

CIRJE-F-822

**Euclidean Revealed Preferences:  
Testing the Spatial Voting Model**

Marc Henry

Université de Montréal

Ismael Mourifié

Université de Montréal

October 2011

CIRJE Discussion Papers can be downloaded without charge from:

<http://www.cirje.e.u-tokyo.ac.jp/research/03research02dp.html>

Discussion Papers are a series of manuscripts in their draft form. They are not intended for circulation or distribution except as indicated by the author. For that reason Discussion Papers may not be reproduced or distributed without the written consent of the author.

# EUCLIDEAN REVEALED PREFERENCES: TESTING THE SPATIAL VOTING MODEL

MARC HENRY AND ISMAEL MOURIFIÉ

ABSTRACT. In the spatial model of voting, voters choose the candidate closest to them in the ideological space. Recent work by (Degan and Merlo 2009) shows that it is falsifiable on the basis of individual voting data in multiple elections. We show how to tackle the fact that the model only partially identifies the distribution of voting profiles and we give a formal revealed preference test of the spatial voting model in 3 national elections in the US, and strongly reject the spatial model in all cases. We also construct confidence regions for partially identified voter characteristics in an augmented model with unobserved valence dimension, and identify the amount of voter heterogeneity necessary to reconcile the data with spatial preferences.

**Keywords:** revealed preference, partial identification, elliptic preferences, voting behavior

**JEL subject classification:** C19, D72

---

*Date:* The present version is of July 28, 2011. Financial support from SSHRC Grant 410-2010-242 is gratefully acknowledged. The authors are also grateful to the co-editor Thierry Magnac, three anonymous referees, Russell Davidson, Romuald Méango, Aureo de Paula, Yves Sprumont, Bernard Salanié, Étienne de Villemeur and participants at the *Montreal Conference on Revealed Preferences and Partial Identification* and the Penn State econometrics seminar for helpful discussions. The usual disclaimer applies. Correspondence address: Département de sciences économiques, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal QC H3C 3J7, Canada. Tel: 1(514)343-2404. Fax: 1(514)343-7221.

## INTRODUCTION

The analysis of voting decisions is an integral part of the revealed preference theory of non-market interactions. A dominant framework in the analysis of voting data is the spatial theory of voting of (Hotelling 1929) and (Downs 1957), which characterizes voters and candidates in an election by their positions in a common ideological space and postulates that voters choose candidates closest to them in that space (see (Hinich and Munger 1997) and (Poole 2005) for accounts of the theory). Two fundamental questions arise regarding the empirical content of this theory.

- (1) Are the distribution of voter ideological positions (hence voter preferences) identified on the basis of voting choices?
- (2) Can the fundamental behavioral assumption of the spatial theory be rejected on the basis of voting data?

Following the work of (Heckman and Snyder 1997) and (Poole and Rosenthal 1997), data on the position of candidates in a two-dimensional ideological space are now widely available (see also (Poole 2005) and references therein). Hence the first question can be reformulated in the following way: does the maintained assumption of spatial voting allow the identification of voter positions in the ideological space on the basis of the knowledge of candidate positions and aggregate voting outcomes? This question is answered affirmatively by (Merlo and de Paula 2009) who also provide a nonparametric estimation strategy with data on repeated elections. The second question is tackled in different guises by (Bogomolnaia and Laslier 2007), (Degan and Merlo 2009) and (Kalandrakis 2010). (Bogomolnaia and Laslier 2007) find the minimal dimension for the ideological space such that voter preference orderings over a finite number of candidates can be represented by spatial utility functions. (Kalandrakis 2010) derives testable restrictions on the positions of voters' ideal points based on a finite number of binary voting choices, when the positions of the alternatives are known. (Degan and

Merlo 2009) establish conditions for the falsifiability of the spatial model in the context of multiple elections, individual voting profile data and observable candidate positions. Falsifiability then derives from the multiplicity of elections; in a single election, a vote for candidate  $j$  is compatible with the voter's ideal point lying in the Voronoi cell of candidate  $j$ 's position (i.e. all points in the space for whom  $j$  is closest among all alternatives). With multiple elections, a voting profile is compatible with the voter belonging to the intersection of the Voronoi cells of respective candidates chosen in that voting profile. When the number of elections is strictly larger than the dimension of the ideological space, some such intersections are empty and the spatial model is falsifiable.

Based on this result, we first provide an econometric methodology to formally test the hypothesis of spatial voting based on observed individual voting profiles from multiple elections, where the ideological position of candidates are known. The main difficulty in setting up a test of the spatial voting model is that the model does not pin down the data generating process for the voting profiles. This is in sharp contrast with the literature on compatibility of discrete choice probabilities and stochastic utility maximization. (Daly and Zachary 1979) and (McFadden 1979) give necessary and sufficient conditions for a discrete choice distribution to be compatible with the maximization of an additively separable random utility. See also (Borsch-Supan 1990) and (Koning and Ridder 2003). The compatibility conditions rely crucially on the coherency of the model, in the sense of (Heckman 1978) and (Gouriéroux, Laffont, and Monfort 1980), and identifiability of its components.

We show that in the present context, despite lack of identification, the null hypothesis of compatibility between the voting data and the spatial model can still be formally tested with an appeal to partial identification techniques recently developed in (Galichon and Henry 2011). This partial identification approach to revealed preference testing relates our work directly to (Blundell, Browning, and Crawford 2008), (Hoderlein and Stoye 2009) and, even more closely (Kawai and Watanabe 2010). The latter partially identify preference parameters in a voting model in which a

fraction of voters incorporate strategic considerations in their voting decisions. A partially identified setting also arises in (Hoderlein and Stoye 2009) who provide a test of WARP based on consumption data from repeated cross sections of heterogeneous consumers. Finally, (Blundell, Browning, and Crawford 2008) also provide empirical implications of revealed preference axioms in the forms of bounds on demand responses.

We test spatial voting in US National elections for years 2000, 2004, and 2008 and the whole data combined. In all three elections, we reject the hypothesis that voters have elliptic preferences over candidates. That is to say, we reject the spatial voting model with heterogeneity of unknown form, both in voter bliss points and in the distances characterizing preferences. This brings new striking evidence to bear on the debate over the adequacy of the spatial voting model in explaining stylized facts on the positioning of political party platforms, and the convergence to the center implied by median voter results (see (Zakharov 2008) for an excellent account). Former empirical analysis and tests of the spatial voting model were conducted assuming knowledge of voter ideological positions (see (Alvarez and Nagler 1998), (Jeong 2008) and references therein). The latter is a reasonable assumption, when analyzing roll call voting in the House and the Senate, but much less so, when analyzing voter behavior in general elections.

Another substantial distinction with roll call voting is the coexistence and competition of two voting logics in general elections, ideology versus performance. The spatial model describes voters' preferences over the candidate's program, whereas preferences over the candidate herself, involving charisma, experience and competence, are typically captured with an additive non spatial term in the utility, generally called *valence*. A fundamental difference between the valence dimensions and the dimensions of the ideological space is that preferences are satiated relative to the latter only. The spatial model is augmented with a valence dimension in (Ansolabehere and Snyder 2000) and

(Groseclose 2001). (Azrieli 2009) axiomatizes the model and (Schofield 2007) shows that incorporating the valence dimension leads to equilibria, where party platforms do not all converge to the center of the electoral distribution.

As in the problem of testing the validity of the spatial model, the problem of estimating and testing valence specifications of the spatial model is complicated by the fact that voter's ideal points are unobserved and the model only partially identifies the distribution of voting profiles. We show how to construct confidence regions for the partial effect of distance to ideal point and of valence characteristics in the utility specification. Of particular empirical relevance are the spatial preference parameter configurations compatible with the smallest values of valence dispersion. These can be interpreted as estimates of the spatial preference parameters that best rationalize the data. A notable finding is that voter differ from candidates in their perception of the relevant ideological dimensions and that the liberal-conservative axis of the standard ideological space dominates the social issues axis in voter preferences. This opens the possibility for parties to gain vote shares by rebalancing political platforms.

The rest of this paper is organized as follows. In the Section 1 we characterize the empirical content of the spatial model of voting. In Section 2, we introduce valence. The data is described in Section 3 and the empirical results in Section 4. The last section concludes.

