Estimating Relative Prevalence of Competing Issue Voting Rules in Multiparty Elections Using Finite Mixture Modeling

Kirill Zhirkov
University of Michigan
kzhirkov@umich.edu

[Draft. Pleast don't cite or circulate]

Abstract

How do citizens of democratic polities translate their policy preferences into voting choices? Proximity and directional theories of issue voting offer different answers to this question that have strong implications for parties' strategies. Controlled scenarios in imaginary two-candidate contests recently gained popularity as a method to estimate proportions of proximity and directional voters in the population. However, this method is not always applicable in comparative research where scholars often have to study multiparty elections with observational data. In the present paper, I demonstrate how relative prevalence of different issue voting rules in such cases can be estimated using finite mixture modeling. Using both a paradigmatic and a recent case, I demonstrate that the mixture model describes observed voting choices better than the alternatives. I also show how finite mixture modeling can be used to study individual-level characteristics of proximity and directional voters using education as an example.

Keywords: directional model, discrete choice, finite mixture modeling, issue voting, multiparty systems, proximity model

1 Introduction

Translation of citizens' policy positions into voting choices is the central question in the study of political representation. It has profound implications for both normative democratic theory and empirical analysis of strategies employed by political parties. The two alternative theories of issue voting offer different mathematical models linking policy positions to electoral behavior. The proximity model is based on the assumption that a voter chooses the party that is closest in terms of policy positions (Davis, Hinich, and Ordeshook 1970; Downs 1957). The directional model, in turn, assumes that citizens have preferences with respect to issue sides and vote for the most extreme of acceptable parties on the same side (Rabinowitz and Macdonald 1989). Therefore, the dominant strategies for parties should be moderation under proximity voting and polarization under directional voting.

A heated debate between supporters of the two models has been largely concluded by acknowledging that none of them could perfectly explain observable behaviors of voters or strategies of parties. As a consequence, authors have proposed alternative models explaining electoral outcomes that are neither perfectly convergent nor perfectly polarized. Examples include accounting for position of policy status-quo (Grofman 1985), and strategic balancing for potential parliamentary coalitions (Kedar 2005). Some models have explicitly attempted to combine directional and proximity elements in voters' choices, such as the policy leadership model (Iversen 1994), and the unified model of voting (Merrill and Grofman 1999). An important feature of all such models is that they assume presence of a relatively sophisticated but uniform voting rule within the population.

Recently, a survey-experimental method to estimate proportions of individuals using different issue voting rules has been proposed (Claassen 2007; Kropko and Banda 2018; Lacy and Paolino 2010). The method employs controlled scenarios: each respondent is offered two options designed so that a choice unambigously reveals the underlying voting rule. Results of the experiment carried out on a U.S. probability sample demonstrate that, even though proximity voting in either simple or discounting form was prevalent, directional voters constitute a non-trivial share of the electorate (Tomz and Houweling 2008). These findings are intuitively appealing. They suggest that individual voters use relatively simple voting rules that, nevertheless, vary within the population. Presence of different voting rules among voters, in turn, leads to aggregate outcomes that deviate from those predicted by the pure proximity model or the pure directional model.

Analytic strategies allowing issue voting rules to vary within the population have important benefits for the study of electoral behavior. Most importantly, they can be used to identify individual-level characteristics associated with different voting rules. At the same time, the experimental method to discriminate between individual issue voting rules is not universally applicable. In multiparty elections, which are prevalent in modern democracies, the numbers of choices are too large to design feasible scenarios that would unambigously discriminate between voting rules used by respondents: proximity and directional models often end up making the same choice predictions for individual voters (Fazekas and Meder 2013). Additionally, the main source of data on multiparty elections around the world is national and international electoral surveys. The only way explore relative prevalence of different voting rules with such (observational) data is to rely on statistical modeling.

The general idea that voters' utility functions ultimately translated into electoral choices are heterogeneous is not new to statistical modeling of voting decisions (Glasgow 2001; Rivers 1988; Stubager, Seeberg, and So 2018). However, most approaches to model this heterogeneity proposed so far assume continuous distributions of parameters—not qualitative differences between different voting rules. In this paper, I suggest that probabilities of different voting rules being employed in multiparty elections can be estimated semi-parametrically using finite mixture modeling (Imai and Tingley 2012; see also McLachlan, Lee, and Rathnayake 2019). The basic premise of the mixture modeling approach applied to the issue voting case is that researchers can assign probabilities to whether individuals' electoral choices are informed by either proximity or directional rule. Then, these probabilities can be modeled with covariates allowing to explore individual-level characteristics of proximity and directional voters. Importantly, this method does not require to deterministically classify respondents by the voting rule that they use—and, as a consequence, to exclude voters for whom the two models make the same predictions. Therefore, it allows maximizing the amount of information used in the analysis.

I apply the proposed method to the 1989 Norwegian parliamentary election, which have been the main testing ground for the proximity and directional voting models, and to the 2012 Dutch general election as a more recent case. I demonstrate that the mixture model is superior to alternatives in terms of fit to the data. I also show how finite mixture modeling can be used to explore individual-level characteristics of proximity and directional voters using education as an example. In conclusion, I discuss promise of the proposed method for comparative research on electoral behavior.

