Some evidence on non-voting models in the spatial theory of electoral competition

*Melvin J. Hinich**

Introduction

This paper presents an empirical analysis of two models of non-voting which Ordeshook and I incorporated into the spatial theory of electoral competition.[1] We called these models abstention from alienation and abstention from indifference since our goal was to formalize established social-psychological hypotheses of non-voting.[2] The data used in this paper is part of the 1968 election survey conducted by the University of Michigan's Survey Research Center.

Spatial theory of electoral competition originated with Downs and Black.[3] Their one dimensional models were generalized by Davis and Hinich.[4] A detailed exposition of spatial theory is given in Chapters 11 and 12 of Riker and Ordeshook.[5]

Let me give a brief review of the basic spatial voting model for a two candidate election where all citizens in the electorate evaluate the candidates in terms of a common set of n basic issues. Assume that every potential voter perceives the candidates as points $\underset{\sim}{\theta}_1$ and $\underset{\sim}{\theta}_2$ in an n-dimensional Euclidean space whose dimensions are these basic issues.

The dimensions are described as salient political issues in previous expositions of the theory, but it is more consistent with empirical studies of voter attitudes to conceive of the dimensions as heuristic factors which are used by a voter to forecast a candidate's behavior with respect to economic and social policy once elected to office.[6] We should expect a voter to simplify the evaluation process by reducing the complexity of the issue space. Since the choice is over representatives and not issues per se, a policy oriented

---

*Department of Economics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia. I wish to thank my graduate assistant, Mr. Thomas Southworth, for assisting me in this research. I also wish to thank Dr. Larry Cahoon for his outstanding insights, patience, and hard work during the years of development of our statistical methodology.

voter must forecast how a candidate will behave in office. It is reasonable to use past performance and past associations as a guide to a candidate's future behavior. Moreover most voters do not have much incentive to invest in information, given the small impact of a single vote and the infrequency of elections. Thus it is rational for voters to use simple rules of thumb based on inexpensive but noisy information to evaluate and choose among competing candidates.

Let us concentrate on the mth citizen in the electorate ($m = 1, \ldots, M$). Assume there exists a unique point $\underset{\sim}{x}_m$ in the space whose coordinates are citizen m's most preferred positions on the basic issues. This point, called the ideal point, is the position for citizen m's ideal candidate. For any other point $\underset{\sim}{\theta}$, assume that the mth citizen's utility function, denoted $u(\underset{\sim}{x}_m, \underset{\sim}{\theta})$, is a monotonically decreasing function of the weighted Euclidean distance:

$$||\underset{\sim}{\theta} - \underset{\sim}{x}_m||_{A(m)} = \left[\sum_{i=1}^{n} a_i(m) (\theta_i - x_{mi})^2 \right]^{1/2} \tag{1}$$

where θ_i and x_{mi} are the ith coordinates of $\underset{\sim}{\theta}$ and $\underset{\sim}{x}$ respectively, and $A(m)$ is a nxn diagonal matrix of *positive* issue weights.[7] If the mth citizen votes according to his preferences for points in the space, these assumptions imply that the mth citizen votes for candidate one if and only if:

$$||\underset{\sim}{\theta}_1 - \underset{\sim}{x}_m||_A < ||\underset{\sim}{\theta}_2 - \underset{\sim}{x}_m||_A . \tag{2}$$

If the inequality is reversed, citizen m votes for candidate two.

1. Abstention models

The two abstention models, originally defined in terms of a vote/abstain dichotomous choice rule, were combined into a general probabilistic voting model by Hinich, Ledyard, and Ordeshook.[8] I will now define them separately in terms of probabilistic choice functions of the utility function $u(\underset{\sim}{x}_m, \underset{\sim}{\theta})$.

Abstention from indifference

The probability that the mth citizen votes for the closest candidate to his ideal point is a monotonically *increasing* function of the absolute difference in utilities:

$$DU_m = |u(\underset{\sim}{x}_m, \underset{\sim}{\theta}_1) - u(\underset{\sim}{x}_m, \underset{\sim}{\theta}_2)| . \tag{3}$$

In order to relate this model with the spatial voting rule given by (2), assume that if citizen m decides to vote when $DU_m = 0$, he votes for candidate one with probability $\frac{1}{2}$. This model implies that in a random sample of the electorate, citizens with large DU's are more likely to vote than citizens with small DU's.[9]

Abstention from alienation
The probability that the mth citizen votes for his closest candidate is a monotonically *increasing* function of the maximum utility for the candidates:

$$MU_m = \max_{1,2}. \{u(\underset{\sim}{x}_m, \underset{\sim}{\theta}_1), u(\underset{\sim}{x}_m, \underset{\sim}{\theta}_2)\} \tag{4}$$

Given a random sample of the electorate, citizens with large MU's are more likely to vote than citizens with small MU's.

One variant of these abstention models which will be discussed later holds when $u(\underset{\sim}{x}_m, \underset{\sim}{\theta})$ is approximately linear in $- || \underset{\sim}{\theta} - \underset{\sim}{x}_m ||_A$.

In this case, expression (3) in the definition of indifference can be replaced by the absolute difference in distance, namely:

$$DD_m = \left| \, || \underset{\sim}{\theta}_2 - \underset{\sim}{x}_m ||_A - || \underset{\sim}{\theta}_1 - \underset{\sim}{x}_m ||_A \, \right|. \tag{3a}$$

For the spatial variant, the probability that the mth citizen votes for his closest candidate is a monotonically *decreasing* function of the minimum distance:

$$\text{Min}D_m = \min_{1,2}. \{ || \underset{\sim}{\theta}_1 - \underset{\sim}{x}_m ||_A , || \underset{\sim}{\theta}_2 - \underset{\sim}{x}_m ||_A \}. \tag{4a}$$

I will now discuss a simple test of these models using "feeling thermometer" scores from the 1968 SRC election survey.

2. Feeling thermometer scores

The individuals interviewed in the 1968 survey were a representative cross-section of voting age citizens living in private households in the continental U.S. The twelve largest metropolitan areas of the country were chosen with certainty.

Interviewing was conducted in two waves. For the first or pre-election wave, the interviewing started in the month of September and continued through the first four days in November. The post-election interviewing began immediately after election day and ended in the latter days of February, 1969. The overall response rate for the pre-election survey was 86.3%. To compensate for an unexpectedly low post-election response rate, a two-page mail questionnaire, inquiring about the 1968 voting behavior, was sent out to 182 post-election non-interviewees for whom a mailing address was available. This mailing, and a subsequent one, brought responses from 36 of the original non-interviewees, thereby increasing the overall post election response rate to 88.5%.

