*The Relationship between Information, Ideology, and Voting Behavior**

Thomas R. Palfrey, *California Institute of Technology*
Keith T. Poole, *Carnegie-Mellon University*

The question of voter sophistication is important for understanding voter and candidate behavior in mass elections. We develop an index of voter information—based on perceptual data—and find that it is significantly related to ideological extremism and voting behavior. Individuals with a high level of information tend to be more extreme than those with low levels and are much more likely to vote.

This paper presents an empirical study of several important relationships between voter information, voter preference, and electoral choice. One relationship involves the connection between voter information and the strength of voter preferences between candidates. Less-informed voters will tend to exhibit more indifference. In the extreme, if a voter is completely uninformed about the candidates' positions on issues and other potentially important characteristics of the candidates, then, even if the voter has relatively strongly held preferences on the issues, the voter has little basis for expressing a strong preference for one candidate over the other. Alternatively, if a voter has "rough" ideological information, such as the candidate's party, but is not well enough informed to understand how this translates (albeit imperfectly) into likely issue positions, we again can hardly expect the voter to express strong preferences between candidates. This latter source of indifference, while due partly to a lack of information is also due partly to the voter's lack of issue consistency. In either case, there is good reason to expect the relative strength of preferences of such voters for candidates to be weak.

Second, we may similarly expect voters who are relatively uninformed about the issues to express either inconsistent or noncomittal preferences on these issues, as compared to well-informed voters. If this is the case, estimation techniques would tend to measure left-right preferences of uninformed voters as moderate or centrist. This is not to say that all informed voters are extremist and all uninformed voters are moderates, rather that there is good reason a priori to expect the

*This work was supported in part by National Science Foundation grants SES-831-390 and IST-8406296. We wish to acknowledge useful conversations with Randall Calvert, Robert Miller, Howard Rosenthal, and Kenneth Singleton.

measured left-right preferences of informed voters to be on average less centrist than the measured left-right preferences of uninformed voters.

Third, if the above relationships hold between the level of a voter's political knowledge and his or her preferences both over specific issues and between candidates, it suggests that we should be able to measure several significant effects of voter information on the individual vote decision. First, by the above indifference argument, uninformed voters will be less likely to vote (Downs, 1957, ch. 14). Second, even if they do vote, they will be less responsive to candidate issue taking, since they do not accurately perceive the candidate's position. This would lead to voting behavior which exhibits more (apparent) randomness, at least when one tries to predict voting behavior based on measured preferences over issues and toward candidates.[1]

Summarizing, this paper studies three sets of relationships between an individual voter's information (or level of political knowledge) and (1) the voter's preference over candidates; (2) the voter's measured preferences on issues and general left-right positions; and (3) vote choice (including nonvoting).

The basic approach we take is based on a spatial model of voter preferences and voter perceptions of candidate positions. We measure the level of information of a voter by comparing the "true" (estimated) spatial locations of the candidates with the voter's perception of these locations. To do this we use the Aldrich-McKelvey (1977) (henceforth A-M) scaling procedure. To study the relationships between information and vote choice, we estimate a conditional logit model of vote choice, including nonvoting as an option,[2] using ICPSR survey data from the 1980 presidential election.

[1] If the preferences and behavior of uninformed voters systematically differ from the preferences and behavior of informed voters, this suggests a number of potentially useful extensions and modifications of current theoretical models of candidate competition (including the spatial model). For example, we might expect candidate competition—at least the issue-taking aspect—to lead to spatial positions which are unrepresentative of the preferences of poorly informed voters. For example, if candidates are typically located on opposite sides of the center, those voters with moderate views have little incentive to gather and process information about the competing candidates, since a voter can rationally expect to be indifferent anyway. Since gathering and processing information is costly, these voters may be rationally uninformed. Candidates, being aware of the lack of information of "moderate" voters may be led to adopt less-centrist positions than those typically predicted in the simple Downsian model of competition. This could result in nonmedian outcomes and help explain nonconvergent candidate behavior. This contrasts sharply with Converse's (1966) hypothesis that "alterations in governing party depend disproportionately on shifts in [the uninformed voters'] sentiment [and] the less-involved and less-informed voters are disproportionately represented" (pp. 136–37).

[2] The estimation of the conditional logit model requires data about nonvoters as well as data about voters. However, it is well known that the reported turnout data from

The paper unfolds as follows: section 1 reexamines the A-M technique for recovering candidates, parties, and voters from perceptual data. We demonstrate that the A-M technique leads to systematic bias (regression to the mean) in scaling *voters* (not candidates) if there is *variation* across voters in how informed they are about the political stimuli. This establishes a purely methodological reason why we would expect *scaled* preferences of less-informed voters to be less extreme. Monte Carlo studies are reported in this section which further demonstrate that, despite this bias, the scaled *stimuli* locations are very accurate even with considerable variation in the information level of voters.

In section 2 we define our index of voter information using estimates of the locations of political stimuli generated by the A-M technique. The index is essentially an average measure of correlation between the (estimated) true location of candidates and parties and a voter's (reported) perception of these locations. This index is compared with a measure of "attitude consistency" similar to the one used by Norman Nie, Sidney Verba, and John Petrocik (1979). Comparisons are reported between informed and uninformed voters regarding their left-right spatial distribution and their strength of preference between candidates. The hypothesized relationships are clearly evidenced in the data.