2 Model Setup

Suppose there are survey data describing electoral choices of voters indexed $i = 1, \ldots, I$. Political parties contesting in the election under consideration are indexed $j = 1, \ldots, J$. Voters and parties are positioned on policy issues indexed $k = 1, \ldots, K$. Each voter is described by individual-specific characteristics indexed $h = 1, \ldots, H$. Let $\mathbf{v}_i = (v_{i1}, \ldots, v_{iK})$ be a vector of length K describing voter i issue placements. Let $\mathbf{c}_i = (c_{i11}, \ldots, c_{iJK})$ be a vector of length $J \times K$ describing voter i evaluations of party (candidate) j issue placements. All issue dimensions are folded around the neutral point to obtain meaningful scalar products. Let $y_i \in \{1, \ldots, J\}$ be a scalar describing the party chosen by voter i. Let z_i be an unobserved scalar describing the issue voting rule used by voter i. Let z_i be an unobserved scalar describing the issue voting rule used by voter i. Let z_i take value 0 for proximity voters and value 1 for directional voters. Let $\mathbf{x}_i = (x_{i1}, \ldots, x_{iH})$ be a vector of length H describing individual-specific characteristics of voter i.

The choice to use voter-specific placement of parties can be criticized for potentially introducing the effects of projection: voters may underestimate the differences in issue positions between themselves and their preferred parties. To account for this potential bias, some authors use average rather than voter-specific issue positions for parties. This solution, however, igniores the fact that parties often strategically choose to present different issue positions to various segments of the population. As a result, averaging can discard meaningful variance in parties' issue positions and introduce additional measurement error (Lewis and King 2000).

Under the proximity issue voting model, the utility for voter i of choosing party j, u_{ij}^p, is defined as:

$$u_{ij}^p = \alpha_j^p + \beta^p \frac{1}{K} \sum_{k=1}^{K} |v_{ik} - c_{ijk}|, \tag{1}$$

where α_j^p is the base utility from voting for party j under the proximity model and β^p is the proximity coefficient. Historically, the proximity model has been specified using either quadratic or absolute loss function, i.e. power (exponent) of issue distances translated into voting utilities. Results from recent studies that have directly estimated the power parameter suggest it may vary across contexts (Berinsky and Lewis 2007; Jackson, Mach, and Markowski 2010; Tiemann 2019). Following these results, I explicitly estimated the power of the loss function under the proximity voting model and found that it was not significantly different from one (i.e., absolute distance specification). See Section 1 in Online Appendix for details.

Under the directional issue voting model, the utility for voter i of choosing party j, u_{ij}^d, is defined as:

$$u_{ij}^d = \alpha_j^d + \beta^d \frac{1}{K} \sum_{k=1}^{K} (v_{ik} \times c_{ijk}), \tag{2}$$

where α_j^d is the base utility from voting for party j under the directional model and β^d is the directional coefficient. The initial formulation of the directional voting model also included the concept known as the "region of acceptability": parties whose issue positions are considered too extreme carry an electoral penalty. However, most tests of the directional model, including ones performed by its authors, have omitted the region of acceptability from setup and estimation (Macdonald, Listhaug, and Rabinowitz 1991; Westholm 1997).[1] The assumption behind this omission is that, in the analyzed election, none of the parties with credible chances of gaining parliamentary representation would be considered too extreme. In my analysis, I follow the same assumption and do not model the region of acceptability.

Let $r \in \{p, d\}$ be any of the two issue voting models (rules) described above. Let $\boldsymbol{\alpha}^r = (\alpha_1^r, \ldots, \alpha_J^r)$ be the vector of length J containing party-specific base utility parameters. For identification purposes, assume $\alpha_1^r = 0$. Let β^r be the model-specific coefficient. The overall value that individual i attaches to voting for party j under any voting model r, denoted $y_{ij}^*(r)$, consists of the utility part, u_{ij}^r, that contains parameters to be estimated and the random (error) part, ε_{ij}^r:

$$y_{ij}^*(r) = u_{ij}^r(\boldsymbol{\alpha}^r, \beta^r) + \varepsilon_{ij}^r. \tag{3}$$

[1] For a recent survey-experimental study on the region of acceptability, see Gallati and Giger (2019).

Assume that, under any voting model r, voter i casts the ballot for party j if and only if overall value voter i assigns to voting for party j is the highest among the alternatives:

$$y_i = j \iff y^*_{ij}(r) = \max\left[\ y^*_{i1}(r), \ldots, y^*_{iJ}(r)\ \right]. \tag{4}$$

