Public Choice **78**: 155–169, 1994.

A test of the predictive dimensions model in spatial voting theory*

JAMES M. ENELOW
MELVIN J. HINICH
Department of Government, University of Texas at Austin, Burdine Hall 536, Austin, TX 78212-1087

Accepted 20 April 1992

Abstract. We model correlated voter-candidate issue data within the framework of the Enelow-Hinich spatial model of predictive dimensions. The empirical consequences of this model of the issue data are surprising and allow for an indirect test of the Enelow-Hinich spatial model. The central prediction of the correlated data model we construct, which depends critically on the underlying spatial model, is tested with issue data from the 1980 NES pre-election interview. The test results are highly supportive of the model's predictions. We conclude both that the spatial model of predictive dimensions is empirically supported and that candidate spatial locations estimated by the model are not an artifact of correlated voter-candidate issue data.

1. Introduction

It has long been noted that significant correlations exist between voter self-placements on public policy issues and voter perceptions of where politicians stand on those same issues (Brody and Page, 1972; Page and Brody, 1972; Page and Jones, 1979). The precise consequences of these correlations for issue voting models are not obvious, nor is it clear how to incorporate this phenomenon into the spatial model of elections. There exist a variety of scaling techniques that use issue data from voter surveys to recover the spatial locations of voters and candidates in some type of reduced policy space. What are the consequences of correlated voter-candidate issue data for the estimated spatial locations that these techniques provide?

In this article, we devise a test which provides strong support for the Enelow-

* The authors wish to thank Ben Page, Eric Devereux, and an anonymous referee for their helpful comments. An earlier draft of the paper was delivered at the 1989 Public Choice Society Meetings, Orlando, Florida. This work was partially supported by the National Science Foundation under Grant # SES 8310591.

Hinich spatial model of predictive dimensions and which at the same time demonstrates that the candidate spatial locations estimated by the Enelow-Hinich model are independent of any correlation between the issue preferences of the voters and the perceived issue positions of the candidates. We thus conclude that estimated candidate spatial locations need not be an artifact of such correlations.

The spatial model developed by Enelow and Hinich (1984) postulates that the voter's perception of a candidate's issue position is a linear function of a fixed set of positions on a small set of underlying predictive dimensions with coefficients that vary by voter. Building on this model, we construct a submodel of these coefficients that allows them to depend on voter issue preferences.

The empirical consequences of this model of the issue data are surprising and allow for an indirect test of the Enelow-Hinich spatial model. The central prediction of the correlated data model we construct is that a factor analysis of the covariance matrix of voter and candidate issue variables yields the same estimated candidate factor loadings as a factor analysis of the candidate variables alone. This prediction depends critically on our model of the data, based on the Enelow-Hinich spatial model. The correlated data model also predicts that the variable representing voter issue positions lies in the same underlying space as the candidate variables. Again, this prediction depends critically on our model of the data and the Enelow-Hinich spatial model.

These predictions can be easily tested. Using the 1980 NES pre-election issue data, we factor analyze the candidate and voter variables together and then the candidate issue variables separately to see whether the inclusion of the extra variable representing the issue positions of the voter affects the candidate loadings in the underlying policy space recovered by factor analysis and, in addition, whether this extra variable lies in the same space as the candidate variables, instead of creating an extra dimension.

Our tests confirm these theoretical predictions. We conclude that the Enelow-Hinich specification of voter issue perceptions is supported empirically, that correlations between voter self-placements and perceived candidate issue positions are correctly measured by the linear coefficients of the model and that valid estimates of candidate positions in a reduced policy space can be obtained from voter survey data, regardless of any correlations between voter preferences and perceived candidate issue positions.

2. The model

For the reader unfamiliar with the Enelow-Hinich spatial model of underlying predictive dimensions (Enelow and Hinich, 1984), we begin by setting out the model's basic elements as simply as possible for the case of two underlying policy dimensions. We then construct a model of correlated voter and candidate issue positions within the framework of the Enelow-Hinich model and develop the consequences of this model of the issue data for estimates of candidate and voter locations in the underlying policy space recovered by factor analysis.

