Integrality and Separability in Multidimensional Voting Models: Ideology and Nationalism in Spanish Regional Elections

Author(s): Rivero Rodriguez, Gonzalo
Date 2011
Type Working Paper
Series Estudios = Working papers / Instituto Juan March de Estudios e Investigaciones, Centro de Estudios Avanzados en Ciencias Sociales 265 (2011)
City: Madrid
Publisher: Centro de Estudios Avanzados en Ciencias Sociales

Center for Advanced Study in the Social Sciences

WORKING PAPERS

INTEGRALITY AND SEPARABILITY
IN MULTIDIMENSIONAL VOTING MODELS:
IDEOLOGY AND NATIONALISM IN
SPANISH REGIONAL ELECTIONS

Gonzalo Rivero

Estudio/Working Paper 2011/265
December 2011

Instituto Juan March de Estudios e Investigaciones

INTEGRALITY AND SEPARABILITY IN MULTIDIMENSIONAL VOTING MODELS: IDEOLOGY AND NATIONALISM IN SPANISH REGIONAL ELECTIONS

Gonzalo Rivero

Estudio/Working Paper 2011/265
December 2011

Gonzalo Rivero is a PhD candidate at the Center for Advanced Study in the Social Sciences, Juan March Institute, and at New York University.

ISSN: 2174-4971

Abstract

This paper analyzes how voters evaluate distance to parties on the ideological and nationalistic dimensions in Spain, using a generalized model that allows me to distinguish the effect of spatial proximity, the weight voters give to each dimension, and the metric (Euclid or Manhattan) that they use to evaluate closeness. I apply this model to post-electoral surveys in Galicia, the Basque Country, and Catalonia —three regions where the political arena is known to be structured along more than one dimension. The results strongly support that voters use a Manhattan metric in all three regions.[1]

Keywords: electoral behavior, Spain, ideology, nationalism, metrics, missing data.

[1] I thank Luis de la Calle, Pablo Fernández-Vázquez, María José Hierro, Jordi Muñoz, and the participants in the seminar of the research group on "Democràcia, eleccions i ciutadania" of the Universitat Autònoma de Barcelona for their helpful comments and discussion. Part of this research was undertaken while I was visiting the Institut d'Anàlisi Econòmica of the Spanish National Research Council (IAE-CSIC). I am most grateful for their hospitality. The conventional disclaimer applies.

INTRODUCTION

Proximity is a fundamental concept to the theory of voting behavior. In the Hotelling-Downs model of party competition (Downs, 1957) voters are assumed to cast their support for the party they feel closest to them, which is another way of saying that voters select the party that best represents them. However, the operationalization of *proximity* is not apparent once we leave the comfortable one-dimensional space underlying the concept of *ideology* in the Downsian setup. Once we attempt to analyze how voters evaluate parties on several dimensions/policies simultaneously, we are forced to introduce assumptions about the way they internally calculate distances, and more specifically about the degree of interdependence between dimensions. Do voters bundle all dimensions together or do they separately calculate distance to parties on each single dimension? Is the effect of distance on one dimension penalized differently relative to the values of the other(s), or are dimensions evaluated independently? Choosing between these alternatives is not without risk for the estimation of voting models, because they may severely affect our predictions about the way voters order parties (*vide infra*).

The question about how voters assess distances lies at the heart of the debate on the relative importance of nationalism versus ideology in the patterns of voting behavior in Spanish nationalities, an area that has been among the most productive in the recent Spanish literature. Starting with the seminal paper by Fernández-Albertos (2002), the research has consistently found that voters are more attentive to ideology than to nationalism (de la Calle, 2005; Balcells, 2007), at least in those regions in which political competition is structured around these two dimensions (Aguilar and Sánchez-Cuenca, 2007). This result is key for our understanding of how voters deal with the trade-off between pushing forward identitarian policies and the pursuit of classical ideological programs, even if we accept that a purely spatial approach has limited scope for the analysis of party strategy (de la Calle, 2005), and it is certainly reductionistic with regard to the motivations of voters (Merrill and Grofman, 1999; Balcells, 2007).

However, previous Spanish research has tended to make very coarse assumptions on the definition of distance between parties and voters. Regardless of the many enhancements over the basic spatial model that have been proposed, the specification of how individuals calculate distance to parties has been settled *via* untested assumptions specifically on the aspects of the model that characterize the relation between the two dimensions. In particular, all papers conjecture that the two dimensions —ideology and nationalism— are *separable*,[2] and some authors (Fernández-Albertos, 2002; Balcells, 2007) also assume that voters penalize the perceived distance to parties with a quadratic term. As it turns out, both assumptions admit an empirical test over their validity.

In this research, I take one step back and look more closely at voters' preferences in order to ascertain the kind of metric —and its characteristics— that they have in mind when making decisions based on spatial proximity on multiple dimensions. Instead of restricting myself to particular specifications, I directly estimate the spatial component of the utility function, by employing a Bayesian alternative to the approach used in Beauchamp (2008). In other words, in this research I seek to estimate all the parameters of the most general specification in order to concede between competing hypothesis regarding the way voters evaluate parties in a spatial electoral arena. It should be emphasized that this research does not reject the significance of the theoretical extensions developed by previous research to the basic spatial model. Rather, I adopt a more modest approach that only targets a part of the empirical specification that has been overlooked.

[2] Aguilar and Sánchez-Cuenca (2007) attempt to check the plausibility of this assumption by checking the correlation between the reported distances on the two dimensions. A more theoretically driven approach is employed here.

The model I analyze below makes various contributions to our understanding of spatial voting in Spain. First, the empirical specification of the voting model explicitly separates the effect of spatial and non-spatial variables in the probability of choosing one party over another. Second, this specification also permits a natural interpretation of the relative weight voters give to nationalism and ideology. And third, the model estimates the parameter that defines the kind of metric that voters use for the evaluation of distance to parties from data. This last point therefore tests the extent to which the *integrality* assumption used in previous research has been a correct choice. Results indicate that separability —represented by a Manhattan metric— should be preferred to integrality —an Euclidean metric— in the three regions in which I have tested the model —Galicia, the Basque Country, and Catalonia. Moreover, adding a dimensional squared penalization seems to have pernicious effects in terms of model fit if we compare it to a more conventional specification.

Finally, this research also relaxes one assumption that has been maintained in the literature. Specifically, I take into account the fact that individuals who do not report their self-location, or any of the parties' location on the ideological or nationalistic scale, tend to have different sociodemographic characteristics than those who do (de la Calle et al., 2010). In other words, I correct the missing data problem, given that there is strong evidence that the MCAR assumption (Rubin, 1976) does not hold. Note that failing to account for missing data may result in biased estimates, and in fact, I find that it produces misleading estimates of one key parameter, namely the parameter that defines the kind of metric used by voters.

THE SPECIFICATION OF DISTANCE IN MULTIDIMENSIONAL SETTINGS

Political competition is known to be structured along two dimensions —ideology and national identification— in three Spanish regions that have alternative national identities to that of Spanish: Galicia, the Basque Country, and Catalonia. In these three cases, national subjective identity has been a key element in the configuration of a particular electoral arena in which parties have to structure the demands and aspirations of regional autonomy —or even independence— as well as to compete in the classical ideological dimension (Montero and Torcal, 1990). In a key contribution, Fernández-Albertos (2002) shows that, assuming that voters behave according to the conventional spatial model (Downs, 1957), both national identification and ideology are significant predictors of voting behavior at the individual level in the Basque Country. More interestingly, he finds that nationalistic proximity has a weaker impact in the voting decision than ideological distance: although citizens evaluate both dimensions when deciding for whom to cast their vote, ideological proximity seems to be the most important dimension to them. Similar results are reported by de la Calle (2005) and Aguilar and Sánchez-Cuenca (2007) for the Basque Country; and Aguilar and Sánchez-Cuenca (2007) and Balcells (2007) for Catalonia.[3]

These results come from a specific conceptualization of voters' behavior: each voter is able to tell how far parties stand from her own ideal position along each dimension, and these distances constitute the cornerstone of their decision at the polling booth. This approach has been routinely accepted for electoral behavior when parties only compete on the ideological dimension, given how instinctive the use of the left-right scale seems to be, even for unsophisticated voters (Laver, 1998). A similar observation could be made about a scale that seeks to

[3] To the best of my knowledge the model has not been applied to Galician data.

account for national identification: it seems natural to assume that voters are able to rank parties according to a perception of their position along the nationalistic stand, and also that voters prefer parties with preferences as close as possible to their own, *cæteris paribus*.[4]

With this setting in mind, all the studies mentioned above use a random utility framework (see Maddala, 1986, for instance) in which the outcome variable —the party each voter has chosen— is assumed to depend on voter's specific characteristics —for instance, sociodemographic traits— and, more importantly, on other information that is specific to each combination of voters and parties —the perceived distance between the voter and each political alternative. Therefore, voters evaluate their perceived distance to each party —calculated from self-placements on 10– or 11–point Likert scales, and the reported perception of each party's location along the same scale—, and the theoretical expectation is that, between two different parties, voters will tend to choose the one that they think is the *closest* to them, *cæteris paribus*.

