A Method of Scaling with Applications to the 1968 and 1972 Presidential Elections

John H. Aldrich; Richard D. McKelvey

Stable URL:
http://links.jstor.org/sici?sici=0003-0554%28197703%2971%3A1%3C111%3AAMOSWA%3E2.0.CO%3B2-R

The American Political Science Review is currently published by American Political Science Association.

A Method of Scaling with Applications to the 1968 and 1972 Presidential Elections

JOHN H. ALDRICH
Michigan State University

RICHARD D. MCKELVEY
Carnegie-Mellon University

Introduction

The analysis of electoral behavior has seen a radical shift in emphasis over the last two decades. Empirically, there has been a major resurgence of interest in analysis of the relationship between issues and electoral behavior.[1] At approximately the same time, the theoretical literature has seen the development of the spatial model of party competition.[2] Based on a rational choice view of politics, this model perceives elections as a strategic contest between candidates who compete for votes by adopting positions in a multidimensional issue space. The further advancement of both the approaches has been hindered by the inability to obtain good empirically based measurements of the positions of the candidates and citizens in a common issue space.

Some recent literature along this line has attempted to estimate candidate and citizen positions by using only individual level *preference* data among candidates. There is a substantial body of psychological literature, culminating in the development of multidimensional proximity scaling methods,[3] which, on the basis of single-peakedness assumptions of individual preferences, provides means for estimating these positions. Despite the attractions of these methods, this approach has some serious drawbacks. Strong assumptions must be made about the nature of individual preferences, and only a relatively small number of all-inclusive and relatively uninterpretable dimensions can be recovered. In addition, by depending on preference rather than perceptual data, these methods end up *assuming* that voters have single-peaked preferences and vote for the candidate closest to them, rather than being able to *test* such assertions.

For the above reasons, we feel that operationalizations of tests of the spatial model have to be based on *perceptual,* as opposed to *preference* data. One method of collecting such data has been the straightforward procedure of simply asking respondents to place themselves as well as the candidates on a common issue continuum. These types of data have been collected, for example, in the 1968 and 1972 SRC election surveys in the form of "seven point scales." Here each respondent is asked to identify the positions of the major candidates and parties on a preselected set of issues. He identifies these perceptions, as well as his own "ideal point," by placing them somewhere on an equal interval scale running from 1 to 7 in which the two endpoints are identified.[4]

Unfortunately, much of the analysis of this type of data indicates that, in general, there is substantial disagreement between different individuals' perceptions of candidates, so it is not clear how to use such data to obtain representations of candidates and voters in a common

[1] See John H. Kessel, "Comment: The Issues in Issue Voting," *American Political Science Review,* 66 (June, 1972), 459–465, for an extensive bibliography of the literature.

[2] See Otto A. Davis, Melvin J. Hinich, and Peter C. Ordeshook, "An Expository Development of a Mathematical Model of the Electoral Process," *American Political Science Review,* 64 (June, 1970), 426–448, for a review of this literature.

[3] See Paul E. Green and Frank J. Carmone, *Multidimensional Scaling and Related Techniques in Marketing Analysis* (Boston: Allyn and Bacon, Inc., 1970) for a review of this literature. Applications of these methods to the 1968 and 1972 presidential elections can be found in Jerrold G. Rusk and Herbert F. Weisberg, "Perceptions of Presidential Candidates: Implications for Electoral Change," *Midwest Journal of Political Science,* 16 (August, 1972), 338–410, Herbert F. Weisberg and Jerrold G. Rusk, "Dimensions of Candidate Evaluation," *American Political Science Review,* 64 (December, 1970), 1167–1185, and Gary A. Mauser, "A Structural Approach to Predicting Patterns of Electoral Substitution," in *Multidimensional Scaling: Theory and Applications in the Behavioral Sciences, Vol. II, Applications,* ed. Roger N.

Shepard et al. (New York: Seminar Press, 1972), pp. 245–287.

[4] These particular data have been analyzed by Benjamin I. Page and Richard A. Brody, "Policy Voting and the Electoral Process: The Vietnam War Issue," *American Political Science Review,* 66 (September 1972), 979–995, and by John H. Aldrich, "Some Results about the 1968 Election Based on the Theory of the Spatial Model of Party Competition" (paper delivered at the 1973 Annual Meetings of the American Political Science Association).

space. Further, a natural interpretation of such data, and an interpretation that has been drawn by some political scientists is that voters simply don't have the necessary information to evaluate and intelligently vote their preferences in an election, as is assumed in the spatial theories.

In this paper, we suggest an alternative interpretation of the above data, and argue that at least part of the confusion which has been attributed to the voter may be attributable purely to methodological difficulties inherent in collecting this type of perceptual data. We propose a model of the possible generation of such data which can be used to eliminate the errors attributable to these methodological difficulties. The same model serves as a scaling procedure which can be used to scale candidates and voters in a common issue space. This yields interval level data on candidate and voter positions which can be used to address various propositions from the spatial modeling literature.

The model we develop assumes that candidates occupy fixed positions in an issue space and that the individual perceptual data arises from this via a two-step process, the first step consisting of "true" error in perception, and the second step consisting of distortion introduced in the actual survey situation. We then derive a least squares solution for the true parameters of this model. The solution turns out to be essentially a principal components solution for the candidate parameters together with a regression estimate of the citizen parameters. We go on to evaluate, by Monte Carlo methods, the statistical properties of these estimators, in the type of situation to which they will be applied. Finally, we use the method to analyze the candidate and citizen positions on two issues in the 1968 and 1972 presidential elections.

The Problem

Before proceeding with the formal development of the model, we shall illustrate, in greater detail, the types of methodological difficulties that can be expected to arise in the analysis of individual level perceptual data of the sort described above. In particular, we consider an example of the type of data that might arise from the "seven-point" SRC scale on Vietnam.

In the example of Table 1, we have illustrated the possible perceptions of three voters. It is evident that although these three voters differ greatly in their placement of the candidates on the Vietnam scale, they seem to agree pretty well on the underlying scale on which the candidates lie. An alternative explanation for the lack of agreement of the voters on the placement of the candidates is that the voters

are simply exhibiting different reactions to the response task. Convincing arguments can be made that this may be occurring at least to some extent. Thus, even though the endpoints of the issue scales are identified, these identifiers are rather vague, and their responses are subject to interpretation. Different voters may be anchoring the scales according to their own interpretation of these endpoints. The fact that voters are also asked to locate their own ideal points on the scale can only serve to accentuate this tendency, for a voter who is himself a hawk is likely to interpret the endpoints of the Vietnam scale in order to accommodate his own ideal point, thus pushing his perceptions of the candidates farther to the left than a dove would. In addition, and associated with the ambiguity of the endpoints, is the problem that different voters may well interpret the intervals on the scale differently. Again, it is reasonable to suspect, for example, that an extreme hawk might see less difference between Nixon and Humphrey than a moderate would. Finally, the forced categorization tends to have additional undesirable effects: Not only does one lose information by forcing voters to ignore small differences, but also voters tend to place their perceptions of candidates, as well as their placement of their own ideal points, more frequently in the "prominent" categories (i.e., 1, 4, and 7) rather than in the "off" categories (i.e., 2, 3, 5, and 6). This tendency leads to curious results when one attempts to analyze data. If one uses the raw data to observe the distribution of ideal points, for example, he or she observes that one of us has likened elsewhere to a "circus tent" effect, obtaining a distribution with modes at the prominent categories. Again, this gravitation of the respondent towards the prominent categories is usually interpreted as meaning that one is asking too much of the respondent—that the voter cannot make such fine distinctions—and one then proceeds to collapse the off categories, losing further information. Here, also, an alternative explanation might be that the gravitation towards the prominent categories is due to the ambiguity of the scale. Individuals use the prominent categories as "natural anchoring" points, but each individual gives his own interpretation to the prominent points.

If the above is an accurate account of the generation of the data, then data of this type would seem to contain contaminating information, in addition to the information they carry about the true candidate positions. For it is possible that there might be complete agreement in the perceptions of the candidates, but that because of different interpretations of the scale, we might be led to believe that there was

Table 1. Some Hypothetical Voters' Scores on Seven-Point Vietnam Scale

	"Dove"	1	2	3	4	5	6	7	"Hawk"
Voter 1		H,J,N			W			S	
Voter 2		H	J		N,S			W	
Voter 3		S		H	J,N			W	

Key: H = Humphrey
J = Johnson
N = Nixon
W = Wallace
S = Self (i.e., Voter's ideal point)

little or no agreement. In fact, one would not expect that *all* of the variation in perceptions of the candidates would be accounted for by the above type of contamination, but one would like to be able to sort out what portion of the variance is due to actual variations in perceptions and what is due to variations in response to the scale. It is this question that we try to answer in the next section. We attempt there to factor out the variations due to differential response to the response task by placing all individuals in a common space such that their perceptions are most in agreement with the common perception of the candidates.

Formal Development of the Model

We assume that the candidates occupy true positions on an issue continuum, and that the information that the citizen gives us on his perception of the candidates is derived from this true position in a two step process. In the first stage, we assume that there is a random disturbance in the citizen's perception of the candidate. This error in perception could arise for several reasons. For example, it may occur because the candidate is unintentionally ambiguous about his position. It may occur because voters only obtain partial information from secondary sources who distort that information in the passing. It may also arise because the voters themselves selectively perceive and distort the information they receive so that it is consistent with their prior information. Whatever the cause, we assume that the first stage, which results in the voters' perceptual space, consists in the voter observing the true space, subject to this error.

The second stage consists of the voter taking what is in his head, i.e., his perceptions, and reporting them to the interviewer. Here, we assume, since there is no common metric for placing the candidates on a scale, that the positions where the citizen reports that he sees the candidates may be an arbitrary linear transformation of his perception of the space.