## 1. EMPIRICAL CONTENT OF THE SPATIAL VOTING MODEL

**1.1. Analytical framework.** The spatial model of voting postulates a common ideological space  $Y \subseteq \mathbb{R}^k$ , where  $k$  is the number of ideological dimensions. Voters face  $m > k$  simultaneous elections, each indexed by  $e \in \{1, \dots, m\}$ . In election  $e$ , each voter chooses exactly one candidate  $j^e \in J^e = \{1, \dots, q^e\}$  among the  $q^e$  candidates competing in election  $e$ . All candidates  $j \in \cup_{e=1}^m J^e$

are characterized by their position  $y^j$  in the ideological space, which is observed by the voters and the analyst. In the rest of this work, we shall consider only two-candidate elections, hence  $q^e = 2$ .

Because of its elegance, simplicity and interpretability, the spatial voting model has dominated a large section of the literature on the analysis of voting choices. An excellent account can be found in (Zakharov 2008). Following the principal component analysis of roll call voting in the US Congress, the two-dimensional NOMINATE Common Space (Poole and Rosenthal 1997) has become a staple of empirical work on the issue. Ideological positions of members of Congress are estimated on the unit square of a two-dimensional space, the first axis of which is usually interpreted as the liberal-conservative axis, measuring economic conservatism, and the second axis of which is usually interpreted as measuring position on social issues. Given the prevalence of this two-dimensional spatial voting model, we will concentrate on the case  $k = 2$  in the rest of this work.

Voters are said to have Euclidean preference (or to “vote ideologically”) if their preferences are satiated at a bliss point  $y^i$  for voter  $i$  in the ideological space and if they maximize a utility function, which is decreasing in the Euclidean distance between their bliss point and the position  $y^j$  of the chosen candidate.

**Definition 1** (Euclidean preferences). *Voter  $i \in I$  has Euclidean preferences (“votes ideologically”) if there exists  $y^i \in Y$  such that voter  $i$  chooses to vote for candidate  $j$  in each election  $e$ , denoted  $v_e^i = j$ , if and only if  $y^j$  minimizes distance  $d(y^i, y; \omega)$  among  $y^{j'}$ ,  $j' \in J^e$ , where  $\omega = (\omega_1, \omega_2)'$ ,  $\omega_2^2 < \omega_1$  and*

$$d(y, y'; \omega) = (y - y')^t \begin{pmatrix} 1 & \omega_2 \\ \omega_2 & \omega_1 \end{pmatrix} (y - y').$$

Figure 1 shows the elliptic indifference curves for a voter, whose ideological bliss point is  $y^i$  and whose utility when candidate with position  $y^j$  is elected is a negative non increasing function of  $d(y^i, y^j; \omega)$ . The dotted circle represents an indifference curve for voter  $i$  when  $\omega_1 = 1$  and  $\omega_2 = 0$ ,

i.e. when the space of reference for candidate coordinates is the effective ideological space for the voter. Note that candidates  $y^j$  and  $y^l$  are both on the circle, and hence are indifferent. The dotted vertical ellipse represents the indifference curve for voter  $i$  when  $\omega_2 = 0$  and  $\omega_1 < 1$ . In that case, the main axes are still the effective ideological dimensions, but in the given units of measurements, the horizontal axis is dominant in the sense that  $y^l$ , closer to  $y^i$  in the horizontal dimension, is preferred to  $y^j$ , closer to  $y^i$  in the vertical dimension. When  $\omega_2 \neq 0$ , the ideological space is different from the reference space.  $y^j$  is now preferred to  $y^l$ .

**1.2. Falsifiability of the model.** Denoting  $v^i = (v_1^i, \dots, v_m^i)$  the voting profile of voter  $i$ , i.e the collection of candidates voter  $i$  chooses in elections  $e = 1, \dots, m$ , the hypothesis that voters have Euclidean preferences is falsifiable if there exist at least one voting profile  $\tilde{v}$ , which cannot be rationalized by the maximization of Euclidean preferences in each election.

**Example 1.** *In figure 2, we illustrate the case with 2 ideological dimensions and 3 simultaneous elections, with 2 candidates each. The black dots represent candidate positions,  $y^{j_1}$  and  $y^{l_1}$  in election 1 and  $y^{j_2}$  and  $y^{l_2}$  in election 2. The lines  $H_e$ ,  $e = 1, 2$  separate the ideological space into the half-space, where voters vote for candidate  $j_e$  and the half-space, where voters vote for candidate  $l_e$ . The intersection of the half-space, where voters vote for  $j_1$  in the first election, say and the half-space, where voters vote for  $l_2$  in the second elections is denoted  $j_1 l_2$  in the figure. All four possible voting profiles are rationalizable by the maximization of Euclidean preferences, hence with only two elections, the hypothesis is not falsifiable. In addition to the half-spaces defined above,  $H_3$  separates the ideological space in two regions, one, where voters prefer (are closer to) candidate  $j_3$  and one, where voters prefer (are closer to) candidate  $l_3$ . The intersection of three half-spaces corresponds to a particular voting profile. For any configuration of candidate positions such that  $y^j \neq y^l$  when  $j \neq l$ , there is exactly one voting profile which is incompatible with Euclidean preference, hence the hypothesis is falsifiable. In figure 2, the half-spaces of voters, who vote for  $j_1$  in election 1, for*



$j_2$  in election 2 and for  $l_3$  in election 3 have empty intersection. Hence, voting profile  $j_1 j_2 l_3$  is incompatible with spatial voting.

More generally, (Degan and Merlo 2009) show that for two candidate elections, falsifiability is equivalent to  $m > k$ . In a small data exploration on US National elections, where voters are faced with  $m = 3$  simultaneous elections (presidential, senate and house) and individual voting profiles and candidate ideological positions in  $\mathbb{R}^2$  are observed, (Degan and Merlo 2009) find evidence of violations of ideological voting as defined in Definition 1 with  $\omega_1 = 1$  and  $\omega_2 = 0$  constant across voters. In the data set that we consider, with the US 2000, 2004 and 2008 elections, we report incidence of violations of the spatial model with no heterogeneity in Table 3. In order to evaluate the statistical significance of these violations and examine alternative specifications, some form of voter heterogeneity needs to be introduced in the utility specification. Unobserved heterogeneity may be entertained within the spatial model in the form of voter specific distance  $d(., .; \omega)$  as we describe in the next section. It may also take the form of a non spatial random utility term, when allowing for voters' response to non ideological characteristics of the candidate. We defer the discussion of the latter form of voter heterogeneity to Section 2 below.

**1.3. Unobserved preference heterogeneity.** Preference heterogeneity may be entertained within the framework of the spatial voting model by assuming that voters may differ in their perception of the relevant ideological dimensions. That is to say, the distance  $d(., .; \omega)$  in the spatial utility function may be voter specific. Thus the requirement that voters have Euclidean preferences according to definition 1 is equivalent to a requirement that all voters have *outward decreasing elliptic preferences* in that each voter  $i$  with position  $y^i$  chooses candidate to minimize  $d(y^i, y^j; \omega^i)$ , where  $\omega^i$  is voter specific. To conduct a revealed preference test of this assumption of elliptic preference maximization, we need to characterize the empirical content of the assumption as follows.

Let  $X$  denote the set of observable variables, which includes the positions of all candidates in the elections. A given voter with bliss point (position in the ideological space)  $y$  and preference parameter  $\omega$  characterizing the shape of her indifference curves is facing  $m = k + 1 = 3$  elections (recalling that  $k = 2$  is the dimension of the ideological space) characterized by the vector  $X$  of positions of all candidates. Given  $(y, \omega, X)$ , the resulting voting profile  $v$  is uniquely determined by the voting model as long as assumptions on the distribution of candidates rule out ties. Denote the unique implied voting profile  $v = g(\omega|X, y)$ , which is the profile of choices of candidates  $j_e$  such that  $y^{j_e}$  minimizes  $d(y, y^{j'_e}; \omega)$  among candidates  $j'_e \in \{j_e, l_e\}$  in each election  $e$ . For instance, in example 1,  $g(\omega|X, y) = l_1 l_2 j_3$  when  $y$  belongs to the central triangle. However, the position  $y$  of voters is unobservable, hence all that utility maximization predicts is that  $v$  lies within the set of compatible voting profiles, which depends on the positions of candidates  $X$  and the realization  $\omega$  of preference heterogeneity. For instance, in example 1, the model only tells us that  $v$  lies in  $\{j_1 j_2 j_3, j_1 l_2 j_3, j_1 l_2 l_3, l_1 j_2 j_3, l_1 j_2 l_3, l_1 l_2 j_3, l_1 l_2 l_3\}$ . We denote  $G(\omega|X)$  this set of compatible voting profiles, i.e.,  $G(\omega|X) = \bigcup_y g(\omega|X, y)$ . The model predicts the following bound on the probability of voting profile being  $V = v$ :

$$\mathbb{P}(V = v|X) = \mathbb{P}(g(\omega|X, y) = v|X) \leq \mathbb{P}(v \in G(\omega|X)|X),$$

since the only model prediction is  $g(\omega|X, y) \in G(\omega|X)$  for any  $y$  in the ideological space. Similarly, for any subset  $B$  of the set of all  $2^m$  possible voting profiles, the model predicts the following bound on the probability of the voting profile  $V$  belonging to  $B$ .

