The Effects of Turnout on Vote Choice: A Simulation Based on Two Multinomial Models[1]

Michael D. Martinez
University of Florida
234 Anderson Hall
Gainesville, FL 32611
martinez@polisci.ufl.edu

Jeff Gill
University of Florida
234 Anderson Hall
Gainesville, FL 32611
jgill@polisci.ufl.edu

April 1, 2002

[1]Prepared for the Annual Meeting of the Southern Political Science Association, November 7-10, 2001, Atlanta.

1 Introduction

1.1 Motivation and Theory

Political scientists face quite a number of relatively simple questions for which we do not yet have satisfactory answers, or at least have some agreement on a clear answer. One of these very basic questions is whether turnout affects election outcomes. Practitioners and pundits often convey the idea that lower turnout (and the factors associated with it) should favor the more conservative party, since non-voters are disproportionately poor and less well-educated people (a group that obviously disproportionately benefits from social welfare programs championed by the liberals).

When those people can be coaxed or prodded to the polls, their support should naturally accrue to the more liberal party. So Democrats in the United States should generally hope for higher turnout races, campaigns which de-emphasize the negatives and focus on the positives (Ansolabehere and Iyengar 1995), liberal registration laws (Wolfinger and Rosenstone 1980), and the option to vote by mail (Southwell and Burchett 2000). Conversely, landslides, mudslinging, and restrictive registration laws and voting procedures should advantage Republicans.[1]

Although these theories are not particularly contested, subsequent empirical investigations in the scholarly literature have varied widely. There is support for the conventional wisdom that higher levels of participation are associated with greater success of Democrats in Presidential elections (Tucker, et al. 1986; Radcliff 1994), and that the finding travels well in cross-national analyses of nineteen western democracies (Pacek and Radcliff 1995), and fifteen post-communist countries (Bohrer et al. 2000). There are also some correlations that indicate *negative* relationships between higher turnout and Democratic vote shares in California Assembly Districts, 1992 Clinton vote shares in U.S. states, and Democratic gains in Congressional elections (Wuffle and Collet 1997). Some analyses suggest that the relationship between turnout and election outcomes are contingent on the party-class linkage (Pacek and Radcliff 1995), or the partisan composition of the district and short-term forces which cause peripheral voters to defect (DeNardo 1980; Grofman 1995; Nagel and McNulty 1996; 2000; see also Zimmer 1985). Finally, Erikson (1995) finds that there is no relationship whatsoever between turnout and Democratic presidential candidates' vote share in non-Southern states. Such a simple and obvious question (which is not even a question to many lay observers) has created a muddle of findings.

Grofman et al. (1999) argue that part of the existing confusion arises from the fact that we are not really addressing a single simple question, but have tangled three logically independent questions. The question of whether peripheral voters have greater Democratic proclivities than core voters is quite different from the question of whether elections with higher turnout should provide more favorable results to Democratic candidates. The same contingent factors that motivate peripheral Democrats to vote may also motivate them to defect from their partisanship, thereby advantaging Republican candidates. In DeNardos' (1980) words, the joke's on the Democrats. Both of those questions are logically independent from the third question: if turnout were increased in some given election, would Democrats have done better? Grofman et

[1]Knack (1994) finds that inclement weather dampens both Democratic and Republican turnout, washing out any partisan advantage.

al. seem satisfied that the literature has provided answers to the first two questions (before: yes, now: no to the first and no to the second), but the third question still awaits a sufficient means to answer it. We will give it a concerted effort in this paper.

1.2 Overview of Estimation

We examine a mechanism to estimate the effects of turnout within the context of the 2000 U.S. presidential election by employing a simulation based on multinomial logit and multinomial probit estimates of the choices made by individual citizens. Our substantive purpose in this paper is to estimate turnout effects by employing a simulation based on model estimates of the choices made by individual citizens.

Our baseline starts with a multinomial logit model as a means of gauging turnout effects under the imposition of the independence of irrelevant alternatives (IIA) assumption. We note the subsequent implications of this specification, and proceed to a more complicated Bayesian multinomial probit model. Multinomial probit is a useful model of nominal choice because it allows a flexible pattern of conditional covariance for the assumed latent utility structures. Originally suggested by Aitchison and Bennett (1970), the multinomial probit model assumes a multivariate normal structure on the errors even though the outcome variables are discrete. In particular this model selection relaxes the assumed strict adherence to the IIA (see Hausman and McFadden (1984) for a discussion) that are a condition of using the much easier to specify multinomial logit model (McFadden 1984). Unfortunately, the multinomial probit model adds additional assumptions to the variance specification in order to be identified and is in general much more difficult to implement in practice due to its computational fragility (Amemiya 1985; Hensher and Johnson 1981; McFadden 1989). A central problem is that estimation for this model requires a careful development of the numerical method for calculating multi-normal orthant integrations in order to obtain the desired probabilities.

Despite the described difficulties, multinomial probit and multinomial logit have been used to estimate the vote choice in multiparty systems in the Netherlands and Great Britain (Whitten and Palmer 1996; Quinn et al. 1999). Alvarez and Nagler (1995, 1998) evaluate the effect of a viable third party candidate (Perot) on vote choice. Lacy and Burden (1999) look at the same question, but find different results. The multinomial probit has been successfully applied to panel data problems (Keane 1994), and to other problems of spatial dependence (Hajivassiliou 1994).

Our substantive approach is similar to Lacy and Burden's, in that we posit that U.S. citizens have three unordered non-nested choices in each election vote Democratic, vote Republican, or abstain. We will first estimate vote choice (including the abstention category) as an unordered multinomial logit function of standard variables associated with both candidate preference and the likelihood of voting. From that estimation, we will derive probabilities for each respondent's selection of each of the three choices. From those probabilities, we simulate several levels of turnout. Higher turnout is simulated by adding to the pool of voters actual abstainers who had the lowest probability of abstaining. Lower turnout is simulated by subtracting from the electorate actual voters who had the highest probability of abstaining.

The use of multinomial logit serves as an excellent baseline, but its dependence on the IIA assumption bothers us here. Therefore we develop a substantially more involved model based on multinomial probit. Our computational method differs from many used thus far. We leverage work by Tanner (1996) and McCulloch (1994) to develop maximum likelihood estimates (along with an associated Hessian matrix) using a Monte Carlo EM algorithm. The idea of the EM algorithm is very simple. In the first step, temporary data that represent a reasonable guess are assigned to unknown quantities such as parameters (the "E-Step"). Then, the parameter estimation proceeds as if we now have a complete-data problem. Once this produces a solution for the parameter estimates, then we use these to update the assignment of the temporary data values with better guesses, and perform again a full model estimation process (the "M-Step"). This two-step process is repeated until the difference in the parameter updates becomes arbitrarily close to zero.

In cases where the E-Step is particularly difficult or time-consuming, it can be simulated without too much trouble by sampling realizations from the distribution for the unknown quantity. This is a substitute step which uses complete-data *conditional* maximum likelihood estimation (CM) where the conditionality is over some convenient function of the parameter estimates. This is therefore termed the *Monte Carlo EM Algorithm* (MCEM). The idea is to replace difficult numerical work with simulation.

2 The Multinomial Logit Model

We start with an analysis using the simple multinomial logit model (Schmidt and Strauss 1975). This model is founded on the *theory of individual choice* developed as: Thurstone's theory of comparative judgment (1927a, 1927b) and Luce's choice axiom (1959). Greatly simplified, the theory states that confronted with a choice set, respondents consult their personalized underlying continuum of utility and comparatively select the choice that maximizes this utility. Thurstone calls this a *discriminal* process to describe the process in which "the organism identifies, distinguishes, discriminates, or reacts to stimuli. . . ."

