Advance Access publication January 24, 2014

Political Analysis (2014) 22:243–257
doi:10.1093/pan/mpt047

Strategic Votes and Sincere Counterfactuals

Alaitz Artabe
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
Departamento de Fundamentos del Análisis Económico II, Universidad del País Vasco, Lehendakari Aguirre 83, 48015 Bilbao, Spain
e-mail: javier.gardeazabal@ehu.es (corresponding author)

Edited by Jonathan Katz

The random utility model (RUM) of voting behavior can account for strategic voting by making use of proxy indicators that measure voter incentives to vote strategically. The contribution of this article is to propose a new method to estimate the RUM in the presence of strategic voters, without having to construct proxy measures of strategic voting incentives. Our method can be used to infer the counterfactual 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%.

1 Introduction

A voter's preferred party frequently has no chance of winning the election. When that is the case, it is often argued that some voters might not vote for their preferred party but for 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, Young, and Turcotte (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."

Strategic voting poses a fundamental question for the estimation of the random utility model (RUM) of voting behavior. Mathematically, a binary relation can represent preferences over parties. Under certain conditions on the binary relation, preferences can be represented by a utility function. If all voters vote for their preferred party, the RUM can be used to investigate how those preferences depend on voter and party characteristics. However, in the case of strategic voting, some voters vote for a party other than their preferred one. If we continue using the RUM of voting behavior, utility no longer represents preferences for all voters, as it measures the utility of actions (votes). In this scenario, utility depends not only on voter characteristics and party attributes but also on strategic incentives. Consequently, RUM estimates using post-electoral surveys that include strategic voters and do not take into account strategic incentives are likely to be biased. For later reference, we will refer to this bias as the *strategic voter bias*. To avoid this bias, Alvarez and Nagler (2000) propose that a set of proxy measures of strategic voting incentives should be included as explanatory variables.

Authors' note: We would like to thank Elena del Rey, Beatriz Gonzalez Lopez-Valcarcel, Thomas Gschwend, Annick Laruelle, Stefan Napel, Federico Valenciano, Donald Wittman, two anonymous referees, and seminar participants at the University of the Basque Country, University of California–Santa Cruz, and the *Simposio de la Asociación Española de Economía 2011* in Malaga. Replication materials are available in the Political Analysis Dataverse dx.doi.org/10.7910/DVN/22991.

© The Author 2014. Published by Oxford University Press on behalf of the Society for Political Methodology.
All rights reserved. For Permissions, please email: journals.permissions@oup.com

https://doi.org/10.1093/pan/mpt047 Published online by Cambridge University Press

The main contribution of this article is to provide a method to estimate the RUM where utility represents voter preferences over parties. In addition, this method can be used to estimate the size of strategic voting. Measuring strategic voting is difficult and controversial, and there are different approaches to doing so. The first approach is the aggregate methodology, which uses aggregate election data to measure individual strategic voting (Spafford 1972; Cain 1978; Galbraith and Rae 1989; 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; Freedman et al. 1991; Achen and Shively 1995; King 1997).

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 whether their vote was strategic (Felsenthal and Brichta 1985; Heath et al. 1991; Lanoue and Bowler 1992; Niemi, Whitten, and Franklin 1992; Franklin, Niemi, and Whitten 1994; Kriesi 1998; Merolla and Stephenson 2007). Fisher (2004) reports 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 *response 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 estimate the RUM of voting behavior. Utility is modeled as a function of personal and party characteristics as well as variables intended to capture strategic incentives (Abramson et al. 1992; Blais and Nadeu 1996; Alvarez and Nagler 2000; Blais et al. 2001; Blais, Young, and Turcotte 2005; Alvarez, Boehmke, and Nagler 2006; Herrmann and Pappi 2008; Lago 2008). 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 RUM 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 that the strategic indicators enter the utility of all voters, whether sincere or strategic, with the same parametric form. Since the response of sincere and strategic voters to those indicators should be different, including indicators of strategic voting incentives that affect the utility of all voters equally forces a biased estimation of the mean response.

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), and is recently expanding into new directions of research (Duch and Palmer 2002; Gschwend 2007; Shikano, Herrmann, and Thurner 2009; MaCuen and Morton 2010; Meffert and Gschwend 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 put forward 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, for example Oaxaca (1973) and Blinder (1973). The idea is to predict what 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 RUM of voting behavior using data from sincere voters and then use the estimated random utility for sincere voters and the personal characteristics of self-declared strategic voters to predict what would have

https://doi.org/10.1093/pan/mpt047 Published online by Cambridge University Press

been their vote 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%.