More formally, we develop the following model: We assume that there are J candidates who occupy the positions $Y_1, Y_2, \ldots, Y_J$ on a one-dimensional continuum, i.e., $Y_j \epsilon R$ for $1 \leqslant j \leqslant J$. Since this scale can only be specified up to a linear transformation, we assume that it is normalized with unit sum of squares, i.e.,

$$\sum_{j=1}^{J} Y_j = 0 \text{ and } \sum_{j=1}^{J} Y_j^2 = 1.$$

Further, we assume that there are n citizens, each of whom has a perception of each candidate. The ith citizen's perception of the jth candidate is denoted Y_{ij}, and we assume that this is distributed randomly around the true candidate position, as illustrated in Figure 1, for four candidates. Thus, for the first stage, we assume that individual perceptions are generated as follows:

$$Y_{ij} = Y_j + u_{ij} \tag{1}$$

for $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant J$,

where u_{ij} is a random variable which satisfies the usual Gauss Markov assumptions, i.e.,

$$E(u_{ij}) = 0 \text{ for all } i, j$$
$$E(u_{ij})^2 = \sigma^2 \text{ for all } i, j. \tag{2}$$
$$E(u_{ij}u_{kl}) = 0 \text{ for all } i, j, k, l \text{ with either } i \neq k \text{ or } j \neq l.[5]$$

---

[5]The scaling model we develop here rests on Gauss Markov type assumptions. There is good reason to question some of these assumptions, and hence attempt to extend the basic results to cover some of these potential violations. One assumption is that of "homoscedasticity" or constant error variance for each respondent and each candidate/stimulus. It may be more reasonable to expect some respondents to have greater perceptual error than others, and some stimuli to have lesser ambiguity in their positioning on issues than others. Another assumption is that of "no covariance" among an individual's perception of different candidates. It may be more realistic to assume that a respondent might identify certain candidates so that a change in perceived position of one candidate carries over to other candidates. A more general model incorporating these objections, would assume that there is a known stochastic variance/covariance matrix for each respondent:
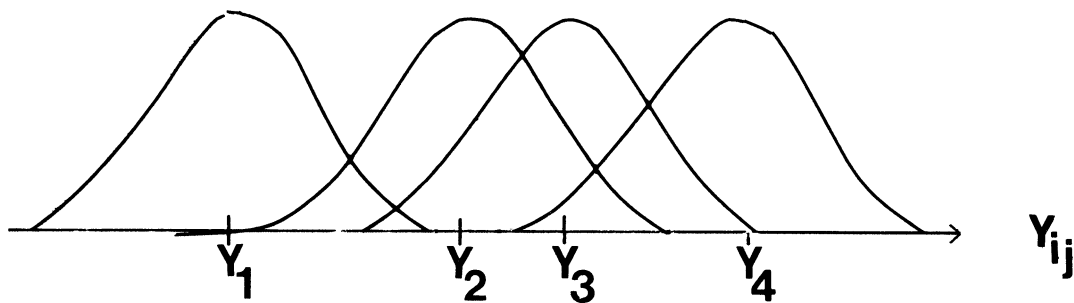
Figure 1. Distribution of Perceptions, Y_{ij}

The second stage of the data generation consists in the reporting of the perceptual space. Here we assume that our observed data consists of a reported position for each candidate, for each voter, and that this consists of some linear transformation of the voter's perceptual space. I.e., we let X_{ij} represent the position where citizen i reports that he sees candidate j, and it is assumed that for each voter, there are scalars, c_i, $w_i \in R$ such that

$$c_i + w_i X_{ij} = Y_{ij} = Y_j + u_{ij} \qquad (3)$$

for $1 \leqslant i \leqslant N$, $1 \leqslant j \leqslant J$, or equivalently,

$$X_{ij} = \frac{1}{w_i}(Y_{ij} - c_i) = \frac{1}{w_i}(Y_j - c_i) + \frac{u_{ij}}{w_i} \qquad (4)$$

Note that the above transformation allows us to account for different anchoring of the scale as well as for different interpretation of the intervals. It does not account for the ordinal nature of the data.

Now, the only data we actually observe is the X_{ij} matrix of reported positions, and from this, we want to recover the true parameters, Y_j, c_i, and w_i for $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant J$. Before proceeding, we establish some matrix notation to make the calculations less burdensome. We set

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_J \end{bmatrix}, \quad X_i = \begin{bmatrix} 1 & X_{i1} \\ 1 & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & X_{iJ} \end{bmatrix}, \quad \beta_i = \begin{bmatrix} c_i \\ w_i \end{bmatrix}.$$

$$E(u_i' u_i) = \begin{bmatrix} \sigma_{1i}^2 & (\sigma_{1i}\sigma_{2i}) \ldots (\sigma_{1i}\sigma_{Ji}) \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ (\sigma_{Ji}\sigma_{1i}) & (\sigma_{Ji}\sigma_{2i}) \ldots \sigma_{Ji}^2 \end{bmatrix}$$

While we have formulated a model of the above sort, we have not yet been able to carry through the mathematical derivations of the estimators for this case.

And, using $\hat{Y}_j$, $\hat{c}_i$, $\hat{w}_i$, etc., to denote our estimates of the true parameters, we set

$$\hat{Y} = \begin{bmatrix} \hat{Y}_1 \\ \cdot \\ \cdot \\ \cdot \\ \hat{Y}_J \end{bmatrix}, \quad \hat{\beta}_i = \begin{bmatrix} \hat{c}_i \\ \hat{w}_i \end{bmatrix}.$$

Then, the perceived candidate positions for citizen i are

$$X_i \hat{\beta}_i = \begin{bmatrix} \hat{c}_i + \hat{w}_i X_{i1} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \hat{c}_i + \hat{w}_i X_{iJ} \end{bmatrix}.$$

The procedure that we will use will be to choose $\hat{Y}$ and $\hat{\beta}_i$, for $1 \leqslant i \leqslant n$ in such a way as to get the best fit, in a least squares sense, between the estimated candidate positions and the citizen's perceptions of them. To do this, we define the vector of estimated residuals, for individual i, as

$$e_i = X_i \hat{\beta}_i - \hat{Y}. \qquad (5)$$

Then, the sum of the squared residuals for a particular voter is

$$e_i' e_i = (X_i \hat{\beta}_i - \hat{Y})' (X_i \hat{\beta}_i - \hat{Y}), \qquad (6)$$

and the total sum of squared residuals is

$$SS(\hat{\beta}_i, \ldots, \hat{\beta}_n, \hat{Y}) = \sum_1^n e_i' e_i. \qquad (7)$$

We want to minimize (7) subject to the constraints that

$$\sum_{j=1}^J \hat{Y}_j = 0 \text{ and } \sum_{j=1}^J \hat{Y}_j^2 = 1. \qquad (8)$$

so, if we let C be a J x 1 matrix of ones, i.e.,

$$C = \begin{bmatrix} 1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix}$$

the constraints in (8) become

$$C'\hat{Y} = 0$$

and   (9)

$$\hat{Y}'\hat{Y} = 1.$$

We set the above up as a Lagrangean multiplier problem, getting

$$L(\hat{\beta}_i, \hat{Y}, \lambda_1, \lambda_2) = \sum_{i=1}^{n} e_i'e_i + 2\lambda_1 \, C'\hat{Y}$$

$$+ \lambda_2(\hat{Y}'\hat{Y}-1) = \sum_{i=1}^{n} (\hat{Y}-X_i\hat{\beta}_i)'(\hat{Y}-X_i\hat{\beta}_i)$$

$$+ 2\lambda_1 C'\hat{Y} + \lambda_2 (\hat{Y}'\hat{Y}-1). \tag{10}$$

Differentiating and setting equal to **0**, we get the $2n + J + 2$ equations:

$$\frac{\partial L}{\partial \hat{\beta}_i} = -2X_i'\hat{Y} + 2X_i'X_i\hat{\beta}_i = 0, \text{ for } 1 \leqslant i \leqslant n \tag{11}$$

$$\frac{\partial L}{\partial \hat{Y}} = -2\sum_{i=1}^{n} X_i\hat{\beta}_i + 2n\hat{Y} + 2C\lambda_1 + 2\lambda_2\hat{Y}$$

$$= 0 \tag{12}$$

$$\frac{\partial L}{\partial \lambda_1} = C'\hat{Y} = 0 \tag{13}$$

$$\frac{\partial L}{\partial \lambda_2} = \hat{Y}'\hat{Y}-1 = 0. \tag{14}$$

Solving (11) for $\hat{\beta}_i$, we get

$$\hat{\beta}_i = (X_i'X_i)^{-1}X_i'\hat{Y}, \tag{15}$$

so that the individual transformation consists of the least-squares regression of the reported on the actual (unknown) positions of the candidates. Now, substituting in (12), gives,

$$\sum_{i=1}^{n} [X_i(X_i'X_i)^{-1}X_i']\hat{Y} - n\hat{Y} - C\lambda_1$$

$$- \hat{Y}\lambda_2 = 0. \tag{16}$$

Setting

$$A = [\sum_{i=1}^{n} X_i(X_i'X_i)^{-1}X_i'], \tag{17}$$

and substituting in (16) we can reduce equations (11) through (14), getting

$$(A-nI)\hat{Y} - C'\lambda_1 - \hat{Y}\lambda_2 = 0 \tag{18}$$

$$C'\hat{Y} = 0 \tag{19}$$

$$\hat{Y}'\hat{Y} = 1. \tag{20}$$

Multiplying (18) by C', we get

$$C'(A-nI)\hat{Y} - J\lambda_1 - (C'\hat{Y})\lambda_2 = 0. \tag{21}$$

But now it is straightforward, by expansion of X_i, to show that, for any $1 \leqslant i \leqslant n$

$$C'[X_i(X_i'X_i)^{-1}X_i'] = C',$$

so that

$$C'A = C'[\Sigma X_i(X_i'X_i)^{-1}X_i'] = nC'$$

and

$$C'(A - nI) = 0. \tag{22}$$

And from (19), it follows that $C'\hat{Y} = 0$, so (21) yields $J\lambda_1 = 0$, or

$$\lambda_1 = 0. \tag{23}$$

Now (18) becomes

$$(A - nI)\hat{Y} = \lambda_2 \hat{Y}. \tag{24}$$

But this simply says that $\hat{Y}$ is a characteristic vector of the matrix $(A - nI)$, which gives us our solution. To determine which characteristic vector to choose, we note that $-\lambda_2$, the negative of the characteristic root, represents the sum of the squared errors associated with the characteristic vector $\hat{Y}$. To see this, we multiply (18) by $\hat{Y}$, getting

$$\hat{Y}'(A - nI)\hat{Y} - \hat{Y}'C\lambda_1 - \hat{Y}'\hat{Y}\lambda_2 = 0,$$

or, applying (19) and (20),

$$\lambda_2 = \hat{Y}'(A - nI)\hat{Y}. \tag{25}$$

But, setting

$$A_i = X_i(X_i'X_i)^{-1}X_i',$$

one can easily show that $(I - A_i)$ is a symmetric, idempotent matrix, and

$$\sum_{i=1}^{n} e_i'e_i = \sum_{i=1}^{n} (\hat{Y} - X_i\hat{\beta}_i)'(\hat{Y} - X_i\hat{\beta}_i)$$

$$= \sum_{i=1}^{n} (\hat{Y} - X_i(X_i'X_i)^{-1}X_i'\hat{Y})'$$

$$(\hat{Y}-X_i(X_i'X_i)^{-1}X_i'\hat{Y})$$

$$= \sum_{i=1}^{n} \hat{Y}'(I - A_i)'(I - A_i)\hat{Y}$$

$$= \sum_{i=1}^{n} \hat{Y}'(I - A_i)\hat{Y} = \hat{Y}'(nI - A)\hat{Y}$$

$$= -\hat{Y}'(A - nI)\hat{Y}.$$

So from (25),

$$-\lambda_2 = \sum_{i=1}^{n} e_i' e_i. \tag{26}$$

Thus, our solution, $\hat{Y}$, is the characteristic vector of the matrix $(A - nI)$ with the highest (negative) nonzero characteristic root. Having obtained a solution for the candidate positions, we can, of course, go back to (15), to obtain the parameters of the individual transformation by performing the least squares regression of the individual's reported positions on the estimated positions of the candidates.