A multinomial logit model with J choices is estimated with respect to a reference category in order to be identified.[2] That is, the resulting coefficient sets (one for each $J-1$ choices distinct from the reference category) provide the relative effect through the logit function of that explanatory variable on the probability that the respondent chose category j rather than the reference category. The result of the estimation process is therefore J different parameter vectors $\boldsymbol{\beta}_j, \quad j = 1 \dots J$, the first of which is all zeros, $\boldsymbol{\beta}_1 = \mathbf{0}$ representing the reference category. So given the design matrix $\mathbf{X}$, the probability that respondent i chooses category j over category 1 is given by:

$$P(y_{ij}) = \frac{\exp(\mathbf{X}_i\boldsymbol{\beta}_j)}{\sum_{k=1}^{J}\exp(\mathbf{X}_i\boldsymbol{\beta}_k)} \tag{1}$$

It should be really clear from (1) that if J=2, this reduces to a standard bivariate normal specification since $\exp[0] = 1$.

[2] What do we mean here exactly? Consider a systematic component to a generalized linear model, $\mathbf{X}\boldsymbol{\beta}$, with an associated link function $g()$ (see Gill 2000, p.30-32). If it were the case that $g(\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\delta}) = g(\mathbf{X}\boldsymbol{\beta})$ for any arbitrary vector $\boldsymbol{\delta}$, which usually occurs due to cancellation, then the model is not identified because unique coefficient estimates for β are obviously now impossible to produce.

Estimation of the multinomial logit is trivial since the log likelihood function,

$$\ell(\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_J) = \sum_{j=1}^{J} \sum_{y_i: y_i = j} \text{exp}(\mathbf{X}_i \boldsymbol{\beta}_j) - \sum_{l=1}^{n} \text{log} \left(1 + \sum \sum_{j=1}^{J} \text{exp}(\mathbf{X}_i \boldsymbol{\beta}_j) \right), \tag{2}$$

is globally concave and the routine is canned in virtually all of the user-friendly software packages. Reliability can be assessed in the usual manner since the maximum likelihood estimate of the $\boldsymbol{\beta}_j$ vector is asymptotically normal with the standard calculation of Fisher information. In finite samples, we assume that the error matrix is multivariate Weibull.

More of a challenge is the substantive interpretation of the resulting coefficient estimates. The most direct way to understand the magnitude of the coefficient estimates is to evaluate the log ratio of the predictors given by:

$$\text{log} \left[\frac{p_{ij}}{p_{i0}} \right] = \mathbf{X}_i \boldsymbol{\beta}_j, \tag{3}$$

which is just the log of the ratio of probability of selecting choice j to the probability of selecting the baseline choice. The fact that it is given by multiplying the j^{th} coefficient vector by the respondents explanatory variable values makes this a particularly easy quantity to obtain. At first the restriction to the baseline comparison seems restrictive, but based on the properties of logs, any desired comparison can be obtained since:

$$\begin{aligned} \mathbf{X}_i(\boldsymbol{\beta}_j - \boldsymbol{\beta}_k) &= \mathbf{X}_i \boldsymbol{\beta}_j - \mathbf{X}_i \boldsymbol{\beta}_k \\ &= \text{log} \left[\frac{p_{ij}}{p_{i0}} \right] - \text{log} \left[\frac{p_{ij}}{p_{i0}} \right] \\ &= \text{log} \left[p_{ij} \right] - \text{log} \left[p_{i0} \right] - \text{log} \left[p_{ik} \right] + \text{log} \left[p_{i0} \right] \\ &= \text{log} \left[\frac{p_{ij}}{p_{ik}} \right], \end{aligned} \tag{4}$$

where j and k represent two arbitrarily chosen choice categories. Therefore we can estimate any relative probabilistic comparison desired.

Our interest actually extends beyond these sorts of calculations to look at the expected behavior of individuals given the estimated coefficients. Using (3) and the underlying individual accounting identity, $\sum_{j=1}^{J} \pi_{ij} = 1$, we obtain the marginal individual vote probabilities:

$$\begin{aligned} p_{i1} &= \left(1 + \sum_{j=2}^{J} \text{exp}[\mathbf{X}_i \boldsymbol{\beta}_j] \right)^{-1} \\ p_{ik} &= \text{exp}[\mathbf{X}_i \boldsymbol{\beta}_j] p_{i1}, \quad \forall j \in [2, \dots, J]. \end{aligned} \tag{5}$$

This approach allows us to provide two types of analyses not typically done with these types of models. First, we compare actual voting and abstaining outcomes to predicted outcomes[3]

[3]In the Bayesian literature one estimates the *posterior predictive distribution* which tests for the reason-

A number of authors have recently provided strong empirical arguments for the utility of the basic MNL model, even though IIA must be assumed (Abramson, et al. 1992; Canache, et al. 1994; Gerber 1996; Iversen 1994; Layman and Carmines 1997; Powers and Cox 1997; Quinn, et al. 1999; Wahlbeck 1997; Whitten and Palmer 1996). Here we will provide MNL results and contrast them with a more general analysis using multinomial probit (MNP).

3 The Data

We estimate our model using data from the 2000 American National Election Study.[4] We excluded forty-three NES respondents who voted for other candidates, as well as those who did not know or refused to report if or for whom they voted. There were 1,839 missing values (3.2%) in the final data set (1470 cases, 39 variables before transformations). We therefore impute the missing data with multiple imputation (Little and Rubin 1983, 1987; Rubin 1987) using an algorithm based on Gibbs sampling (Gill 2002, code freely available). The multiple imputation process creates more that one set of imputations based on draws from the estimated posterior distribution for the missing data, and estimation is performed on all imputed datasets (usually 7-10) with an adjustment required for measures of uncertainty. This is an explicitly Bayesian process, but can be applied in general settings.

In Table 1, we show the distribution of the outcome variable with three categories: abstention (including those who reported not voting in the election and those who voted in the election but did not vote for president), Gore voters, and Bush voters. As is usually the case, NES survey self-reports of turnout are much higher than the actual turnout in the presidential election, due to a combination of differences between voters and nonvoters in misreports, sample selection probabilities, and panel mortality. Among voters, the 95% confidence intervals of the reported distribution of votes for Gore (49% to 55%) and Bush (45% to 51%) easily include their respective actual national two-party vote levels (50.3% and 49.7%).

Since the outcome variable represents both the decision to vote and the vote choice (Bush or Gore), we chose some explanatory variables that we expect to be associated with turnout or candidate preference, or both. To capture the variety of motivations to affect turnout, we included appropriate demographic variables (age and education), mobilization variables (contact by each party), social connectedness variables (married, children living in the household, church attendance, number of political discussants named[5], and attitudes toward the campaign (knowledge

ableness of the final estimates through simulated data according to the posterior specification (Gill 2002, Chapter 6). Starting with the *prior predictive distribution* of a new data value, x_{new} before observing the full dataset: $p(x_{new}) = \int_{\boldsymbol{\beta}} p(x_{new}, \boldsymbol{\beta}) d\boldsymbol{\beta} = \int_{\boldsymbol{\beta}} p(x_{new}|\boldsymbol{\beta}) p(\boldsymbol{\beta}) d\boldsymbol{\beta}$, which is the marginal distribution of an unobserved data value is the product of the prior for $\boldsymbol{\beta}$ and the single variable PDF or PMF, integrating out this parameter (Rubin 1984). More usefully, from a diagnostic perspective, is the distribution of a new data point, x_{new} *after* the full iid dataset, $\mathbf{X}$, has been observed: the posterior predictive distribution, $p(x_{new}|\mathbf{x}) = \int_{\boldsymbol{\beta}} p(x_{new}, \boldsymbol{\beta}|\mathbf{x}) d\boldsymbol{\beta} = \int_{\boldsymbol{\beta}} \frac{p(x_{new}, \boldsymbol{\beta}|\mathbf{x})}{p(\boldsymbol{\beta}|\mathbf{x})} p(\boldsymbol{\beta}|\mathbf{x}) d\boldsymbol{\beta} = \int_{\boldsymbol{\beta}} p(x_{new}|\boldsymbol{\beta}) p(\boldsymbol{\beta}|\mathbf{x}) d\boldsymbol{\beta}$.

