Assessing the Electoral Impact of Voting Advice Applications
Experimental Evidence from the Italian Election of 2013

Diego Garzia and Alexander H. Trechsel
European University Institute

<VERY PRELIMINARY VERSION – DO NOT QUOTE>

Introduction[1]

Voting Advice Applications (VAAs) have literally taken Europe by storm in the last decade (Cedroni and Garzia, 2010; Trechsel and Mair, 2011). Millions of voters consult voting recommendations of VAA websites prior to casting their ballots. A growing body of literature supports the idea that VAA bear a non-trivial effect on users' behavior in terms of participation and choice (for a review, see: Garzia, 2010; Garzia and Marschall, 2012). However, observational studies of VAA effects bear severe limitations due to self-selection issues associated with VAA usage, and results are therefore still inconclusive. To overcome the inherent selection bias in VAA usage, we conducted a randomized field experiment in Italy to deduce causal effects. This is embedded in a representative, pre/post electoral web-panel fielded in cooperation with the Italian National Election Study (ITANES). To the best of our knowledge, this is the first time an experimental VAA platform is tested on a nationally representative sample in the context of a real election. In this paper, we will leave aside the issue of party choice to focus on turnout – *To what extent did VAA usage increase*

[1] Paper prepared for delivery at the EUDO Dissemination Conference, San Domenico di Fiesole, 28-29 November 2013.

users' propensity to cast their vote in elections? In order to answer this question, we will briefly review the available literature on VAAs and mobilization. We will then introduce the design of our Italian experiment. The results of the statistical analysis are presented and briefly discussed in the concluding section, where we acknowledge the limitations of this preliminary analysis and highlight further steps to strengthen the solidity of our findings.

Literature review

Early studies of the impact of *Wahl-O-Mat* usage in German Federal elections consistently find one *Wahl-O-Mat* user out of ten declaring to feel more motivated to turnout because of having used that VAA (Marschall, 2005; Marschall and Schmidt, 2010). Further evidence in this direction comes from the Swiss case. An analysis of *Smartvote* 2007 opt-in survey data found about forty per cent of respondents declaring that using the VAA had a "decisive or at least slight influence on their decision to go to the polls" (Ladner and Pianzola, 2010). According to Fivaz and Nadig (2010), the overall turnout in that election could have been about 5 per cent lower had the *smartvote* platform not made available to Swiss voters. A critical issue with the aforementioned studies lies with their exclusive reliance on opt-in surveys administered to users right after having been exposed to the VAA. In other words, the influence exerted by the VAA on users is measured through self-assessment and *only* among those who are willing to fill the opt-in survey. Apart from being subject to a heavy self-selection bias, this kind of data do not even assure that subjective estimates of impact will match with actual changes in terms of preferences and behavior. Indeed, Walgrave *et al.* (2008) found that the reported intention of changing behavior as a result of having used a VAA is not always (nor often) matched with actual changes in voting behavior.

In order to address this critical issue, VAA scholars have turned to mass survey data. Marschall and Schultze (2012) take advantage of a pre-electoral wave of the German

Longitudinal Election Study and find a 6 per cent increase in the probability to cast a ballot among VAA users as compared to non-users. However, their study suffers a low external validity because the dataset employed consists of a quota sample of the German online population. Moreover, the dependent variable is measured before the election, so one cannot be sure whether turnout intentions get actually converted into electoral participation.

To overcome these limitations, a growing number of studies have resorted to national election study data. Working with nationally representative samples increases substantially the external validity of the findings. At the same time, the structure of post-election surveys allows for factual measures of VAA usage (rather than subjective assessments of impact) and actual voting behavior. Ruusuvirta and Rosema's (2009) analysis of the Dutch election of 2006 based on NKO data finds that the massive usage of VAAs among the voting population increased turnout at that election by three per cent (Ruusuvirta and Rosema, 2009: 18). Another analysis by Dinas, Trechsel and Vassil (2011) on EES data shows that even after controlling for a wide set of socio-structural, attitudinal and behavioral variables, the individual-level probability to cast a vote in the EP election of 2009 was 14 percentage points higher for VAA users as compared to non-users.