Section 3 presents the conditional logit analysis of the effect of information on turnout and vote choice. Information is found to have a significant positive effect on turnout, as expected. We then analyze the effect of voter information levels on classification accuracy in order to identify a relationship between voter information and "randomness" of the voting decision.

The Aldrich-McKelvey Scaling Procedure with Heteroscedastic Error

Description of the Aldrich-McKelvey Procedure

The A-M scaling procedure is a technique for constructing estimates of candidate locations and voter ideal points on a dimension (or issue), using only data about voter perceptions of the candidate locations. A loose interpretation of the final estimate of a candidate's location is that it is a weighted average of the voters' perceptions of the candidate's "left-right" position on the issue. Because different voters may report their perceptions of the left-right scale differently (e.g., some voters may have a tendency to exaggerate perceived differences when responding to surveys), the method requires simultaneous estimation of these individual

---

surveys are unreliable, due to heavy overreporting of turnout. Fortunately, the individual turnout responses in the 1980 survey we analyze were directly validated. We compare the estimation results using both the unvalidated and the validated survey responses.

"distortion parameters." John Aldrich and Richard McKelvey demonstrate that the resulting two-step method is essentially equivalent to "a principal components solution to the candidate estimates together with a regression estimate of the citizen parameters" (1977, p. 112).

The basic model used by A-M assumes that individual i's perception of candidate ("stimulus") j is given by

$$Y_{ij} = Z_j + e_{ij}$$

where Z_j is the true location of j and e_{ij} is a random variable which has zero expectation, positive variance which is independent of i and j (homoscedastic), and zero covariance across the i's and j's.

The rationale for the (perceptual) error structure in the model is that lack of information about candidates on the part of voters leads to misperceptions. However, Aldrich and McKelvey assume that all voters are *equally* imperfectly informed about the candidates, an assumption which we find implausible.

Aldrich and McKelvey introduce two distortion parameters, c_i, w_i, which transform the perceived candidate position into a reported candidate position, Z_{ij}, according to:

$$Z_{ij} = \frac{1}{w_i}(Y_{ij} - c_i).$$

A least-squares minimization procedure is then used to obtain estimates of

$$\{Z_j\}_{j=1}^{J} \text{ and } \{w_i,\ c_i\}_{i=1}^{I}.$$

It is possible that some w_i's will be estimated to be negative, which may occur when a voter reports a perception of the stimuli that is negatively correlated with the estimated locations of the stimuli.[3] After obtaining estimates of the locations of the stimuli and the individual distortion parameters, the latter are then applied to transform self-reported locations, thereby obtaining estimates of the individual voter's "ideal points."

Monte Carlo Results with Heteroscedastic Error

Aldrich and McKelvey (1977) performed a Monte Carlo study of their procedure under the assumption of homoscedastic error and found that the "candidate positions are recovered exceptionally well" (p. 118). We echo this finding with heteroscedastic error. In other words, their procedure, which assumes homoscedastic error, is effective at recovering candidate locations, even with heteroscedastic error.

[3] A negative weight means that the individual reverses most positions of the stimuli—or, as Aldrich and McKelvey (1977, p. 116) put it, the individual perceives a "mirror image."

In our Monte Carlo study we used six stimuli and 500 individuals because there were six stimuli (Reagan, Carter, Anderson, Kennedy, Democratic party, Republican party) in common to nine seven-point scales in the 1980 election study and 500 or more scalable respondents per scale.[4] In our first set of experiments each individual's perceptual variance was drawn randomly from a uniform distribution.[5] The error was homoscedastic across each individual's judgments of the stimuli but heteroscedastic across individuals. We found that, even at enormous error levels (e.g., when over 90 percent of the individuals had error standard deviations greater than the distance spanned by the stimulus configuration), the stimulus configuration was accurately recovered.

In our second set of experiments we randomly selected each individual's perceptual variance from a uniform distribution and then weighted it, for each stimulus, by the individual's "true" squared distance to the stimulus. The error was, therefore, heteroscedastic within and across individuals. We again found that, even at very high levels of error, the stimulus configuration was recovered well.

Given the robustness of the procedure in recovering the stimuli configuration at reasonable noise levels, we focused our attention in our third set of experiments on the accuracy of the recovery of the individuals. To do this we split the individuals into two groups: (1) informed, those with low variance in perceptual error, and (2) uninformed, those with high variance in perceptual error. The error within each group was homoscedastic. That is, within each group, the error was homoscedastic within and across individuals. We then varied the levels of "low" and "high" variance and the ratio of informed to uninformed. We set the overall noise level as closely as possible to that encountered in actual seven-point scale data. In particular, we used a noise level that resulted in approximately 18 percent of the individuals with estimated weights ($\hat{w}_i$) that were negative—the same as that found for the ICPSR 1980 defense-spending scale.