[4] The principal investigators for the 2000 American National Election Study were Nancy Burns, Donald R. Kinder, Steven J. Rosenstone, Virginia Sapiro, and the National Election Studies. The data were made available to us by the Inter-University Consortium for Political and Social Research (ICPSR Study 3131). Neither the principal investigators nor ICPSR bear any responsibility for our analyses and interpretations.

[5] See, for example, Robert Huckfeldt, et al. 2001.

Table 1: Voting Outcomes, 2000 ANES

Choice:	Abstain	Gore	Bush
Votes in Survey	426	550	507
Percent	29	37	34

of the issues, interest, and caring about the outcome of the election) and the political system (belief that there are important differences between the parties, internal efficacy, external efficacy, and political trust). Similarly, we selected a variety of variables that might affect candidate preference, including partisanship, retrospective evaluations of the economy and the Clinton's job performance, demographic variables (dummies for Black, Latino, Catholic, Born Again Protestant, Black Born Again Protestant), homogeneous discussion networks, issue preferences (on race, government services, moralism, and environment), and evaluations of each candidates' integrity, competence, and empathy. We also include dummy variables for interview type, NES's attempt to convert a refusal, and a question version indicator. Appendix 1 provides a full description of the variable construction.

4 Results from the MNL Model

Table 2 shows the estimation of the model. Our multinomial logit model generates estimates of the effects of each explanatory variable on the probabilities of voting for Gore and Bush compared to the baseline category of abstention. Large, positive t-statistics for both the Gore and Bush coefficients for any single explanatory variable indicate that respondents reporting that characteristic were more likely to vote (for either candidate) than not to vote. To no one's great surprise, we find that people with higher levels of political efficacy, education, age, and caring about the election outcome were more likely to vote for either candidate than to abstain.

To capture the variety of motivations to affect turnout, we included appropriate demographic Our partisan contact variables showed a somewhat more surprising result. Republican contact appeared to stimulate turnout for Bush, but it was also associated with turnout for Gore. Likewise, Democratic contact was associated with turnout for both candidates, and the Bush coefficient was larger than the corresponding Gore coefficient. These somewhat anomalous patterns underscore the effects of mobilization (Rosenstone and Hansen 1993; Gerber and Green 2000), and they probably reflect intense efforts on both sides to convert voters in a close election.

Large, positive t-statistics for Gore coefficients only are indicative of variables that tended to promote voting for Gore. Surprisingly, those included number of discussants and interest (which we expected to be associated with turnout for both candidates). Less surprising, voting for Gore was associated with being Black, Democratic partisanship, approval of Clinton's job performance, perceiving Gore as empathic, and having only Gore supporters in one's political discussion network.

Voting for Bush was associated with being Catholic or born again, Republican partisanship, and perceptions of Bush as competent. Moralist issue positions, having only Bush supporters

Table 2: MNL Model Summary

	GORE/ABSTAIN		BUSH/ABSTAIN	
	Estimate	95% CI	Estimate	95% CI
Intercept	-3.3701	[-5.5352:-1.2050]	-6.7196	[-9.0224:-4.4168]
Type of Interview	0.1051	[0.0009: 0.2094]	0.0882	[-0.0176: 0.1940]
Question Wording, Retrospective	0.1356	[-0.2453: 0.5165]	0.0409	[-0.3537: 0.4355]
Initial Refusal	-0.4519	[-0.9889: 0.0852]	-0.4619	[-1.0158: 0.0920]
Economic Evaluation	0.0272	[-0.1484: 0.2027]	0.0230	[-0.1606: 0.2066]
Clinton On Economy	-0.2656	[-0.4391:-0.0921]	-0.0190	[-0.1667: 0.1287]
Age	0.0174	[0.0039: 0.0309]	0.0169	[0.0026: 0.0312]
Black	0.9915	[0.2999: 1.6830]	-0.1799	[-1.3220: 0.9622]
Black & Born-Again	-0.3048	[-1.5168: 0.9071]	-1.0306	[-3.4981: 1.4369]
Born-Again	0.1324	[-0.4974: 0.7622]	0.8791	[0.3063: 1.4519]
Bush Cares	-0.4514	[-0.7713:-0.1315]	0.2436	[-0.1072:-0.5944]
Bush Competent	-0.5284	[-0.8649:-0.1918]	0.6704	[0.2475: 1.0933]
Bush Interview	-0.0167	[-0.3521: 0.3186]	0.2071	[-0.1732: 0.5874]
Bush Only	-1.0947	[-1.6906:-0.4988]	0.2584	[-0.2285: 0.7453]
Care	0.6459	[0.1890: 1.1027]	0.8811	[0.3916: 1.3706]
Catholic	0.3771	[-0.1059: 0.8601]	1.0992	[0.6017: 1.5967]
Church	0.3701	[-0.1723: 0.9124]	0.3665	[-0.2015: 0.9345]
Contacted by Dem.	0.5578	[0.1139: 1.0016]	0.7315	[0.2551: 1.2079]
Contacted by Rep.	0.5992	[0.1453: 1.0532]	0.5151	[0.0371: 0.9931]
Education	1.6027	[0.7276: 2.4778]	1.3774	[0.4732: 2.2816]
Environment	-0.0388	[-0.9714: 0.8937]	0.4676	[-0.5283: 1.4635]
External Effic.	0.9761	[0.0442: 1.9079]	1.2809	[0.2811: 2.2807]
Gore Cares	0.4286	[0.0956: 0.7616]	-0.3024	[-0.6621: 0.0573]
Gore Competent	0.1422	[-0.2296: 0.5140]	-0.3641	[-0.7558: 0.0276]
Gore Interview	-0.0242	[-0.3746: 0.3262]	-0.2038	[-0.5384: 0.1308]
Gore Only	0.5071	[0.0266: 0.9876]	0.0426	[-0.5713: 0.6565]
Parties Different	-0.0334	[-0.4426: 0.3757]	0.0315	[-0.4042: 0.4672]
Internal Efficacy	-0.0624	[-0.9347: 0.8098]	0.5299	[-0.3971: 1.4569]
Interest	0.9271	[0.2779: 1.5763]	0.1732	[-0.5245: 0.8709]
Children	-0.1528	[-0.5987: 0.2932]	-0.0983	[-0.5641: 0.3675]
K-Issues	0.5321	[-0.1303: 1.1945]	0.6031	[-0.0989: 1.3051]
Latino	-0.1052	[-0.9307: 0.7202]	-0.7261	[-1.6427: 0.1905]
Married	0.3659	[-0.0397: 0.7714]	0.5155	[0.1034: 0.9276]
Moral	-1.0212	[-1.7294:-0.3131]	-0.3470	[-1.1148: 0.4208]
Number of Discussants	0.2881	[0.1384: 0.4379]	0.1140	[-0.0400: 0.2680]
Race	-0.1299	[-0.8671: 0.6072]	0.3478	[-0.4356: 1.1312]
Service	0.5205	[-0.6122: 1.6533]	0.7187	[-0.4011: 1.8385]
Trust	0.5686	[-0.3085: 1.4457]	0.0692	[-0.8537: 0.9921]
Strong Democrat	0.6336	[-0.0849: 1.3521]	-1.4432	[-2.5388:-0.3476]
Moderate Democrat	0.7775	[0.0934: 1.4616]	-0.6115	[-1.4192: 0.1962]
Weak Democrat	0.3580	[-0.3290: 1.0450]	-0.0108	[-0.7498: 0.7282]
Weak Republican	-0.4146	[-1.2705: 0.4413]	0.7438	[0.0497: 1.4379]
Moderate Republican	-0.4484	[-1.2880: 0.3912]	0.8599	[0.1861: 1.5337]
Strong Republican	-1.5407	[-2.8865:-0.1949]	0.9979	[0.1727: 1.8231]