The question used in this analysis concerns the feeling which the respondents held towards twelve major political figures, including Humphrey, Nixon, and their running mates. This question was asked in the post-election wave. The respondents were told to use a score of 100° to indicate a very warm or favorable feeling for the politician. In contrast, a score of 0° was to indicate a very cold or unfavorable feeling for the politician. A response of 50° was specified to mean "No feeling at all for the politician." An inspection of the distribution of scores in the sample suggested that many respondents were using a 50° response when they did not know much about a politician. Presumably these respondents were unwilling to take the initiative and admit their ignorance.[10]

Let T_{jm} denote the mth respondent's score of the jth politician. Let me make the critical assumptions that a) the maximum of T_{Nm} and T_{Hm} for each respondent is proportional to the maximum utility of MU_m, and b) that the absolute difference in thermometer scores for Nixon and Humphrey, $|T_{Nm} - T_{Hm}|$, is proportional to DU_m for these two candidates.

Table 1. Means and standard deviations of thermometer statistics of the post-election sample.

	Total	Voters Nixon/Humph	Non-voters with Wallace voters	Non-voters excluding Wallace voters
N	1391	920	471	260
Max T_m	80.3 (17.5)	84.4 (14.1)	72.5 (20.6)	82.5 (17.2)
$\lvert T_{Nm} - T_{Hm} \rvert$	31.7 (23.4)	33.7 (22.7)	27.7 (24.1)	31.6 (23.8)
T_{Nm}	65.8 (23.5)	68.6 (22.6)	60.3 (24.2)	65.8 (24.5)
T_{Hm}	63.3 (27.3)	66.4 (26.4)	57.1 (28.0)	67.5 (26.2)

These assumptions seem reasonable for respondents who have moderate preferences for or against Nixon and Humphrey. The assumptions are questionable for respondents who give 0° or 100° to one or both of the candidates. It is impossible to determine the intensity of negative feelings toward a candidate indicated by a 0° score. The finiteness of the range of scores distorts the utility response in an unknown manner. Since there is no common unit for utilities, moreover, the mapping of utility to thermometer scores is idiosyncratic. The proportionality constants for the relationships between thermometer scores and the abstention models vary among respondents.

There are other problems with the assumptions connecting thermometer scores with utilities. As was previously mentioned, there seems to be a confounding of 50° scores with "don't knows." Some respondents might give non 50° scores on some random basis to politicians in the party which is not their preferred party. In addition, some respondents may be unwilling to express their true preferences for political figures they are cognizant of.[11]

The next section presents an analysis of the abstention models using the thermometer scores. The data is consistent with the existence of both types of abstention models, assuming the relationship between the thermometer scores and utilities as stated above. After the first data results are presented, I will discuss the causal relations between the models and the data in light of the assumptions.

3. Analysis of thermometer scores

I tabulated the thermometer scores for the 1391 respondents out of the post-election sample of 1481 who gave scores to *both* Nixon and Humphrey. There were 920 respondents who stated they voted for either Nixon or Humphrey.

The Wallace voters, however, posed a tactical problem for testing the abstention models. It is hard to believe that many Wallace voters expected Wallace to win, or even expected that Wallace would receive enough electoral voters to have a significant impact on social policy. Given the nature of the 1968 campaign, it seems reasonable to assume that Wallace voters did not vote strategically. In light of these assumptions, two definitions of abstention were used. In one, the 211 respondents who stated they voted for Wallace or said they would have had they voted were classified as non-voters. In a second classification, these respondents were excluded. The means and standard deviations of T_{Nm}, T_{Hm}, $|T_{Nm} - T_{Hm}|$, and the maximum of the pair $\{T_{Nm}, T_{Hm}\}$ are displayed in Table 1.[12]

Since $DU_m = |T_{Nm} - T_{jm}|$ by assumption, it follows that the mean absolute difference in thermometer scores should be lower for non-voters than for voters if the *indifference* hypothesis holds. Similarly, if the *alienation* hypothesis holds, the mean maximum score should be lower for non-voters than for voters. Table 2 presents the t-values, degrees-of-freedom, and P values for two sample t tests of differences in the means of voters and non-voters using the Welch treatment of the Bahrens-Fisher problem.[13]

Table 2. One sided P values for (Welch) t tests of differences in the means

Whole post-election sample		
Wallace voters included		
Max T_m	$t_{703} = 11.12$	$P = 0$
$\lvert T_{Nm} - T_{Hm} \rvert$	$t_{902} = 4.49$	$P = 3 \times 10^{-6}$
Wallace voters excluded		
Max T_m	$t_{364} = 1.62$	$P = 0.053$
$\lvert T_{Nm} - T_{Hm} \rvert$	$t_{402} = 1.26$	$P = 0.104$
Republicans and Independents leaning to Rep's in the subsample		
Wallace voters included		
Max T_m	$t_{66} = 3.75$	$P < 0.001$
$\lvert T_{Nm} - T_{Hm} \rvert$	$t_{96} = 2.52$	$P < 0.01$
Wallace voters excluded		
Max T_m	$t_{38} = 0.66$	$P = 0.27$
$\lvert T_{Nm} - T_{Hm} \rvert$	$t_{44} = 2.24$	$P < 0.025$
Democrats and Independents leaning to Dem's in the subsample		
Wallace voters included		
Max T_m	$t_{91} = 5.30$	$P < 0.001$
$\lvert T_{Nm} - T_{Hm} \rvert$	$t_{116} = -0.57$	Wrong sign
Wallace voters excluded		
Max T_m	$t_{43} = 1.76$	$P < 0.05$
$\lvert T_{Nm} - T_{Hm} \rvert$	$t_{52} = -0.55$	Wrong sign

The results for the case with Wallace voters defined as non-voters support the cross-tabulation findings of Page and Brody.[14] Both the abstention from indifference and alienation effects are significant at the 0.05 level. When the Wallace voters are excluded, however, the differences in the means have the "correct" sign, but they are not statistically significant at the 0.05 level. The Wallace voters gave much lower scores on average to Nixon and Humprey than the rest of the sample. The correlation coefficient between max T_m

and $|T_{Nm} - T_{Hm}|$ in the whole sample is 0.43. The correlation coefficient is 0.42 for the Humphrey-Nixon voters.

