Does Personalized Information on Issue Congruence Affect Voting Preferences? Evidence from Field Experiments in Five Countries

Kostas Gemenis* Micha Germann† Fernando Mendez‡

May 19, 2021

Abstract

Voting Advice Applications (VAAs) are online tools which provide voters with personalized information on the extent to which their policy views match those of parties or candidates. VAAs have become increasingly prominent features of electoral campaigns in recent years, especially in multi-party systems. However, to what extent whether voters actually make use of VAAs to inform their voting preferences remains unclear. We present new field experimental evidence on the effects of VAAs on voting preferences from five European countries. We find consistent evidence that exposure to VAA advice leads voters to update their voting preferences in line with the information provided. Furthermore, we find partial evidence that VAAs have stronger effects on the preferences of less politically interested and undecided voters. Overall, our results point to the value of VAAs as a mechanism to strengthen democratic representation and accountability.

Word count: 6,100

*Max Planck Institute for the Study of Societies. Email: gemenis@mpifg.de

†Department of Politics, Languages & International Studies, University of Bath. Email: m.germann@bath.ac.uk.

‡Department of Law, University of Zurich. Email: fernando.mendez@uzh.ch

Earlier versions of this paper were presented in seminars at the University of Bath, the University of Cologne, and the Max Planck Institute for the Study of Societies. We would like to thank Iulia Cioroianu, Hilde Coffé, Constantinos Djouvas, Vasileios Manavopoulos, Oliver Proksch, Katerina Vráblíková, Ana Weeks, and Jonathan Wheatley for helpful comments and criticisms.

Introduction

Among the most most established findings in Political Science is that voters tend to have low levels of political information (e.g., Delli Carpini and Keeter 1996). A worrying implication is reduced democratic accountability: effective democratic control requires that citizens conduct detailed evaluations of parties' and candidates' policy platforms and then cast their votes for the party or candidate whose issue positions are closest to their own (Enelow and Hinich 1984). However, acquiring information about the policy stances of parties and candidates is costly and many citizens do not have strong incentives to incur this cost. As a result, many voters do not vote for the party or candidate which best matches their policy interests (Lau et al. 2014). In this article, we study a proposal for a partial remedy to the problem of the uninformed voter: online voter information tools known as Voting Advice Applications (VAAs).

VAAs match voters with parties or candidates based on their policy views. Typically launched during election campaigns, their stated mission is voter education and information. As such, VAAs perform a function similar to the traditional mass media during election campaigns (Krouwel et al. 2014). However, VAAs go beyond newspapers, TV, and radio because they provide voters with *personalized* information on the congruence between their policy preferences and the programs of political parties or candidates. In that sense, VAAs are more similar to campaigning materials, such as leaflets or political ads. Yet contrary to the latter, VAAs are nonpartisan and their developers tend to strive to scientific accuracy. For example, VAA developers often spend considerable energy when it comes to the identification of relevant policy issues (cf. Walgrave et al. 2009) and the coding of party or candidate positions (cf. Garzia et al. 2017; Gemenis 2015).

VAAs have proliferated across advanced democracies in recent years and become integral parts of electoral campaigns, especially in multi-party systems. For example, between 10% and 20% of eligible voters turned to VAAs in the run-up to recent national elections in Canada, Finland, Germany, Greece, New Zealand, and Switzerland, and even larger numbers before recent elections in Denmark and the Netherlands (Germann and Gemenis 2019). Given their significant popularity, it is no surprise that VAAs are increasingly attracting the attention of political scientists, communication scholars, psychologists, and even computer scientists (for a recent review of VAA research cf. Garzia and Marschall 2019). Nevertheless, the answer to

what is perhaps the most foundational question in VAA research—whether voters actually use them to inform their voting preferences—remains unclear.

In this article, we report new field experimental evidence on the effects of VAAs on voting preferences. Contrary to most existing studies, we integrate our experiments directly into actual VAAs, which has the key advantage of allowing us to study the effects of VAAs in the context of natural, real-world VAA usage. As we explain below, our design also allows us to address several other methodological limitations of prior studies including low compliance with treatment allocation and lack of access to the "voting advice" (i.e., the issue congruence scores) shown to users. Our study covers a total of five countries from Eastern (Bulgaria, Romania), Southern (Greece, Spain), and Western (UK) Europe as well as two electoral contexts (supranational and national), rendering this the most extensive experimental evidence on the effects of VAAs on voting preferences to date.

Across all contexts we find consistent evidence that VAA usage causes voters to update their voting preferences in line with the information provided. Furthermore, we find evidence that VAAs have stronger effects on the voting preferences of less politically interested voters and undecided voters. However, the latter findings do not emerge in all cases and we fail to find support for several other sources of effects heterogeneity proposed in the literature including the age and education of VAA users. As discussed in the conclusion, this suggests that individual-level causal heterogeneity remains poorly understood. Still, our results provide clear evidence that voters are influenced by VAAs and use them to inform their voting preferences during election campaigns. Overall, our study suggests that VAAs can make a small but valuable contribution to democratic quality.

