Lines in the sand on the Australian political beach[*]

Sinclair Davidson[a], Lisa Farrell[b] and Tim R. L. Fry[a]

[a] School of Economics and Finance
Royal Melbourne Institute of Technology
GPO Box 2476V
Melbourne, Victoria 3001

[b] Department of Economics
The University of Melbourne
Parkville, Victoria 3010, Australia

Abstract: Spatial models of voting behaviour are the dominant paradigm in political science. Consistent with this approach, it will be the case that, *ceteris paribus*, voters should vote for the party nearest to them on the political spectrum. A key question is how we measure nearness or distance. We investigate this issue by estimating discrete choice models for voting outcomes using the 2001 Australian Election Study survey data. The evidence supports the proposition that it is perceived and not actual distance that performs best. Our findings also suggest that where a voter locates on the political spectrum is almost as good a predictor of their voting outcome as how close they are to the parties.

Keywords: Spatial models, voter choice.
J.E.L. Classification: **C25, D72**

---

1.   Introduction

Hotelling (1929) established the basic idea of spatial competition. This concept has become particularly important in the Public Choice and Political Science literature, where it is known as the proximity model. Consider two soft drink vendors on a mile long beach.[1] A social planner, wishing to minimise consumer transport costs, would place each vendors at ¼ and ¾ mile positions. In a competitive market, however, each of the vendors would locate themselves back to back in the centre of the beach (1/2 mile) and attempt to "poach" consumers from the other vendor. Hotelling makes the argument at the end of his paper that his model could very easily explain political behaviour. In Hotelling's view individual voters will donate money and vote for the party closest to them on the "political beach."[2]

Thurner (2000) describes the proximity model as the "dominant paradigm in mathematical political theory." Mueller (2003) indicates this model is based on a series of "unrealistic assumptions" (p. 232); however, the predictive power of the model (and its extensions) "seems sufficiently strong" (p. 666). Lewis and King (1999) state that "countless papers" simply assume the proximity model is true while testing other hypotheses. This paper is concerned with the empirical applications of the model. We are not particularly interested in testing the theory itself rather we are interested in how "distance" is measured. In some respects our paper is similar to Lewis and King (1999) that compares the directional and proximity models. Their results suggest that the measure of party location can have "important side effects for estimation" (p. 27).

In this paper we estimate a number of models that use different measures of location and of distance and compare their accuracy. The models considered include two models that test the directional spatial model described by Rabinowitz and MacDonald (1989). Our results indicate that when distance is measured as the difference between voter location and mean party location that the directional and

---

[1] Textbook treatments of the Hotelling location theorem use beaches. Hotelling (1929) himself did not use beaches but rather "Main Street."
[2] Hotelling (1929: 55) writes, "The Democratic party, once opposed to protective tariffs, moves gradually to a position almost, but not quite identical with that of the Republicans. It need have no fear of fanatical free traders since they will still prefer it to the Republican party, and its advocacy of a continued high tariff will bring it money and votes of some intermediate groups."

proximity models are empirically equivalent. When distance is measured as the difference between voter location and voter placement of the party that the proximity model empirically dominates the directional model and the absolute measure dominates the quadratic measure.

The plan of the rest of this paper is as follows. In the next section we describe the measures of distance. Section 3 describes the data that we use in our analysis. We then discuss the models (section 4) and the results of our analysis (section 5). Finally, section 6 provides some conclusions.

2.   Measures of Distance

Political parties compete in "ideological space" and voting outcomes could be influenced either by an individual's own position in "ideological space" or by the distance between the individual voter and the party in such "ideological space". Rabinowitz and MacDonald (1989) introduce the idea of "directional spatial voting". The notion is that individuals who define themselves as being on the left will only vote for parties that they perceive to be on the left, i.e. the voter will not vote for the nearest party if they perceive that party to be "right" rather they may vote for a party further away from themselves as long as that party is "left." In this case it is the individual's own position in "ideological space" that will *ceteris paribus* influence the voting outcome.

We now consider the situation where voting outcomes are influenced by the distance between the individual voter and the party in such "ideological space". For simplicity we consider a political beach with two political parties L and R. Each party locates at position z_L and z_R. We have voters who are located across the beach. The median voter theorem says that any voter i will vote for the party (L, R) closest to her own position on the beach (z_i). There are two possibilities concerning the voting rule that might be used.

