

ISSN 1988-088X



Department of Foundations of Economic Analysis II
University of the Basque Country UPV/EHU
Avda. Lehendakari Aguirre 83
48015 Bilbao (SPAIN)
<http://www.dfaeii.ehu.es>

DFAE-II WP Series

2013-05

Alaitz Artabe & Javier Gardeazabal

Strategic Votes and Sincere Counterfactuals

STRATEGIC VOTES AND SINCERE COUNTERFACTUALS

Alaitz Artabe and Javier Gardeazabal*

University of The Basque Country

October 3, 2012

Abstract

The random utility model assumes voters cast their vote for their preferred party. Strategic voters vote for a party other than their preferred party, as it is not likely to win the elections. Therefore, the existence of strategic voters poses a problem when estimating a random utility model of voting behavior. In this paper, we propose a new method to estimate a random utility model of voting behavior in the presence of strategic voters. Our method can be used to infer the sincere vote of those who vote strategically and provides an estimate of the size of strategic voting. We illustrate the procedure using post-electoral survey data from Spain. Our calculations indicate that strategic voting in Spain is about 2.19 per cent.

Keywords: Strategic voting; self-reported vote; random utility model; voter preferences; multinomial logit.

*We would like to thank Donald Wittman and seminar participants at the University of The Basque Country, University of California at Santa Cruz and the Simposio de la Asociación Española de Economía 2010 in Málaga. Alaitz Artabe acknowledges financial support from the Basque Government through “Programas de Ayudas para la Formación y Perfeccionamiento del Personal Investigador” and from the Spanish Ministry of Science and Innovation (ECO2009-11213). Javier Gardeazabal acknowledges financial support from the Spanish Ministry of Science and Innovation (ECO2009-09120) and the Basque Government (IT-313-07).

Mailing address: Departamento de Fundamentos del Análisis Económico II, Universidad del País Vasco, Lehendakari Aguirre 83, 48015 Bilbao, Spain. E-mail: alaitz.artabe@ehu.es javier.gardeazabal@ehu.es

1 Introduction

A voter's preferred party has frequently no chances of winning the elections. When this is the case, it is often argued that some voters might not vote for their preferred party but another option: the so called strategic voting. Fisher (2004) defines "tactical voters [as those]... who vote for a party they believe is more likely to win than their preferred party, to best influence who wins in the constituency." A more general definition of strategic voting is given by Blais et al. (2005): "The strategic vote is a vote for a party that is not the preferred one, motivated by the intention to affect the outcome of the election".

The analysis of strategic voting rises three very interesting questions. The first one is the issue of the size of strategic voting. A large group of strategic voters might change the outcome of an election while few strategic voters cannot. Although several estimates have been provided in the literature, these estimates remain controversial, leaving some room for further research. A second intriguing question is to know what strategic voters would have voted, had their vote been sincere. Such inference requires a counterfactual estimate that is difficult to obtain. Researchers have used different methods which lead to contentious estimates. A third important issue is the implication of strategic voting to estimate the random utility model of voting behavior. This model assumes individuals vote for their preferred party. However, the existence of strategic voters, who vote for alternatives which do not provide them with the highest utility, does not correspond with the assumed behavior. As a consequence, post-electoral surveys including data on strategic voters cannot be used to estimate random utility models. Alternatively, if such data are used, estimates are bound to be biased. For later reference, we will refer to this bias as the *strategic voter bias*.

In this paper we design a new methodology to estimate strategic voting and contribute answers to the aforementioned three questions. First, our method provides a quantitative assessment of strategic voting. Second, the method can be used to infer the counterfactual

sincere vote of those who vote strategically. Third, the methodology allows us to use survey data to estimate random utility model in the presence of strategic voters, thus avoiding the strategic voter bias.

Measuring strategic voting is difficult and controversial. There are different approaches to measuring strategic voting. The first approach is the aggregate methodology, which uses aggregate election data to measure individual strategic voting (Spafford 1972, Cain 1978, Galbraith and Rae 1988 and Johnston and Pattie 1991). This methodology has been criticized on the grounds that it makes inferences about specific individuals based solely upon the aggregated data for the group, so there is an ecological inference problem or ecological fallacy in that it is assumed that individuals of a group have the average characteristics of the group (Robinson 1950, Achen and Shively 1995, King 1997 and Freedman 1999).

The second approach is the self-reporting methodology, also known as the direct approach. The idea is to use particular items of post-electoral surveys where voters are asked to manifest whether their vote was strategic (Felsenthal and Brichta 1985, Franklin, Niemi and Whitten 1994, Heath, Curtice and Jowell 1991, Kriesi 1998, Lanoue and Bowler 1992, Merolla and Stephenson 2007 and Niemi, Whitten and Franklin 1992). Fisher (2004) is in favor of this procedure arguing that "... tactical voting should ideally be measured directly, rather than indirectly [...] the Heath et al. measure of voting is *a priori* the best available". An advantage of the self-reporting approach is that it provides an easy way to estimate the size of strategic voting by simply calculating the percentage of respondents who declare their vote as strategic. On the negative side, it has been argued that the self-reporting approach hides a bias "...in favor of finding increased levels of strategic voting the further the interview is conducted from election day" (Alvarez and Nagler, 2000). We will refer to this as the *inflation bias*.

A third way of measuring strategic voting is termed the inference approach, also known as the indirect approach. This methodology uses post-electoral surveys to esti-

mate random utility models of voting behavior. Vote choices are modeled as a function of personal and party characteristics as well as variables intended to capture strategic considerations (Abramson et al. 1992, Alvarez and Nagler 2000, Alvarez et al. 2006, Blais and Nadeu 1996, Blais et al. 2001; 2005, Herrmann and Pappi 2008, Lago 2008 and Young and Turcotte 2005). An advantage of this methodology is that, as it includes indicators intended to capture strategic behavior, it can potentially eliminate the *strategic voter bias*. In addition, the estimated random utility party preference model can be used to predict which respondents would have voted differently in the absence of strategic considerations. These properties of the inference approach come at the cost of having to specify indicators measuring strategic voting incentives. To the extent that these indicators are proxy measurements of strategic voting incentives, this procedure introduces a measurement error in the estimation process. Another drawback of this approach is suggested by Blais et al. (2005) who argue that "... 60 per cent of the individuals who are categorized differently by the two approaches, that is, who are deemed to be strategic according to one method and to be sincere according to the other, are respondents whose vote choice is incorrectly predicted by the multinomial probit equation."

Recent in-lab experimental evidence, Meffert and Gschwend (2011), suggests that strategic voting exists and its size is smaller than previously thought. Strategic voting is an active field of study, both theoretical (Myatt 2007) and empirical (Fieldhouse, Shryane and Pickless 2007, Kim and Kostadinova 2011, Kselman and Niou 2010) and is recently expanding into new directions of research (Abramson, Aldrich, Blais, Diamond, Diskin, Indridason, Lee and Levine 2010, Duch and Palmer 2002, Gschwend 2007, MaCuen and Morton 2010, Meffert and Gschwend 2011, Shikano, Herrmann and Thurner 2009, García-Viñuela and Artés 2011).

Our contribution to this expanding field of research is a new estimation method. We combine some of the advantages of the self-reporting and inference methods to provide what we consider to be a better measurement of strategic voting. We term our method,

the counterfactual approach. A similar method has long been used in labor economics to analyze the gender wage gap, e.g. Oaxaca (1973) and Blinder (1973). The idea is to predict the wage women would earn, had they been paid like men. This counterfactual estimate uses the observed characteristics of women and the estimated wage function for men to estimate the wage that women had earned if they had been paid according to the wage function for men. Similarly, we intend to estimate what strategic voters would have voted, had they voted sincerely. In a nutshell, we estimate the random utility model of voting behavior using data from sincere voters and then use the estimated model to predict what would have been the vote of strategic voters, had they voted for their most preferred party. The method is illustrated using post-electoral survey data from Spain. We estimate strategic voting in the 2008 Spanish General Elections to be 2.19 per cent.

Additionally, our research contributes a method that can also be applied in fields other than strategic voting. When some survey respondents vote strategically, traditional estimation methods of the random utility model of voting behavior result in an *strategic voting bias*. Our analysis contributes a methodological recommendation for estimation of the random utility model in the presence of strategic voters: use survey items to identify potentially strategic voters, drop them from the estimation sample and estimate the random utility model with the remaining sample. To the extent that strategic voting is a generalized phenomenon across countries, this recommendation has a great potential for future applications in applied voting analysis.