Existing Research

The first VAA was developed in the late 1980s in the Netherlands and took the form of a paper-and-pencil test. However, VAAs only really took off with the advent of widespread personal computer and internet usage in the early and mid-2000s (Garzia and Marschall 2019). Soon after, the first studies of the effects of VAAs on voting preferences started to emerge, generally reporting evidence in favor of an effect of VAA usage on voting preferences (e.g., Alvarez et al. 2014; Kamoen et al. 2015; Ruusuvirta and Rosema 2009; Walgrave et al. 2008;

Wall et al. 2014). But while early studies exhibited significant methodological creativity, a key weakness was that their observational research designs could not establish whether changes in voting preferences can, in fact, be attributed to VAA usage (cf. Gemenis and Rosema 2014; Germann and Gemenis 2019; Kleinnijenhuis et al. 2019; Pianzola 2014a). To address endogeneity concerns, VAA research has increasingly turned towards randomized experiments in recent years (Enyedi 2016; Garry et al. 2019; Mahéo 2016; Munzert et al. 2021; Pianzola et al. 2019; Vassil 2011). In stark contrast to the earlier observational studies, most experimental studies found no evidence for an effect of VAAs on voting preferences (cf. the meta-analysis by Munzert and Ruiz 2021).

Despite these frequent null findings, the effects of VAA usage on voting preferences remain unclear. Most existing experimental studies employed encouragement designs whereby a randomly selected subset of subjects is encouraged to use a VAA between two survey waves, each of which includes questions on voting preferences. The key advantage of this approach is that the encouragement to use a VAA is exogeneous, which goes a long way towards disentangling the effects of VAA usage on policy preferences from potential confounders. However, encouragement designs have several important limitations when it comes to the study of VAAs. First, encouragement experiments estimate complier effects, that is, the impact of VAA usage among subjects who used the VAA solely because they were encouraged to do so, and would not have used a VAA otherwise (Eckles et al. 2016). Notably, the encouragement to use a VAA can involve a financial incentive (e.g., Mahéo 2016; Munzert et al. 2021). Perhaps especially in the latter case, it remains unclear to what extent complier effects can be generalized to naturally occurring VAA usage. Notably, this limitation applies even when, as is the standard in the existing literature, subjects are encouraged to use real-world VAAs.

Second, encouragement experiments can run into serious problems due to low compliance with treatment allocation, which in turn can threaten the experiment's internal validity (Germann and Gemenis 2019). Problems with treatment allocation frequently occur because not all subjects who are encouraged to use a VAA also proceed to use a VAA (never-takers). Furthermore, subjects may make use of the VAA even if they are assigned to the control group (always-takers). Always-takers are likely to constitute a significant problem only in countries where VAAs are highly popular. Yet, it is precisely these countries–such as Canada, Germany,

and Switzerland—that have been predominantly studied in the experimental VAA literature. Low compliance with treatment allocation can be countered using instrumental variables estimation. However, especially in countries where VAAs are popular, treatment assignment is bound to be only weakly correlated with VAA use and this can result in biased point and standard error estimates (Murray 2006).

Third, a related problem is that most existing experimental studies have relied on self-reports of compliance with treatment allocation. An issue this raises is social desirability bias: because they were previously asked to use a VAA, subjects in the treatment group may indicate that they used the VAA even if they did not, causing VAA effect estimates to be downward biased (Munzert et al. 2021). Again, this problem may be particularly acute when financial incentives are offered.

Finally, most existing VAA experiments did not have access to the issue congruence scores shown to subjects. Instead, some studies chose to ask subjects what party or candidate was recommended to them in the aftermath of VAA usage. Such recall measures of the "advice" given by VAAs have been shown to be unreliable and possibly influenced by third variables, such as partisanship (Walgrave et al. 2008; Wall et al. 2014). Therefore, reliance on recall measures could bias causal estimates of the effects of VAAs on voting preferences. Other studies have chosen to focus on the effects of VAA usage versus non-usage while ignoring the nature of the advice entirely. While eschewing bias due to endogenous recall measures, this approach makes it difficult to evaluate the full range of the possible effects of VAAs on voting preferences.

In this study, we build on an alternative experimental design recently proposed by Garry et al. (2019) that addresses the above-mentioned concerns. Rather than encouraging VAA use, the central idea is to integrate an experimental manipulation directly into real-world VAAs by randomizing the time at which users are asked about their voting preferences—either immediately before or shortly after receiving the VAA advice. This approach has several important advantages. First, it makes it possible to study VAA effects among actual VAA users instead of individuals who only used a VAA because they were encouraged to do so. Second, it minimizes problems related to compliance with treatment allocation and, third, it ensures access to the voting advice (i.e., the issue congruence scores) shown to users. Evidently, there are also costs to integrating the experiment directly into VAAs because one can only establish short-term

effects. However, randomization of the time when users are asked about their voting preferences can establish the short-term effects of exposure to VAA advice with significantly higher internal, external, and ecological validity.

Using this approach, Garry et al. (2019) found evidence that VAAs affect their users' voting preferences. This could suggest that the frequent null findings in other experimental studies are due to the above-mentioned problems with the currently dominant encouragement design. However, while Garry et al.'s findings certainly point in this direction, their evidence is limited to a single regional election in Northern Ireland. To be sure, Garry et al. are by no means an exception in a literature generally dominated by single country (or region) studies. Nevertheless, generalizability remains a key concern. In this paper, we significantly expand the range of electoral contexts studied and use the approach pioneered by Garry et al. to investigate the effects of exposure to VAA advice across national and supranational elections in five different European countries. The breadth of cases studied easily makes this the most comprehensive study of the impact of VAAs on voting preferences to date.