Also, assume that errors are independent and have the standard type I generalized extreme value distribution, also known as the standard Gumbel distribution:

$$\varepsilon^r_{ij} \sim \text{Gumbel}(0, 1). \tag{5}$$

Then (for proof, see McFadden 1973):

$$\Pr(y_i = j) = f^r_i(j \mid \boldsymbol{\alpha}^r, \beta^r) = f^r_i(j) = \frac{\exp\left[u^r_{ij}(\boldsymbol{\alpha}^r, \beta^r)\right]}{\sum_{j'=1}^{J} \exp\left[u^r_{ij'}(\boldsymbol{\alpha}^r, \beta^r)\right]}, \tag{6}$$

where the formula for u^r_{ij} can correspond to Equation 1 or Equation 2, depending on the issue voting model. The resulting functional form translating utilities into voting choices is known as conditional (multinomial) logit and belongs to a larger class of random utility models (Thurner 2000). Its parameters can be easily estimated using maximum likelihood methods. The likelihood function is:

$$L(\boldsymbol{\alpha}^r, \beta^r) = \prod_{i=1}^{I} \prod_{j=1}^{J} f^r_i(j \mid \boldsymbol{\alpha}^r, \beta^r)^{\mathbf{1}(y_i = j)}. \tag{7}$$

Under the mixture model, voter i can use either proximity ($z_i = 0$) or directional ($z_i = 1$) choice rule depending on the value of unobserved scalar z_i. One of the key benefits of finite mixture modeling is possibility to explore relative prevalence of the two voting rules and individual-level characteristics of proximity vs. directional voters. Let w_i be voter i's propensity to use directional (rather than proximity) voting rule depending on covariates x_{ih} defined as:

$$w_i = \gamma + \sum_{h=1}^{H} \delta_h x_{ih}, \tag{8}$$

where γ is the rule-ratio parameter and $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_H)$ is the vector of length H containing covariate-specific coefficients. Define linear predictor for using directional voting rule z_i^* consisting of the propensity part, u_{ij}^r, that contains parameters to be estimated and the random (error) part, e_i:

$$z_i^* = w_i(\gamma, \boldsymbol{\delta}) + e_i. \tag{9}$$

Assume that:

$$z_i = 1 \iff z_i^* > 0 \qquad \text{and} \qquad z_i = 0 \iff z_i^* \leq 0. \tag{10}$$

Also assume that errors are independent and have standard logistic distribution:

$$e_i \sim \text{Logistic}(0, 1). \tag{11}$$

Then:

$$\Pr(z_i = 1) = \pi_i(\gamma, \boldsymbol{\delta}) = \pi_i = \frac{1}{1 + \exp\big[-w_i(\gamma, \boldsymbol{\delta})\big]}. \tag{12}$$

Note that $z_i \in \{0, 1\}$, so:

$$\Pr(z_i = 0) = 1 - \Pr(z_i = 1) = 1 - \pi_i(\gamma, \boldsymbol{\delta}). \tag{13}$$

Then, using the law of total probability, it is possible to define overall probability of voter i choosing party j as a sum of the corresponding probabilities under the two alternative issue voting rules weighted by probabilities of using the respective rules:

$$\Pr(y_i = j) = \Pr(z_i = p)\ \Pr(y_i = j \mid z_i = p) + \Pr(z_i = d)\ \Pr(y_i = j \mid z_i = d) \tag{14}$$

$$= \pi_i(\gamma, \boldsymbol{\delta})\ f_i^d(j \mid \boldsymbol{\alpha}^d, \beta^d) + \big[1 - \pi_i(\gamma, \boldsymbol{\delta})\big]\ f_i^p(j \mid \boldsymbol{\alpha}^p, \beta^p). \tag{15}$$

Conceptually, the mixture model can be seen as a compound process in which an individual first chooses the voting rule and then applies it to cast a ballot for a party. In order to achieve identification, constrain base utilities to be equal for proximity and directional voters: $\boldsymbol{\alpha}^p = \boldsymbol{\alpha}^d = \boldsymbol{\alpha}$. Let $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \beta^p, \beta^d, \gamma, \boldsymbol{\delta})$ be the vector of length $J + 3 + H$ containing all mixture model parameters to be estimated. The observed-data likelihood function with the latent voting rule variable integrated out is:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{I} \left(\pi_i \prod_{j=1}^{J} f_i^d(j)^{\mathbf{1}(y_i=j)} + [1 - \pi_i] \prod_{j=1}^{J} f_i^p(j)^{\mathbf{1}(y_i=j)} \right), \tag{16}$$

where expressions for $f_i(j)$ and π_i correspond to Equation 6 and Equation 12 respectively.

3 Data and Method

In the 1990s, the 1989 Norwegian Election Study became the major testing ground for the competing theories of issue voting. Therefore, it was a natural choice as the source of data for my analysis. The 1989 parliamentary election in Norway was based on party-list proportional representation in nineteen multi-member constituencies corresponding to country's administrative regions (counties). Seven major political parties, from the radical left to the radical right, had credible chances to win seats in the parliament (see Section 2 in Online Appendix for the full list of parties). Respondets were asked to report which of the parties they voted for in the studied election.

Respondents were also asked to position themselves and the parties on political issues: government support for agriculture, environmental protection, immigration policy, privatization of healthcare, and regulation of alcohol trade (see Section 3 in Online Appendix for exact question wordings). All issue scales ranged from 1 to 10 and were centered around the middle point to obtain meaningful directional scalar products (i.e., recoded scales took values from -4.5 to 4.5).