The Enelow-Hinich model assumes that electoral competition can be analyzed in either of two spaces: the candidate space, on which we have data, and an unobservable, underlying predictive or factor space, which is a condensation of the candidate variables. The data consist of perceived issue positions of the candidates, where perceptions vary over voters and issues. For simplicity, we will develop the model for a single issue scale, which means that if there are n voters who can locate j candidates on this issue scale, there are n cases for each of j variables. Issues and voters are sampling units; the variables are the candidates. If observations are given over m issues for n voters, we would have mn cases for each of these same j candidate variables.

For a single voter, i, let $C_i = (c_{i1}, \ldots, c_{ij})^T$ be a j x 1 vector, denoting where voter i perceives candidates $1, \ldots, j$ to be located on a single policy issue. Also let $V_i = (v_{i1}, v_{i2})^T$ be a 2 x 1 vector measuring the slope coefficients in a linear model that expresses the connection that voter i makes between candidate positions on two underlying predictive dimensions and candidate positions on the given policy issue. Finally, let $P = (P_1, P_2)$ be a j x 2 matrix of positions for the j candidates on the two underlying dimensions, where $P_1 = (p_{11}, \ldots, p_{j1})^T$ and $P_2 = (p_{12}, \ldots, p_{j2})^T$. Then, the Enelow-Hinich model (with an error term representing lack of fit of the model) postulates that candidate positions on a given policy issue are related to candidate positions on the predictive dimensions of the campaign by the identity

$$C_i = P\,V_i + e_i \tag{1}$$

where e_i is a j x 1 vector of random errors that are distributed *independently of V_i and voter i's ideal point x_i* on the given policy issue. Assume that the covariance matrix of e_i is diagonal and denote it by D. As equation (1) shows, the model assumes that differences in voter perceptions of candidate issue positions occur only through the coefficients V_i.

To allow for perceptual variation across voters of the issue position of some candidate r (if one exists) for whom $p_{r1} = p_{r2} = 0$, we may have written equation (1) as $C_i = b_i + P\,V_i + e_i$, where b_i is a j x 1 vector of idiosyncratic intercept terms and $b_{ir} + e_{ir}$ is i's perception of r's issue position. For simplicity of exposition, we omit b_i from the model, but it is easily included if we define $Z_i = C_i - b_i$, in which case Z_i can be substituted for C_i in the following analysis.

Intuitively speaking, the Enelow-Hinich model assumes that the voter uses the candidate's predictive label as a shorthand device to estimate the candidate's positions on the issue of the campaign. The reasons for assuming that voters take this shortcut are straightforward. As Downs (1957) states, it is rational for many voters not to invest in acquiring information about the issue positions of the candidates. The theory of predictive dimensions provides a means by which voters can acquire such information at a low cost. Given a candidate label, such as "Reagan Republican", the voter can make inferences about the candidate's issue stands without expending the resources to observe them directly. The connections the voter makes between candidate labels and candidate issue positions come from his own understanding of politics. The predictive dimensions model formalizes the nature of this connection by assuming that the voter, as statistican, uses "least-squares" assumptions to structure his knowledge of politics.

Substantively speaking, v_{ik} represents i's perception of the issue difference between two candidates who are one unit apart on the k^{th} predictive dimension. The size of v_{ik} depends on the issue-significance of differences in candidate predictive labels. A voter who believes that there isn't a "dime's worth of difference" between Republicans and Democrats will have a v_{ik} much closer to zero than a voter who sees major issue differences between the two parties.

To give the reader some idea of how the model of equation (1) is estimated, the vector of observations C_i is given by the issue data (in this paper obtained from the 1980 NES pre-election seven-point survey questions), P is the matrix of estimated loadings obtained from a factor analysis of the sample covariance or correlation matrix constructed from the C_i of all the voters with nonmissing data, and V_i are the factor scores derived from P, the sample covariance or correlation matrix, and C_i. Further details will be provided in the data analysis section.

We now build an explicit model of correlated voter-candidate issue data within the Enelow-Hinich model outlined above by assuming that

$$V_i = V' + W_i(x_i) \qquad (2)$$

where $V' = (v_1, v_2)^T$ is the mean voter linkage between candidate positions on the two predictive dimensions and candidate positions on the policy issue in question. $W_i(x_i) = (w_{i1}(x_i), w_{i2}(x_i))^T$ is a 2 x 1 vector of idiosyncratic differences that depend on x_i and whose functional form varies over voters. W_i measures the degree of dependence between v_{ik} and x_i for $k = 1, 2$. Since $E(V_i) = V'$, $W_i(x_i)$ measures the variation of voter linkages around the mean linkage between dimensions and issues. This mean linkage may or may not be "true" in some objective sense.