The results show a statistical relationship between the thermometer scores and non-voting. Provided that most respondents revealed their issue oriented preferences in their scores, the results support the non-voting models presented above. But there is another plausible explanation of the results. It might have been the case that respondents who did not vote gave equal scores to Humprey and Nixon, and Wallace voters gave them low scores In other words, respondents who never vote or who are ignorant about politics give either equal scores to the major candidates, or they give them low scores. These competing models can be resolved by a panel study over several elections, but a partial resolution can be obtained by analyzing a subsample of the 1968 respondents. The next section presents an analysis of a specially selected subsample of 756 respondents.

4. A subsample analysis

Assume that respondents who (1) do not score all twelve politicians, or (2) who give a score of 50° to four or more politicians, tend to be politically ignorant or apathetic. These respondents were removed from the sample. Another group of ten respondents were removed because they did not reveal their voting decision. Table 3 shows the difference between the total and subsample income and education distributions, and also gives a breakdown of reported voting statistics for the two groups. The subsample has higher income and education than does the whole sample, and contains fewer non-voters.

From the remaining 756 respondents two groups were analyzed. The first group consisted of those respondents who identified themselves as strong Democrats, weak Democrats, or Independents who leaned toward the Democrats. This latter group was included since a review of the data indicated that their preferences were strongly Democratic. The second group were those respondents who identified themselves as strong Republicans, weak Republicans, or Independents who leaned toward the Republicans (see Table 4).

The t statistics for the differences in means using the subsample scores are shown on the bottom part of Table 2.

For the Republican identifiers, the t value for the difference in the means of $|T_{Nm} - T_{Hm}|$ between voters and non-voters is significant at the 0.05 level whether the Wallace voters are included or excluded from the non-voter category. Thus, the indifference mode is supported by the Republican results. The alienation effect is washed out when the Wallace voters are excluded.

For the Democratic identifiers, on the other hand, the alienation model is supported by the data since the t values of the differences of the $\text{Max}\, T_m$ means are significant at the 0.05 level. The indifference model is rejected

Table 3. Income distribution and educational level of respondents; voting statistics.

Income distribution of respondents

Income	Percent of population sample in each income group	Percent of subsample in each income group
Less than $1,999	9.0	4.2
$2,000-3,999	14.2	10.4
$4,000-5,999	13.4	11.1
$6,000-7,999	18.4	19.6
$8,000-9,999	13.0	16.3
$10,000-11,999	11.1	11.9
$12,000-14,999	9.5	12.2
$15,000-19,999	6.0	7.0
$20,000-24,999	2.0	2.6
$25,000 or more	3.3	4.6

Educational levels of respondents

Education level	Percent of population sample in each educational group	Percent of subsample in each educational group
Eight grades or less	21.3	13.1
Between eight and twelve	40.0	35.2
Some or all of college	34.6	45.3
Advanced degrees	4.1	6.5

Non-voting

	Population Sample	Subsample
Voted	75.8%	86.5%
Abstained	24.2%	13.5%

since the means of $|T_{Nm} - T_{Hm}|$ for non-voters is greater than the means of voters.

These tests only make use of three basic assumptions: 1) Max T_m is proportional to MU_m for each m, 2) $|T_{Nm} - T_{Hm}|$ is proportional to DU_m, and 3) the decision to vote depends on the utilities in the form described in Section 1. No use has been made of the *spatial model*. I will present an analysis using the spatial model of utility.

Table 4. Means and standard deviations of thermometer statistics.

	Total	Voters Nixon/Humph	Non-voters with Wallace Voters	Non-voters excluding Wallace voters
Republicans and Independents leaning to Republicans				
N	243	191	52	31
Max T_m	83.4 (14.8)	85.6 (13.0)	75.4 (13.1)	83.7 (14.5)
$\|T_{Nm} - T_{Hm}\|$	38.5 (25.2)	40.4 (25.9)	31.5 (21.4)	30.5 (22.3)
Democrats and Independents leaning to Democrats				
N	312	239	73	39
Max T_m	82.1 (16.0)	85.1 (13.1)	71.8 (20.1)	79.2 (20.2)
$\|T_{Nm} - T_{Hm}\|$	31.0 (22.8)	30.6 (22.7)	32.4 (23.3)	32.7 (21.8)

5. Using the spatial model

The scores which a respondent gave to the other candidates such as Johnson, Muskie, Reagan, etc. contain information about the ideal points of the respondents. The whole set of scores also can be used to estimate the positions of the candidates in a political space, under certain assumptions about common voter perceptions and issue weights. Cahoon, Hinich, and Ordeshook have developed a technique to "map" ideal points and candidate positions using a parametric spatial model.[15] The CHO technique is based on the following spatial model for the thermometer scores. Recalling the definition of weighted Euclidean distance (1), assume that the mth respondent's thermometer score for politician j takes the form:

$$T_{jm} = 100 - \| \theta_j - x_m \|_A^r + \epsilon_{jm}, \tag{5}$$

where ϵ_{jm} is a stochastic error term, and at least one of the n dimensions is a valence issue.[16]

Citizens may define the space with a particular sensitivity to positions that are "far" from their ideal points or alternatively they may be sensitive, i.e., perceive differences as substantively meaningful, only if positions are "near" their ideal points. To accommodate the several possibilities, the exponent r is allowed to vary. If $r = 2$, citizen m is more sensitive to positions that are far from $\underset{\sim}{x}_m$. If $r = 1$, the sensitivity is uniform. If $r = 1/2$, citizens are more cognizant of differences near their ideal points. For a given population, r is a parameter that either must be estimated in terms of goodness-of-fit, or is preselected.

Assuming that the thermometer score is given by (5) for each respondent, the CHO goal is to identify the dimensionality of the issue space, n, each politician's position in the space, $\underset{\sim}{\theta}_j$, each citizen's ideal point, $\underset{\sim}{x}_m$, and the matrix of issue weights, A. The CHO method, in a nutshell, transforms the covariance matrix of thermometer scores and applies the principal components version of factor analysis (see Appendix). The covariance matrix must be modified since a factor analysis is appropriate only if T_{jm} is a linear function of $\underset{\sim}{x}_m$, whereas from (5) it is clear that T_{jm} has a non-linear term $\underset{\sim}{x}_m' \underset{\sim}{A} \underset{\sim}{x}_m$. In addition to the spatial model itself, the basic assumptions underlying the method are: (1) all respondents have the same perception of each candidate; (2) $a_i(m)$ is independent of m for each i, that is, all respondents weight the issues in an identical fashion; and (3) the ideal points in the sample have considerable variation. The first two assumptions, while implicit in metric scaling techniques and while possessing a long history in spatial analysis, are restrictive. Without the assumption of some structure, however, estimation is impossible.