The analyzed dataset was structured according to the formal setup presented in the previos section. The dependent variable, voting choice, contained information on the party chosen by each voter. It was a nominal variable that took values from 1 to 7 corresponding to the parties included in the analysis. Proximity terms were calculated as absolute differences between voter i and party j positions on each available issue. Directional terms were calculated as products of voter i and party j positions. These values for each voter–party pair were stored in the dataset as separate variables, i.e. in the "wide" format. In the analysis, choice utilities were modeled as sums of issue distances (proximity rule, see Equation 1) or scalar products (directional rule, see Equation 2) with different issues weighted equally.[2] Translation of utilities into choice probabilities was done via logistic function (see Equation 6).

As it often happens in survey studies, some voters did not take positions on all issues. Excluding all such voters from the analysis would decrease the sample size and statistical power. To avoid this problem, I defined a number of issue dimensions that voter i took positions on, denoted K_i. Then, I calculated issue-related components of the utility functions using the available issue positions (with appropriate scaling), instead of omitting all respondents with $K_i < 5$.

Similarly, there were voters in the data who did not place some of the parties on the five issues in question. Again, instead of omitting such people from the analysis, I estimated their utility functions using existing choice options (parties) if at least two alternatives were available.

All models presented in the paper were estimated using PythonBiogeme (Bierlaire 2016). It used maximum likelihood estimation and feasible sequential quadratic programming (FSQP; Tits 2009), an iterative gradient-based procedure designed especially for highly complex functions, as the optimization method.

[2] Equality constraints are necessary to achieve identification of the finite mixture model. Analyses with unconstrained coefficients across different issues show that, for both proximity model and directional model, all used issues significantly contribute to utility on the 99.9% confidence level and the corresponding coefficients have comparable magnitudes (see Section 4 in Online Appendix).

Table 1. Model comparison

	Pure proximity	Pure directional	Mixture
Proximity coefficient	-1.08^{***}		-1.66^{***}
	(0.04)		(0.15)
Directional coefficient		1.16^{***}	0.34^{**}
		(0.04)	(0.11)
Rule-ratio parameter			-1.34^{***}
			(0.33)
Log-likelihood	−2153.6	−2332.0	−2123.8
Parameters	7	7	9
Likelihood-ratio test			
H_0: Mixture model does not fit better			
Test statistic	59.6	416.4	
Degrees of freedom	2	2	
p value	$< .001$	$< .001$	

$N = 1736$. Standard errors in parentheses

For the full sets of parameter estimates for each model, see Sections 7, 8, and 9 in Online Appendix

* $p < .05$, ** $p < .01$, *** $p < .001$

4 Results

Comparison of the models—proximity, directional, and mixture—in terms of fit is presented in Table 1. The mixture model clearly outperforms the alternatives according to the log-likelihood ratio tests. According to the mixture model, both directional and proximity coefficients are significant and in the expected direction. Value of the rule-ratio parameter term suggests that, on average, approximately 21% of respondents used the directional voting rule (percentages obtained by translating the estimated rule-ratio parameter to the probability scale). In other words, both proximity and directional voting rules were present in the 1989 Norwegian parliamentary election and accounting for their coexistence significantly improved the prediction of voters' behavior.[3]

[3] As a robustness check, I re-estimate these models (a) with demographic controls and (b) with listwise deletion of missing values. Results show that, independently of specification, finite mixture model fits significantly better than the alternatives and values of the rule-ratio parameter are not significantly different. See Sections 5 and 6 in Online Appendix.

Table 2. Predicting voting rule

	Model 1	Model 2	Model 3
	Choice predicting part		
Proximity coefficient	-1.72^{***}	-1.70^{***}	-1.73^{***}
	(0.17)	(0.16)	(0.16)
Directional coefficient	0.37^{***}	0.37^{***}	0.38^{***}
	(0.10)	(0.10)	(0.10)
	Rule predicting part		
Education			
Binary	-1.09^{***}		
	(0.35)		
Interval		-0.24^{***}	
		(0.06)	
Years			-0.23^{***}
			(0.06)
Rule-ratio constant	-0.93^{**}	-1.15^{***}	-1.18^{***}
	(0.31)	(0.34)	(0.35)

$N = 1718$. Standard errors in parentheses

For the full sets of parameter estimates for each model, see Sections 10, 11, and 12 in Online Appendix

* $p < .05$, ** $p < 0.01$, *** $p < .001$

An important benefit of the finite mixture modeling approach is ability to include individual-level covariates for choice rules employed by voters. Here, I use education as an example: it has been argued that proximity voting rule, being more cognitively demanding, is more likely to be applied by relatively sophisticated voters (Macdonald, Rabinowitz, and Listhaug 1995). Therefore, I re-estimated the mixture model with education as a predictor of the directional vs. proximity voting rule being used. Results are presented in Table 2 for three different measures of education: binary (college or higher vs. less than college), interval (8-point from elementary to complete tertiary), and total years of full-time education. Non-binary measures of education were centered around median to obtain meaningful rule-ratio parameters.

According to my results, education was indeed significantly and robustly related to issue voting rules: highly educated respondents were less likely to use the directional rule (vis-a-vis the proximity one). The effects were also substantial in terms of magnitude. For instance, according to the model using the binary measure of education predicted probability of using the directional voting rule was approximately 28% for a person without college education and only 12% for a college graduate.