According to Hinich and Munger (1997) the voting rule is written in terms of squared distances as follows:

$$E[(z_i - z_R)^2] < E[(z_i - z_L)^2] \Rightarrow \text{Vote R} \qquad (1a)$$

$$E[(z_i - z_R)^2] > E[(z_i - z_L)^2] \Rightarrow \text{Vote L} \qquad (1b)$$

Drazen (2000), on the other hand, formulates the decision rule in terms of absolute distances as follows:

$$E[|z_i - z_R|] < E[|z_i - z_L|] \Rightarrow \text{Vote R} \qquad (2a)$$

$$E[|z_i - z_R|] > E[|z_i - z_L|] \Rightarrow \text{Vote L} \qquad (2b)$$

We view the question of whether the decision rule is written in terms of squared or absolute distances as an empirical one that we consider in our analysis.

The next issue is to define z_L and z_R. Here the issue of voter uncertainty arises. Candidates[3] for elections make promises regarding their behaviour and choices once elected. Belief or disbelief of those promises is a simple issue, however, voters may be uncertain as to believe or disbelieve election promises. Lupia and McCubbins (1998) provide a set of theories that show how, when and why voters may choose to believe or not believe candidate promises. The problem for voters is that the exact position of any given party is not known with certainty. To the extent that parties locate near each other, but not exactly next to each other, voters may have difficulty deciding who the nearest party is. A further problem facing empirical research is that we also do not know where the party is located with any certainty.

Empirical research usually uses some form of survey to rate respondents on a Left-Right scale and then also candidates (or their policy positions) on the same Left-Right scale. As an empirical issue, Alvarez (1997) uses, and recommends, the sample mean of candidate placements as a measure of the "true" position of the parties. He argues that this approach is widely employed, is simple and the data readily available. Hinich and Munger (1997) show that the sample mean is appropriate if we assume voters have rational expectations. In this case for any given party p its' position is represented as $z_p = \bar{z}_p + e_p$ where e_p is a random error term with mean zero. In this paper in the situation that the position of the party is taken as the sample mean we define "measured distance" as the difference between the voters' own position and the

---

[3] We use the terms candidate and party interchangeably in this paper.

sample mean position of the party. Measured distance could itself be defined using either squared or absolute distance.

It seems unusual to ask survey respondents where they place themselves on the political beach (a subjective assessment), to compare that to an objective (assuming rational expectations) candidate position and then "forecast" who that respondent would vote for. This approach also abstracts from the voter uncertainty as to where a candidate locates. To account for this uncertainty would require it to be explicitly modelled. Furthermore, while it does overcome some statistical issues related to missing observations (Alvarez (1997)), it does use up degrees of freedom in calculating the mean and may induce an error in variables problem. Powell (1989) has shown that error can occur due to respondents guessing where parties are also from instrument-induced error. Further Lewis and King (1999) indicate that even if the party location could be measured without any error, parties have incentives to emphasise different policies to different voters and/or groups of voters.

In short, this approach requires candidates to expend some resources establishing their position on the beach and/or voters need to expend resources discovering where parties are on the beach. As Hotelling (1929) makes clear, parties prefer not to be too specific about their exact location. Similarly the literature indicates that individual voters do not expend too many resources when making political decisions (see Lupia and McCubbins (1998)). It is for these reasons that we argue that this approach is inconsistent with the underlying theory and our knowledge of voter behaviour. An alternate approach would be to compare the voters' subjective assessment of their own location on the beach and then to compare it to their own subjective assessment of where they place the parties on the beach. The smaller that distance the more likely, *ceteris paribus*, they are to vote for that party. Thus in our analysis we define "perceived distance" as the difference between the voters' own position and the voters perception of the position of the party. Perceived distance could itself be defined using either squared or absolute distance.

This section has shown the that either an individuals' own location on the beach or their measured distance from the parties or their perceived distance from the parties can, *ceteris paribus*, influence voting outcome. Moreover, both of the distance

4

measures can be defined either in terms of a squared or absolute value metric. These measures are compared in section 5.

3.   Data

We employ data from the 2001 Australian Federal Election. At that election the entire House of Representatives and half the Senate were elected. Government is formed by the party (or parties) that command a majority in the House of Representatives. In some respects Australia is well suited to investigate the proximity model and appropriate measures of distance. First, voting is compulsory and issues such as alienation and voter turnout are unlikely to bias results. Second we have data to calculate location for both incumbents and also for (most) challengers. At a minimum we can calculate location for all major political parties that contested the House of Representatives. We cannot calculate location for individual challengers (incumbents) in some seats and for extreme fringe parties such as the Shooters Party and the No-Airport Noise Party.[4] Factors, which may reduce the desirability of Australian data, include the preferential voting system and a coalition government. These factors are known to undermine the proximity model.