The procedure described below can be interpreted as an observational study that analyzes the effect of a treatment in a nonexperimental setting. This interpretation views strategic voters as treated subjects and sincere voters as untreated ones, so the treatment is "strategic considerations." The procedure proposed in this article estimates the probability of voting for a party as a function of personal and party characteristics using a sample of untreated subjects (sincere voters) and then computes the counterfactual vote of treated subjects (strategic voters) had the treatment not been imposed on them. Then, we can estimate the treatment effect on the treated by simply comparing the observed and counterfactual vote of treated individuals.

An additional feature of our method is that it can be applied in fields other than strategic voting. When some survey respondents vote strategically, estimation of the RUM of voting behavior requires including strategic incentive indicators; otherwise, the estimates include a *strategic voting bias*. Our analysis opens another route for estimation of the RUM in the presence of strategic voters: use survey items to identify self-declared strategic voters, drop them from the estimation sample, and estimate the RUM 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 article is organized as follows. First, we describe the aggregate 2008 Spanish General Elections results, the strategic incentives that arise in the Spanish electoral arena, and the post-electoral survey data used in the empirical analysis. Second, we depict the counterfactual approach and compute estimates of the magnitude of strategic voting. Third, we discuss the results obtained and compare them with those obtained using alternative methodologies. Finally, we summarize the results.

2 The Case

Table 1 shows the results of the 2008 Spanish General Elections. The Socialist Party (PSOE) and People's Party (PP) obtained more than 80% 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 throughout Spain, whereas regional parties have their vote concentrated in some provinces (constituencies). Although only PSOE and PP had real changes of winning the elections, minority parties also play an important role in Congress, as they are usually pivotal.

Although strategic voting has been typically analyzed in Single Member District Plurality (SMDP) systems, a Proportional Representation (PR) system, such as the Spanish one, also leads to strategic voting, for example Bargsted and Kedar (2009), Cox and Shugart (1996),

Table 1 The 2008 Spanish General Elections

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

https://doi.org/10.1093/pan/mpt047 Published online by Cambridge University Press

Garcia Viñuela and Artés (2012), Gschwend and Stoiber (2012), Lago (2005, 2008, 2012), and Slinko and White (2010). In a PR system, even if your first-choice candidate does not have a chance of winning, your candidate might still get a seat in the legislature. Therefore, strategic voting incentives in PR systems might appear to be weaker than in First-Past-the-Post (FPTP) systems. However, there are reasons to believe that strategic voting incentives in PR systems might be strong. To sustain this claim, Abramson et al. (2010) argue that "... only voting systems that meet the conditions of Arrow's theorem are strategy-proof. Thus, since dictatorship is the only preference aggregation rule that satisfies Arrow's conditions, it is logically possible that strategic voting will be found in PR systems much like strategic behavior is found in extensively examined FPTP systems." Moreover, Abramson et al. (2010) report empirical evidence suggesting that strategic voting is more common in PR systems than FPTP systems.

In the case of the Spanish General Elections, what is really at stake is winning the elections at the national level; therefore, a voter whose preferred party has no chances of winning at the national level might have incentives to cast a strategic vote. The arguably largest potential source of strategic voting may come from voters whose preferred party is IU. Despite the fact that voting for IU might help IU to obtain an extra representative in any given constituency, some truly IU voters might have incentives to vote for PSOE or PP, as these are the only parties who have real chances of winning the elections at the national level. In fact, stalwart IU voters are more likely to strategically vote for PSOE than PP, as PSOE is closer to IU in the left-right dimension. Garcia-Viñuela and Artés (2009, 2012) and Lago (2005) provide evidence that this source of strategic voting is sizable. Another potential source of strategic voters may accrue from voters whose preferred party is a regional party: BNG in Galicia, ERC and CIU in Catalunya, CC in Canary Islands, and PNV in the Basque Country. These voters might have incentives to vote for PSOE or PP, which are the only parties with a significant probability of winning the general election at the national level.