5 Replication

The 1989 Norwegian parliamentary election was chosen as the main application of the proposed method due to its prominence in the original proximity vs. directional voting debate. In order to explore the broader applicability of the finite mixture modeling to estimate relative prevalence of the two voting models, I replicated my analysis using a more recent election in a different country: the 2012 Dutch general election. The list of political parties and policy issues included in the 2012 Dutch parliamentary election study are presented in Sections 13 and 14 of Online Appendix. Data preparations and analyses were exactly the same as in the main study.

Results of this additional analyses are presented in Sections 15 and 16 of Online Appendix. They fully replicate the findings reported using the 1989 Norwegian election. First, the finite mixture model clearly outperforms the alternatives according to the likelihood-ratio tests. Second, non-trivial shares of both proximity and directional voters are present in the data. Third, less educated voters are more likely to use the directional voting rule compared to the proximity one.

6 Conclusion

In this paper, I propose a method to estimate relative prevalence of proximity and directional issue voting rules in multiparty elections based on finite mixture modeling. I apply this method to the 1989 Norwegian parliamentary election, a classic case in directional vs. proximity voting debate, and to the 2012 Dutch general election as a more recent case. I demonstrate that the mixture model fits the data significantly better than the alternatives. My analysis also highlights a major benefit of finite mixture modeling, namely the ability to predict application of voting rules using individual-level covariates. Specifically, I show that less educated voters are more likely to cast ballots using the directional voting rule.

The key benefit of the proposed method is its applicability to observational data from national and international electoral surveys. Therefore, it can be used in comparative studies on relative prevalence of directional vs. proximity voting in a country and its relationships to electoral institutions and party systems. Additionally, it might be interesting to see whether the found relationship between education and issue voting rules is replicated across countries. Scholars can also look for other individual-level characteristics that distinguish directional voters from proximity voters.

Finite mixture modeling can be also applied to a number of other questions related to issue voting. For instance, it can be used to identify groups of voters who choose parties/candidates on the basis of issues and those who cast their ballots for some other reasons (non-issue voters). Similarly, finite mixture modeling can be used to contrast individuals who care about different policy areas, such as "economic" vs. "cultural" voters. Altogether, finite mixture modeling is a powerful and flexible analytical tool that deserves more attention from scholars of electoral behavior, especially in comparative perspective.

References

Berinsky, Adam J., and Jeffrey B. Lewis. 2007. "An Estimate of Risk Aversion in the U.S. Electorate." *Quarterly Journal of Political Science* 2 (2): 139–54.

Bierlaire, Michel. 2016. "PythonBiogeme: A Short Introduction." Technical report. http://transp-or.epfl.ch/pythonbiogeme/documentation/pythonfirstmodel/pythonfirstmodel.pdf.

Claassen, Ryan L. 2007. “Ideology and Evaluation in an Experimental Setting Comparing the Proximity and the Directional Models.” *Political Research Quarterly* 60 (2): 263–73.

Davis, Otto A., Melvin J. Hinich, and Peter C. Ordeshook. 1970. “An Expository Development of a Mathematical Model of the Electoral Process.” *American Political Science Review* 64 (2): 426–48.

Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York: Harper.

Fazekas, Zoltan, and Zsombor Zoltan Meder. 2013. “Proximity and Directional Theory Compared: Taking Discriminant Positions Seriously in Multi-party Systems.” *Electoral Studies* 32 (4): 693–707.

Gallati, Laura, and Nathalie Giger. 2019. “Proximity and Directional Voting: Testing for the Region of Acceptability.” *Electoral Studies.* Published ahead of print. https://doi.org/10.1016/j.electstud.2019.02.015.

Glasgow, Garrett. 2001. “Mixed Logit Models for Multiparty Elections.” *Political Analysis* 9 (2): 116–36.

Grofman, Bernard. 1985. “The Neglected Role of the Status Quo in Models of Issue Voting.” *Journal of Politics* 47 (1): 230–37.

Imai, Kosuke, and Dustin Tingley. 2012. “A Statistical Method for Empirical Testing of Competing Theories.” *American Journal of Political Science* 56 (1): 218–36.

Iversen, Torben. 1994. “Political Leadership and Representation in West European Democracies: A Test of Three Models of Voting.” *American Journal of Political Science* 38 (1): 45–74.

Jackson, John E., Bogdan W. Mach, and Radoslaw Markowski. 2010. "Party Strategies and Electoral Competition in Post-Communist Countries: Evidence from Poland." *Electoral Studies* 29 (2): 199–209.

Kedar, Orit. 2005. "When Moderate Voters Prefer Extreme Parties: Policy Balancing in Parliamentary Elections." *American Political Science Review* 99 (2): 185–99.

Kropko, Jonathan, and Kevin K. Banda. 2018. "Issue Scales, Information Cues, and the Proximity and Directional Models of Voter Choice." *Political Research Quarterly* 71 (4): 772–87.

Lacy, Dean, and Philip Paolino. 2010. "Testing Proximity versus Directional Voting Using Experiments." *Electoral Studies* 29 (3): 460–71.