The model of equation (2) is deliberately general so as not to tie our results to a specific functional form between voter and perceived candidate issue positions. A linear specification of this relationship is $v_{ik} = v_k + w_{ik}x_i$. In this case, w_{ik} is the slope coefficient measuring the strength and sign of the relationship between x_i and v_{ik} ($k = 1, 2$). In Enelow (1986), it was found that the average v_{ik} varied systematically across electoral subgroups for most of the issues included in the 1980 NES survey. For a fixed dimensional difference between Carter and Reagan, Democrats felt, on average, that Reagan would increase defense spending by a smaller amount relative to Carter than Republicans believed would occur. On average, Republicans desired more defense spending than Democrats. Thus, a reasonable hypothesis for these data is that v_{ik} ($k = 1, 2$) is linear in x_i and that the averge w_{ik} is significantly different between Democrats and Republicans.

An alternative hypothesis is that $w_{ik}(x_i)$ is *nonlinear* in x_i. Suppose, for example, that $v_{ik} = v_k + w_{ik}x_i^2$, where w_{ik} is the coefficient of the quadratic term. If this coefficient is positive, then v_{ik} increases at an increasing rate as x_i increases. Individuals whose perceptions fit this model attribute a greater and greater issue difference to the same pair of candidates as their own most preferred policy becomes more extreme. To give an example, the policy difference attributed to Ted Kennedy and Ronald Reagan on the issue of abortion may be nonlinear in the voter's own most preferred abortion policy, with this policy difference being exponentially magnified by voters with increasingly extreme views on the subject. In such an event, a significant linear correlation between x_i and v_{ik} may exist even though the relationship between w_{ik} and x_i is nonlinear.

From equation (1), $c_{ij} = p_{j1}v_{i1} + p_{j2}v_{i2} + e_{ij}$; so, substituting for v_{i1} and v_{i2} from equation (2), we have that

$$c_{ij} = a_j + p_{j1}w_{i1}(x_i) + p_{j2}w_{i2}(x_i) + e_{ij} \tag{3}$$

where $a_j = p_{j1}v_1 + p_{j2}v_2$. Equation (3) re-expresses our model of the data to make it clear that the variation in perceptions of candidate j's issue stands are a direct function of the w_{ik}'s of the voters.

Substituting for V_i, equation (1) can now be expressed as

$$C_i = P\ V' + P\ W_i(x_i) + e_i. \tag{4}$$

In order to compute the covariance matrix of C_i, we need to assume some structure on the v_i's, keeping in mind that for simplicity we are developing the model for the case of a single issue. From equation (2), we have that $v_{ik} = v_k + w_{ik}(x_i)$, where $k = 1, 2$ indexes the underlying dimensions and i the voters. Then, we make the following two identification assumptions:

(1) $w_{i1}(x_i)$ and $w_{i2}(x_i)$ are uncorrelated, i.e. $E[w_{i1}(x_i)w_{i2}(x_i)] = 0$.
(2) $E\{[w_{ik}(x_i)]^2\}$ over i is constant for all k. Since units are

arbitrary, set the variance of $w_{ik}(x_i)$ equal to one.

We do not have a substantive theory about the structure of the w_{ik}'s, and so we cannot make an a priori specification of the covariance matrix of these terms. The correlation between $w_{i1}(x_i)$ and $w_{i2}(x_i)$ is unknown and can either be negative or positive. Standard practice in factor analysis (Morrison, 1976: 262; Anderson, 1989: 553) is to identify the model by assuming that the correlation is zero. Such an approach is consistent with confirmatory analysis (Long, 1983: 40–41) given that we do not have a causal model of the w_{ik}'s.