The rest of the paper is organized as follows. The Spanish post-electoral survey data used is first described. Then, we depict the self-reported, inference and counterfactual approaches respectively and compute estimates of the magnitude of strategic voting. Finally, we discuss and summarize the results obtained.

2 The case

This paper analyses strategic voting in the 2008 Spanish General Elections. Table 1 shows the results of the 2008 Spanish General Elections. Socialist Party (PSOE) and People's Party (PP) obtained more than 80 per cent of the total number of votes and 323 seats out of the 350 total seats in Congress. The Spanish electoral system results in a relatively small representation of United Left (IU) in terms of seats, as compared with regional parties. IU votes are spread out all over Spain, whereas regional parties have their vote concentrated in some provinces (constituencies).¹ Although only PSOE and PP had real chances to win the elections, minority parties also play an important role in congress, as they are usually pivotal. Therefore, the distribution of seats among parties, and particularly among the minority parties, is likewise very important.

The first, and arguably the largest, source of strategic voting in Spanish General Elections may come from voters whose preferred party is IU and might have incentives to vote for PSOE, which is closer than PP in the left-right dimension (García-Viñuela and Artés 2009, 2011, Lago Peñas 2005). Conceivably, those whose preferred party is IU could also vote strategically for PP, but this seems to be less likely. A second sizable source of strategic voters may also accrue from voters whose preferred party is a regional party and have incentives to vote for PSOE or PP, which are the only parties with a significant probability of winning the general election. A third, albeit smaller, source of strategic voting may correspond to voters with a strong anti-nationalist sentiment, who might vote for PSOE or PP, not because they are their first option, but to avoid a regional party obtaining an additional representative. With all these options available, strategic voting in Spain is a real possibility worth exploring.

The empirical evidence reported in this paper makes use of post-electoral survey (code

¹Each constituency (province) has a minimum of two seats (except Ceuta and Melilla with one seat). In this way, 102 out of 350 seats are assigned to the constituencies. The remaining seats are assigned to constituencies in proportion to population. Seats are assigned to parties using the D'Hont method.

2757) carried out by CIS (*Centro de Investigaciones Sociológicas*)² in 50 provinces.³ The analysis is restricted to parties with at least one seat in congress and 15 or more respondents who declared they voted for that party. Using this criterion results in a sample of individuals who declared they voted for the national parties with candidates in all constituencies (PSOE, PP, IU and UPyD), regional parties in some constituencies (CIU and ERC in Catalonia, PNV in the Basque Country, BNG in Galicia and CC in Canary Islands) as well as those who abstained. In order to give an idea of the location of parties in a left-right scale, the last column of Table 1 reports the average scores of voters on the left-right scale, with 1 at the extreme left and 10 at the extreme right. For instance, individuals who voted for PSOE had, on average, a score of 3.62 in the left-right scale and those who voted for PP 6.56.

The survey data includes 6,083 interviews and records individual characteristics and electoral choices. Only 5,240 respondents declared their vote. Many respondents did not answer other items that we use for inference, leaving a working sample of 2,146 observations with a declared political choice and data on all variables used as explanatory. Therefore, the question arises whether the results obtained using the working sample can be extrapolated to the Spanish population. Table 2 shows that the non-response issue does not fundamentally alter vote shares, except for the under representation of abstention in the working sample. A detailed description of all variables used in the analysis is given in the appendix.

3 The self-reporting approach

This approach uses a direct way of identifying strategic voting. It makes use of survey items to identify strategic voters. In our particular application of this method to the 2008 Spanish General Elections, we use items 36a and 36b of the post-electoral survey. Re-

²The data are available at <http://www.cis.es>.

³The two North-African Spanish constituencies Ceuta and Melilla are not covered in this survey.

spondents who manifested they voted for PSOE or PP were asked the following question: “Which is the main reason for voting PSOE / PP?”. Respondents had to choose one of the following options:

1. Because of José Luis Rodríguez Zapatero / Mariano Rajoy,
2. Because it is my party and I always vote for it,
3. Because it is the most qualified party to govern,
4. Because it is the party that best represents my ideas,
5. Because of its policy conduct during the last four years,
6. To avoid PP / PSOE winning the elections,
7. Other reasons,
8. No answer.

Applying the self-reporting method, we identify strategic voting as those who selected option 6, i.e. to avoid the other party winning the elections. Table 3 shows that 188 out of the 2146 individuals voted strategically. Strategic voting amounts to 8.76 per cent of the working sample. According to this table, PSOE received 113 strategic votes, and PP 75. Alvarez and Nagler (2000) argue that there seems to exist a post-electoral bias in self-reported declarations in favor of finding increasing levels of strategic voting the further the survey from the election day. This is what we term *inflation bias*. If this bias actually exists, the self-reporting approach estimates of strategic voting would be upward biased.

4 The inference approach

The inference approach uses the random utility model to infer strategic voting, thus our terminology “inference approach”. The random utility model assumes that, given a set of

alternatives $A = \{1, \dots, J\}$ which includes abstention, the utility individual i obtains from alternative j , U_{ij} , is given by

$$U_{ij} = X_i\beta_j + Z_{ij}\theta + \varepsilon_{ij}, \quad (1)$$

where X_i is a $1 \times K$ vector of individual characteristics; Z_{ij} is a $1 \times M$ vector of alternative j attributes from the point of view of individual i ; β_j and θ are parameter vectors and ε_{ij} is a zero-mean random disturbance. According to this model, individual i votes for alternative j if and only if $U_{ij} > U_{il}$ for all $l \neq j$ (Alvarez and Nagler 1995). This model defines the “preferred alternative”, so strategic voting would deviate from the prediction of this model.

Alvarez and Nagler (2000) modify the random utility model to account for strategic behavior by adding a third group of variables to capture the incentives voters might have to vote strategically. They include two strategic voting indicators, a measure of how far voter’s first option is from the most voted party and a measure of the distance between the two most voted parties. Their specification is

$$U_{ij} = X_i\beta_j + Z_{ij}\theta + W_{ij}\gamma + \varepsilon_{ij} \quad (2)$$

where $W_{ij} = [W1_{ij}, W2_{ij}, W1_{ij} \times W2_{ij}]$ is a vector consisting of the two variables capturing strategic voting incentives as well as their interaction.

For each individual i , define the sets of alternatives H_i and \bar{H}_i as follows. Set H_i includes the two parties expected to receive most votes in the constituency of voter i , and set \bar{H}_i is the union of H_i and abstention. Variable $W1_{ij}$ is defined as

$$W1_{ij} = \begin{cases} \max_l S_{il} - S_{ij} & \text{if } j \notin \bar{H}_i \\ 0 & \text{if } j \in \bar{H}_i \end{cases}$$

where S_{ij} is the expected vote share of party j and $\max_l(S_{il})$ is the maximum expected vote share in the constituency of individual i . This variable measures how far behind the voter's first choice is in the voter's constituency. In single-member district plurality systems, $W1_{ij}$ measures the percentage vote needed by parties to obtain the seat in the constituency. Although in the case of a multiple-member district plurality system, as in the Spanish General Elections, $W1_{ij}$ does not have the same interpretation, it still measures how far behind alternative j is in the constituency of voter i . The further the distance of alternative j to winning the election, the higher the incentive to vote strategically is. Therefore, this variable should enter the estimating equation with a negative sign.

Variable $W2_{ij}$ is defined as

$$W2_{ij} = \begin{cases} \frac{1}{|S_{i1} - S_{i2}|} & \text{if } j \notin H_i \\ 0 & \text{if } j \in H_i \end{cases}$$

where S_{i1} and S_{i2} are the two largest expected vote shares. The term $|S_{i1} - S_{i2}|$ measures how close the race is expected to be in the constituency of voter i . Note that individuals whose preferred option is abstention may be interested in voting for one of the two most voted parties to avoid the other party winning the elections. The closer the race, the higher the incentives to vote strategically are, and therefore, the lower the incentives to vote for alternative j . Hence, intuition calls for a negative effect of $W2_{ij}$ on the probability of voting for alternative j .