As a partial resolution of the problem the Democratic and Republican party identifiers were analyzed separately. These groups are sufficiently homogeneous with respect to perceptions and issue weights to make assumptions 1 and 2 reasonable, but have considerable dispersion of ideal points. The CHO technique cannot be used for groups, such as the Blacks, whose thermometer scores have little variation in the orderings of the politicians.

The feeling thermometer scores from each group were used by CHO to estimate a joint space containing politicians and respondents. As is discussed in detail in the cited working paper, the data supports an assumption that the space is two-dimensional (with a third valence dimension for the less well-known politicians). The best fit for the exponent r is obtained for $r = 1/2$.

The estimated thermometer score for Nixon given by the mth respondent is:

$$\hat{T}_{Nm} = 100 - [\hat{a}_1(\hat{\theta}_{N1} - \hat{x}_{m1})^2 + \hat{a}_2(\hat{\theta}_{N2} - \hat{x}_{m2})^2]^{1/4}, \qquad (6)$$

where $\hat{a}_1$ and $\hat{a}_2$ are the estimated salience weights, $\hat{\theta}_{N1}$ and $\hat{\theta}_{N2}$ are the estimates of Nixon's position coordinates, and $\hat{x}_{m1}$ and $\hat{x}_{m2}$ are the estimated ideal point coordinates. The scores for extremists can be less than 0 since $\hat{T}_{Nm}$ is not constrained to be non-negative. The estimated thermometer score for Humphrey is the same as (6) with $\hat{\theta}_{H1}$ and $\hat{\theta}_{H2}$ instead of $\hat{\theta}_{N1}$ and $\hat{\theta}_{N2}$. The average correlation between computed scores and the raw thermometer scores is about 0.38. The weights $\hat{a}_1$ and $\hat{a}_2$, although not constrained to be positive, were positive for the two groups. This result provides strong support for the spatial model of the scores.

There are no significant differences in the candidate maps for the three

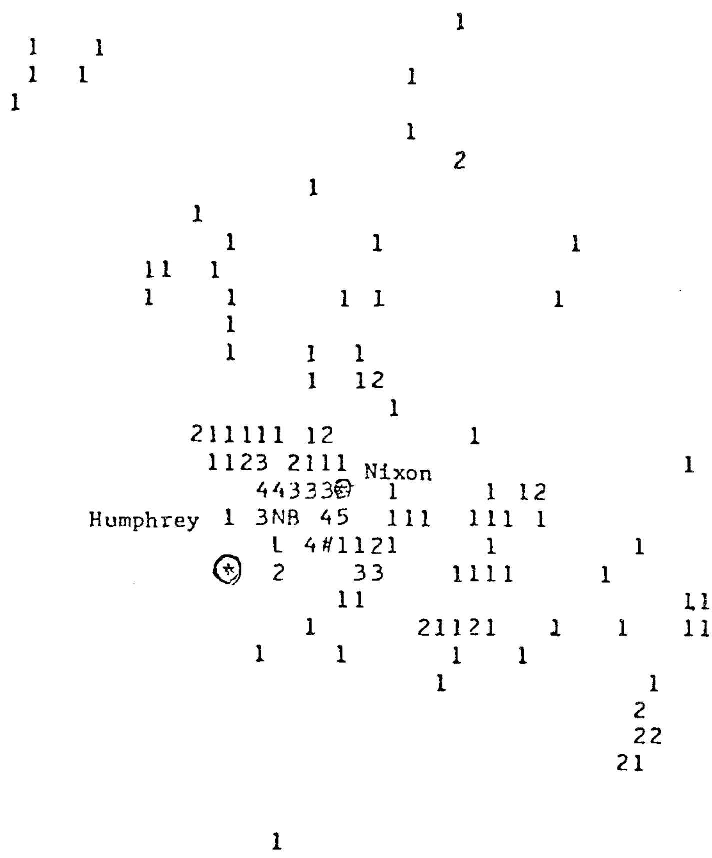

Figure 1. Distribution of Republican ideal points, # = 10, A = 11, B = 12, etc.

groups. The major differences among the estimated parameters between the groups were in the mean ideal points and the issue weights. For example, $\hat{a}_2/\hat{a}_1 = 1.74$ for the Republicans and $\hat{a}_2/\hat{a}_1 = 5,43$ for the Democrats, and the magnitude of the valence issue of the Democrats was 2/3 of the magnitude for the Republicans. These results are discussed in detail in the cited working paper. The results are sufficiently positive to support the use of candidate-ideal point map to test the non-voting models for the subsample.

The estimated ideal points for the Republican identifiers and non-voters are shown in Figures 1 and 2, where the coordinates have been adjusted by the salience weights so that any distance between points in these figures is *simple Euclidean distance*. As a consequence, all the ideal points which are closer to Nixon in Figure 1 should prefer Nixon to Humphrey and vice-versa. For Republican identifiers who stated they voted, such a decision rule gave