in one's network, and perceiving Bush as empathic were associated with demobilizing support for Gore (large, negative coefficient mobilization for Bush.[6] Finally, some variables did not have pronounced independent effects on voting for either candidate. Those included knowledge of the candidates' relative positions on issues, political trust, internal efficacy, issue preferences related to the level of government services, the environment and race, and the two candidates' integrity. We do not claim that these issues were unimportant in the election, only that collinearity with other factors in the model may undermine our efforts to uncover their effects reliably.

5 Simulation of Turnout Based on the MNL Model

Using the estimated effects from the multinomial logit model, we calculated the probabilities for each respondent abstaining, voting for Gore, and voting for Bush. Of course, since we excluded third-party voters and non-respondents to the vote choice question, the sum of those three probabilities equals one for every respondent. The sums of each of these probabilities *across* respondents are extremely close to their actual weighted distributions in the NES sample, as seen in Table 3.

Table 3: ACTUAL VS. ESTIMATED VOTING, 2000 ANES

Choice:	Actual Percent	Estimated Percent
Abstain	27.7	27.4
Gore	37.6	38.0
Bush	34.7	34.6
Weighted N	1457	1457

For each respondent, we also calculated a probability of voting for each candidate, excluding the probability of abstention by:

$$p(Gore|Vote) = \frac{p(Gore)}{1 - p(Abstain)}$$

$$p(Bush|Vote) = \frac{p(Bush)}{1 - p(Abstain)}$$

To simulate the effects of varying levels of turnout on the aggregate vote choice, we sum $p(Gore|Vote)$ and $p(Bush|Vote)$ across different sets of respondents. For the 72% of NES respondents who reported voting, the sums of these probabilities (52.4% for Gore, 47.6% for Bush) are again very close to the actual reported votes (52.0% for Gore, 48.0% for Bush).

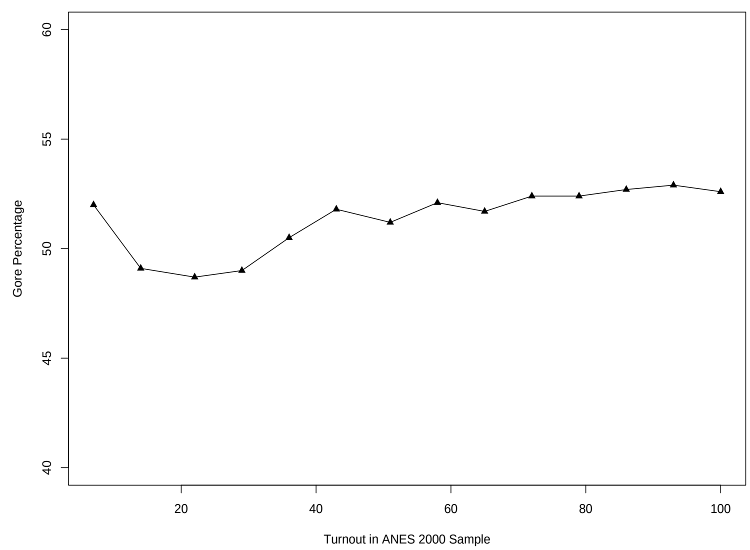
[6]Being married produced somewhat ambiguous results. Married people were more likely to have voted for Bush than to have abstained, but we cannot rule out the possibility that the turnout effects of being married are similar for both candidates.

Higher levels of turnout are simulated by the set of respondents who actually voted plus the non-voters who had the lowest probability of abstaining. Thus, an electorate with a 79% turnout rate is represented by those who actually voted (72%) plus the non-voters who had no more than a 0.3124 probability of abstention. An electorate with a 86% turnout rate includes those who actually voted plus the non-voters who had no more than a 0.5851 probability of abstention. And so on.

Lower levels of turnout are simulated by sequentially removing voters with the highest probability of abstaining. Thus, an electorate with a 65% turnout rate is simulated by the set of voters who had no more than a 0.5104 probability of abstention. Lower levels of turnout are simulated by lowering the thresholds of abstention probabilities.

For each of fourteen subsets of respondents (ranging from 7% turnout to universal turnout), we calculate the vote distribution between Gore and Bush by summing the quantities $p(Gore|Vote)$ and $p(Bush|Vote)$, and calculating percentages based on those sums.

Figure 1: Simulated Turnout Effect



Does the level of turnout affect the distribution of the votes between Gore and Bush? Figure 1 arrays the simulated vote percentages for Gore against the fourteen simulated levels of turnout, and shows that the level of support for Gore is remarkably stable. (We were stunned.) Only at extreme values of turnout is there any perceptible effect on vote preference. Gore's highest level of support is at 93% turnout, and his lowest level of support is at 22% turnout, and the difference between those conditions (3.2%) is a lot in a close election. But within less extreme variations from current conditions, Gore's level of support hardly budges at all. Aggregate voter choice barely moves over turnout levels from 58% of respondents to 86% of respondents. Not only is the line relatively flat, it is non-monotonic (Gore does slightly better at 43% turnout than he does at 51% turnout, and slightly better at 58% turnout than at 65% turnout). Thus, any general admonitions that Democrats should always hope for (or Republicans should fear) higher levels of turnout appear to be unwarranted.

These results suggest to us that marginal changes in the costs of voting would not have dramatic consequences for the partisan outcome of most elections. This presidential election was unusually close, and would have been decided differently by the addition of a few well placed votes.

A few caveats are in order. Obviously, partisan mobilization can matter a lot more than our results suggest. Increasing turnout *among Democrats* or *among Republicans* can dramatically alter election outcomes much more than simply making it a little easier or a little harder to vote. Moreover, our motivation was to determine whether turnout might have mattered in the 2000 presidential election in the United States, and conditions in other elections in other settings (including a highly energized American election) could be very different.

These results also depend on the plausibility and fit of our multinomial logit model, including the assumption that voters and non-voters weigh their considerations in voter preferences similarly (see Gant and Lyons 1993), and that preference orders between choices are not affected by the inclusion (or elimination) of other choices. This may or may not be a realistic foundation. To test the robustness of these findings we continue by specifying a multinomial probit model and contrasting the findings. This has shown to be a revealing process in other cases (Quinn, et al. 1999).