These basic mechanics have been complemented by several extensions in recent years. The model proposed by Rabinowitz and Macdonald (1989) argues that voters do not treat distances symmetrically, and that they instead divide the ideological scale according to the side in which the party is located. Under this model, a voter located between two equally separated parties would tend to prefer the one that is more extremist than himself. Besides, voters are also propelled to trade-off spatial proximity with classical evaluations of the performance of the government (Sánchez-Cuenca, 2008), even if parties can take advantage of these "ideological glasses" in order to avoid electoral accountability (Maravall and Przeworski, 2001). Moreover, voters may behave strategically with respect to the bargaining balance in the post-electoral landscape, and vote for the party that maximizes the chances of a given policy/dimension being pushed forward (Kedar, 2009). In other words, the basic spatial model can be enriched to account for more complex calculations. However, all of these extensions continue to share a common core of the basic spatial premise: voters prefer closer rather than distant parties. I will abstract from these issues for the remainder of this study in an attempt to focus on the simplest, bare-bones scenario of pure proximity voting.

In a unidimensional setting, the operationalization of *closeness* to parties is uncontroversial: the absolute value metric (that is, the difference between the location of the voter and the location of the party, arises naturally as the proper way to calculate distances).[5] However, this is no longer the case when voters have to evaluate parties on several dimensions simultaneously (Eguia, 2009). Researchers are then forced to choose between several

[4] Aguilar and Sánchez-Cuenca (2007) argue that the notion of nationalism is not well suited for a pure spatial analysis. In their paper, they draw on Serrano's (1998) findings about the cohabitation of a civic-territorial nationalism with another of ethnic component in both Catalonia and the Basque Country, to argue that both an ordered, as well as a dichotomous measurement of nationalism should be included in our specifications. In their account, we can think of the civic-territorial nationalism as creating a position on a continuous scale that orders policies according to the weight they put in the expansion of administrative decentralization or on the promotion of a common identity. However, the other type of nationalism splits alternatives into an opposed pair over which no gradation is possible: parties and voters either are nationalist or they are not. In any case, it is likely that, when asked to locate parties on a given continuous nationalistic scale, survey respondents will identify the latter mode as corresponding to an extreme of the former. Therefore, even with all the possible caveats about the interpretation of the nationalistic scale, we can imagine that voters are capable of giving it an intelligible meaning.

[5] A different question is how to recover those positions from survey data. The existence of projection and persuasion effects (Page and Jones, 1979; Feldman and Conover, 1983) distorts the relation between the true distance and the distance that is reported by respondents. Some authors have suggested using the mean of the location attributed by all respondents or by those respondents who have voted for a given party (Rabinowitz and Macdonald, 1989; Quinn et al., 1999) as a proxy for the true location of that party. However, this approach is not without criticism (Westholm, 1997).

reasonable alternatives. Previous research in the literature of nationalism *versus* ideology in Spain has found a common ground in the assumption of dimensional *separability* (Garner, 1974) when it has to deal with multiple dimensions. In other words, scholars presume that voters are able to tell movements apart on ideological and nationalistic dimensions, and that they actually care for them separately. This is a rather natural starting point (Westholm, 1997), because it implies that voters can differentiate a policy that increases the nationalistic standing of one party from other policies that fall into the conventional left-right spectrum, and that the evaluation on one dimension does not affect the perception of the other.

However, separable preferences are not the only alternative. Voters might evaluate distances to parties, not dimension-wise, but by taking the space of evaluations as a joint bundle, and assessing separation in a straight line. In this setting, the set of combinations that make the voter indifferent would now be homologous to the set of locations that individuals perceive as equidistant from their position in a local physical space (Laver and Humphreys, 2009).

The distinction between the two ideas is easier to understand if we take a look at the specification of the spatial model of voting behavior. Let's denote by x_i^n voter's i position in each of $n = 1, 2, \dots, N$ issues, and by x_{ik}^n the perceived location of party $k = 1, 2, \dots, K$ to voter i on issue n. Therefore, we are interested in a model of the following class

$$\Pr(y_i = k) = f(\theta \|x_i - x_{ik}\|_\rho + Z_i \boldsymbol{\omega}) \quad (1)$$

where $\boldsymbol{\omega}$ is the vector of coefficients that capture all non-spatial factors and individual-specific effects Z_i that affect voting behavior, like valence issues, the personal evaluation of the economy, individual income, ... Here, f is the function that transforms distances to parties into voting probabilities. In all the papers above, the model linking distances to voting probabilities is assumed to be a McFadden's conditional logit (Alvarez and Nagler, 1998). The most important information for the spatial model is contained in $\|\cdot\|_\rho$, that stands for the way in which voters are assumed to calculate distance with respect to parties.

There are several alternatives for the specification of $\|\cdot\|_\rho$ but the empirical literature in political science typically focuses on the so-called Minkowskian ρ-metrics, given by

$$\|x_i - x_{ik}\|_\rho = \left(\sum_n |x_i^n - x_{ik}^n|^\rho \right)^{1/\rho}$$

Special emphasis has been placed on the cases with $\rho = 1$, which gives rise to the *city-block* or *Manhattan* metric, and $\rho = 2$ or *Euclidean* metric. A conventional generalization, in order to test the differential effect of each dimension in the final voting decision, involves the addition of a set of dimension-specific weight parameters $\alpha_n, n = 1, 2, \dots, N$ such that

$$\|x_i - x_{ik}\|_\rho = \left(\sum_n \alpha_n |x_i^n - x_{ik}^n|^\rho \right)^{1/\rho}, \quad (2)$$

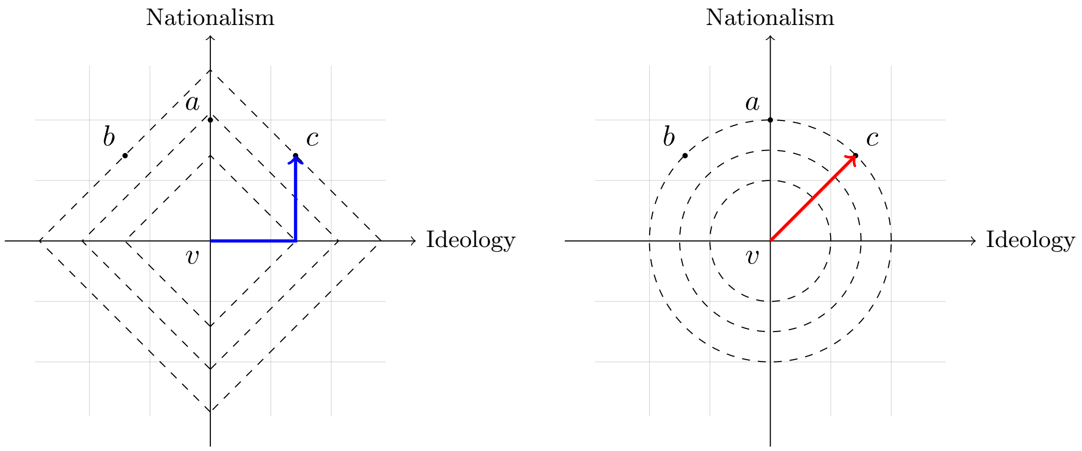
where it is convenient to assume that $\sum_n \alpha_n = 1$.

The interpretation of the parameters α and θ in equations 1 and 2 is straightforward. On the one hand, α reflects the relative importance the voter gives to each dimension. On the other, θ captures the impact of spatial proximity with respect to the rest of the non-spatial variables Z_i. That is, θ puts the spatial model in context.

Figure 1 depicts an illustration of the implications of assuming one or other metric. Let us assume, for the sake of argument, that voters deterministically vote for the closest party to them on two dimensions: nationalism (*nacl*) and ideology (*ideol*). The figure represents indifference curves for a voter with a saddle point at $(0, 0)$. In panel (b) of Figure 1 I have represented with dashed lines, three indifference curves derived from Euclidean preferences ($\rho = 2$) on two dimensions, that is, points that are equidistant from $(0, 0)$. Here it is assumed that voters'

FIGURE 1: Representation of indifference curves for Manhattan (left plot) and Euclidean (right plot) preferences.

