley, Neil Matthews, Fernando Mendez and Jonathan Wheatley. 2018. "Does receiving advice from Voter Advice Applications (VAAs) affect public opinion in deeply divided societies? Evidence from a field experiment in Northern Ireland." *Party Politics* 0(0):1–8.

Garzia, Diego, Alexander H Trechsel and Andrea De Angelis. 2017. "Voting advice applications and electoral participation: a multi-method study." *Political Communication* 34(3):424–443.

Garzia, Diego and Andrea De Angelis. 2014. The Impact of Voting Advice Applications on Electoral Participation. In *Matching Voters with Parties and Candidates. Voting Advice Applications in a Comparative Perspective*, ed. Diego Garzia and Stefan Marschall. ECPR Press chapter 8, pp. 105–114.

Gemenis, Kostas. 2018. "The Impact of Voting Advice Applications on Electoral Turnout: Evidence from Greece." *Statistics, Politics and Policy* 9(2):161–179.

Gemenis, Kostas and Martin Rosema. 2014. "Voting advice applications and electoral turnout." *Electoral studies* 36:281–289.

Germann, Micha and Kostas Gemenis. 2019. "Getting Out the Vote With Voting Advice Applications." *Political Communication* 36(1):149–170.

Harrer, Mathias, Pim Cuijpers, Toshi Furukawa and David Daniel Ebert. 2019. *dmetar: Companion R Package For The Guide 'Doing Meta-Analysis in R'*. R package version 0.0.9000.

Heinsohn, Till, Jonas Israel, Stefan Marschall and Martin Schultze. 2016. “Online-Wahlhilfen in Wahlkämpfen.” *Zeitschrift für Politikwissenschaft* 26(3):253–277.

Israel, Jonas, Stefan Marschall and Martin Schultze. 2016. “Kognitive Dissonanz und die Effekte von Voting Advice Applications–Ergebnisse eines Quasi-Experiments anlässlich des Einsatzes des Wahl-O-Mat zur Europawahl 2014 (Cognitive Dissonance and the Effects of Voting Advice Applications–Results from a Quasi-Experiment during Use of the Wahl-O-Mat before the 2014 European Elections).” *Politische Psychologie* (1):115–135.

Israel, Jonas, Stefan Marschall and Martin Schultze. 2017. “Cognitive dissonance and the effects of Voting Advice Applications on voting behaviour: evidence from the European Elections 2014.” *Journal of Elections, Public Opinion and Parties* 27(1):56–74.

Kamoen, Naomi, Bregje Holleman, André Krouwel, Jasper Van de Pol and Claes De Vreese. 2015. “The effect of voting advice applications on political knowledge and vote choice.” *Irish Political Studies* 30(4):595–618.

Klein Kranenburg, Laurens. 2015. The impact of voting advice applications on party choice in Dutch national and provincial elections. Master’s thesis University of Twente.

Kleinnijenhuis, Jan, Jasper van de Pol, Anita MJ van Hoof and André PM Krouwel. 2019. “Genuine effects of vote advice applications on party choice: Filtering out factors that affect both the advice obtained and the vote.” *Party Politics* 25(3):291–302.

Ladner, Andreas, Jan Fivaz and Joëlle Pianzola. 2012. “Voting advice applications and party choice: evidence from smartvote users in Switzerland.” *International Journal of Electronic Governance* 5(3-4):367–387.

Ladner, Andreas and Joëlle Pianzola. 2010. Do voting advice applications have an effect on electoral participation and voter turnout? Evidence from the 2007 Swiss Federal Elections. In *International Conference on Electronic Participation*. Springer pp. 211–224.

Mahéo, Valérie-Anne. 2016. “The Impact of Voting Advice Applications on Electoral Preferences: A Field Experiment in the 2014 Quebec Election.” *Policy & Internet* 8(4):391–411.

Mahéo, Valérie-Anne. 2017. “Information Campaigns and (Under) Privileged Citizens: An Experiment on the Differential Effects of a Voting Advice Application.” *Political Communication* 34(4):511–529.

Manavopoulos, Vasilis, Vasiliki Triga, Stefan Marschall and Lucas Constantin Wurthmann. 2018. “The Impact of VAAs on (non-Voting) Aspects of Political Participation: Insights from Panel Data Collected During the 2017 German Federal Elections Campaign.” *Statistics, Politics and Policy* 9(2):105–134.

Marschall, Stefan and Christian K Schmidt. 2008. Preaching to the converted or making a difference? Mobilizing effects of an internet application at the German general election 2005. In *Non-party actors in electoral politics*. Nomos Verlagsgesellschaft mbH & Co. KG pp. 259–278.

Marschall, Stefan and Christian K Schmidt. 2010. “The impact of voting indicators: the case of the German Wahl-O-Mat.” *Voting advice applications in europe. The state of the art* pp. 61–86.

Marschall, Stefan and Martin Schultze. 2012*a*. “Normalisierung oder Mobilisierung?–Die Auswirkungen politischer Online-Kommunikation auf die Wahlbeteiligung am Beispiel einer Internet-Applikation zur Bundestagswahl 2009.” *Politische Vierteljahresschrift* pp. 444–466.