The weights estimated for the uninformed group were all near zero—especially at higher noise levels. Because the estimated individual transformation coefficients, $\hat{c}_i$ and $\hat{w}_i$, are recovered from a regression of the reported on the estimated candidate locations, $\hat{w}_i$ tends to zero as the perceptual error increases. The individual's scaled position is found by applying the linear transformation to the individual's reported ideal

[4] The nine scales were: Liberal-Conservative (preelection and postelection); government spending; inflation; defense spending; government aid to minority groups; Russia; women's equal role; and government guaranteed jobs.

[5] A detailed explanation of the Monte Carlo study is available from the authors on request.

point. If $\hat{w}_i$ is near zero, the individual is recovered near the center of the space. Thus, the uninformed group, regardless of their "true" distribution, will be mapped toward the center of the space.

Although we found that the recovery of the informed group was much better than the uninformed group, neither recovery was totally free of bias. As Aldrich and McKelvey (1977) point out, because of regression toward the mean, "extremists will be estimated as somewhat less extreme than they actually are" (p. 120). In addition, because the scales have labeled endpoints, individuals who consider themselves more extreme than the labels have to place themselves at the endpoints. This, coupled with the mapping of the uninformed toward the center of the dimension, leads to a tendency of the procedure to recover a less-dispersed distribution of scaled ideal points.

To illustrate this problem we ran a fourth set of experiments in which the true distribution of the individual ideal points was bimodal. Figure 1 shows the true and recovered distributions for all, informed, and uninformed individuals, respectively. The plots are from an experiment performed with an overall noise level similar to that encountered in actual seven-point scale data.

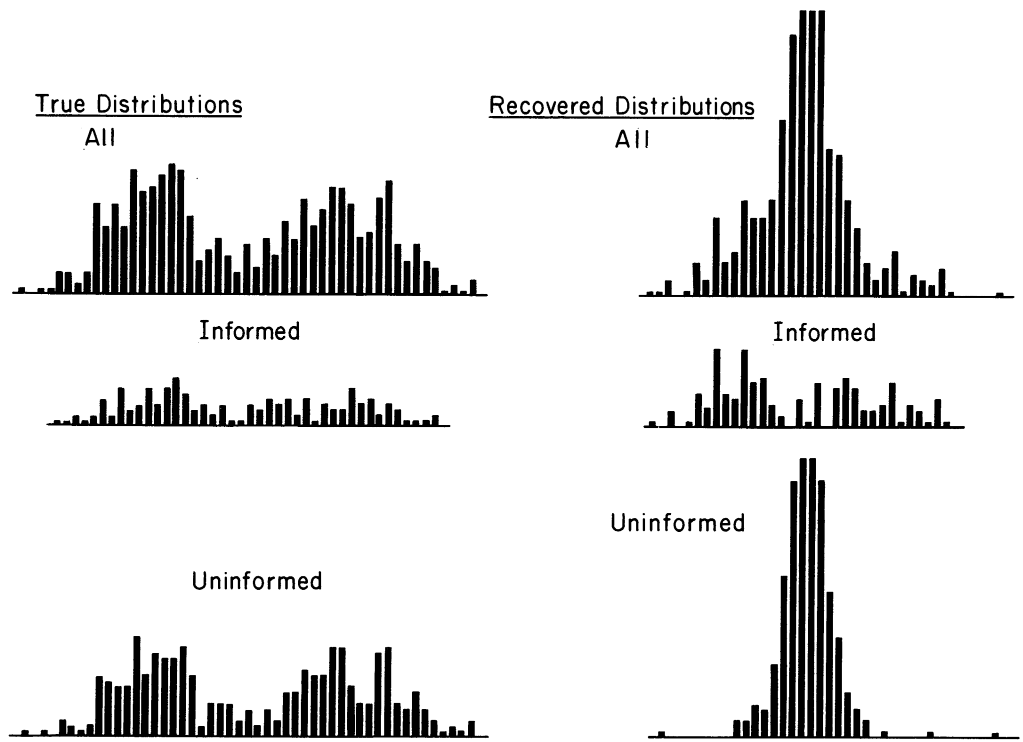
In sum, our Monte Carlo experiments, in conjunction with those performed by Aldrich and McKelvey, show that the recovery of the candidate coordinates is very robust and, at the error levels encountered in ICPSR data, is highly reliable. However, the recovered individual locations must be analyzed with some caution. In particular, the regression to the mean effect is very severe for relatively uninformed voters. This can easily lead to a recovered unimodal distribution of ideal points when the "true" distribution is bimodal.

The Information Index

Given the quality of the recovery of the "true" stimuli configuration by the A-M scaling technique, a reasonable choice for a proxy variable to measure an individual's level of political information is the correlation between the individual's perceived location of the candidates/parties on a seven-point scale and the scaled estimate of the candidates/parties locations. The higher the correlation, the more informed the individual is about the "true" candidate locations.

To lend some stability to our information index, we used only those respondents who were scaled on at least three of the nine seven-point scales asked in 1980. We computed the mean correlation for each respondent, and we used this as our proxy variable for the level of political information.