In addition, we contribute with an extended study of individual-level causal heterogeneity. Past work has identified effects heterogeneity along characteristics such as age, education, and political interest (e.g., Alvarez et al. 2014; Kleinnijenhuis et al. 2019; Mahéo 2016). However, most of the existing evidence on individual-level causal heterogeneity is based on observational research designs. Furthermore, to the best of our knowledge commonly suggested sources of individual-level causal heterogeneity have never been tested in a cross-national framework. We fill this lacuna and use our experimental data to systematically test the following five potential sources of individual-level causal heterogeneity: age, education, political interest, whether voters already have a vote intention, and whether issue positions are an important consideration to voters.

Hypotheses

Our first expectation is that the advice provided by VAAs affects their users' voting preferences. The rationale for this hypothesis is simple: VAAs provide information to voters on how close they are to parties or candidates on a large number of salient political issues. In keeping with standard assumptions from issue voting theory, we expect that VAA users use this information

to re-evaluate their voting preferences (Enelow and Hinich 1984). If users are informed that they are close to a party or candidate on political issues, they will be more supportive of that party or candidate; if they are informed that a party or candidate is far away from them in terms of political issues, they will be less supportive.

H1: VAA users align their voting preferences with the advice they receive.

However, not all users of VAAs may be equally likely to be affected by the advice. First, age has been argued to moderate the relationship between VAAs and voting preferences because younger voters tend to have lower political knowledge and less solidified voting preferences (Pianzola 2014b; Vassil 2011). In keeping with existing literature, we therefore expect that VAAs have stronger effects on the voting preferences of younger voters. Similarly, education and political interest have been argued to act as moderators because less educated and less politically interested voters often have comparatively low levels of political information (Alvarez et al. 2014) and simultaneously because more interested and more highly educated voters may be more critical of the information provided by VAAs (Kamoen et al. 2015). Therefore, we expect that less educated and less politically interested voters are more likely to be influenced by VAAs.

Prior research suggests that some VAA users already have a relatively firm voting intention when they use the tool while others do not, be it because they do not have any strong preference or because they are oscillating between different candidates or parties (Wall et al. 2014; Van de Pol et al. 2014). As voters who have already made up their mind may have a lower need for new information and more generally may be unlikely to revisit their decision, we follow the existing literature in expecting that VAAs have stronger effects on the preferences of undecided voters (Garry et al. 2019; Kamoen et al. 2015; Kleinnijenhuis et al. 2019). Finally, issue positions may be the most important factor determining vote choice for some voters compared to others. For example, some voters may lay greater emphasis on the perceived competence of candidates, their gender, or their ethnic identity. Since VAAs provide information on issue congruence, we expect that they affect voters for whom issues are an important consideration more strongly than voters for whom other considerations are more important (Vassil 2011).

H2: Exposure to VAA advice affects the voting preferences of the following groups more strongly: (a) younger voters, (b) less educated voters,

(c) voters with low political interest, (d) undecided voters, and (e) self-perceived issue voters.

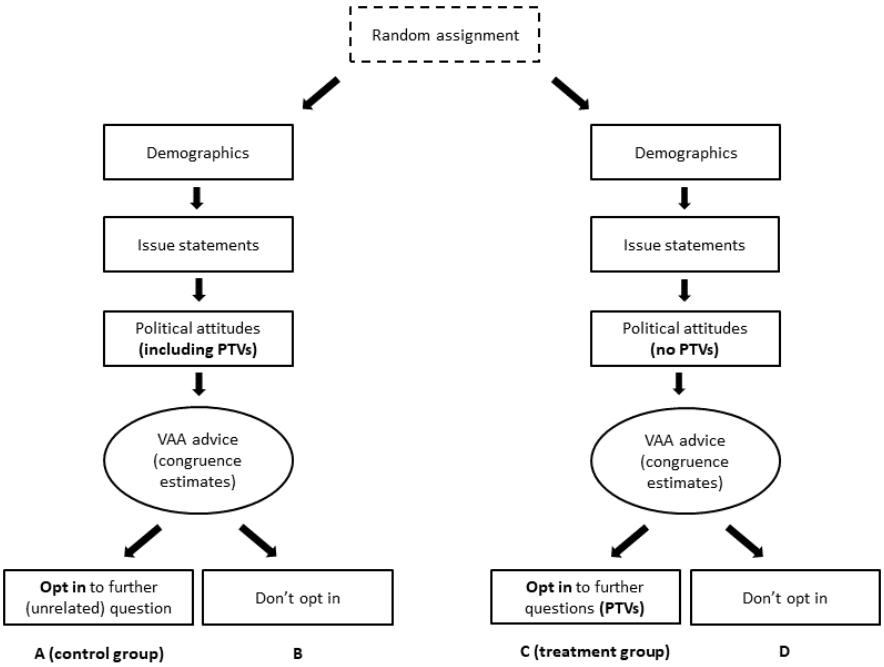
Experimental Design

To test our hypotheses, we integrated randomized experiments into VAAs that were deployed in the weeks prior to the May 2019 elections to the European Parliament (EP) in five countries from Eastern (Bulgaria, Romania), Southern (Greece, Spain), and Western (UK) Europe. For replication in the context of a national election, we repeated the same experiment in the December 2019 UK general election (GE). The VAAs were made freely available online and promoted via print, broadcast, online, and social media.[1] The total number of responses (see Table 1), after removing repeated attempts by the same users and other invalid entries,[2] was between 4,000 (Bulgaria) and 57,000 (UK EP). However, for reasons we detail below, we analyze only a subset of all valid responses.