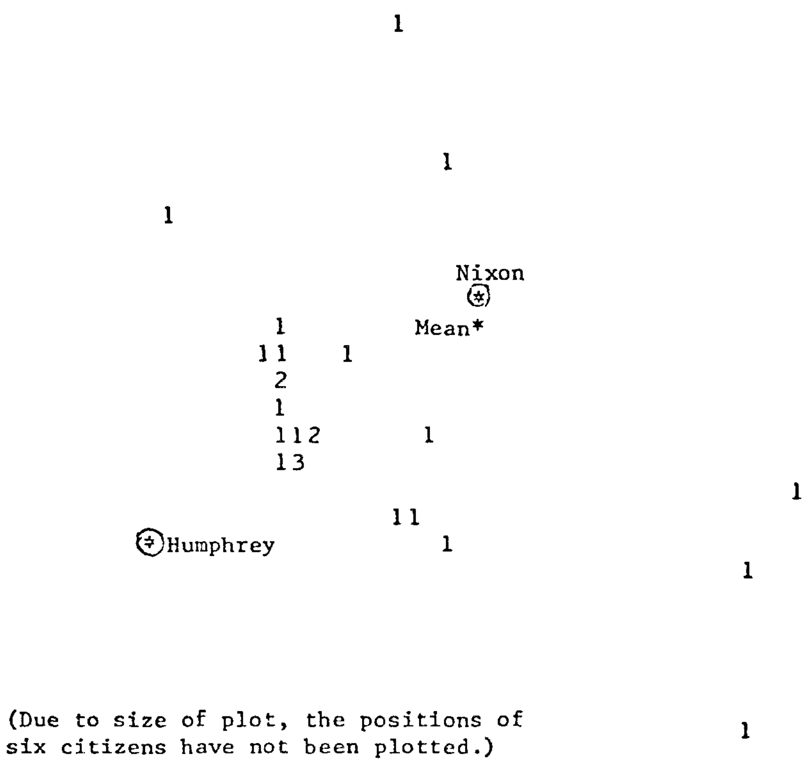

Figure 2. Republicans 1968: non-voters.

Note. *Mean denotes average ideal point for Republicans.

95.9% correct predictions using a bias correction discussed by CHO.

It was not possible to map Wallace in the space since Wallace received a zero score from about 43% of the sub-sample. These zero swamped the sample covariance between positive Wallace scores and other candidate scores, causing Wallace (and LeMay) to be placed by the procedure on a separate dimension. If respondents were allowed to give negative scores, we might have been able to place these two candidates at the extreme position where we subjectively felt them to be located.

These spatial maps of Republican and Democratic party identifiers can be used to test the spatial variant of the abstention models. The statistics for DD_m and MinD_m, as defined by expressions (3a) and (4a) are given in Table 5. The alienation effect for these voters seen in Table 2 is lost in the spatial data. When the Wallace voters are excluded, the two sample t statistic is significant at the 0.05 level for the indifference model. *The alienation model is rejected for the Republican subsample.*

Table 5. Statistics for mean differences in distances to Nixon and Humphrey and mean minimum distance.

	Voters	Non-voters incl. Wallace voters	t-value	Non-voters excl. Wallace voters	t-value
Republican Identifiers					
N	191	52		31	
DD_m	15.5 (12.4)	15.6 (12.4)	$t_{80} = -0.06$ Not significant	10.9 (10.8)	$t_{44} = 2.15$ $P < 0.025$
MinD_m	44.3 (44.4)	41.8 (32.6)	$t_{108} = -0.45$ Not significant	34.5 (23.6)	$t_{71} = 1.85$ Wrong sign
Democratic Identifiers					
N	239	73		39	
DD_m	11.7 (7.9)	13.4 (8.9)	$t_{109} = -1.50$ Wrong sign	13.6 (8.6)	$t_{49} = -1.31$ Wrong sign
MinD_m	49.2 (61.9)	76.6 (76.6)	$t_{102} = 2.78$ $P < 0.005$	65.2 (77.6)	$t_{46} = 1.23$ Not significant

For the Democratic subsample, *the indifference model is rejected*. It appears that the Democratic Wallace voters exhibit the alienation effect. When they are excluded, the difference between the mean of MinD_m for voters and non-voters has the correct sign for alienation, but the t value is not statistically significant at the 0.05 level. These results are consistent with the Democratic thermometer results given in Table 2.

There is some confounding between the two models. When $DD_m = 0$, $\tilde{x}_m$ lies on the line which bisects the line segment connecting $\tilde{\theta}_N$ and $\tilde{\theta}_H$. An individual whose ideal point lies near either extreme of this bisecting line has a small DD_m and a large MinD_m. Thus one cannot distinguish between the models for individuals at the extreme of the bisector, but there are relatively few such respondents in those regions for the two subsamples.

6. Conclusion

The results viewed as a whole support the hypothesis that the alienation and indifference effects are present in the population. The Wallace voters seem to be an alienated group, especially for the Democrats. The indifference model is supported by the data from the Republican sample, but not for the Democrats.

The spatial results presented in this paper provide a first step in the process of testing the spatial model of voting. An analysis of the 1972 and 1976 election surveys should provide more support for spatial theory.

Appendix

Suppose that M respondents give thermometer scores T_{jm} to $p + 1$ political figures whom I will call candidates for sake of a better label. Thus the data set is an $M \times (p+1)$ array. The estimation procedure begins by transforming this array into one whose entries are linear in θ_j and $\tilde{x}_m$. This is done by raising the scores to the rth power and then subtracting the entries for the $p + 1$st candidate from the other p. In order to simplify the exposition, I will present the methodology for the case when $r = 1$, and for a two dimensional mapping ($n = 2$).

Before going into the details of the statistical method, let me try to clarify the treatment of the valence issue. Suppose there is a third dimension, such as candidate honesty or executive ability, for which all citizens have identical ideal points. With no loss of generality let this ideal point be *zero*. It is impossible to use the scores to estimate the candidate positions θ_{j3} when there is no variation in the ideal positions on this dimension. As a consequence, the θ_{j3} positions must be chosen by the data analyst.

In our analysis of the 1968 survey, we decided that the average respondent knew more about Johnson, Nixon, R. Kennedy and Humphrey than the others, especially McCarthy and Romney. We grouped Robert Kennedy with the other three because we believed that many respondents scored him as if he were JFK (RFK was killed five months before the survey). We grouped Rockefeller and Reagan with the lesser known figures on the premise that they lacked national stature although they had strong regional reputations. The Republican results were unaltered when we grouped Rockefeller with J, N, K, and H. Given our subdivision, we set $\theta_{j3} = 0$ for J, N, K, and H and $\theta_{j3} = 1$ for the rest. The dimension weight a_3 was estimated for each group, yielding $\hat{a}_3/\hat{a}_1 = 0.39$ for the Democrats and 0.79 for the Republicans.