Lewis, Jeffrey B., and Gary King. 2000. "No Evidence on Directional vs. Proximity Voting." *Political Analysis* 8 (1): 21–33.

Macdonald, Stuart Elaine, Ola Listhaug, and George Rabinowitz. 1991. "Issues and Party Support in Multiparty Systems." *American Political Science Review* 85 (4): 1107–31.

Macdonald, Stuart Elaine, George Rabinowitz, and Ola Listhaug. 1995. "Political Sophistication and Models of Issue Voting." *British Journal of Political Science* 25 (4): 453–83.

McFadden, Daniel L. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, edited by Paul Zarembka, 105–42. New York: Academic Press.

McLachlan, Geoffrey J., Sharon X. Lee, and Suren I. Rathnayake. 2019. "Finite Mixture Models." *Annual Review of Statistics and Its Application* 63: 355–78.

Merrill, Samuel, III, and Bernard Grofman. 1999. *A Unified Theory of Voting: Directional and Proximity Spatial Models.* Cambridge, MA: MIT Press.

Rabinowitz, George, and Stuart Elaine Macdonald. 1989. "A Directional Theory of Issue Voting." *American Political Science Review* 83 (1): 93–121.

Rivers, Douglas. 1988. "Heterogeneity in Models of Electoral Choice." *American Journal of Political Science* 32 (3): 737–57.

Stubager, Rune, Henrik Bech Seeberg, and Florence So. 2018. "One Size Doesn't Fit All: Voter Decision Criteria Heterogeneity and Vote Choice." *Electoral Studies* 52: 1–10.

Thurner, Paul W. 2000. "The Empirical Application of the Spatial Theory of Voting in Multiparty Systems with Random Utility Models." *Electoral Studies* 19 (4): 493–517.

Tiemann, Guido. 2019. "The Shape of Utility Functions and Voter Attitudes towards Risk." *Electoral Studies.* Published ahead of print. https://doi.org/10.1016/j.electstud.2019.102051.

Tits, Andre L. 2009. "Feasible Sequential Quadratic Programming." In *Encyclopedia of Optimization*, edited by Christodoulos A. Floudas and Panos M. Pardalos, 2nd ed., 1001–05. New York: Springer.

Tomz, Michael, and Robert P. Van Houweling. 2008. "Candidate Positioning and Voter Choice." *American Political Science Review* 102 (3): 303–18.

Westholm, Anders. 1997. "Distance versus Direction: The Illusory Defeat of the Proximity Theory of Electoral Choice." *American Political Science Review* 91 (4): 865–83.

Online Appendix

1 Distance function

I estimated the proximity model with the unconstrained loss function:

$$u_{ij}^{p} = \alpha_{j}^{p} + \beta^{p} \frac{1}{K} \sum_{k=1}^{K} |v_{ik} - c_{ijk}|^{q} ,$$

where q was the power parameter. According to the estimation results, the value of q was 0.91 (see full results in the table below). A likelihood-ratio test also demonstrated that this value was not significantly different from 1 ($\chi_1^2 = 2.66$, $p = .103$). So, I proceeded assuming the absolute loss function for issue distances (i.e., $q = 1$).

Parameter	Estimate	SE	z-statistic	p-value
Proximity coefficient	−1.26	0.13	−10.28	< .001
Basic (non-issue) utilities				
Socialist Left Party	−0.88	0.09	−10.10	< .001
Liberal Party	−2.16	0.13	−16.98	< .001
Christian People's Party	−1.30	0.10	−12.78	< .001
Center Party	−2.08	0.12	−17.41	< .001
Conservative Party	−0.38	0.07	−4.82	< .001
Progress Party	−0.65	0.11	−6.07	< .001
Loss function power	0.91	0.05	17.65	< .001
Number of observations = 1736				
Number of estimated parameters = 8				
Log-likelihood = −2152.2				
SE = standard error				
Non-issue utility for the Labour Party constrained to 0				

2 List of parties

Name	Family	Election, %	Data, %
Socialist Left Party	Radical left	10.1	12.2
Labour Party	Social democratic	34.3	33.6
Liberal Party	Liberal	3.2	4.2
Christian People's Party	Christian democratic	8.5	8.7
Centre Party	Agrarian	6.5	6.0
Conservative Party	Conservative	22.2	21.9
Progress Party	Radical right	13.0	11.5

Percentages do not sum up to 100 due to those who voted for other parties or cast invalid/blank votes

3 List of issues

Agricultural policy	
	1 = Agriculture must suffice without public support
	10 = Maintain current support scheme
Nature and environment	
	1 = Increase environmental protection even if it means lower living conditions for everyone
	10 = Environmental protection should not be carried on if it affects living conditions
Immigration policy	
	1 = Want to make it easier to achieve access to Norway
	10 = Even stricter limitations on ammounts of immigrants to Norway
Private health	
	1 = Considerable larger elements of private health services
	10 = Health services should only be a public responsibility
Alcohol policy	
	1 = Free sale and large price reduction on alcoholic beverages
	10 = Sales and production of alcohol shoud be regulated even stricter