The covariance matrix of v_{ik} is the same as for $v_{ik} - \dot{v}_k$. Thus, the covariance matrix of v_{ik} is the same as for $w_{ik}(x_i)$. Given assumptions (1) and (2) and recalling from equation (2) that $E[w_{ik}(x_i)] = 0$; the covariance matrix of the v_{ik}'s is the 2 x 2 identity matrix. Given this result *and the independence of* e_{ik} *and* $w_{ik}(x_i)$, we can then express the covariance matrix of C_i as

$$E[(PW_i(x_i) + e_i)(PW_i(x_i) + e_i)^T] = PE[W_i(x_i)W_i(x_i)^T]P^T + D = PP^T + D, \tag{5}$$

where $D = E(e_ie_i^T)$ is a diagonal matrix of specific variances. The covariance of $(x_iC_i^T)$ is

$$E[(x_i - x_m)(PW_i(x_i) + e_i)^T] = HP^T, \tag{6}$$

where $x_m = E(x_i)$ and H is a 1 x 2 vector of covariances

$$H = E[(x_i - x_m)(W_i(x_i))^T] = [\mathrm{cov}(x_i, w_{i1}(x_i)), \mathrm{cov}(x_i, w_{i2}(x_i))] \tag{7}$$

Thus, the covariance matrix of (C_i, x_i) =

$$\begin{bmatrix} P\,P^T & P\,H^T \\ H\,P^T & H\,H^T \end{bmatrix} + R + D' = (P^T, H^T)^T(P^T, H^T) + R + D', \tag{8}$$

where D' is a $(j+1)$ x $(j+1)$ matrix whose j x j submatrix is D and whose $(j+1)$st row and $(j+1)$st column are vectors of 0's.

The matrix R is $(j+1)$ x $(j+1)$ with only 0's except the element in the $(j+1)$st row and $(j+1)$st column which is $\mathrm{var}(x_i) - H\,H^T$. This $(j+1)$st term, which is the specific variance of x_i, is $\mathrm{var}(x_i) - [\mathrm{cov}^2(x_i, w_{i1}(x_i)) + \mathrm{cov}^2(x_i, w_{i2}(x_i))]$. This identity provides an internal validity check on our model, since we have two estimates of the specific variance of x_i: one directly from the factor analysis of (C_i, x_i) and the other from this identity. The sample variance of x_i estimates $\mathrm{var}(x_i)$. From (2), $\mathrm{cov}(x_i, w_{ik}(x_i)) = \mathrm{cov}(x_i, v_{ik})$, so the sample covariance of x_i and the computed scores on each factor estimate this quantity. If we used scores from the factor analysis of (C_i, x_i), the identity between the estimated specific variance of x_i and the theoretical predictor of this specific variance would necessarily hold. By using factor scores from a factor analysis of the candidate variables alone, the x_i's do not directly affect the computation of the scores. With these "pure" scores, the difference between the estimated specific variance of x_i and $\text{var}(x_i) - H\,H^T$ measures the degree of fit between our data set and our model of the issue data.

The first term in (8) is the factorization of the $(j+1) \times (j+1)$ covariance matrix of (C_i, x_i), yielding a $(j + 1) \times 2$ loading matrix which is a consistent estimator of $(P^T, H^T)^T$ up to an orthogonal rotation. The first j rows of this matrix estimate P, the loadings of the j candidates on the two predictive dimensions. The $(j + 1)$st row estimates H, the loadings of the voter ideal point variable on these same two dimensions. *In general, the j x 2 submatrix of candidate loadings will not be the same as the candidate loadings derived from a factorization of C_i.* The structure of the covariance matrix of (C_i, x_i) will not be the same as for C_i. The model given by equation (2) provides the simple relationship between the factorization of C_i and that of (C_i, x_i) that we derive, a relationship not expected by chance. The R and D' matrices are diagonal and get estimated by the specific variance matrix. The loadings matrix is not rotated.

Let us put in words the basic result of equation (8). If our model of the data specified by equations (1) and (2) is correct, we will obtain the same candidate loadings when we factor analyze the j candidate variables as when we factor analyze the $j+1$ candidates plus voter variables. If, on the other hand, the voter ideal points were correlated with the model errors e_i, then these two factor analyses would *not* yield the same estimated candidate loadings. We will also obtain different candidate loadings if P were a random variable correlated with x_i.

Another result is worth noting. Since the factor space for the $j + 1$ variables (C_i, x_i) has the same number of dimensions as that for the j variables C_i, the extra variable x_i lies in the same space as C_i. The addition of the voter ideal point variable does not create a new dimension in the underlying space.