Marschall, Stefan and Martin Schultze. 2012*b*. “Voting Advice Applications and their effect on voter turnout: the case of the German Wahl–O–Mat.” *International Journal of Electronic Governance* 5(3-4):349–366.

Marschall, Stefan and Martin Schultze. 2012*c*. “Voting Advice Applications and their effect on voter turnout: the case of the German Wahl–O–Mat.” *International Journal of Electronic Governance* 5(3-4):349–366.

Munzert, Simon, Pablo Barberá, Andrew M. Guess and JungHwan Yang. Forthcoming. "Do Online Voter Guides Empower Citizens? Evidence from a Field Experiment with Digital Trace Data." *Public Opinion Quarterly* .

Mykkänen, Juri and Tom Moring. 2006. "Dealigned Politics Comes of Age? The Effects of Online Candidate Selectors on Finnish Voters." Paper presented at the Conference of Politics on the Internet: New Forms of Media for Political Action, Tampere, 24-25 November, 2006.

Nuytemans, Michiel, Stefaan Walgrave and Kris Deschouwer. 2010. Do the vote test: the Belgian voting aid application. In *Voting advice applications in Europe, the state of the art/Cedroni, Lorella [edit.]; ea*. pp. 125–142.

Olkin, Ingram, Issa J Dahabreh and Thomas A Trikalinos. 2012. "GOSH–a graphical display of study heterogeneity." *Research synthesis methods* 3(3):214–223.

Ooms, Jeroen. 2018. *writexl: Export Data Frames to Excel 'xlsx' Format*. R package version 1.1.
URL: *https://CRAN.R-project.org/package=writexl*

Pianzola, Joëlle. 2014*a*. "Selection biases in voting advice application research." *Electoral Studies* 36:272–280.

Pianzola, Joëlle. 2014*b*. "Swing voting due to smartvote use? Evidence from the 2011 Swiss federal elections." *Swiss Political Science Review* 20(4):651–677.

Pianzola, Joëlle, Alexander H Trechsel, Kristjan Vassil, Guido Schwerdt and R Michael Alvarez. 2019. "The impact of personalized information on vote intention: Evidence from a randomized field experiment." *The Journal of Politics* 81(3):833–847.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
URL: *https://www.R-project.org/*

Ramos, Javier, Javier Padilla and Enrique Chueca. 2019. "Abstentionism, Voting Advice Applications and Voting Activation." *Statistics, Politics and Policy* 10(1):55–85.

Ruusuvirta, Outi and Martin Rosema. 2009. Do online vote selectors influence electoral participation and the direction of the vote? In *ECPR General Conference 2009, Potsdam*. pp. 1–22.

Schultze, Martin. 2013. "Effekte des Wahl-O-Mat auf politisches Wissen über Parteipositionen." *ZPol Zeitschrift für Politikwissenschaft* 22(3):367–391.

Schultze, Martin. 2014. "Effects of voting advice applications (vaas) on political knowledge about party positions." *Policy & Internet* 6(1):46–68.

van de Pol, Jasper. 2016. *Voting wiser: The effect of Voting Advice Applications on political understanding*. University of Amsterdam: PhD Thesis.

Vassil, Kristjan. 2011. *Voting Smarter? The impact of voting advice applications on political behavior*. European University Institute: PhD Thesis.

Viechtbauer, Wolfgang. 2010. "Conducting meta-analysis in R with the metafor package." *Journal of Statistical Software* 36(3):1–48.

Walgrave, Stefaan, Peter Van Aelst and Michiel Nuytemans. 2008. "'Do the vote test': the electoral effects of a popular vote advice application at the 2004 Belgian elections." *Acta Politica* 43(1):50–70.

Wall, Matthew, André Krouwel and Thomas Vitiello. 2014. "Do voters follow the recommendations of voter advice application websites? A study of the effects of kieskompas. nl on its users' vote choices in the 2010 Dutch legislative elections." *Party Politics* 20(3):416–428.

Wang, Austin. 2016. The Connection and Effectiveness of iVoter in Taiwan's 2012 Legislative Election. In *Political Behavior and Technology*. Springer pp. 137–155.

Westle, Bettina, Christian Begemann and Astrid Rütter. 2015. "The "Wahl-O-Mat" in the course of the German Federal Election 2013–Effects of a German VAA on users' election-relevant political knowledge." *Zeitschrift für Politikwissenschaft* 24(4):389–426.

Wickham, Hadley. 2015. *stringr: Simple, Consistent Wrappers for Common String Operations*. R package version 1.0.0.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wickham, Hadley and Evan Miller. 2018. *haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files*. R package version 1.1.1.

Wickham, Hadley and Jennifer Bryan. 2018. *readxl: Read Excel Files*. R package version 1.1.0.
URL: *https://CRAN.R-project.org/package=readxl*

Wickham, Hadley and Romain Francois. 2015. *dplyr: A Grammar of Data Manipulation*. R package version 0.4.3.