In practice, computation of $W1_{ij}$ and $W2_{ij}$ is carried out using actual vote shares in the previous elections as proxy measurements of expected vote shares. Strategic voting is estimated in four steps. First, estimates of the parameters in equation (2), say $\hat{\beta}, \hat{\theta}, \hat{\gamma}$, are obtained by multinomial probit or logit. Second, we compute the *predicted strategic vote* of individual i as the alternative j_i^{stra} that maximizes the (estimated) probability of

observing that choice, mathematically

$$j_i^{stra} = \arg \max_{l \in A} \Pr [j_i = l | X_i, Z_{il}, W_{il}, \hat{\beta}_l, \hat{\theta}, \hat{\gamma}]. \quad (3)$$

Third, we compute *predicted sincere vote* of individual i as the alternative j_i^{sinc} such that

$$j_i^{sinc} = \arg \max_{l \in A} \Pr [j_i = l | X_i, Z_{il}, 0, \hat{\beta}_l, \hat{\theta}, \hat{\gamma}], \quad (4)$$

where the variables measuring strategic voting are set equal to zero. Fourth, strategic voters are identified as those whose predicted strategic and sincere votes are different, that is, $j_i^{stra} \neq j_i^{sinc}$.

Estimation of the random utility model requires specification of the probability function. There are two estimators available: the multinomial logit and probit. Table 4 reports the multinomial probit and logit estimates of the random utility model (2) when we restrict the analysis to PSOE, PP, IU and abstention. The results are very similar quantitatively and in terms of statistical significance and goodness of fit. Henceforth we use the multinomial logit model. In addition to the similarity of results obtained in our particular application, there are other reasons for our choice. First, Dow and Endersby (2004) compare the multinomial logit and probit in the context of voter choice in multiparty elections. They conclude that the multinomial logit is often preferable to the multinomial probit arguing that although the multinomial logit model imposes the restriction of Independence of Irrelevant Alternatives (IIA), “its simplicity in estimating the model prevails the sensitivity of multinomial probit to a number of estimation problems.” Second, the multinomial logit model allows us to increase the number of alternatives available for each individual voter. This is not feasible with the multinomial probit which requires numerical integration in a space of dimension equal to the number of alternatives. Third, a larger set of alternatives allows us to determine whether some of the strategic vote comes from regional parties. Fourth, increasing the number of available alternatives makes the assumption of IIA less

of a concern.

Table 5 reports the multinomial logit estimates including regional parties. PSOE is the reference alternative, thus all coefficients are to be interpreted relative to PSOE. Union membership and gender are not significant for any of the parties considered. Income dummies are not significant with the exception of the upper income class dummy for PNV, which is marginally significant and has negative sign. Years of education is significant for UPyD and marginally for CIU. Not surprisingly, religion is a significant determinant of the PNV vote, as PNV was a founder of the Christian Democrat International. Personal positioning on the left-right scale is significant for all the center-to-right-wind parties: PP, UPyD, CIU, PNV and CC; and also significant but with different sign for ERC. Interestingly, those who abstain tend to self-position themselves to the right of the reference alternative, PSOE. A positive economic evaluation tends to increase incumbent's vote (PSOE), while it decreases the probability of voting for PP, IU and CIU. A positive political evaluation tends to increase vote for PSOE and CC and decreases vote for PP. Among the party specific covariates, we find that the vote in the previous election appears with a positive and significant coefficient, indicating that there is persistence in voting decisions.⁴ A dummy variable (*Closest party*) indicating for each individual the closest party with respect to the most important topic of the electoral campaign has a positive and significant coefficient estimate. A count variable (*Campaign*) recording the number of ways parties reached each elector during campaign turns out to be significant. Finally, the results for the strategic voting indicators are as follows. Coefficient estimates on *W1*, *W2* and their interaction are negative as expected, although they are not statistically significant.

Indicators *W1* and *W2* were specially designed by Alvarez and Nagler (2000) to fit the purpose of measuring the incentives voters might have to vote strategically in the context of the single-member district case of the British elections. It might be argued that in

⁴In fact, 78.4 per cent of the individuals vote for the party they voted in the previous general elections.

our application to the Spanish parliamentary elections, with a multiple-member district system, those indicators might not be the most suitable ones to measure strategic voting incentives. However, using alternative indicators of strategic behavior, more suitable for the case of Spanish elections, turn out to yield similar results.⁵

Table 6 assesses goodness of fit. Entries in this table show the number of individuals who voted for the party indicated by the column and the estimated model predicts a vote for the party indicated by the corresponding row. The largest numbers are along the main diagonal of the table, indicating that most votes are correctly predicted. The estimated model predicts 87.98 per cent of the cases correctly.

The inference approach uses the estimated model to generate *predicted strategic votes* and *predicted sincere votes* as defined by equations (3) and (4). Table 7 shows the source and direction of strategic votes. Rows indicate the source of the strategic votes and columns indicate the beneficiaries of those strategic votes. PSOE receives 23 strategic votes, most of them from IU, and PP receives 7 strategic votes. Overall, 1.4 per cent of the individuals in the working sample voted strategically. Our estimate of strategic voting in Spain using the inference approach is not very different from previous estimates. For instance, Garcia Viñuela and Artés (2011) estimate that strategic voting in Spain ranges from 1.15 to 1.24.

5 The counterfactual analysis

In this section we propose a methodology that combines features of the self-reported and inference approaches. The counterfactual method has four steps. First, we identify sincere voters using survey items. Second, we fit the random utility model (1) to those individuals identified as sincere voters in the first step. Third, we identify potential strategic voters also using survey items. Fourth, we use the estimated model to predict the sincere vote

⁵These results are available upon request.

of the potential strategic voters identified in the third step. We identify as strategic voters those potential strategic voters whose actual and predicted sincere vote are different. Next we describe these steps in detail.

Step 1. We identify sincere voters using survey items where individuals were asked to manifest their preferred party. In particular, items 44 and 44a of the survey state: Could you tell me if you feel close to any party? Which party do you feel close to? We classify individuals as sincere voters if they voted for their preferred party. Note that we identify those who we suspect voted sincerely, as opposed to the self-reporting approach which identifies strategic voters. Our objective is to eliminate all strategic voters from the sample. In doing this, some sincere voters might be dropped from the sample, but this is of no concern at this stage, the important point is to retain only those who manifest a sincere vote. Later on, in the fourth step, we will be able to identify those sincere voters who were mistakenly dropped from the sample in the first step. To the extent that a post-electoral bias (inflation bias) seems to exist in self-reported declarations in favor of finding increasing levels of strategic voting, we are on the safe side.

Step 2. We next estimate the coefficients of the random utility model (1) using individuals classified as sincere in the first step. This model indicates how sincere voters' utility depends on individual and party specific characteristics. Note that this procedure is not subject to the criticism commented above whereby estimates of the random utility party preference model (1) could yield biased estimates if the sample used includes strategic voters. If step one correctly identifies sincere voters, there are no strategic voters in the sample of individuals used in the estimation.

Table 8 shows the estimation results. Years of education has a positive and significant coefficient estimate for PP, UPyD, CIU and ERC. Religion turns out to be significant for PNV. The personal positioning on the left-right segment also influences significantly voting behavior, with PP, UPyD, CIU, PNV, CC voters and abstainers to the right of PSOE, and IU voters to the left of PSOE. A positive political evaluation benefits the

incumbent, PSOE, with respect to PP. We also find vote persistence, as it is more likely to vote for a party if you voted for it in the previous elections. The variable “campaign” is marginally significant, which suggests that stronger campaigns pay out. Finally, it is more likely to vote for the closest party with respect to the most important topic of the electoral campaign.

Table 9 evaluates goodness of fit. Each cell reports the number of individuals who voted for a party indicated by the corresponding column and the model predicts a vote for a party indicated by the corresponding row. The numbers along the main diagonal indicate those cases where the model predicts correctly. Numbers off diagonal correspond to cases where the model predicts incorrectly. Overall, the estimated model predicts 92.07 per cent of the votes correctly.