With regard to the estimates of the individual perceptions, we note that $\hat{Y}_i = X_i \hat{\beta}_i$ is an estimate of the ith voter's perceptions of the candidate positions. Ideally, we would hope that the average perception of a candidate's position would correspond to the estimate of his position. In vector notation, we would want

$$\frac{\sum \hat{Y}_i}{n} = \hat{Y}.$$

In fact, this is not the case. Rather, because of a "regression towards the mean" on individual's candidate perceptions, we get

$$\frac{\sum \hat{Y}_i}{n + \lambda_2} = \hat{Y} \tag{27}$$

To see this, we note that

$$\sum \hat{Y}_i = \sum X_i \hat{\beta}_i$$
$$= \sum X_i (X_i' X_i)^{-1} X_i' \hat{Y}$$
$$= A \hat{Y}$$
$$= (n + \lambda_2) \hat{Y}, \tag{28}$$

from which (27) follows.

Because of this relation, in our empirical applications, we will actually present the voter's perceptions of the candidate in terms of the expanded transformation of equation (27), which differs from the least squares estimators by a factor of $\frac{n}{n+\lambda_2}$. This has the effect of normalizing the solution with respect to the mean perceptions of the candidates, and makes possible more direct comparisons with the unscaled data.

Note that from (26), it follows that the expression

$$\frac{-\lambda_2}{nJ} = \frac{\sum_{i=1}^{n} e_i' e_i}{nJ} \tag{29}$$

represents the average squared deviations of the observed from the true candidate positions, and we can use this as an estimate of σ^2. Formally, we set

$$\hat{\sigma}^2 = \frac{-\lambda_2}{nJ}. \tag{30}$$

Since the estimated scale positions of the candidates are normalized to have unit sum of squares, $\hat{\sigma}^2$ can also be used as a measure of the "goodness of fit" of the model. Actually, we will see later that $\hat{\sigma}^2$ is generally a biased estimator of σ^2 providing a substantial underestimate of σ^2. This can be partially corrected for by computing the sum of the squared error in the expanded perceptual space described above. This results in the formula

$$\hat{\sigma}^2 = \frac{-\lambda_2}{nJ} \left(\frac{n}{n+\lambda_2}\right)^2 = -\frac{n\lambda_2}{J(n+\lambda_2)^2} \tag{31}$$

which is the formula we will actually use. Even with this adjustment there is substantial bias left, as we shall see. We have not yet been able to correct for this bias, however, and use σ^2 as defined above for the present, realizing that it must be interpreted cautiously.

Another point of caution regards the estimate of the individual transformations in (15). Note that no constraints are placed on $\hat{\beta}_i$. In particular, no constraint is placed on $\hat{w}_i$, so it is assumed that $\hat{w}_i$ could be negative for some voters. In applications to real data, this means that voters who perceive the candidates in a "mirror image" space will be estimated as having good fits to the true model, but with negative weights. A voter who sees things backwards then contributes to a better fit to the "true" space, and this accounts for some of the underestimation of σ^2 mentioned above.

Given that the endpoints of the scales are identified in the empirical data to which we actually apply the techniques, it is not clear that one would want to treat such voters as we have above. It would be more reasonable, perhaps, to assume that the parity of the scale is given, and that any misperception of this parity is due to error in perception. In terms of the model, this would correspond to an additional set of constraints, i.e., that $\hat{w}_i > 0$ for all voters. We have not done this for several reasons. First, the additional mathematical complexities which are introduced by this modification are substantial. Second, although the problem of negative weights is serious if one only has a small number of candidates, it should be less so as the number of candidates increases. Thus, with large numbers of candidates, the probability of obtaining a mirror-image set of observations purely by chance becomes smaller. Finally, the procedure we have developed above at least has the virtue of identifying the voters with negative weights so that one can treat them separately if need be.

All of the above analysis has dealt with scaling of the candidate positions. We have not yet discussed the treatment of individual ideal points, but this procedure is straightforward. To obtain the individual's ideal point in the common space, we merely subject it to the same transformation that his perceptions of the candidates are subjected to. Thus, if X_{i0} represents the ith individual's placement of his ideal point, then

$$\hat{Y}_{i0} = \hat{c}_i + \hat{w}_i X_{i0} \tag{32}$$

is our estimate of his ideal point in the common space.

Before proceeding, it will be worthwhile to point out some similarities between the specification of the model that we have developed in this section and the usual factor analytic model. For these purposes, we set

$$X = \begin{bmatrix} X_{11} \dots X_{1J} \\ \cdot \\ \cdot \\ \cdot \\ X_{n1} \dots X_{nJ} \end{bmatrix}, A = \begin{bmatrix} a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{bmatrix}$$

$$F = \begin{bmatrix} F_1 \dots F_n \end{bmatrix}$$

$$D = \begin{bmatrix} d_1 & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & d_n \end{bmatrix}, \text{ and } U = \begin{bmatrix} U_{11} \dots U_{1J} \\ \cdot \\ \cdot \\ \cdot \\ U_{n1} \dots U_{nJ} \end{bmatrix} \tag{33}$$

Here, X represents the matrix of observed data, where the X_{ij} are defined above. F represents a common factor of scale positions (similar to Y above), A is a vector of individual transformations, D a diagonal matrix of scalars, and U a matrix of errors.[6] Then the usual one-factor model can be written

$$X = AF + DU \tag{34}$$

or

$$X_{ij} = a_i F_j + d_i u_{ij},$$

which can be compared to (4) to note the similarities. The differences are that in (34), only stretching and shrinking of the original space is allowed. More important, however, the factor model generally treats F as a random variable rather than as a parameter to be estimated.[7] Hence, although one can obtain estimates of the factor scores, one does not obtain sampling distributions of these estimates. Since the candidate positions are of primary interest, we are particularly concerned about the accuracy with which they are recovered, and we would want a model that treats them as parameters rather than random variates.

In addition to the differences in the specification of the model, one should note that if the factor-analytic formulation above is used, computational problems arise in applying usual factor-analytic procedures for obtaining a solution because the usual roles of the observation and the variable are reversed. Thus a variable, under this representation, is a respondent, while an observation is the vector of individual perceptions of a particular candidate. Because of this reversal of roles, one would end up factor analyzing a matrix of, say, four observations and 1000 variables, leading to unmanageable correlation matrices.

Despite the differences between the two formulations, it can be shown that the solution we have derived above for the candidate parameters is mathematically equivalent to extracting the first principal component of the correlation matrix XX'.

Monte Carlo Results

The last section has derived a least-squares solution for our scaling problem. In order to assess the adequacy of the solution, and its performance in a given situation, however, we should know something about the statistical properties of the estimators. Thus, unless we know the theoretical sampling distributions of the estimators, it is difficult to know how much confidence to place in the results. Since our estimators have been derived in the rather complex manner described above, analytical determination of their sampling distributions is exceptionally difficult. Although we cannot obtain mathematical derivations of these distributions, we can obtain an indication of their properties by conducting Monte Carlo type experiments.

[6]See Harry H. Harman, *Modern Factor Analysis* (Chicago: The University of Chicago Press, 1967), for a more complete development of the factor-analytic model.

[7]See e.g., D. N. Lawley and A. E. Maxwell, *Factor Analysis as a Statistical Method* (London: Butterworths, 1963) for a discussion of this point.

The purpose of a Monte Carlo experiment is to generate artificial data according to a specified probabilistic model. In our case, we can specify true parameters and a stochastic term of known size and then generate data according to the model of the previous section to see how well the technique recovers these parameters. Typically, we choose, as known parameters, ones that will be as realistic as possible. To do so, we have used real data estimates to serve as parameters, and in particular, we have used those determined from the 1968 Vietnam scale data to be discussed below. Therefore, we hope to investigate the adequacy of the technique in the sort of situation we will be faced with in the real data. The true positions of the stimuli were chosen to correspond to the estimated positions of Johnson, Humphrey, Nixon, and Wallace, respectively, and the variance, σ^2, of the perceptions around the true positions was assumed to be equal to the largest variance in the estimated solution. Thus, we have $\sigma = .388$, or $\sigma^2 = .1505$. (This value of σ is actually somewhat different from the results reported below, because the Monte Carlo experiment was based on a preliminary estimation of the 1968 data.) Finally, the individual parameters, c_i and w_i, were chosen by taking a random sample of the 1968 respondents, and using their estimated values, $\hat{c}_i$ and $\hat{w}_i$ for the Monte Carlo experiment. Using these "true" parameters, we generated 25 samples with an n of 100 each, according to the model described in the previous section.

The results of the estimation of the parameters is reported in Table 2, and we are led to conclude that all the parameters describing the candidate positions are recovered exceptionally well. Figure 2 illustrates the distribution of perceptions in the assumed true model, while Figure 3 illustrates the sampling distributions of the estimated candidate positions. We see that even with a substantial amount of misperception in the original data, the technique recovers the candidate positions very well. The average correlation between the 25 estimated candidate vectors and the true candidate vector was .9977, with the average estimate of each candidate being nearly identical to the true position of the candidate. The mean estimate for each candidate is well within one standard deviation of the true position; thus if there is any bias in these estimators, it is insignificant in relation to their standard error.