(a) $\rho = 1$ (b) $\rho = 2$



utility is

$$U(x_i, x_{ik}^n) = -\Big(\alpha(x_i^{ideol} - x_{ik}^{ideol})^2 + (1-\alpha)(x_i^{nacl} - x_{ik}^{nacl})^2\Big)^{1/2} \quad (3)$$

where, in the particular example in the figure, both dimensions weigh exactly the same, so $\alpha = 0.5$.[6]

On the other hand, we could have assumed $\rho = 1$ (panel (a) in Figure 1), which would imply that voter's utility takes the following form

$$U(x_i, x_{ik}^n) = -\Big(\alpha|x_i^{ideol} - x_{ik}^{ideol}| + (1-\alpha)|x_i^{nacl} - x_{ik}^{nacl}|\Big) \quad (4)$$

where, again, $\alpha = 0.5$. Now, distance is calculated as the weighted sum of the coordinate-by-coordinate lengths of the vector that goes from the saddle point to the party; while in the previous case the relevant value was the length of the vector going from the position of the voter to any particular point.

Note an important substantive difference between the two metrics: in the Manhattan case, both dimensions are assumed to be independent, while in the Euclidean case they are not. To put it differently, under the assumption of Manhattan preferences, variations on the ideological distance are not affected by the existing distance on the nationalistic dimension in the decision of choosing party k. Thus, the impact of a unit increase in the distance on one dimension is the same for whichever distance we have on the other dimensions. Nonetheless, this is no longer the case in the Euclidean world: the effect of a small change on one dimension depends on the existing separation in the others.

We can use Figure 1 to illustrate the implication of using one metric or another. Suppose that three different parties a, b, and, c are evaluated by a voter v on two dimensions simultaneously, and that we know that both dimensions are equally important to him. Under the assumption of Euclidean preferences, all three parties a, b and c report v the same utility. Nevertheless, using Manhattan preferences *with the same* α, a is clearly preferred to b and c, that are still indifferent, as is shown in panel (a). The implication is clear, if we knew how much the voter cares about the two dimensions but have assumed a wrong metric, our inferences about his voting decision would be incorrect. Now assume that we need to estimate α from data, and that we know that the three parties are indifferent to the voter, as in Figure

[6] In Fernández-Albertos (2002), this value would be equivalent to $P = 1$.

1: an assumption of Euclidean preferences would return $\alpha = 0.5$, however, in order to make the three parties indifferent under Manhattan preferences, we would need to stretch the ideological dimension, thereby increasing the weight of nationalism to a value $\alpha < 0.5$. Therefore, the distinction between Manhattan and Euclidean metrics carries substantive implications that are most relevant.

However, despite the aforementioned differences, it has been common in the literature (Grynaviski and Corrigan, 2006) to focus on the estimation of either

$$\Pr(y_i = k) = f\Big(\eta_1(x_i^{ideol} - x_{ik}^{ideol})^2 + \eta_2(x_i^{nacl} - x_{ik}^{nacl})^2 + Z_i\Omega\Big) \quad (5)$$

or

$$\Pr(y_i = k) = f\Big(\eta_1|x_i^{ideol} - x_{ik}^{ideol}| + \eta_2|x_i^{nacl} - x_{ik}^{nacl}| + Z_i\Omega\Big) \quad (6)$$

There are three things that differentiate these alternatives from the model based on equations 1 and 2. First, there appears to be a systematic confusion in the Spanish empirical literature between Euclidean and Manhattan metrics, given that the articles mentioned above tend refer to equation 5 as a model with Euclidean preferences, while in fact it is a case of *quadratic* Euclidean preferences (Eguia, 2009). Second, despite its appeal, equation 5 presents a theoretical indeterminacy: it can be interpreted either as a Manhattan metric in which each dimension is penalized with a quadratic term, or as an Euclidean distance with quadratic penalization on the whole metric term. And third, the decomposition of the η_1 and η_2 parameters in the equations 5 and 6 into components α (dimensional weight) and θ (effect of distance) in equation 2 is an easier interpretation of the outcome of the estimation. The fact that α is normalized to 1 makes it more convenient for the interpretation of the relative weight of each dimension, while the parameter θ allows the researcher to target the marginal effect of the pure spatial component in the voting decision. Note that these two different elements are dumped into coefficients η_1 and η_2 in the conventional models above, which means that even though we could factorize them to recover the α and θ parameters from them, we could not say anything about their statistical significance.

Previous research has generally found that separable preferences outperform Euclidean preferences in multidimensional voting models, usually by drawing on U.S. data (Grynaviski and Corrigan, 2006; Berinsky and Lewis, 2007; Beauchamp, 2008). However, there are some significant differences with respect to the research presented here. Particularly interesting is the research by Grynaviski and Corrigan (2006), who investigate the performance of two different models separately, namely a model with quadratic Euclidean and another with city-block preferences. However, they do not attempt to directly estimate the non-linear coefficient defining the metric, that could award between the competing theories directly. In this sense, the approach I take here, consisting in a direct estimation of the metric component of the utility function, is essentially a replication of the research design used by Beauchamp (2008) for the U.S. case, but in a context —the political arenas of Galicia, the Basque Country and Catalonia— characterized by multiparty competition and structured around political aggregates such as ideology and nationalism, rather than policies or issues.

In sum, the aim of this paper is to overcome the limitations mentioned above by directly estimating a voting model in which distance is based on equation 2, that is, a model that estimates the exponent ρ from data. This strategy will allow me to check the extent to which the ubiquitous assumption of Manhattan metric is adequate. Besides, by decomposing the coefficients that reflect the weight that voters give to each dimension into α and θ, we will have a more intuitive interpretation of voting behavior.

EMPIRICAL DESIGN AND DATA

I have drawn on data from the CIS post-electoral studies after the autonomic elections of 2009 in Galicia (study 2796), 2009 in Basque Country (study 2795), and 2006 in Catalonia (study 2660). In each of these three surveys, respondents were asked to report the party they voted for in the last regional election, their self-placement, and the position they attributed to all parties running in the election on two Likert scales, one for ideology (1 for extreme left, 10 for extreme right) and another for nationalism (1 for less nationalist, 10 for more nationalist). I have omitted any consideration of projection or persuasion effects (Page and Jones, 1979; Feldman and Conover, 1983) and have used in all cases the individual subjective distance to parties (the reported self-location minus the reported location of parties) as the distance voters take into account when making a decision, instead of using the estimated mean from each of the individual locations (Rabinowitz and Macdonald, 1989; Quinn et al., 1999). Given that we are evaluating an equivalent model, neither of these cognitive effects should make a difference for the selection of the appropriate metric. Voters will apply dissonance in the same way.

I have modeled the voting decision using a conditional logit in order to replicate the literature. That is, I have assumed that the probability that individual i chooses party k is given by:

$$\Pr(y_i = k) = \frac{\exp(g(x_k, Z_i, \mathbf{\Gamma}))}{\exp(\sum_k g(x_{ik}, Z_i, \mathbf{\Gamma}))}$$

where

$$\begin{aligned} g(x_k, Z_i, \mathbf{\Gamma}) = \theta[\alpha(|x_i^{ideol} - x_{ik}^{ideol}|)^\rho \\ + (1-\alpha)(|x_i^{nacl} - x_{ik}^{nacl}|)^\rho]^{1/\rho} \\ + Z_i \boldsymbol{\omega} \end{aligned} \tag{7}$$

For simplicity, throughout this research I have omitted all other factors Z_i that affect voting behavior. This last assumption implies that I have modelled the voting decision of voter i solely on the basis of the evaluation of the perceived distance to the party k on the ideological and nationalistic dimensions. The decision to use a very stylized voting equation is likely to result in incorrect estimates of the taste parameters α and, obviously, θ. However, this design also fully exploits the distortions that arise in the event that the researcher makes a wrong assumption about ρ.