The design of the VAAs was similar to other VAAs previously deployed in the countries we analyze as well as other countries. Upon accessing our tools, voters were first asked to answer a few general questions on their demographics. Subsequently, users were asked to indicate their preferences on a series of up to 30 policy statements (e.g., "Privatisation leads to a more efficient provision of public services"). Issue statements were carefully selected to reflect salient political issues across a number of policy areas (e.g., economy, immigration, climate change) and varied across countries. Users were asked to indicate their preferences on five-point scales ranging from "completely disagree" to "completely agree", with an additional "no opinion" option. Voter-party issue congruence was estimated by comparing the answers provided by users to the positions of the various parties as estimated by political scientists who examined primary sources through an expert survey. The results were shown in the form of a bar chart indicating the degree of congruence between the user and the various political parties. Congruence scores ranged from -100 (complete disagreement) to +100 (complete agreement), and used a traffic light system to convey the degree of voter-party congruence: scores below 0 were shown in red and flagged as negative matches; scores between 0 and 40 were shown in amber and flagged

[1] In two countries (Greece and UK), the VAAs were promoted using paid advertising on Facebook.
[2] See section 2 of the Supplementary Material for details.

Figure 1: Experimental design



as weak matches; and scores above 40 were shown in green and flagged as strong matches. In practice, scores close to the -100 to +100 extremes were rarely achieved. We provide additional details on the design of our tools including screenshots, the selection of policy issues, the formula used for calculating issue congruence, and the coding of party positions in section 1 of the Supplementary Material.

The experimental manipulation consisted of the *time* when users were asked to provide information on their voting preferences. Users were randomly assigned to a control or treatment group upon accessing the online tool. The control group was asked about their support for parties *before* seeing their policy congruence scores. By contrast, users in the treatment group were asked about their support for parties *after* seeing their issue congruence scores (see Figure 1). This was achieved by a pop-up window prompt that appeared 30 seconds into the results screen, to which we refer as the opt-in page. We measured voting preferences using a battery of 'propensity to vote' (PTV) questions asking how likely it is, on a scale of 0 to 10, that users would vote for the different parties contesting the election (Van der Eijk et al. 2006).

The randomization of the time when respondents were asked to answer PTVs allows for between-subject comparisons of the effects of exposure to new information on issue congruence. However, an issue we are facing is attrition. Some respondents in the treatment group may have already left the website by the time the opt-in page was shown, while others declined to answer

Table 1: Sample descriptives

	Valid responses		Opt-in completed		%		χ^2 test (p-value)
	CT	TR	CT	TR	CT	TR	
Bulgaria (EP)	1895	1935	841	841	44	43	0.57
Greece (EP)	7632	7877	2933	2946	38	37	0.19
Romania (EP)	7140	7388	1553	1583	22	21	0.64
Spain (EP)	5277	5431	1720	1771	33	33	0.99
UK (EP)	28480	28972	10528	10675	37	37	0.77
UK (GE)	23575	24283	6657	6724	28	28	0.18

Note: The null hypothesis in the χ^2 test is that respondents in the treatment and control groups are equally likely to complete the opt-in questionnaire. CT = Control; TR = Treated.

Table 2: Relationship between treatment group status and covariates among opt-in takers

	Mean age (years)		Female (%)		Degree (%)		High interest (%)		Mean left-right (0-10)		Un-decided (%)		Issue voter (%)		F-test (p-value)
	CT	TR	CT	TR	CT	TR	CT	TR	CT	TR	CT	TR	CT	TR	
Bulgaria (EP)	35	37	37	34	76	79	42	42	6.8	6.8	29	25	45	46	0.59
Greece (EP)	39	39	39	39	70	66	36	36	4.3	4.5	25	23	52	54	0.14
Romania (EP)	34	34	37	38	75	75	53	53	6.3	6.3	10	10	48	51	0.73
Spain (EP)	40	39	28	29	69	67	76	74	2.9	2.9	15	17	67	69	0.42
UK (EP)	45	45	47	46	72	72	51	51	3.7	3.7	20	20	69	70	0.22
UK (GE)	43	43	53	52	66	65	56	57	4.2	4.3	22	22	48	47	0.24

Note: F-test = test of the overall significance of a linear regression of treatment assignment on age, gender, education, high political interest, left-right self-placement, being undecided what party to vote for, and being a self-declared issue voter. CT = Control; TR = Treated.

the PTV questions. Therefore, we showed an analogous pop-up window prompt to users in the control group featuring an unrelated question about turnout in the upcoming election. In all our analyses, we restrict the sample to users who completed the opt-in pages (groups A and C in Figure 1). This ensures that attrition is random across the experimental groups analyzed and therefore not a threat to the internal validity of the experiment. As Table 1 shows, opt-in completion rates vary from around 20% to 44%, depending on the country. However, there are no statistically significant differences in opt-in completion rates between control and treated groups. Furthermore, opt-in takers in control and treated groups are balanced in terms of key demographics and political attitudes (see Table 2).