6 The Multinomial Probit Model

A perhaps more realistic model for multichotomous vote choices is the multinomial probit model, which substitutes for the assumption of iid Weibull distributed error terms in the MNL model with the assumption of multivariate normal error terms. The result of this change of assumption is a model that is much more difficult to estimate but is free of the required IIA assumption of the multinomial logit model (Hensher and Johnson 1981, Chapter 5; Davidson and MacKinnon 1993, p.533). The multinomial probit model has received extensive treatment in the econometric literature: Amemiya (1985), Bunch (1991), Daganzo (1979), Dansie (1985), Hausman and Wise (1978), Keane (1992), Manski and McFadden (1981), Recently a number of authors in political science have used the multinomial probit in useful ways: Alvarez and Nagler (1995, 1998), Burden and Lacy (1999), Quinn, et al. (1999).

The theoretical superiority of the MNP model over the MNL model is not just a function of freedom from the IIA assumption. Adams (1997) showed that normally distributed errors emerge from very general assumptions in his simulation study. This is basically an expression of the persistence of the central limit theorem, but it highlights the fact that normally distributed errors are not only more tied to mathematical-statistics theory, they also emerge empirically.

6.1 Model Specification

Suppose there exist N respondents in the dataset with c choices observed for each respondent: $\boldsymbol{\omega}_i = [\omega_{i1}, \omega_{i2}, \ldots, \omega_{ic}]$, where all but one of these vector values is zero with the remaining value equal to one indicating individual selection. It is standard and convenient to assume that $\boldsymbol{\omega}_i$ is the observable manifestation of an underlying continuous measure of utility, $\mathbf{U}_i = [U_{i1}, U_{i2}, \ldots, U_{ic}]$, in which j^{th} value of $\boldsymbol{\omega}_i$ is equal to one because the associated latent measure has the greatest utility to person i of all alternatives: $U_{ij} > U_{ik}, \ \ \forall k \neq j$. We further assume that these utilities are generated by the distribution:

$$\mathbf{U}_i \sim \mathcal{N}(\mathbf{Z}'\boldsymbol{\gamma}, \Omega_{\mathbf{z}}), \tag{6}$$

where: $\mathbf{Z}$ is a $c \times k$ data matrix, $\boldsymbol{\gamma}$ is a $k \times 1$ coefficient vector and $\Omega_{\mathbf{z}}$ is a $c \times c$ covariance matrix.

That is, the underlying motivation for the model is multivariate Gaussian-normal. This assumption is really for convenience than for any strong theoretical reason.

As expressed (6) is not identified, and it is again necessary to set a reference category and express the $J-1$ choices comparatively (Bunch 1991, Dansie 1985). Thus we reexpress from absolute utilities for person i, U_{ij}, to relative utilities, $y_{ij} = U_{ij} - U_{i1}$, where this relative to the arbitrary baseline category as in the MNL model. This produces the assumed model:

$$\mathbf{y}_i \sim \mathcal{N}(\mathbf{X}'\boldsymbol{\beta}, \Omega_{\mathbf{x}}), \qquad \text{where: } \mathbf{X} \text{ is } (c-1) \times k, \ \boldsymbol{\beta} \text{ is } k \times 1, \ \Omega_{\mathbf{z}} \text{ is } (c-1) \times (c-1). \tag{7}$$

The result of this specification is that the error terms in the model are now multivariate normal distributed, rather than Weibull distributed as in the MNL model.

Now introduce a new variable $W_{ij} = I(y_{ij} > 0, y_{ij} = \max(y_{i\cdot}))$, and: $W_{i1} = 1$, $W_{i2:J} = 0$ if all values of y_{ij} are negative (McCulloch 1994). This indicator function makes the estimation of the coefficients much easier. The MNP likelihood is now the simple form:

$$\ell(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_J) = \prod_{j=1}^{J} \prod_{i=1}^{N} \pi_{ij}^{W_{ij}}, \tag{8}$$

where π_{ij} is the probability that the i^{th} individual selects choice j with the obvious constraints that $\pi_{ij} > 0, \ \forall j$, and $\sum_{j=1}^{J} \pi_{ij} = 1$.

As setup, this model is not identified because the scale of the relative utilities, y_{ij} is indeterminate. Various authors have dealt with this in various ways, some of which are quite restrictive. Alvarez and Nagler (1995, 1998) and Lacy and Burden (1999), for instance, set all posterior variances to unity (Burden and Lacy also set one covariance equal to zero). The result of this change to unity along the diagonal is to make the covariance matrix a correlation matrix, which works well when the off-diagonal elements are of prime interest. Others (Quinn, et al. 1999, for instance) are less restrictive and merely confine the first diagonal term in the covariance matrix to be unity. Following Geweke et al.'s (1994) advice, we follow the latter convention and set the first diagonal element of $\Omega_{\mathbf{x}}$ to a constant.

6.2 Motivation for Bayesian Estimation with MCEM

A second serious issue, and one that leads to great agony in practice, is the complexity of the required joint-normal integral:

$$\begin{aligned}\pi_c &= p(U_c > U_{c-1} > \ldots > U_1)\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{U_c}\cdots\int_{-\infty}^{U_c} f(U_1,\ldots,U_{c-1},U_c)dU_1\ldots dU_{c-1}dU_c\end{aligned} \tag{9}$$

Amemiya (1985, p.308). For $c > 3$ this can be prohibitive. From the normal consequence of (7) we know that:

$$\pi_{ic} \propto -\int_{\mathbf{y}_i}\frac{1}{2}(\mathbf{y}-\mathbf{X}_i\boldsymbol{\beta})'\Omega_x^{-1}(\mathbf{y}-\mathbf{X}_i\boldsymbol{\beta})I(y_{ij}>0,\ y_{ij}=\max_j(y_{ik})d\mathbf{y}_i. \tag{10}$$

Except for cases where Ω_x is greatly restricted, this integral has no analytical solution. Our solution is to use a variant of the EM algorithm in which the E-Step computations are performed with Gibbs sampling. This is called Monte Carlo EM (MCEM) and reduces the computation burden by substituting an analytical expectation calculation with a simulation based estimate.

6.3 Explaining MCEM

The EM algorithm was formalized by Dempster, Laird, and Rubin in their seminal 1977 article Expectation-Maximization (EM) is a very flexible and popular technique for so-called incomplete data problems. Incomplete data in this context is either missing values in the data set or unknown parameters as in our case. In the first step, the "E-Step", temporary data that represent a reasonable guess are assigned to replace the missing values. Then, the parameter estimation proceeds normally, the "M-Step", as if we now have a complete-data problem (complete in the sense that observed and missing data are now both "available" in the analysis). Once this produces a new solution for the parameter estimates, then we use these to update the assignment of the temporary data values with better guesses, and perform again a full model estimation process. This two-step process is repeated as often as required until the difference in the parameter updates becomes arbitrarily close to zero and we therefore have convergence. See Gill (2002, Chapter 8) for a lengthy discussion of the theory and application of the EM algorithm.

Starting with data $\mathbf{X} = [\mathbf{X}_{mis}, \mathbf{X}_{obs}]$, a likelihood function $f(\mathbf{X}|\boldsymbol{\beta})$, and arbitrary starting values for the vector $\boldsymbol{\beta}^{(k)}$,

- **[E-Step:]** compute $Q(\boldsymbol{\beta}^(k+1)|\boldsymbol{\beta}^{(k)}) = \int \ell(\boldsymbol{\beta}|\mathbf{X}_{obs},\mathbf{X}_{mis})f(\mathbf{X}_{mis}|\mathbf{X}_{obs},\boldsymbol{\beta}^{(k)})d\mathbf{X}_{mis}$,
- **[M-Step:]** choose the value for $\boldsymbol{\beta}$ that maximizes $Q(\boldsymbol{\beta}^{(k+1)}|\boldsymbol{\beta}^{(k)})$,
- repeat these steps until the difference between $\boldsymbol{\beta}^{(k+1)}$ and $\boldsymbol{\beta}^{(k)}$ is arbitrarily small.

