DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

TWO-STAGE ESTIMATION OF NON-RECURSIVE CHOICE MODELS

R. Michael Alvarez
California Institute of Technology

SOCIAL SCIENCE WORKING PAPER 905

October 1994

TWO-STAGE ESTIMATION OF NON-RECURSIVE CHOICE MODELS

R. Michael Alvarez

Abstract

Questions of causation are important issues in empirical research on political behavior. Most of the discussion of the econometric problems associated with multi-equation models with reciprocal causation has focused on models with continuous dependent variables (e.g Markus and Converse 1979; Page and Jones 1979). Since many models of political behavior involve discrete or dichotomous dependent variables, this paper turns to two techniques which can be employed to estimate reciprocal relationships between dichotomous and continuous dependent variables. One technique which I call two-stage probit least squares (2SPLS) is very similar to familiar two-stage instrumental variable techniques. The second technique, called two-stage conditional maximum likelihood (2SCML), may overcome problems associated with 2SPLS, but has not been used in the political science literature. First I show the properties of both techniques using Monte Carlo simulations. Then, I apply these techniques to an empirical example which focuses on the relationship between voter preferences in a presidential election and the voter's uncertainty about the policy positions taken by the candidates. This example demonstrates the importance of these techniques for political science research.

TWO-STAGE ESTIMATION OF NON-RECURSIVE CHOICE MODELS *

R. Michael Alvarez

1 Introduction

Many interesting aspects of political behavior involve dichotomous decisions. For example, a potential voter decides whether to go to the polls on election day (Wolfinger and Rosenstone 1980); activists decide to donate time or resources to a campaign (Verba and Nie 1972); and potential candidates decide to enter particular races in certain political contexts (Banks and Kiewiet 1989; Canon 1990; Jacobson and Kernell 1981; Schlesinger 1966). One of the most interesting and studied choices, though, occurs once a citizen has entered the voting booth, since in American national elections voters essentially have two ways to cast their ballot — for the Democrat or the Republican. While sometimes there are other viable choices in each of these examples, much of the empirical research in political behavior has examined binary choices.

The practical econometric difficulties associated with dichotomous dependent variables are now well known in political research. Given that ordinary-least squares does not perform well when the dependent variable is binary, researchers now turn to either linear probability models or to logit and probit models (Achen 1986; Aldrich and Nelson 1984). In either framework, under certain assumptions, the dichotomous nature of the dependent variable is not an obstacle to unbiased estimation of model coefficients.

However, researchers using these dichotomous dependent variable models have not made much progress toward including binary choice models in larger non-recursive models of political behavior. In fact, the prominent examples of non-recursive models in the literature either have introduced surrogate variables for binary candidate choices (e.g. Page and Jones 1979) or have resorted to least squares estimation of a binary choice equation (e.g. Markus and Converse 1979). In only a few instances have researchers attempted to deal with the problems of endogeneity in discrete choice models in political

*John Aldrich and John Brehm provided important comments and advice. The John M. Olin Foundation provided support for this research.

science, with the most notable being models of party identification (Franklin and Jackson 1983; Fiorina 1982). But even these models have not involved binary choices.

In this paper I examine two techniques which can be used to estimate reciprocal relationships between dichotomous and continuous dependent variables. One technique which I call two-stage probit least squares (2SPLS) is very similar to familiar two-stage instrumental variable techniques. The second technique, called two-stage conditional maximum likelihood (2SCML), overcomes many of the problems associated with 2SPLS, but has not seen adoption in the political science literature. The next section discusses the two models. Then I turn to a presentation of the properties of each model, by examining Monte Carlo simulations. Thereafter I show the applicability of both models to a problem of contemporary interest.