The estimators for the variance in perception show a slightly different story. Here we get an overall estimate of σ of .2845, a significant underestimate of the true parameter, $\sigma = .388$. The estimates of the error in perceptions around each candidate are similar in magnitude

and in their negative bias. This negative bias reveals one of the potential drawbacks of the technique, i.e., since the least-squares procedure attributes as much of the error as possible to variations in reactions to the scale, the procedure cannot recognize someone who actually perceives all candidates (say) to the left of their true positions. Such a voter is seen as perceiving the space correctly; and consequently, we underestimate to a certain extent the variance in perceptions. Similarly, as discussed in the previous section, voters with negative weights contribute to underestimation of σ. Both of these sorts of underestimation should become less severe as the number of candidates increases.

Turning to the individual parameters, we now must view the experiment as 100 individual parameter pairs, each being estimated by 25 observations, or estimates. Rather than presenting the estimated sampling distribution of each of these 200 estimators, we present, in Table 3, some summary statistics of these sampling distributions across all 100 voters. The first set of figures gives the distribution of the true parameters, so that we have a base against which to evaluate the estimators. The next two statistics give an indication of the average bias of the estimators. First is a measure of the average (signed) bias, and the fact that these are so close to zero indicates that there is no consistent bias in one direction or the other across the 100 estimators. The second statistic is the root mean-square of the bias over all 100 estimators. This gives a better indication of the average magnitude of the bias. The third figure gives the average standard error of the 100 estimators. These indicate that the bias is generally insignificant in comparison with the standard error of the estimators. Unlike the candidate estimates, however, the standard error is fairly large. To get an idea of the amount of error that is represented in the estimators of the c_i and w_i, one can compare these figures (the average standard error figures) with the standard deviations of the distributions of the true c_i and w_i. The average standard error figures are on the order of half the size of these standard deviations. This means that although these estimators may perform well on the average, in any given sample there can be a substantial amount of error in the estimation of a particular voter's transformation parameters.

It is difficult to assess the seriousness of the error in the individual estimates when they are expressed in the above form. This is because we are not really interested in the transformation parameters themselves, but are interested in them so that we can determine how accurate is the recovery of arbitrary points on the individu-

Summary of Monte Carlo Experiment
Distribution of Estimated Parameters
(In 25 samples, n = 100)

Table 2. Distribution of Candidate Estimators

	True Parameters	Candidate Parameters		Citizen Perceptions	
		Mean	Standard Deviation	Mean σ_j	Standard Deviation σ_j
Y_1	−.321	−.320	.040	.289	.021
Y_2	−.424	−.422	.035	.290	.020
Y_3	−.096	−.096	.036	.304	.013
Y_4	.841	.840	.010	.253	.021
σ	.388	.285	.013		

Average correlation of Y_j with $\hat{Y}_j$ = .9977 (over 25 samples)
Average correlation of Y_{ij} with $\hat{Y}_j$ = .868
Correlation of Y_j with $\hat{Y}_j$ = .999997.

Table 3. Distribution of Voter Transformation Parameters and Ideal Points

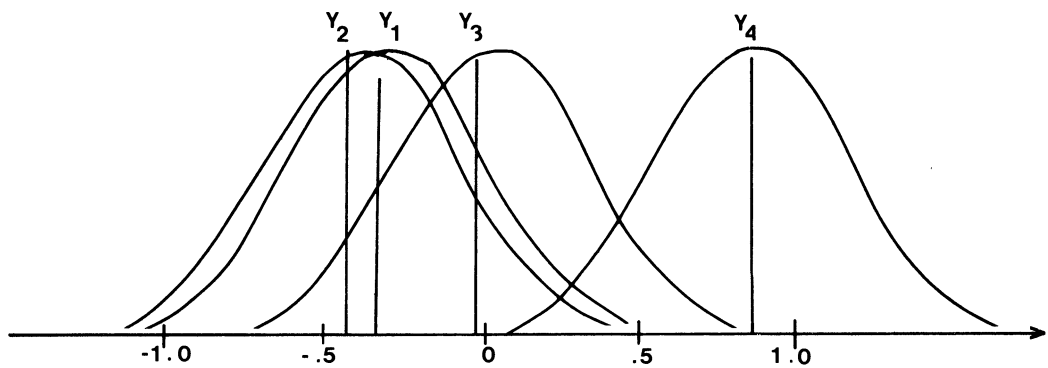|		c_i	w_i	Y_{i0}
True Distribution	Mean	−.1544	.368	−.209
	Standard Deviation	(2.242)	(.525)	1.757
Estimators	Average Bias	−.0897	.0217	−.189
		(.356)	(.083)	(1.376)
	RMS Bias	.1338	.0072	
	Average Standard	1.080	.2511	.486
		(.865)	(.279)	(.422)

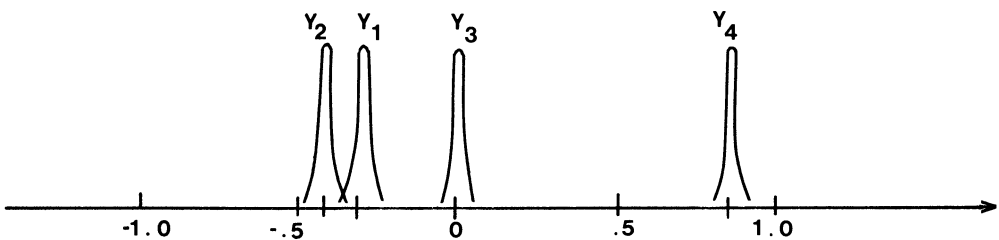Figure 2. Assumed True Model (Distribution of Perceptions)

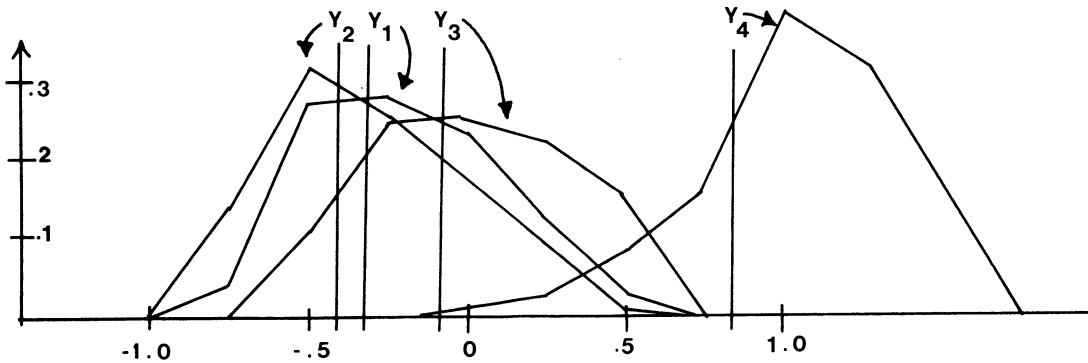Figure 3. Estimated Model (Sampling Distribution of Candidate Estimators)

Figure 4. Estimated Model (Recovered Distribution of Perceptions)

al's perceptual scale. In particular, in our applications below, we will be interested in how well we can recover the individual's ideal point. To investigate this problem we assume that each of the 100 voters has an ideal point Y_{i0}, which consists of a point in his perceptual space. Unlike his perceptions of the candidates, this is not subject to error. We then want to discover how well, for each of the 100 voters, we recover the ideal point over the course of the 25 samples. As above, we should end up with 100 estimators of 100 true parameters (in this case the Y_{i0}), each estimator based on a sample size of 25.

Just as the c_i and w_i were taken as the estimated values from an application to the 1968 Vietnam scale, the Y_{i0} were determined by setting $Y_{i0} = c_i + w_i X_{i0}$, where X_{i0} is the reported ideal point of the ith voter and c_i and w_i his estimated coefficients in the 1968 application. This computation should result in a distribution of true values much like that which we would expect to find in the actual applications. For each of the 100 voters, then, we get an estimator, $Y_{i0} = c_i + w_i X_{i0}$, for which we have 25 observations, and it is this estimator that interests us. In Table 3, we have presented information on the distributions of these estimators. We note that the average standard error of these estimators is .486, which is substantial, but comparing it to the distribution of the true ideal points, with a standard deviation of 1.757, we conclude that we can tell at least in what general area of the distribution the citizen's ideal point falls. Another way of looking at this result is to note that a 95 per cent confidence interval for an individual ideal point will be about one-half of a standard deviation either way of the estimated value. Although the amount of error may appear to be substantial, each of the individual ideal point estimates is based on a regression on four data points, and one might consider it surprising that the error is not larger than it is.

These estimates do seem to reflect a systematic bias. Points close to the mean of the true distribution are recovered with relatively greater accuracy than are those at the extremes, as reflected by the correlation of .135 between the standard error and the true value squared, and by the correlation of −.447 between the bias and the true value squared. Thus, there is a larger bias toward the mean and a larger standard error in the estimates of extremist ideal points, leading to a greater total mean squared error for these voters. This "regression toward the mean" suggests that a good portion of the above error in the individual estimators may be accounted for by the extremists. From our point of view, i.e., for predicting voting behavior, we are more interested in getting good estimates of moderate ideal points than of extremists, for it is the moderate whose vote is most difficult to predict. Thus, the bias here works in our favor, giving us better estimators where we need them. As a final note, the "regression toward the mean" implies that extremists will be estimated as somewhat less extreme than they actually are. Undoubtedly this same phenomenon accounts for the skewing of the distributions of the perceptions of a candidate, as observed in Figure 4 above.

This section is rather brief, since the important findings are best presented in tabular form. This brevity does not mean that the results obtained are unimportant. On the contrary, they strongly indicate the scaling approach developed in the previous section will yield reasonably accurate estimates of the unknowns, provided that the model underlying the technique fairly describes the data being scaled. The estimators of the candidate parameters are remarkably accurate, the only deficiency being in an underestimation of the variance in individual perceptions. The recovery of the individual parameters is considerably less accurate, but there is very little bias in these estimates, and one can expect that in large

samples, the method will recover these parameters well on the average. With these results in mind, we can proceed to apply the technique to some real data.

Empirical Results—Introduction

The scaling procedure we have developed provides a solution to a general scaling problem. In this section we apply this technique to some electoral data, the type of problem that catalyzed our interest in the scaling procedure in the first place. The basic data will be the two seven point issue scales concerning urban unrest and Vietnam that were asked in the SRC's 1968 election survey and again in 1972. Following a description of the data, we will explain the candidate position estimates and the remaining variance of individual perceptions of these issues in the two very different elections. We will use the 1968 results to define a two-dimensional issue space in which we can locate both candidates and citizens. This space can be used to predict voting behavior and to demonstrate the improvement in such predictions from using the scaling results instead of the unscaled seven-point issue data. Unfortunately, the urban unrest scale was given to only a random half sample in the 1972 survey. Therefore, we are not able to relate the 1972 scaling results to the vote. Finally, we briefly report the scaling estimates for the 10 seven point issue scales that were asked of the whole sample in 1972.

