Multidimensional spatial voting with non-separable preferences

Lukas F. Stoetzer University of Mannheim [*]

Steffen Zittlau University of Mannheim [†]

[*]lstoetze@mail.uni-mannheim.de

[†]zittlau@uni-mannheim.de

A B S T R A C T

In most multidimensional spatial models, the systematic component of agent utility functions is specified as additive separable. We argue that this assumption is too restrictive, at least in the context of spatial voting in mass elections. Here, assuming separability would stipulate that voters do not care about how policy platforms combine positions on multiple policy dimensions. We present a statistical implementation of Davis, Hinich and Ordeshook's (1970) Weighted Euclidean Distance model that allows for the estimation of the direction and magnitude of non-separability from vote choice data. We demonstrate in a Monte-Carlo experiment that conventional separable model specifications yield biased and/or unreliable estimates of the effect of policy distances on vote choice probabilities in the presence of non-separable preferences. In three empirical applications, we find voter preferences on economic and socio-cultural issues to be non-separable. If non-separability is unaccounted for, researchers run the risk of missing crucial parts of the story. The implications of our findings carry over to other fields of research: Checking for non-separability is an essential part of robustness testing in empirical applications of multidimensional spatial models.

I. Introduction

The spatial model of voting is *the* work-horse for theories and empirical models in many fields of political science research, such as the equilibrium analysis in mass elections (e.g., Schofield, 1978; McKelvey, 1986; Calvert, 1985; Lin et al., 1999), the estimation of legislators ideal points (e.g., Poole and Rosenthal, 1985; Clinton et al., 2004) and the study of voting behavior (e.g., Kedar, 2005; Dow and Endersby, 2004; Quinn et al., 1999; Alvarez and Nagler, 1998). Since Downs' (1957) seminal work, the theory has come a long way. Its generalization to the multidimensional policy space, the Weighted Euclidean Distance (WED) model (Davis et al., 1970; Enelow and Hinich, 1984; Hinich and Munger, 1997) forms the stable theoretical foundation upon which nearly all present variations, extensions and applications of multidimensional spatial voting rest.

While all these contributions have advanced spatial theory and methodology, we argue that an important concept that used to be an integral part of the multidimensional spatial model was somehow lost along the way: the idea that policy preferences on multiple dimensions may be non-separable. Non-separability means that utility derived from policy distance on one dimension is dependent on policy distances on other dimensions. Empirical as well as formal models commonly rely on *additive separable* specifications of the spatial utility function, which preclude this possibility as utility is only a function of the sum of dimension-specific policy distances. In separable specifications, the dimensions have "nothing to do with each other" (Ordeshook, 1986, 90). In this paper we make the case that we should not stick to model specifications that preclude non-separability a priori. "There is nothing perverse about this preference rule" (Hinich and Munger, 1997, 86), and there are good reasons why real-world voter utility functions may be non-separable (Milyo, 2000).

The original mathematical formulation of the WED model explicitly allows for the possibility of non-separability, which is modeled as the product of dimension-specific directed distances (Davis et al., 1970; Enelow and Hinich, 1984; Hinich and Munger, 1997). The textbook example for non-separability is a scenario of committee voting, where committee members vote sequentially on multiple issues (Hinich and Munger, 1997, 60). Here, preferences on one issue are conditional on the outcome of voting on another issue, if the two issues are non-separable. Non-separability has also been studied in similar contexts, such as legislative voting on multiple issues (Kadane, 1972; Kramer, 1972; Schwartz, 1977), voting in multiple simultaneous elections or referenda (Brams et al., 1997, 1998; Lacy and Niou, 2000), voting for multiple candidates (Cox, 1984; Benoit and Kornhauser, 1994; Lacy and Niou, 1998), in models of committee agenda control (Denzau and Mackay, 1981; Mackay and Weaver, 1981; Enelow and Hinich, 1984) and EU council bargaining, where actors' spending preferences are conditional on expected policy outcomes (Finke, 2009; Finke and Fleig, 2013). A major contribution is also Lacy's model of survey responses, which explains item instability and question order effects by the non-separability of the underlying policy preferences (Lacy, 2001a,b).

The concept of non-separability has not yet been applied to the logic of multidimensional spatial voting in mass elections, in which voters choose policy platforms by evaluating their policy positions on multiple relevant policy dimensions. In the context of mass elections, non-separability means that a voter's evaluation of a platform on one policy dimension is conditional upon the position of this platform on other policy dimensions. If voters have non-separable utility functions, they no longer only evaluate platforms by their multidimensional distance from their ideal point, but also take into consideration how platforms combine directed distances over dimensions. As we show, these combinations, which we call policy packages, then have distinctive properties that voters care about. We suggest and provide evidence that accounting for non-separability might in fact be essential to our understanding of political choice. As all real-world policy platforms *only* come as policy packages, packaging might matter to voters.

Using the generalized Weighted Euclidean Distance model as the starting point, we discuss the theoretical foundations and implications of non-separability in mass elections. We show how non-separability can be incorporated and estimated in standard discrete choice models. In a Monte Carlo experiment we study the statistical consequences if the separability assumption is violated. We find that separable specifications then yield biased and/or unreliable estimates. In three empirical applications to national and presidential elections in the Netherlands, the US and Germany we demonstrate that accounting for non-separability can lead to very different conclusions about the substantive role of policy preferences in explaining voting behavior. Lastly, we discuss how testing for non-separability should be an essential part of robustness testing in all empirical applications of multidimensional spatial models.

II. Non-separability in the Weighted Euclidean Distance Model

The canonical WED model (Davis et al., 1970; Enelow and Hinich, 1984; Hinich and Munger, 1997) explicitly allows for non-separable voter utility functions. The spatial loss function for voter i and policy platforms j in a d-dimensional policy space is

$$U_{ij} = -\sqrt{[\boldsymbol{p_j} - \boldsymbol{v_i}]^T \boldsymbol{A} [\boldsymbol{p_j} - \boldsymbol{v_i}]} \qquad (1)$$

where $\boldsymbol{v}$ is a coordinate vector of voter ideal point of length d, and $\boldsymbol{p}$ is a coordinate vector of policy platform positions of length d. $\boldsymbol{A}$ is a $d \times d$ weighting matrix.[1] Its diagonal entries are weights expressing the importance, or salience, voters attach to distances on the policy dimensions. Off-diagonal entries contain the separability terms. Preferences are separable iff $\boldsymbol{A}$ is a diagonal matrix, i.e., all off-diagonal entries are zero. $\boldsymbol{A}$ is subject to an important constraint: It is a symmetric positive definite matrix (Davis et al., 1970, 433).[2] Positive definiteness guarantees that the quadratic form $[\boldsymbol{p}-\boldsymbol{v}]^T\boldsymbol{A}[\boldsymbol{p}-\boldsymbol{v}]$ is positive for all $p_j - v_i \neq 0$. The substantive meaning of the symmetric property is that non-separability does not depend on which dimension is evaluated first. Positive non-separability parameters indicate a substitutional relationship between dimensional preferences, negative a complementary relationship.