Cross-sectional survey data has however its limitations when it comes to causal inference. VAA use is not randomly assigned to individuals. It is the respondent, rather than the researcher, who decides whether to use a VAA for the elections or not, thus self-selecting themselves into the "treatment condition" (in this case, using a VAA). If the decision to become a VAA user and the decision to go to the polls have common determinants that are either unmeasured or unknown, estimates from a regular regression model will be biased. This problem could be obviated through the use of Heckman selection models, but it might be extremely complicated to find the required exogenous variables to identify the selection/outcome equations (e.g., Vassil, 2011; Pianzola and Ladner, 2011).

On the basis of the previous discussion, the ideal scenario would seem the random assignment of the treatment in a proper experimental setting. Through randomization of treatment assignment, the researcher can simply compare treatment and control groups. As of now, only two experimental studies of VAA effects have been made available to the research community. Vassil's (2012) analysis of the 2009 Estonian election to the European Parliament finds very weak effects of VAA usage on participation. As his study population consists exclusively of university students, however, the findings are of limited external validity. The only experimental analysis of VAA effects involving a nationally representative sample of voters is that by Pianzola *et al.* (2012) in the context of the Swiss Federal election of 2011. Unfortunately, their study suffers of a very low "first stage". If access to a VAA is open to the public (as it was the case with the Swiss VAA *smartvote* employed in the study) one cannot exclude the possibility that subjects in the control group could take the treatment independently from the experiment. Indeed, over 70 percent of participants in the control group reported to have used *smartvote* independently from the experiment.

The Italian experiment of 2013

To overcome all the limitations stemming from the existing studies, we have set up an experiment in the context of the most recent national elections held in Italy, on February 24th 2013. The Italian case can be considered an ideal "laboratory" for the assessment of VAA-effects in the context of real-world elections. The country is in fact characterized by a very scarce availability of VAAs to voters.[2] Concerns with respect to the first-stage are further minimized by our decision to resort to a "mock" VAA platform. Through an invited accessibility design**,** the experimental VAA platform was in fact accessible only to the

[2] Only one actual VAA was made available to Italian voters prior to the 2013 election: *Voi Siete Qui,* an application whose success in terms of advices provided (i.e., about 700.000) can be considered modest in comparison to the size of the voting population in the country.

respondents in the treatment group. In this way, we were able to overcome the main shortcoming of Pianzola et al.'s study.

The VAA was developed by the EUI research team in collaboration with the University of Siena.[3] The issue statements were based on the salient issues of the campaign as chosen by the research team. The final selection of 30 statements out of the initial 40 was guided by the aims of maximizing variation across parties (that is, we excluded all those questions with a lower discriminating power across party positions) and to be as comprehensive as possible in terms of policy domains [economy: 8, environment: 4, Europe and foreign policy: 4, judiciary: 4, institutions: 3, values: 3, immigration: 4]. The positioning of parties on the various statements was achieved on the basis of a hierarchy of available data sources. Party manifestos were obviously the main source of information. When information about specific issues was not available in party manifestos, we resorted to party websites' content and declaration of party leaders. If none of these sources proved useful, we made use of previous expert positioning endeavours conducted on the Italian case (i.e., ITANES Expert Survey 2011). All parties already represented in Parliament, as well as those with a reasonable chance to attain representation in the 2013 legislature were coded by the research team, for a total of 14 parties included in the VAA. The experimental VAA platform invited respondents to offer their reaction to the 30 issue statements with one of five responses, ranging from "completely agree" to "completely disagree" plus a "no opinion" option.

[3] The EUI research team consists of: Andrea De Angelis, Lorenzo De Sio, Diego Garzia, Davide Morisi, Alexander Trechsel and Kristjan Vassil. The team at the University of Siena helped with the coding of party positions. It was coordinated by Paolo Bellucci and consisted of: Pellegrino Cammino, Simone Cresti, Alba Giovannetti Cicala and Francesco Visconti.

Figure 1. Example of a VAA question (a); the 'voting advice' provided in the results screen (b)

The calculation algorithm was based on the *city-block* method. The visual outcome was the classic match-list, at the top of which stands the party closest to the respondent's policy preference (see Figure 1).