2 Non-Recursive Two-Stage Choice Models

I begin this discussion with a simple two-variable non-recursive system:

$$y_1 = \gamma_1 y_2 + \beta_1 X_1 + \mu_1 \quad (1)$$
$$y_2^* = \gamma_2 y_1 + \beta_2 X_2 + \mu_2 \quad (2)$$

where y_1 is a continuous variable, X_1 and X_2 are independent variables, μ_1 and μ_2 are error terms, γ and β are parameters to be estimated, and:

$$y_2^* = \begin{cases} 1 & \text{if } y_2 > 0 \\ 0 & \text{if } y_2 \leq 0 \end{cases}$$

From these, the reduced-form equations are:

$$y_1 = \pi_1 X_1 + \pi_2 X_2 + \nu_1 \quad (3)$$
$$y_2^* = \pi_3 X_1 + \pi_4 X_2 + \nu_2 \quad (4)$$

Within this simple set-up, a number of strategies could be followed to estimate the parameters of interest in Equations 1 and 2. First, if we were to assume that the usual OLS assumptions held for Equation 1 and the usual assumptions for the probit model held for Equation 2, we could proceed with independent estimation of each equation. However, estimation of Equation 1 by OLS and 2 by probit would produce consistent estimates of the actual coefficients only by imposing the following restrictions on the model:

$$E(r_{y_2\mu_1}) = \frac{1}{N} E \sum y_{2n}\mu_{1n} = 0$$
$$E(r_{y_1\mu_2}) = \frac{1}{N} E \sum y_{1n}\mu_{2n} = 0$$

To put it in words, only if the endogenous variable on the right-hand side of each equation is uncorrelated with the error term in that equation can we expect OLS or probit to produce consistent estimates of the coefficients of interest in either equation.

Of course, whether these assumptions are met in practice is dependent upon the substantive problem being modeled. However, there are good reasons to believe that assumptions like these are rarely met in practice. First, if common factors are left out of the specification of the model, and these factors influence each dependent variable, then these assumptions will be violated. Notice that this requires that we be very confident in the "correct" specification of both equations; if even one variable is left out of the right-hand side of each equation, estimation of these equations by OLS or probit will yield incorrect results. Second, if we do not correctly measure the endogenous variables, that measurement error can itself lead to the violation of these assumptions. Clearly, these two reasons alone should underscore the desirability of models which avoid these assumptions.

However, two-stage estimation of models which avoid these strict assumptions has been discussed in the literature, with two different techniques advocated (Achen 1986; Amemiya 1978; Maddala 1983). In the first approach, two-stage probit least squares (2SPLS), reduced-form equations for each endogenous variable are estimated initially. The reduced-form equation for the continuous variable (Equation 3) is estimated in the usual fashion, using ordinary least squares, while the reduced-form equation for the binary choice variable (Equation 4) is estimated via probit analysis. The parameters from the reduced-form equations are then used to generate a predicted value for each endogenous variable, and these predicted values are then substituted for each endogenous variable as they appear on the right-hand side of the respective equation (i.e., Equations 1 and 2).[1] Then the equations are estimated, with the predicted values from the reduced-forms serving as instruments on the right-hand sides of the equations. It has been shown that the estimates obtained in this second stage are consistent (Achen 1986; Amemiya 1978).[2]

But, the estimated standard errors of 2SPLS are likely to be biased. For the continuous variable equation, since they are estimated by two-stage least squares, the standard errors can be easily corrected by multiplying the estimated standard errors by an appropriate weighing factor, as summarized in Achen (1986: 43). However, for the binary choice equation, the standard errors cannot be easily corrected (Achen 1986: 49). The asymp-

[1]Note that the predicted value from the probit reduced-form is the linear predictor, βX_i, not a transformed probability for each voter.

[2]The use of two-stage, or limited-information models, instead of full-information models, can be justified on two grounds. First, limited-information models are easier to estimate and interpret than their full-information counterparts. Derivation of a full-information likelihood function for the model presented later in this paper yielded an exceptionally complex function, which made estimation computationally difficult (an example of the complexity of the FIML case can be seen in King [1989, section 8.2]). Second, full-information models, while theoretically more efficient since they utilize information in the data more fully, can be quite problematic if even one of the equations in the model is misspecified since the biases associated with specification errors will be distributed throughout the model. Limited-information models are not problematic in this regard, since they ignore information about the joint distribution of the error terms across the equations, which leads to a loss of potential efficiency.

totic covariance matrix of the probit estimates has been derived by Amemiya (1978), but is exceptionally complex and computational difficult. Indeed, those in the political science literature who have utilized the 2SPLS methodology have been willing to settle with consistent estimates and possibly incorrect standard errors, due to this computational difficulty (Fiorina 1982; Franklin and Jackson 1983).