Kselman and Niou (2010) show that, in a SMDP system, "... a necessary condition for strategic voting is that the voter's second-most preferred candidate has a better chance of winning than does his/her most preferred candidate." In a PR system, however, strategic voting might occur at a particular constituency even if the strategic voter's preferred party is running first, but at the national level has no chances of winning. What are the strategic incentives in this case? By voting strategically for a party that is running second or lower in your constituency, but has a significant chance of winning the election at the national level, the likely winners at the national level might have the marginal representative. This is in sharp contrast to SMDP systems where deviating from voting for your preferred party would not make sense if it is running first in your constituency, as the likely winning party at the national level would have no marginal representative to win. This type of strategic voting may be present in the Spanish General Election, where certain regional parties (PNV and CIU) often run first in polls and actual votes in some constituencies, but some of their partisan voters might decide to vote for PSOE or PP instead, which are the only parties with real chances of winning the elections at the national level. Another, arguably small, source of strategic voting may be voters with a strong antinationalist sentiment, who might vote for PSOE or PP 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 article makes use of a post-electoral survey carried out by CIS (*Centro de Investigaciones Sociológicas*) in fifty provinces. The analysis is restricted to parties with at least one seat in Congress and fifteen or more respondents who declared they voted for that party. Using this criterion results in a sample of individuals who declared that 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 Catalunya, PNV in the Basque Country, BNG in Galicia, and CC in Canary Islands) as well as those who abstained.[1]

The survey data includes 6083 interviews and records individual characteristics and electoral choices. Only 5240 respondents declared their vote. Many respondents did not answer other items

[1]The data set as well as STATA programs to reproduce the results reported in this article are available in Artabe and Gardeazabal (2013).

https://doi.org/10.1093/pan/mpt047 Published online by Cambridge University Press

Table 2 Vote shares in the survey and working subsample

	Individuals who declared their vote	*Working subsample*
Number of respondents	5240	2146
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

that we use for inference, leaving a working sample of 2146 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 nonresponse 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 Counterfactual Analysis

It is our understanding that much of the debate in the empirical literature on strategic voting inference has to do with the fact that defenders of the self-reported methodology and the inference approach neglect the other method as a valid way to estimate strategic voting. 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 RUM to those individuals identified as sincere voters in the first step. Third, we identify self-declared strategic voters likewise using survey items. Fourth, we use the estimated model to predict the sincere vote of the self-declared strategic voters identified in the third step. We identify as strategic voters those self-declared strategic voters whose actual and predicted sincere votes are different. Next we describe this procedure in detail and illustrate it using Spanish post-electoral survey data.

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 closest party. As a result of this, 1438 respondents were identified as sincere voters. Note that we identify those who we suspect voted sincerely, as opposed to the self-reporting approach that identifies strategic voters. Our objective is to eliminate all strategic voters from the sample. In doing so, 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 *response 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 RUM using individuals classified as sincere in the first step. The RUM assumes that, given a set of alternatives $A = \{1, ..., J\}$ that 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

https://doi.org/10.1093/pan/mpt047 Published online by Cambridge University Press

random disturbance. According to this model, individual i votes for alternative j if and only if $U_{ij} > U_{il}$ for all $l \neq j$. This model indicates how sincere voters' utility depends on individual and party-specific characteristics. This model defines the "preferred alternative," so strategic voting would deviate from the prediction of this model. Note that despite the fact that we do not include strategic voting indicators in equation (1), this procedure is not subject to the criticism commented above, whereby estimates of the random utility party preference model could yield biased estimates if the sample used includes strategic voters. If Step 1 correctly identifies sincere voters, there are no strategic voters in the sample of individuals used in the estimation.

There are two estimators available: the multinomial logit and probit. As pointed out by Dow and Endersby (2004), although the multinomial logit model imposes the restriction of Independence of Irrelevant Alternatives (IIA), the multinomial logit model allows us to increase the number of alternatives available for each individual voter, whereas this is difficult or even unfeasible with the multinomial probit, which requires numerical integration in a space of dimension equal to the number of alternatives. Should the number of alternatives be smaller, multinomial probit estimation could perfectly be used in this step.