A possible concern with our reliance on opt-in takers is that these could be different from the typical user, limiting the external validity of the experiment. However, as shown in section

3 of the Supplementary Material, the differences between those who opted-in and those who did not are minor in terms of a large number of variables including age, gender, education, and political interest. Another concern could be that our samples tend to over-represent younger, male, highly educated, and highly politically interested voters (see Table 2). But VAAs are commonly used disproportionately by these demographic groups (Marschall and Schmidt 2008; Van de Pol et al. 2014; Vassil 2011). While our samples are therefore clearly not representative of general voter populations, they are representative of VAA user populations. Finally, despite the sample restrictions we need to employ, the relative popularity of our VAAs ensures sufficient statistical power, with the number of analyzable responses ranging from 1,700 (Bulgaria EP) to more than 20,000 (UK EP).

Results

Average Effects Conditional on VAA Advice

We begin by investigating the average effects of VAAs on voting preferences depending on the issue congruence scores shown to users (i.e., the "voting advice"). To do so, we estimate a total of six linear regressions, one for each experiment. The dependent variable in all regressions is the PTV, i.e., the propensity to vote for a given party (0–10). Depending on the electoral context, users were asked to rate between five and nine parties. Accordingly, we perform all analyses on stacked datasets where the unit of analysis is the user x party combination.

The list of independent variables includes a binary indicator of the treatment group to which a user was assigned, a user's VAA issue congruence score for a given party, and the interaction between treatment status and congruence score. Note that as a result of randomization, treated and control subjects received the exact same issue congruence scores on average (see section 4 of the Supplementary Material for supporting evidence). However, only users in the treatment group had been exposed to the congruence scores when indicating their voting preferences. Therefore, a significant-positive interaction term indicates that users in the treatment group aligned their voting preferences with the VAA advice. Because the same individuals are observed multiple times in our data, we cluster standard errors at the subject level. Figure 2 visualizes the results. The regression output is reported in section 5.1 of the Supplementary

Material.

A remarkably consistent picture emerges. Across all cases we find a positive and statistically significant interaction between the treatment indicator and the VAA congruence scores ($p < 0.001$, except for Romania where $p = 0.0021$), suggesting that exposure to VAA advice had a causal effect on users' voting preferences. As Figure 2 shows, the effect sizes are broadly similar across all cases. Specifically, we find that exposure to information that a party constitutes a strong (green) match led to an average increase in the propensity to vote for that party by 0.25 to 0.5 points. At the same time, a negative (red) match decreased the average propensity to vote for that party by up to 0.5 points—though in Bulgaria and Greece the effect of negative matches is statistically significant only at relatively extreme values

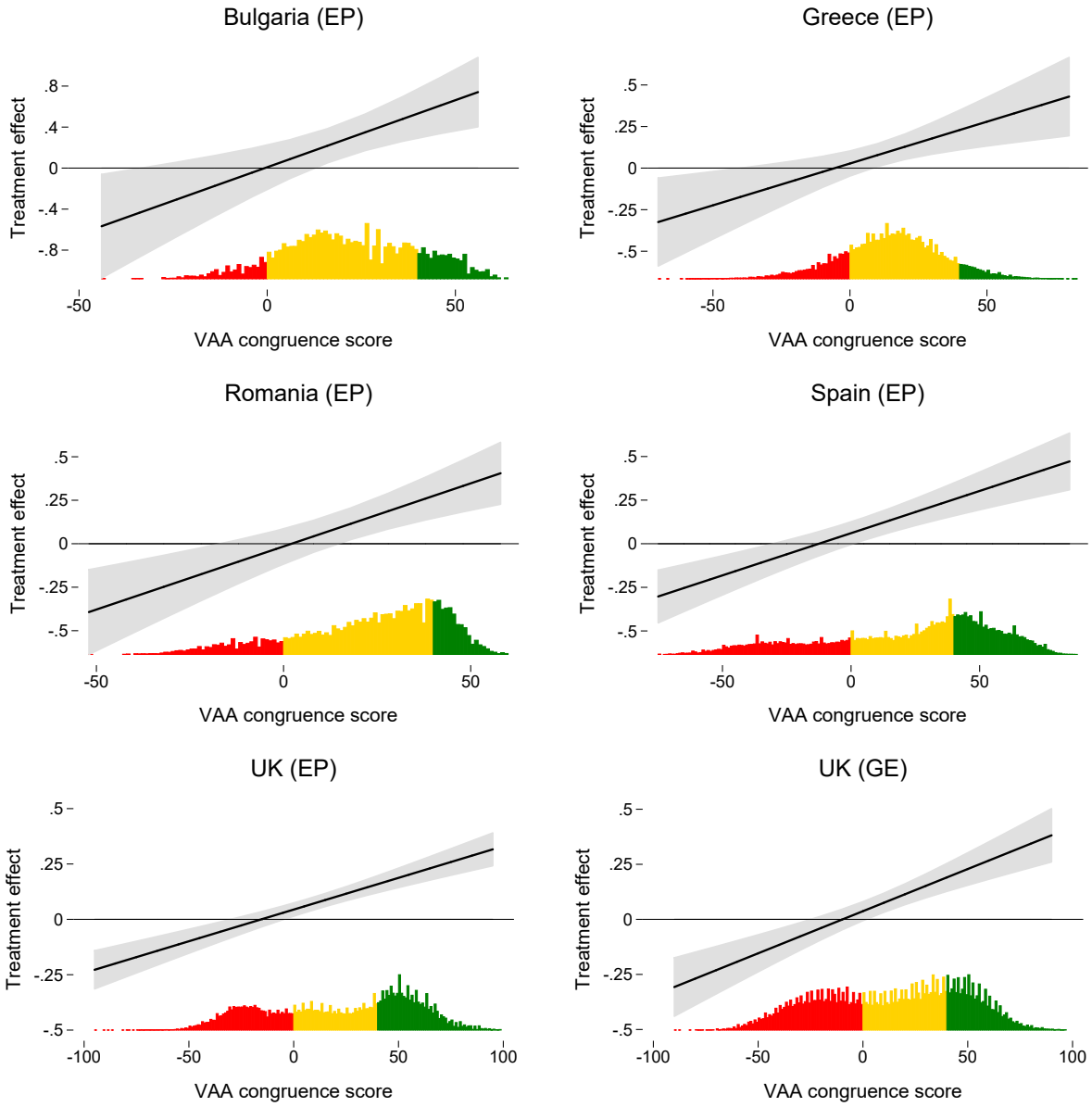
Overall, these results suggest that in line with *H1*, voters use VAAs to inform their voting preferences. Notably, this finding holds across a variety of democracies from Eastern, Southern, and Western Europe. Furthermore, in the case of the UK, the effect estimates are virtually identical in the VAA that was deployed before the EP elections and the VAA that was deployed in the run-up to the general election, suggesting that voters draw on VAAs to inform their voting preferences both in the context of national and supra-national elections. Finally, the results reported here are notably also consistent with Garry et al. (2019), who used the same design to study the impact of VAAs on voting preferences in the context of a regional election in Northern Ireland.