It is important to understand that this iterative process gives a consecutive updating of the parameter estimate that will under most circumstances to converge to the maximum conditional likelihood value (Beale 1977, p.23) and is guaranteed to converge to at least a stationary point (Wu 1983). By cycling between the E-Step and the M-Step we progressively move closer to the posterior mode.

In cases where the E-Step is particularly difficult or time-consuming, it can be simulated by sampling M realizations from the distribution $X^m \sim f(\mathbf{X}_{mis}|\mathbf{X}_{obs}, \boldsymbol{\beta}^{(k+1)})$ and calculating:

$$Q(\boldsymbol{\beta}^{(k+1)}|\boldsymbol{\beta}^{(k)}) = \frac{1}{M}\sum_{m=1}^{M} \ell(\boldsymbol{\beta}|\mathbf{X}_{obs}, \mathbf{X}^m), \tag{11}$$

a substitute step which uses complete-data *conditional* maximum likelihood estimation (CM) where the conditionality is over some convenient function of the parameter estimates. This is the basic *Monte Carlo EM Algorithm* (MCEM) (Booth and Hobert 1999; Chan and Ledolter 1995; Celeux and Diebold 1985; Guo and Thompson 1992; Wei and Tanner 1990). Since the Monte Carlo step is performed at every iteration of the EM algorithm, it is important to determine a value of M that represents a good compromise between efficiency and accuracy. We perform this step using Gibbs sampling to produce empirical values of $\ell(\boldsymbol{\beta}|\mathbf{X}_{obs}, \mathbf{X}^m)$.

6.4 Applying MCEM to the MNP Model

Our approach is based roughly on that of McCulloch (1994) and McCulloch and Rossi (1994), where the central idea is to modify the EM algorithm to produce maximum likelihood estimates with a Gibbs step to ease calculations. Treat the $\mathbf{y}_i$ and $\mathbf{X}_i$ in (10) as the complete data, and β, Ω_x as the information to be filled in. This gives as the E-Step where we obtain the conditional expectation of the unknown $\boldsymbol{\beta}$ parameters:

$$\begin{aligned} Q(\boldsymbol{\beta}^{(k+1)}|\boldsymbol{\beta}^{(k)}, \Omega_x^{(k)}) &= E[(\mathbf{y} - \mathbf{X}_i\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}_i\boldsymbol{\beta})|\boldsymbol{\beta}^{(k)}, \Omega_x^{(k)}] \\ &= VAR[\mathbf{y}_i|\mathbf{W}_i, \boldsymbol{\beta}^{(k)}, \Omega_x^{(k)}) \\ &\quad + (E(\mathbf{y}_i|\mathbf{W}_i, \boldsymbol{\beta}^{(k)}, \Omega_x^{(k)}) - \mathbf{X}_i\boldsymbol{\beta})'(E(\mathbf{y}_i|\mathbf{W}_i, \boldsymbol{\beta}^{(k)}, \Omega_x^{(k)}) - \mathbf{X}_i\boldsymbol{\beta}). \end{aligned} \tag{12}$$

Then in the M-Step we maximize the interim (k^{th} step) log likelihood:

$$\ell(\boldsymbol{\beta}, \Omega_x) \propto -\frac{n}{2}\text{log}||\Omega_x|| - \frac{1}{2}\text{tr}(\Omega_x^{-1})Q(\boldsymbol{\beta}^{(k+1)}|\boldsymbol{\beta}^{(k)}, \Omega_x^{(k)}) \tag{13}$$

with respect to $\boldsymbol{\beta}$ and Ω_x. Actually this must be done in two steps: maximize with respect to $\boldsymbol{\beta}$ holding Ω_x constant, and then maximizing with respect to Ω_x holding $\boldsymbol{\beta}$ constant. This process is not as bad as it may seem. The first maximization is equivalent to calculating the general least squares estimator:

$$\boldsymbol{\beta}^{(k+1)} = \left(\sum_{i=1}^{n} \mathbf{X}_i'(\Omega_x^{(k)})^{-1} E(\mathbf{y}_i|\mathbf{W}_i, \boldsymbol{\beta}^{(k)}, \Omega_x^{(k)})\right) \left(\sum_{i=1}^{n} \mathbf{X}_i'(\Omega_x^{(k)})^{-1}\mathbf{X}_i\right)^{-1}, \tag{14}$$

which is Meng and Rubin's (1993) equation 2.2 with the expectation of $\mathbf{y}_i$ replacing $\mathbf{y}_i$ itself. The second maximization simply substitutes the new $\boldsymbol{\beta}^{(k+1)}$ into (13) and maximizes with respect to Ω_x. This conditional M-Step is called Monte Carlo Expectation Conditional Maximization (MCECM) by Meng and Rubin because of the additional conditioning process.

6.5 Conference Update

We now have the MCEM/MNP procedure coded and working for small contrived datasets where we know the answer that the estimation procedure is supposed to produce. However, with the full ANES 2000 subset used in the MNL procedure (1470 cases, 44 explanatory variables) the MCMC element of the process is very slow and we cannot now confirm convergence of the chain. It is our expectation that through a number of well-known tricks such as variance stabilization reparameterization and embedded Metropolis-Hastings steps (Chen, et al. 2000, Robert and Casella 1999) that we can significantly speed up this process.

7 References

Abramson, Paul R., John H. Aldrich, Phil Paolino, David W. Rohde. 1992. "Sophisticated" Voting in the 1988 Presidential Primaries." *The American Political Science Review* 86, 55-69.

Adams, James. 1997. "Condorcet Efficiency and the Behavioral Model of the Vote." *Journal of Politics* 59, 1252-63.

Aitchison, J., and J. Bennett. 1970. "Polychotomous Quantal Response by Maximum Indicant." *Biometrika* 57, 253-62.

Alvarez, R. Michael, and Jonathan Nagler. 1995. "Economics, Issues, and the Perot Candidacy: Voter Choice in the 1992 Elections." *American Journal of Political Science* 39, 714-44.

Alvarez, R. Michael, and Jonathan Nagler. 1998. "When Politics and Markets Collide: Estimating Models of Multiparty Elecitions." *American Journal of Political Science* 42-55.

Amemiya, Takeshi. 1985. *Advanced Econometrics.* Cambridge, MA: Harvard University Press.

Ansolabehere, Stephen, and Shanto Iyengar. 1995. *Going Negative: How Political Advertising Shrinks and Polarizes the Electorate.* New York: Free Press.

Beale, E. M. L. 1977. "Discussion of the Paper by Professor Dempster, Professor Laird and Dr. Rubin." *Journal of the Royal Statistical Society, Series B* 139, 22.

Bohrer, Robert E., II, Alexander C. Pacek, and Benjamin Radcliff. 2000. "Electoral Participation, Ideology, and Party Politics in Post-Communist Europe." *Journal of Politics* 62, 1161-72.

Bolduc, D. 1992. "Generalized Autoregressive Errors in the Multinomial Probit Model." *Transportation Research* 26B, 155-70.