Non-separability has far-reaching consequences for our understanding of spatial voting. Non-separability "requires that voters consider *all* issue positions before choosing *any*" (Hinich and Munger, 1997, 85). In effect, this means that voters evaluate policy packages and not the separate positions platforms take on each of the relevant policy dimensions. To illustrate the electoral consequences of non-separability, we confine our analysis to a two-dimensional policy space. Imagine the policy space to be defined by an economic left-right and a socio-cultural liberal-conservative di-

[1] $\boldsymbol{A}$ may be individual-specific or, as we assume here, homogeneous in the population. For a detailed discussion of the homogeneity assumption see Rivers (1988).

[2] A symmetric matrix is positive definite if all its eigenvalues are positive. A 2×2 matrix is positive definite if the product of the diagonal elements is larger than the product of the off-diagonal elements.

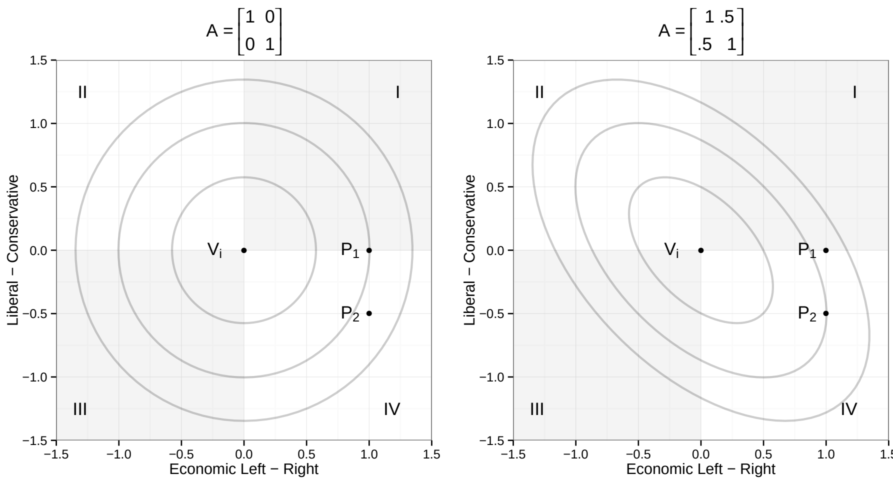

Figure 1: Exemplary choice scenario. Gray lines depict indifference contours. Left panel: Voter i with ideal point V and separable preferences prefers platform $P1$ over platform $P2$. Right panel: Voter i with non-separable preferences prefers $P2$ over $P1$.

mension. Dropping the matrix notion,

$$U_{ij} = -\sqrt{a_{11}[p_{j1} - v_{i1}]^2 + a_{22}[p_{j2} - v_{i2}]^2 + 2a_{12}[p_{j1} - v_{i1}][p_{j2} - v_{i2}]}, \qquad (2)$$

where a_{11} and a_{22} are the dimension-specific salience parameters and $2a_{12}$ are the symmetric separability parameters.[3] Figure 1 depicts an exemplary spatial configuration in a Cartesian coordinate system, in which combinations of more leftist and more socially conservative positions are found in quadrant II, more rightist and conservative combinations in I, and so on.

When voter policy preferences are separable (left panel of Figure 1), i will prefer platform 1 over platform 2, as P_1 is located on a higher utility curve than P_2. i's preference ordering over platforms is reversed in the non-separable case, depicted in the right-hand panel of Figure 1. Here, platform 2 is at a higher utility curve even

[3] Note that assuming squared Euclidean metric, and $a_{12} = 0$, yields the commonly used specification of the spatial model: $U_{ij} = -a_{11}[p_{j1} - v_{i1}]^2 - a_{22}[p_{j2} - v_{i2}]^2$.

though both platforms are equidistant on the economic dimension, and platform 1 is congruent with i's liberal-conservative ideal point. The reversal of voter i's preference order over platforms is not a result of differences in proximity or salience, but due to the fact that the policy package offered by platform 2 yields higher utility than the package offered by platform 1. This property is inherent to the policy package. Platform 2 simply yields higher utility because it combines dimensional distances in a way that conforms with the direction of the voter's non-separability terms.[4]

If spatial preferences are non-separable, policy packages hold properties of their own. All real-world policy platforms only come as policy packages, and packaging, the way in which policy platforms combine policies, might matter to voters. We think there is no good reason why it should not. Thus taking non-separability into account has the potential of offering a more realistic picture of spatial voting. Whether policy preferences are non-separable is, at this stage, an empirical question. We proceed to present a statistical WED model, which allows for the consistent estimation of the non-separability parameter.

III. A Conditional Logit Model with Non-Separable Preferences

McFadden's (1974) conditional logit is widely considered an appropriate discrete choice model to study spatial voting in multi-party systems (Alvarez and Nagler,

[4]The consequences of non-separability may also be analyzed as sequential voting over individual dimensions. Keeping a party's positions on the economic dimension fixed, the voter ideal point on the liberal-conservative dimension shifts if preferences are non-separable. The new *conditional* ideal point is $v_2^*(p_1)$ is $v_2 - \frac{a_{12}}{a_{22}}(p_1 - v_1)$ (Enelow and Hinich, 1984). Although conditional ideal points shift, this does not mean that *unconditional* voter ideal points are no longer fixed. Only the context changes. Voters still have an ideal package - their unconditional ideal point, but "there is no 'best' unique issue-by-issue ideal point" (Hinich and Munger, 1997, 61).

1998; Dow and Endersby, 2004).[5] Taking the conditional logit as our starting point, we propose a non-separable specification of the systematic component that follows from the WED model, and account for the positive definite constraint over $\boldsymbol{A}$. Unlike conventional specifications, which specify $\boldsymbol{A}$ as diagonal, we specify $\boldsymbol{A}$ to be symmetric and positive definite.