Yet it is important to be able to estimate reliable standard errors, so that we can conduct statistical hypothesis tests. The second estimation technique advanced in the literature should mitigate the problems with incorrect standard errors, and by using this technique we should not have to resort to ad hoc corrections of the coefficient standard errors. Rivers and Vuong (1988) developed what they term the two-stage conditional maximum likelihood (2SCML) approach to obtaining consistent and asymptotically efficient estimates for the probit equation. This approach assumes that we are only interested in recovering the structural parameters for the probit equations.[3] Then, to estimate the probit coefficients and their variances in the 2SCML method, we first estimate the reduced forms for the continuous variable equations, obtain the residuals from the reduced form regression, and add these residuals to the probit equation for the binary choice variable as two additional variables with corresponding parameters to be estimated. Rivers and Vuong demonstrate that this method produces consistent and asymptotically efficient estimates.[4] Rivers and Vuong also demonstrate that the joint statistical significance of the parameters of the reduced-form errors is a robust exogeneity test.

Thus, there are two different techniques which can be used to estimate consistently the coefficients for the binary choice equation (Equation 2) — 2SPLS and 2SCML. The major difference between the two techniques is that the 2SCML technique should produce standard error estimates which are more efficient than 2SPLS; both should provide consistent coefficient estimates. So, by estimating the binary choice equation using both techniques we may be better able to evaluate the second-stage probit estimates of parameters and standard errors. If the two methodologies produce similar results for the binary choice equations, we can greatly increase our confidence in the results.

Notice, though, that the 2SCML model does not, by assumption, deal with the continuous variable equation. If we are interested in recovering those parameters as well, the only viable approach is to use an instrument taken from the reduced-form for the binary choice model (equation 4) and estimate the coefficients of Equation 1 with this instrument on the right-hand side (RHS). Given that this is almost identical to textbook two-stage least squares, in subsequent discussion I call this component of 2SPLS the 2SLS technique.

[3]This assertion is not problematic, since it is possible using the usual 2SPLS method to obtain consistent and efficient estimates of the coefficients in the continuous variable equations.

[4]They show that the 2SCML estimators are clearly asymptotically efficient when in the probit equation, the right-hand side endogenous variables are actually exogenous, or when the probit equation is just identified. However, their Monte Carlo evidence shows that the 2SCML estimator is more efficient than the other classes of simultaneous probit estimators even if these conditions are not met.

While these techniques for non-recursive choice models should be employed frequently in political science research, they are not. The 2SPLS estimator has seen limited applications in the political science (Alvarez 1992; Fiorina 1982; Franklin and Jackson 1983), while the 2SCML estimator has not been used in published political science work. Also, little is known about the performance of these estimation techniques, other than the Monte Carlo work in Rivers and Vuong (1988). Accordingly, to evaluate the performance of these two estimation techniques for non-recursive choice models, the next section of this paper presents the results from Monte Carlo simulations. Then, I present a substantive example of a problem in which endogeneity is suspected in a system of equations. The application of both techniques to this problem, focusing on the relationship between voter choice and voter uncertainty about candidate policy stances, underscores the importance of these techniques for political science research.

3 Monte Carlo Results

3.1 Experimental Methodology

To examine the properties of these two statistical models, Monte Carlo simulations were performed. The simulations were based on 1000 replications of a 300 observation data set, the sample size being chosen to closely approximate the size of datasets typically employed in political science research.[5] First, 300 observations of three normally-distributed variables were randomly drawn.[6] These constituted the systematic component (the X matrix) for a simple two equation model:

$$y_2 = x_1 + 2x_2 \tag{5}$$
$$y_1 = 2y_2 + x_2 + 1.5x_3 \tag{6}$$

Then, an error term for the first equation (y_2) was drawn from a normal distribution with mean zero and unit standard deviation. An error term for the second equation was constructed by drawing another normal variate with mean zero and unit standard deviation, and transforming it with the following equation:

$$\varepsilon_1 = \lambda(\mu_1) + N(1,0) \tag{7}$$

where N(1,0) is the newly-drawn random variate, μ_1 is the error for the first equation, and ε_1 is the error term for the second equation. By changing the values of λ I can simulate different degrees of correlation between these two error terms, and hence observe the properties of the models in response to varying degrees of error correlation. Each

[5]The Rivers and Vuong Monte Carlo analysis was based on 1000 replications of a 100 observation dataset.