Response options from 2 to 9 not labeled

4 Results with issue-specific coefficients

Issue	Proximity	Directional
Agriculture	-0.23^{***}	0.20^{***}
	(0.02)	(0.02)
Environment	-0.29^{***}	0.27^{***}
	(0.02)	(0.03)
Immigration	-0.22^{***}	0.17^{***}
	(0.02)	(0.03)
Healthcare	-0.33^{***}	0.40^{***}
	(0.02)	(0.02)
Alcohol	-0.28^{***}	0.25^{***}
	(0.02)	(0.03)

Standard errors in parentheses

* $p < .05$, ** $p < .01$, *** $p < .001$

5 Model comparison: controls

	Proximity	Directional	Mixture
Proximity coefficient	-1.09^{***}		-1.63^{***}
	(0.04)		(0.16)
Directional coefficient		1.17^{***}	0.38^{***}
		(0.04)	(0.12)
Rule-ratio parameter			-1.36^{***}
			(0.37)
Log-likelihood	−2001.1	−2169.0	−1980.9
Parameters	7	7	9
Likelihood-ratio test			
H_0: Mixture model does not fit better			
Test statistic	40.4	159.2	
Degrees of freedom	2	2	
p value	$< .001$	$< .001$	

$N = 1663$. Standard errors in parentheses

* $p < .05$, ** $p < .01$, *** $p < .001$

6 Model comparison: listwise deletion

	Proximity	Directional	Mixture
Proximity coefficient	-1.40^{***}		-1.88^{***}
	(0.06)		(0.24)
Directional coefficient		1.81^{***}	0.84^{***}
		(0.08)	(0.21)
Rule-ratio parameter			-1.33^{***}
			(0.53)
Log-likelihood	−1086.6	−1162.8	−1083.1
Parameters	7	7	9
Likelihood-ratio test			
H_0: Mixture model does not fit better			
Test statistic	7.0	159.2	
Degrees of freedom	2	2	
p value	.030	$< .001$	

$N = 914$. Standard errors in parentheses

* $p < .05$, ** $p < .01$, *** $p < .001$

7 Full results of the pure proximity model

Parameter	Estimate	SE	z-statistic	p-value
Proximity coefficient	−1.08	0.04	−30.77	$< .001$
Basic (non-issue) utilities				
Socialist Left Party	−0.87	0.09	−10.04	$< .001$
Liberal Party	−2.15	0.13	−17.01	$< .001$
Christian People's Party	−1.30	0.10	−12.77	$< .001$
Center Party	−2.07	0.12	−17.42	$< .001$
Conservative Party	−0.38	0.08	−4.90	$< .001$
Progress Party	−0.63	0.11	−5.92	$< .001$
Number of observations = 1736				
Number of estimated parameters = 7				
Log-likelihood = −2153.6				
SE = standard error				
Non-issue utility for the Labour Party constrained to 0				

8 Full results of the pure directional model

Parameter	Estimate	SE	z-statistic	p-value
Directional coefficient	1.16	0.04	27.66	$< .001$
Basic (non-issue) utilities				
Socialist Left Party	−1.37	0.09	−15.71	$< .001$
Liberal Party	−2.18	0.13	−17.43	$< .001$
Christian People's Party	−1.50	0.10	−15.16	$< .001$
Center Party	−2.20	0.12	−18.40	$< .001$
Conservative Party	−0.28	0.07	−3.84	$< .001$
Progress Party	−1.61	0.11	−15.29	$< .001$
Number of observations = 1736				
Number of estimated parameters = 7				
Log-likelihood = −2332.0				
SE = standard error				
Non-issue utility for the Labour Party constrained to 0				

9 Full results of the mixture model (no predictors)

Parameter	Estimate	SE	z-statistic	p-value
Directional coefficient	0.34	0.11	3.21	$< .001$
Proximity coefficient	−1.66	0.15	−10.76	$< .001$
Basic (non-issue) utilities				
Socialist Left Party	−0.96	0.10	−10.00	$< .001$
Liberal Party	−2.31	0.14	−16.71	$< .001$
Christian People's Party	−1.41	0.11	−12.50	$< .001$
Center Party	−2.26	0.14	−16.60	$< .001$
Conservative Party	−0.46	0.09	−5.38	$< .001$
Progress Party	−0.80	0.13	−6.29	$< .001$
Rule-predicting part				
Constant	−1.34	0.33	−4.06	$< .001$
Number of observations = 1736				
Number of estimated parameters = 9				
Log-likelihood = −2123.8				
SE = standard error				
Non-issue utility for the Labour Party constrained to 0				

10 Predicting voting rule using binary education

Parameter	Estimate	SE	z-statistic	p-value
Proximity coefficient	−1.72	0.17	−10.34	< .001
Directional coefficient	0.37	0.10	3.61	< .001
Basic (non-issue) utilities				
Socialist Left Party	−0.95	0.10	−9.81	< .001
Liberal Party	−2.33	0.14	−16.64	< .001
Christian People's Party	−1.43	0.11	−12.51	< .001
Center Party	−2.27	0.14	−16.47	< .001
Conservative Party	−0.49	0.09	−5.59	< .001
Progress Party	−0.86	0.13	−6.62	< .001
Rule-predicting part				
Education (binary)	−1.09	0.35	−3.10	< .001
Constant	−0.93	0.31	−2.96	.002
Number of observations = 1718				
Number of estimated parameters = 10				
Log-likelihood = −2096.7				
SE = standard error				
Non-issue utility for the Labour Party constrained to 0				