To reiterate, we have constructed a model of functionally related candidate and voter issue positions that leads to several verifiable results. Given data on voter perceptions of the issue positions of a set of candidates and voter self-placements on these same issue scales, a factor analysis of j candidate variables and a factor analysis of $j + 1$ candidate plus voter variables should yield the same candidate loadings on the underlying factors. The extra voter variable should lie in the same factor space as the j candidate variables. We now proceed to test these propositions.

3. Data analysis

The data we use to test our theoretical propositions are survey responses to the five issue scale questions contained in the pre-election wave of the 1980 NES Pre-Post interview. These scales concern the issues of defense spending, government services, inflation/unemployment, abortion, and a tax cut. The first three scales are seven points in length, the abortion scale has four points, and the tax cut scale five points. To increase comparability across scales, the abortion and tax cut scales were lengthened to make the end points 1 and 7 with equal spacing in between.

Issue data exist on seven candidates: Carter, Reagan, Anderson, Kennedy, the Republican party, the Democratic party, and the Federal government. To create an idiosyncratic intercept term for the linear function expressed by equation (1), the Federal government variable is subtracted from each of the other variables, leaving us with six candidate variables. An observation on a single variable is a respondent's perception of where a candidate (or the respondent) is located on a given issue relative to what the respondent perceives as the position of the Federal government. There are 1614 respondents in the pre-election wave of the 1980 Study. If observations range over respondents and over issues (five), a maximum of 8070 cases exist on each variable. If observations range only over respondents, the maximum number of cases is reduced to 1614.

To establish the empirical plausibility of equation (2), we began our data analysis by computing sample correlations between respondent self-placements on a single issue scale and factor scores on the same issue. As explained in Enelow (1986), the factor scores calculated for each case estiamte the v_i's in the Enelow-Hinich model. More specifically, if cases consist of respondents' perceptions of candidate positions on a single issue, then v_{i1} and v_{i2} are estimated by the factor scores computed for case i on factors 1 and 2. Note that the respondents themselves are not part of this factor analysis, so that the v_i's are not partially computed from the x_i's. For a full discussion of the estimation of factor scores for both fixed and random factor models, see Anderson (1984).

Table 1 reports ten bivariate correlations between the respondent's self-placement on one of the five pre-election issue scales and each of the two factor scores computed for the respondent on the same issue. In theoretical terms, we are examining the correlations between x_i and v_{i1} and x_i and v_{i2} for a single issue. The w_{ik}'s of equation (2) determine the strength of this correlation. As shown in Table 1, the sample correlations are all significantly positive between self-placement and first factor scores. The sample correlations with the second factor scores are weaker.

An explanation for this result is suggested by the relative loadings of the candidates on the two derived factors. On factor 1, Carter and Reagan load posi-

Table 1. Sample correlations between respondent self-placement and factor score on five issues

	All respondents	Carter supporters	Reagan supporters
F1D	.67	.57	.81
	(N = 526)	(N = 196)	(N = 227)
	(p = .00)	(p = .00)	(p = .00)
F2D	.14	.46	−.08
	(N = 526)	(N = 196)	(N = 227)
	(p. = 00)	(p = .00)	(p = .12)
F1G	.60	.39	.82
	(N = 545)	(N = 198)	(N = 221)
	(p = .00)	(p = .00)	(p = .00)
F2G	.17	.60	−.27
	N = 545)	(N = 198)	(N = 221)
	(p = .00)	(p = .00)	(p = .00)
F1I	.67	.40	.87
	(N = 387)	(N = 153)	(N = 152)
	(p = .00)	(p = .00)	(p = .00)
F2I	.04	.45	−.20
	(N = 387)	(N = 153)	(N = 152)
	(p = .22)	(p = .00)	(p = .01)
F1A	.35	.43	.30
	(N = 306)	(N = 113)	(N = 122)
	(p = .00)	(p = .00)	(p = .00)
F2A	.08	.32	−.33
	(N = 306)	(N = 133)	(N = 122)
	(p = .08)	(p = .00)	(p = .00)
F1T	.48	.17	.69
	(N = 242)	(N = 86)	(N = 109)
	(p = .00)	(p = .06)	(p = .00)
F2T	.00	.32	−.10
	(N = 242)	(N = 86)	(N = 109)
	(p = .48)	(p = .00)	(p = .14)