The respondents sampled in 1968 were asked to locate themselves and four "candidates," Humphrey, Nixon, Wallace, and Johnson on the two dimensions. We limit our attention to those respondents who placed all four candidates and themselves on both scales and who reported their voting behavior. Further we remove any citizen who placed all candidates on the same point on an issue, because these individual parameter estimates are undefined. For these "no-variance" people,

the effect would be the same if we assigned them w_i parameter values of 0, since all candidates were seen to take the same position on that issue. These restrictions leave us with an n of 885, or about 64 per cent of the 1384 respondents asked the questions.[8] Citizens were included in the 1972 scalings if they placed themselves and all five "candidates"— McGovern, Nixon, Wallace and the two political parties—on the seven-point scale and saw at least some variance in the candidates' positions on that issue. The sample sizes are 1045 for the Vietnam scale and 519 for urban unrest. While the term "candidates" is broadly defined in these examples, it is clear that the "candidate" stimuli are all relevant to the elections and to these particular issue dimensions.

Candidate Position Estimates

The scaling estimates of the candidates' positions, the $\hat{Y}_j$, are presented in Table 4 along with the standard deviation of the citizens' perceptions of each candidate, the $\hat{Y}_{ij}$. Note that the mean of the distribution of $\hat{Y}_{ij}$ is equal to $\hat{Y}_j$. Included as well are the mean and standard deviation of the estimated distributions of citizens' ideal points (denoted by "I" in the Table). The reader should keep in mind that it is not meaningful to compare candidate positions among dimensions, even for the 1968 estimates which are based on the same sample of citizens. The "unit of measurement" of each dimension has been arbitrarily set to have a mean of zero and unity sum of squares of candidate positions, reflecting our assumption

[8]The issue scales were asked on the postelection wave of the 1968 survey. Only 1384 individuals responded to this wave, down from an original N of 1557 in the initial, self-weighting, cross-sectional sample. Fifty citizens saw no differences between the candidates on Vietnam, while 14 (including 7 of the original 50) were no variance respondents on urban unrest.

Table 4. Scaled Estimates of Candidate Positions on Vietnam and Urban Unrest

1968	LBJ	HHH	RMN	GCW	I
Vietnam	−.321	−.424	−.096	.841	−.238
(Standard deviation)	(.208)	(.302)	(.401)	(.403)	(1.033)
Urban Unrest	−.394	−.402	−.003	817	−.131
	(.239)	(.249)	(.327)	(.276)	(.645)

1972	RMN	McG	GCW	Dem.	Rep.	I
Vietnam	.361	−.705	.376	−.355	.326	.046
	(.290)	(.338)	(.430)	(.327)	(.277)	(.694)
Urban Unrest	.180	−.602	.670	−.374	.126	.243
	(.348)	(.446)	(.500)	(.363)	(.329)	(1.164)

that the dimensions are unique only up to a positive linear transformation. We will be able to make some cross-dimensional comparisons when we consider the individual's ideal point distributions.

The scaling results indicate that Wallace was distinctly the most conservative candidate on these two important issues in 1968. While Humphrey was the most liberal on urban unrest and dovish on Vietnam, the distance between him and the remaining two candidates is much less than that separating Wallace and Nixon.

In 1972, Wallace is once again the most right/hawkish candidate. In this election, however, Wallace and Nixon are estimated to be very similarly hawkish, while a relatively large distance remains between these two candidates on urban unrest. Balancing Wallace on the right and more, McGovern appears to be very liberal on both dimensions, relative to the positions of the other candidates. The Democratic party appears to be relatively liberal on the two issues. Nonetheless, there is a discernible gap between this party and its nominee, the party appearing more moderate. The same is not true of the competing party. The Republican party is very close to, but somewhat more liberal than, the President on both dimensions. Perhaps the most notable characteristic of these placements is the consistency of at least the ordinal properties of the four scales (we will find some examples where the ordinality is violated for the other 1972 dimensions). This ordinal consistency supports the notion that there might be a single dimension underlying the two issues in each election. The interval placements, however, are not completely consistent with this view (consider the relative placements of Nixon and Wallace in 1972).

Perceptual Variation

We will return to the consideration of the candidate point-estimations later, after we have investigated the distributions of perceptions and citizens' ideal points. The overall variance to perceptions in the scaled data, $\hat{\sigma}^2$, was defined in equation (31) of Section 3, and is reported in Table 5. These figures can be used as an indication of the overall "goodness of fit" of the model and data. An alternative way of looking at the "goodness of fit" is to provide a benchmark basis of comparison to indicate the amount of reduction of the variance of the scaled over the unscaled data. To make this comparison, we normalized the seven-point scale data so that the average perceived candidate position on the seven-point scales has the same mean of zero and sum of squares of one as the scaled estimates of the candidates' positions. This restandardization of the seven point scale data leads to an average variance to perceptions that can be compared with $\hat{\sigma}^2$. The ratio of $\hat{\sigma}^2$ to the average variance just described gives an indication of the reduction of variance of perceptions accomplished by the scaling technique. These figures, also found in Table 5, indicate substantial reductions in variance for all four dimensions. These range from about 31 per cent of the variance in the original data for the 1972 Vietnam scale to only 7 per cent for the 1968 Vietnam issue. Actually, in light of our Monte Carlo results, $\hat{\sigma}^2$ is probably an underestimate of the true stochastic component to perceptions. But even allowing for considerable bias in these estimators, it is clear that the scaling has effected a considerable reduction of the variance in perceptions. Our original suspicions seem to be confirmed: a substantial portion of the observed variance in perceptions seems to be due simply to different reactions to the interview response task, and not all variance is due to error in perceptions.

Figure 5 and Figure 6 graph the individual distributions of perceptions of each candidate, (the $\hat{Y}_{ij}$'s), on the four dimensions.[9] The first observation that strikes one is the general similarity of these distributions to those derived

[9] These graphs were drawn to scale by determining the individual frequency distributions of Y_j. The scaling dimension was divided into 16 categories, each spanning a range of .25, and the proportion of $\hat{Y}_j$ in each category determined. This is the same procedure as used in the Monte Carlo experiment.

Table 5. Goodness-of-Fit Measures for Scaling Solutions

	$\hat{\sigma}^2$	Average Var. 7 Point Scale Data (0,1)	Ratio of $\hat{\sigma}^2$/Average Var.	% Negative Weights
1968				
Urban Unrest	.075	.285	.264	7.5%
Vietnam	.127	1.913	.066	21.3
1972				
Urban Unrest	.162	.607	.267	7.8
Vietnam	.113	.362	.313	14.5

Urban Unrest 1968

Humphrey

Wallace

Johnson

Nixon

Vietnam War 1968

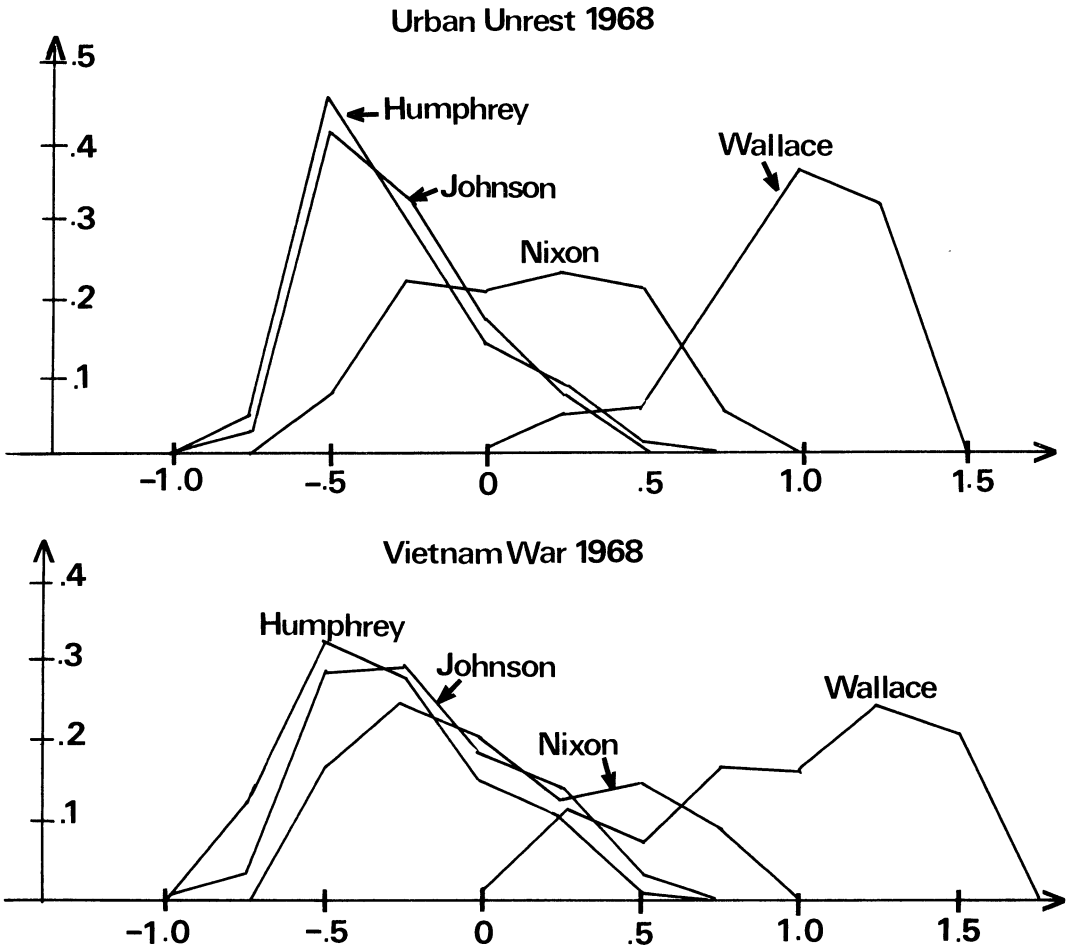Humphrey

Johnson

Wallace

Nixon

Figure 5. Distribution of Scaled Perceptions of Candidates in 1968

from the Monte Carlo experiment. In particular, there seems to be a similar skewing of the distribution of the more extreme candidates, indicative perhaps of a "regression toward the mean" phenomenon. Second, there seems to be a greater clarity of the perceptual distributions for the urban unrest dimension in 1968 and Vietnam in 1972. This relative clarity seems most reasonable in the contexts of the two elections. In contrast, the urban unrest dimension in 1972 in particular is rather jumbled at best. In fact, the two extreme candidates, McGovern and Wallace, have perceptual distributions with long tails toward the center of the candidate distribution and beyond, which overlap with each other and seem rather clearly not unimodal. Large portions of the sample, however, make a distinction between these two extreme candidates—a distinction that had been only slightly clearer in the 1968 version of this issue. While the perceptions of Wallace are