Returning to the statistical model, suppose we reorder the candidates so that $\theta_{p+1,\,3}$

= 0. Define the transformed score:

$$Y_{jm} = D_{jm}^2 - D_{p+1,m}^2 - \overline{D_j^2} + \overline{D_{p+1}^2}, \tag{A1}$$

$$D_{jm} = 100 - T_{jm}$$

where $\overline{D_j^2}$ denotes the *sample average* of D_{jm}^2. Thus from (5)

$$Y_{jm} = -2\underset{\sim}{\theta}_j'(\underset{\sim}{x}_m - \overline{\underset{\sim}{x}}) + 2\,\|\underset{\sim}{\theta}_j - \underset{\sim}{x}_m\|_A\, \epsilon_{jm} \tag{A4}$$

$$-2\,\|\underset{\sim}{x}_m\|_A\, \epsilon_{p+1,m} + \epsilon_{jm}^2 - \epsilon_{p+1,m}^2 + 0(M^{-1}),$$

where:

$$\overline{\underset{\sim}{x}} = M^{-1} \sum_{m=1}^{M} \underset{\sim}{x}_m \text{ and } \|\underset{\sim}{x}_m\|_A = [\sum_{m=1}^{M} a_i x_{mi}^2]^{1/2}.$$

Assume that the errors ϵ_{jm} are uncorrelated across respondents and candidates. In other words, the expected covariance structure of the Y_{jm} is completely specified by the spatial model. Assume in addition that $E\epsilon_{jm} = 0$, $E\epsilon_{jm}^3 = 0$, and $E\epsilon_{jm}^4 = 3\sigma_j^4$ where $\sigma_j^2 = E\epsilon_{jm}^2$ (the variance of the ϵ_{jm}) for each j and m. The assumptions about the third and fourth moments of ϵ_{jm} hold if ϵ_{jm} is normally distributed. It is proven in Cahoon, L. and Hinich, M. J., "Locating Targets Using Range Only,"*IEEE Trans. on Information Theory IT-22*, March (1976), that the expected value of the pxp covariance matrix of the Y_{jm} is:

$$4\underset{\sim}{\theta}'A\Sigma A\underset{\sim}{\theta} + c\underset{\sim}{1}\underset{\sim}{1}' + \Psi. \tag{A5}$$

In this expression, $\underset{\sim}{\theta} = (\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2, \ldots, \underset{\sim}{\theta}_p)$ is the $2xp$ matrix of candidate positions, Σ is the true nxn covariance matrix of the ideal points, $\underset{\sim}{1}$ is a $px1$ vector of ones, Ψ is a *pxp diagonal* matrix with jth diagonal element $4[\tilde{E}(D_{jm}^2) - (\sigma_j^2/2)]\ \sigma_j^2$, and the constant:

$$c = 4[E(D_{p+1,m}^2) - (\sigma_{p+1}^2/2)]\ \sigma_{p+1}^2. \tag{A6}$$

Since there are no natural units for the underlying dimensions, it is impossible to use this data base to estimate both A and Σ separately. The method estimates $A\Sigma A$ and the orientation of the coordinate system under the assumption that $A\Sigma A$ is diagonal. If not, the method rotates the coordinate system to match the eigenvectors of $A\Sigma A$. In the special case when $A\Sigma A$ is a scalar times the identity, the covariance matrix of the Y_{jm} is invariant under rotation. In order to estimate the components of A, assume that Σ is the identity matrix. This is not a bad working assumption given our lack of evidence about the distribution of the ideal points.

The covariance matrix (A5) can be rewritten as $\Lambda C\Lambda' + \Psi$, where Λ is the $px3$ matrix $(2\underset{\sim}{\theta}', \underset{\sim}{1})$, and C is the 3×3 diagonal matrix:

$$C = \begin{pmatrix} A^2 & 0 \\ 0 & c \end{pmatrix}$$

The *sample* covariance matrix of the Y_{jm} is then factored to give an estimate of $\Lambda C^{1/2}\Gamma$ where Γ is a 3×3 unknown orthogonal rotation. A factor of the sample covariance yields an estimate of Ψ as well as of $\Lambda C^{1/2}\Gamma$.

The next step is to estimate the rotation Γ and the matrix C by least squares fits. To simplify the notation let $\hat{M} = (\hat{m}_{jk})$ be the estimate of the $p \times 3$ matrix $\Lambda C^{1/2}\Gamma$ obtained from the factor analysis. Let $\hat{\Psi} = (\psi_{jk})$ denote the estimate of Ψ. The diagonal element $\hat{\psi}_{jj}$ estimates $4[E(D_{jm}^2) - (\sigma_j^2/2)]\sigma_j^2$. Using $\overline{D_j^2}$ as an estimate of $E(D_{jm}^2)$, the quadratic equation:

$$4[\overline{D_j^2} - (\hat{\sigma}_j^2/2)]\hat{\sigma}_j^2 = \hat{\psi}_{jj} \tag{A7}$$

is solved to obtain an estimate $\hat{\sigma}_j^2$ of σ_j^2. The smaller root $\overline{D_j^2} - [(\overline{D_j^2})^2 - (\hat{\psi}_{jj}/2)]^{1/2}$ is the correst estimate since the larger root would have $\hat{\sigma}_j^2 > \overline{D_j^2}$, which cannot hold in the limit as $M \to \infty$.

A brief note is in order at this point concerning the identifiability of Γ. Multiplying any row or column of Γ by -1 leaves it orthogonal. Because of this, together with the nature of the observations, and the method of constructing the regressions used to estimate Γ, the target map has certain ambiguities. These ambiguities consist of 90° and 180° rotations of all points and reflections across either axis. These ambiguities can be easily resolved if we have compass quadrant information for any target.

To estimate Γ, note that it can be written as the product of three orthogonal matrices $\Gamma = \Gamma_3\Gamma_2\Gamma_1$ where:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\delta_1 & -\sin\delta_1 \\ 0 & \sin\delta_1 & \cos\delta_1 \end{bmatrix} = \Gamma_1$$

$$\begin{bmatrix} \cos\delta_2 & 0 & -\sin\delta_2 \\ 0 & 1 & 0 \\ \sin\delta_2 & 0 & \cos\delta_2 \end{bmatrix} = \Gamma_2$$

$$\begin{bmatrix} \cos\delta_3 & -\sin\delta_3 & 0 \\ \sin\delta_3 & \cos\delta_3 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \Gamma_3. \tag{A8}$$

The first least squares fit estimates Γ_1 and Γ_2 by utilizing our knowledge that the third column of the matrix Λ is the vector $\underset{\sim}{1}$. This will simultaneously provide an estimate of c. The second fit estimates A and Γ_3 by using the $\overline{D_j^2} - \overline{D_{p+1}^2}$ as the dependent variable values.