We believe, however, that these two factors are less likely to introduce bias to the results. First the outcome of the election was not determined by preference votes and second the coalition government (Liberals and National parties) is an *ex ante* coalition. The parties "have been in permanent coalition since 1923" (McAllister 2003) and voters know in advance that should they win the majority of seats that these two parties will form government.

Our analysis of the 2001 election is based on the Australian Electoral Study (AES) (Bean *et al* 2002).[5] The AES is conducted after every federal election and surveying for the 2001 election occurred between 12 November 2001 and 5 April 2002. The sample is drawn from the electoral roll which, given Australia's mandatory voting

---

[4] These latter parties won very few votes and their appeal is extremely limited. On the other hand some individuals, not affiliated with any party did win some seats.
[5] A full description can be found at http://assda.anu.edu.au/codebooks/aes2001/title.html

regime, is kept up to date and is reliable. In total there are 2010 cases (individual voters) and 379 variables per case in the final data set.

Importantly for our purposes the AES asks voters to rank themselves on a "Left-Right" spectrum and also to rank the various political parties on the same "Left-Right" spectrum. The ranking ranges from 0 (very left) to 10 (very right). Most voters (41.8%) rank themselves as being in the centre (5 on the scale). Table 1 summarises the "Left-Right" spectrum results.

Table 1: Left-Right Spectrum.

	Left-Right Score		
	Mean	Median	Std.Dev
Own	5.30	5	1.98
Liberal	6.49	7	2.59
National	6.31	7	2.56
Coalition	6.40	7	2.38
Labor	4.71	5	2.36
Democrat	4.39	5	2.00
One Nation	5.90	6	3.76
Greens	3.65	4	2.40

Results are a ranking of the voter perception of self and the various political parties on a "Left-Right" spectrum. We see that the Greens are seen as the left-most party in Australia. Interestingly, the One Nation Party – popularly thought to be on the right is seen by voters to be to the "left" of each of the coalition partners (Liberals and Nationals). This result is due to the bimodal distribution of the data for that party. It is probable that this bimodality is a result of the fact that the One Nation Party has a combination of policies that could be described as being "left" from an economic perspective or "right" from a social perspective.

4.   Modelling Voting Outcomes

4.1   Discrete Choice Modelling

We now consider the individual voter's decision of which party to vote for. The AES asked respondents which party they voted for in the House of Representatives in the 2001 Federal Election. In our modelling we estimate a discrete choice model for the probability that voter i votes for party j. Underlying this decision is a random utility model:

$$U_{ij} = V_{ij} + e_{ij}, \quad i = 1,\ldots,n; \, j = 1,\ldots,J.$$

This model postulates that voter i derives "utilities" (U_{ij}) from voting for each of the J parties. These utilities are decomposed into two components: a deterministic component (V_{ij}) and a stochastic component (e_{ij}). The voter is assumed to vote for the party that yields them maximum utility. The functional form for the probability that voter i votes for party j is determined by the choice of distribution for the stochastic component e_{ij}. V_{ij} is taken to be a linear-in-parameters function of the characteristics of the parties (in our case the distance that party j is from voter i) and the characteristics of the voter (e.g. income) with outcome specific coefficients ($V_{ij} = d_{ij}a + x_i^{'} b_j$).

The random utility maximization model is made operational by the choice of the distribution for the stochastic component. We begin by assuming that the e_{ij} in the random utility model are independent, identically distributed Extreme Value. This yields the Multinomial Logit (MNL) form for the outcome probabilities:

$$\Pr(voter \ i \ votes \ for \ party \ j) = \frac{\exp(V_{ij})}{\sum_{k=1}^{J} \exp(V_{ik})}.$$

To complete the specification an identifying restriction is required such that one of the coefficient vectors (b_j) is normalised to zero[6]. The model is estimated by maximum likelihood.

---

[6] In our case the normalisation is to set the coefficient vector for Coalition to zero.

4.2   Determinants of Voting Behaviour

The focus of this paper is the investigation of the different specifications of "distance". To this end we fit a number of MNL models for individual voter choice between Coalition, ALP, Democrats, Greens, One Nation and Others. In total six models are estimated, each one having a different specification of "distance". However, to ensure that our results are comparable across the models we employ a common set of "control" variables in each of the models[7]. We now describe these control variables.