$$\mathbb{P}(V \in B|X) \leq \mathbb{P}(G(\omega|X) \cap B \neq \emptyset|X). \quad [\text{EC}]$$

The inequalities above specify a set of bounds on the distribution of unobserved heterogeneity. As shown above, if voters' choices are compatible with spatial preferences with heterogeneity, the inequalities in [EC] are necessarily satisfied. Conversely, if the inequalities in [EC] are satisfied, then

voters' choices can be rationalized by spatial preferences with heterogeneity. In that precise sense, the inequalities in [EC] define *sharp bounds* on the distribution of unobserved heterogeneity. The converse statement is a corollary of Theorem 1 in (Galichon and Henry 2011) (see also (Beresteanu, Molchanov, and Molinari 2011)). One way to gain insight into the proof of this result is to characterize rationalizability of voters' choices by spatial preferences as the existence of an *assignment* of voting profiles  $v$  to unobserved heterogeneity values  $\omega$  satisfying the constraints  $v \in G(\omega|X)$ . By the Marriage Lemma (see (Berge 1962)), such an assignment exists if and only if there is no *over demanded* set of  $\omega$ 's, which in the present setting means there is no subset  $B$  of voting profiles such that  $\mathbb{P}(B|X) > \mathbb{P}(G(\omega|X) \cap B \neq \emptyset|X)$ . The discussion above can be summarized in the following theorem, which gives the characterization of the empirical content of the model.

**Theorem 1** (Empirical content). *The empirical content of the spatial voting model is characterized by the inequalities [EC] for each subset  $B$  of the set of voting profiles.*

Theorem 1 tells us that a test of the spatial voting model with heterogeneity in the shape of the voters' indifference curves, i.e., in  $\omega$ , is equivalent to testing that the inequalities [EC] hold. Letting the distribution of unobserved heterogeneity  $\omega$  be characterized by the parameter vector  $\theta$ , we consider the set  $\Theta_I$  (possibly empty) of values of  $\theta$ , such that [EC] hold, noting that the right-hand side of [EC] depends on the distribution of  $\omega$ , hence on  $\theta$ .  $\Theta_I$  is called the *identified set*.

**Definition 2.** (*Identified set*) *We call identified set  $\Theta_I$  the set of parameter values such that the moment inequalities in [EC] hold.*

By theorem 1,  $\Theta_I$  is exactly the set of parameters  $\theta$  such that the spatial voting model with heterogeneity is not rejected.  $\Theta_I$  is sometimes called *sharp identified set* to emphasize the fact that all values of  $\theta$  in  $\Theta_I$  are observationally equivalent: no value in  $\Theta_I$  can be rejected on the basis of the information contained in the spatial model and the true distribution of voting profiles.

As a result, a test of the inequalities of Theorem 1 is a classical revealed preference test of the spatial voting model. The way we implement the test is to construct a confidence region for the identified set, using the methodology proposed in (Henry, Méango, and Queyranne 2010) and described in detail in Appendix A. In the data set described in Section 3, we find a 99% confidence region for the identified set  $\Theta_I$  to be empty, hence we reject the spatial voting model specification with distance heterogeneity at the 99% level of significance (see Section 4 for details). In other words, the data cannot be rationalized by a model with heterogenous elliptic preferences.

## 2. INTRODUCING VALENCE

The rejection of the spatial model leads to the consideration of a non spatial component in preferences. As mentioned in the introduction, there is a large literature in political science that attempts to reconcile voting models with observed (non converging) distributions of political party platforms by combining two logics of voting in voter preferences, the logic of ideology, in a spatial term, and the logic of performance, in a non spatial non satiated valence term. We now adopt this approach in our empirical investigation of the determinants of voting choices. The specification generally considered in the literature is the following: voter  $i$  maximizes utility

$$U_i(j) = -d(y^i, y^j; \omega)^2 + \epsilon_i^j \quad (2.1)$$

in each election over  $j \in J^e$  candidates. The valence term  $E = (\epsilon_i^{j_1}, \epsilon_i^{l_1}, \dots, \epsilon_i^{j_m}, \epsilon_i^{l_m})$  is independently distributed conditionally on  $X = (y^{j_1}, y^{l_1}, y^{j_2}, y^{l_2}, y^{j_3}, y^{l_3})$ . Let the distribution of the valence term be parameterized with parameter  $\theta$ .

As in the case of heterogeneity in the distance characterizing preferences in Section 1, we denote by  $G(\epsilon|X; \omega)$  the set of compatible profiles for a given realization  $\epsilon$  of unobserved valence heterogeneity.

The same reasoning applies to show that the inequalities

$$\mathbb{P}(V \in B|X) \leq \mathbb{P}(G(\epsilon|X; \omega) \cap B \neq \emptyset|X; \theta) \quad [\text{ECval}]$$

for each subset  $B$  of voting profiles, characterize the empirical content of the spatial model with valence heterogeneity (note that now  $\epsilon$  is the random unobserved heterogeneity, whereas  $\omega$  is treated as a deterministic parameter vector). The characterization of the empirical content of the model in  $[\text{ECval}]$  still involves a large number of inequalities, namely up to one for each subset  $B$  of the set  $V^m$  of voting profiles. It turns out that in the case of valence heterogeneity, we can achieve a dramatic dimension reduction to 2 inequalities. First we show the following lemma.

**Lemma 1** (Incompatible profiles). *Let  $m = k + 1$  elections with 2 candidates in each election, then for almost all  $X$ , there is a pair of profiles  $\bar{v}(X)$  and  $\underline{v}(X)$  such that for all  $\epsilon$ ,  $G^c(\epsilon|X; \omega) = \{\bar{v}(X)\}$  or  $G^c(\epsilon|X; \omega) = \{\underline{v}(X)\}$ .*

In other words, for a given  $X$  characterizing the positions of all candidates, there is exactly one voting profile incompatible with Euclidean preferences and it belongs to a pair  $\{\bar{v}(X), \underline{v}(X)\}$ . This pair is independent of  $\epsilon$ , so that for different values of unobserved heterogeneity  $\epsilon$ , the unique incompatible profile can only take value  $\bar{v}(X)$  or  $\underline{v}(X)$ . In example 1 and figure 2, the two profiles that are potentially incompatible with spatial voting are profiles  $l_1l_2j_3$  and  $j_1j_2l_3$ . The formal proof is given in the appendix, but it is very easy to understand its heuristics in figure 2:  $j_1j_2l_3$  is the only incompatible profile and  $l_1l_2j_3$  is the only profile with compact support, which is compatible with spatial voting. The slopes of separating hyperplanes are independent of  $\epsilon$ . Hence non compact profile supports cannot disappear, and the only profile that can disappear when  $\epsilon$  shifts the location of the lines in the figure is  $l_1l_2j_3$ . The latter disappears and is replaced by  $j_1j_2l_3$  in any of the following three cases:  $H_1$  moves sufficiently to the left,  $H_2$  moves sufficiently to the right or  $H_3$

moves sufficiently to the right. We are now in a position to show our simple characterization of the empirical content of the model.

**Theorem 2** (Empirical content of model with valence). *The empirical content of the spatial voting model is characterized by exactly 2 inequalities*

$$\mathbb{P}(\bar{v}(X)|X) \leq \mathbb{P}(\bar{v}(X) \in G(\epsilon|X; \omega)|X; \theta) \text{ and } \mathbb{P}(\underline{v}(X)|X) \leq \mathbb{P}(\bar{v}(X) \in G(\epsilon|X; \omega)|X; \theta). \quad (2.2)$$

In other words, if the two inequalities are satisfied, then all inequalities in [ECval] are satisfied and the spatial model with valence heterogeneity is compatible with the true voting profile distribution. Conversely, if for some  $X$ , one of these two inequalities is violated, then the true voting distribution is incompatible with the spatial voting model with valence heterogeneity.