As indicated in the previous section, this specification exhibits several interesting features. First, θ captures the effect of distance with respect to the party (how much voters value proximity), while α gathers the weight voters give to each dimension (which dimension matters the most to voters). This information was dumped into η_1 and η_2 in the models previously used. Besides, if $\rho = 1$, then we would be in the 1-norm (Manhattan) case, while if $\rho = 2$, then, the 2-norm (Euclid) would be more appropriate. Note that it would be possible to further generalize the model in equation 7 by allowing the exponent $1/\rho$ to be multiplied by a term γ that would capture the penalization factor (Beauchamp, 2008). Under that model, $\rho = 1$ and $\gamma = 1$ would produce the city-block metric; $\rho = 2$ and $\gamma = 1$, the Euclidean metric; and $\rho = 2$ and $\gamma = 2$, the quadratic Euclidean model. I followed a different approach here. I estimated the model with penalization and compared its fit to the canonical model resulting from equation 7, to avoid an amalgamation between the cognitive metric and the additional *ad hoc* term not directly related to it.

Due to sample size, I restricted my attention only to the major parties in each region. Specifically, I selected those parties that received the (declared) vote of at least 5% of the regional sample. These are the parties that enter into the dataset for each region:

- Galicia: PSdG-PSOE, PP, and BNG.
- Basque Country: PNV-EAB, PSE-EE, PP and Aralar.
- Catalonia: CiU, ERC, ICV, PP, PSC-PSOE.

In all the analysis shown below, I only used those individuals who reported their vote in the last regional elections for one of the options above. However, this decision does not avoid the issue of incomplete data in the dataset. The frequency of each missing data pattern is represented in Table 1.

The missingness rate —incomplete cases over total observations— is mild, reaching 30% in Galicia, 25% in the Basque Country, and 20% in Catalonia. Besides, in all three regions, the probability for an individual to report her ideological/nationalistic distance to any party (see Table 2) increases with her education, and decreases with her age —except in Catalonia. Women are also less likely to report than men.[7] This result is consistent with de la Calle et al. (2010), and rules out the possibility that data is missing completely at random (MCAR) (Little and Rubin, 1987): individuals who hide information from the observer are different from those for whom we have complete information. Although there is some evidence of reporting bias in the case of the Basque Country (Urquizu-Sancho, 2006), the *ignorability* assumption for the missing data mechanism (Schafer, 1997) is compelling in this dataset, and I relied on it throughout the remainder of the analysis. Therefore, I used an ignorable procedure to impute missing data on the reported distance from voters to parties on the ideological and nationalistic dimension in order to correct the bias that might arise in a complete-case analysis. Specifically, I have used the IP method developed by Schafer (1997), that iteratively simulates the conditional distribution of missing data and the parameters governing the joint distribution of complete and incomplete data —in this case, a multivariate normal distribution. As additional information to feed the imputation model I included, along with the self-reported distance to each of the parties on each of the dimensions and the voting decision, two sociodemographic variables (gender and education level of the individual in a 7 point scale), the evaluation of the incumbent[8] and the main challenger at the regional level, and that of the Prime Minister José Luís Rodríguez Zapatero on 0–10 Likert scales. To summarize, the models below[9] have been estimated by employing all the individuals for whom we know their past vote, independently of whether they have reported all the spatial information to the interviewer or not. When this last piece of information is missing, I have replaced by using a simulation procedure that recovers the information from the joint data distribution.

A graphical representation of the distribution of self-locations of individuals by the party they voted for[10] is presented in Figure 2 for Galicia (left figure), Basque Country (center figure), and Catalonia (right figure). It is interesting to note that the nationalistic dimension consistently shows higher variation than the ideological one, which points to some indefinition in the concept the question is trying to capture. I did not re-escalate the distance from voters to parties in the two dimensions (*cf.* Aguilar and Sánchez-Cuenca, 2007) in order to preserve a comparable metric between them.

All estimations were performed using `JAGS 2.1`. Models were run with one MCMC chain and convergence was checked using Geweke diagnostics implemented in the `coda 0.13-5` package in `R 2.11.1`. For η_1, η_2, and θ parameters, I chose a prior normal distribution centered at 0 with precision 10^{-5}. Party-specific intercepts (β_k parameters) also have a prior normal distribution centered at 0 with precision 10^{-5}. For α, I used a uniform distribution between 0 and 1. ρ coefficients have prior

[7] `gender` takes value 1 if the respondent is female. `age` is measured in years from the age of 18. `education` takes the value of the highest academic degree achieved by the interviewee on a 7-point scale.

[8] I have taken as main incumbent, the leader of the opposition party with the most seats in the regional parliament.

[9] Results for the complete-case dataset are available upon request. Table 6 compares the effect of missing data on the taste parameters regarding nationalism and ideology.

[10] I have omitted ICV in the plot from the Catalonian sample for ease of presentation. This party is included in all models below.

TABLE 1: **Missing data structure of the varying-choice dataset**

Pattern	Galicia	Basque Country	Catalonia
00	0.704	0.749	0.810
10	0.071	0.115	0.055
01	0.098	0.061	0.065
11	0.125	0.073	0.069

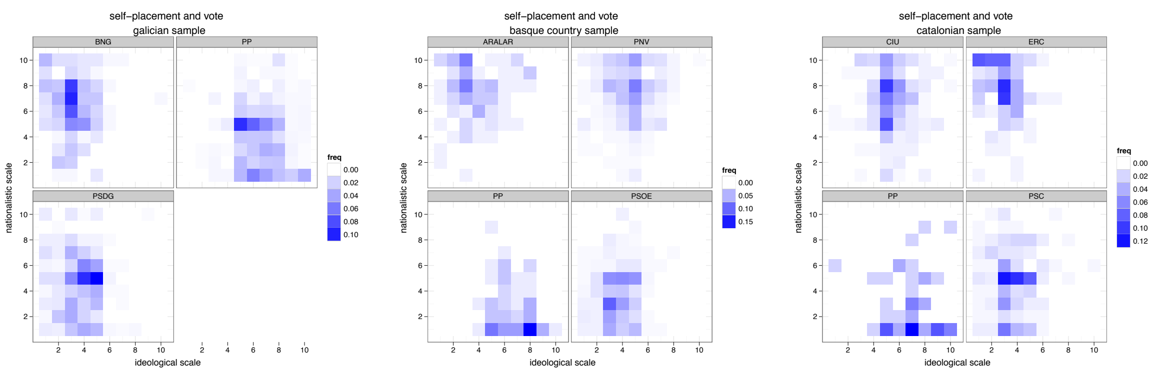
Notes: 0 indicates that the variable was observed. The first variable is the subjective ideological distance to each party. The second variable is the subjective nationalistic distance.

TABLE 2: **Models for the prediction of non-response**

	Galicia	Basque Country	Catalonia
	Mean 95% HPD	Mean 95% HPD	Mean 95% HPD
`Intercept`	-0.866	-1.226	-1.024
	[-1.315, -0.368]	[-1.917, -0.519]	[-1.933, -0.130]
`gender`	0.665	0.850	0.801
	[0.444, 0.898]	[0.529, 1.162]	[0.362, 1.266]
`age`	0.013	0.011	-0.001
	[0.005, 0.020]	[0.000, 0.022]	[-0.015, 0.014]
`education`	-0.305	-0.272	-0.371
	[-0.382, -0.228]	[-0.364, -0.181]	[-0.512, -0.233]
BIC	1158.535	660.001	464.121
N	1857	1131	951

Notes: The dependent variable takes value 1 if the respondent has not reported her distance to any of the parties on any of the two dimensions. The table shows the mean, the 2.5 quantile and the 97.5 quantile of the marginal posterior distribution of the coefficient of interest.

FIGURE 2: **Distribution of self-placement and vote for the three samples.**



uniform distributions[11] in the interval $[0.5, 2.5]$.

RESULTS

This section is structured in the following way. First, I replicate the two models that have been used in the literature: a random utility model with Manhattan preferences with dimension-by-dimension penalization (equation 5); and a model using pure Manhattan preferences (equation 6). Then, I estimate a pure Euclidean model like the one shown in equation 3. Whenever it was possible, I decomposed the coefficients η_1 and η_2 into α (dimension weight) and θ (distance weight). Finally, I estimate equation 7 directly.

Table 3 shows a replication of the model in equation 5 using a conditional logit: a model in which voters penalize distance with respect to parties using a squared factor on each (separable) dimension. The table shows the mean, the 2.5, and the 97.5 percentiles of the posterior distribution of the two relevant parameters. Therefore, a parameter can be considered to be different from zero if the zero does not fall within the 95% probability interval shown in the tables. As can be seen, results are standard. Ideology weighs more than nationalism in the voting decision in each region. With a 95% probability, the two coefficients are different from each other and they are also different from zero.