rather distinct on Vietnam in 1968, the other three distributions are nearly identical, and even the Wallace distribution has considerable overlap with the other three. If, as Page and Brody (1972) have argued, Humphrey and Nixon (and they argue perhaps even Wallace) were purposely vague on this issue, they certainly succeeded in confusing their Vietnam policy proposals with each other's in the minds of the electorate. This finding contrasts with the relative clarity with which the Y_{ij} distributions of Nixon and McGovern are distinguished in 1972 on the Vietnam issue. The much larger overlap of the perceptual distributions of Nixon and Wallace on both issues (but especially Vietnam) in 1972 than in 1968 is noteworthy. We will shortly indicate whether this change was due to Wallace's being perceived as less conservative in 1972, or to Nixon's being viewed as more conservative. Finally, the central location and relatively large dispersal of the
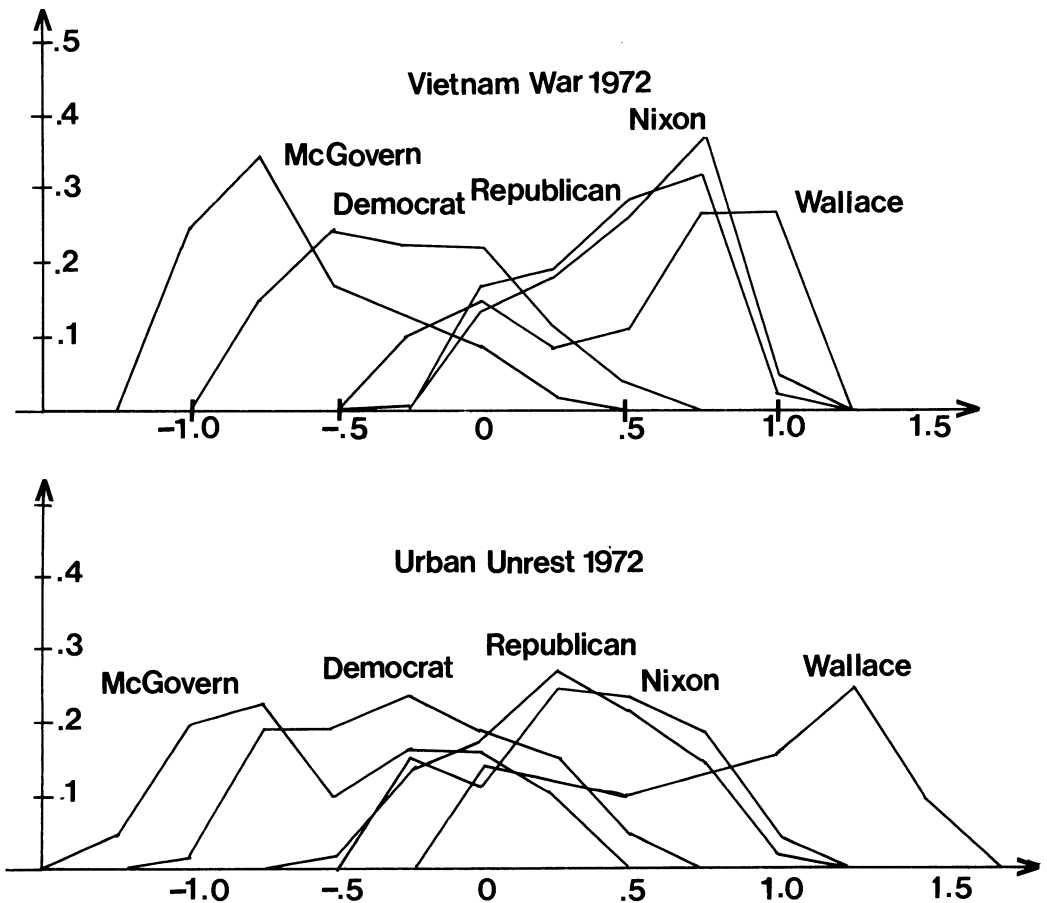
Figure 6. Distribution of Scaled Perceptions of Candidates, 1972

distributions of perceptions of Nixon in 1968 support the argument of calculated ambiguity on his part.

The entire set of candidate position estimates seems quite reasonable and helps to confirm the use of the scaling procedure for all four issues. The more refined analysis, particularly of Y_{ij}, further strengthens the argument for the "reasonableness" of the scaling estimations in our opinion. All indications, especially the distribution of perceptions, seem to point toward a greater clarity of candidate positions, and quite possibly a set of scaling estimates of higher quality, on the issues of urban unrest in 1968 and Vietnam in 1972 than on the other two dimensions. Our reading of the two elections leads us to believe that these two issues were discussed at length in the elections and that the major candidates took clear and distinguishable positions on them. This was not so obviously the case for Vietnam in 1968, when candidates were less clear in their statements, and for urban unrest in 1972, when the extent of discussion of the issue was lower and

the immediacy of the riots present in 1968 was four years removed by 1972. These expectations are clearly supported by the scaling estimated distributions.

The argument of the higher quality of the scaling estimations of the issues of urban unrest in 1968 and Vietnam in 1972 receives additional support from the proportion of negative "weighting" parameters, w_i, estimated. Recall that these negative weights indicate that the citizen's perceptions can best be "fit" with the candidate position estimates by a *negative* linear transformation. In 1968, a quite small number, less than 8 per cent are estimated for the urban unrest issue (see Table 5), as is also the case for the Vietnam issue dimension in 1972. However, 14.5 per cent are estimated for urban unrest in 1972 and an unfortunately large 21.3 per cent are estimated for the 1968 Vietnam dimension. Taking the later case for example, to a fairly large extent, citizens who responded by perceiving themselves as hawks have become transformed to doves and vice versa. Perhaps the most typical case would be a

hawkish respondent who claimed that Wallace is the most dovish candidate, while Humphrey or Johnson was most hawkish. Such an individual is likely to have a negative w_i and to have perceptions that correlate highly (but negatively) with the estimated candidate positions. The effect of the scaling estimations on such an individual is to treat him as if he were a dove, placing his estimated ideal point closest to Humphrey and Johnson. Thus, while it reverses his stated position, it places him closest to those candidates he claims to perceive as being closest to him.

Our interest in the individual parameter estimates, c_i and w_i, is primarily in their use in estimating the citizen's ideal point location on the set of common dimensions. The mean and variance of the distribution of ideal points for each dimension is found in Table 4. While we will be shortly putting the two 1968 distributions together to determine the distribution of citizens in the two dimensional issue space and relating their and the candidates' positions to the citizens' voting behavior, we can also use the aggregate distributions to make some cross-dimensional comparisons. The results here can only be considered tentative, especially for the 1972 comparisons which are based on nonidentical sets of respondents. With this in mind, however, we can use the variance of ideal point distributions on each dimension to determine the relative dispersion of candidate positions on them. In particular, the estimated dimensions are unique only up to a positive linear transformation. At present, the "unit of measurement" is determined by setting the distribution of candidate positions at the "(0,1)" standardization. Alternatively, we can linearly transform each dimension with respect to the distribution of ideal points so that all four have the same mean (say zero) and variance (in this case, we set all dimensions by the ideal point distribution of *urban unrest* in 1968). Therefore, all ideal point variances are set equal to .416. The new transformation of candidate position estimates would then represent the position of the candidate relative to the distribution of citizen's ideal points. Assuming that the distribution of ideal points remains constant from 1968 to 1972, we can then not only make cross dimensional comparisons within a given year but also look at the movement of candidates between elections.

The results of the above normalization, displayed in Figure 7, tend to agree with the journalistic interpretation of the candidate positions. In 1968, Humphrey, Johnson, and Nixon were very similar and close to the average citizen on both dimensions, but especially so on Vietnam. Wallace, however, was particularly extreme on urban unrest and only somewhat less so on Vietnam. In 1972, McGovern was closer to the average citizen on urban unrest than the Democratic candidates in 1968. McGovern, however, was quite extreme on Vietnam, and the Democratic party was also more extreme than its nominees were in 1968. Nixon is estimated to have been more conservative than the average citizen in 1968, but in 1972, he is estimated to have been even more so. More dramatic movement was displayed by Wallace who appeared to have moved toward the center on both dimensions in 1972. In 1968, he was located beyond one (marginal) standard deviation of the distribution of citizens from their mean on both dimensions. By 1972, he was estimated to be well within one deviation.

Candidates appear to have been more widely dispersed on urban unrest in 1968, especially the three major party figures. By 1972, the dispersal was much greater on Vietnam. Further, while all candidates appeared to be approximately distributed along a straight line in 1968, such a "unidimensional" distribution of candidates was much less adequate in 1972.

These over-time comparisons do lend support to the scaling methodology employed. Issues that appear to us to reflect important concerns on which relatively clear positions are taken lead to better scaling estimations. Moreover, the scaling estimates conform reasonably well with a priori expectations based on (not impartially) observing the two elections.