We are interested in the set  $\Xi_I$  of parameter values  $\xi = (\omega, \theta)$  of the model (possibly empty) such that (2.2) holds. This set is the *identified set* in the model with valence heterogeneity. Our goal is to build a confidence region of asymptotic level CL for the identified set which is defined as a region  $\hat{\Xi}$  satisfying  $\mathbb{P}(\Xi_I \subseteq \hat{\Xi}) \geq \text{CL}$  asymptotically. The methodology derived from (Henry, Méango, and Queyranne 2010) is detailed in appendix A, where a step-by-step account of the procedure is given. The procedure involves few, relatively simple steps and is computationally efficient. Once the confidence region  $\hat{\Xi}$  is obtained, we can directly test specifications of the spatial model at the same level of significance. Recall that  $\hat{\Xi}$  is a set of values of the parameter vector  $\xi$  that are not rejected. Suppose for illustration purposes, that  $\epsilon$  is normal with mean zero and variance  $\sigma^2$ . Suppose, moreover that the region  $\hat{\Xi}$  does not contain any value of  $\xi$  with  $\sigma < 2$ , then  $\sigma^2 = 2$  is a lower bound on the variance of unobserved valence necessary to rationalize the data with model (2.1). Suppose further that the region  $\hat{\Xi}$  does not contain any value of  $\xi$  with  $(\omega_1, \omega_2) = (1, 0)$ , then the spatial model with no distance distortion is rejected. If in  $\hat{\Xi}$  we always have  $\omega_2 > 0$  and  $\omega_1 > 1$ , we

can reject the hypothesis that the NOMINATE Common Space first coordinate (liberal-conservative axis) matters more to voters than the second coordinate (social issues).

The partial identification approach adopted here is particularly well suited to the revealed preference problem at hand. Indeed, we wish to test to what extent the spatial model rationalizes the data. We have no information about the position of voters in the ideological space, hence it is undesirable to predicate rejection of the spatial model on ad-hoc assumptions on the latter. This would be the case, if we parameterized the distribution of voter positions and valence heterogeneity, imposed additional restrictions for identification of this nonlinear model, estimated with maximum likelihood and tested for the presence of unobserved heterogeneity with a version of (Chesher 1984).<sup>1</sup>

### 3. DATA

Data are drawn from the following sources. The candidate positions are drawn from the Poole and Rosenthal NOMINATE Common Space data set<sup>2</sup> ((Poole and Rosenthal 1997), (Poole 2005)), which gives the position of candidates on a two dimensional ideological space based on individual roll call voting of members of Congress. Only candidates who have held office are included in the data set. A fundamental assumption here is that voters observe the true ideological position of all candidates, whereas the econometrician only observes positions of candidates included in the NOMINATE Common Space data set. Each candidate that has not held office is assigned the position among all candidates in his party and district, which is most favorable to the hypothesis of spatial voting, to ensure that rejection of the model is not driven by missing data issues. This is done by choosing the position, which minimizes the number of voters in the district, who choose one of the pair of forbidden profiles  $\{\bar{v}(X), \underline{v}(X)\}$ .

---

<sup>1</sup>We are grateful to the co-editor Thierry Magnac for this suggestion.

<sup>2</sup>Available at <http://voteview.com>

As for voting choices, they are obtained from the American National Elections Study (ANES), which represents the best and most widely used source of individual-level data on electoral participation and voting in United States<sup>3</sup>. For each election year, ANES contains individual voting decisions in presidential and congressional elections of a nationally representative sample of the voting age population. In addition, the ANES contains information on the congressional district where each individual resides, the identity of the Democratic and the Republican candidate competing for election in his or her congressional district, and, in the event that a Senate election is also occurring in his or her state, the identity of the candidates competing in the Senate race. For each election we consider a sub-sample, which contains only voters, who vote in the three simultaneous elections. Hence, all voters face three simultaneous elections. As in (Degan and Merlo 2009), we match each voter in the ANES sample with the position of the presidential, senate and house candidates. Districts are indexed by  $h = 1, \dots, 435$  and states by  $s = 1, \dots, 50$ . We consider election years 2000, 2004 and 2008.

**3.1. Candidate positions.** Summary statistics on the distribution of candidate positions on the NOMINATE Common Space are described in Table 2. Table 2 shows that democratic candidates are somewhat more dispersed than republican candidates in election year 2000, whereas republican candidates are somewhat more dispersed in election years 2004 and 2008. The overall dispersion is about twice as large as within party dispersion, but significantly smaller than between party dispersion. The 0 in the latter table indicates that in election year 2000, one of the republican candidates shared their ideological position with one of the democratic candidates, indicating that they may have had an identical roll call voting record. The dispersion of elected candidates is greater than within party dispersion, but smaller than the overall dispersion of candidates. On the first axis, sometimes called *liberal-conservative* axis, democratic and republican candidates are clearly separated, in the sense that for every election year as well as for the aggregate data, mean  $\pm$

---

<sup>3</sup>The ANES is available on-line at <http://www.electionstudies.org/>



two standard deviations confidence intervals for democratic and republican candidates are disjoint. This feature is also very visible on the scatter plots of figure 3. On the second axis (also referred to as the *social issues* axis), however, the distributions of democratic and republic candidates are not clearly distinguishable.

**3.2. Voting profiles.** The distributions of voting profiles for elections years 2000 to 2008 are given in Table 1. A large majority of voters choose candidates from the same party in all three elections (75% in 2000, 76% in 2004 and 78% in 2008). Otherwise, no clear pattern arises among the remaining voting profiles.

#### 4. EMPIRICAL RESULTS

**4.1. Distance heterogeneity.** We considered a normal specification for  $\omega_2$  with mean  $\mu_2$  and variance  $\sigma_2^2$  and a log-normal specification for  $\omega_1$  with mean  $\mu_1$  and variance  $\sigma_1^2$ .  $\omega_1$  and  $\omega_2$  were assumed independent. We conditioned on the event  $\omega_2^2 < \omega_1$ , so that  $(\omega_1, \omega_2)$  characterizes a distance. We restricted the parameter space to  $(\mu_1, \mu_2) \in [-5, 5] \times [0, 10]$  and  $(\sigma_1^2, \sigma_2^2) \in [0, 13]^2$ . The 99% level confidence region for the parameter vector  $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$  is empty, so that the spatial voting model with distance heterogeneity is rejected at the 99% level of significance. We conclude that voters' choices cannot be rationalized by spatial voting, even if we let the shape of the indifference ellipses be voter specific.

**4.2. Valence heterogeneity.** We considered two parametric specification for the valence term  $\epsilon$  in specification 2.1. We first modeled  $E = (\epsilon^{j_1}, \epsilon^{l_1}, \epsilon^{j_2}, \epsilon^{l_2}, \epsilon^{j_3}, \epsilon^{l_3})$  as a vector of independent mean zero binary variables taking values  $\eta$  and  $-\eta$ . Some features of the resulting confidence regions are given in figure 4 for the aggregate data, and results from individual elections are available on request. We then considered the specification, where  $E$  is a vector of independent normal variables with mean zero and variance  $\sigma^2$ . Some features of the resulting confidence regions are given in figure 4 for the

aggregate data. Figure 4(c) shows the 3-dimensional confidence region for the parameters of the spatial model 2.1  $(\omega, \eta)$ . It is a 95% confidence region for the set of values of the parameter vector under which the spatial model is valid. Figures 4(a) and (b) show the effect of an increase in the admissible valence dispersion on the set of rationalizable distance distortions  $\omega$ . The pure Euclidean model with  $\omega_2 = 0$  and  $\omega_1 = 1$  (such that indifference curves are circles) lies outside the cut in confidence region in figure 4(a), so it is not rationalizable, but it lies inside the cut in figure 4(b), so it becomes rationalizable for this higher level of admissible valence dispersion. The minimum value of  $\eta$  in the region is 0.19. This indicates that the spatial model can only be rationalized by adding a non ideological term in the utility of magnitude at least 0.19. This is to be compared with the distribution of squared distances between candidates in Table 2. The minimum valence needed to rationalize the valence-augmented spatial model is of the order of the mean of squared distances between democratic candidates, half the mean of squared distances between all candidates and a quarter of the mean of squared distances between democratic and republican candidates. Moreover, for that minimal non spatial utility term, the model can only be rationalized for a specific distance distortion, namely  $\omega = (\omega_1, \omega_2) = (0.33, -0.11)$ . This corresponds to a much greater emphasis on the liberal-conservative axis than on the social axis. This can be seen more clearly on figure 4(d) which shows the tilt of the major axis of the elliptic indifference curves of voters compatible with the minimum valence magnitude. The value is close to 100 degrees, which indicates a situation similar to the dotted vertical ellipse of figure 1. The NOMINATE Common Space is shared by voters, except that greater importance is given to changes on the liberal-conservative axis. We see on figure 4(d) that the tilt of the indifference ellipses remains between 45 and 135 degrees for any value of the non ideological disturbance  $\eta$  below 0.22. So the conclusion on the relative weights of each axis remains.