Step 3. In this step we identify potential strategic voters making use of survey items. A possibility would be to use the same items used in step 1 to identify sincere voters. However, those individuals not classified as sincere voters in step 1 are not necessarily strategic voters. They vote for a party which is not their preferred party, but need not be strategic voters. They could be voting for the winner or casting a punishment vote. Strategic voting requires not only voting for a party different from your preferred party, but also the intention to influence the outcome of the election. Therefore, it is necessary to identify among those who did not vote for their preferred party, those who had strategic motivations. For this purpose we used survey items where respondents were asked to give the reason of their vote, identifying as strategic voters those who voted to avoid other party winning. In particular, we identified as potential strategic voters those who were not selected among sincere voters in step 1 and chose answer (6) in items 36a or 36b (see Section 3 for a description of the choices). By doing so we select a set of individuals who did not vote for their preferred party (a necessary condition for strategic voting) and manifested intention to affect the result of the election. As a result of this procedure, 127 respondents were identified as potential strategic voters.

Step 4. We next conduct the counterfactual analysis. Let $\tilde{\beta}$ and $\tilde{\theta}$ be the estimated parameter values in step 2. The counterfactual sincere vote of potential strategic voter i is j_i^{*S} defined as⁶

$$j_i^{*S} = \arg \max_{l \in A} \Pr [j_i = l | X_i, Z_{il}, \tilde{\beta}, \tilde{\theta}].$$

Strategic voters are identified as those whose real vote and predicted sincere vote are different. If the real and predicted sincere vote are equal, then the individual is considered a sincere voter. Note that these sincere voters should be those that we excluded from the analysis in the first step as we suspected they could be strategic voters, but our method predicts they are not. Table 7 reports the number of strategic votes and their sources; PSOE receives 32 strategic votes and PP 15. The counterfactual approach suggests that a total of 47 individuals, 2.19 per cent of the working sample, voted strategically in the elections.

6 Discussion

Table 10 shows a comparison of the total strategic vote estimates according to the three methodologies. Strategic behavior according to the self-reporting methodology reaches 8.76 per cent. The inference approach results in a 1.40 per cent and the counterfactual methodology estimates a 2.19 per cent of strategic votes.

The estimate of strategic behavior obtained using the self-reporting methodology is higher than the estimates obtained using the other two methods. As argued by Alvarez and Nagler, when voters confess the reason why they voted for a party, there is a tendency to observe increasing values of strategic vote estimates as the survey is conducted further away from the election day. In our particular application to the 2008 Spanish General Election, there is only one post-electoral survey, which prevents us from confirming Alvarez and Nagler's findings. However, we have some indirect evidence that this is in

⁶Note that, potential strategic voters were not used to estimate the random utility model.

fact the case. Among the 113 individuals who confessed they voted for PSOE to avoid PP winning the elections, the counterfactual model predicts that only 32 of them voted strategically, while the remaining 80 were sincere PSOE voters. Similarly, among the 75 electors who indicated that they voted PP to avoid PSOE winning the elections, only 15 voted strategically according to the counterfactual method, the other 60 were sincere PP voters.

The inference and counterfactual approaches provide a low estimate of strategic voting and somewhat similar sources of strategic voting. There is, however, one main difference between these estimates. The counterfactual approach suggests that a large portion of the strategic vote comes from abstention, whereas this source of strategic voting is not present in the inference approach.

The counterfactual approach provides an overall estimate of strategic voting of 2.19 per cent, much lower than the self-reporting estimate of 8.76 per cent, despite the fact that the counterfactual approach uses the same survey information used in the self-reporting approach.

The counterfactual approach yields a higher estimate of strategic voting than the inference approach. It can be argued that, if there are strategic voters whose behavior is not well captured by the strategic vote indicators, strategic vote estimates might be downward biased. A beauty of the counterfactual methodology is that we avoid the problem of not being able to properly model strategic behavior.

The counterfactual estimate is not subject to the criticisms applied to the other methods. First, it is not subject to the inflation bias obtained in self-reported strategic vote estimates because we use self-reported motivations to identify sincere rather than strategic voters. Second, the counterfactual method is not subject to measurement problem of the inference approach because it does not make use of strategic voting indicators. Third, it avoids the strategic voter bias, because strategic voters are not used in estimating the random utility model.

A potential caveat is the issue of whether, in politics with a significant amount of partisan dealignment, survey items like the ones used to identify sincere voters, may identify as sincere voters an increasingly atypical group of voters: truly partisan voters. This is of no concern per se, as those who we drop in the first step of the procedure, are used in the fourth step to see if they are really strategic or sincere. The problem is that the utility preference model is estimated only with those observations corresponding to truly partisan voters who might not be representative of the “average voter”. While we understand the importance of this caveat, there are reasons to believe that this caveat is unfounded. If the selected group of sincere voters was an atypical group of voters, then our estimates of strategic voters would be upward biased, as we would be classifying sincere but dealigned voters as strategic. A reason to believe that this dealigned voter issue is not that important is the fact that our estimate of strategic voters is only slightly above 2 per cent. Had partisan dealignment been a problem, our estimate of strategic vote would have been greater. On the contrary, the significant degree of partisan dealignment in today’s politics could very well explain why a significant amount of the strategic vote we find comes from abstention, that is, voters whose first choice is not one of the potentially winning parties.

As a by-product, our method can be used in any empirical application of the the random utility model where the presence of strategic voters is suspected. If estimation of the size of strategic voting is not attempted, our method boils down to two steps: first, use survey items to identify sincere voters and, second, estimate the random utility model with the sample of sincere voters. Typical applications of the random utility model go directly to the second step, without removing from the sample potentially strategic voters. If we skip the first step and estimate the random utility model with all individuals in the sample, the resulting estimates are those reported in Table 11. Comparing these results with those reported in Table 8 indicates that coefficient estimates are fairly different, suggesting the presence of a strategic voter bias. In addition, significance of coefficients in Table 11 is lower, which suggests that eliminating strategic voters from the sample helps to identify

vote determinants.

Conclusions

This paper contributes to the literature on strategic voting estimation in several aspects. First, we review the existing methodologies by highlighting their virtues and drawbacks. The self-reported methodology uses survey items on vote motivation in order to measure strategic vote. Despite its simplicity, this method is known to exhibit a post-electoral inflation bias in self-reported strategic voting. The inference approach uses a random utility model to indirectly measure the level of strategic behavior. The inference approach uses indicators of the incentives that individuals might have to vote strategically. This procedure is potentially able to avoid the strategic voter bias, but might incur in an error of measurement problem.

Second, we propose an alternative methodology, the counterfactual approach, that makes use of the strengths of previous methodologies and avoids their problems. Instead of identifying those who self-report strategic vote, as in the self-reporting approach, we identify those who manifest that their vote is sincere. Therefore, we do not incur the risk of misidentifying individuals who manifest their vote is strategic but it is not. So the procedure avoids the inflation bias. On the other hand, we estimate the random utility model using data on those individuals identified as sincere voters. The procedure does not face the error of measurement problem, as we do not have to include the indicators used in the inference approach. Moreover, we also avoid the strategic voter bias, due to the fact that we do not include strategic voters in the estimation of the random utility model. The counterfactual method estimates the predicted sincere vote of potentially strategic voters, which can be used to determine what electors voted strategically or sincerely. The method provides estimates of the size and the direction of the strategic vote.

Third, our methodology is able to capture types of strategic behavior that the other

methods are not able to capture. Voting behavior could be important in elections where individuals have a large set of alternatives and different ones depending on the constituency of the voter. Voters whose preferred alternative is only available in a subset of constituencies, might vote strategically. The inference approach proposed by Alvarez and Nagler makes use of the multinomial probit estimation method which has the advantage of satisfying the IIA axiom. However, this methodology is not feasible when the set of alternatives is large. Instead, we use the multinomial logit model which can handle this case. We show how the multinomial logit model, although it does not satisfy the IIA axiom, generates estimates very similar to the probit model. By allowing for a large set of alternatives, the multinomial logit captures strategic behavior among voters whose sincere choice would be regional parties.

Fourth, we contribute a method that can be used in any empirical application where the random utility model of voting behavior is estimated. Even if researchers are not interested in strategic voting, steps 1 and 2 of the counterfactual approach still provide directions for practitioners to avoid the strategic voter bias.

Strategic vote estimates vary depending on the methodology used. The counterfactual method proposed in this paper estimates that the strategic vote in the 2008 Spanish General Election was 2.19 per cent.

Appendix A: Data description

Personal characteristics and alternative-specific attributes are defined next and their descriptive statistics are reported in Table 12.