It is important to note that these results are roughly similar to those found in previous studies. For instance, Fernández-Albertos (2002) reports point estimates of η_1 between -0.156 and -0.201, and of η_2 between -0.064 and -0.096 for different post-electoral studies in the Basque Country between 1993 and 2001, while in my dataset and with the same model, the means of the posterior distributions are 0.158 and -0.074. Similarly, Balcells (2007) reports η_1 coefficients between -0.135 and -0.202, and between -0.064 and -0.066 for η_2 in Catalonia in 1999 and 2003, while in my dataset they are distributed with means in -0.191 and -0.067 for the 2006 elections. Therefore, my dataset reveals the same structural relations that have been found by previous studies if we impose the same assumptions that have been used. Therefore, differences between the estimations reported here and those that appear in the previous literature can be attributed to the estimation procedure and not to different structural conditions operating in my dataset.

In this sense, a comparison of the estimates with and without the MCAR assumption (see Figure 6) indicate that the violation of this condition has minimal effects for this specific model. The bias is remarkably small, and only seems to make some negligible difference in the case of the coefficient for ideological distance (η_1) in Catalonia.

The way coefficients η_1 and η_2 are specified in the previous model makes them difficult to interpret. In order to give a clearer perspective on the substantive meaning of these results I applied these estimates to a fake dataset that simulates the effect of increasing the distance that separates the voter from the party on each dimension. Specifically, Figure 3 shows the expected probability of voting for one party in each region, in this case the regional brand of the PSOE, for a voter that reports a fixed nationalistic distance of 3 points while we allow the ideological distance to vary throughout its range[12] (simulation A); and the same probability when we instead fix the ideological distance to 3 points and the nationalistic distance changes between 0 and 10 (simulation B). It is now easier to see the strong effect of distance on both dimensions for all regions: starting from a 40%/50% probability of voting for the PSOE when voter and party are in the same ideological position, it significantly drops to less than 10% when this distance is increased to 4 points, and to virtually

[11] The issue associated with the lack of invariance to transformations for prior uniform distributions (see Gill, 2007, for instance) is not relevant in this case given that ρ and α have a natural metric that do not need to be modified.

[12] For simplicity, I restricted the plot to positive values of distance. Note that the effect is symmetric by construction.

TABLE 3: Estimation of models with separation and penalization

	Galicia	Basque Country	Catalonia
	Mean	Mean	Mean
	95% HPD	95% HPD	95% HPD
η_1	-0.143	-0.158	-0.191
	[-0.157, -0.128]	[-0.177, 0.137]	[-0.216, -0.168]
η_2	-0.047	-0.074	-0.067
	[-0.056, -0.039]	[-0.085, -0.064]	[-0.079, -0.057]
BIC	2372.738	1524.998	1947.595
N	1944	1160	995

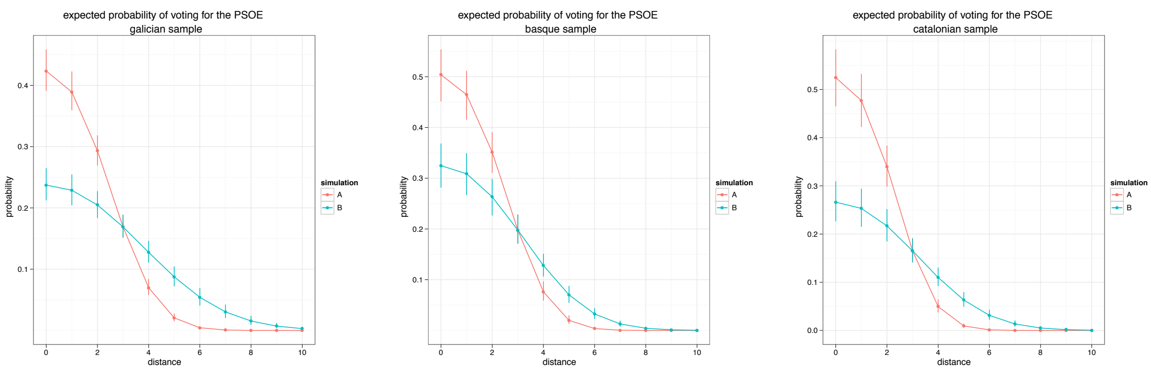
Notes: Estimation of model in equation 5. The table shows the mean, the 2.5 quantile and the 97.5 quantile of the marginal posterior distribution of the coefficient of interest. η_1 (resp., η_2) represents the posterior distribution of the ideological (resp., nationalistic) distance. I have omitted party-specific intercepts for ease of presentation.

TABLE 4: Model fit from model in Table 3

Galicia	Mean	2.5%	97.5%	Basque Country	Mean	2.5%	97.5%	Catalonia	Mean	2.5%	97.5%
BNG	0.73	0.63	0.91	Aralar	0.45	0.36	0.50	CiU	1.23	1.14	1.32
PP	1.10	1.07	1.13	PNV	1.15	1.12	1.18	ERC	0.89	0.76	1.06
PSdeG	1.01	0.92	1.08	PP	0.81	0.70	0.94	ICV	0.31	0.27	0.43
				PSE	1.08	1.03	1.13	PP	1.06	0.80	1.33
								PSC	1.19	1.09	1.37
Correct	0.770	0.756	0.781	Correct	0.780	0.774	0.784	Correct	0.647	0.638	0.664

Notes: The table shows the distribution (mean, 2.5 and 97.5 quantiles) of cases that are predicted to vote for a given party according to the model over the observed number of votes in the survey. The last row summarizes the percentage of correctly classified cases.

FIGURE 3: Expected probability of voting for the PSOE according to the models in Table 3.



0 when the ideological distance reaches its maximum. Something similar can be said about distance on the nationalistic dimension, although both its absolute effect in terms of the expected probability of voting for the PSOE, and its penalization, measured as the rate of change of this probability when we increase distance, are now smaller.

It is interesting, for the purpose of comparing models, to check model fit by the rate of correct classifications. This information is shown in Table 4. Specifically, the table shows both the overall fit, measured by the total rate of correct classifications, and how well it predicts voting for each party. In general, models in Table 3 seem to perform well if we take into account the fact that only two variables correctly classify between 65% and 78% of the voting decision. Nonetheless, there is a wide variation in party-specific predictions: the model tends to under- or over-represent a significant number of parties. In fact, with a 95% probability the model biases the predictions for the BNG (in Galicia), Aralar and PP (in the Basque Country), and ICV (in Catalonia) downwards, while it biases upwards PP (in Galicia), PNV and PSE (in the Basque Country) and CiU (in Catalonia). In other words, the model shows systematic biases in predictions for 8 out of 12 political parties, which means that we should treat it with the convenient caution, given that the good fit is in fact produced by systematic errors in classification.

We can compare these results with the estimation of a conventional Manhattan model (equation 6) using my dataset (Table 5). The interpretation of the model is easy and consistent with the intuitive interpretation of the previous table: voters seem to heavily discount ideological distance —and at a higher rate than nationalistic distance. Again, with a 95% probability, the coefficients for nationalism and ideology are different between them, and also different from zero.

However, the substantive impact of these estimations (Figure 4) is now slightly different. By construction, the predicted probability of voting for a given party now falls rapidly even at very short reported distances, while in the previous models the effect of distance was at first more moderated and then more accentuated. This is a purely mechanical effect of the penalization rate on each dimension. But despite this built-in feature, the substantive effect of distance is considerably different between the two models. With the results of the model in Table 5, the predicted probability of voting for the PSOE for an individual who considers himself to be at zero ideological points and three nationalistic points from the PSOE is around 65% in Galicia, and 70% in the Basque Country and Catalonia. The comparison with the effect of using a penalization for each dimension is striking: the model in Table 3 returned a 40% in Galicia, 50% in the Basque Country and 52% in Catalonia. At a distance of two ideological points, the differential effect in probability of voting for the PSOE between the two models is considerably reduced. In other words, the penalization factor strongly reduces the effect of short distances in the probability of choosing a given party.

The violation of the MAR assumption (Table 6) now makes a more relevant difference, although none of the main qualitative results change. Ignoring missing data in this case would lead us to biased estimations of the main coefficients: the imputation procedure used here reduces the impact of ideology in Galicia and the Basque Country, while it increases the effect of this same variable in Catalonia.