In the conditional logit, choice probabilities for voters $i \in (1, \dots, n)$ choosing between policy platforms $j \in (1, \dots, k)$ take the form

$$P_{ij} = \frac{e^{V_{ij}}}{\sum_{j=1}^{k} e^{V_{ij}}}. \tag{3}$$

V_{ij} is the systematic component of the voter utility function, which we specify as

$$V_{ij} = \theta_j + \boldsymbol{X_i \delta_j} - \sqrt{[\boldsymbol{p}_j - \boldsymbol{v}_i]^T \boldsymbol{A} [\boldsymbol{p}_j - \boldsymbol{v}_i]}.^6 \tag{4}$$

θ_j is a platform-specific constant that captures non-policy aspects, oftentimes labelled party or candidate valence. $\boldsymbol{X_i}\delta_j$ captures the effect of non-spatial individual-specific covariates on choice probabilities. The negative square root is the multidimensional spatial voting part as conceptualized in the WED model. $\boldsymbol{A}$ is a symmetric positive definite matrix. To incorporate this constraint in the maximum likelihood framework, we re-parameterize $\boldsymbol{A}$ as its Cholesky decomposition. This is a common procedure to solve numerically difficult optimization problems, such as the estima-

[5] Conditional logit, like multinomial logit, assumes the random error to be independently and identically distributed Type-1 extreme value. An undesirable feature of conditional logit is its reliance on the independence of irrelevant alternatives (IIA) (for a detailed discussion see e.g. Rivers, 1988; Alvarez and Nagler, 1998; Dow and Endersby, 2004). Multinomial probit has been considered as a solution. Specifying the systematic component in multinomial probit models as non-separable works the same way. Nevertheless, we opt for conditional logit because of its continuing popularity and since its computational convenience facilitates our Monte Carlo experiments.

[6] δ_j and θ_j are choice-specific parameters, while $\boldsymbol{A}$ is assumed to be homogeneous over choices and individuals.

tion of variance-covariance matrices (Pinheiro and Bates, 1996). $\boldsymbol{A}$ is parameterized as a lower triangular matrix $\boldsymbol{L}$, with $\boldsymbol{A} = \boldsymbol{L}^T\boldsymbol{L}$.

For a 2×2 $\boldsymbol{A}$ matrix, $\boldsymbol{L}$ contains three parameters.

$$\boldsymbol{L} = \begin{bmatrix} l_1 & 0 \\ l_{12} & l_2 \end{bmatrix} \tag{5}$$

$\boldsymbol{A}$ is restored post-estimation after maximizing likelihood with respect to $\boldsymbol{L}, \theta_k, \delta_k$.[7]

IV. The consequences of misspecification: A Monte Carlo Experiment

Failing to account for non-separability if it is part of the true data generating process constitutes a misspecification of the functional form of how voter and party platform positions enter into the utility function. We study the consequences of misspecification using Monte Carlo methods. In order to obtain conservative estimates and to facilitate interpretation, we opt for a very basic design: Political choice in a policy space with two equally salient orthogonal policy dimensions.[8]

In separable specifications (6a), dimension-specific policy distances enter utility additively, in the non-separable specification (6b) they enter utility additively *and*

[7] The likelihood function is given by the product over all realized probabilities. In order to identify this model, θ_k and δ_k are set to zero, for a baseline platform $j = k$. We use Broyden-Fletcher-Goldfarb-Shanno (BFGS) iterative numerical algorithm to maximize log-likelihood directly, using R's optim() function. In order to assure convergence on global maxima, maximization is repeated multiple times using randomly drawn starting values.

[8] There is no indication that the implications of our findings do not apply to higher-dimensional spaces as well. We suspect that the consequences of misspecification may become more pronounced as the number of dimensions increases and with larger differences in dimensional salience weights.

multiplicatively .

$$U(v,p) = -\sqrt{a_{11}[p_1 - v_1]^2 + a_{22}[p_2 - v_2]^2} \tag{6a}$$

$$U^*(v,p) = -\sqrt{a_{11}[p_1 - v_1]^2 + a_{22}[p_2 - v_2]^2 + 2a_{12}[p_1 - v_1][p_2 - v_2]}. \tag{6b}$$

If the data generating process follows (6b), a conventional model (6a) is misspecified. Misspecification in choice models can result in biases that are analogous to omitting important variables (Signorino and Yilmaz, 2003). Omitting non-separability terms when non-separability is present has the potential of leading to biased estimates of the salience parameters a_{11} and a_{22}, which express the importance of policy dimensions in the voters' choice calculus - the parameters of interest.

Under which circumstances are conventional estimates biased, in which direction and how severely? We show that the magnitude and direction of bias depends on the magnitude and direction of the non-separability parameter, and the distribution of platform positions in the policy space, relative to the distribution of voter ideal points. In order to test our intuition, we analyze the conditions under which the two expressions are not equivalent in expectation, i.e., $E[U(v,p)] \neq E[U^*(v,p)]$. If non-separability plays a role in the true data generating process ($a_{12} \neq 0$), the expressions are equivalent if $E([p_1 - v_1][p_2 - v_2]) = 0$. Without loss of generality, assume $E(v_{i1}) = E(v_{i2}) = 0$, which would be for example the case if voter ideal points are distributed independently multi-variate normal around the origin of a Cartesian coordinate system.[9] At this point, let us recall the properties of the variables v and p. While voter ideal points vary between voters, platform positions are fixed in a

[9] Voter ideal points are also assumed to be uncorrelated over dimensions and platform positions on one dimension are independent of voter ideal points on the second dimension.

given sample. With $E(v)$ at $[0,0]$, $E([p_1 - v_1][p_2 - v_2])$ can become either negative or positive, depending on p's position relative to $[0,0]$. If $a_{12} > 0$, the multiplicative non-separability term is positive if the platform is in quadrant I or III, negative in quadrant II and IV, and vice versa if $a_{12} < 0$. The omitted term can therefore enter utility positively or negatively. As voters choose between multiple platforms, one has to consider the direction the omitted term has in expectation, over all platforms in the choice set. This is determined by the directed distance of platform positions relative to the expected voter ideal point. If platforms are positioned in a systematic way in the policy space relative to the expected voter ideal point, utility derived from policy distances under non-separability and separability rule is systematically different, if $a_{12} \neq 0$. The pattern of platform positions can be summarized by their correlation coefficient on the two dimensions. A positive correlation would indicate that positions along the first angle bisector of the Cartesian coordinate system are more likely, and negative along the second angle bisector. If platform positions are uncorrelated, positive and negative omitted terms cancel each other out in expectation. While conventional estimates would still be unbiased in expectation, one should expect that non-separability in this case increases the variance of the sampling distribution.