Personal characteristics

Income_1 This variable takes value 1 if individual income is less than 900 euros and 0 otherwise

Income_2 This variable takes value 1 if individual income is between 901 and 1800 euros, and 0 otherwise

Income_3 This variable takes value 1 if individual income is between 1801 and 3000 euros, and 0 otherwise

Income_4 This variable takes value 1 if individual income is 3001 euros or higher and 0 otherwise

Years_of_education The number of years that an individual needs to attain certain level of study. It takes the following values:

5 if the individual has less than 5 years of schooling

12 if the maximum level of study is “Educación Primaria” (Primary Education) or equivalent

16 if the maximum level of study is “ESO” (Compulsory Secondary Education) or equivalent

18 if the maximum level of study is “FP grado medio” (Middle Level Vocational Training) or equivalent

18 if the maximum level of study is “Bachillerato” (Upper Secondary) or equivalent

20 if the maximum level of study is “FP grado superior” (Higher Level Vocational Training) or equivalent

21 if the maximum level of study is “Arquitectura Técnica”/ “Ingeniería Técnica” (3-year Quantity Surveyor or Engineering Degree) or equivalent

21 if the maximum level of study is “Diplomatura” (3-year Degree) or equivalent

22 if the maximum level of study is “Licenciatura” (4 or 5-year Degree) or equivalent

23 if the maximum level of study is “Arquitectura”/ “Ingeniería” (5-year Quantity Surveyor or Engineering Degree) or equivalent

26 if the maximum level of study is “Estudios oficiales de Postgrado”/ “Doctorado” (Post-graduate official studies / Doctorate) or equivalent

Religion It takes value one if the individual is a believer of a religion and practices religion at least once a month, 0 otherwise

Left-Right Self-positioning score on a left to right scale with 1 at the extreme left and 10 at the extreme right.

Economic_evaluation This variable takes value 1 if the individual evaluates the Spanish economic situation as fair, good or very good, and 0 if bad or very bad.

Political_evaluation This variable takes value 1 if the individual evaluates the Spanish political situation as fair, good or very good, and 0 if bad or very bad.

Party_affiliation This variable takes value 1 if the individual is a member of a party, 0 otherwise.

Union_affiliation This variable takes value 1 if the individual is a member of a labor union, 0 otherwise.

Age The age of the individual.

Gender It takes value 1 if the individual is a woman and value 0 if the individual is a man.

Alternative-specific attributes

Closest_party_{*ij*} It takes value 1 if for individual *i*, *j* is the alternative with the most similar ideas with respect to the main topic of the electoral campaign, and 0 otherwise.

Campaign_{*ij*} It indicate how many different ways alternative *j* has used to inform individual *i* during the electoral campaign. It takes value 0 if *j* has not used any way to inform *i*, 1 if only has used one way, 2 if it has used two ways, and so on.

Previous_vote_{*ij*} It takes value 1 if *j* is the alternative that individual *i* voted for in the 2004 Spanish General Elections, and 0 otherwise.

References

- [1] Abramson P.R., Aldrich, J.H., Paolino, P., Rohde, D.W., 1992. "Sophisticated" Voting in the 1988 Presidential Primaries. *American Political Science Review*, 86(1): 55-69
- [2] Abramson, P.R., Aldrich, J.H., Blais, A., Diamond, M., Diskin, A., Indridason, I.H., Lee, D.J. and Levine, R., 2010. Comparing Strategic Voting Under FPTP and PR. *Comparative Political Studies* 43(1): 61-90
- [3] Achen, C., Shively, P., 1995. *Cross-Level Inference*. The University of Chicago Press. Chicago
- [4] Alvarez, R.M., Nagler, J., 1995. Economics, Issues and the Perot Candidacy: Voter Choice in the 1992 Presidential Election. *American Journal of Political Science*, 39: 714-744.
- [5] Alvarez, R.M., Boehmke, F.J., Nagler, J., 2006. Strategic Voting in British Elections. *Electoral Studies*, 25:1-19
- [6] Alvarez, R.M., Nagler, J., 2000. A New Approach for Modelling Strategic Voting in Multiparty Elections. *British Journal of Political Science*, 30:57-75
- [7] Barreiro, B., 2002. La progresiva desmovilización de la izquierda en España: un análisis de la abstención en las elecciones generales de 1986-2000. *Revista Española de Ciencia Política*, 6: 183-206
- [8] Blais, A., Nadeu, R., 1996. Measuring Strategic Voting: A Two-Step Procedure. *Electoral Studies*, 15(1): 39-52.
- [9] Blais, A., Nadeu, R., Gidengil, E., Nevitte, N., 2001. Measuring strategic voting in multiparty elections. *Electoral Studies*, 20: 343-352.

- [10] Blais A., Young R., Turcotte M., 2005. Direct or indirect? Assessing two approaches to the measurement of strategic voting. *Electoral Studies*, 24: 163-176.
- [11] Blinder, Alan S., 1973. Wage Discrimination: Reduced Form and Structural Estimates. *Journal of Human Resources* 8(4): 436-455
- [12] Boix, C., Riba, C., 2000. Las bases sociales y políticas de la abstención en las elecciones generales española: Recursos individuales, movilización estratégica e instituciones electorales. *Revista Española de Investigaciones Sociológicas*, 90: 97-129
- [13] Dow, J., Endersby, J.W., 2004. Multinomial Probit and Multinomial Logit: a Comparison of Choice Models for Voting Research. *Electoral Studies*, 23: 107-122
- [14] Duch R. M., Palmer, H.D., 2002. Strategic Voting in Post-Communist Democracy?. *British Journal Political Science*, 32: 63-91.
- [15] Evans, G., Heath A., 1993. A Tactical Error in the Analysis of Tactical Voting: A Response to Niemi, Whitten and Franklin. *British Journal of Political Science*, 23: 131-137.
- [16] Felsentahl, D.S., Brichta, A., 1985. Sincere and Strategic Voter: an Israeli Study. *Political Behavior*, 7(4): 311-324
- [17] Fieldhouse, E., Shryane, N., Pickles, A., 2007. Strategic Voting and Constituency Context: Modelling Party Preference and Vote in Multiparty Elections. *Political Geography* 26: 159-178
- [18] Fisher, S.D., 2004. Definition and Measurement of Tactical Voting: The Role of Rational Choice. *British Journal of Political Science*, 34:152-166.
- [19] Franklin, M.N., Niemi, R.G., Whitten, G., , 1994. The Two Faces of Tactical Voting. *British Journal of Political Science*, 24: 549-557.

- [20] Freedman, D.A., Klein, S.P., Sacks, J., Smyth, C.A., Everett, C.G., 1991. Ecological Regression and Voting Rights. *Evaluation Review*, 15: 673-711.
- [21] Garcia Viñuela, E., Artés, J., 2009. Una estimación del voto estratégico de Izquierda Unida al Partido Socialista en la elecciones generales del período 200-2008. *Revista Española de Investigaciones Sociológicas*, 128: 35-55.
- [22] García Viñuela, E., Artés, J., 2012. Strategic Voting and Effective Representation thresholds: Evidence from three Spanish General Elections. *European Journal of Political Research*, 51: 289-315
- [23] Gschwend, T., 2007. Ticket-splitting and strategic voting under mixed electoral rules: Evidence from Germany. *European Journal of Political Research*, 46: 1-23.
- [24] Heath, A., Curtise, J., R. Jowell, et al., 1991. *Understanding Political Change: The British Voter 1964-1987*. Pergamon Press, Oxford.
- [25] Heath, A., Evans, G., 1994. Tactical Voting: Concepts, Measurement and Findings. *British Journal of Political Science*, 24(4): 557-561
- [26] Herrmann, M., Pappi, F.U., 2008. Strategic Voting in German Constituencies. *Electoral Studies*, 27: 228-244.
- [27] Johnston, R.J., Pattie, C.J., 1991. Tactical Voting in Great Britain in 1983 and 1987: An Alternative Approach. *British Journal of Political Science*, 21: 95-128.
- [28] Kim, H.M., Kostadinova, T., 2011. Does Tactical Voting Matter? The Political Impact of Tactical Voting in Canadian Elections. *International Area Studies Review* 14(1): 49-72
- [29] King, G., 1997. *Solution to the Ecological Inference Problem*. Princeton University Press, Princeton.