[6]These three normal variates were distributed with zero means and unit standard deviations. All were generated using SST v. 2 on a Gateway 486/66.

observation of y_1 was then recoded into a binary variable, depending upon whether the observation was above or below the mean value of y_1.

Next, an instrumental variables regression for y_2 was estimated, using the three independent variables, and predicted values for y_2 were calculated from the reduced-form estimates. 2SPLS and 2SCML probit equations were then estimated for y_1, with the appropriate instrument for y_2 from the reduced-form equations substituted on the right-hand side of the equation for 2SPLS or with the reduced-form residuals added to the RHS for 2SCML. This procedure, starting from the drawing of the first error term, was then replicated 1000 times for 13 different values of λ (producing error correlations ranging from -.95 to .95), and summary statistics for the coefficients of these replications were calculated.[7] To evaluate the performance of the 2SPS model for estimation of the continuous variable equation, this methodology was altered. Each observation for y_2 was recoded into a binary variable, and a probit instrument was estimated from the reduced-form equation of y_2, and this instrumental variable (the linear predictor of y_2) was substituted for the actual value of y_2 in the continuous variable equation (here y_1). This was also replicated 1000 times for the same 13 values of λ; summary statistics for the coefficients across all of these replications were calculated.

3.2 Monte Carlo Results

The results of the Monte Carlo simulations are presented in three different ways. First, I show the average estimated coefficients for 2SPLS, 2SCML, and 2SLS graphically over the range of error correlations (Figures 1-3). Second, I give the mean square errors of each estimator, which is a measure of the average squared distance between the estimated coefficients and their true values.[8] The mean squared error estimates are given in Tables 1-3. Last, I present in Tables 4-6 estimates of the average estimated standard errors relative to the actual standard errors of the coefficients to examine specifically the estimated coefficient standard errors of the three estimators.

First, perhaps the most pressing question concerns the ability of the 2SPLS and 2SCML techniques to recover the actual values of the coefficients in the binary choice equation. The relative abilities of each technique to estimate the true coefficients are shown quite clearly in Figures 1 and 2. There, the average estimated coefficients are plotted across the range of error correlations.

---

[7]An excellent summary of the Monte Carlo simulation approach is in Hanushek and Jackson (1977). Fishman (1986) and Mitrani (1982) demonstrate how the Monte Carlo approach can be applied to discrete data. The Monte Carlo models here were estimated using SST v. 2 (Dubin and Rivers 1992).

[8]The mean square error criteria has both decision theoretic and statistical interpretations. In terms of decision theory, the mean square error criteria can be thought of as a measure of the loss from using the particular estimator, or as the risk of using the estimator (taking the expectation of the loss). In terms of statistical theory, the mean square error criteria is useful since it incorporates both the precision and bias of the estimator (Judge et al., 1988).

Recall that the actual values of these three parameters are $\gamma_1=2$, $\beta_2=1$, and $\beta_3=1.5$. With these in mind, the ability of each technique to estimate each parameter is obvious. In the 2SPLS results (Figure 1), when the correlations between the two error terms is quite large (approaching 1 or -1), it is apparent that this technique does not come very close to the actual value of any of the three parameters. In these cases, there is a substantial *negative* bias in the 2SPLS estimates; the average estimated coefficients here are approximately two-thirds of their actual values. However, when the error term correlations approach zero from either side this technique does a relatively good job at estimating parameters quite close to their actual values. Last, only in a very narrow region around error correlations of zero does the 2SPLS model do a good job of reproducing the actual coefficients.

But the 2SCML model (Figure 2) produces estimates of the coefficients which are close to their true values across the entire range of error correlations. The fact that the 2SCML technique produces estimates of the model parameters quite close to their actual values, no matter the error correlation, does provide evidence for the utility of this model over the 2SPLS model.