The experiment was embedded in a multi-wave CAWI panel of the Italian National Election Study. The panel design of the study was especially useful for the purposes of the experiment as it allowed not only to measure the outcomes of interest after the election, but also to measure baseline attitudes and behavior before participants' exposure to the treatment.[4] The experimental protocol consisted in three stages, summarized as follows:

Pre-treatment measurement. The pre-treatment measurement was carried out on the entire sample population (N=908) on January 2013 in the context of a longer panel study began by ITANES in 2011. This was actually the fourth wave of the study. The survey included items about respondents' baseline political attitudes (e.g., party identification, left-right orientations, candidate/leader evaluations, economic assessments, and valence-related

[4] The main drawback of CAWI technology lies in the slightly biased demographics of those who tend to respond to online questionnaire invitations. Indeed, youngsters were slightly over-represented in our sample (mean age is 45.5 as compared to the 49.4 in the CATI post-electoral survey) and so were respondents with high educational level (university graduates are 23.4 per cent of the sample as opposed to 12.9 per cent among CATI respondents).

considerations) and behaviour (e.g., willingness to participate in the forthcoming national election, and voting intention).

Randomization and treatment assignment. The sample has been randomly split in halves (N=454). Only the treatment group received, on February 15th, an invitation to take part in the experiment. Upon acceptance, respondents were redirected to our server and asked to perform the VAA test. Response rate was a noteworthy 95.6 percent (N=434). The few respondents that did not perform the VAA test before the opening of the ballots (N=20) have been excluded from the sample.

Post-treatment measurement. The post-treatment measurement was carried out in the context of the fifth, post-electoral wave of the ITANES panel (fieldwork: late February/early March). From the point of view of our experiment, this involved again the entire sample population. The key attitudinal questions remained identical from those in the previous wave in order to achieve full comparability. As to the two core behavioural items, turnout and voting intention have been replaced with their behaviour-recall counterparts (i.e., "did you vote in the last national election?", "what party did you vote for?"). In order to isolate the independent effect of the treatment on users' political behaviour, the post-electoral survey featured a question about VAA-usage during the campaign, in order to exclude from the control group all those users (N=69) who used an actual VAA website during the campaign independently from the experiment.

Table 1 shows the treatment and control group distributions. The statistical analysis that follows will be performed only on those respondents properly "grounded" in their respective condition (i.e., 434 respondents from the treatment group and 385 from the control group, for a total N of 819).

Table 1. Treatment and control group distribution

	Taken	Not taken	Total
Treatment group	**434**	20	454
Control group	69	**385**	454
Total	503	405	908

Data analysis

The focus of the analysis is on the mobilization potential of VAAs. The dependent variable is thus *mobilization* – the difference in the intention to participate in the forthcoming national election as measured in pre-treatment survey and the reported turnout as measured in the post-treatment survey.[5] The variable measures the VAA's capacity to mobilize those who intend to abstain from the elections but subsequently still vote. Therefore, it is coded '1' in all those cases in which the respondent aims to abstain from the elections at t-1, but then decides to participate in elections. The variable is coded '0' for those to whom the intention to participate was equivalent to the reported behaviour after elections (that is, planned to vote and voted, and correspondingly, did not plan to vote and did not vote). There were few observations who intended to vote, but subsequently did not. These 'demobilized' voters (N=23) where coded as missing.[6]

[5] In the pre-treatment survey, respondents' intention to turnout was measured through a 4-point scale ranging from "very likely" to "not at all likely". In order to achieve comparability with the actual turnout variable (dichotomous) we recoded the turnout intention variable coding '1' all respondents declaring themselves "very likely" or "fairly likely" to participate in the forthcoming election. All others have been coded '0' except for the DKs and the NAs which we coded as missing. Due to missing values on either wave, 45 respondents have been excluded from the anlysis.

[6] When including the 'demobilized' to the reference group, the estimation results remained undistinguishable from those presented in this analysis.

Table 2. Patterns of mobilization across treatment/control groups

	Not mobilized	Mobilized	Total
Treatment group	85.5%	14.2%	100%
	(338)	(56)	(394)
Control group	93.6%	6.4%	100%
	(334)	(23)	(357)

The bivariate analysis presented in Table 2 highlights a higher percentage of mobilized voters in the treatment group (14.2 percent) as compared to the control group (6.4 percent). The t-test is statistically significant and points in the expected direction:

t=3,490, p=.001
mean difference = 7,77
95% confidence level=(3,39 – 12,14)

Respondents' pattern of electoral mobilization across the campaign would seem statistically different across treatment and control groups.