- [30] Kriesi, H., 1998. Straightforward and Strategic Voting in the Elections for the Swiss Council of States in 1995. *Electoral Studies*, 17(1): 45-59.
- [31] Kselman, D., Niou, E., 2009. Strategic Voting in Plurality Elections. *Political Analysis* 18: 227-244.
- [32] Lago Peñas, I., 2005. *El voto estratégico en las elecciones generales en España, 1977-2000*. Madrid: Centro de Investigaciones Sociológicas
- [33] Lago Peñas, I., 2008. Rational Expectations or Heuristics?: Strategic Voting in Proportional Representation Systems. *Party Politics*, 14(1): 31-49.
- [34] Lanoue, D.J., Bowler, S., 1992. The Sources of Tactical Voting in British Parliamentary Elections, 1983-1987. *Political Behavior*, 14(2): 141-157.
- [35] McCuen, B., Morton, R.B., 2010. Tactical Coalition Voting and Information in the Laboratory. *Electoral Studies* 29: 316-328.
- [36] Meffert, M.F., Gschwend, T., 2011. Polls, Coalition Signals and Strategic Voting: An Experimental Investigation of Perceptions and effects. *European Journal of Political Research*, 50: 636-667.
- [37] Merolla, J.L., Stephenson, L.B., 2007. Strategic Voting in Canada: A Cross Time Analysis. *Electoral Studies*, 26: 235-246.
- [38] Myatt, D.P., 2007. On the Theory of Strategic Voting. *Review of Economic Studies* 74: 255-281.
- [39] Niemi, R.G., Whitten, G., Franklin, M.N., 1992. Constituency Characteristics, Individual Characteristic and Tactical Voting in the 1987 British General Election. *British Journal of Political Science*, 22: 229-240.

- [40] Niemi, R.G., Whitten, G., Franklin, M.N., 1992. People Who Live in Glass Houses: A Response to Evans and Heath's Critique of our Note on Tactical Voting. *British Journal of Political Science*, 23: 549-563.
- [41] Oaxaca R., 1973. Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, 14(3): 693-709
- [42] Robinson, W.S., 1950. Ecological Correlations and the Behavior of Individuals. *American Sociological Review*, 15: 351-357
- [43] Shikano, S., Herrmann, M., Thurmer, P.W., 2009. Strategic Voting under Proportional Representation: Threshold Insurance in German Elections. *West European Politics* 32 (3): 634-659

Table 1: The 2008 Spanish General Elections

	Number of		Left/Righth	
	Votes	Vote shares	Seats	Score
Partido Socialista Obrero Español (PSOE)	11,289,335	43.87	169	3.62
Partido Popular (PP)	10,278,010	39.94	154	6.56
Izquierda Unida (IU)	969,946	3.77	2	2.67
Unión Progreso y Democracia (UPyD)	306,079	1.19	1	4.93
Convergència i Unió (CIU)	779,425	3.03	10	5.20
Esquerra Republicana per Catalunya (ERC)	298,139	1.16	3	2.80
Partido Nacionalista Vasco (PNV)	306,128	1.19	6	4.82
Bloque Nacionalista Gallego (BNG)	212,543	0.83	2	3.33
Coalición Canaria (CC)	174,629	0.68	2	4.91
Total number of electors	35,073,179	100.00		
Votes for a party	25,900,442	73.85		
Abstention	9,172,737	26.15		

The Left/Right score is computed from post-electoral survey. It reports the average of respondents' self assessment of their position in a left (1) to right (10) scale. The other data are from actual elections.

Table 2: Vote shares in the survey and working subsample

	Individuals who declared their vote	Working subsample
Number of respondents	5,240	2,146
Partido Socialista Obrero Español (PSOE)	46.85	52.80
Partido Popular (PP)	27.32	27.83
Izquierda Unida (IU)	4.10	5.06
Unión Progreso y Democracia (UPyD)	1.26	1.48
Convergència i Unió (CIU)	1.56	1.30
Esquerra Republicana per Catalunya (ERC)	0.76	1.30
Partido Nacionalista Vasco (PNV)	1.18	1.12
Bloque Nacionalista Gallego (BNG)	0.80	0.94
Coalición Canaria (CC)	0.53	0.67
Abstention	15.61	3.68

Table 3: Self-reporting approach results

	PSOE	PP	Other	Total
Identified as strategic	113	75	-	188
Identified as sincere	1,065	546	347	1,958
Total	1,178	621	347	2,146

Number of respondents in each category.

Table 4: Multinomial Probit and Logit estimates

Personal characteristics	PROBIT			LOGIT		
	PP	IU	Abstention	PP	IU	Abstention
<i>Income_2</i>	0.369 (0.283)	0.0558 (0.343)	0.0988 (0.314)	0.646* (0.371)	-0.0201 (0.493)	0.0997 (0.336)
<i>Income_3</i>	0.245 (0.318)	0.471 (0.356)	-0.179 (0.353)	0.598 (0.419)	0.485 (0.509)	-0.145 (0.380)
<i>Income_4</i>	0.227 (0.432)	-0.416 (0.511)	-0.409 (0.523)	0.573 (0.579)	-0.627 (0.733)	-0.234 (0.574)
<i>Years_of_education</i>	-0.00618 (0.0225)	0.0166 (0.0261)	-0.0242 (0.0266)	-0.00839 (0.0286)	0.0153 (0.0369)	-0.0315 (0.0292)
<i>Religion</i>	0.202 (0.235)	-0.587 (0.425)	-0.106 (0.313)	0.265 (0.298)	-0.767 (0.623)	0.0882 (0.338)
<i>Left-Right</i>	0.418*** (0.0671)	0.0914 (0.0871)	0.101 (0.0791)	0.681*** (0.0858)	-0.157 (0.109)	0.228*** (0.0804)
<i>Economic_evaluation</i>	- 0.548*** (0.199)	-0.462** (0.226)	-0.204 (0.233)	- 0.746*** (0.264)	-0.543* (0.326)	-0.241 (0.256)
<i>Political_evaluation</i>	- 0.641*** (0.244)	-0.550* (0.299)	-0.324 (0.312)	- 0.834*** (0.321)	-0.262 (0.441)	-0.431 (0.334)
<i>Party_affiliation</i>	0.102 (0.512)	-0.0884 (0.470)	-0.901 (0.809)	0.0128 (0.742)	0.172 (0.630)	-0.951 (1.079)
<i>Union_affiliation</i>	0.0459 (0.270)	-0.106 (0.274)	0.351 (0.299)	0.0255 (0.360)	0.00830 (0.372)	0.498 (0.330)
<i>Age</i>	0.0118* (0.00691)	0.00702 (0.00746)	-0.0125 (0.00804)	0.0155* (0.00880)	0.00207 (0.0110)	-0.00782 (0.00874)
<i>Gender</i>	-0.116 (0.187)	-0.0644 (0.205)	0.0595 (0.209)	-0.121 (0.245)	-0.0135 (0.291)	0.127 (0.231)
<i>Constant</i>	- 2.022*** (0.733)	-0.569 (0.829)	0.0623 (0.785)	- 3.462*** (0.926)	0.414 (1.188)	-0.148 (0.864)
Alternative-specific						
<i>W1</i>		-0.00355 (0.00795)			-0.0186 (0.0117)	
<i>W2</i>		-0.261** (0.119)			-0.286** (0.137)	
<i>W1 × W2</i>		0.00423 (0.00476)			0.00431 (0.00659)	
<i>Closest_party</i>		1.229*** (0.0985)			1.395*** (0.0897)	
<i>Campaign</i>		0.162 (0.137)			0.207 (0.165)	
<i>Previous_vote</i>		1.723*** (0.113)			2.002*** (0.0932)	
Goodness of fit		0.8691			0.871	

PSOE is the reference alternative. Standard errors in parentheses. One, two and three stars indicate significance at the 10, 5 and 1 per cent significance levels. Goodness of fit defined as the fraction of correct predictions.