Although we are not particularly interested in the multinomial logit estimates, we report them to show that signs of parameter estimates and their statistical significance make sense. Table 3 shows the estimation results. Income dummies are significant in a number of cases. PNV receives less votes and BNG more votes from the second income class than the reference alternative (PSOE), whereas CC tends to receive less votes from the upper income classes than PSOE. The upper income class tends to abstain less relatively to voting PSOE. Years of education have a positive and significant coefficient estimate for PP, UPyD, CIU, and ERC. Not surprisingly, religion is a significant determinant of the PNV vote, as PNV was a founder of Christian Democrat International. 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 and ERC voters to the left of PSOE. A positive political evaluation benefits the incumbent, PSOE, with respect to PP. PSOE voters tend to be younger than PP, UPyD, CIU, ERC, and CC voters and abstainers. CIU does better among men than PSOE, and PNV does better among women than PSOE. We also find vote persistence, as an individual is more likely to vote for a party if they voted for it in the previous elections. A count variable (*Campaign*) recording the number of ways parties reached each elector during the campaign turns out to be marginally significant, which suggests that stronger campaigns pay out. Finally, a voter is more likely to vote for the party closest to the most important topic of the electoral campaign.

Table 4 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. Numbers along the main diagonal indicate those cases where the model predicts correctly. Numbers off the diagonal correspond to cases where the model predicts incorrectly. Overall, the estimated model predicts 92.07% of the votes correctly.

Step 3: We here identify self-declared 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 that is not their preferred party, but they 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, voters with strategic motivation need to be identified among those who did not vote for their preferred party. For this purpose, we used survey items where respondents were asked to give the reason of their vote, where self-declared strategic voters were identified as those who voted to avoid the other party winning. Respondents who manifested that they voted for PSOE or PP were asked the following question (items 36a or 36b): "Which is the main reason for voting PSOE/PP?" Respondents had to choose from 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;

https://doi.org/10.1093/pan/mpt047 Published online by Cambridge University Press

https://doi.org/10.1093/pan/mpt047 Published online by Cambridge University Press

Table 3 RUM: The counterfactual approach

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

PSOE is the reference alternative. Standard errors in parentheses. One, two, and three stars indicate significance at the 10%, 5%, and 1% levels. A zero entry indicates that the corresponding coefficient was restricted to be equal to zero because no observations fell in that category. Goodness of fit: 1336/1438 = 0.9207.

Table 4 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

4. because it is the party that best represents my ideas;
5. because of its policy performance during the last four years;
6. to avoid PP/PSOE winning the elections;
7. other reasons;
8. no answer.

In particular, we identified as self-declared strategic voters those who were not selected among sincere voters in Step 1 and chose answer 6 in item 36a or 36b. These items are the most suitable ones to identify strategic voting from among all those available in the survey. Using these items, we select a set of individuals who did not vote for their preferred party (a necessary condition for strategic voting) and manifested an intention to affect the result of the election. As a result of this procedure, 127 respondents were identified as self-declared strategic voters. Let I be the set of all those identified as self-declared strategic voters.

Step 4: We next conduct the counterfactual analysis. For each self-declared strategic voter i in I, we compute the probabilities of voting for party l as $\Pr\left[j_i = l \mid X_i, Z_{il}, \tilde{\beta}, \tilde{\theta}\right]$, where $\tilde{\beta}$ and $\tilde{\theta}$ are the estimated parameter values in Step 2 (only with sincere voters), and repeat this for all l in A. The counterfactual sincere vote of self-declared strategic voter $i \in I$ is

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

Let j_i be the actual vote of self-declared strategic voter i. Strategic voters are identified as those whose real vote and predicted sincere vote are different, that is, $j_i^* \neq j_i$. If the real and predicted sincere votes 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 5 reports the number of strategic votes and their sources; PSOE receives thirty-two strategic votes and PP fifteen. The counterfactual approach suggests that a total of forty-seven individuals, 2.19% of the working sample, voted strategically in the elections. In order to provide a confidence interval of our estimate of strategic voting, we generated one thousand subsamples of size fifty from the set of self-declared strategic voters and then computed an estimate of the share of strategic voters. As a result of this procedure, a 95% confidence interval for strategic voting runs from a low 1.42% to a high 2.96%.[2]

[2]Note that this confidence interval does not take into account the uncertainty accruing from estimation of the multinomial logit. Accounting for such a source of uncertainty would require estimating the multinomial logit using subsamples, some of which would include too few voters for minority alternatives, making estimation unfeasible.

https://doi.org/10.1093/pan/mpt047 Published online by Cambridge University Press

Table 5 Strategic voting

	Counterfactual approach		*Inference approach*	
	Actual votes		*Predicted strategic vote*	
	PSOE	*PP*	*PSOE*	*PP*
Predicted sincere				
PSOE	–	6	–	0
PP	2	–	1	–
IU	13	1	15	3
UPyD	0	0	1	3
CIU	1	0	0	0
ERC	4	0	5	0
PNV	0	0	0	0
BNG	1	0	0	1
CC	0	0	0	0
Abstain	11	8	1	0
Total	32	15	23	7