Alvarez, Nagler and Willette (2000) argue that there is an "incumbent bias". That is, when it is felt that the financial situation of the household and/or the general economic situation in the country has improved in the recent past or will improve in the future then the government is more likely to retain the support of the household. Only if the "economy" (household and/or country) is performing badly does support for opposition parties improve. Thus the way that we control for economic performance has been captured is through the "hip pocket" questions on the household and national performance compared to 12 months ago.

A number of economic issues, however, can play a role in their own right. It is often felt that unemployment, industrial relations, taxation, education and health policies and performance can play a role in determining electoral outcomes. Indeed, many parties campaign actively on such issues. Parties also focus upon other policy issues such as the environment, immigration, national security and, more recently, terrorism. To control for these issue based factors we use indicator variables for the importance that an individual places on: Health & Education, Tax & GST, Defence & Terrorism and Immigration & Refugees.

Whilst economic and other issues undoubtedly contribute to electoral outcomes, personal characteristics of the voters themselves also play a role. "Conservative" parties tend to be supported by older voters; "educated" voters might favour "Green"

---

[7] In the notation above these control variables are elements of the x_i vector.

parties and so on. Thus we also use a number of socio-economic and demographic characteristics of the voter (gender, age, income, educational attainment and marital status) as controls that influence voting outcomes. The final control variable is the type of electoral division that the voter resides in (Inner Metropolitan, Outer Metropolitan, Provincial or Rural).

Controlling for the factors discussed above we hypothesise that voting outcomes could be influenced either by an individual's own position in "ideological space" or by the distance between the individual voter and the party in such "ideological space". An individuals' positioning of themselves in Left-Right space will be denoted z_i. Such a variable is a characteristic of the individual and forms part of the x_i component of the (expected) utility function V_{ij}. In our modelling we allow for two possibilities. That the coefficients associated with z_i vary freely across voting outcomes (parties) or that the coefficients associated with z_i are equal across voting outcomes (parties).

The discussion in section 2 above suggested that there were two possible ways of defining distance (Measured and perceived distance) and that each of these could be measured using squared or absolute value metrics. If the distance that voter i is from party j is defined by d_{ij}, an individuals' positioning of themselves in Left-Right space by z_i, the individuals' placement of party j by z_{ij} and the mean (over all respondents) placement of party j as $\bar{z}_j$ we define measured distance by $d_{ij} = \left| z_i - \bar{z}_j \right|$ or $d_{ij} = \left(z_i - \bar{z}_j \right)^2$. Perceived distance is defined by $d_{ij} = \left| z_i - z_{ij} \right|$ or $d_{ij} = \left(z_i - z_{ij} \right)^2$. Plots of measured distance, absolute value and perceived distance, absolute value can be found in Figures 1 and 2[8].

FIGURES 1 and 2 about here

---

[8] In Figure 2 the points plotted are the mean perceived distance for the sub-sample of individuals at that point in the political spectrum.

Both measures clearly show that on the left of the beach voters are closer to the Australian Greens and on the right of the beach to the Coalition. It is also the case that the gaps between the parties are wider for the perceived distance measure.

Thus in our empirical analysis we estimate six models. Each of these models has a different variable to reflect either position or distance. Table 2 summarises the models:

Table 2: Models and Measures

Model	Measure
M1	Own LR position [coefficients unrestricted]
M2	Own LR position [coefficients equal across outcomes]
M3	Measured distance, absolute value
M4	Measured distance, squared value
M5	Perceived distance, absolute value
M6	Perceived distance, squared value

We now turn to a discussion of the results of estimating the six models.

5.   Results

5.1   Comparing the Models

All six models were estimated by maximum likelihood methods. As our emphasis is on comparing the empirical performance of the six competing measures we do not present the coefficient estimates rather we focus on which model (measure) performs best[9]. We note, however, that the parameter estimates on the control factors are robust across the six models.

We use a number of criteria to select the "best" model. The models are not nested, consequently we cannot formally test which is the best model. Instead we make use of two model selection criteria – the Akaike Information Criterion (AIC) (Akaike (1973)) and the Schwarz or Bayesian Criterion (SC) (Schwarz (1978)). These criteria are based upon the value of the maximized log-likelihood function adjusted for the estimation sample size and penalized for the number of parameters fitted. For both criteria the preferred model is the one that has the lowest vale of the criterion. Table 3

---

[9] Full estimation results are available on request from the authors.

gives the values of AIC and SC. From this table we see that Model 5: Perceived distance, absolute value is the "best" model.