To construct the first least squares fit, note that:

$$\Lambda C^{1/2}\Gamma = \Lambda C^{1/2}\Gamma_3\Gamma_2\Gamma_1$$

where Γ_3 transforms the matrix $\Lambda C^{1/2}$ but leaves the third column unchanged. Next Γ_2 acts on $\Lambda C^{1/2}\Gamma_3$ in a similar way, leaving the second column unchanged, and similarly for Γ_1. Thus since the third column of $\Lambda C^{1/2}$ is $\underset{\sim}{1}c^{1/2}$, we also know the third column of $\Lambda C^{1/2}\Gamma_3$ is $\underset{\sim}{1}c^{1/2}$. If we post-multiply our estimate of $\Lambda C^{1/2}\Gamma$ by $\Gamma_1'\Gamma_2'$, we obtain an estimate of $\Lambda C^{1/2}\Gamma_3$, the third column of which is $\underset{\sim}{1}c^{1/2}$. Since:

$$\Gamma_1'\Gamma_2' = \begin{matrix} \cos\delta_2 & 0 & \sin\delta_2 \\ -\sin\delta_1 \sin\delta_2 & \cos\delta_1 & \sin\delta_1 \cos\delta_2 \\ -\cos\delta_1 \sin\delta_2 & -\sin\delta_1 & \cos\delta_1 \cos\delta_2 \end{matrix}$$

the fact that the third column of $\Lambda C^{1/2}\Gamma = (m_{jk})$ is $\underset{\sim}{1}c^{1/2}$ yields the p linear equations:

$$1 = \beta_1 m_{j1} + \beta_2 m_{j2} + \beta_3 m_{j3} \qquad \text{(A9)}$$

where the parameters:

$$\beta_1 = c^{-1/2} \sin\delta_2$$

$$\beta_2 = c^{-1/2} \sin\delta_1 \cos\delta_2$$

$$\beta_3 = c^{-1/2} \cos\delta_1 \cos\delta_2$$

are estimated by fitting the p equations by least squares, using the estimates $\hat{m}_{jk}$ in place of the m_{jk}.

Having obtained the least squares estimates $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$, we estimate c by $\hat{c} = (\hat{\beta}_1^2 + \hat{\beta}_2^2 + \hat{\beta}_3^2)^{-1}$. Then $\hat{c}$ is used to estimate σ_{p+1}^2 with $j = p + 1$. Also, $\hat{\beta}_1\hat{c}^{1/2}$ estimates $\sin\delta_2$ and $\hat{\beta}_2/\hat{\beta}_3$ estimates $\tan\delta_1$. From these we obtain estimates of $\cos\delta_2$ and $\sin\delta_1$ through the identities $\cos\delta_1 = \pm(1 - \sin^2\delta^1)^{1/2}$ and $\sin\delta_2 = \pm(1 - \cos^2\delta_2)^{1/2}$. The choice of sign is arbitrary since the various choices represent reflection and 90° rotations in the estimated positions of the θ_j.

Once these quantities have been estimated, we have estimates $\hat{\Gamma}_1$ and $\hat{\Gamma}_2$ of Γ_1 and Γ_2, respectively. The matrix $\hat{M}\hat{\Gamma}_1'\hat{\Gamma}_2'$ is formed to estimate $\Lambda C^{1/2}\Gamma_3$. It then follows that, by deleting the third column which estimates $\underset{\sim}{1}c^{1/2}$, we have an estimate of $2\underset{\sim}{\theta}'AR$ where:

$$R = \begin{bmatrix} \cos\delta_3 & -\sin\delta_3 \\ \sin\delta_3 & \cos\delta_3 \end{bmatrix} \qquad \text{(A10)}$$

This estimate will now be used to estimate R and A.

Recalling (5) and (A1), we have:

$$\overline{D_j^2} - \overline{D_{p+1}^2} = \underset{\sim}{\theta}_j' A \underset{\sim}{\theta}_m - 2\underset{\sim}{\theta}_j' A \underset{\sim}{x}_j + \delta_j$$

where δ_j is an error term and $E\delta_j = \sigma_j^2 - \sigma_{p+1}^2$. We first correct for the nonzero mean in the error term by subtracting $\hat{\sigma}_j^2 - \hat{\sigma}_{p+1}^2$ and then use the resulting expression as the dependent variable in the regression, i.e.,

$$\overline{D_j^2} - \overline{D_{p+1}^2} - \hat{\sigma}_j^2 + \hat{\sigma}_{p+1}^2 = \underset{\sim}{\theta}_j' A \underset{\sim}{\theta}_j - 2\underset{\sim}{\theta}_j' A \underset{\sim}{x}_j + u_j \qquad \text{(A11)}$$

where, from the asymptotic theory of factor analysis, the errors u_j are of order $0(M^{-1/2})$.

Let $\hat{N}$ denote the estimate of $\theta' AR$. Since A is diagonal and $R^{-1} = R'$, with a small additive error we have:

$$\underset{\sim}{\theta}_j' A \underset{\sim}{\theta}_j = \sum_{i=1}^{2} a_i^{-1} \left(\sum_{h=1}^{2} \hat{n}_{jh} r_{ih} \right)^2 \tag{A12}$$

where $\hat{N} = (\hat{n}_{jh})$ and $R = (r_{ih})$. Moreover:

$$\theta_j' A x = \sum_{i=1}^{2} \sum_{h=1}^{2} \hat{n}_{jh} r_{ih} \bar{x}_i \tag{A13}$$

where the centroid $\underset{\sim}{\bar{x}} = (\bar{x}_1, \bar{x}_2)$, is unknown. Expanding and combining these equations, we have p equations:

$$\overline{D_j^2} - \overline{D_{p+1}^2} - \hat{\sigma}_j^2 + \hat{\sigma}_{p+1}^2 = \alpha_0 \hat{n}_{j1}^2 + \alpha_1 \hat{n}_{j2}^2 + \alpha_2 \hat{n}_{j1} \hat{n}_{j2} + \alpha_3 \hat{n}_{j1} + \alpha_4 \hat{n}_{j2} \tag{A14}$$

where, with a small additive error:

$$\begin{aligned}
\alpha_0 &= a_1^{-1} \cos^2 \delta + a_2^{-1} \sin^2 \delta \\
\alpha_1 &= a_1^{-1} \sin^2 \delta + a_2^{-1} \cos^2 \delta \\
\alpha_2 &= 2(a_2^{-1} - a_1^{-1}) \sin \delta \cos \delta \\
\alpha_3 &= -2(\bar{x}_1 \cos \delta + \bar{x}_2 \sin \delta) \\
\alpha_4 &= -2(-\bar{x}_1 \sin \delta + \bar{x}_2 \cos \delta).
\end{aligned} \tag{A15}$$

For notational convenience we have written $\delta = \delta_3$.