Table 6 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%

4 Discussion

Table 6 compares our strategic vote estimates with those obtained using the self-reporting methodology and the inference approach. The self-reported approach makes use of survey items to identify strategic voters. In our particular application of this method to the 2008 Spanish General Elections, we use item 36a or 36b and identified strategic voting as those who selected option 6, that is, ''to avoid the other party winning the elections.'' According to Table 6, 188 individuals (8.76% of the working sample) voted strategically; PSOE received 113 strategic votes and PP 75. Likewise for comparison purposes, we implemented the inference approach using the indicators of strategic voting incentives defined in Appendix B. According to the inference approach estimates, thirty individuals (1.4% of the working sample) voted strategically; PSOE received twenty-three strategic votes and PP seven. This estimate of strategic voting in Spain using the inference approach is consistent with the previous estimates. For instance, Garcia Viñuela and Artés (2012) estimate that strategic voting in Spain ranges from 1.15% to 1.24%.

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 Elections, 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 could be the case. Among the 113 individuals who confessed they voted for PSOE to avoid PP winning the elections, the counterfactual model predicts that only thirty-two of them voted strategically, whereas the remaining eighty were sincere PSOE voters.

https://doi.org/10.1093/pan/mpt047 Published online by Cambridge University Press

Similarly, among the seventy-five electors who indicated that they voted PP to avoid PSOE winning the elections, only fifteen voted strategically according to the counterfactual method; the other sixty were sincere PP voters.

According to Table 6, the inference and counterfactual approaches provide low estimates of strategic voting. The counterfactual approach estimate of strategic voting is slightly higher than the inference approach estimate. However, the 95% confidence interval obtained using the counterfactual approach (1.42, 2.96) does not include the strategic voting estimate obtained using the inference approach. Table 5 indicates that both methods suggest somewhat similar sources of strategic voting. There is, however, one main difference between these sources of strategic voting. 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.

A beauty of the counterfactual methodology is that we avoid the problem of finding good indicators of strategic incentives. Moreover, even if the indicators of strategic incentives used in the implementation of the inference approach measure strategic incentives correctly, truly partisans should not react to those indicators. Assuming that the effects of strategic indicators in the inference approach is constant across sincere and strategic voters should result in biased parameter estimates.

Our estimation procedure relies on information from survey items. Instead of using a single item to determine what respondents self-declare strategic, we combine the information provided by two items. One of these items is used to identify respondents who voted for a party other than their closest party, a necessary condition for strategic voting. A second item is used to determine whether the reason why some voters did not vote for their closest party was strategic. Of course, the internal validity of the procedure depends crucially on the consistency of the items used for identification of the set of self-declared strategic voters. As using survey items that indicate respondents voted to influence the outcome is known to engender response bias, our procedure predicts what their vote would have been according to their observed characteristics and the utility function estimated for sincere voters. This counterfactual estimate allows us to eliminate those self-declared strategic voters that our model predicts are not strategic voters and would otherwise bias the strategic voting estimate upward as in the self-reported approach.

A potential caveat is the issue of whether, in politics with a significant amount of partisan dealignment, survey items such as 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 whether 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. Our estimate of strategic voters is only slightly above 2%. Had partisan dealignment been a problem, our estimate of strategic vote would have been greater. On the contrary, the high 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.

Related to the previous caveat is the degree of covariate balance among the sincere and self-declared strategic voters. In order to make the counterfactual approach more credible, a covariate balancing procedure could be used to select a subset of sincere voters with covariate values similar in a statistical sense to those of the group of self-declared strategic voters. Such a subset of sincere voters would then be used to estimate the RUM.

As a by-product, our method can be used in any empirical application of the RUM 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 RUM with the sample of sincere voters.

https://doi.org/10.1093/pan/mpt047 Published online by Cambridge University Press

5 Conclusion

This article contributes to the literature on strategic voting by proposing an alternative methodology, the counterfactual approach, which 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 that their vote is strategic but it is not. So, the procedure avoids the response bias. On the other hand, we estimate the RUM 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 strategic incentives indicators used in the inference approach. Moreover, we also avoid the strategic voter bias, as we do not include strategic voters in the estimation of the RUM. The counterfactual method estimates the predicted sincere vote of self-declared 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. The counterfactual method proposed in this article estimates that the strategic vote in the 2008 Spanish General Elections was 2.19%.