Note:
F1D = score on first factor for respondent-defense cases.
F2D = score on second factor for respondent-defense cases.
F1G, F2G, F1I, F2I, F1A, F2A, F1T, F2T defined identically for govt. services (G), inflation/unemp. (I), abortion (A), and tax cut (T).

tively with Reagan more positive than Carter. On factor 2, Reagan loads negatively and Carter positively. For a Reagan supporter to positively correlate his issue position with Reagan's, the second factor score must be negative. Thus, we expect a negative correlation between second factor scores and self-placement for Reagan supporters and a positive correlation between second factor scores and self-placement for Carter supporters. Table 1 also reports these correlations, and our prediction is largely confirmed.

So far, then, we find empirical support for the model of the data postulated

Table 2. Sample correlation matrix and principal results of factor analysis of candidate variables and of voter plus candidate variables based on 1980 pre-election issue data

	Sample correlation matrix						
	C	R	A	K	RE	DE	V
Carter	1.00	.35	.45	.52	.36	.70	.35
Reagan		1.00	.58	.31	.90	.35	.63
Anderson			1.00	.49	.59	.50	.53
Kennedy				1.00	.32	.62	.34
Rep Party					1.00	.34	.62
Dem Party						1.00	.39
Voter							1.00

	Unrotated loadings from candidate factor analysis		Unrotated loadings from candidate and voter factor analysis	
Variable	Factor 1	Factor 2	Factor 1	Factor 2
Carter	.50	.59	.50	.58
Reagan	.93	−.18	.93	−.19
Anderson	.67	.22	.68	.22
Kennedy	.45	.53	.45	.53
Rep Party	.93	−.19	.93	−.19
Dem Party	.52	.73	.52	.73
Voter	–	–	.68	.05

	Eigenvalues of initial factor matrix	
Factor	Candidate variables	Cand plus voter variables
1	3.46	3.94
2	1.23	1.30
3	.51	.51
4	.41	.46
5	.28	.41
6	.10	.28
7	–	.10

in equation (2). We now turn to a direct test of our major theoretical proposition, namely that a factor analysis of candidate variables yields the same factor loadings for the candidates as a factor analysis of candidate variables with a variable for the voters' ideal points also included. In addition, we expect the voter ideal point variable to lie in the same factor space as the candidate variables.

Table 2 reports the results of two factor analyses, one of the candidate variables and one of the candidate plus voter variables. Observations range across

Table 3. Eigenvalues and factor loadings from factor analysis of six and six plus one variables based on a scrambled matrix of 1980 pre-election issue data

Eigenvalues of initial factor matrix

Factor	Six variables	Six plus one variables
1	1.71	1.71
2	1.23	1.57
3	.99	.99
4	.83	.95
5	.70	.71
6	.53	.56
7	–	.50

	Unrotated loadings from six var factor analysis		Unrotated loadings from six plus one var factor analysis	
Variable	Factor 1	Factor 2	Factor 1	Factor 2
one	–.26	.31	.37	–.25
two	.16	.33	.18	.13
three	.49	.34	.08	.42
four	.78	–.03	–.01	.88
five	–.41	.31	.08	–.37
six	.15	.10	.28	.15
seven	–	–	.94	.01

respondents and issues. After listwise deletion of missing data, we are left with 2006 cases. Table 2 also reports the sample correlation matrix of the candidate and voter variables. The method of factoring used is Rao's (1955) canonical method which is asymptotically equivalent to maximum likelihood factor analysis.

The robustness of maximum likelihood factor analysis has been demonstrated under a variety of conditions. Anderson (Sec. 14.3.3, 1984) shows that if the observed data are (multivariate) normally distributed, the maximum likelihood estimator yields a consistent estimator of the loading matrix whose large sample standard errors are proportional to $N^{-1/2}$. Anderson and Amemiya (1988) extend this result to show that the asymptotic properties of maximum likelihood factor analysis are maintained when the normality assumption is violated, as long as the distributions of the factor vector and the error vector have finite second moments.