Last, in Figure 3 I present the results for the 2SLS Monte Carlo results — where a predicted value for the first equation was obtained using a probit reduced-form model, which was substituted in the continuous variable equation for the right-hand endogenous binary variable. As is apparent in Figure 3, when the error correlations are high (greater than .6 or less than -.6) the coefficient on the instrument (γ) clearly has a positive bias. And this bias is substantial: at error correlations of greater than .8 or less than -.8 the estimates of γ are inflated by at least a factor of 2. However for the other two coefficients, the 2SLS model does a remarkable job of reproducing the true parameter values no matter the magnitude of the error correlation.

Next, the mean squared error (MSE) calculations give another way to evaluate these estimators. These are given in Tables 1 (2SPLS), 2 (2SCML) and 3 (2SLS). We can directly evaluate the differences between 2SPLS and 2SCML with the MSE criteria. At extreme error correlations the MSE for the 2SPLS models are quite large and many times the size of those for the 2SCML model: with error correlations of .95 or -.95 the MSE for the 2SPLS estimates are about 10 times the estimates for 2SCML; with error correlations of .70 or -.70, the MSE for the 2SPLS are about 3 times the MSE for the 2SCML estimates. But when the error correlations are inside those extremes, the two estimators are roughly equivalent in their ability to reproduce the true model parameter values.

Next, in Table 3 we see the MSE estimates for the 2SLS model. Compared to the earlier models, this technique does a poor job of reproducing the coefficient on the RHS endogenous variable (γ) in a broader range (greater than .45 or less than -.45). Within that interval, this technique is relatively accurate. Also, it is again clear that the other two coefficients are estimated close to their true values across the entire range of error correlations.

The last set of results from these Monte Carlo simulations focuses on the estimated coefficient standard errors. In Tables 4-6 I present the magnitude of the average estimated standard error relative to the actual standard error of the average estimated coefficient. These calculations will tell whether the estimated standard errors are larger or smaller than the actual standard error of the coefficients; if the quotient is greater (lesser) than unity, then the estimated standard error is larger (smaller) than the actual standard error.

The 2SPLS standard errors are generally less than unity, except for those in the last two rows (Table 4). Thus, the estimated standard errors in the 2SPLS model are *smaller* than the actual standard errors, although not by a great extent. The general pattern across the range of error correlations, though, shows that the quotient is nearest unity at high error correlations, and that it falls to the minimum at error correlations near zero. This is true for all three parameters in the probit equation.

But a similar pattern in the 2SCML standard errors is difficult to spot (Table 5). Here, the estimated standard errors (like the estimated coefficients) do not seem sensitive to the magnitude of the error correlations. Furthermore, the 2SCML standard errors are similar to the 2SPLS standard errors in that both are *smaller* than the actual standard errors. However, in both of these techniques for estimating the binary choice equation, it is clear that the estimated standard errors are smaller than the actual standard errors.

The last set of standard errors to examine are for the 2SLS model. These estimated standard errors (Table 6) are quite close to zero, showing that they are generally much smaller than the actual coefficient standard errors. Additionally, there is a pattern across the range of error correlations where the quotients are closest to unity at error correlations around zero; as the error correlations increase, the estimated standard errors get relatively smaller.

So what conclusions can be drawn from these Monte Carlo simulations? First, the 2SCML technique does a better job of accurately estimating the model parameters than its rival, 2SPLS. This conclusion is warranted by the fact that no matter what the error correlation, this model can estimate coefficients closer to the true values than the alternative technique, 2SPLS. But, when the error correlations are small, the 2SPLS model does accurately estimate model parameters. Second, the technique for estimating the coefficients of the continuous variable equation (2SLS) is accurate in estimating the non-endogenous independent variables across possible error correlations. However, 2SLS does a poor job in accurately estimating the endogenous variable coefficient when the error correlations are large.

Third, while the 2SPLS and 2SLS standard error estimates are quite sensitive to the magnitude of the error correlation, again the 2SCML standard error estimates are not. But all of these techniques under-estimate the standard errors of the coefficients. Clearly, when estimating these models in small samples, there is cause for concern about the estimated standard errors. This is a problem with these two-stage models, and this may necessitate the use of corrections to the estimated coefficient variances.