Strategic voting could be important in elections where individuals have a large set of alternatives and different ones depending on the constituency of the voter. The multinomial probit estimation method is not feasible when the set of alternatives is large. Instead, we use the multinomial logit model, which can handle the case of a large set of alternatives and captures strategic behavior among voters whose sincere choice would be regional parties.

The counterfactual method can be used in any empirical application where the RUM 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 without having to construct strategic incentives indicators.

Funding

A. A. acknowledges financial support from the Basque government through *Programas de Ayudas para la Formación y Perfeccionamiento del Personal Investigador* and from the Spanish *Ministerio de Ciencia e Innovación* (ECO2009-11213). J. G. acknowledges financial support from the Spanish *Ministerio de Economía y Competitividad* (ECO2012-35820) and the Basque government (IT783-13).

Appendix A: Data Description

Post-electoral survey conducted by *Centro de Investigaciones Sociológicas* (survey code 2757). Personal characteristics and alternative-specific attributes are defined next, and their descriptive statistics are reported in Table A1.

Personal Characteristics

Income_1

equals 1 if income is less than 900 euros, and 0 otherwise.

Income_2

equals 1 if individual income is between 901 and 1800 euros, and 0 otherwise.

Income_3

equals 1 if individual income is between 1801 and 3000 euros, and 0 otherwise.

Income_4

equals 1 if individual income is 3001 euros or higher, and 0 otherwise.

https://doi.org/10.1093/pan/mpt047 Published online by Cambridge University Press

Table A1 Descriptive statistics

Personal characteristics					
	Observations	*Mean*	*Standard deviation*	*Minimum*	*Maximum*
Income_2	2146	0.4471	0.4972	0	1
Income_3	2146	0.3094	0.4622	0	1
Income_4	2146	0.0875	0.2826	0	1
Years_of_education	2146	16.4580	4.9630	0	26
Religion	2146	0.1915	0.3935	0	1
Left-Right	2146	4.4545	1.8837	1	10
Economic_evaluation	2146	0.6089	0.4880	0	1
Political_evaluation	2146	0.8016	0.3988	0	1
Party_affiliation	2146	0.0566	0.2311	0	1
Union_affiliation	2146	0.1378	0.3447	0	1
Age	2146	45.4210	16.2031	18	92
Gender	2146	0.4627	0.4986	0	1

Mean of alternative-specific attributes										
	PSOE	*PP*	*IU*	*UPyD*	*CIU*	*ERC*	*PNV*	*BNG*	*CC*	*Abst.*
Closest_party	0.519	0.309	0.046	0.010	0.082	0.086	0.165	0.053	0.032	0.074
Campaign	0.583	0.548	0.364	0.036	0.833	0.710	0.701	0.366	0.340	0.000
Previous_vote	0.522	0.245	0.052	0.000	0.093	0.089	0.237	0.100	0.064	0.075

Years_of_education

The number of years that an individual needs to attain a certain level of study. It takes the following values: 5 (less than 5 years of schooling), 12 (''Educación Primaria''/Primary Education or equivalent), 16 (''ESO''/Compulsory Secondary Education or equivalent), 18 (''FP grado medio''/Middle Level Vocational Training or equivalent), 18 (''Bachillerato''/Upper Secondary or equivalent), 20 (''FP grado superior''/Higher Level Vocational Training or equivalent), 21 (''Arquitectura Técnica''/''Ingeniería Técnica''/3-year Quantity Surveyor or Engineering Degree or equivalent), 21 (''Diplomatura''/3-year degree or equivalent), 22 (''Licenciatura''/4 or 5-year degree or equivalent), 23 (''Arquitectura''/''Ingeniería''/5-year Quantity Surveyor or Engineering degree or equivalent), 26 (''Estudios oficiales de Postgrado''/''Doctorado''/Post-graduate official studies/Doctorate or equivalent).

Religion

equals 1 if believer of a religion and practices religion at least once a month, and 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

equals 1 if evaluation of the Spanish economic situation is fair, good, or very good, and 0 if bad or very bad.