Two-Dimensional Distribution of Ideal Points—1968

The final stage in the analysis concerns the location of the citizen's ideal point in the common space. We have already examined the ideal-point distribution for each individual dimension. In this section, we will look at the distribution of ideal points for the two 1968 dimensions in greater detail. Beyond our concern with the extent of negative weight parameters estimated, w_i, our interest in the individual coefficients is primarily directed at ideal point placement, which results from solving equation (32). Figure 8 provides a scatterplot of the citizens' ideal-point locations in the common space for 1968. It is clear that most positions are estimated to be rather centrally located. Recalling the Monte Carlo results concerning the ideal point estimates, we found that ideal points were recovered rather well over all, and that this was particularly true for less extreme estimations. Thus, this sort of distribution in the real data is likely to be somewhat more precise than a distribution with a larger number of extreme cases.
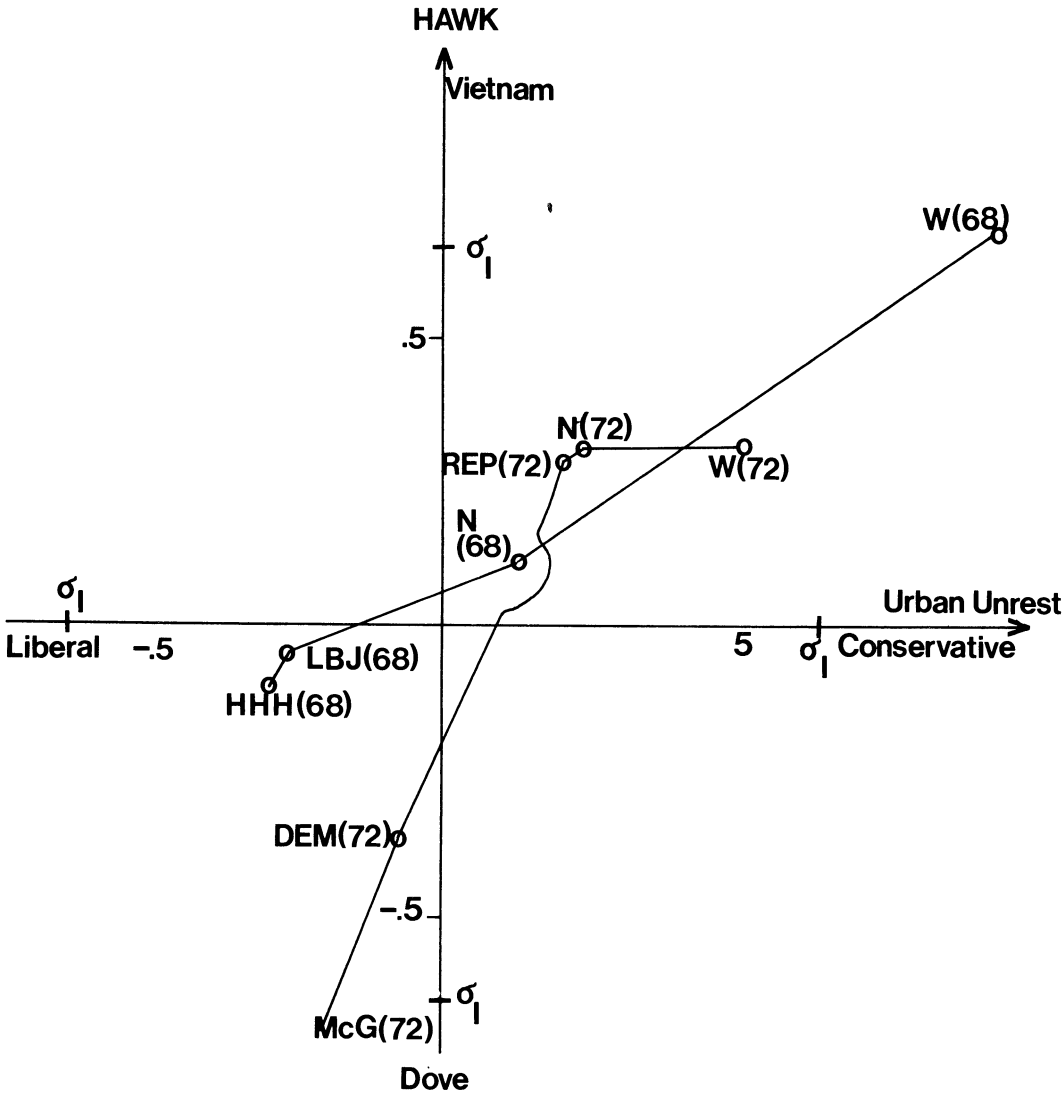
[a] σ_I indicates one (marginal) standard deviation of the ideal-point distribution from its mean of zero.

Figure 7. Candidate Positions on Urban Unrest and Vietnam in 1968 and 1972 with Dimensions Set so that Ideal Point Distributions Have Mean of Zero and Variance Equal to .416 (Based on Urban Unrest 1968)

The scaling estimated distribution of ideal points differs quite extensively from the distribution obtained using the raw data. In the seven-point scale data, the individual is constrained to be located at one of the $(7 \times 7) = 49$ points of ordered pairs of positions on the two dimensions, leading to a distribution of ideal points which looks like a "circus tent" As in the scaling estimated distribution, there is some concentration of citizens towards the center, the global mode being located at the point $(4,4)$, where one would place the large center pole of a circus tent. There were subsidiary modes, however, for each dimension at the end

points of 1 and 7. This resulted in local modes at the four corners and at the "center edges" of such pairs of positions as $(1,7)$ and $(4,7)$, thus heightening the tent-like appearance of the distribution. In the transformed space, the "lumpiness" has been considerably smoothed — the distribution looking, if anything more nearly unimodal. While the distribution is not exactly unimodal or symmetric, it is greatly altered from its unscaled counterpart.

Predicting the 1968 Vote

The above estimations of the candidate positions exhibit a substantial amount of sub-
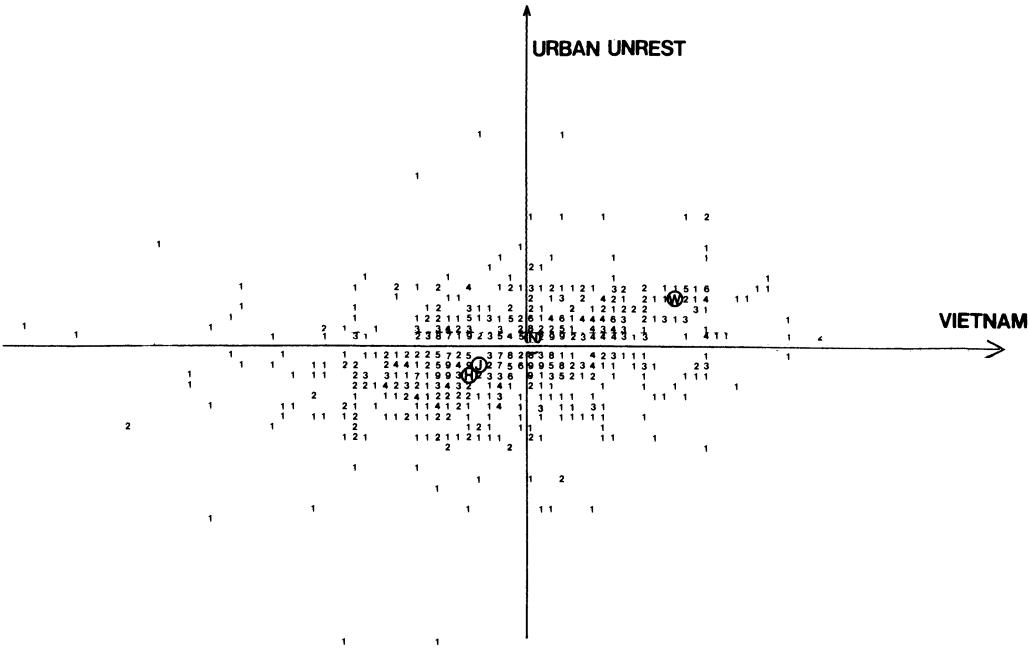
Figure 8. Two Dimensional Ideal Point Distribution, 1968

stantive reasonableness. Further, we have seen that the voter may be less confused over the candidate positions than would appear at first glance. These conclusions suggest that the assumptions of the spatial theories may constitute a reasonable model of voter behavior. The argument would be strengthened, however, if we could demonstrate that the scaling placements also lead to better predictions of electoral behavior.

To test this question, we make a simple prediction of the vote. In both the scaled and unscaled cases, the predicted vote is determined by computing the two-dimensional Euclidean distance between the voter's ideal point and the candidate position. The candidate position in the unscaled data is taken to be simply the mean perception of the sample on the seven point scales, while we of course use the estimated position for the scaled instance. We have also computed two forms of the actual vote. First, we use the actually reported vote itself and examine the relationship in two way contests between pairs of candidates. In this case, we look only at those voters who actually voted for one of the pair. We use a second measure based on the SRC's 100-point-thermometer measures of candidate evaluation. Here, the "vote" is measured by assuming that the individual would vote for whichever candidate in the pair stands higher on the thermometer preference measure. Finally, it should be

pointed out that this spatial prediction is a special case of the Downsian type spatial model, where it is assumed that the citizen votes, if at all, for the candidate closest to him in space.

All forms of the predictions are uniformly high, as reported in Table 6. It is noteworthy that voters seem to do remarkably well in conforming to the predictions of even this simple Downsian model, as witnessed particularly by the very high proportions of accurate predictions in those situations involving Wallace, where the citizens are generally presented two alternatives which are more easily distinguished. In this case, more than three quarters of the vote is correctly predicted.

Even with the uniformly high percentage of votes correctly predicted, the scaling based predictions consistently outperform the unscaled data. The marginal improvements run anywhere from about 2 to 8 per cent. To make sure that the improvement of the scaling results was not entirely due to the method of handling the negative weighting parameters, we reran the predictions including only those respondents who were estimated to have positive weights. These results, summarized in Table 6, indicate that the predictions continued to be improved consistently by the scaling estimations. The predictions were also improved in the unscaled data with the negative weights removed, although the scaling predictions still outperform

Table 6. 1968 Vote Predictions, Comparing the Scaled and Unscaled Data

	Entire Sample				Negative Weights Removed			
	Vote		Thermometers		Vote		Thermometers	
	Unscaled	Scaled	Unscaled	Scaled	Unscaled	Scaled	Unscaled	Scaled
H-N	66%	68%	65%	69%	68%	70%	67%	72%
N-W	79	87	76	82	85	89	83	85
N-W	75	78	79	83	77	81	84	85
	Both Issues Salient							
	Vote		Thermometers		Vote		Thermometers	
	Unscaled	Scaled	Unscaled	Scaled	Unscaled	Scaled	Unscaled	Scaled
H-N	67	77	69	78	72	78	74	79
H-W	86	94	81	87	93	96	86	88
N-W	78	80	83	86	81	82	88	87
	Both Issues of Low Saliency							
	Vote		Thermometers		Vote		Thermometers	
	Unscaled	Scaled	Unscaled	Scaled	Unscaled	Scaled	Unscaled	Scaled
H-N	63	61	62	64	67	66	67	68
H-W	74	80	75	81	79	83	80	85
N-W	76	77	79	80	80	82	83	82

the seven point scale based ones. This finding illustrates the often noted tendency of people to vote their perceptions more often than they vote on the basis of the "true" positions of the candidates.