4 Modeling Votes and Uncertainty

4.1 The Theoretical Model

The substantive example concerns the relationship between voter choice in a presidential election and voter uncertainty about the positions of the candidates on various issues. First, as has been shown in the positive theory of voter decision making under uncertainty, given the assumption that voters are risk averse (an assumption implied by the assumption that voter utility functions for candidates are single-peaked and concave), uncertainty about the positions of candidates on policy issues should depress a voter's utility for a candidate (Alvarez 1992; Bartels 1986; Enelow and Hinich 1984; Franklin 1991). This is an important insight into how imperfect information influences voter decision making, since the more uncertain a voter is about a candidate, the less likely the voter should be to support that candidate. Thus, it is critical for the positive theory of voter decision making under uncertainty that this implication be tested rigorously. Additionally, it is important to understand the variation in candidate uncertainty across voters in a presidential election, especially if we want to think about implementing electoral reforms which make citizens better informed about the choices they make each election year.

The two primary components of the model are a voter's preferences over presidential candidates, and their uncertainty about the policy positions taken by the candidates. Beginning with the former, the functional form for a voter's preferences when there is uncertainty about the candidate's policy positions can be easily understood using the spatial theory of elections (Enelow and Hinich 1984). So, the utility the voter expects to obtain from candidates G and J is:

$$\begin{aligned} E[U_{iJ}] &= c_{iJ} - (p_J - x_i)^2 - \sigma^2_{iJ} \\ E[U_{iG}] &= c_{iG} - (p_G - x_i)^2 - \sigma^2_{iG} \end{aligned} \tag{8}$$

where p_K denotes the position of candidate K on a policy issue, x_i the position of the voter, σ^2_{iK} the voter's uncertainty about the candidate's position on the issue, and c_{iK} other non-policy factors entering into the voter's utility evaluation. If the election involves these two candidates, then the decision rule for the voter is simple: vote for candidate J if $E[U_{iJ}] \geq E[U_{iG}]$. Or:

$$\begin{aligned} c_{iJ} - (p_J - x_i)^2 - \sigma^2_{iJ} \geq \\ c_{iG} - (p_G - x_i)^2 - \sigma^2_{iG} \end{aligned} \tag{9}$$

Here it is clear that if $c_{iJ} = c_{iG}$ and $(p_J - x_i)^2 = (p_G - x_i)^2$, then the voter's decision hinges on the relative magnitudes of σ^2_{iJ} and σ^2_{iG}. Thus if the voter evaluates the two candidates identically on non-policy dimensions, *and* the voter is the same distance from *both* candidates on the issue, then they will support candidate J only if $\sigma^2_{iJ} \leq \sigma^2_{iG}$ which is true only when they are more (or equally) certain of J's position on the issue.

The second component of the model focuses on the determinants of a voter's uncertainty about a candidate's policy positions. Three factors account for voter uncertainty of the candidate's policy positions: their personal information costs, their exposure to information, and the flow of information during the campaign. Basically, the more costly it is for a voter to obtain, process, and store information, the more uncertain they should be about the candidate's position; the less exposed to information, and the less attentive and interested the voter is, the greater their uncertainty about the position of the candidate; and the greater the amount of information available about the candidates, the less the uncertainty a voter will have regarding the positions of the candidates (Alvarez 1992). With these variables a model of uncertainty can be constructed under certain assumptions about the relationship between these independent effects and voter uncertainty.

Then, this uncertainty should directly influence the voter's evaluation of the candidate, controlling for other policy and non-policy factors relevant to the voter's calculus. The uncertainty measure is thus an important explanatory variable in the determinants of candidate evaluation and choice, as well as an important endogenous variable. This causal process relating uncertainty and candidate evaluation and choice is usually depicted in the literature as a hierarchical model (Bartels 1986; Franklin 1991). This hierarchical model can be shown as two equations:

$$v_{iJ} = \beta_1 + \beta_{11}X_{1i} + \beta_{12}X_{2i} + \tau_{11}u_{iJ} + \xi_{1i} \tag{10}$$

$$u_{iJ} = \beta_2 + \beta_{21}X_{1i} + \beta_{23}X_{3iJ} + \tau_{21}v_{iJ} + \xi_{2i} \tag{11}$$

where v_{iJ} if voter i's uncertainty about candidate J, X_{1i} are demographic variables, X_{2i} are variables measuring voter i's information costs and exposure to political information, X_{3iJ} are variables relating to i's evaluation of candidate J's policy and non-policy attributes, u_{iJ} is the utility or evaluation of the voter for candidate J, the β's and τ's are parameters to be estimated, and ξ' are error terms in each model.[9]