Table 3: Model Selection Criteria Values.

Model	Sample	Parameters	lnL	AIC	SC
M1 Own Left-Right: MNL	1246	80	-1217.91	2515.81	3006.03
M2 Own Left-Right: Constrained MNL	1246	76	-1226.55	2529.11	2994.81
M3 Measured Distance: Absolute Value	1246	76	-1226.45	2528.91	2994.61
M4 Measured Distance: Squared Value	1246	76	-1219.18	2514.35	2980.06
M5 Perceived Distance: Absolute Value	1013	76	-901.49	1878.97	2328.94
M6 Perceived Distance: Squared Value	1013	76	-916.51	1909.01	2358.99

We also compare the models on their ability to correctly classify the observed voting outcomes. Table 4 tells us how accurate the six models are in classifying each of the voting outcomes and their overall accuracy. Again we see that Model 5: Perceived distance, absolute value is the "best" model.

Table 4: Percentage Correct from Classification Hit-Miss Tables.

	M1	M2	M3	M4	M5	M6
Coalition	79.66	79.66	79.66	79.66	85.02	84.39
A.L.P.	67.18	66.74	66.96	66.96	72.14	71.87
Democrat	0.00	0.00	0.00	0.00	10.61	4.55
Greens	16.42	13.43	22.39	22.39	28.33	23.33
O.N.P.	0.00	0.00	0.00	0.00	6.67	3.33
Other	0.00	0.00	0.00	0.00	0.00	0.00
Total	62.60	62.28	62.84	62.84	67.92	66.73

5.2   Some scenarios

Our comparison of the models and measures has clearly indicated that Model 5 that uses perceived distance, absolute value is the "best" model. In this section we report the results of two particular scenarios to illustrate how the probabilities of voting for each of the parties vary as we move across the Australian political beach from left to right. In the first scenario we step through the Left-Right spectrum holding all the controls at their sample means and using the mean values of distance (perceived distance, absolute vale) at each point in the spectrum. The results of this are found in Figure 3. For the second scenario we step through the Left-Right spectrum setting all the controls and distance (perceived distance, absolute vale) to their sample means at each point in the spectrum.

Figures 3 and 4 ABOUT HERE

Both figures tell a similar story that the probability of voting for the Coalition (Labor) increases (decreases) as we move from left to right. On the left, at the value of 2 on the scale there is a fall in the probability of voting Labor due to an increase in the probability of voting Green.

6. Conclusions

Spatial models of voting behaviour are the dominant paradigm in political science. Consistent with this approach, it will be the case that, *ceteris paribus*, voters should vote for the party nearest to them on the political spectrum. A key question is how we measure nearness or distance. We investigate this issue using the 2001 Australian Election Study survey data. We estimate a number of models that use different measures of location and of distance and compare their accuracy. The models considered include two models that test the directional spatial model. Our results indicate that when distance is measured as the difference between voter location and mean party location that the directional and proximity models are empirically equivalent. When distance is measured as the difference between voter location and voter placement of the party that the proximity model empirically dominates the directional model and the absolute measure dominates the quadratic measure.

7.   References

Akaike, H. (1973), "Information theory and an extension of the maximum likelihood principle", in B.N. Petrov and F. Csaki, eds., *Second International Symposium on Information Theory*, 267 – 281, Budapest: Akademiai Kaido.

Alvarez, R. M. (1997), *Information and elections*. Ann Arbor: University of Michigan Press.

Alvarez, R.M, Nagler, J. and Willette (2000), "Measuring the relative impact of issues and the economy in democratic elections", *Electoral Studies*, **19**, 237-253.

Bean, C, Gow, D. and I. McAllister (2002), *Australian Election Study, 2001 [computer file]*, Canberra: Social Science Data Archives, The Australian National University.

Downs, A. (1957), *An economic theory of democracy*. New York: Harper and Row.

Drazen, A. (2000), *Political economy in macroeconomics*. Princeton: Princeton University Press.

Hinich, M. and M. Munger (1997), *Analytical politics.* Cambridge: Cambridge University Press.

Hotelling, H. (1929), "Stability in competition", *Economic Journal*, **39**, 41 – 57.

Lewis, J. and G. King. (1999), "No evidence on directional vs. proximity voting", *Political Analysis*. **8**, 21 – 33.