This rather intuitive analysis of the implications of violating the separability assumption motivates our design of a Monte Carlo experiment.[10] We study the conse-

[10] An analytical solution is not easily tractable for the outlined choice model. Studies that are concerned with specification in choice models find omitted variable bias to be a more challenging problem than in the linear case (Yatchew and Griliches, 1985; Wooldridge, 2002). In probit models the estimates of a coefficient are generally biased downwards even if omitted variables are not correlated with other variables (Cramer, 2005). However, results from binary probit models do not straightforwardly carry over to unordered choice models. Lee (1980) explicitly studies omitted variable biases in the multinomial-logit context. His results of direction and strength of the bias are restricted to the case where omitted variables can be expressed as a linear function of other covariates with normal error. This is not the case here, since the omitted non-separability term can be expressed as a function of the distance terms. Moreover, the omitted variable bias is more complicated for the conditional logit model compared to the multinomial logit model.

quences of violating the separability assumption for the unbiasedness and sampling variance of the parameters of interest, the dimension-specific saliences, in three scenarios in which we vary the distribution of platforms in the policy space. In the first scenario, platform positions on the two dimensions are correlated. In the second scenario, platform positions on the two dimensions are uncorrelated, meaning that platforms are scattered unsystematically in the policy space. Platform positions are again heavily correlated in our third scenario, this time however negatively. For each of these three scenarios we draw 1000 voter ideal points and four party positions. Platform positions are drawn from a bivariate normal with variance terms of 0.5 and scenario-specific covariance terms.[11] Voter ideal points are drawn from a bivariate standard normal distribution. Voters choose between platforms according to the non-separable specification (Eq. 6b), assuming constant equal weights to both dimensions ($a_{11} = a_{22} = 1$) and varying degrees of separability. For each scenario, we vary the separability parameter a_{12} in 11 steps over the interval that meets the positive-definite constraint, $[-1, 1]$. 500 random samples for each combination of platform scenario and value of the non-separability parameter are drawn. For each of the resulting 16,500 unique datasets, we estimate a conventionally-specified separable (Eq. 6a) as well as a fully-specified non-separable model (Eq. 6b).

For each specific subscenario, we approximate the sampling distribution of the salience parameters by their empirical distribution in the 500 Monte Carlo samples. As we expect misspecification to render salience estimates inconsistent and/or inefficient, we report both the mean and 90% range of the bias in salience parameter estimates (Figure 2). The upper horizontal panel displays the additive separable model estimates, the lower horizontal panel the estimates obtained from the

[11] Correlations are set to .8, 0, and negative .8 respectively.

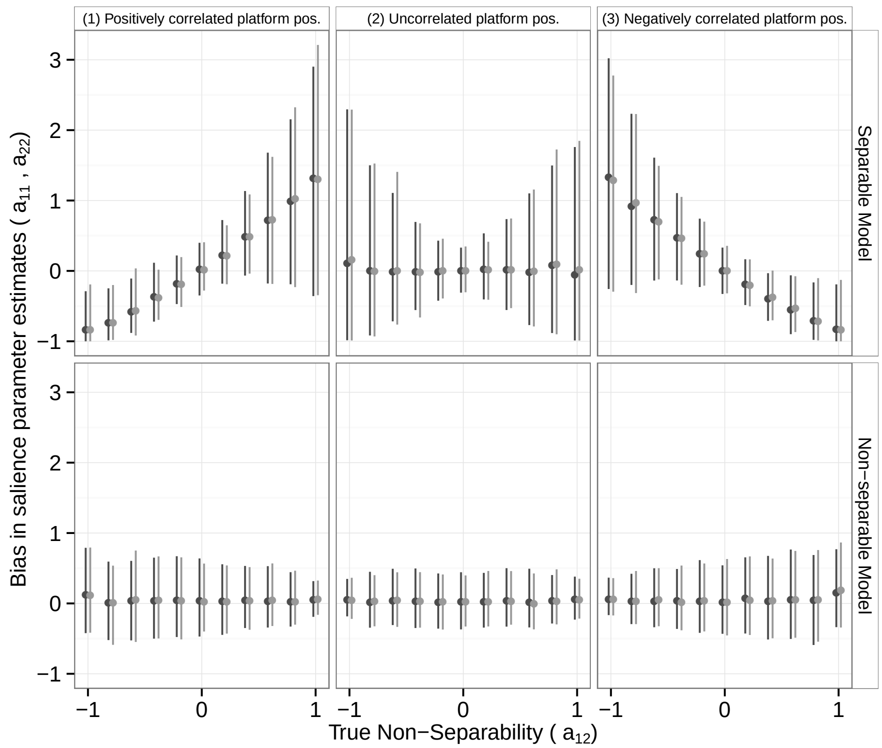

Figure 2: Monte Carlo Experiment: Bias in salience parameter estimates due to omitted non-separability. True salience parameters are both set at 1. Dots indicate mean estimates of salience parameters a_{11} (dark gray) and a_{22}(light gray). Vertical bars depict the 90% range of all estimates from Monte Carlo samples. (1) Platform positions randomly drawn from bivariate normal with positive covariance terms ($\rho = .8$), (2) Platform positions randomly drawn from bivariate normal with zero covariance terms ($\rho = 0$), (3) Platform positions randomly drawn from bivariate normal with negative covariance terms ($\rho = -.8$). Voter ideal points are drawn from a bivariate standard normal distribution. The non-separability parameter a_{12} is consistently estimated by the fully-specified model in all scenarios.

non-separable model. Vertical panels indicate the three main scenarios, in which platform positions were either correlated, not correlated, or negatively correlated. The upper horizontal panel shows that the misspecified model yields either biased and/or more unreliable estimates depending on the distribution of policy platforms in the policy space. If platform positions are positively correlated, and dimensional distances are complements ($a_{12} < 0$), the misspecified models underestimates the salience of both dimensions. If dimensions are substitutes ($a_{12} > 0$), their salience is considerably overestimated. The bias is not negligible: Even with moderately negative non-separability ($a_{12} = -0.6$), the salience parameters (a_{11} and a_{22}) are estimated at only around 50% of their true value, and are inflated by around 80% in the presence of moderately positive non-separability parameters (0.6). In case platform positions are negatively correlated, the direction of the bias is reversed, and positive non-separability parameters lead to a downward bias and negative to an upward bias. There is no theoretical bias if platform positions are uncorrelated, for all values of non-separability. However, the sampling variance increases considerably, as non-separability increases. This renders conventional estimates unreliable. As the lower horizontal panel indicates, a fully-specified model reliably recovers the true salience parameters in all scenarios.[12]. The non-separability parameter a_{12} is consistently estimated in all scenarios by the fully-specified model (see Supplementary Materials)

The message of our Monte Carlo experiment is clear: In the presence of non-separability, the statistical properties of a non-separable model are preferable to conventional, separable specifications. Dependent on party positions, a separable salience estimator is inconsistent and/or inefficient. The size and direction of the bias is dependent on the relative distribution of voters and policy platforms in the policy space. Even in a most basic case of two equally salient dimensions we find that bias can become severe and is not easily tractable. Given real-world data, it is therefore hardly ever apparent whether conventional models will run into problems and if they do, how severe these are.

In light of these insights, it is advisable to test the robustness of conventional mod-

[12]For a comparison of Root Mean Square Error (RMSE) and correctly predicted cases see the Supplementary Materials, which are available at http://dx.doi.org/10.7910/DVN/VCSRMX.

els to potentially omitted non-separability. To demonstrate that non-separability is relevant not only in highly stylized Monte Carlo experiments but in explaining empirical phenomena, we proceed to three empirical applications in which we compare estimates obtained by separable and non-separable specifications.