FIGURE 1
Monte Carlo Simulation



In order to verify whether our measure is indeed capturing what we want, we performed several checks. In the first, a probit equation was estimated, using the "who controlled the House of Representatives" question[6] in the postelection study as the dependent variable (correct, incorrect) and our information index as the independent variable. The response "Democrat" was coded as "correct"; the responses "Republican" and "Don't Know" were both coded as "incorrect." The estimated probit equation was

Who controls House?

$$\begin{matrix}\text{Correct} & <0 \\ \text{Incorrect} & >0\end{matrix} = \begin{matrix}1.37 \\ (9.88)\end{matrix} - \begin{matrix}0.81 \\ (4.87)\end{matrix} \begin{bmatrix}\text{info} \\ \text{level}\end{bmatrix}.$$

The estimated R-square (see McKelvey and Zavonia, 1975) was .132, and the t-statistics are shown in parentheses. Our information-level index, which is in fact a measure of how well individuals perceive political actors' positions on issues, is significantly related to objective knowledge of the political system.

As a second check, we regressed the information index on a measure of issue consistency. We defined an individual's level of issue consistency[7] as the standard deviation of the self-placements on the nine seven-point scales.[8] The OLS equation was

$$\begin{matrix}\text{issue} \\ \text{consistency}\end{matrix} = \begin{matrix}1.68 \\ (34.39)\end{matrix} - \begin{matrix}0.56 \\ (7.73)\end{matrix} \begin{bmatrix}\text{info} \\ \text{level}\end{bmatrix}.$$

The R-square was .086, and the t-statistics are shown in parentheses. The information index is significantly related to the coherence of an individual's attitudes as well as to his or her objective knowledge of the political world.

As a third check, we computed the mean information index for the three categories of respondents identified by Lee Sigelman (1982): (1) actual voters—those respondents who reported they had voted and actually did; (2) misreporters—those who reported they had voted and actually did not; and (3) admitted nonvoters—those who reported they

[6] Variable number 1027.

[7] Echoing Nie, Verba, and Petrocik (1979, p. 123), we regard the labels "attitude consistency," "attitude coherence," and "attitude constraint" as interchangeable with "issue consistency."

[8] We subtracted the self-placements from 7 for the government spending and inflation scales because, in our judgment, in contrast to the other seven scales, scale position 1 was the "conservative" position on these two scales.

had not voted and actually did not.[9] The mean information value for actual voters was .351 with a standard deviation of .074 ($N = 459$), for misreporters the mean was .347 with a standard deviation of .067 ($N = 117$), and for admitted nonvoters the mean was .309 with a standard deviation of .073 ($N = 140$). Evidently the information index is also related to *reported* turnout.

Confidence in the index is further bolstered by the fact that the order of the group means is the same as that found by Sigelman (1982) using a discriminant analysis of a battery of socioeconomic variables. In addition, we also found that in our conditional logit model of voting behavior below, the estimated candidate parameters were essentially the same with or without the misreporters in the analysis. Based on the analysis of the 1978 validated study, Sigelman concluded that "our understanding of the factors that influence voting appears to be largely unaffected by the misreporting phenomenon" (p. 55). We echo this finding as it applies to *conditional* (on turnout) vote choice. However, as we shall see below, misreporting significantly affects the estimates of several turnout parameters and therefore may have serious consequences for our understanding of important factors (such as information) that influence the turnout decision.

Using the information index, it is straightforward to address our first hypothesis about the relationship between voter information and indifference. We measured indifference in two ways: (1) the absolute value of the difference between the Carter and Reagan feeling thermometers and (2) the absolute value of the difference between the respondent's distance to Reagan and the respondent's distance to Carter on the preelection liberal-conservative seven-point scale (see next section). Our hypothesis is confirmed. The correlation between the information index and indifference as measured by thermometer differences was .15. The correlation for the liberal-conservative seven-point scale indifference measure was .27.

Another question that can be investigated using this measure is whether the distribution of preferences of informed voters is the same or different from the distribution of uninformed voters. We find that there is an important qualitative difference between the two distributions, which is consistent with some of the theoretical issues raised earlier in this paper. *Informed voters tend to be located at the extremes of the political spectrum. Uninformed voters tend to be in the center.* The top histogram

[9] None of the respondents in the missing fourth category—those who reported they did not vote and actually voted—had enough responses to the seven-point scales to calculate an information index.

in Figure 2 shows the distribution of our sample of 644 respondents, using the locations recovered by applying the A-M procedure to preelection self-reported locations on the liberal-conservative seven-point scale. The first column of Table 1 shows the actual distribution of self-reports across this dimension. Similarly, the middle histogram in Figure 2 and the second column of Table 1 show the corresponding distributions of 154 "informed" (defined as $r > .95$) respondents, and the bottom histogram and the third column of Table 1 show the distributions of 143 "uninformed" (defined as $r < .50$) respondents. With Figure 1 we illustrated the fact that the scaled distribution of all individuals is misleading because of the presence of the uninformed. We encounter the same possibility here.