In order to control for potential confounding factors that could affect voters' patterns of electoral mobilization, we estimated a number of regression models in order to isolate the *average treatment effect* (ATE) of VAA usage on mobilization (see Table 3). Model 1 estimates the effect of VAA-usage on the dependent variable without any further statistical control. The effect is strong (b=.875) and highly significant ($p < .01$). Most importantly, this effect proves robust to the inclusion of a wide set of controls. In Model 2 we add users' socio-demographic characteristics: age (plus a squared term to control for possible quadratic effects), gender and educational level (three categories from low to high). None of these

Table 3. Logistic regression analysis

	(1)	(2)	(3)	(4)
ATE	**.875**** (,259)	**.878**** (,262)	**.762**** (,271)	**.811**** (,276)
Socio-demographics				
Age	-	-.055 (,046)	-.047 (,047)	-.033 (,048)
Age sq.	-	.000 (,001)	.000 (,001)	.000 (,001)
Education	-	-.087 (,196)	-.051 (,204)	.031 (,206)
Gender (0=male)	-	.441 (,248)	.107 (,263)	-.041 (,275)
Social-psychological resources				
Belongs to a religion	-	-	.764 (,454)	.712 (,460)
Trade union member	-	-	.228 (,445)	.184 (,450)
Strength of PartyID	-	-	-.804** (,149)	-.751** (,155)
Cognitive resources				
Interest in politics	-	-	-	-.411* (,207)
TV news exposure	-	-	-	-.735* (,320)
Constant	-2.673** (,216)	-1.129 (1,103)	-1.044 (1,209)	-.487 (1,233)
Nagelkerke's R-squared	.033	.064	.161	.185
N	750	750	750	750

Note: Cell entries are logistic regression estimates (standard errors in parentheses). ** $p < .01$ * $p < .05$

variables achieve statistical significance and the ATE's coefficient remains virtually unchanged. In Model 3 we control further for socio-structural sources of mobilization. We include two dummies capturing whether respondents are members of a trade union and whether they belong to a religion. Finally, we include the strength of their identification with a political party (if any). Only the latter variable achieve conventional levels of statistical significance. The negative coefficient indicates that the stronger the partisan tie, the lower the probability to be mobilized during the campaign. To our purposes, it is worth noting that the ATE's coefficient barely budges. Inclusion of further controls related to respondents' cognitive resources (i.e., interest in political matters and frequency of TV news watching) do not affect this conclusion. More politically interested appear (similarly to strong partisans) less likely to be mobilized in the course of the campaign. Notwithstanding the inclusion of statistical controls, the average effect of VAA-usage on mobilization remains strong and highly significant.

A few concluding remarks

Our experiment contributes to the research on the mobilizing potential of VAAs through an "ideal" design: an experiment on a nationally representative sample of voters in the context of a real election. The scattered diffusion of VAAs in the Italian context provided the conditions to test VAA effects in a sort of nationwide laboratory. Our results find a significant effect of the treatment of users' patterns of mobilization. Among non-users, only 6 percent would appear to have been mobilized through the campaign. Among users, this percentage hits almost 15 percent. The independent effect of VAA usage, quantifiable in *circa* 7 percentage points, proved robust to multiple statistical controls.

These preliminary findings will benefit from the inclusion of further controls to test other confounding factors. Efforts should be also devoted to the mechanism underlying the

mobilizing effect of VAAs. In this analysis, we only estimated average treatment effects (ATE) assuming a constant effect among subgroups. Yet there are grounds to believe that certain subgroups are particularly "vulnerable" to the mobilizing effect of VAA usage.[7] Looking at other outcomes of interest (e.g., party choice, knowledge effects) is also on our research agenda.

[7] For instance, we found that the exclusion from the sample of all respondents refusing to place themselves on the left-right scale leads the ATE's coefficient close to statistical insignificance (p=.046). This finding is in line with the idea that VAA effects are especially pronounced among unaligned voters, and calls for an assessment of *local average treatment effects* (LATE).