In case of normal specification of the valence term, the 95% confidence region for the set of parameters  $(\omega, \sigma)$  for which the valence-augmented spatial model is rationalizable is similar to the

previous case. The minimal standard deviation needed to rationalize the valence-augmented spatial model is 0.13. The corresponding indifference curves are ellipses with major axis tilted at 100 degrees again, so that similar conclusions apply. The liberal-conservative axis remains dominant for all values of the valence standard deviation below 0.18. As seen on Table 4, those conclusions remain for data on individual elections, with however a clear trend towards the spatial model, as the lowest valence term necessary to rationalize the data drops from 0.12 in the 2000 elections, to 0.09 in the 2004 election and finally 0.06 in the 2008 elections.

Overall results strongly support the hypothesis that voters' choices are driven by a combination of ideological and competence considerations, as the magnitude of valence dispersion needed to rationalize voter choices is very close to the mean of the distribution of squared ideological distances between candidates from the same party and a fifth of the mean of squared ideological distances between candidates of opposing parties. Results also strongly support the hypothesis that voters give more weight to the liberal-conservative axis in the NOMINATE Common Space. The estimated distance distortion gives candidates a direction in which to rebalance their political program in order to increase their vote share. However, results do not support the hypothesis that the social issues axis is irrelevant in voters' decisions. Indeed, the minimum valence dispersion needed to rationalize voters' choices with a single ideological dimension (liberal-conservative axis alone, namely  $\omega_1 = \omega_2 = 0$ ) is 50% larger than the dispersion needed to rationalize voters' choices with two ideological dimensions (see Table 4).

## CONCLUSION

We have considered the spatial model of voting and provided a methodology for conducting revealed preference tests of spatial preferences. Falsifiability of the model is driven by the existence of voting profiles in multiple elections that are incompatible with maximization of spatial preferences

in each election. The main difficulty in testing the hypothesis of spatial voting is that the latter only specifies a set of voting profiles compatible with spatial voting, and hence only partially identifies the distribution of voting profiles. It is shown here how to circumvent this fundamental characteristic of revealed preference tests with an appeal to recent results in partial identification. The hypothesis of spatial preferences is strongly rejected in a sample of voting profiles from three US elections, and confidence regions are constructed for the set of parameters compatible with the spatial voting model augmented with an unobserved non spatial component. A robust finding from those confidence regions is that the ideological dimension generally associated with economic conservatism dominates the dimension associated with social issues. The methodology is currently being extended to the estimation of revealed spatial preferences over competing characteristics in social networks with homophily, to complement work in (Christakis, Fowler, Imbens, and Kalyanaraman 2010), (Galichon and Salanié 2010) and (Chiappori, Gandhi, Salanié, and Salanié 2009).

#### APPENDIX A. INFERENCE METHODOLOGY

The methodology is detailed for inference on the identified set  $\Xi_I$  in case of valence heterogeneity. It also applies to inference on the identified set  $\Theta_I$  in case of distance heterogeneity with a trivial change of notation,  $\Theta_I$  for  $\Xi_I$  and  $\omega$  for  $\epsilon$ , and replacing the two inequalities of Theorem 2 by the inequalities in [EC].

We are interested in constructing a random region  $\hat{\Xi}$  such that the true identified set  $\Xi_I$  is contained in  $\hat{\Xi}$  with at least probability CL. Given the sample of observations  $((V_1, X_1), \dots, (V_n, X_n))$  for a sample of  $n$  voters, we construct data dependent functions  $\underline{P}(\cdot|x)$  such that the following statement is true with probability tending to no less than CL:  $\underline{P}(\tilde{v}|X_i) \leq P(\tilde{v}|X_i)$  for all  $X_i$ , and  $\tilde{v} \in \{\bar{v}(X_i), \underline{v}(X_i)\}$ . Once this is achieved, the region  $\hat{\Xi} = \{\xi = (\omega, \theta) \mid \underline{P}(\tilde{v}|X_i) \leq \mathbb{P}(\tilde{v} \in G(\epsilon|X_i, \omega)|X_i; \theta), i = 1, \dots, n, \tilde{v}(X) \in \{\bar{v}(X), \underline{v}(X)\}\}$  satisfies the required conditions: indeed,

if  $\xi \in \Xi_I$ ,  $\xi$  satisfies (2.2), so with probability no less than CL, it also satisfies  $\underline{P}(\tilde{v}|X_i) \leq \mathbb{P}(\tilde{v} \in G(\epsilon|X_i, \omega)|X_i; \theta)$ ,  $i = 1, \dots, n$ ,  $\tilde{v}(X) \in \{\bar{v}(X), \underline{v}(X)\}$ , hence it belongs to  $\hat{\Xi}$ . The confidence region is computed by checking every parameter value  $\xi = (\omega_1, \omega_2, \theta)$  on a regular grid of  $10^6$  points. There remains to explain how  $\mathbb{P}(\tilde{v} \in G(\epsilon|X_i, \omega)|X_i; \theta)$  and  $\underline{P}(\cdot|X_i)$  are computed.

**A.1. Computation of  $\mathbb{P}(\tilde{v} \in G(\epsilon|X_i, \omega)|X_i; \theta)$ :** For each value of  $X_i$ ,  $i = 1, \dots, n$  and  $\xi = (\omega, \theta)$  on the grid, draw  $N = 999$  values of valence  $\epsilon^l$ ,  $l = 1, \dots, N$ . For each  $\epsilon^l$ , check whether  $\tilde{v} \in G(\epsilon|X_i, \omega)$  as in Section A.2. Approximate  $\mathbb{P}(\tilde{v} \in G(\epsilon|X_i, \omega)|X_i; \theta)$  with the Monte Carlo probability  $(1/N) \sum_{l=1, \dots, N} 1\{\tilde{v} \in G(\epsilon^l|X_i, \omega)\}$ .

**A.2. Checking whether  $\tilde{v} \in G(\epsilon|X_i, \omega)$ :** In election  $e$  with candidates  $j_e$  and  $l_e$ , voter with position  $y$  and valence perception  $E$  will choose candidate  $j_e$  if  $-d(y, y^{j_e}; \omega)^2 + \epsilon^{j_e} > -d(y, y^{l_e}; \omega)^2 + \epsilon^{l_e}$  which is equivalent to  $y \cdot_{\omega} (y^{l_e} - y^{j_e}) > \frac{1}{2}(\|y^{j_e}\|_{\omega}^2 - \|y^{l_e}\|_{\omega}^2 + \epsilon^{l_e} - \epsilon^{j_e})$ , where “ $a \cdot_{\omega} b$ ” denotes the inner product  $a \cdot_{\omega} b = a^t W b$ ,  $W = \begin{pmatrix} 1 & \omega_2 \\ \omega_2 & \omega_1 \end{pmatrix}$  and  $\|a\|_{\omega} = \sqrt{a \cdot_{\omega} a}$  denotes the corresponding norm. Calling  $\lambda_e = y^{l_e} - y^{j_e}$  and  $\mu_e = \frac{1}{2}(\|y^{j_e}\|_{\omega}^2 - \|y^{l_e}\|_{\omega}^2 + \epsilon^{l_e} - \epsilon^{j_e})$ , the hyperplane  $H_e = \{y \in Y | \lambda_e^t W y = \mu_e\}$  separates the ideological space into two regions  $Y^{j_e} = \{y \in Y | \lambda_e^t W y > \mu_e\}$  and  $Y^{l_e} = \{y \in Y | \lambda_e^t W y < \mu_e\}$ . In  $m$  elections, a profile  $v = (j_1, \dots, j_m)$  corresponds to a voter with ideological position  $y$  in the intersection  $Y^{j_1} \cap \dots \cap Y^{j_m}$ . If this intersection is non empty, the profile is compatible with the voting model and  $\tilde{v} \in G(\epsilon|X_i, \omega)$ .