FIGURE 2
1980 Liberal-Conservative A-M Scaled Locations

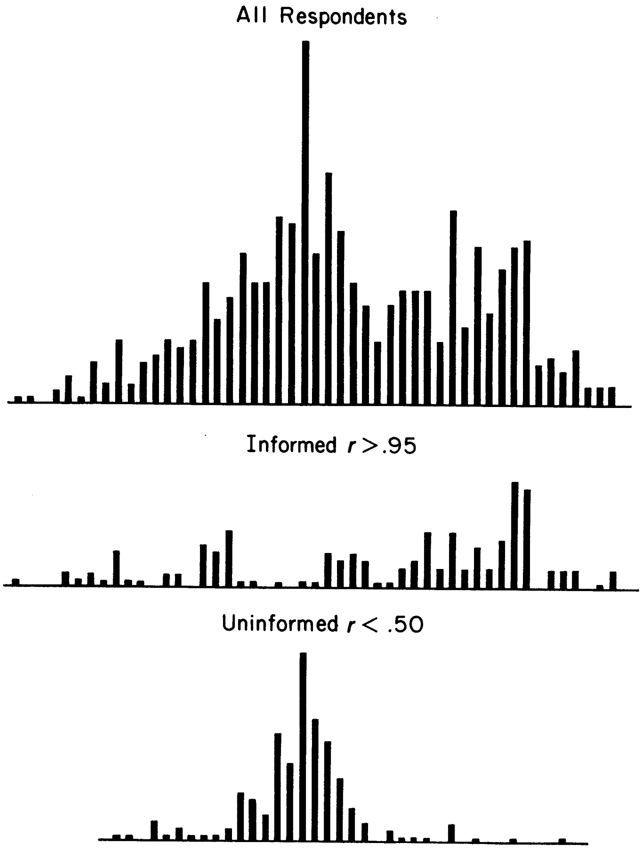


TABLE 1

Distributions across 1980
Preelection Liberal-Conservative Seven-Point Scale
(In Percentages)

		All	Informed*	Uninformed**
Liberal	(1)	1.4	2.0	1.4
	(2)	7.8	11.0	4.2
	(3)	12.1	11.0	12.0
	(4)	31.6	16.2	48.6
	(5)	21.9	25.3	21.1
	(6)	22.6	32.5	9.9
Conservative	(7)	2.6	2.0	2.8
	$N=$	644	154	143

NOTE: *Correlation greater than .95; **correlation less than .50.

The overall recovered distribution shows a large mass of individuals in the center of the space; no such large mass is present in the distribution of the informed. However, the similar distributional pattern in Table 1, based on actual self-reports, indicates that the strong empirical relationship between information and polarization is not just a methodological artifact of the A-M scaling procedure.

The Effect of Voter Information on Turnout and Choice

Information and turnout should be positively correlated. The more an individual knows about politics the likelier it is that the individual will vote. However, given that an individual has decided to vote, information should play a lesser role in candidate choice.

To test these propositions we shall use the postelection validated truthful choices of the 1980 respondents as the dependent variables in a conditional logit model. Specifically, for each individual we shall estimate four probabilities corresponding to the four choices in the 1980 presidential election: not voting, Reagan, Anderson, Carter. We will use maximum likelihood to obtain estimates of these probabilities.

The logit model assumes that the probability that respondent i did not vote is

$$q_{i4} = Pr(\text{not vote}) = e^{\delta' X_i} / (1 + e^{\delta' X_i}) \quad (1)$$

where δ is a vector of coefficients which we estimate and X_i is the vector of the respondent's values on a set of independent variables. The probability that respondent i voted for Reagan is the product of the conditional probability that, given that i decides to vote, i votes for Reagan times the probability that i votes. That is

$$q_{i1} = Pr(\text{vote for Reagan}) = \left[\frac{e^{\beta_1' X_i}}{1 + e^{\beta_1' X_i} + e^{\beta_2' X_i}}\right]\left[\frac{1}{1 + e^{\delta' X_i}}\right]. \tag{2}$$

Similarly, for Anderson and Carter:

$$q_{i2} = Pr(\text{vote for Anderson}) = \left[\frac{e^{\beta_2' X_i}}{1 + e^{\beta_1' X_i} + e^{\beta_2' X_i}}\right]\left[\frac{1}{1 + e^{\delta' X_i}}\right] \tag{3}$$

and

$$q_{i3} = Pr(\text{vote for Carter}) = \left[\frac{1}{1 + e^{\beta_1' X_i} + e^{\beta_2' X_i}}\right]\left[\frac{1}{1 + e^{\delta' X_i}}\right] \tag{4}$$

where β_1, β_2, and δ are vectors of coefficients which we estimate. Our likelihood function is therefore

$$L(\delta, \beta_1, \beta_2) = \prod_{i=1}^{N} \prod_{k=1}^{4} q_{ik}^{c_{ik}} \tag{5}$$

where $c_{ik} = 1$ if i chooses k and 0 otherwise.

The logit model we estimate utilizes seven independent variables—our information measure and six variables related to voter preferences: the absolute strength of party identification; utility for Reagan; utility for Anderson; utility for Carter; alienation; and indifference. The reason we included these variables and excluded others, such as socioeconomic status, is that our model of voting behavior suggests that the relevant variables to include are measures of information and preferences. While other variables such as reported income and sex may improve the "fit" slightly, our simple theoretical model based on information and preferences suggests no a priori reason to include these variables.[10]

The strength of party identification is computed by taking the absolute value of 3 minus the ICPSR strength of party ID variable which ranges from 0 (strong Democrat) to 6 (strong Republican). We measured utility in two ways: (1) a constant minus the distance between an individual's self-placement and his/her reported perceived locations of the candidates on the preelection liberal-conservative seven-point scale and (2) the candidate feeling thermometers.[11] Both measures were

[10] For completeness, we tried including a number of additional independent variables to see if they would have an effect. For the most part, our suspicion that these would have little effect was confirmed. There were two minor exceptions. Efficacy had a very small effect and did not improve classification accuracy. Education also had a relatively small effect. As expected, as education decreases, the probability of not voting increases.