Table 5: Random utility model: the inference approach

Personal characteristics	PP	IU	UPyD	CIU	ERC	PNV	BNG	CC	Abstain	Alternative-specific	Alternative-specific
<i>Income_2</i>	0.262 (0.366)	-0.150 (0.428)	0.283 (0.791)	0.312 (0.902)	-1.452 (1.844)	-0.685 (1.033)	1.301 (1.214)	-0.258 (0.769)	0.222 (0.484)	<i>W1</i>	-0.0103 (0.00890)
<i>Income_3</i>	0.463 (0.413)	0.496 (0.439)	0.690 (0.809)	0.500 (1.120)	-1.391 (1.923)	-0.254 (1.238)	1.615 (1.363)	-1.011 (1.401)	0.445 (0.541)	<i>W2</i>	-0.0826 (0.0912)
<i>Income_4</i>	0.270 (0.632)	-0.662 (0.771)	0.165 (0.924)	-0.0331 (1.263)	-0.217 (2.472)	-2.494* (1.439)	0 (0)	-0.819 (1.382)	0.125 (0.909)	<i>W1</i> × <i>W2</i>	- (0.000105)
<i>Years of education</i>	-0.00458 (0.0264)	0.00448 (0.0369)	0.207*** (0.0685)	0.181* (0.0969)	0.0779 (0.126)	0.111 (0.0803)	-0.0499 (0.120)	0.0699 (0.0804)	0.0379 (0.0418)	<i>Closest party</i>	1.750*** (0.103)
<i>Religion</i>	0.151 (0.314)	-0.775 (0.518)	0.119 (0.463)	0.200 (0.795)	0 (0)	3.054*** (0.804)	-0.168 (0.857)	1.349 (0.948)	-0.0355 (0.476)	<i>Campaign</i>	0.509*** (0.185)
<i>Left-Right</i>	0.797*** (0.116)	-0.123 (0.124)	0.605*** (0.146)	0.611*** (0.233)	-0.496** (0.242)	0.525 (0.346)	-0.0569 (0.309)	0.500*** (0.179)	0.270** (0.129)	<i>Previous vote</i>	2.131*** (0.108)
<i>Economic evaluation</i>	- 0.774*** (0.287)	-0.671* (0.343)	-0.392 (0.412)	-1.753* (0.949)	-0.666 (0.652)	0.123 (1.266)	-0.00153 (0.678)	-1.346 (0.828)	-0.314 (0.364)		
<i>Political evaluation</i>	-0.761** (0.348)	-0.270 (0.479)	-0.610 (0.479)	0.524 (1.368)	1.160 (0.895)	-1.471 (1.405)	-0.993 (1.049)	3.567*** (0.954)	-0.645 (0.437)		
<i>Party affiliation</i>	0.685 (0.750)	0.304 (0.542)	0 (0)	1.810** (0.745)	-1.046 (1.016)	0 (0)	-0.247 (0.655)	3.075** (1.525)	0 (0)		
<i>Union affiliation</i>	-0.131 (0.400)	0.0153 (0.404)	0.434 (0.428)	0.591 (1.103)	0.845 (0.882)	0.692 (0.812)	0.149 (0.676)	-1.312 (0.842)	0.0731 (0.441)		
<i>Age</i>	0.0129 (0.00826)	0.00145 (0.0138)	0.0232* (0.0125)	0.0285 (0.0203)	-0.00792 (0.0201)	0.00792 (0.0300)	-0.0173 (0.0248)	0.0266 (0.0184)	0.0103 (0.0124)		
<i>Gender</i>	-0.0247 (0.257)	- (0.298)	-0.557 (0.403)	0.516 (0.770)	-0.0428 (0.550)	-1.394 (0.905)	0.657 (0.671)	-0.291 (0.561)	0.169 (0.322)		
<i>Constant</i>	- 3.786*** (0.997)	0.529 (1.310)	- 7.314*** (2.206)	- 7.857*** (2.636)	0.722 (2.426)	-2.853 (2.096)	2.179 (2.656)	- 6.416*** (1.892)	-2.948** (1.253)		

PSOE is the reference alternative. Standard errors in parentheses. One, two and three stars indicate significance at the 10, 5 and 1 per cent significance levels. A zero entry indicates that the corresponding coefficient was restricted to be equal to zero because no observations fell in this category. Results for alternative-specific characteristics are reported in an additional column. Goodness of fit: $1,888/2,146 = 0.8798$.

Table 6: Goodness of fit for the inference approach

		Actual vote										
		PSOE	PP	IU	UPyD	CIU	ERC	PNV	BNG	CC	Abstain	TOTAL
Predicted Vote	PSOE	1127	31	41	13	6	8	2	6	2	30	1266
	PP	22	576	3	12	1	0	0	1	2	16	633
	IU	14	1	64	0	0	0	0	0	0	4	83
	UPyD	0	0	0	5	0	0	0	0	0	0	5
	CIU	1	2	1	1	22	0	0	0	0	0	27
	ERC	1	0	1	0	0	21	0	0	0	1	24
	PNV	1	0	1	0	0	0	22	0	0	3	27
	BNG	3	0	0	0	0	0	0	14	0	1	18
	CC	0	1	0	0	0	0	0	0	10	0	11
	Abstain	9	10	2	2	0	0	1	0	1	27	52
	TOTAL	1178	621	113	33	29	29	25	21	15	82	2146

Table 7: Strategic voting

		Inference approach		Counterfactual approach	
		Predicted Strategic Vote		Actual vote	
		PSOE	PP	PSOE	PP
Predicted Sincere	PSOE	-	0	-	6
	PP	1	-	2	-
	IU	15	3	13	1
	UPyD	1	3	0	0
	CIU	0	0	1	0
	ERC	5	0	4	0
	PNV	0	0	0	0
	BNG	0	1	1	0
	CC	0	0	0	0
	Abstain	1	0	11	8
	Total	23	7	32	15

Table 8: Random utility model: the counterfactual approach

Personal characteristics	PP	IU	UPyD	CIU	ERC	PNV	BNG	CC	Abstain
<i>Income_2</i>	-0.745 (0.698)	0.172 (0.880)	-0.0185 (0.990)	-0.860 (0.940)	-2.216 (1.955)	-2.445* (1.304)	2.996*** (1.127)	-0.921 (0.984)	0.642 (0.663)
<i>Income_3</i>	0.246 (0.796)	0.945 (0.909)	0.768 (1.030)	-0.386 (0.972)	-1.498 (2.037)	-2.033 (1.297)	2.294* (1.316)	-2.658** (1.085)	1.315* (1.085)
<i>Income_4</i>	-1.467 (1.472)	-1.546 (1.503)	-0.234 (1.417)	-0.390 (1.208)	-4.129 (2.537)	0 (0)	0 (0)	-4.023*** (1.516)	-4.023*** (1.516)
<i>Years_of_education</i>	0.153** (0.0631)	0.0600 (0.0666)	0.438*** (0.142)	0.464*** (0.0882)	0.529*** (0.175)	0.115 (0.107)	0.163 (0.142)	0.265*** (0.0647)	0.265*** (0.0647)
<i>Religion</i>	0.682 (0.549)	-0.463 (0.658)	0.125 (1.018)	-0.380 (0.596)	0 (0)	4.037*** (1.101)	1.000 (1.312)	-0.915 (0.950)	-0.915 (0.950)
<i>Left-Right</i>	1.525*** (0.258)	-0.603*** (0.218)	1.099*** (0.266)	0.964*** (0.284)	-0.163 (0.281)	1.442*** (0.304)	1.065*** (0.265)	1.373*** (0.231)	1.373*** (0.231)
<i>Economic_evaluation</i>	-0.123 (0.559)	-0.529 (0.504)	-0.616 (0.771)	-1.327* (0.683)	-0.283 (0.673)	0.564 (1.394)	1.958** (0.854)	0.985 (0.834)	0.985 (0.834)
<i>Political_evaluation</i>	-1.332** (0.675)	-0.586 (0.729)	-0.182 (0.868)	0.826 (0.843)	0.770 (0.814)	0.905 (1.445)	-1.029 (1.117)	2.897* (1.485)	2.897* (1.485)
<i>Party_affiliation</i>	1.085* (0.582)	0.0827 (0.755)	0 (0)	2.023*** (0.732)	-1.479* (0.862)	0 (0)	0.486 (1.009)	5.949*** (1.120)	5.949*** (1.120)
<i>Union_affiliation</i>	0.301 (0.692)	-1.103 (0.683)	-0.0652 (0.719)	-0.616 (0.808)	0.647 (1.114)	1.480 (1.099)	0.253 (0.775)	0.0950 (0.967)	0.0950 (0.967)
<i>Age</i>	0.0333* (0.0180)	-0.0246 (0.0183)	0.0450** (0.0225)	0.0767*** (0.0208)	0.0557** (0.0282)	0.0344 (0.0295)	-0.0775* (0.0423)	0.0986*** (0.0205)	0.0986*** (0.0205)
<i>Gender</i>	0.437 (0.489)	-0.0355 (0.444)	-0.158 (0.717)	2.639*** (0.560)	-0.204 (0.605)	-5.592*** (1.145)	0.367 (0.829)	-1.022 (0.794)	-1.022 (0.794)
Alternative-specific									
<i>Closest_party</i>					2.538*** (0.184)				
<i>Campaign</i>					0.548* (0.284)				
<i>Previous_vote</i>					3.128*** (0.215)				
<i>Constant</i>	-10.60*** (2.309)	2.412 (1.722)	-14.35*** (2.756)	-16.57*** (2.484)	-9.615*** (2.570)	-7.375** (3.332)	-6.176 (3.817)	-16.23*** (2.797)	-5.752*** (1.827)