Political_evaluation

equals 1 if evaluation of the Spanish political situation is fair, good, or very good, and 0 if bad or very bad.

https://doi.org/10.1093/pan/mpt047 Published online by Cambridge University Press

Party_affiliation

equals 1 if member of a party, and 0 otherwise.

Union_affiliation

equals 1 if member of a labor union, and 0 otherwise.

Age

The age of the individual.

Gender

equals 1 for women and 0 for men.

Closest_party$_{ij}$

equals 1 if individual i feels closest to alternative j with respect to the main topic of the electoral campaign, and 0 otherwise.

Campaign$_{ij}$

The number of ways alternative j has used to inform individual i during the electoral campaign.

Previous_vote$_{ij}$

equals 1 if individual i voted alternative j in the 2004 Spanish General Elections, and 0 otherwise.

Appendix B: The Inference Approach

This appendix shows how we construct the indicators of strategic voting incentives used in the inference approach. We construct two indicators of strategic voting incentives, $W1_{ij}$ and $W2_{ij}$. For each individual i, define the set of alternatives H_i that includes the two parties expected to receive the most votes in the constituency of voter i, and set $\bar{H}_i$ that is the union of H_i and abstention. Variable $W1_{ij}$ is defined as

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

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 in the voter's constituency is. In SMDP systems, $W1_{ij}$ measures the percentage vote needed by parties to obtain the seat in the constituency. Although in the case of a PR 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. Therefore, this variable should enter the estimating equation with a negative sign. Variable $W2_{ij}$ is defined as

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

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,

https://doi.org/10.1093/pan/mpt047 Published online by Cambridge University Press

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.

References

Abramson, P. R., J. H. Aldrich, A. Blais, M. Diamond, A. Diskin, I. H. Indridason, D. J. Lee, and R. Levine. 2010. Comparing strategic voting under FPTP and PR. *Comparative Political Studies* 43(1):61–90.

Abramson, P. R., J. H. Aldrich, P. Paolino, and D. W. Rohde. 1992. "Sophisticated" voting in the 1988 presidential primaries. *American Political Science Review* 86(1):55–69.

Achen, C., and P. Shively. 1995. *Cross-level inference*. Chicago: University of Chicago Press.

Alvarez, R. M., F. J. Boehmke, and J. Nagler. 2006. Strategic voting in British elections. *Electoral Studies* 25:1–19.

Alvarez, R. M., and J. Nagler. 2000. A new approach for modelling strategic voting in multiparty elections. *British Journal of Political Science* 30:57–75.

Artabe, A., and J. Gardeazabal. 2013. *Replication data for: Strategic votes and sincere counterfactual*. http://dx.doi.org/10.7910/DVN/22991.

Bargsted, M. A., and O. Kedar. 2009. Coalition-targeted Duvergerian voting: How expectations affect voter choice under proportional representation. *American Journal of Political Science* 53:307–23.

Blais, A., and R. Nadeu. 1996. Measuring strategic voting: A two-step procedure. *Electoral Studies* 15(1):39–52.

Blais, A., R. Nadeu, E. Gidengil, and N. Nevitte. 2001. Measuring strategic voting in multiparty elections. *Electoral Studies* 20:343–52.

Blais, A., R. Young, and M. Turcotte. 2005. Direct or indirect? Assessing two approaches to the measurement of strategic voting. *Electoral Studies* 24:163–76.

Blinder, Alan S. 1973. Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources* 8(4):436–55.

Cain, B. E. 1978. Strategic voting in Britain. *American Journal of Political Science* 22(3):639–55.

Cox, G. W., and M. S. Shugart. 1996. Strategic voting under proportional representation. *Journal of Law, Economics & Organization* 12(2):299–324.

Dow, J., and J. W. Endersby. 2004. Multinomial probit and multinomial logit: A comparison of choice models for voting research. *Electoral Studies* 23:107–22.

Duch, R. M., and H. D. Palmer. 2002. Strategic voting in post-communist democracy? *British Journal Political Science* 32:63–91.

Felsentahl, D. S., and A. Brichta. 1985. Sincere and strategic voter: An Israeli study. *Political Behavior* 7(4):311–24.

Fieldhouse, E., N. Shryane, and A. Pickles. 2007. Strategic voting and constituency context: Modelling party preference and vote in multiparty elections. *Political Geography* 26:159–78.

Fisher, S. D. 2004. Definition and measurement of tactical voting: The role of rational choice. *British Journal of Political Science* 34:152–66.