[11] A third possible measure of utility could be constructed by using the distance between an individual's *scaled* location and each candidate's *scaled* location, where these scaled locations are recovered using the A-M procedure discussed in section 2. However, the Monte Carlo results reported in section 3 indicate that this will lead to systematic error

normalized to have a range from 0 to 1. Alienation was taken to be the maximum of the three candidate utilities, and indifference was defined as the absolute difference between the utility for Reagan and the utility for Carter. Alienation and indifference were used only in the nonvoting equation. The results for the estimations using both measures of utility are reported in Tables 2–7.[12]

TABLE 2

Estimates of Distance-Utility Model

	Parameter	Estimate		*t*-Statistic	
		Reported	Validated	Reported	Validated
	Constant	−0.35	−0.50	−0.34	−0.40
	Info	1.05	0.79	2.45	1.58
	Party	−0.32	−0.20	−1.68	−0.89
Reagan (β_1)	Utility RR	6.79	7.50	8.98	6.96
	Utility JA	−0.54	0.06	−0.79	0.06
	Utility JC	−4.87	−5.99	−5.29	−5.08
	Constant	−1.12	−0.71	−0.85	−0.50
	Info	0.79	0.45	1.37	0.69
	Party	−0.55	−0.50	−2.62	−2.06
Anderson (β_2)	Utility RR	0.90	1.39	1.23	1.68
	Utility JA	3.80	4.05	3.12	3.02
	Utility JC	−3.92	−4.77	−3.76	−3.74
	Constant	2.56	1.33	2.56	1.65
	Info	−1.63	−1.11	−4.07	−3.53
	Party	−0.39	−0.35	−2.45	−3.12
Nonvoting (δ)	Utility RR	−1.35	−1.54	−1.47	−2.37
	Utility JA	−0.99	−1.06	−1.13	−1.71
	Utility JC	−1.54	−0.68	−1.75	−1.05
	Alienation	0.76	2.19	0.39	1.49
	Indiff	−2.62	−1.43	−2.46	−1.93
	Total log likelihood	−480.09	−521.31		
	N	521	514		

in the utility measure especially for low information voters. As a check, we estimated the model using this scaled utility measure. Not surprisingly, the results were poor (both in terms of fit and classification accuracy) compared to the two measures we report in the text.

[12] We used the Berndt, Hall, Hall, and Hausman (1974) method to estimate the parameters.

A well-known problem with the survey data we are using is that many respondents fail to report correctly their actual vote decision. While the privacy of a secret ballot prevents us from detecting whether respondents who actually voted correctly report for whom they voted, turnout data on an individual level is a matter of public record. Fortunately, the individual turnout responses for the 1980 survey we are using were validated (i.e., corrected by examining election records). To give unfamiliar readers an indication of the magnitude of this problem, approximately 40 percent of nonvoters in the sample we used inaccurately reported that they had voted when, in fact, they did not vote. Since one of the most important aspects of our empirical study involves the turnout decision, this could potentially have a major impact on our results. For completeness we present estimation results for both the validated and the unvalidated data. The results are, for the most part, quite similar, but a few interesting differences were detected and are discussed below.

The parameter estimates in Tables 2 and 3 must be interpreted with some caution. Because we are using a conditional logit analysis, the parameter estimates in the Reagan, Anderson, and nonvoting equations cannot be directly interpreted in the same way as a coefficient in a linear regression; some additional computation is required to make statements about the effect of a marginal change in one of the independent variables on the probabilities of voting for Anderson, Carter, Reagan, or nonvoting. To see why this additional computation is needed, notice that the vote probability equations (2), (3), and (4) include the nonvoting parameters (δ). However, the t-statistics of the estimated parameters are meaningful. Furthermore, by taking partial derivatives of equations (1)–(4) one obtains simple formulas to compute estimates of the marginal effects of the independent variables from the estimated nonvoting and voting parameters. As an aid to interpretation, we display in Tables 4 and 5 these estimated first partial derivatives for each of the four probabilities, evaluated at the sample means of the independent variables.

As an example of how to compute the estimated marginal effect (partial derivative) of changing one of the independent variables on the probability of one of the choices, consider the effect of an increase in X_1 (information) on q_3 (the probability of voting for Carter), for an "average" voter (i.e., a voter for whom X_i equals the sample mean). Using equation (4), we first compute the partial derivative $\partial q_3/\partial X_1$ treating β and δ as fixed. This gives

$$\frac{\partial q_3}{\partial X_1} = -\frac{\beta_{11}e^{\beta_1' X} + \beta_{21}e^{\beta_2' X}}{(1 + e^{\beta_1' X} + e^{\beta_2' X})^2(1 + e^{\delta' X})} - \frac{\delta_1 e^{\delta' X}}{(1 + e^{\beta_1' X} + e^{\beta_2' X})(1 + e^{\delta' X})^2} \quad (4')$$