**A.3. Construction of  $\underline{P}(\cdot|X)$ :** To construct  $\underline{P}(\cdot|X)$  we first compute a nonparametric estimator  $\hat{P}(\cdot|X)$  for  $P(\cdot|X)$  (see (Li and Racine 2008) for the procedure and its properties). Heuristically, we want  $\underline{P} \leq P$ , i.e.  $\underline{P} \leq \hat{P} + (P - \hat{P})$  with probability CL. The natural approach is to choose  $\underline{P}$  equal to  $P$  plus the CL-quantile of the distribution of  $P - \hat{P}$ . However  $\hat{P} - P$  is a random function, hence the quantile of its distribution is not defined. Instead, following (Henry, Méango, and Queyranne 2010), we use the following generalized quantile notion: we draw  $B$  bootstrap

samples  $((V_1^b, X_1^b), \dots, (V_n^b, X_n^b))$ ,  $b = 1, \dots, B$  and compute for each the nonparametric estimator  $\hat{P}^b(\cdot|X)$ . Let  $\phi^b$  be the minimum over  $i = 1, \dots, n$  and  $\tilde{v}(X) \in \{\bar{v}(X), \underline{v}(X)\}$  of the quantity  $\hat{P}^b(\tilde{v}|X_i) - \hat{P}^b(\bar{v}|X_i)$ , and let  $\phi^*$  be ranked  $[B \times \text{CL}]$ 'th in decreasing order among the  $\phi^b$ . Then for all  $i = 1, \dots, n$  and  $\tilde{v} = \underline{v}(X)$  or  $\tilde{v} = \bar{v}(X)$ , set  $\underline{P}(\tilde{v}|X_i) = \hat{P}(\tilde{v}|X_i) + \inf_b \{\hat{P}(\tilde{v}|X_i) - \hat{P}^b(\tilde{v}|X_i) \mid \phi^b > \phi^*\}$ . See (Henry, Méango, and Queyranne 2010) for discussion of the method and its properties.

## APPENDIX B. PROOFS OF RESULTS IN THE MAIN TEXT

*Proof of lemma 1.* As in section A.2, we call  $\lambda_e = y^{l_e} - y^{j_e}$  and  $\mu_e = \frac{1}{2}(\|y^{j_e}\|_\omega^2 - \|y^{l_e}\|_\omega^2 + \epsilon^{l_e} - \epsilon^{j_e})$  and  $H_e = \{y \in Y \mid \lambda_e^t W y = \mu_e\}$  the hyperplane that separates the ideological space into two regions  $Y^{j_e} = \{y \in Y \mid \lambda_e^t W y > \mu_e\}$  and  $Y^{l_e} = \{y \in Y \mid \lambda_e^t W y < \mu_e\}$ . The  $k$  first hyperplanes define a system of  $k$  linear equations in  $k$  variables  $\Lambda y^t = \mu$  where  $\Lambda = [\lambda_1 \dots \lambda_k]$  is a  $k \times k$  matrix. The rank of  $\Lambda$  is equal to  $k$  due to the linear independence of the vectors  $\lambda_e$ ,  $e = 1, \dots, m$ . The equation of the last hyperplane in this same space is define by  $\lambda_{k+1} y' = \mu_{k+1}$ . Transform the ideological space with the change of coordinates  $x^t = \Lambda y^t - \mu$ . The center of the new space is  $y_*^t = \Lambda^{-1} \mu$ , which is well defined since  $\Lambda$  is full rank. We transform also the  $(k+1)^{th}$  hyperplane. It is defined by  $\mathcal{H} = \{x \in X : \lambda_{k+1} \Lambda^{-1} x^t = \mu_{k+1} - \lambda_{k+1} \Lambda^{-1} \mu\}$ . Call  $v^+$  (resp.  $v^-$ ) the orthant defined by  $\{x \in X : \text{sign}(x) = \text{sign}(\lambda_{k+1} \Lambda^{-1})\}$  (resp.  $\{x \in X : \text{sign}(x) = -\text{sign}(\lambda_{k+1} \Lambda^{-1})\}$ ), where the function  $\text{sign}(x)$  is understood element by element when applied to a vector. We then have  $\forall x \in v^+$ ,  $\lambda_{k+1} \Lambda^{-1} x^t \geq 0$  and  $\forall x \in v^-$ ,  $\lambda_{k+1} \Lambda^{-1} x^t \leq 0$ . If  $\mu_{k+1} - \lambda_{k+1} \Lambda^{-1} \mu > 0$ ,  $\mathcal{H}$  cannot partition orthant  $v^-$ , whereas if  $\mu_{k+1} - \lambda_{k+1} \Lambda^{-1} \mu < 0$  it is  $v^+$  which cannot be partitioned. But never both at the same time. Then, since the number of profiles which are not incompatible is  $\rho(k+1, k) = 2^{k+1} - 1$ , all the orthants except  $v^+$  and  $v^-$  are always partitioned, and this for all values of  $(\mu_1, \dots, \mu_{k+1})$ . The existence of an incompatible voting profile is due to the fact that one of two profiles  $\{\bar{v}, \underline{v}\}$  is

not partitioned. Then, the incompatible profile is  $\bar{v}$  if orthant  $v^+$  is not partitioned, and  $\underline{v}$  if orthant  $v^-$  is not partitioned. This is complete our proof.  $\square$

*Proof of Theorem 2.* Call  $\nu$  the distribution of unobserved heterogeneity  $\epsilon$ . All profiles except  $\{\bar{v}(X), \underline{v}(X)\}$  belong to the equilibrium correspondence  $G(\epsilon | X)$  for all  $\epsilon$ . Hence, for any profile  $v \notin \{\bar{v}(X), \underline{v}(X)\}$ , we have  $\nu(G^{-1}(v | X)) = 1$ . In addition,  $\nu(G^{-1}(\bar{v} \cup \underline{v} | X)) = \nu(G^{-1}(\bar{v} | X)) + \nu(G^{-1}(\underline{v} | X))$ , since exactly one of the two belongs to  $G$ . Suppose (2.2) holds. Take any subset  $A$  of voting profiles. If  $A \setminus \{\underline{v}(X), \bar{v}(X)\} \neq \emptyset$  or if  $\{\underline{v}(X), \bar{v}(X)\} \subseteq A$ , then  $\nu(G^{-1}(A | X)) = 1 \geq P(A|X)$ . Otherwise,  $A = \{\underline{v}(X)\}$  or  $A = \{\bar{v}(X)\}$  and  $P(A|X) \leq \nu(G^{-1}(A | X))$  by assumption. Hence (EC) holds, and the proof is complete.  $\square$

## REFERENCES

- ALVAREZ, M., and J. NAGLER (1998): “When politics and models collide: estimating models of multiparty elections,” *American Journal of Political Science*, pp. 55–96.
- ANSOLABEHRE, S., and J. SNYDER (2000): “Valence politics and equilibrium in spatial elections models,” *Public choice*, 51, 327–336.
- AZRIELI, Y. (2009): “An axiomatic foundation for multidimensional spatial models of elections with a valence dimension,” unpublished manuscript.
- BERESTEANU, A., I. MOLCHANOV, and F. MOLINARI (2011): “Sharp identification regions in models with convex predictions: games, individual choice, and incomplete data,” forthcoming in *Econometrica*.
- BERGE, C. (1962): *The Theory of Graphs and its Applications*. John Wiley: New York.
- BLUNDELL, R., M. BROWNING, and I. CRAWFORD (2008): “Best nonparametric bounds on demand responses,” *Econometrica*, 76, 1227–1262.