In terms of model fit, penalized and non-penalized models perform quite differently. Now, the pure Manhattan model only underestimates the vote for Aralar (in the Basque Country), and ICV (in Catalonia) while it overestimates the PNV and PSE (in the Basque Country) and PSC (in Catalonia). It means that we have systematic biases only in 5 out of 12 parties, even though the overall fit is not completely different from what we found above: around 80% —in mean— of the cases are correctly classified in the Basque

TABLE 5: Estimation of models with separation

	Galicia		Basque Country		Catalonia	
	Mean		Mean		Mean	
	95% HPD		95% HPD		95% HPD	
η_1	-0.790		-0.814		-0.817	
	[-0.857,	-0.728]	[-0.918,	-0.722]	[-0.896,	-0.740]
η_2	-0.305		-0.483		-0.425	
	[-0.356,	-0.258]	[-0.542,	-0.422]	[-0.485,	-0.366]
BIC	2194.313		1382.969		1841.905	
N	1944		1160		995	

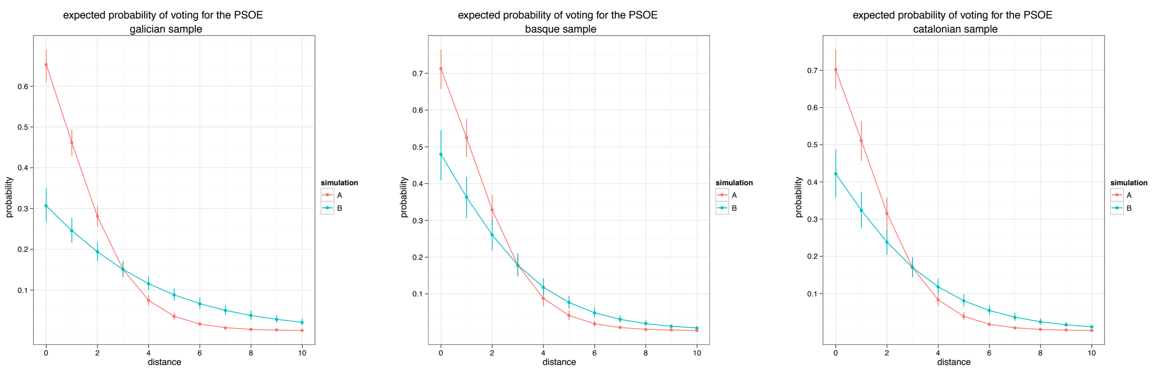
Notes: The table shows the mean, the 2.5 quantile and the 97.5 quantile of the marginal posterior distribution of the coefficient of interest. η_1 (resp., η_2) represents the posterior distribution of the ideological (resp., nationalistic) distance. I have omitted party-specific intercepts for ease of presentation.

TABLE 6: Model fit from models in Table 5

Galicia				Basque Country				Catalonia			
	Mean	2.5%	97.5%		Mean	2.5%	97.5%		Mean	2.5%	97.5%
BNG	0.88	0.80	1.03	Aralar	0.65	0.51	0.74	CiU	1.06	0.97	1.12
PP	0.98	0.96	1.02	PNV	1.06	1.02	1.11	ERC	0.98	0.86	1.08
PSdeG	1.09	0.98	1.16	PP	0.87	0.76	1.00	ICV	0.60	0.48	0.78
				PSE	1.11	1.05	1.16	PP	0.99	0.70	1.13
								PSC	1.18	1.05	1.32
Correct	0.795	0.787	0.800	Correct	0.789	0.777	0.795	Correct	0.679	0.665	0.691

Notes: The table shows the distribution (mean, 0.025 and 0.975 quantiles) of cases that are predicted to vote for a given party according to the model over the observed number of votes in the survey. The last row summarizes the percentage of correctly classified cases.

FIGURE 4: Expected probability of voting for the PSOE according to the models in Table 5.



Country and Galicia, and 68% in Catalonia. Hence, similar overall predictions are here due to a more realistic behavior of the model at party level.

More importantly, differences between models are also translated into differences in terms of BIC, which in the Bayesian framework becomes a natural alternative for the comparison of non-nested models (Kass and Raftery, 1995). This value penalizes the deviance of the model —the overall fit— by the number of parameters and number of cases that enter into the estimation, and gives a rough approximation to the logarithm of the Bayes factor. Specifically, the BIC is defined as $-2 \ln L + k \ln n$, where L is the maximized value of the likelihood function for the estimated model, k represents the number of parameters, and n is the number of observations. Therefore, between two competing models we should prefer, according to this criterion, the one with the smaller BIC. In the case of models in Table 3 and 5, the BIC is considerably smaller for Galicia and the Basque Country but only slightly so in Catalonia. Therefore, by all standards we should prefer the conventional Manhattan metric over the mixed model with dimension-by-dimension penalization. To put it differently, it is hardly the case that checking the effect of squaring the distance to parties constitutes a "robustness" test, given that it implies a different set of assumptions which, at least in the sample I used, lead to two undesirable results: poor overall fit and systematic errors in predictions. Thus, to the theoretical cautions that were raised above against the squared penalization for each dimension in the theoretical section, we can add now an empirical inadequacy.

However, the main aim of the paper is to check the assumption of separability —in detriment of integrality— that is implicit in previous research. In order to do this, in Table 7 I show the results of applying a conventional Euclidean model (equation 3) to data. In terms of equation 7, the aim now is to estimate a model in which $\rho = 2$ is assumed. The model is show in terms of α and θ, that were not identifiable in previous estimations, but that carry a more intuitive interpretation.

The results partially agree with what has been shown in previous tables. In the case of Galicia and Catalonia ideology is found to consistently have a higher weight than nationalism (above 70%), but this is no longer the case for the Basque Country. According to results in Table 7 most of the probability of the α parameter for the Basque Country is concentrated around 0.5, that is, it indicates that there is evidence that voters weigh both dimensions equally. As for θ coefficients, in a model with no non-metric covariates they do not have strong substantive implications. However, it is interesting to note how a marginal increase in distance (in whichever dimensions) translate into more than one unit net effect at least for the Basque Country and Catalonia. Furthermore, in all cases, the coefficient is negative, as predicted by the theory.

In terms of party-specific predictions, the model underperforms the Manhattan model in Table 5: the Euclidean model shows systematic deviations (with a 95% probability) for BNG and PSdeG (in Galicia), Aralar and PSE (in Basque Country), and CiU, ICV, and PSC (in Catalonia). However, biases are either similar or relatively small: with respect to the Manhattan model it improves the classification probability for PNV and Aralar —even although it is still biased—, and deteriorates the classification for BNG, PSE, and ICV. As before, these systematic errors do not translate into a general misclassification and in fact the predictive ability only decreases —in mean— by 1% in Galicia and 2.5% in the Basque Country. In addition, the BIC of the models for the three regions worsens with respect to the pure Manhattan model.

Hence, we have reasons to suspect that integrality is not a good assumption and that in fact voters do separate their evaluation of nationalism and ideology wh en deciding which party to vote for. However, we can strengthen the test to award between Manhattan and Euclidean metrics by taking advantage of the fact that the coefficient

TABLE 7: Estimation of models with integrality

	Galicia	Basque Country	Catalonia
	Mean 95% HPD	Mean 95% HPD	Mean 95% HPD
α	0.789	0.585	0.684
	[0.721, 0.845]	[0.472, 0.681]	[0.599, 0.757]
θ	-1.075	-1.314	-1.251
	[-1.116, -0.902]	[-1.441, -1.195]	[-1.359, -1.139]
BIC	2231.686	1433.157	1860.044
N	1944	1160	995

Notes: The table shows the mean, the 2.5 quantile and the 97.5 quantile of the marginal posterior distribution of the coefficient of interest. α is the coefficient of the ideological distance. θ is the coefficient on the metric component. I have omitted party-specific intercepts for ease of presentation.

TABLE 8: Model fit from model in Table 7

Galicia				Basque Country				Catalonia			
	Mean	2.5%	97.5%		Mean	2.5%	97.5%		Mean	2.5%	97.5%
BNG	0.84	0.77	0.95	Aralar	0.66	0.57	0.76	CiU	1.07	1.02	1.12
PP	0.99	0.97	1.01	PNV	1.03	0.99	1.08	ERC	0.96	0.84	1.06
PSdeG	1.10	1.03	1.15	PP	0.92	0.78	1.17	ICV	0.59	0.44	0.75
				PSE	1.13	1.07	1.17	PP	1.09	0.86	1.30
								PSC	1.19	1.06	1.33
Correct	0.793	0.783	0.799	Correct	0.777	0.761	0.785	Correct	0.673	0.662	0.682

Notes: The table shows the distribution (mean, 2.5 and 97.5 quantiles) of cases that are predicted to vote for a given party according to the model over the observed number of votes in the survey. The last row summarizes the percentage of correctly classified cases.