V. Empirical Applications

Empirical studies of spatial voting in mass elections ultimately rely on estimates of voter ideal points and party platform positions. Voter ideal points are commonly inferred from voter surveys, which ask respondents to locate themselves on various policy or issue scales. Platform positions are inferred from either where respondents place policy platforms on these policy scales, or from outside sources such as expert surveys, roll call votes or analyses of platform manifestos. Over the years, a multitude of approaches has developed, each addressing some of the difficulties of estimating reliable ideal point and platform position estimates (e.g. Aldrich and McKelvey, 1977; Poole and Rosenthal, 1985; Kedar, 2005; Bartels, 2006; Jessee, 2009; Lo et al., 2013).

For our purposes we deem factor-analytic techniques which have been employed by Quinn et al. (1999), Schofield et al. (1998) and Schofield and Zakharov (2009) as most appropriate. Using expert and voter survey responses on multiple concrete issue scales, this approach allows for the placement of voters and platforms in a common multidimensional Euclidean policy space.[13] In effect, factor analytic methods approximate the structure of the policy space by analyzing the structure of voter survey responses. Policy dimensions are not defined a priori, but are rather uncovered using empirical data. The procedure can be described as follows: First, survey re-

[13]We closely follow the procedure outlined in Quinn et al. (1999), we kindly ask the reader to refer to this article for details.

sponses on a number of issue items are used to estimate an explanatory factor model. In these issue items respondents are asked to what degree they agree or disagree with specific policy statements, such as "Politics should abstain from intervening in the economy.". From the factor analytic solution, the number of underlying dimensions is determined that sufficiently describes the policy space. In all cases, we find a two-factor model to be most appropriate. We identify the first factor as an "economic left-right" dimension, and the second as capturing socio-cultural preferences, which we label as the "liberal-conservative" dimension.[14] Secondly, we conduct a two-factor confirmatory factor analysis (CFA) with uncorrelated factors. Using the CFA factor loadings we locate respondents in the policy space. In the third step policy platforms are projected into the same space, using the factor loading from the voter CFA. In two of our applications, positions of policy platforms on these issue items were not available. Here scales in expert surveys are identified that thematically match the policy scales of survey items as closely as possible. This necessarily involves a substantial degree of discretion, as survey items and expert scales are not identical (see Supplementary Materials).

We apply this method to three different empirical applications. The first uses the Euro-Barometer 11 dataset (Commission of the European Communities, 2012) on vote intention in 1979 in the Netherlands employed by Quinn et al. (1999).[15] The

[14]While this finding is hardly controversial for the Netherlands and Germany, where the two-dimensionality of the policy space is well established (see e.g. Benoit and Laver, 2006; Schofield et al., 1998), it may raise some eyebrows in the US application. Here the working consensus seems to be that inter-party political conflict is virtually unidimensional (McCarty et al., 2006; Poole and Rosenthal, 2007; Aldrich et al., 2014). However, unidimensionality of political conflict among political elites does not imply that voter policy preferences are unidimensional as well. Work on the structure of political ideology among Americans finds, very much in line with our findings, that voter preferences are structured by an economic and a social policy dimension (Shafer and Claggett, 1995; Treier and Hillygus, 2009; Klar, 2014).

[15]Replication data is well-documented and available online at `http://adm.wustl.edu/replication.php`. The model estimated in the original article however does not allow, unlike our WED model, for dimension-specific weights. Our results are therefore not readily comparable.

second application is concerned with voting in the 2008 US presidential election. Voter data is from the 10th wave of the American National Election Panel Study 2008-2009 (American National Election Studies, 2009), in which respondents were asked a battery of eight policy issue questions ranging from immigration to health care to taxation. Respondents were asked to locate themselves and the Democratic and Republican candidate on each of these 7-point scales, ranging from "strongly opposed" to "strongly in favor." After projecting both voter and candidate placements into the policy space, we average over the candidate positions to obtain a robust measure of candidate positions. The third application analyzes vote intention in the 2009 German federal election. Voter data is from the European Election Survey 2009 (EES) (van Egmond et al., 2013), which includes seven issue scales capturing attitudes towards immigration, extent of public services, state intervention in the economy, redistribution of wealth, criminal punishment and homosexuality. In order to locate party platforms on these scales, we identify seven issue scales in the Chapel-Hill Expert Survey 2011 (CHES) that match the EES scales. For a more detailed description of the question wording, highest density plot and factor loadings, see the Supplementary Materials.

In accordance with standard model specifications (see e.g. Dow and Endersby, 2004; Kedar, 2005; Quinn et al., 1999), we include individual-specific control variables such as gender, age, education, religion, income or party identification. For each application we specify two vote choice models: a normal WED model that allows for dimension-specific weights, but assumes separability, and a non-separable WED model that allows for non-separability. We estimate the two models according to the conditional logit specification outlined above.

VI. Results

Table 1 compares the spatial parameter estimates obtained from the separable and non-separable model specifications. We report estimated saliences of distances on the economic left-right dimension and on the liberal-conservative dimension, and the estimated separability parameter for the non-separable models.[16] As salience parameters are constrained to be positive, 95% confidence intervals are used to quantify estimation uncertainty.[17] How severe non-separability is in an estimated $\boldsymbol{A}$ matrix is not immediately obvious. We therefore report an intuitive measure of the degree of non-separability, that sufficiently summarizes both the direction and the degree of non-separability in two-dimensional policy spaces. This measure utilizes the positive-definite constraint to scale the separability parameter to the interval $[-1, 1]$, where -1 indicates perfect complements, and 1 perfect substitutes.[18]

Does modeling non-separability make a difference in the three applications we present here? If it does not, both model specifications should yield similar salience estimates. Salience estimates should not only have about the same magnitude, but should also not vary in their relative magnitude. Model fit is a second criterion. If non-separability is not an issue, the non-separable model should not exhibit a better model fit. As the models are nested, Likelihood Ratio tests are appropriate, which we report in the last row of Table 1. We also report expected Percentage Correctly Pre-

[16] The paramter estimates of the individual-specific control variables are reported in the Supplementary Materials.

[17] Confidence intervals may be non-symmetrical due to the constraint induced by the Cholesky decomposition. The separability parameter is not subject to the positive constraint, is however constrained by the positive-definiteness of the $\boldsymbol{A}$ matrix.

[18] As $\boldsymbol{A}$ is a symmetric positive definite 2×2 matrix, the condition $a_{11} \cdot a_{22} - a_{12}^2 \geq 0$ holds. By rearranging we see that a_{12} is bounded between $\pm\sqrt{a_{11} \cdot a_{22}}$. Therefore $\frac{a_{12}}{\sqrt{a_{11} \cdot a_{22}}}$ is bounded between $[-1, 1]$. In order to convey the estimation uncertainty associated with the measure, we approximate 95% confidence intervals by calculating the degree of separability for repeated draws from the sampling distribution of $\boldsymbol{L}$.