Fitting the p equations, we obtain estimates of the five α_i. We may then solve the equations in (A15) for the various parameters. Solving for δ, a_1^{-1}, and a_2^{-1} we obtain:

$$\tan 2\delta = \frac{\alpha_2}{\alpha_1 - \alpha_0} \tag{A16}$$

$$2a_1^{-1} = \alpha_1 + \alpha_0 - (\sin 2\delta)^{-1} \alpha_2$$

$$2a_1^{-1} = \alpha_1 + \alpha_0 + (\sin 2\delta)^{-1} \alpha_2$$

provided $\alpha_1 \neq \alpha_0$ and $\delta \neq 0$ or $\pm\pi/2$. These cases are covered by performing two tests immediately after the regression has been run, viz., the tests for $\alpha_2 = 0$ and for $\alpha_1 = \alpha_0$. The coordinates of the mean ideal point $(\bar{x}_1, \bar{x}_2)$ are estimated by combining α_3 and α_4 with the estimates of δ, a_1, and a_2 found from (A16).

Notes

1.

Hinich, M. J. and Ordeshook, P. C. "Abstentions and Equilibrium in the Electoral Process," *Public Choice 8*, 1969, pp. 81–100.

2.

The concepts underlying these models is presented on pages 323–326, Riker, W. and Ordershook, P. C. *An Introduction to Positive Political Theory*, Englewood Cliffs, N. J.: Prentice Hall, 1973.

3.

Downs, A. *An Economic Theory of Democracy*, New York: Harper and Row, 1957. Black, D. *The Theory of Committees and Elections*, Cambridge: Cambridge University Press, 1958.

4.

Davis, O. A. and Hinich, M. J. "A Mathematical Model of Policy Formation in a Democratic Society," in Berndt, J. L. (ed), *Mathematical Applications in Political Science II*, Dallas: S. M. U. Press, 1966.

5.

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

6.

An issue oriented rational voter who is faced with a choice among candidates for an executive office or a legislative seat must try to forecast how the candidates will function in office. It is rational for a voter to consider personality factors when choosing a representative. In a referendum election, on the other hand, voters only have to understand the issue and imagine the future consequences of the competing positions involved in the election. Elections for representatives require a greater cognitive effort and information investment from rational voters.

7.

This assumption about the utility functions implies that their indifference contours are ellipsoids whose axes are parallel to the spatial coordinate system. If the matrix $A(m)$ had non-zero elements off the diagonal, the indifference ellipsoids for the mth citizen are rotated with respect to the spatial coordinate system.

I have used a notation for A which indicates that the weights can be idiosyncratic. Most results in spatial theory, however, require that the weights are the same for all voters. This homogeneity assumption is relaxed somewhat in the following papers: Davis, O. A. and Hinich, M. J. "On the Power and Importance of the Mean Preference in a Mathematical Model of Democratic Choice," *Public Choice 5*, 1968, pp. 59–72, Davis, O. A. and Hinich, M. J. "Some Extensions to a Mathematical Model of Democratic Choice," in Lieberman, B. (ed) *Social Choice*, New York: Gordon and Breach, 1971, pp. 320–347.

8.

The probablistic choice model is discussed in detail in the paper Hinich, M. J., Ledyard, J. O., and Ordeshook, P. C. "A Theory of Electoral Equilibrium: A Spatial Analysis Based on the Theory of Games," *Journal of Politics*, 35, 1973, pp. 154–193.

9.

This form for abstention was called cross pressures in the original papers.

10.

See the discussion about this point in Section 6.2, Cahoon, Lawrence S. "Locating a Set of Points Using Range Information Only," unpublished PhD dissertation in statistics, Carnegie-Mellon University, July, 1975. The twelve politicians are Nixon, Humprey, Agnew, Muskie, Reagan, Romney, Rockefeller, McCarthy, Robert Kennedy, Johnson, Wallace, and LeMay.

11.

Although there is a lot of noise in the data, several empirical studies have shown that the spatial assumption for thermometer scores makes sense in the aggregate. Building on the work of Rusk and Weisberg, George Rabinowitz explicitely used the geometry inherent in the spatial assumption to "map" the politicians and voters in a joint space. Rusk, J. G. and Weisberg, H., "Perceptions of Presidential Candidates," *Midwest Journal of Political Science 16*, 1972, pp. 338–410. Rabinowitz, G. B. *Spatial Models of Electoral Choice*, Chapel-Hill: University of North Carolina Monograph, 1974.

The basic assumptions of spatial voting have been tested by Cahoon, Hinich, and Ordeshook using the thermometer scores from both the 1968 and 1972 election surveys. The 1968 results are contained in the working paper: (Cahoon, L. S., Hinich, M. J., and Ordeshook, P. C. "A Multidimensional Statistical Procedure for Spatial Analysis," V.P.I., 1975.

12.

The numbers in the table are rounded to the first decimal place.

13.

The Welch treatment of the Bahrens-Fisher problem concerning the two sample t-test with unequal variances is given on pages 300–301 of *Statistical Theory and Methodology* by Brownlee, K. A. New York: Wiley, 2nd edition, 1965.

14.

Page, B. and Brody, R. A. "Indifference, Alienation and Rational Decisions: The Effects of Candidate Evaluations on Turnout and the Vote," *Public Choice 15*, 1973, pp. 1–17. They do not discuss how they handled the Wallace voters. Also see, Kelley, S. Jr. and Mirer, T. W. "The Simple Act of Voting," *American Political Science Review 68* (2), 1974, pp. 272–291. Their work supports the indifference hypothesis.

15.

Details of this technique are given in the working paper, "A Multidimensional Statistical Analysis." This working paper can be obtained from Hinich at V.P.I. or Ordeshook at Carnegie-Mellon University. The paper presents the mathematics of the procedure, an analysis of its statistical properties using artifical data, and also presents the results of the use of the method for the subsample scores for the 1968 survey. The mathematic statistics of the method is presented in full detail in Cahoon's PhD thesis, op. cit.

16.

A valence issue is defined to be an issue for which all citizens have the same ideal point. See Stokes, D., "Spatial Models of Party Competition," *American Political Science Review* 54, 1963, pp. 368–377.