The similarity between the two sets of loadings is striking. As predicted by the model, the candidate loadings exhibit nonsignificant differences in the two factor analyses. In our case, the proportionality constant for the standard

Table 4. Results of factor analysis of candidate variables and of voter plus candidate variables based on 1980 pre-election issue data factor analyzed one issue at a time

		Unrotated loadings from candidate factor analysis		Unrotated loadings from candidate and voter factor analysis	
Variable	Issue	Factor 1	Factor 2	Factor 1	Factor 2
Carter	Defense	.36	.66	.38	.65
Reagan	Defense	.89	−.17	.89	−.21
Anderson	Defense	.70	.19	.71	.17
Kennedy	Defense	.43	.49	.45	.48
Rep Party	Defense	.92	−.24	.90	−.26
Dem Party	Defense	.39	.82	.41	.80
Voter	Defense	–	–	.70	.14
Carter	Govt serv	.43	.55	.45	.54
Reagan	Govt serv	.89	−.17	.89	−.20
Anderson	Govt serv	.64	.28	.64	.26
Kennedy	Govt serv	.38	.63	.39	.61
Rep Party	Govt serv	.91	−.17	.90	−.19
Dem Party	Govt serv	.41	.67	.44	.66
Voter	Govt serv	–	–	.63	.19
Carter	Inf/unemp	.53	.49	.53	.48
Reagan	Inf/unemp	.90	−.22	.90	−.24
Anderson	Inf/unemp	.79	.20	.78	.19
Kennedy	Inf/unemp	.48	.56	.48	.55
Rep Party	Inf/unemp	.90	−.24	.90	−.24
Dem Party	Inf/unemp	.50	.72	.51	.72
Voter	Inf/unemp	–	–	.70	.05
Carter	Abortion	.72	.22	.71	.24
Reagan	Abortion	.70	−.50	.72	−.50
Anderson	Abortion	.62	.25	.63	.27
Kennedy	Abortion	.62	.20	.61	.22
Rep Party	Abortion	.71	−.50	.72	−.47
Dem Party	Abortion	.82	.40	.81	.41
Voter	Abortion	–	–	.37	.07
Carter	Tax cut	.54	.61	.54	.62
Reagan	Tax cut	.88	−.25	.88	−.24
Anderson	Tax cut	.57	.09	.57	.09
Kennedy	Tax cut	.44	.35	.44	.36
Rep Party	Tax cut	.87	−.28	.87	−.28
Dem Party	Tax cut	.54	.67	.54	.66
Voter	Tax cut	–	–	.50	.01

Note: The variable-issue pair Carter-Defense specifies that cases consist of respondent perceptions of Carter's position on that issue. Other variable-issue pairs are defined similarly. There is one factor analysis per variable set for a given issue, yielding five factor analyses for the candidate variables and five factor analyses for the voter plus candidate variables.

errors is of the order of one. For the results of Table 2, N = 2006, so the large sample standard deviation is approximately .022. Inspection of Table 2 shows that the two sets of loadings differ by no more than .01. Since our data are quantized between 1 and 7, the Anderson and Amemiya large sample results apply.

Also as predicted, the voter variable does not create an extra dimension in the factor solution. This point is established more precisely by comparing the two sets of eigenvalues of the two initial factor matrices.

At this point, it is useful to demonstrate empirically that it is unexpected to find the estimated loadings of the candidate variables unchanged by the inclusion of an additional variable. To establish this point, we employed a partial resampling method, randomly permuting the rows and columns of the data matrix (1614 x 1 values at a time) on which the results of Table 2 are based. We then ran two factor analyses, one on the first six variables (columns) and one on all seven variables in this scrambled matrix. Table 3 reports the results. The loadings of the first six variables are significantly different in the two factor analyses.

To test the robustness of the results in Table 2, we ran separate factor analyses on each of the five issues taken separately. The maximum number of cases is now reduced from 8070 to 1614. With listwise deletion of missing data, the sum of the cases across the five issues (N(Defense) = 526; N(Govt. Services) = 545; N(Infla./Unemp.) = 387; N(Abortion) = 306; N(Tax Cut) = 242) equals 2006. Table 4 reports the results of five factor analyses of (C_i, x_i) and five factor analyses of C_i.

Examining thirty pairwise comparisons of 1 x 2 candidate loading vectors, we conclude that the results in Table 2, based on averaging across issues and voters, are not due to chance variation. Adding a voter ideal point variable to the factorization has no significant effect on the loadings of the candidate variables. Furthermore, the voter variable does not add an extra factor to any of the solutions.