- BOGOMOLNAIA, A., and J.-F. LASLIER (2007): “Euclidean preferences,” *Journal of Mathematical Economics*, 43, 87–98.
- BORSCH-SUPAN, A. (1990): “On the compatibility of nested logit models with utility maximization,” *Journal of Econometrics*, 43, 371–387.
- CHESHER, A. (1984): “Testing for neglected heterogeneity,” *Econometrica*, 52, 865–872.
- CHIAPPORI, P.-A., A. GANDHI, B. SALANIÉ, and F. SALANIÉ (2009): “Identifying preferences under risk from discrete choices,” *American Economic Review*, 99, 356–362.
- CHRISTAKIS, M., B. FOWLER, G. IMBENS, and K. KALYANARAMAN (2010): “An empirical model for strategic network formation,” unpublished manuscript.
- DALY, A., and S. ZACHARY (1979): “Improved multiple choice models,” in *In D. Hensher and Q. Dalvi(eds), Determinants of Travel Choice*, pp. 337–357. London: Teakfield.
- DEGAN, A., and A. MERLO (2009): “Do voters vote ideologically?,” *Journal of Economic Theory*, 144, 1868–1894.
- DOWNES, A. (1957): “An economic theory of political action in a democracy,” *Journal of Political Economy*, 65, 135–150.
- GALICHON, A., and M. HENRY (2011): “Set identification in models with multiple equilibria,” *Review of Economic Studies*; doi: 10.1093/restud/rdr008.
- GALICHON, A., and B. SALANIÉ (2010): “Matching with trade-offs: revealed preferences over competing characteristics,” unpublished manuscript.
- GOURIÉROUX, C., J.-J. LAFFONT, and A. MONFORT (1980): “Coherency conditions in simultaneous linear equations models with endogenous switching regimes,” *Econometrica*, 48, 675–695.
- GROSECLOSE, T. (2001): “A model of candidate location when one candidate has a valence advantage,” *American Journal of Political Science*, 45, 862–886.



- HECKMAN, J. (1978): “Dummy endogenous variables in a simultaneous equations system,” *Econometrica*, 46, 931–960.
- HECKMAN, J., and J. SNYDER (1997): “Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators,” *RAND Journal of Economics*, 28, S142–S189.
- HENRY, M., R. MÉANGO, and M. QUEYRANNE (2010): “Bootstrap inference in models of interactions,” unpublished manuscript.
- HINICH, M., and M. MUNGER (1997): *Analytical politics*. Cambridge University Press: Cambridge.
- HODERLEIN, S., and J. STOYE (2009): “Revealed preference in heterogeneous populations,” unpublished manuscript.
- HOTELLING, H. (1929): “Stability in competition,” *Economic Journal*, 39, 41–51.
- JEONG, G.-H. (2008): “Testing the predictions of the multidimensional spatial voting model with roll call data,” *Political Analysis*, 16, 179–196.
- KALANDRAKIS, T. (2010): “Rationalizable voting?,” *Theoretical Economics*, 5, 93–125.
- KAWAI, K., and Y. WATANABE (2010): “Inferring strategic voting,” unpublished manuscript.
- KONING, R., and G. RIDDER (2003): “Discrete choice and stochastic utility maximization,” *Econometrics Journal*, 6, 1–27.
- LI, Q., and J. RACINE (2008): “Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data,” *Journal of Business and Economic Statistics*, 26, 423–434.
- McFADDEN, D. (1979): “Econometric models of probabilistic choice,” in *In C. Manski and D. McFadden (eds), Structural analysis of discrete data with econometric application*, pp. 198–272. Cambridge: MIT Press.

- MERLO, A., and A. DE PAULA (2009): "Identification and estimation of preference distributions when voters are ideological," unpublished manuscript.
- POOLE, K. (2005): *Spatial models of parliamentary voting*. Cambridge University Press.
- POOLE, K., and H. ROSENTHAL (1997): *Congress: a political economic history of roll call voting*. Oxford University Press.
- SCHOFIELD, N. (2007): "The mean voter theorem: necessary and sufficient conditions for convergent equilibrium," *Review of Economic Studies*, 74, 965–980.
- ZAKHAROV, A. (2008): "Spatial voting theory: a review of the literature," unpublished manuscript.

TABLE 1. Voting Profiles in U.S Presidential, Senate and House elections.

Voting Profiles	2000	2004	2008	Aggregate
DDD	246	182	277	705
DDR	36	14	44	94
DRD	21	13	16	50
DRR	22	54	11	87
RDD	21	0	19	40
RDR	20	21	6	47
RRD	31	0	17	48
RRR	205	149	124	478
Total	602	433	514	1549

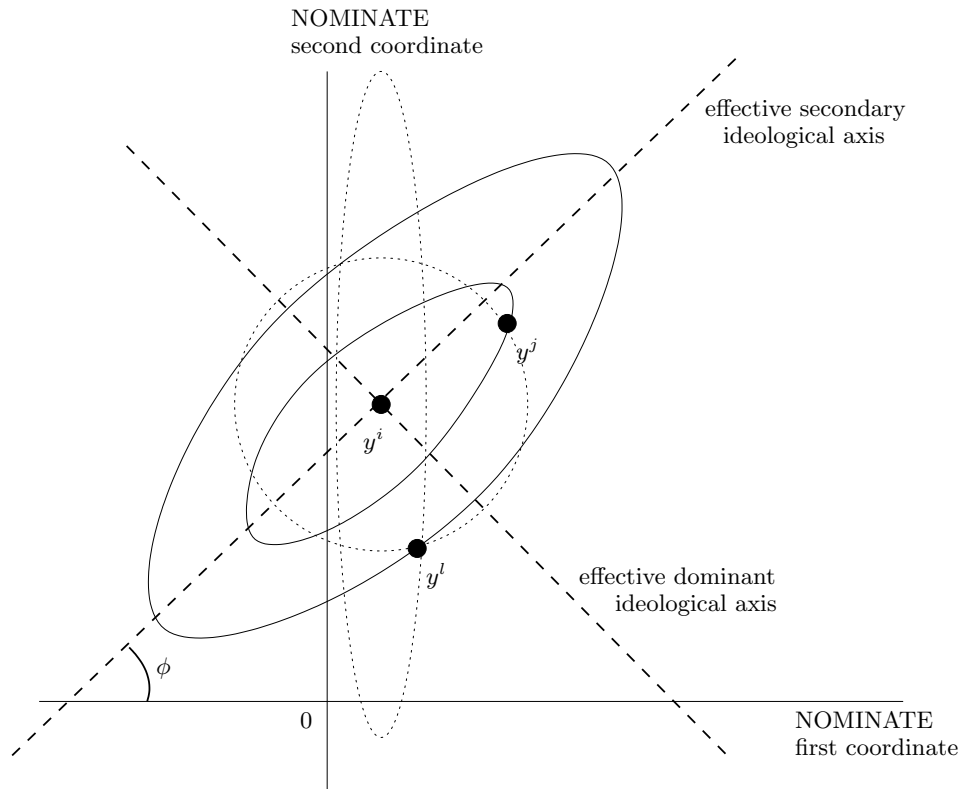
FIGURE 1. Indifference curves for spatial preferences with distance  $d(., .; \omega)$ .

TABLE 2. Distribution of ideological square distances between candidates.

Democratic candidates				
	2000	2004	2008	Aggregate
Mean	0.2113	0.1326	0.1029	0.1711
Std dev.	0.2648	0.1719	0.1121	0.2209
Max	2.1579	1.3474	0.6309	2.7159
Republican candidates				
	2000	2004	2008	Aggregate
Mean	0.1311	0.1656	0.1126	0.1406
Std dev.	0.1544	0.2201	0.1639	0.3275
Max	1.4227	1.6869	0.8461	1.8529
All candidates				
	2000	2004	2008	Aggregate
Mean	0.4038	0.3649	0.3565	0.3880
Std dev.	0.3368	0.3249	0.2993	0.3275
Max	2.1759	1.9907	1.4703	2.7159
Between democrats and republicans				
	2000	2004	2008	Aggregate
Mean	0.6313	0.6130	0.5981	0.6212
sdt dev.	0.2681	0.2591	0.1981	0.2565
Min	0	0.0056	0.1882	0
Max	2.1759	1.9907	1.4703	2.1759
Between elected candidates				
	2000	2004	2008	Aggregate
Mean	0.3259	0.3904	0.3333	0.3464
sdt dev.	0.3121	0.3534	0.2431	0.3050
Min	0.0057	0.0014	0.0030	0.0014
Max	1.3824	1.8855	1.1449	1.8855

TABLE 3. Incidence of violations of spatial voting. The unit of observation is a voting profile.