Booth, J. G., and J. P. Hobert. 1999. "Maximizing Generalized Linear Mixed Model Likelihoods with an Automated Monte Carlo EM Algorithm." *Journal of the Royal Statistical Society, Series B* 61, 265-85.

Bunch, D. S. 1991. "Estimability in the Multinomial Probit Model." *Transportation Research* 25B, 1-12.

Canache, Damarys, Jeffery J. Mondak, Annabelle Conroy. 1994. "Politics in Multiparty Context: Multiplicative Specifications, Social Influence, and Electoral Choice." *Public Opinion Quarterly* 58, 509-38.

Chan, K. S., and Johannes Ledolter. 1995. "Monte Carlo EM Estimation for Time Series Models Involving Counts." *Journal of the American Statistical Association* 90, 242-52.

Chen, Ming-Hui, Qi-Man Shao, and Joseph G. Ibrahim. 2000. *Monte Carlo Methods in Bayesian Computation.* New York: Springer-Verlag.

Celeux, G., and J Diebolt. 1985. "The SEM Algorithm: A Probabilistic Teacher Algorithm Derived from the EM Algorithm. for the Mixture Problem." *Computational Statistics* 2, 73-82.

Daganzo, C. 1979. *Multinomial Probit.* New York: Academic Press.

Dansie, B. 1985. "Parameter Estimability in the Multinomial Probit Model." *Transportation Research* 19B, 526-8.

Davidson, Russell, and James G. MacKinnon. 1993. *Estimation and Inference in Econometrics.* Oxford: Oxford University Press.

Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. "Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society, Series B* 39, 1-38.

DeNardo, James. 1980. "Turnout and the Vote: The Joke's on the Democrats." *American Political Science Review* 74, 406-420.

Erikson, Robert S. 1995. "State Turnout and Presidential Voting - a Closer Look." *American Politics Quarterly* 23, 387-96.

Gant, Michael M., and William Lyons. 1993. "Democratic Theory, Nonvoting, and Public Policy: The 1972-1988 Presidential Elections." *American Politics Quarterly* 21, 185-204.

Geoffrey C. Layman, Edward G. Carmines. 1997. "Cultural Conflict in American Politics: Religious Traditionalism, Postmaterialism, and U.S. Political Behavior." *Journal of Politics* 59, 751-77.

Gerber, Elisabeth R. 1996. "Legislative Response to the Threat of Popular Initiatives." *American Journal of Political Science* 40, 99-128.

Geweke, John, Michael Keane, and David Runkle. 1994. "Alternative Computational Approaches to Inference i9n the Multinomial Probit Model." *Review of Economics and Statistics* 76, 609-32.

Gill, Jeff. 2000. *Generalized Linear Models: A Unified Approach.* Thousand Oaks, CA: Sage.

Gill, Jeff. 2002. *Bayesian Methods for the Social and Behavioral Sciences.* New York: Chapman & Hall. Available temporarily at: `http://web.clas.ufl.edu/~jgill/Bayes.Dir`.

Grofman, B. 1985. "The Neglected Role of the Status Quo in Models of Issue Voting." *Journal of Politics* 47, 230-37.

Grofman, B., G. Owen, and C. Collet. 1999. "Rethinking the Partisan Effects of Higher Turnout: So What's the question?" *Public Choice* 99, 357-376.

Guo, S. W., and E. A. Thompson. 1991. "Monte Carlo Estimation of Variance Component Models for Large Complex Pedigrees." *IMA Journal of Mathematics Applied in Medicine and Biology* 8, 171-89.

Hajivassiliou, V. A. 1994. "A Simulation Estimation Analysis of the External Debt Crises of Developing Countries." *Journal of Applied Econometrics* 9, 109-131.

Hausman, J. A., and D. McFadden. 1984. "Specification Tests for the Multinomial Logit Model." *Econometrica* 52, 1219-40.

Hausman, J. A., and D. A. Wise. 1978. "A Contional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Inderdependence and Heterogeneous Preferences." *Econometrica* 46, 403-26.

Hensher, David A., and Lester W. Johnson. 1981. *Applied Discrete-Choice Modelling.* John Wiley & Sons.

Huckfeldt, Robert, Jeanette Morehouse, and Tracy Osborn. 2001. "Disagreement, Ambivalence, and Engagement: The Political Consequences of Heterogeneous Networks." Paper prepared for delivery at the 2001 Annual Meeting of the American Political Science Association, San Francisco, September, 2001.

Iversen, Torben. 1994. "Political Leadership and Representation in Western European Democracies: A Test of Three Models of Voting." *American Journal of Political Science* 38, 45-74.

Keane, Michael P. 1992. "A Note on Identification in the Multinomial Probit Model." *Journal of Business and Economic Statistics* 10, 193-200.

Knack, Stephen. 1994. "Does Rain Help the Republicans? Theory and Evidence on Turnout and the Vote." *Public Choice* 79, 187-209.

Lacy, Dean, and Barry C. Burden. 1999. "The Vote-Stealing and Turnout Effects of Ross Perot in the 1992 U.S. Presidential Elections." *American Journal of Political Science* 42, 233-55.

Little, Roderick J. A., and Donald B. Rubin. 1983. "On Jointly Estimating Parameters and Missing Data by Maximizing the Complete-Data Likelihood." *The American Statistician* 37, 218-20.

Little, Roderick J. A., and Donald B. Rubin. 1987. *Statistical Analysis with Missing Data.* New York: John Wiley & Sons.

Luce, R. D. 1959. *Individual Choice Behavior.* John Wiley & Sons.

Manski, C. F., and D. McFadden. 1981. *Structural Analysis of Discrete Data with Econometric Applications.* Cambridge, MA: MIT Press.

McCulloch, Charles E. 1994. "Maximum Likelihood Variance Components Estimation for Binary Data." *Journal of the American Statistical Association* 89, 330-5.

McCulloch, Robert, and Peter E. Rossi. 1994. "An Exact Likelihood Analysis of the Multinomial Probit Model." *Journal of Econometrics* 1994, 207-40.

McFadden, D. 1984. "Econometric Analysis of Qualitiative choice Models." In *Handbook of Econometrics*, Vol.II, Z. Griliches and M. D. Intriligator (eds.). Amsterdam: North-Holland.

McFadden, D. 1989. "A Method of Simulated Moments for Estimation of Discrete repsonse Models Without Numerical Integration." *Econometrica* 57, 995-1026.

Meng, Xiao-Li., and Donald B. Rubin. 1993. "Maximum Likelihood Estimation via the ECM Algorithm: A General Framework." *Biometrika* 80, 267-78.

Nagel, Jack H., John E. McNulty. 1996. "Partisan Effects of Voter Turnout in Senatorial and Gubernatorial Elections." *American Political Science Review* 90, 780-93.

Nagel, Jack H., John E. McNulty. 2000. "Partisan Effects of Voter Turnout in Presidential Elections." *American Politics Quarterly* 28, 408-429.

Pacek, Alexander, Benjamin Radcliff. 1995. "The Political Economy of Competitive Elections in the Developing World." *American Journal of Political Science* 39, 745-759.

Powers, Denise V., and James H. Cox. 1997. "Echoes from the Past: The Relationship between Satisfaction with Economic Reforms and Voting Behavior in Poland." *American Political Science Review* 91, 617-33.

Quinn, Kevin M., Andrew D. Martin, and Andrew B. Whitford. 1999. "Vote Choice in Multi-Parti Democracies: A Test of Competing Theories and Models." *American Journal of Political Science* 43 1231-47.

Radcliff, Benjamin. 1994. "Turnout and the Democratic Vote." *American Politics Quarterly* 22, 259-276.