TABLE 3

Estimates of Thermometer-Utility Model

	Parameter	Estimate		t-Statistic	
		Reported	Validated	Reported	Validated
	Constant	2.73	2.67	2.29	2.02
	Info	−0.52	−0.73	−0.97	−1.22
	Party	−0.23	−0.28	−1.06	−1.15
Reagan (β_1)	Utility RR	10.63	10.33	9.84	8.83
	Utility JA	−0.80	−0.97	−0.83	−0.87
	Utility JC	−11.80	−10.88	−8.68	−7.58
	Constant	−0.14	−0.47	−0.09	−0.30
	Info	0.33	0.04	0.49	0.05
	Party	−0.85	−0.82	−3.72	−3.07
Anderson (β_2)	Utility RR	2.37	2.42	2.27	2.09
	Utility JA	7.54	7.69	5.30	4.90
	Utility JC	−8.85	−8.38	−6.03	−5.29
	Constant	1.95	1.53	2.95	2.72
	Info	−1.73	−1.08	−5.82	−4.44
	Party	−0.41	−0.32	−3.52	−3.40
Nonvoting (δ)	Utility RR	−0.24	−0.75	−0.42	−1.64
	Utility JA	−0.08	−0.82	−0.13	−1.82
	Utility JC	0.04	−0.11	0.07	−0.26
	Alienation	−2.09	0.08	−1.65	0.08
	Indiff	−0.01	−0.24	−0.02	−0.48
	Total log likelihood	−524.31	−597.13		
	N	686	675		

where β_{11} and β_{21} are, respectively, the Reagan and Anderson coefficients on X_1. Then (4′) is evaluated at $X = \overline{X}$ and $(\beta, \delta) = (\hat{\beta}, \hat{\delta})$. For the unscaled distances model using validated responses, this number is −.033 (see Table 4). This means that if an average individual's information index were to increase by .10, the probability that individual would vote for Carter would decline slightly from .1897 to .1863, a difference of .0033. The first probability (.1897) is computed by substituting $\overline{X}$ into equation (4); it can be interpreted as the estimated probability

TABLE 4

Distance-Utility Model: (Estimated) Partial Derivatives

Reported

		Info	Absolute Strength Party ID	Utility Reagan	Utility Anderson	Utility Carter	Alienation	Indiff
Reagan	∂q_1	.280	−.012	1.477	−.272	−.713	—	—
Anderson	∂q_2	.027	−.025	−.298	.372	−.040	—	—
Carter	∂q_3	−.135	.078	−1.035	.005	.916	—	—
Nonvoting	∂q_4	−.172	−.041	−.143	−.105	−.162	.080	−.276

Validated

		Info	Absolute Strength Party ID	Utility Reagan	Utility Anderson	Utility Carter	Alienation	Indiff
Reagan	∂q_1	.250	.037	1.364	−.070	−.654	—	—
Anderson	∂q_2	.025	−.019	−.244	.343	−.037	—	—
Carter	∂q_3	−.033	.058	−.788	−.044	.838	—	—
Nonvoting	∂q_4	−.241	−.076	−.332	−.230	−.147	.474	−.309

that the average individual in the sample will vote for Carter. The other partial derivatives reported in Tables 4 and 5 are obtained in a similar fashion.

For the most part, voter information has the predicted effects on voting behavior. As the level of information increases, the probability of not voting declines. The information coefficient is significant in the nonvoting equation in both models but does not significantly impact the Reagan, Anderson, and Carter voting probabilities in the thermometer model—the derivatives are small (and positive). In the distance-utility model, information significantly affects the Reagan vote probability, and the computed derivative is positive and quite large—the probability of voting for Reagan increases with information level. The opposite is true for Carter. Replacing the distance-utility measures with thermometers eliminates any significant effect of information on the probabilities of voting for Reagan or Anderson.

The absolute strength of party ID is significant in the Anderson and the nonvoting equations in both models, and the derivatives have

TABLE 5

Thermometer-Utility Model: (Estimated) Partial Derivatives

Reported

		Info	Absolute Strength Party ID	Utility Reagan	Utility Anderson	Utility Carter	Alienation	Indiff
Reagan	∂q_1	.091	.055	1.500	−.500	−1.315	—	—
Anderson	∂q_2	.067	−.036	−.399	.510	.059	—	—
Carter	∂q_3	.086	.040	−1.064	.001	1.251	—	—
Nonvoting	∂q_4	−.245	−.058	−.034	−.011	.066	−.295	−.002

Validated

		Info	Absolute Strength Party ID	Utility Reagan	Utility Anderson	Utility Carter	Alienation	Indiff
Reagan	∂q_1	.096	.054	1.273	−.278	−.957	—	—
Anderson	∂q_2	.050	−.023	−.277	.422	.029	—	—
Carter	∂q_3	.099	.041	−.825	.043	.954	—	—
Nonvoting	∂q_4	−.245	−.072	−.170	−.187	−.025	.019	−.055

the predicted sign—the probabilities of not voting and the probability of voting for Anderson both decline as strength of party identification increases.

Finally, alienation does not have a significant effect on nonvoting in either model, while indifference has a significant effect in the distance-utility model.