Sub-Group Results

Next, we consider whether the effects of exposure to VAA advice are conditioned by users' age (measured in years), education (university degree vs no degree), and political interest (high vs low). Furthermore, we investigate whether the effects of exposure to VAA advice depend on whether users already had a vote intention before seeing the VAA advice; and whether they see themselves as issue voters. We count users as issue voters if they indicated that the reason for their vote intention was that they are close to the party on political issues, as opposed to other reasons including leader competence or tactical voting. All moderators were measured pre-treatment (see Figure 1).

To investigate individual-level causal heterogeneity, we estimate a total of 30 linear re-

Figure 2: Average effects of VAAs on voting preferences depending on advice



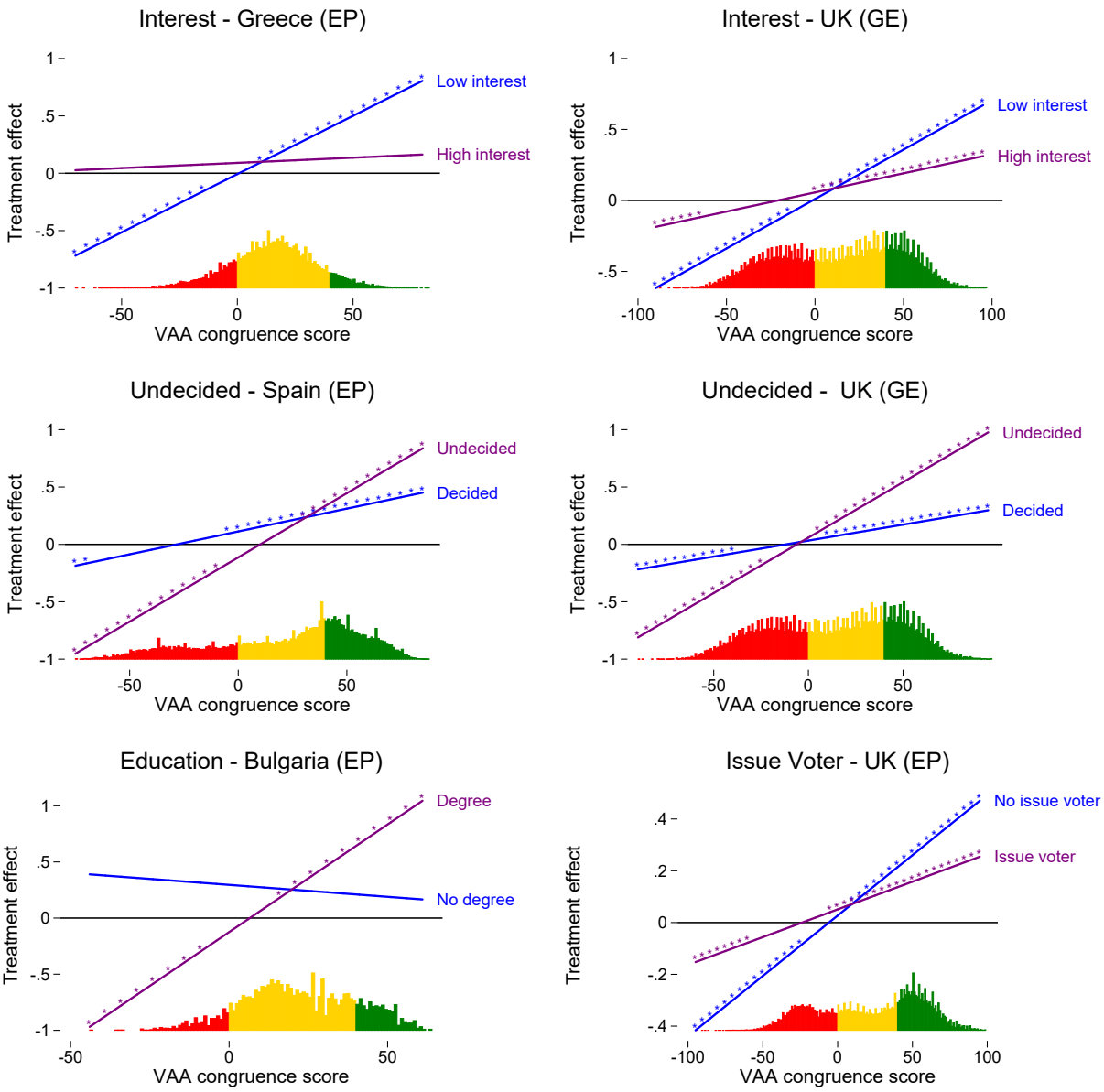
Note: The black lines give point estimates with 95% confidence intervals in gray. The histograms show the number of observations at different levels of agreement with a party. Green = strong match; gold = weak match; red = negative match.