FIGURE 5: Expected probability of voting for the PSOE according to the models in Table 7.

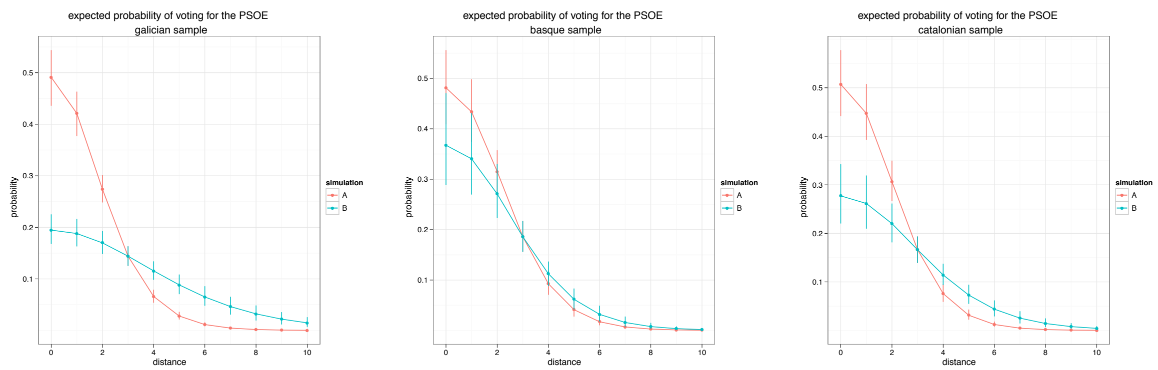


TABLE 9: Estimation of the general model

	Galicia	Basque Country	Catalonia
	Mean	Mean	Mean
	95% HPD	95% HPD	95% HPD
ρ	1.025	0.925	1.177
	[0.851, 1.217]	[0.774, 1.122]	[0.962, 1.435]
α	0.723	0.621	0.677
	[0.685, 0.764]	[0.557, 0.662]	[0.632, 0.724]
θ	-1.090	-1.298	-1.236
	[-1.171, -1.009]	[-1.410, -1.188]	[-1.334, -1.135]
BIC	2195.243	1382.461	1841.927
N	1298	1160	995

Notes: Estimation of the model in equation 7. The table shows the mean, the 2.5 quantile and the 97.5 quantile of the marginal posterior distribution of the coefficient of interest. ρ is the exponent of the metric. α is the coefficient of the ideological distance. θ is the coefficient on the metric component. I have omitted party-specific intercepts for ease of presentation.

TABLE 10: Model fit from model in Table 9

Galicia				Basque Country				Catalonia			
	Mean	2.5%	97.5%		Mean	2.5%	97.5%		Mean	2.5%	97.5%
BNG	0.88	0.79	1.04	Aralar	0.64	0.56	0.74	CiU	1.06	1.00	1.12
PP	0.98	0.96	1.02	PNV	1.06	1.02	1.09	ERC	0.98	0.86	1.07
PSdeG	1.09	0.98	1.16	PP	0.88	0.76	1.00	ICV	0.60	0.46	0.76
				PSE	1.11	1.06	1.15	PP	1.02	0.73	1.30
								PSC	1.18	1.05	1.33
Correct	0.796	0.785	0.803	Correct	0.786	0.777	0.795	Correct	0.679	0.668	0.688

Notes: The table shows the distribution (mean, 0.025 and 0.975 quantiles) of cases that are predicted to vote for a given party according to the model over the observed number of votes in the survey. The last row summarizes the percentage of correctly classified cases.

ρ in equation 7 that differentiates the two metrics, can be estimated from data. The result of this estimation is shown in Table 9. Note again that this is the most general model we can aspire to test that does not include penalizations.

The result of the estimation of ρ is clear in all three regions, and does not deserve further comments. We can accept, with a 95% probability, that $\rho = 1$ in all three regions, and in fact the distribution is very close to being centered around 1, with a sharp mean of the posterior distribution of 1.025 in Galicia, and a 0.925 in the Basque Country. Only in Catalonia is the value not obvious, even though 1 falls within any reasonable probability interval we can fix around the mean of the posterior distribution. However, it is important to note that the correction for missing data is most significant in this last region. In the results shown in Table 9, the distribution of ρ is centered around 1.177 but close enough to 1 for us to accept the Manhattan metric. However, the value 1 falls outside the 95% probability interval (mean in 1.304) if we assume that the dataset is MCAR, this is, if we ignore the missing data problem. In other words, in the full specification, omiting a correction for missing data seems to have noticeable impacts on a key variable to the extent that it might make us reject the hypothesis of a Manhattan metric.

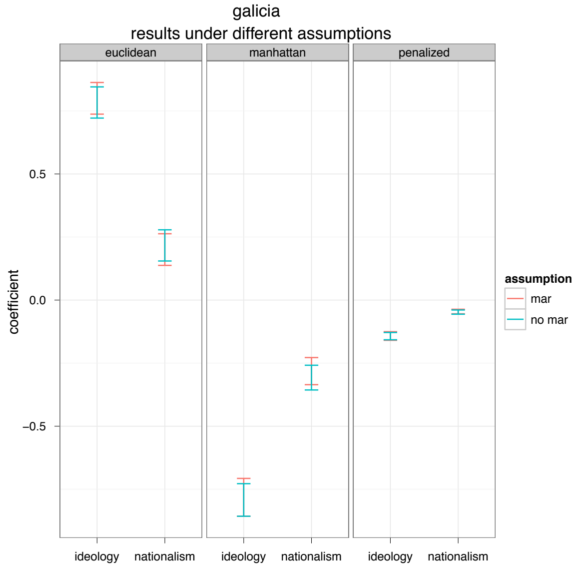
CONCLUSIONS

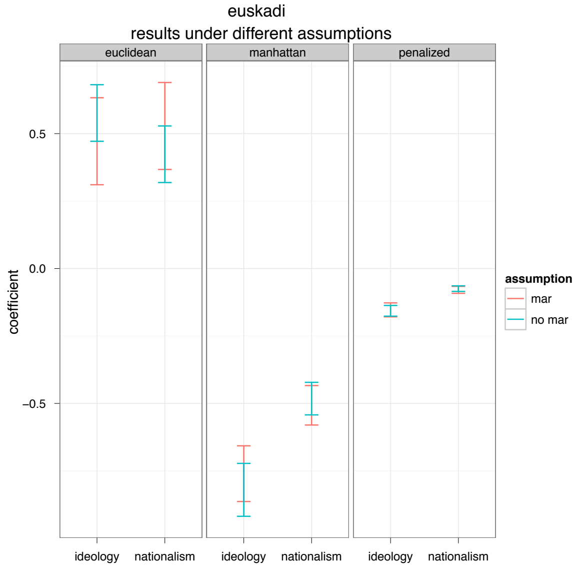
Previous research on the spatial determinants of voting behavior in those Spanish regions in which national identification plays a role in the political arena, found that voters tend to be more attentive to the ideological distance they perceive with respect to parties, rather than to their nationalistic stand. However, this finding was grounded in an untested empirical assumption, particularly about the way voters evaluate distance to parties. By employing a Bayesian alternative to the approach suggested by Beauchamp (2008), I sought to overcome this limitation by analyzing the effect of different assumptions on the specification of the utility function for voters in regional elections in Galicia, the Basque Country and Catalonia. My results advocate a strong preference for a conventional Manhattan metric, which is the specification that —fortunately— most of the literature has favored. Interestingly enough, this extremely stylized model performs remarkably well in terms of prediction. This last observation is important if we compare it with the effect of assuming an incorrect metric: although the overall fit is roughly similar to what we obtained from the full model, a wrong assumption about the cognitive metric results in non-ignorable biases in the allocation of voters to parties.