Without heterogeneity ( $\omega_1 = 1$ and $\omega_2 = 0$ )				
	2000	2004	2008	Aggregate
violations	28	17	5	50
total sample	602	433	514	1549
percentage	4.7%	3.9%	0.7%	3.2%

TABLE 4. Minimum valence dispersion to rationalize voting choices (i), to rationalize voting choices so that social issues dominate (ii), to rationalize voting choices with a single ideological dimension (iii).

## (i) Rationalization of voting choices

	2000	2004	2008	Aggregate
Minimum $\eta$	0.12	0.09	0.06	0.13
Corresponding $\omega$	(0.33,-0.11)	(0.22,0.33)	(0.33,-0.11)	(0.33,-0.11)
Minimum $\sigma$	0.19	0.07	0.05	0.13
Corresponding $\omega$	(0.33,0.11)	(0.222,0.333)	(0.33,-0.11)	(0.33,-0.11)

## (ii) Reversal of ideological axes.

	2000	2004	2008	Aggregate
Minimum $\eta$	0.18	0.18	0.13	0.22
Minimum $\sigma$	0.18	0.18	0.13	0.18

## (iii) Rationalization with a single dimension.

	2000	2004	2008	Aggregate
Minimum $\eta$	0.12	0.13	0.09	0.18
Minimum $\sigma$	0.13	0.17	0.13	0.22

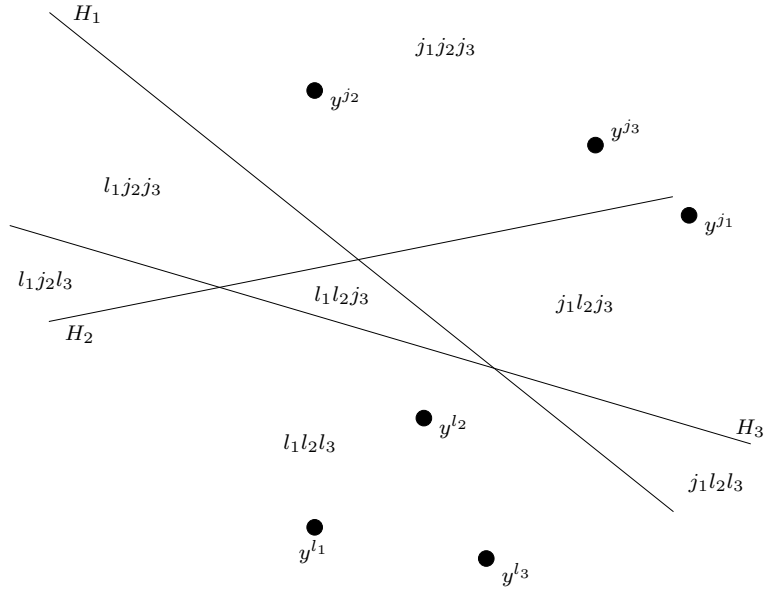


FIGURE 2. Three elections and two ideological dimensions.

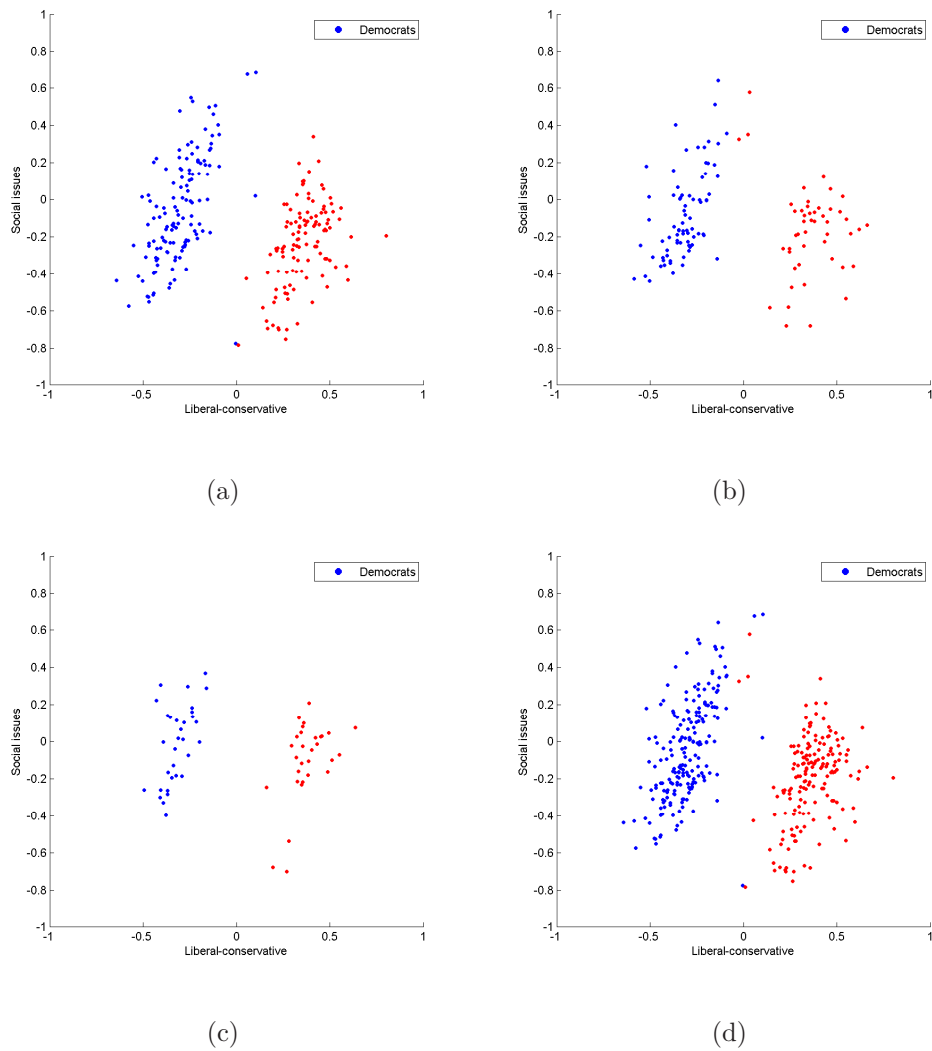
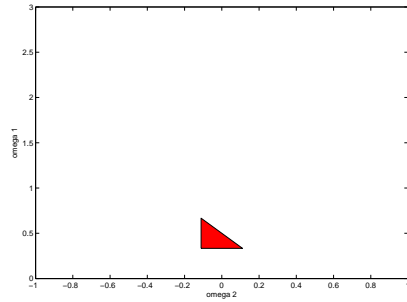
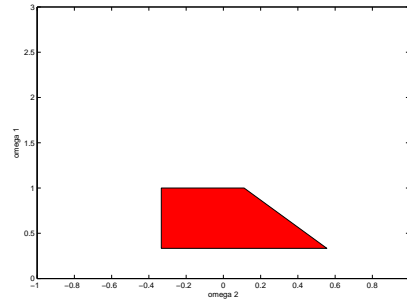


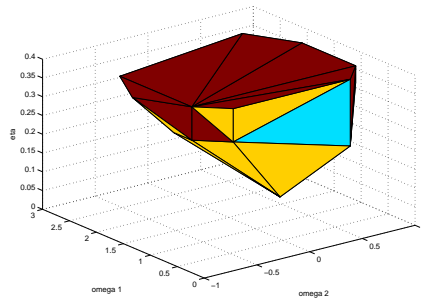
FIGURE 3. Ideological positions of candidates in the Pooler and Rosenthal NOMINATE Common Space. Republicans in red and democrats in blue. (a) 2000 election, (b) 2004 elections, (c) 2008 elections and (d) aggregate data.



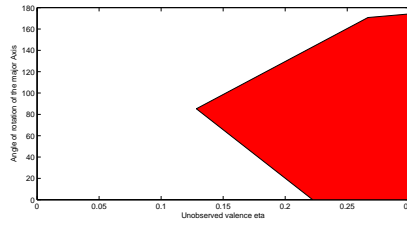
(a)



(b)



(c)



(d)

FIGURE 4. Confidence region in case of binary specification of valence term for aggregate data. (a) Section of the 95% level confidence region for  $(\omega, \eta)$  at  $\eta = 0.18$ ; (b) at  $\eta = 0.22$ . (c) Confidence region for  $(\eta, \omega)$  and (d) confidence region for  $(\eta, \phi)$ .