gressions, each including a three-way interaction between the treatment indicator, the VAA congruence score, and one of our five moderators (5 moderators x 6 experiments = 30 models). In all models, the dependent variable remains the propensity to vote and the unit of analysis the user x party combination. Standard errors are clustered at the user level. Evidence for causal heterogeneity emerges when the three-way interaction term is statistically significant ($p < 0.05$). Figure 3 visualizes the results of all models where this is the case. The complete regression output including non-significant results is reported in section 5.2 of the Supplementary Material.

Overall, we find only limited evidence for individual-level causal heterogeneity. The only partial exceptions emerge in the case of our hypotheses about political interest (*H2c*) and undecided voters (*H2d*). Specifically, consistent with expectations we find that exposure to VAA advice affects users with low political interest more strongly in the cases of Greece and the UK; and that undecided voters are more strongly affected in the cases of Spain and the UK (see Figure 3). Taken together, this suggests that voters with lower ex-ante levels of political information as well as voters with unclear preferences are more likely to adjust their voting preferences as a result of VAA usage. However, it is important to note that these results do not replicate in all cases. Furthermore, it is worth noting that while the point estimates suggest that VAAs affect less politically interested and undecided voters more strongly in both UK cases, the differences are statistically significant only in the case of the VAA we deployed before the general election.

At the same time, we have to reject all other hypotheses about individual-level causal heterogeneity. Contrary to *H2a*, we do not find any evidence that younger voters are more affected by VAA advice. Turning to education, we find that the differences between voters with and without a university degree fail conventional levels of statistical significance in 5 of the 6 contexts we study. The only exception is Bulgaria, where we find that exposure to VAA advice had a stronger effect on the preferences of voters with a university degree. The latter result directly contradicts *H2b*, which predicted that voters with higher education should be less affected by VAAs. Similarly, we do not find statistically significant differences between self-declared issue voters and voters who stated that other considerations, such as leader competence, are more important to them in 5 of the 6 contexts we study. The only

Figure 3: Statistically significant three-way interactions



Note: The solid lines give point estimates (* $p < 0.05$). The histograms show the number of observations at different levels of agreement with a party. Green = strong match; gold = weak match; red = negative match.

exception emerges in the case of the UK (EP), where we find that exposure to VAA advice had a weaker effect on self-declared issue voters. Based on the existing literature, we expected the exact opposite and therefore have to reject our *H2e*.

Robustness Checks

We report a series of robustness checks in section 6 of the Supplementary Material including models adjusting for a large set of covariates; models dropping users from the analysis who rushed through our VAAs in super-human speed; and non-linear interaction models. Estimates of the the average effect of VAAs conditional on the VAA advice remain similar irrespective of specification and estimator choices. Substantively similar conclusions also emerge in the case of the sub-group analysis.

Conclusion

VAAs significantly reduce the cost of acquiring information about issue congruence with political parties and candidates. The results of this study suggest that voters in different electoral contexts and from different European countries engage with this information and use VAAs to inform their voting preferences. The average effects of exposure to VAA advice we found may be modestly sized. However, since VAAs are used by millions of voters in some countries even small individual-level effects could have a substantial aggregate-level impact. Furthermore, we found some evidence to suggest that certain groups of voters (i.e., less politically interested and undecided voters) are more strongly affected.

Overall, our results suggest that VAAs make a small but valuable contribution to democratic representation and accountability. Notably, this is contrary to many existing experimental studies, which have tended to report null results. Yet as we have argued, most prior experiments suffered from low compliance with treatment allocation, did not have access to the actual VAA congruence scores, and did not study naturally occurring VAA usage. A limitation of our design remains that we have no means of telling how durable VAAs' effects on voting preferences are. While we believe that our findings increase the plausibility of long-term effects, future research should revisit the durability of VAA effects including the question whether VAAs affect electoral choices.