Table 1: Empirical applications: Parameter estimates

DV: Vote choice	Netherlands 1979		United States 2008		Germany 2009	
	Sep.	Non-sep.	Sep.	Non-sep.	Sep.	Non-sep.
Economic Left-Right (a_{11})	0.73	0.92	0.81	0.44	0.8	0.74
	(0.43; 1.15)	(0.56; 1.38)	(0.42; 1.35)	(0.21; 1.07)	(0.38; 1.35)	(0.35; 1.46)
Liberal-Conservative (a_{22})	0	0.38	1.77	0.98	0.14	0.18
	(0; 0.2)	(0.19; 0.63)	(1.13; 2.56)	(0.43; 1.79)	(0.02; 0.41)	(0.04; 0.44)
Separability term (a_{12})	0	0.59	0	0.44	0	0.17
	.	(0.33; 0.93)	.	(0.19; 0.64)	.	(0.02; 0.38)
Degree of Separability	0	1	0	0.67	0	0.49
	.	(0.99; 1)	.	(0.19; 1)	.	(0.06; 0.85)
Number of Observations	529	529	1133	1133	619	619
ePCP	0.49	0.5	0.75	0.76	0.6	0.6
logLikelihood	−498	−483.5	−514.1	−510.9	−512.6	−510.6
Likelihood Ratio Test	6.7e-08		1.1e-02		0.04	

Note: Table reports point estimates and 95% confidence intervals in parentheses.

dicted (ePCP) (Herron, 1999) as an additional measure of predictive fit. In the first application, concerned with vote intention in 1979 in the Netherlands, the separable model suggests that only distances on the economical left-right dimension are relevant for vote choice. The coefficient of the second, liberal-conservative dimension is virtually zero, from which we would conclude that policy distance on this dimension is not associated with vote choice probabilities. This interpretation changes when we run a model that accounts for non-separability. Here we find distances on the second dimension to play a still subordinate, but noticeable role in explaining vote choices. The separability parameter estimate is significantly larger than zero. As the measure of the degree of separability indicates, policy distances on the two dimensions are estimated to be perfect substitutes. The Likelihood Ratio test indicates that the non-separable model fits the data considerably better. ePCP reveals a small increase in predictive accuracy when comparing the two models. In the application to voting in the US presidential election 2008, the separable model suggests that policy

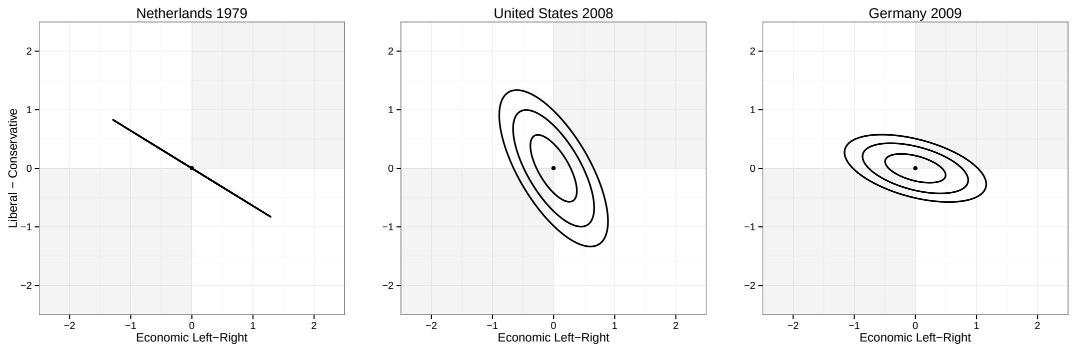

Figure 3: Indifference contours for A matrix point estimates.

distance on the socio-cultural liberal-conservative dimension is more important to voters than distance on the economic dimension (95% C.I. for the difference in a_{11} and a_{22}: $[-1.69, -0.37]$). In contrast, the difference in coefficients becomes indistinguishable from zero in the non-separable model (95% C.I. $[-1.07, 0.09]$), indicating that policy distances on both dimensions play an equally important role in the vote choice mechanism. Again, the separability term is significantly larger than zero, as policy distances on the two dimensions are estimated to be partial substitutes. The degree to which preferences are substitutes is associated with substantial estimation uncertainty [(0.19, 1]. The Likelihood Ratio test indicates that the non-separable model fits the data better, ePCP shows a small increase in predictive accuracy. In the German case, non-separability seems to be less of an issue. Salience estimates are relatively robust to changes in the specification of the spatial model. The separability term has again a positive sign, but is only barely distinguishable from zero. The non-separable model fits the data only slightly better.

Our findings indicate that non-separability makes a substantive difference for our understanding of voting behavior. In two out of three applications, allowing for non-separability has led to different conclusions about the role of dimensional distances in the voters' choice rationale. We find that in the Netherlands case, policy distances on the two dimensions are actually functional equivalents for voters. Figure 4 displays the shape of the indifference contours for point estimates of the spatial parameters. In the Netherlands case, where dimensional distances are perfect substitutes, the indifference ellipsoid practically collapses into a line: Even though voters hold meaningful preferences on two dimensions, preferences are translated into vote choices using a one-dimensional concept. Effectively, non-separability leads to a reduction in the dimensional complexity of political choice. In the application to voting in US presidential elections, separable models would have found that voters assign a higher weight to preferences on the liberal-conservative dimension than to the economic dimension. Preferences on the two dimensions have about the same importance to American voters, if we allow for non-separability. Although the degree of non-separability varies considerably, our results are remarkably consistent. Preferences on the two dimensions under investigation are substitutes in all three electoral contexts we have analyzed. A voter located at the center of the policy space, choosing among two parties which take more rightist position than the voter on the economic left-right dimension prefers the party that is more liberal on the cultural dimension to one that holds his preferred position on the cultural dimension. This is because the more liberal party compensates distance on the economic dimension by distance in the opposite direction on the cultural dimension.

VII. Conclusion

In this paper we have advocated to bring back a concept that once used to be an integral part of the multidimensional spatial framework: Multidimensional spatial preferences may not be independent of each other, they may be non-separable. In mass elections, non-separability means that voters do not separately evaluate the positions policy platforms take on multiple dimensions. They rather care about the policy packages that platforms offer. If the way in which platforms combine positions matters to voters, we need to allow for non-separability in our models. We present a consistent way of including non-separability in vote choice models. We find that failing to allow for non-separability can seriously undermine the validity of empirical tests of spatial theory. Our Monte Carlo experiments show that conventional salience estimates are biased and/or unreliable in the presence of non-separability. The magnitude and direction of bias depends on a non-trivial interaction between the degree and direction of non-separability and the distribution of policy platforms relative to voter ideal points in the policy space. If faced with real-world data, it is therefore not apparent whether omitted non-separability might be problematic in a statistical sense. Even more so when voter preferences are not limited to a low number of latent policy dimensions, but are defined over a high-dimensional issue space, where we might reasonably suspect non-separability to be the rule rather than the exception. Thus, to be on the safe side, careful researchers should test the robustness of obtained estimates with non-separable model specifications. Researchers who want to rely on linear predictor functions fit for out-of-the-can statistical programs can accommodate for non-separability by adding the products of all combinations of *directed* dimension-specific distances to the systematic component. In a two-dimensional policy space and using squared Euclidean metric, voter utility can be specified as $U_{ij} = \beta_1(p_{j1} - v_{i1})^2 + \beta_2(p_{j2} - v_{i2})^2 + \beta_3(p_{j1} - v_{i1}) \times (p_{j2} - v_{i2})$. β_3 then can be interpreted as the non-separability parameter. If $\hat{\beta}_3$ is significantly different from zero, non-separability is an issue.