Robert, Christian P., and George Casella. 1999. *Monte Carlo Statistical Methods.* New York: Springer-Verlag.

Rubin, Donald. 1984. "Bayesianly Justifiable and Relevant Frequency Calculations for the Applied Statistician." *Annals of Statistics* 12, 1151-72.

Rubin, Donald. 1987. *Multiple Imputation for Nonresponse in Surveys.* New York: John Wiley & Sons.

Schmidt, P. and R. P. Strauss. 1975. "Estimation of Models with Jointly Dependent Qualitative Variables: A Simultaneous Logit Approach." *Econometrica* 44, 745-55.

Southwell, Priscilla L., and Justin I. Burchett. 2000. "The Effect of All-Mail Elections on Voter Turnout." *American Politics Quarterly* 28, 72-79.

Tanner, Martin A. 1996. *Tools for Statistic Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions.* New York: Springer.

Thurstone, L. L. 1927a. "A Law of Comparative Judgement." *Psychological Review* 34, 278-86.

Thurstone, L. L. 1927b. "Psychophysical Analysis." *American Journal of Psychology* 38, 368-89.

Tucker, Harvey J., Arnold Vedlitz, and James DeNardo. 1986. "Does Heavy Turnout Help Democrats in Presidential Elections?" *American Political Science Review* 80, 1291-1304.

Wahlbeck, Paul J. 1997. "The Life of the Law: Judicial Politics and Legal Change." *Journal of Politics* 59, 778-802.

Wei, G. C. G., and M. A. Tanner. 1990. "A Monte Carlo Implementation of the EM Algorithm and the Poor Man's Data Augmentation Algorithm." *Journal of the American Statistical Association* 85, 699-704.

Whitten, Guy D., and Harvey D. Palmer. 1996. "Heightening Comparativists' Concern for Model Choice: Voting Behavior in Great Britain and the Netherlands." *American Journal of Political Science* 40, 231-60.

Wolfinger, Raymond E., and Steven Rosenstone. 1980. *Who Votes.* New Haven: Yale University Press.

Wu, C. F. Jeff. 1983. "On the Convergence Properties of the EM Algorithm." *Annals of Statistics* 11, 95-103.

Wuffle, A., and C. Collet. 1997. "Why Democrats Shouldn't Vote." *Journal of Theoretical Politics* 9, 137-140.

Zimmer, Troy A. 1985. "Political Competition, Voter Turnout, and Party Vote." *Social Science Journal* 22, 103-110.

8 Appendix 1: Table of Explanatory Variables

Table 4: Variable Summary

		Range	Turnout	Vote Choice
	Retrospective Economic Evaluation			
v000491	Now thinking about the economy in the country as a whole, would you say that OVER THE PAST YEAR the nation's economy has gotten worse, stayed about the same, or gotten better? [Experimental version: gotten worse, stayed about the same, or gotten better?] Both versions: Much better/worse or somewhat better/worse?	1-5	-.054	.275
V000005e	Method: v000491 standard or experimental	1-2		
V000503	Do you approve or disapprove of the way Bill Clinton is handling the economy? Do you approve/disapprove strongly or not strongly?	1-5	-.058	.534

Table 5: VARIABLE SUMMARY (CONT.)

		Range	Turnout	Vote Choice
	Partisanship			
V000523 dummies	Generally speaking, do you think of yourself as a Republican, a Democrat, an Independent, or what? Would you call yourself a strong Democrat/Republican or a not very strong Democrat/ Republican? Do you think of yourself as closer to the Republican Party or to the Democratic party?	0-1	.036	.752
	Issues			
moral	Index built from Abortion (v000694), Gays in Military (v000727), Gays Adopt (v000748). Abortion was double weighted. (Alpha = .63)	0-1	-.027	.349
service	Index built from v000614 (Health Care), v000681 (spending on social security), v000693 (surplus to protect social security), v000550 (spending services scale) (alpha = .63)	0-1	.156	.400
race	Index built from v000645 (Aid Blacks), v000674 (affirmative action), v000745 (English only) (alpha = .45)	0-1	.047	.317
env	Index built from v000713 (jobs/environment scale), v000776 (environmental regulation scale), v000682 (spending on environmental protection) (alpha = .66)	0-1	.057	.326
	Mobilization and contact			
contactd	Contacted or received mail from Democrats or both parties (recoded from v001220 and v001223)	0-1	.301	-.107
contactr	Contacted or received mail from Republicans or both parties (recoded from v001220 and v001223)		.337	.138
ndiss	Number of discussants (from v001699 to v001702)	0-4	.297	.011
goreonly	At least one discussant voted for Gore and none voted for Bush (from v001710 v001718 v001726 and v001734)	0-1	.169	-.424
bushonly	At least one discussant voted for Bush and none voted for Gore (from v001710 v001718 v001726 and v001734)	0-1	.088	.485
church	Frequency of church attendance (recoded from v000877 and v000879)	0-1	.215	.140

Table 6: VARIABLE SUMMARY (CONT.)

		Range	Turnout	Vote Choice
	Retrospective Economic Evaluation			
v000491	Now thinking about the economy in the	1-5	-.054	.275
	Demographics			
educ	Education recoded from v000913	0-1	.320	.059
v000908	Age	18-97	.200	-.006
Black	from v001006a or v001006b or v001006c	0-1	-.004	-.284
Latino	from v001006a or v001006b or v001006c	0-1	-.093	-.053
Catholic	from v000882 or v000883	0-1	.083	.029
Bornagin	from (v000882 or v000883 or v000899) and v000903	0-1	.072	.157
blckborn	Black * Bornagin	0-1	-.009	-.159
married	v000909	0-1	.194	.152
kids	from v001024	0-1	-.087	.065
inteff	from v001516 through v001519 (alpha =.80)	0-1	.287	.044
exteff	from v001527, v001528, v001538, v001539 (alpha = .72)	0-1	.262	-.010
trust	from v001534 to v001537 (alpha = .63)	0-1	.077	-.086
imptdiff	from v001435	0-1	.269	.039
kissues	correct relative Gore and Bush issue placements from abortion, spending and services, aid to blacks, and environmental protection scales (alpha = .67)	0-1	.332	.035
interest	Interest in campaign (from v000301)	0-1	.368	-.019
care	Care about outcome (from v000302)	0-1	.374	-.010
goreint	from v000524 (Gore moral) and v000528 (Gore dishonest)	1-4	.025	-.493
gorecare	from v000525 (Gore cares about people like you) and v000530 (Gore out of touch)	1-4	.040	-.540
gorecomp	from v000526 (Gore knowledgeable), v000527 (Gore strong leader), and v000529 (Gore intelligent)	1-4	-.018	-.500
bushint	from v000531 (Bush moral) and v000535 (Bush dishonest)	1-4	.122	.444
bushcare	from v000532 (Bush cares about people like you) and v000537 (Bush out of touch)	1-4	.033	.558
bushcomp	from v000533 (Bush knowledgeable), v000534 (Bush strong leader), and v000536 (Bush intelligent)	1-4	-.011	.544
clinint	from v000855 (Clinton moral) and v000859 (Clinton dishonest)	1-4	-.158	-.450
clincare	from v000856 (Clinton cares about people like you) and v000861 (Clinton out of touch)	1-4	-.004	-.584
clincomp	from v000857 (Clinton knowledgeable), v000858 (Clinton strong leader), and v000860 (Clinton intelligent)	1-4	.089	-.477
v00022	refusal conversion indicator	0-1	-.046	.016