Another important avenue for future research that our analysis points to is the question of individual-level causal heterogeneity. First, the differences between users with low and high political interest as well as users who do or do not already have a vote intention did not replicate in all contexts. Second, we found no support at all for several other common expectations, such as that VAAs impact younger and less educated voters more strongly. Overall, this suggests that individual-level causal heterogeneity remains poorly understood. Future research should therefore theorize in more detail what kind of voter is most likely to be influenced by VAAs. In our view, a promising way forward would be to think less about individual characteristics and more about combinations of characteristics. VAAs should be most likely to influence individuals who are both in need of *and* receptive to information on issue congruence. Accordingly, our ability to capture individual-level causal heterogeneity may profit from a turn to multi-dimensional measures combining indicators of uncertainty, political interest, and political efficacy (cf. Van de Pol et al. 2014).

Finally, a key implication of our study is that the developers of VAAs need to adhere to the highest ethical and methodological standards. Methodological choices such as what policy statements are included, how they are formulated and aggregated, and how party/candidate positions are measured are known to influence the quality of VAA congruence scores (Gemenis and Ham 2014; Germann et al. 2015; Walgrave et al. 2009). If, as our study suggests, it is true that VAAs exert a measurable influence on voters' preferences, then it is all the more important that VAA developers follow the best practices established in the literature and do so in a transparent way that can be scrutinized by the public as well as the scientific community.

References

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

Delli Carpini, M.X., and S. Keeter. 1996. *What Americans know about politics and why it matters*. New Haven: Yale University Press.

Eckles, Dean, René F. Kizilcec, and Eytan Bakshy. 2016. "Estimating peer effects in networks with peer encouragement designs." *Proceedings of the National Academy of Sciences* 113 (27): 7316–7322.

Enelow, James M., and Melvin J. Hinich. 1984. *The Spatial Theory of Voting: An Introduction*. Cambridge: 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.

Garry, John, James Tilley, Neil Matthews, Fernando Mendez, and Jonathan Wheatley. 2019. "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* 25 (6): 854–861.

Garzia, Diego, and Stefan Marschall. 2019. "Voting Advice Applications." In *Oxford Research Encyclopedia*, ed. by W. Thompson. New York: Oxford University Press.

Garzia, Diego, Alexander Trechsel, and Lorenzo De Sio. 2017. "Party Placement in Supranational Elections: An Introduction to the euandi 2014 Dataset." *Party Politics* 23 (4): 333–341.

Gemenis, Kostas. 2015. "An iterative expert survey approach for estimating parties' policy positions." *Quality & Quantity* 49:2291–2306.

Gemenis, Kostas, and Carolien van Ham. 2014. "Comparing methods for estimating parties' positions in voting advice applications." In *Matching Voters to Parties and Candidates*, ed. by Diego Garzia and Stefan Marschall, 33–47. Colchester: ECPR Press.

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.

Germann, Micha, Fernando Mendez, Jonathan Wheatley, and Uwe Serdült. 2015. "Spatial maps in voting advice applications: The case for dynamic scale validation." *Acta Politica* 50 (2): 214–238.

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.

Kleinnijenhuis, Jan, Jasper van de Pol, Anita M.J. van Hoof, and André P.M. 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.

Krouwel, André, Thomas Vitiello, and Matthew Wall. 2014. "Voting advice applications as campaign actors: Mapping VAA's interactions with parties, media and voters." In *Matching Voters with Parties and Candidates*, ed. by Diego Garzia and Stefan Marschall, 67–78. Colchester: ECPR Press.

Lau, Richard R., Parina Patel, Dalia F. Fahmy, and Robert R. Kaufman. 2014. "Correct Voting Across Thirty-Three Democracies: A Preliminary Analysis." *British Journal of Political Science* 44 (2): 239–259.

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.

Marschall, Stefan, and Christian 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*, ed. by David Farrell and Rüdiger Schmitt-Beck, 259–278. Baden-Baden: Nomos.

Munzert, Simon, and Sebastian Ramirez Ruiz. 2021. "Meta-analysis of the effects of voting advice applications." *Political Communication,* forthcoming.

Munzert, Simon, Pablo Barberá, Andrew Guess, and JungHwan Yang. 2021. "Do online voter guides empower citizens? Evidence from a field experiment with digital trace data." *Public Opinion Quarterly,* forthcoming.

Murray, Michael P. 2006. "Avoiding invalid instruments and coping with weak instruments." *Journal of Economic Perspectives* 20 (4): 111–132.

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

Pianzola, Joëlle. 2014b. "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." *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?*" Paper presented at the 2009 ECPR General Conference, Potsdam, 10–12 September.*

Van de Pol, Jasper, Bregje Holleman, Naomi Kamoen, André Krouwel, and Claes De Vreese. 2014. "Beyond young, highly educated males: A typology of VAA users." *Journal of Information Technology & Politics* 11 (4): 397–411.

Van der Eijk, Cees, Wouter van der Brug, Martin Kroh, and Mark Franklin. 2006. "Rethinking the dependent variable in voting behavior: On the measurement and analysis of electoral utilities." *Electoral Studies* 25:424–447.

Vassil, Kristjan. 2011. "Voting smarter? The impact of voting advice applications on political behavior." PhD thesis, European University Institute.

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.

Walgrave, Stefaan, Michiel Nuytemans, and Koen Pepermans. 2009. "Voting advice applications and the effect of statement selection." *West European Politics* 32 (6): 1161–1180.

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.