On a more positive note, addressing the issue of non-separability can help reduce bias and/or increase the precision of spatial estimates. Our findings are therefore potentially relevant for all empirical applications of multidimensional spatial voting models that base their inferences on spatial salience estimates. Most prominently in studies of voting behavior that compare the importance voters assign to various issues or dimensions, such as the question whether economic issues trump "moral" issues or vice-versa in U.S. presidential elections (Bartels, 2006; Gelman, 2008), how party system compactness relates to relative issue importance (Alvarez and Nagler, 2004), or which role attitudes towards Europe play in explaining electoral behavior in European elections, relative to left-right preferences (De Vries et al., 2011; Hobolt et al., 2009; Lo et al., 2013). Here, the validity of spatial estimates, and the conclusion that one dimension is relevant, not relevant, or more relevant for political choice than other dimensions may depend on whether voter utility functions are specified as separable or non-separable.

Apart from being a safeguard against statistical pitfalls, caring about non-separability opens up new interesting perspectives on the structure of voter preferences in the multidimensional policy space. Analogous to a resource or budget constraint (Milyo, 2000), non-separability in mass elections can be imagined as an *ideological* constraint (Converse, 1964). The constraint determines which policy packages are more attractive to voters. In all of our empirical applications we find a substitutional relationship between economic and social policy preferences. Such a relationship would indicate that the two policy dimensions share at least to some degree the same function and fulfill the same voter needs. If preferences are perfect substitutes, it becomes hard to argue that voters really care about individual policy dimensions. Although voters have well-defined preferences on these policy dimensions, what they really care about when choosing representatives is a lower-dimensional concept such as a single ideological dimension. In such a case, non-separability leads to a reduction in the effective dimensionality of the policy space, linking multidimensional mass preferences to unidimensional inter-party competition at the elite level.

Non-separability therefore not only informs the study of voting behavior but also the formal analysis of party competition. How do equilibrium configurations change if parties maneuver in a multidimensional policy space with non-separable voter preferences (Merrill and Adams, 2001; Schofield and Sened, 2005), and can non-separability explain the empirical phenomenon that in many countries parties align along a single axis of competition even though the policy space is two-dimensional (Shikano, 2008)? Non-separability can also be brought to bear on the empirical question whether political polarization in the US has increased in recent decades (Fiorina et al., 2008; Levendusky, 2009; Aldrich et al., 2014). Polarization may express itself not only in changes in voter preferences, issue partisanship, or issue alignment, but also in increasing non-separability of voter utility functions.

The implications of non-separability are multifaceted. Additional research is required to deepen our understanding of non-separability, and to thereby deepen our understanding of spatial voting in multidimensional spaces. To be sure, we do not suggest that multidimensional representations of voter preferences are generally preferable to one-dimensional representations. But we argue that if empirical researchers opt for multidimensional spatial representations, the potential non-separability of spatial preferences needs to be addressed. We hope that the findings and methods presented in this paper can serve as a guideline for future research to bring non-separability back into the fold of spatial theory.

References

Aldrich, J. H. and R. D. McKelvey (1977). A Method of Scaling with Applications to the 1968 and 1972 Presidential Elections. *The American Political Science Review 71*(1), 111–130.

Aldrich, J. H., J. M. Montgomery, and D. B. Sparks (2014). Polarization and Ideology: Partisan Sources of Low Dimensionality in Scaled Roll Call Analyses. *Political Analysis 22*(4), 435–456.

Alvarez, R. M. and J. Nagler (1998). When Politics and Models Collide: Estimating Models of Multiparty Elections. *American Journal of Political Science 42*(1), 55–96.

Alvarez, R. M. and J. Nagler (2004). Party system compactness: Measurement and consequences. *Political Analysis 12*(1), 46–62.

American National Election Studies (2009). American National Election Studies (ANES) Panel Study, 2008-2009. ICPSR29182-v1. http://doi.org/10.3886/ICPSR29182.v1.

Bartels, L. M. (2006). What's the Matter with What's the Matter with Kansas? *Quarterly Journal of Political Science 1*(2), 201–226.

Benoit, J.-P. and L. A. Kornhauser (1994). Social Choice in a Representative Democracy. *The American Political Science Review 88*(1), 185–192.

Benoit, K. and M. Laver (2006). *Party Policy in Modern Democracies*. London: Routledge.

Brams, S. J., D. M. Kilgour, and W. S. Zwicker (1997). Voting on referenda: the separability problem and possible solutions. *Electoral Studies 16*(3), 359–377.

Brams, S. J., D. M. Kilgour, and W. S. Zwicker (1998). The paradox of multiple elections. *Social Choice and Welfare 15*(2), 211–236.

Calvert, R. L. (1985). Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence. *American Journal of Political Science 29*(1), 69–95.

Clinton, J., S. Jackman, and D. Rivers (2004). The Statistical Analysis of Roll Call Data. *American Political Science Review 98*(2), 1–16.

Commission of the European Communities (2012). Eurobarometer 11 (Apr 1979). http://dx.doi.org/10.4232/1.10866.

Converse, P. E. (1964). The Nature of Belief Systems in Mass Publics. In D. Apter (Ed.), *Ideology and Discontent*, pp. 206–261. New York: Free Press.

Cox, G. W. (1984). Electoral equilibrium in double member districts. *Public Choice 44*(3), 443–451.

Cramer, J. S. (2005). Omitted Variables and Misspecified Disturbances in the Logit Model.

Davis, O. A., M. J. Hinich, and P. C. Ordeshook (1970). An Expository Development of a Mathematical Model of the Electoral Process. *American Political Science Review 64*(2), 426–448.

De Vries, C. E., W. van der Brug, M. H. van Egmond, and C. van der Eijk (2011). Individual and contextual variation in EU issue voting: The role of political information. *Electoral Studies 30*(1), 16–28.

Denzau, A. T. and R. J. Mackay (1981). Structure-Induced Equilibria and Perfect-Foresight Expectations. *American Journal of Political Science 25*(4), 762–779.