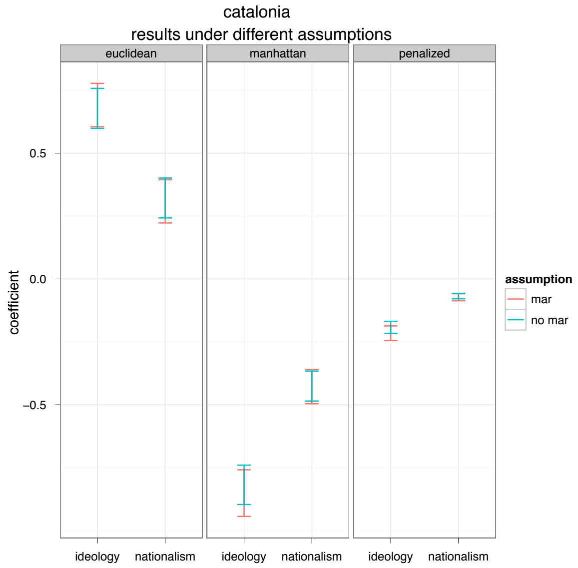
An additional advantage of the specification that I employed in this research is that it allowed me to decompose the conventional parameters that appear in the literature into more meaningful values that have an intuitive interpretation. Future research which explores the mechanisms that complement the basic spatial model in more detail, can benefit from this more general specification, particularly the separation between θ and α. The results presented here also highlight the potential negative effects of ignoring missing data in this setting, even though it does not seem to have much of an effect on the estimation of the taste parameters which are the quantities of interest in the substantive research.

APPENDIX

FIGURE 6: Comparison of the coefficients between different specifications. Coefficients are presented in their original metric. In red, model assuming MAR; in green, model not assuming MAR.







REFERENCES

Aguilar, P. and I. Sánchez-Cuenca (2007). ¿Gestión o representación? Los determinantes del voto en contextos políticos complejos. *Revista Española de Investigaciones Sociológicas 11*, 61–86.

Alvarez, M. R. and J. Nagler (1998). When politics and models collide: estimating models of multicandidate elections. *American Journal of Political Science 42*(1), 55–96.

Balcells, L. (2007). ¿Es el voto nacionalista un voto de proximidad o un voto de compensación? Una nueva aproximación "espacial" al voto en dos dimensiones. *Revista Española de Ciencia Política 16*, 61–88.

Beauchamp, N. (2008). How do we combine issues? Estimating spatial metrics and utility functions using maximum likelihood. Unpublished manuscript. New York University.

Berinsky, A. and J. Lewis (2007). An estimate of risk aversion in the US electorate. *Quarterly Journal of Political Science 2*(2), 139–154.

de la Calle, L. (2005). Cuando la proximidad deja de ser importante: modelos espaciales y voto en la política vasca. 1994-2001. *Revista Española de Ciencia Política 12*, 21–52.

de la Calle, L., L. Orriols, and Á. Martinez (2010). Voting without ideology. Evidence from Spain (1978-2004). *Revista Española de Investigaciones Sociologicas 129*, 107–129.

Downs, A. (1957). *An Economic Theory of Democracy.* New York: Harper and Row.

Eguia, J. (2009). Utility representations of risk neutral preferences in multiple dimensions. *Quarterly Journal of Political Science 4*(4), 379–385.

Feldman, S. and P. Conover (1983). Candidates, issues and voters: The role of inference in political perception. *The Journal of Politics 45*(04), 810–839.

Fernández-Albertos, J. (2002). Votar en dos dimensiones: el peso del nacionalismo y la ideología en el comportamiento electoral vasco, 1993-2001. *Revista Española de Ciencia Política 6*, 153–181.

Garner, W. R. (1974). *The Processing of Information and Structure.* Wiley New York.

Gill, J. (2007). *Bayesian Methods. A Social and Behavioral Sciences Approach* (2nd ed.). London: Chapman and Hall.

Grynaviski, J. and B. Corrigan (2006). Specification issues in proximity models of candidate evaluation (with issue importance). *Political Analysis 14*(4), 393–420.

Kass, R. E. and A. E. Raftery (1995). Bayes factors. *Journal of the American Statistical Association 90*(430), 773–795.

Kedar, O. (2009). *Voting for Policy, Not Parties: How Voters Compensate for Power Sharing.* Cambridge University Press Press.

Laver, M. (1998). *Private desires, political action.* Sage.

Laver, M. and M. Humphreys (2009). Spatial models, cognitive metrics, and majority rule equilibria. *British Journal of Political Science 40*, 11–30.

Little, R. and D. Rubin (1987). *Statistical Analysis with Missing Data.* Wiley New York.

Maddala, G. S. (1986). *Limited-dependent and qualitative variables in econometrics.* Cambridge University Press.

Maravall, J. M. and A. Przeworski (2001). Political reactions to the economy: the Spanish experience. In S. C. Stokes (Ed.), *Public Support for Market Reforms in New Democracies*, pp. 35–76. Cambridge University Press.

Merrill, S. and B. Grofman (1999). *A unified theory of voting: Directional and proximity spatial models.* Cambridge Univ Press.

Montero, J. R. and M. Torcal (1990). Autonomias y comunidades autónomas en España: preferencias, dimensiones y orientaciones políticas. *Revista de Estudios Políticos 70*, 33–91.

Page, B. I. and C. C. Jones (1979). Reciprocal effects of policy preferences, party loyalties, and the vote. *The American Political Science Review 73*, 1071–1789.

Quinn, K., A. Martin, and A. Whitford (1999). Voter choice in multi-party democracies: A test of competing theories and models. *American Journal of Political Science 43*(4), 1231–1247.

Rabinowitz, G. and S. Macdonald (1989). A directional theory of issue voting. *The American Political Science Review 83*(1), 93–121.

Rubin, D. (1976). Inference and missing data. *Biometrika 63*, 581–592.

Sánchez-Cuenca, I. (2008). How can governments be accountable if voters vote ideologically? In J. M. Maravall and I. Sánchez-Cuenca (Eds.), *Controlling Governments: Voters, Institutions and Accountability*, pp. 45–81. Cambridge University Press.

Schafer, J. (1997). *Analysis of Incomplete Multivariate Data.* Chapman & Hall/CRC.

Serrano, A. (1998). Manifestaciones étnicas y cívico-territoriales de los nacionalismos. *Revista Española de Investigaciones Sociológicas 82*, 97–125.

Urquizu-Sancho, I. (2006). The non-declared vote in the surveys: The Spanish case in the 1980s. *Electoral Studies 25*, 103–128.

Westholm, A. (1997). Distance versus direction: The illusory defeat of the proximity theory of electoral choice. *The American Political Science Review 91*(4), 865–883.

Center for Advanced Study in the Social Sciences (http://www.march.es/ceacs/ingles/ceacs.asp)

WORKING PAPERS

Series Editor: Andrew Richards

Most recent titles:

Working Paper	Author	Title
2009/246	Amat, F. and Wibbels, E.	*Electoral Incentives, Group Identity and Preferences for Redistribution.*
2009/247	Alonso, S. and Claro da Fonseca, S.	*Immigration, Left and Right.*
2009/248	Amat, F., Jurado, I., and León, S.	*A Political Theory of Decentralization Dynamics.*
2010/249	Kselman, D. M.	*Electoral Institutions, Legislative Accountability, and Political Corruption.*
2010/250	Pop-Eleches, G. and Tucker, J. A.	*After the Party: Legacies and Left-Right Distinctions in Post-Communist Countries.*
2010/251	Urquizu, I.	*Moderate Voters and Internal Party Factionalism.*
2010/252	Kuo, A. and Margalit, Y.	*Measuring Individual Identity: Experimental Evidence.*
2010/253	Kuo, A.	*Explaining Employer Coordination: Evidence from the US Inter-War Period.*
2011/254	Falcó-Gimeno, A.	*Portfolio Allocation and Time out of Office in Coalition Governments.*
2011/255	Gorodzeisky, A. and Semyonov, M.	*Occupational Incorporation of Immigrants in Western European Countries.*
2011/256	Brader, T. A. and Tucker, A.	*Follow the Leader: Party Cues, Policy Opinion, and the Power of Partisanship in Three Multiparty Systems.*
2011/257	De la Calle, L. and Roussias, N.	*Independents and Vote Choice: Spatial or Performance Voting?*
2011/258	Lapuente, V.	Trade-Offs in Corporate Governance.
2011/259	Lago, I.	Why (Not) National Party Systems?
2011/260	Burgoon, B.	Immigration, Integration and Support for Redistribution in Europe.
2011/261	Kelly, J., Hamann, K., and Johnston, A.	Unions against Governments: Explaining General Strikes in Western Europe, 1980-2006.
2011/262	Riera, P.	Electoral Coordination in Mixed-Member Systems: Does the Level of Democratic Consolidation Matter?
2011/263	Aguilar, P.	Authoritarian Repression, Judicial System and Transitional Justice: The Spanish Case in Comparative Perspective.
2011/264	Cebolla, H. and Finotelli, C.	Integration beyond Models: An Empirical Outlook to the Impact of Integration Models.