Tables 6 and 7 show the classification of the choices for our model. We correctly classify 63 percent of the reported voting (and abstaining) decisions with the distance-utility model and 69 percent with the thermometer-utility model. A voter is classified according to the choice with the highest estimated probability for that voter, using the conditional logit estimates. Since there are four possible categories, a purely random model would correctly classify 25 percent of voters. To consider a more generous baseline comparison, using the maximum thermometer value of the three candidates to classify a voter yields a 45 percent success rate (note that all the nonvoters are errors). This is not surprising, since it has been known for some time that thermometers

TABLE 6

Distance-Utility Model: Classification of Choices

Reported Choices

		Reagan	Anderson	Carter	Nonvoting	Total
Predicted Choices	Reagan	201	21	30	36	288
	Anderson	1	8	7	2	18
	Carter	28	24	102	23	177
	Nonvoting	9	2	10	17	38
	Total	239	55	149	78	521

Percentage correctly classified = 63.0

Validated Choices

		Reagan	Anderson	Carter	Nonvoting	Total
Predicted Choices	Reagan	149	14	16	58	237
	Anderson	0	0	0	0	0
	Carter	11	12	68	35	126
	Nonvoting	30	18	28	75	151
	Total	190	44	112	168	514

Percentage correctly classified = 56.8

are excellent predictors of candidate vote choice, *given that an individual reports that he voted* (Weisberg and Rusk, 1970). A far more difficult task is to predict nonvoting. It is worth noting that the classification of nonvoters is far more accurate using the validated choices, despite the fact that overall classification accuracy is substantially lower in the validated sample.

As we discussed earlier, we hypothesize that if a person voted, the probability that we classify him or her correctly in the conditional logit model should be an increasing function of information level. To test this we estimated a probit model with classification accuracy as the dependent variable and the information index as the independent variable. The probability of correctly classifying with the distance model was significantly affected by the information level. This supports our hypothesis. The estimates are given below:

TABLE 7

Thermometer-Utility Model: Classification of Choices

		Reported Choices				
		Reagan	Anderson	Carter	Nonvoting	Total
	Reagan	261	18	13	63	355
Predicted	Anderson	7	31	7	12	57
Choices	Carter	17	12	159	38	226
	Nonvoting	12	5	9	22	48
	Total	297	66	188	135	686
	Percentage correctly classified = 69.0					

		Validated Choices				
		Reagan	Anderson	Carter	Nonvoting	Total
	Reagan	194	12	5	88	299
Predicted	Anderson	3	19	4	10	36
Choices	Carter	7	8	100	51	166
	Nonvoting	36	14	31	93	174
	Total	240	53	140	242	675
	Percentage correctly classified = 60.1					

$$\begin{array}{lll} \text{Correctly classify} & < 0 & = -.085 - .816\ [\text{info}] \\ \text{Incorrectly classify} & > 0 & \quad\ (-.816)\ (-3.464) \end{array}$$

with an estimated R-square of .05. In contrast, a similar probit run with the thermometer model produced an estimated R-square of .00 and an insignificant information coefficient.

Conclusions

We have five main conclusions based on our findings from the Monte Carlo experiments and the conditional logit analysis of 1980 presidential election survey data. First, as a methodological finding, the Aldrich-McKelvey scaling procedure produces a significant bias in the estimation of individual voter's spatial locations when perceptual error is heteroscedastic across voters. The nature of this bias causes the scaled distribution to be very centrally tended (unimodal), even in strongly bimodal populations. The bias is especially severe for poorly informed voters.

Second, we construct from perceptual data an information index which appears to be a meaningful measure of an individual's knowledge and sophistication regarding the political world. Third, voter information measured in this way is positively correlated with ideological "extremism" and negatively correlated with indifference between candidates. Highly informed voters consist of a significantly more polarized subset of the electorate than uninformed voters. Fourth, the information level of a voter has a strong positive effect on the likelihood the voter will vote. Fifth, conditional on voting, informed voters are significantly more predictable (at least by a "spatial model") in their voting behavior than uninformed voters.

Manuscript submitted 29 July 1985
Final manuscript received 26 June 1986

REFERENCES

Aldrich, John H., and Richard D. McKelvey. 1977. A method of scaling with applications to the 1968 and 1972 presidential elections. *American Political Science Review,* 71:111–30.

Berndt, Ernst K., Bronwyn H. Hall, Robert E. Hall, and Jerry A. Hausman. 1974. Estimation and inference in nonlinear structure models. *Annals of Economic and Social Measurement,* 3, 4:653–65.

Converse, Philip E. 1966. Information flow and the stability of partisan attitudes. In Angus Campbell, Philip Converse, Warren Miller, and Donald Stokes, eds., *Elections and the political order.* New York: Wiley.

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

McKelvey, Richard D., and William Zavoina. 1975. A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology,* 4:103–20.

Nie, Norman H., Sidney Verba, and John R. Petrocik. 1979. *The changing American voter.* Cambridge: Harvard University Press.

Sigelman, Lee. 1982. The nonvoting voter in voting research. *American Journal of Political Science,* 19:393–417.

Weisberg, Herbert, and Jerrold G. Rusk. 1970. Dimensions of candidate evaluation. *American Political Science Review,* 64:1167–85.