Finally, Table 5 reports the predicted vs. the actual estimates of the specific variance of the voter ideal point variable for the five issue-at-a-time factor analyses. As explained in the previous section, $\mathrm{var}(x_i) - H\, H^T = \mathrm{var}(x_i) - [\mathrm{cov}^2(x_i, w_{i1}(x_i)) + \mathrm{cov}^2(x_i, w_{i2}(x_i))] = \mathrm{var}(x_i) - [\mathrm{cov}^2(x_i, v_{i1}) + \mathrm{cov}^2(x_i, v_{i2})]$ is a theoretical predictor of the specific variance of x_i, where v_{i1} and v_{i2} are factor scores computed from a factor analysis of the candidate variables alone. As Table 5 shows, this theoretical predictor provides a close fit with the specific variance estimated from a factor analysis of the voter and candidate variables. These results provide further evidence of the empirical validity of our model of correlated voter-candidate issue data.

Table 5. Theoretical vs. actual estimates of specific variance of voter ideal point variable (as proportion of total variance)

Issue	Est specific variance	Var (x_i) − H H^T
Defense	.50	.56
Govt serv	.57	.65
Inf/unemp	.51	.57
Abortion	.86	.88
Tax cut	.75	.79

Note: The issue specifies that cases consist of individual perceptions of candidate and own position on the given issue. For example, Defense denotes a factor analysis of seven variables, six candidate variables and the variable "voter", where each variable consists of perceptions of where a candidate or oneself stands on the issue of defense spending. H is estimated from factor scores that derive from an analysis of only the candidate variables. For a fuller explanation, see text.

4. Conclusions

We have found that the Enelow-Hinich spatial model is supported by an unusual empirical test. Postulating that voter-candidate issue correlations are measured by the translation coefficients of a linear spatial model leads to a strong and unexpected empirical prediction; that the estimated candidate loadings obtained from a factor analysis of the candidate variables are unchanged when a voter ideal point variable is added to the analysis. This result does not hold if the voter ideal points are correlated with the model errors or with any other systematic variation in the data not captured by the v_{ik}'s in the model.

Futhermore, the independence of the candidate loadings from voter policy preferences means that the estimated candidate map is not an artifact of voter preferences. This is a reassuring result, since it implies that the Enelow-Hinich conception of a stable set of candidate positions on an underlying set of dimensions is consistent with the issue data.

Voters are frequently confused about a candidate's issue stands. This is not surprising since the average voter does not find it rational to spend scarce resources learning a great deal about the candidates competing for his vote. Given the disincentives to gather issue information about the candidates, we should not be surprised to find correlated voter-candidate issue data. What we have shown is that these correlations are completely consistent with a spatial model in which candidates have fixed, stable locations on a set of underlying predictive dimensions.

References

Anderson, T.W. (1984). *An introduction to multivariate statistical analysis.* Second Edition. New York: Wiley.

Anderson, T.W. and Amemiya, Y. (1988). The asymptotic normal distributions of estimators in factor analysis under general conditions. *The Annals of Statistics* 16(2): 759–771.

Brody, R. and Page, B. (1972). Comment: The assessment of policy voting. *American Political Science Review* 66: 450–458.

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

Enelow, J. (1986). Measuring the linkage between predictive dimensions and candidate issue positions: An examination of group differences. *Political Behavior* 8(3): 245–261.

Enelow, J. and Hinich, M. (1984). *The spatial theory of voting: An introduction.* New York: Cambridge University Press.

Long, J.S. (1983). *Confirmatory factor analysis.* Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-033. Beverly Hills and London: Sage Publications.

Morrison, D. (1976). *Multivariate statistical methods.* Second Edition. New York: McGraw-Hill.

Page, B. and Brody, R. (1972). Policy voting and the electoral process: The Vietnam War issue, *American Political Science Review* 66: 389–400.

Page, B. and Jones, C. (1979). Reciprocal effects of policy preferences, party loyalties and the vote. *American Political Science Review* 73: 1071–1089.

Rao, C.R. (1955). Estimation and tests of significance in factor analysis. *Psychometrika* 20: 93–111.

Tukey, J. (1977). *Exploratory data analysis.* Reading MA: Addison-Wesley.