References list

Cedroni, L. and Garzia, D. (eds.) (2010) *Voting Advice Applications in Europe: The State of the Art*, Scriptaweb, Napoli.

Dinas, E., Trechsel, A. H. and Vassil, K. (2011) 'A Look into the Mirror. Preferences, Representation and Electoral Participation'. Paper presented at the *Conference Internet, Voting and Democracy*, 12-13 May 2011. Laguna Beach.

Fivaz, J. and Nadig, G. (2010) 'Impact of Voting Advice Applications (VAAs) on Voter Turnout and Their Potential Use for Civic Education', *Policy & Internet*, Vol. 2 No. 4, pp. 167-200.

Garzia, D. (2010) 'The Effects of VAAs on Users' Voting Behaviour: An Overview', in Cedroni, L. and Garzia, D. (Eds.), *Voting Advice Applications in Europe. The State of the Art*, ScriptaWeb, Napoli, pp. 13-33.

Garzia, D. and Marschall, S. (2012). 'Voting Advice Applications under Review: The State of Research', *International Journal of Electronic Governance* Vol. 5 No. 3/4, pp. 203-222.

Ladner, A. and Pianzola, J. (2010) 'Do Voting Advice Applications Have an Effect on Electoral Participation and Voter Turnout? Evicence from the 2007 Swiss Federal Elections', in Tambouris, E., Macintosh, A. and Glassey, O. (Eds.), *Electronic Participation. Second IFIP WG 8.5 International Conference, ePart 2010, Lausanne, Switzerland, August 29–September 2, 2010. Proceedings*, Springer, Berlin, pp. 211-224.

Marschall, S. (2005) 'Idee und Wirkung des Wahl-O-Mat', *Aus Politik und Zeitgeschichte*, Vol. 55, No. 51-52, pp. 41-46.

Marschall, S. and Schmidt, C. K. (2010) 'The Impact of Voting Indicators: The Case of the German Wahl-O-Mat', in Cedroni, L. and Garzia, D. (Eds.), *Voting Advice Applications in Europe. The State of the Art*, ScriptaWeb, Napoli, pp. 65-104.

Marschall, S. and Schultze, M. (2012). 'Voting Advice Applications and their effect on voter turnout: the case of the German Wahl-O-Mat', *International Journal of Electronic Governance* Vol. 5 No. 3/4, pp. 349-366.

Pianzola, J. and Ladner, A. (2011) 'Tackling Self-Selection into Treatment and Self-Selection into the Sample Biases in VAA Research'. Paper presented at the *European Consortium for Political Research (ECPR) General Conference*. 24-27 August 2011. Reykjavik, Iceland.

Pianzola, J.; Trechsel, A.; Schwerdt, G.; Vassil, K. and Alvarez, M. (2012). 'The Effect of Voting Advice Applications (VAAs) on Political Preferences'. Paper presented at the *American Political Science Association Annual Meeting*.

Ruusuvirta, O. and Rosema, M. (2009) ‘Do online vote selectors influence electoral participation and the direction of vote?’. Paper presented at *The European Consortium for Political Research (ECPR) General Conference*. 10-12 September 2009. Potsdam, Germany.

Trechsel, A. H., Mair, P. and Alvarez, R. M. (2010) ‘The impact of Voting Advice Applications on Voting’. Paper presented at the *Conference Internet, Politics, Policy 2010: An Impact Assessment*. 16-17 September 2010. Oxford, Great Britain.

Vassil, K. (2011) ‘Role of Self Selection in Estimating the Effects of Voting Advice Applications: Empirical Evidence on the Basis of Swiss Smartvote Data’. Paper presented at the *European Consortium for Political Research (ECPR) General Conference*. 24-27 August 2011. Reykjavik, Iceland.

Vassil, K. (2012) ‘Voting Smarter? The Impact of Voting Advice Applications on Political Behavior’, Unpublished doctoral dissertation.

Walgrave, S., van Aelsta, P. and Nuytemans, M. (2008) ‘’Do the Vote Test’: The Electoral Effects of a Popular Vote Advice Application at the 2004 Belgian Elections’, *Acta Politica*, Vol. 43 No. 1, pp. 50-70.