11 Predicting voting rule using interval education

Parameter	Estimate	SE	z-statistic	p-value
Proximity coefficient	−1.70	0.16	−10.68	$< .001$
Directional coefficient	0.37	0.10	3.65	$< .001$
Basic (non-issue) utilities				
Socialist Left Party	−0.95	0.10	−9.82	$< .001$
Liberal Party	−2.33	0.14	−16.69	$< .001$
Christian People's Party	−1.42	0.11	−12.50	$< .001$
Center Party	−2.26	0.14	−16.50	$< .001$
Conservative Party	−0.49	0.09	−5.69	$< .001$
Progress Party	−0.86	0.13	−6.68	$< .001$
Rule-predicting part				
Education (interval)	−0.24	0.06	−4.25	$< .001$
Constant	−1.15	0.34	−3.39	$< .001$
Number of observations = 1718				
Number of estimated parameters = 10				
Log-likelihood = −2090.8				
SE = standard error				
Non-issue utility for the Labour Party constrained to 0				

12 Predicting voting rule using education in years

Parameter	Estimate	SE	z-statistic	p-value
Proximity coefficient	−1.73	0.16	−10.56	$< .001$
Directional coefficient	0.38	0.10	3.83	$< .001$
Basic (non-issue) utilities				
Socialist Left Party	−0.96	0.10	−9.86	$< .001$
Liberal Party	−2.34	0.14	−16.71	$< .001$
Christian People's Party	−1.42	0.11	−12.45	$< .001$
Center Party	−2.26	0.14	−16.52	$< .001$
Conservative Party	−0.50	0.09	−5.72	$< .001$
Progress Party	−0.86	0.13	−6.71	$< .001$
Rule-predicting part				
Education (in years)	−0.23	0.06	−4.39	$< .001$
Constant	−1.18	0.35	−3.36	$< .001$
Number of observations = 1718				
Number of estimated parameters = 10				
Log-likelihood = −2085.6				
SE = standard error				
Non-issue utility for the Labour Party constrained to 0				

13 List of parties: Netherlands 2012

Name	Family
Christian Democratic Appeal	Christian democratic
Labour Party	Social democratic
People's Party for Freedom and Democracy	Conservative liberal
Democrats 66	Social liberal
Party for Freedom	Radical right
Socialist Party	Radical left

14 List of issues: Netherlands 2012

Income differences:
"Some people think that the differences in incomes in our country should be increased. Others think that they should be decreased. How would you place [yourself/party] on a line from 1 to 7, where 1 means differences in income should be increased and 7 means that differences in income should be decreased?"

European unification:
"Some people and parties think that the European unification should go further. Others think that the European unification has already gone too far. Where would you place [yourself/party] on a line from 1 to 7, where 1 means that the European unification should go even further and 7 that the unification has already gone too far?"

Euthanasia:
"Some people think that euthanasia should always be forbidden. Other people think that euthanasia should be possible if the patient asks for this. Of course there are also people who have an opinion that lies in between. Where would you place [yourself/party] on a line from 1 to 7, where 1 means that euthanasia should be forbidden and 7 means that euthanasia should be possible?"

Foreigners:
"In the Netherlands some think that foreigners should be able to live in the Netherlands while preserving their own culture. Others think that they should fully adapt to the Dutch culture. Where would you place [yourself/party] on a line from 1 to 7, where 1 means preservation of own culture for foreigners and 7 means that they should fully adapt?"

15 Model comparison: Netherlands 2012

	Proximity	Directional	Mixture
Proximity coefficient	-1.58^{***}		-2.75^{***}
	(0.06)		(0.45)
Directional coefficient		1.03^{***}	0.74^{***}
		(0.05)	(0.08)
Rule-ratio parameter			−0.12
			(0.29)
Log-likelihood	−1540.7	−1580.9	−1517.4
Parameters	7	7	9
Likelihood-ratio test			
H_0: Mixture model does not fit better			
Test statistic	46.6	127.0	
Degrees of freedom	2	2	
p value	$< .001$	$< .001$	

$N = 1235$. Standard errors in parentheses

* $p < .05$, ** $p < .01$, *** $p < .001$

16 Predicting voting rule: Netherlands 2012

	Model 1	Model 2
Choice predicting part		
Proximity coefficient	-2.60^{***}	-2.68^{***}
	(0.38)	(0.36)
Directional coefficient	0.79^{***}	0.79^{***}
	(0.09)	(0.08)
Rule predicting part		
Education		
Binary	-0.98^{**}	
	(0.41)	
Interval		-0.47^{**}
		(0.18)
Rule-ratio constant	0.16	0.30
	(0.30)	(0.32)

$N = 1184$. Standard errors in parentheses

* $p < .05$, ** $p < 0.01$, *** $p < .001$