Lupia, A. (1994), "Shortcuts versus encyclopedias: Information and voting behavior in California reform elections", *American Political Science Review*. **88**, 63 – 76.

Lupia, A. and M. McCubbins (1998), *The democratic dilemma: Can citizens learn what they need to know?* Cambridge: Cambridge University Press.

McAllister, I. (2003), "The Federal election in Australia, November 2001", *Electoral Studies*, **22**, 381 – 387.

Mueller, D. (2003), *Public Choice III*. Cambridge: Cambridge University Press.

Powell, L. (1989), Analyzing misinformation: Perceptions of congressional candidates' ideologies. *American Journal of Political Science*. **33**, 272 – 293.

Rabinowitz, G. and S. MacDonald (1989), "A directional theory of voting", *American Political Science Review*, **83**, 93 – 121.

Smithies, A. (1941), "Optimum location in spatial competition", *Journal of Political Economy*, **49**, 423 – 439.

Schwarz, G. (1978), "Estimating the dimension of a model", *Annals of Statistics*, **6**, 461 – 464.

Thurner, P.W. (2000), "The empirical application of the spatial theory of voting in multiparty systems with random utility models", *Electoral Studies*, **19**, 493-517.

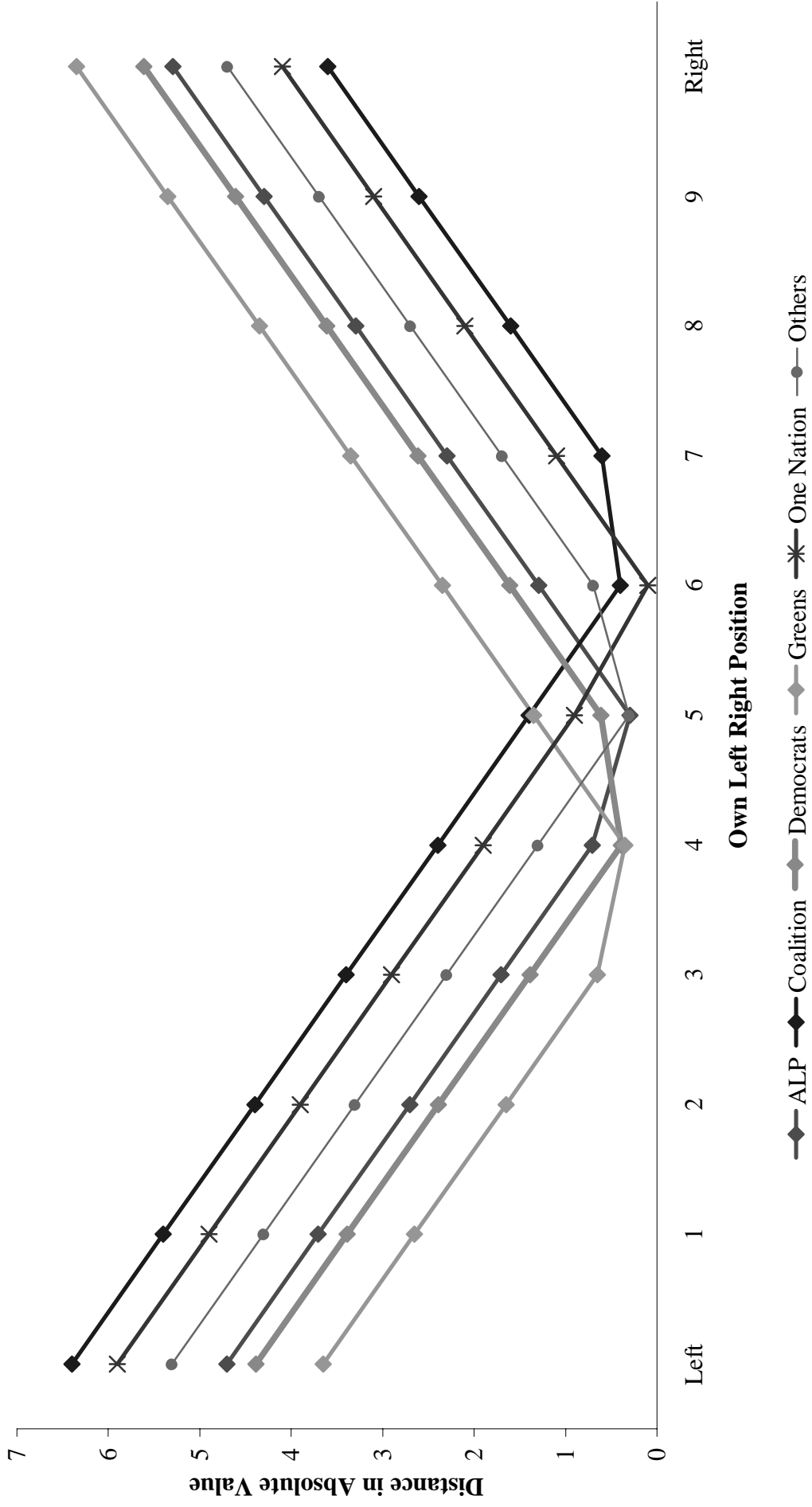**Figure 1: Measured Distances from the Parties**