34

PSOE is the reference alternative. Standard errors in parentheses. One, two and three stars indicate significance at the 10, 5 and 1 per cent significance levels. A zero entry indicates that the corresponding coefficient was restricted to be equal to zero because no observations fell in this category. Goodness of fit: 1,336/1,438=0.9207.

Table 9: The Goodness of fit for the counterfactual approach

	Actual vote											
	PSOE	PP	IU	UPyD	CIU	ERC	PNV	BNG	CC	Abstain	TOTAL	
Predicted Vote	PSOE	769	1	16	5	0	0	1	1	0	18	811
	PP	5	368	0	2	0	1	0	0	0	7	383
	IU	5	0	61	0	0	1	0	1	0	4	72
	UPyD	0	0	0	8	0	1	0	0	0	0	9
	CIU	0	0	0	0	20	0	0	0	0	0	20
	ERC	1	0	1	0	0	21	0	0	0	0	23
	PNV	0	0	0	0	0	0	17	0	0	2	19
	BNG	1	0	0	0	0	0	0	13	0	0	14
	CC	0	0	0	0	0	0	0	0	8	0	8
	Abstain	15	8	2	2	0	0	0	0	1	51	79
	TOTAL	796	377	80	17	20	24	18	15	9	82	1438

Table 10: Comparison between methods

	Strategic vote		
	Self-reported	Inference	Counterfactual
PSOE	113	23	32
PP	75	7	15
Total	188	30	47
Percentage	8.76%	1.40%	2.19%

Table 11: Random Utility Preference Model with all voters

Personal characteristics	PP	IU	UPyD	CIU	ERC	PNV	BNG	CC	Abstain
<i>Income_2</i>	0.582 (0.361)	-0.0619 (0.453)	0.441 (0.812)	0.840 (0.823)	-1.328 (1.586)	-0.842 (0.839)	0.899 (1.133)	-0.0886 (0.724)	-0.00445 (0.338)
<i>Income_3</i>	0.619 (0.407)	0.529 (0.466)	0.727 (0.834)	0.873 (0.971)	-1.285 (1.684)	-0.579 (1.057)	1.328 (1.307)	-0.910 (1.334)	-0.191 (0.382)
<i>Income_4</i>	0.467 (0.587)	-0.520 (0.738)	0.196 (0.940)	0.482 (1.152)	-0.702 (2.071)	-2.484** (1.140)	0 (0)	-0.751 (1.314)	-0.274 (0.601)
<i>Years_of_education</i>	-0.00636 (0.0245)	0.0152 (0.0378)	0.208*** (0.0692)	0.167* (0.0885)	0.0889 (0.0984)	0.0842 (0.0680)	-0.104 (0.151)	0.0713 (0.0715)	-0.0384 (0.0290)
<i>Religion</i>	0.262 (0.282)	-0.755 (0.499)	0.153 (0.453)	0.225 (0.668)	0 (0)	2.215*** (0.685)	-2.011 (2.699)	1.403 (0.884)	0.0689 (0.342)
<i>Left-Right</i>	0.696*** (0.0955)	-0.156 (0.112)	0.554*** (0.129)	0.609*** (0.204)	-0.443** (0.218)	0.635** (0.261)	-0.0124 (0.309)	0.418** (0.164)	0.238** (0.0954)
<i>Economic_evaluation</i>	-0.841*** (0.255)	-0.597* (0.329)	-0.459 (0.402)	-1.692* (0.883)	-0.410 (0.599)	-0.109 (0.914)	-0.155 (0.717)	-1.273 (0.777)	-0.286 (0.280)
<i>Political_evaluation</i>	-0.687** (0.308)	-0.269 (0.452)	-0.474 (0.458)	0.430 (1.196)	0.829 (0.745)	-1.087 (1.086)	-0.788 (0.958)	3.240*** (0.910)	-0.322 (0.345)
<i>Party_affiliation</i>	0.547 (0.643)	0.317 (0.492)	0 (0)	1.887*** (0.691)	-0.505 (0.819)	0 (0)	-0.0664 (0.647)	2.859* (1.469)	0 (0)
<i>Union_affiliation</i>	-0.0884 (0.354)	0.00647 (0.380)	0.445 (0.419)	0.352 (1.074)	0.739 (0.820)	0.743 (0.703)	0.207 (0.655)	-1.317 (0.843)	0.452 (0.335)
<i>Age</i>	0.0122 (0.00753)	0.00196 (0.0126)	0.0197 (0.0121)	0.0230 (0.0182)	-0.00767 (0.0167)	0.00283 (0.0215)	-0.0277 (0.0263)	0.0221 (0.0180)	-0.0109 (0.00950)
<i>Gender</i>	-0.0671 (0.229)	0.00552 (0.280)	-0.598 (0.396)	0.432 (0.710)	0.0396 (0.540)	-1.046 (0.747)	0.337 (0.714)	-0.457 (0.526)	0.0946 (0.243)
Alternative-specific									
<i>Closest_party</i>					1.558*** (0.0879)				
<i>Campaign</i>					0.359** (0.155)				
<i>Previous_vote</i>					2.028*** (0.0909)				
<i>Constant</i>	-3.544*** (0.887)	-0.219 (1.201)	-7.820*** (2.158)	- (2.334)	7.706*** (2.011)	-2.651 (1.973)	3.287 (2.871)	-5.904*** (1.787)	0.113 (0.943)

PSOE is the reference alternative. Standard errors in parentheses. One, two and three stars indicate significance at the 10, 5 and 1 per cent significance levels. A zero entry indicates that the corresponding coefficient was restricted to be equal to zero because no observations fell in this category.

Table 12: Descriptive statistics

Personal characteristics											
	Observations	Mean	Standard deviation	Minimum	Maximum						
<i>Income_2</i>	2146	0.4471	0.4972	0	1						
<i>Income_3</i>	2146	0.3094	0.4622	0	1						
<i>Income_4</i>	2146	0.0875	0.2826	0	1						
<i>Years_of_education</i>	2146	16.4580	4.9630	0	26						
<i>Religion</i>	2146	0.1915	0.3935	0	1						
<i>Left-Right</i>	2146	4.4545	1.8837	1	10						
<i>Economic_evaluation</i>	2146	0.6089	0.4880	0	1						
<i>Political_evaluation</i>	2146	0.8016	0.3988	0	1						
<i>Party_affiliation</i>	2146	0.0566	0.2311	0	1						
<i>Union_affiliation</i>	2146	0.1378	0.3447	0	1						
<i>Age</i>	2146	45.4210	16.2031	18	92						
<i>Gender</i>	2146	0.4627	0.4986	0	1						
Mean of alternative-specific attributes											
	PSOE	PP	IU	UPyD	CIU	ERC	PNV	BNG	CC	Abst.	
<i>Closest_party</i>	0.519	0.309	0.046	0.010	0.082	0.086	0.165	0.053	0.032	0.074	
<i>Campaign</i>	0.583	0.548	0.364	0.036	0.833	0.710	0.701	0.366	0.340	0.000	
<i>Previous_vote</i>	0.522	0.245	0.052	0.000	0.093	0.089	0.237	0.100	0.064	0.075	