Franklin, M. N., R. G. Niemi, and G. Whitten. 1994. The two faces of tactical voting. *British Journal of Political Science* 24:549–57.

Freedman, D. A., S. P. Klein, J. Sacks, C. A. Smyth, and C. G. Everett. 1991. Ecological regression and voting rights. *Evaluation Review* 15:673–711.

Galbraith, J. W., and N. C. Rae. 1989. A test of the importance of tactical voting: Great Britain, 1987. *British Journal of Political Science* 19(01):126–36.

Garcia Viñuela, E., and J. Artés. 2009. Una estimación del voto estratégico de Izquierda Unida al Partido Socialista en la elecciones generales del período 2000–2008. *Revista Española de Investigaciones Sociológicas* 128:35–55.

———. 2012. Strategic voting and effective representation thresholds: Evidence from three Spanish General Elections. *European Journal of Political Research* 51:289–315.

Gschwend, T. 2007. Ticket-splitting and strategic voting under mixed electoral rules: Evidence from Germany. *European Journal of Political Research* 46:1–23.

Gschwend, T., and M. Stoiber. 2012. Strategic voting in proportional systems: The case of Finland. Universität Mannheim Working Paper.

Heath, A., J. Curtice, R. Jowell, et al. 1991. *Understanding political change: The British voter 1964–1987*. Oxford: Pergamon Press.

Herrmann, M., and F. U. Pappi. 2008. Strategic voting in German constituencies. *Electoral Studies* 27:228–44.

Johnston, R. J., and C. J. Pattie. 1991. Tactical voting in Great Britain in 1983 and 1987: An alternative approach. *British Journal of Political Science* 21:95–128.

Kim, H. M., and T. Kostadinova. 2011. Does tactical voting matter? The political impact of tactical voting in Canadian elections. *International Area Studies Review* 14(1):49–72.

King, G. 1997. *Solution to the ecological inference problem*. Princeton, NJ: Princeton University Press.

Kriesi, H. 1998. Straightforward and strategic voting in the elections for the Swiss Council of states in 1995. *Electoral Studies* 17(1):45–59.

Kselman, D., and E. Niou. 2010. Strategic voting in plurality elections. *Political Analysis* 18:227–44.

https://doi.org/10.1093/pan/mpt047 Published online by Cambridge University Press

Lago, I. 2005. *El voto estratégico en las elecciones generales en España, 1977–2000*. Madrid: Centro de Investigaciones Sociológicas.
———. 2008. Rational expectations or heuristics? Strategic voting in proportional representation systems. *Party Politics* 14(1):31–49.
———. 2012. Strategic voting in proportional representation systems: Evidence from a natural experiment. *Party Politics* 18:653–65.
Lanoue, D.J., and S. Bowler. 1992. The sources of tactical voting in British parliamentary elections, 1983–1987. *Political Behavior* 14(2):141–57.
McCuen, B., and R. B. Morton. 2010. Tactical coalition voting and information in the laboratory. *Electoral Studies* 29:316–28.
Meffert, M. F., and T. Gschwend. 2011. Polls, coalition signals and strategic voting: An experimental investigation of perceptions and effects. *European Journal of Political Research* 50:636–67.
Merolla, J. L., and L. B. Stephenson. 2007. Strategic voting in Canada: A cross-time analysis. *Electoral Studies* 26:235–46.
Myatt, D. P. 2007. On the theory of strategic voting. *Review of Economic Studies* 74:255–81.
Niemi, R. G., G. Whitten, and M. N. Franklin. 1992. Constituency characteristics, individual characteristic and tactical voting in the 1987 British General Election. *British Journal of Political Science* 22:229–40.
———. 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* 23:549–63.
Oaxaca, R. 1973. Male-female wage differentials in urban labor markets. *International Economic Review* 14(3):693–709.
Robinson, W. S. 1950. Ecological correlations and the behavior of individuals. *American Sociological Review* 15:351–57.
Shikano, S., M. Herrmann, and P. W. Thurmer. 2009. Strategic voting under proportional representation: Threshold insurance in German elections. *West European Politics* 32(3):634–59.
Slinko, A., and S. White. 2010. Proportional representation and strategic voters. *Journal of Theoretical Politics* 22:301–32.
Spafford, D. 1972. Electoral systems and voters' behavior. *Comparative Politics* 3(1):115–25.

https://doi.org/10.1093/pan/mpt047 Published online by Cambridge University Press