Figure 2: Perceived Distances from the Parties

Own Left Right Position

Distance in Absolute Value

ALP — Coalition — Democrats — Greens — One Nation — Others
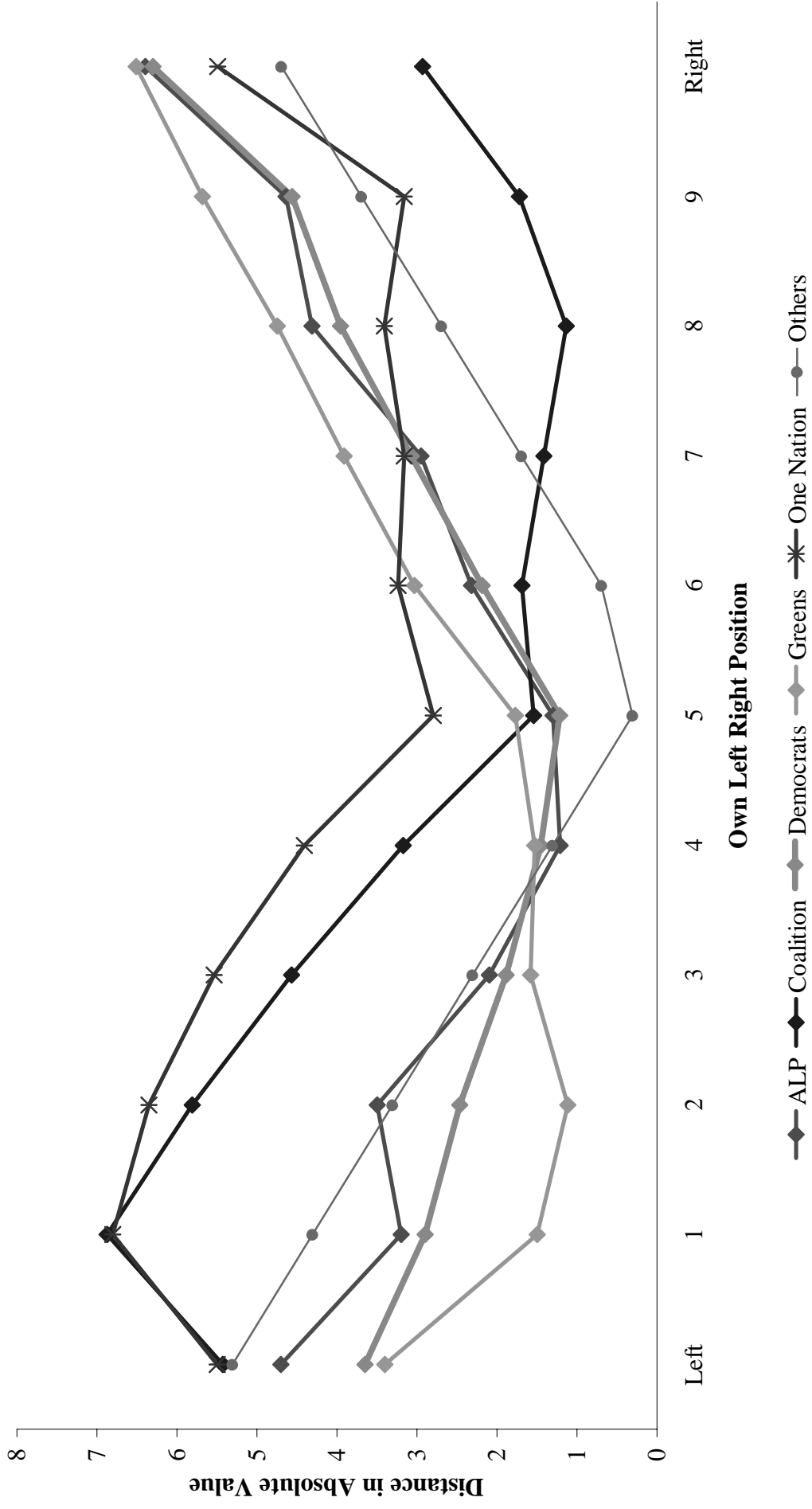
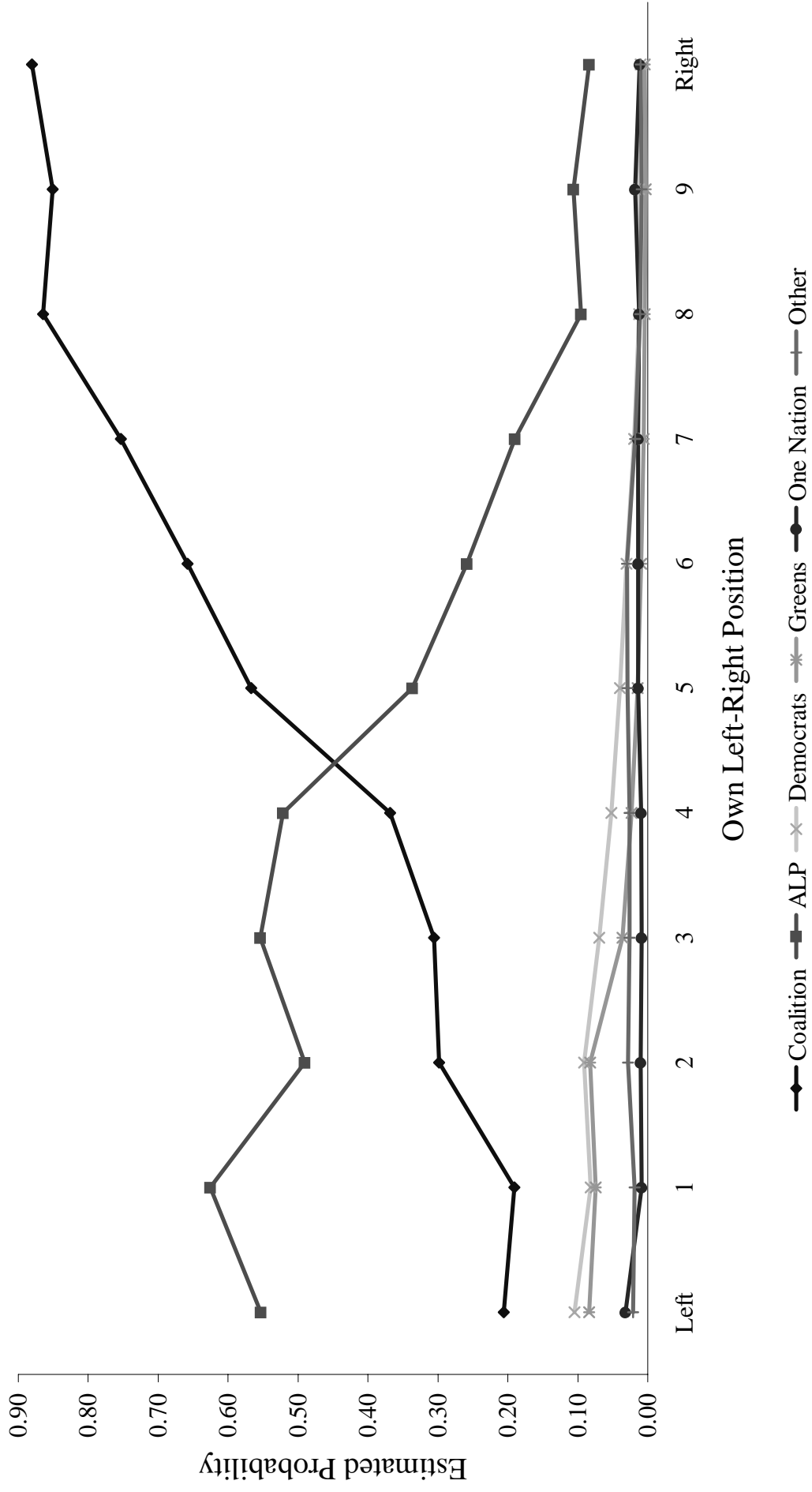**Figure 3: Probability of Voting from Model 5, Scenario 1**

Figure 4: Probability of Voting from Model 5, Scenario 2