The high level of success in predicting the vote is based on a very simple model relating distance to electoral behavior. Nonetheless, some error in the prediction may result from the implicit assumption that the two issue dimensions are of equal importance. Citizens differ, of course, in the importance they attach to different issues (an argument made most forcefully by Repass[10]). Issues that the citizen believes to be of little importance to him are likely to have a smaller impact on his behavior than more salient issues. To check this possibility and any effect it might have on our results we applied a simple control. In particular, we looked at (a) the subset of citizens who claimed that both issues were either "very important" or "the single most important thing in the election" in the two questions following the seven point scale responses, and (b) the subset of citizens who claimed that neither issue was very important. We then reran the same predictions for these two groups, both for the whole subsets and for the subsets with negative weights removed. These results again conform to all our expectations. The scaling-based predictions are consistently higher, and

[10]See David E. Repass, "Issue Salience and Party Choice," *American Political Science Review*, 65 (June, 1971), 389–400.

sometimes greatly so, in all but a couple of minor instances, and the saliency control improved the predictions uniformly when both issues were thought to be of importance. In fact, in the "clearest" case of predicting the actual vote between Humphrey and Wallace using the scaling estimates with high saliency control and no negative weights, 96 per cent of the vote was accurately predicted (and 94 per cent when negative parameters were included).

Scaling Estimates—1972

We conclude this paper with a brief report of the scaling estimates for the 10 seven-point issue scales administered to the entire sample in 1972. The estimates are summarized in Table 7 and Figure 9.

The most prominent feature of the estimations is the general overall similarity of the candidates' relative positioning on most dimensions. Without exception, McGovern is clearly the most liberal candidate, followed by the estimated location of the Democratic party. As we saw earlier, McGovern is always separated from his party by a noticeable distance. Obviously, the Democrats nominated a candidate in 1972 who was not seen as a typical party member. The distance between McGovern and the Democratic party, perceptible though it may be, is always less than the sometimes very large relative distance between the Democrats and the third most liberal candidate. Thinking of the mean of the candidate distribution (i.e., 0) as the center of gravity, we see that it takes the three remaining candidates on the opposite
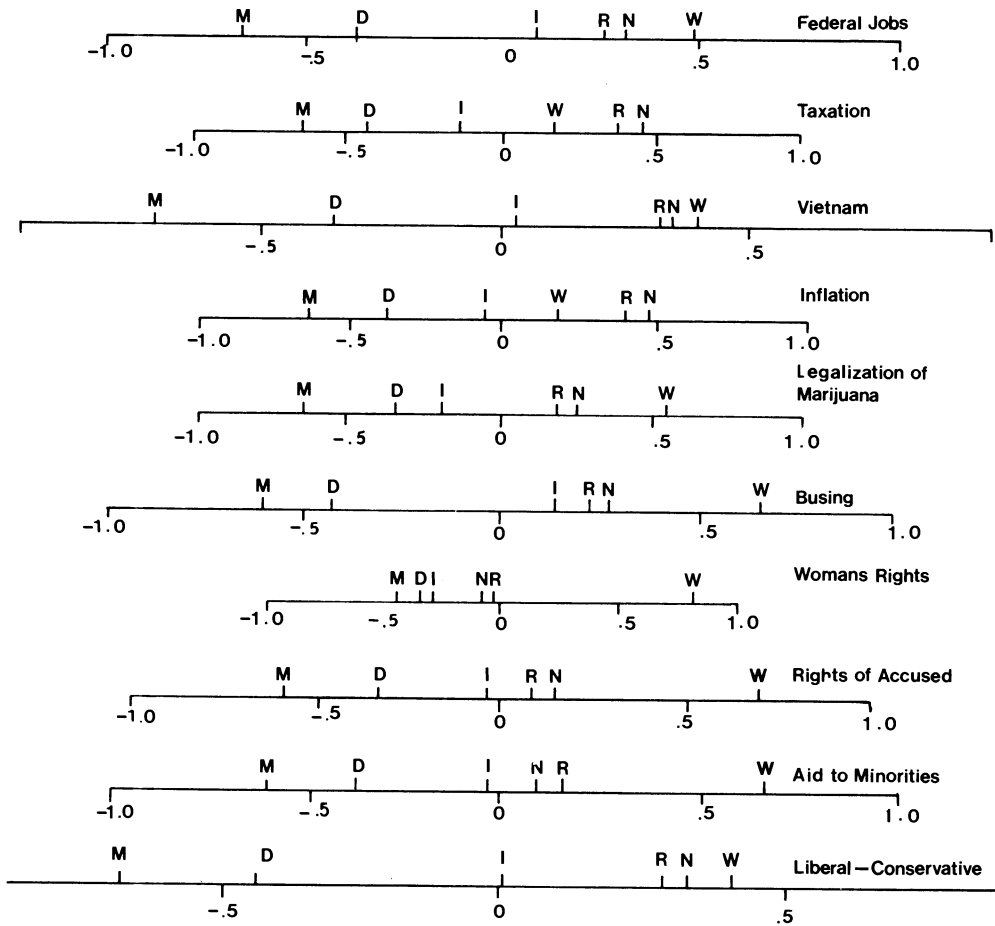
Figure 9. Candidate Positions in 1972

side of the mean to balance the Democratic Party and its nominee, excepting only the issue of women's rights. This contrasts with the estimates in 1968, when Wallace was estimated to be so extreme that the two Democratic candidates were not sufficient to balance him. In opposition to the distance separating McGovern and his party, the Republican party and Nixon are estimated as very similar on all ten dimensions. We cannot tell from these data whether this means that Nixon is seen as a (perhaps *the*) typical Republican or whether it means that a "typical" party nominee, especially an incumbent president, defines much of what the party stands for in the eyes of the electorate. The consistent tendency (19 out of 20 instances) to find the party closer to the mean of the distribution than its nominee may imply that parties evoke less clear cut perceptions than single individuals. This perception may be true in the eyes of each citizen, or it may be based on the obvious regional and other background differences within each party that evoke different perceptions, tending to "cancel out" extreme positionings.

The position of Wallace presents the most striking cross-dimensional differences. On most issues, Wallace is the most conservative candidate (although, as we have seen, perhaps not so obviously in 1972 as in 1968). On such issues as aid to minorities, the rights of the accused, busing, and especially women's rights, the distance separating Wallace from the two Republican stimuli is very large. The difference is much smaller on such issues as Vietnam and the general liberal/conservative continuum. Most importantly, on the two economic issues, Wallace is perceived to be only slightly on the conservative side of the mean, in a middle position dividing the Democrats from the Republicans. Again, the scaling estimated positions seem to reflect rather well the nature of the 1972 election. For example, the positioning of Wallace on the issues seems most reasonable in some detail. If he is viewed as a populist governor, his moderate economic stance could be expected.

Table 7. 1972 Candidate Position Estimates

	RMN	McG	GCW	Dem.	Rep.	I
Federal Jobs	.307[a]	−.663	.474	−.401	.283	.106
	.315	.383	.438	.345	.308	.871
Taxation	.462	−.647	.166	−.402	.421	−.110
	.377	.462	.494	.424	.383	1.099
Vietnam	.361	−.705	.376	−.355	.326	.046
	.290	.338	.430	.327	.277	.694
Inflation	.457	−.628	.198	−.436	.409	−.028
	.402	.454	.515	.431	.396	1.100
Legalization of Marijuana	.290	−.672	.551	−.354	.185	−.159
	.307	.410	.394	.355	.322	1.140
Busing for Integration	.240	−.587	.634	−.420	.133	.248
	.359	.406	.391	.365	.322	.857
Women's Rights	−.031	−.464	.827	−.314	−.018	−.320
	.382	.437	.559	.358	.365	1.433
Rights of Accused	.139	−.571	.714	−.370	.089	−.045
	.344	.414	.448	.325	.324	.927
Aid to Minorities	.116	−.606	.699	−.337	.128	−.008
	.303	.361	.406	.326	.297	.876
Liberal/Conservative Continuum	.343	−.669	.415	−.405	.315	.006
	.261	.348	.438	.311	.272	.597

[a]Entries are the scaled candidate or average ideal-point position on top and standard deviation underneath.

The distribution of ideal points has been indicated in Table 7 and Figure 11. Since the distribution is only an aggregate measure, the "average citizen" not surprisingly is located rather near the center on all dimensions, dividing the two parties and nominees. Fairly consistently, then, as would be expected, the Democratic party and candidate are seen as at least somewhat liberal (this holding in 1968 as well), while the Republicans tend to be somewhat more conservative on all issues (again, the same is true in 1968). In 1972, the average citizen is consistently much closer to Nixon than McGovern. The voters tend to be, on average, slightly on the conservative side of the mean of the candidate distribution, the principal exceptions being women's rights, taxation and the legalization of marijuana. This is not to say, of course, that the average citizen leans towards the legalization of marijuana. In fact, they are strongly conservative on the seven-point scale data. They simply see the three conservative candidates as even more conservative than themselves. These aggregate figures are, of course, only very general indicators. Our principal interest in the distribution of ideal points in this section is to rescale the unit of measurement of the dimensions, as we did previously. Recall that this would allow us to make cross-dimensional comparisons, at least to the extent of indicating the dispersal of candidates through the distribution of citizens. The relationship between the variance of ideal points and the dispersion of candidates should be inverse. That is, the smaller the variance of ideal points estimated in the scaling technique, the more widely dispersed the candidates relative to the citizens. Therefore, we have taken the issue with the smallest ideal-point variance (the liberal/conservative dimension) and set all other dimensions to have the same variance, keeping the candidate positions the same on the "base" issue dimension and narrowing the distance on all others by the appropriate proportion. As Figure 11 shows, the liberal/conservative dimension and Vietnam are quite similar in their relatively large dispersion of candidates (and especially McGovern). On the other hand, on the issue of women's rights, the candidates assume virtually identical positions. This issue has "shrunk" to less than one-half of its original size.

Conclusion

This article has attempted to apply a probabilistic model of the individual's response to questions on candidate perception to "factor out" the influences due to variations in reaction to the response task. This method then estimates the candidate positions by using the common part of individual perceptions. We have applied the method to the 1968 and 1972 election studies with reasonable success, in that the estimates correspond to a great degree with the a priori expectations, and explain voting behavior with a high degree of accuracy.