Dow, J. K. and J. W. Endersby (2004). Multinomial probit and multinomial logit: a comparison of choice models for voting research. *Electoral Studies 23*(1), 107–122.

Enelow, J. M. and M. J. Hinich (1984). *The Spatial Theory of Voting: An Introduction*. New York: Cambridge University Press.

Finke, D. (2009). Estimating the Effect of Nonseparable Preferences in Eu Treaty Negotiations. *Journal of Theoretical Politics 21*(4), 543–569.

Finke, D. and A. Fleig (2013). The merits of adding complexity: non-separable preferences in spatial models of European Union politics. *Journal of Theoretical Politics 25*(4), 546–575.

Fiorina, M. P., S. A. Abrams, and J. C. Pope (2008). Polarization in the American Public: Misconceptions and Misreadings. *The Journal of Politics 70*(2), 556–560.

Gelman, A. (2008). *Red state, blue state, rich state, poor state : why Americans vote the way they do*. Princeton: Princeton Univ. Press.

Herron, M. (1999). Postestimation Uncertainty in Limited Dependent Variables Models. *Political Analysis 8*(1), 83–98.

Hinich, M. J. and M. C. Munger (1997). *Analytical Politics*. Cambridge, New York, Melbourne: Cambridge Univ. Press.

Hobolt, S. B., J.-J. Spoon, and J. Tilley (2009). A Vote Against Europe? Explaining Defection at the 1999 and 2004 European Parliament Elections. *British Journal of Political Science 39*(01), 93–115.

Jessee, S. A. (2009). Spatial voting in the 2004 presidential election. *American Political Science Review 103*(01), 59–81.

Kadane, J. B. (1972). On division of the question. *Public Choice 13*(1), 47–54.

Kedar, O. (2005). When Moderate Voters Prefer Extreme Parties: Policy Balancingin Parliamentary Elections. *American Political Science Review 99*(2), 185–199.

Klar, S. (2014). A Multidimensional Study of Ideological Preferences and Priorities among the American Public. *Public Opinion Quarterly 78*, 344–359.

Kramer, G. H. (1972). Sophisticated voting over multidimensional choice spaces. *The Journal of Mathematical Sociology 2*(2), 165–180.

Lacy, D. (2001a). A Theory of Nonseparable Preferences in Survey Responses. *American Journal of Political Science 45*(2), 239–258.

Lacy, D. (2001b). Nonseparable Preferences, Measurement Error, and Unstable Survey Responses. *Political Analysis 9*(2), 95–115.

Lacy, D. and E. M. Niou (2000). A Problem with Referendums. *Journal of Theoretical Politics 12*(1), 5–31.

Lacy, D. and E. M. S. Niou (1998). Elections in Double-Member Districts with Nonseparable Voter Preferences. *Journal of Theoretical Politics 10*(1), 89–110.

Lee, L.-f. (1980). Specification Error in Multinomial Logit Models: Analysis of the Omitted Variable Bias. Discussion Paper No. 80 - 131. Center for Economic Research, University of Minnesota.

Levendusky, M. S. (2009). The microfoundations of mass polarization. *Political Analysis 17*(2), 162–176.

Lin, T.-m., J. M. Enelow, and H. Dorussen (1999). Equilibrium in multicandidate probabilistic spatial voting. *Public Choice 98*(1-2), 59–82.

Lo, J., S.-O. Proksch, and T. Gschwend (2013). A Common Left-Right Scale for Voters and Parties in Europe. *Political Analysis 22*(2), 205–223.

Mackay, R. J. and C. L. Weaver (1981). Agenda control by budget maximizers in a multi-bureau setting. *Public Choice 37*(3), 447–472.

McCarty, N., K. T. Poole, and H. Rosenthal (2006). *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge, Mass: MIT Press.

McKelvey, R. D. (1986). Covering, Dominance, and Institution-Free Properties of Social Choice. *American Journal of Political Science 30*(2), 283–314.

Merrill, S. and J. Adams (2001). Computing Nash Equilibria in Probabilistic , Multiparty Spatial Models with Nonpolicy Components. *Political Analysis 9*(4), 347–361.

Milyo, J. (2000). Logical deficiencies in spatial models: a constructive critique. *Public Choice 105*(3–4), 273–289.

Ordeshook, P. C. (1986). *Game theory and political theory*. New York: Cambridge University Press.

Pinheiro, J. C. and D. M. Bates (1996). Unconstrained parametrizations for variance-covariance matrices. *Statistics and Computing 6*(3), 289–296.

Poole, K. T. and H. Rosenthal (1985). A Spatial Model for Legislative Roll Call Analysis. *American Journal of Political Science 29*(2), 357–384.

Poole, K. T. and H. Rosenthal (2007). *Ideology & Congress*. New Brunswick, NJ: Transaction Publishers.

Quinn, K. M., A. D. Martin, and A. B. Whitford (1999). Voter Choice in Multi-Party Democracies: A Test of Competing Theories and Models. *American Journal of Political Science 43*(4), 1231–1247.

Rivers, D. (1988). Heterogeneity in Models of Electoral Choice. *American Journal of Political Science 32*(3), 737–757.

Schofield, N. (1978). Instability of simple dynamic games. *Review of Economic Studies 45*(3), 575–594.

Schofield, N., A. D. Martin, K. M. Quinn, and A. B. Whitford (1998). Multiparty electoral competition in the Netherlands and Germany: A model based on multinomial probit. *Public Choice 97*(2), 257–293.

Schofield, N. and I. Sened (2005). Multiparty Competition in Israel, 1988–96. *British Journal of Political Science 35*(4), 635.

Schofield, N. and A. Zakharov (2009). A stochastic model of the 2007 Russian Duma election. *Public Choice 142*(1-2), 177–194.

Schwartz, T. (1977). Collective Choice, Separation of Issues and Vote Trading. *The American Political Science Review 71*(3), 999–1010.

Shafer, B. E. and W. J. M. Claggett (1995). *The Two Majorities: The Issue Context of American Politics*. Baltimore: Johns Hopkins University Press.

Shikano, S. (2008). Die Eigendynamik zur Eindimensionalität des Parteienwettbewerbs: Eine Simulationsstudie. *Politische Vierteljahresschrift 49*(2), 229–250.

Signorino, C. S. and K. Yilmaz (2003). Strategic misspecification in regression models. *American Journal of Political Science 47*(3), 551–566.

Treier, S. and D. S. Hillygus (2009). The nature of political ideology in the contemporary electorate. *Public Opinion Quarterly 73*(4), 679–703.

van Egmond, M., W. van der Brug, S. Hobolt, M. Franklin, and E. V. Sapir (2013). European Parliament Election Study 2009, Voter Study. http://dx.doi.org/10.4232/1.11760.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, Mass: The MIT press.

Yatchew, A. and Z. Griliches (1985). Specification Error in Probit Models. *The Review of Economics and Statistics 67*(1), 134–139.