Past research regarding voter uncertainty of candidate policy positions has assumed that τ_{11} is zero, implying that a voter's evaluation of the candidate does not influence their uncertainty of the candidate. Under this assumption, there is no reason to suspect a correlation between ξ_{1i} and ξ_{2i}, and therefore, the estimates of the parameters of each model have the usual properties under certain other assumptions. But if τ_{11} is *not zero*, then the two error terms are likely to be correlated, and the error term in each of the equations is likely to be correlated with right-hand side variables in each equation. As a consequence of this endogeneity, the estimates of the parameters in this model are likely to be biased.

[9]The demographic variables (X_{1i}) are in the first equation since I expect that two demographic groups, minorities and females, might be more uncertain than others. It is possible that other demographic groupings might be useful, like income and socio-economic status, but these are concepts which surveys are not well-suited to measure. These same variables are in the second equation to control for non-policy and candidate variations across individuals in their candidate preference.

But are there theoretical reasons to suspect that a voter's evaluation of a candidate might influence the amount of uncertainty they have about the candidate? Assume for a moment that the situation is the typical two-candidate presidential race — under what conditions might we expect that voter uncertainty about the candidates is conditional on their respective evaluations of the two candidates? Downs, in his chapter "The Process of Becoming Informed," argues:

> Three factors determine the size of his planned information investment. The first is the value to him of making a correct decision as opposed to an incorrect one, i.e., the variation in utility incomes associated with the possible outcomes of his decision. The second is the relevance of the information to whatever decision is being made . . . The third factor is the cost of the data (Downs 1957: 215-216).

Take a voter to whom the value of making a correct decision is quite high, and to whom the relevance of the available campaign information is quite high, but the cost of obtaining and utilizing this information is quite low. It is reasonable to argue that such a voter would attempt to minimize the uncertainty associated with both candidates, regardless of their prior evaluations of each candidate, since the value of being correct is high, and the costs are low.

But what of a voter to whom the value of being correct is quite low, but the costs of information are high and relevant information is quite difficult to obtain? It is reasonable to argue that such voters might be attentive to, or process, only information about their preferred candidate, and avoid or ignore information about the other, less preferred candidate. This type of process is similar to information processing strategies discussed in the political cognition literature — termed "top-down" or "theory-driven" processing by Rahn (1990), or schema-based processing by Fiske and Pavelchak (1986), or those described in the literature on how the media influences voter information processing (Graber 1988; Lazarsfeld, Berelson and Gaudet 1944; Patterson 1980). In any case, there are strong theoretical reasons to believe that the uncertainty voters possess about candidates might be contingent not only on their information costs, awareness and attentiveness, and the information made available by the campaign, *but also upon their existing evaluations of the candidates.*

Thus, there are theoretical reasons to suspect that τ_{11} might be non-zero, and that a simultaneous relationship exists between candidate evaluations and voter uncertainty. This means that independent estimation of these uncertainty equations is inappropriate, and would lead to incorrect estimates of the coefficients in each equation. Rather, the endogeneity between these two variables must be appropriately modeled so that consistent empirical results can be obtained.

The terms in the model can be expanded to clarify the general statistical model of uncertainty and evaluations:

$$v_{iJ} = \beta_1 + \beta_{11}X_{1i} + \beta_{12}X_{2i} + \tau_{11}u_{iJ} + \xi_{1i} \quad (12)$$
$$u_{iJ} = \beta_2 + \beta_{21}X_{1i} + \beta_{23}X_{3iJ} + \beta_{24}X_{4iJ} + \tau_{21}v_{iJ} + \xi_{2i} \quad (13)$$

where v_{iJ} is the voter's uncertainty of the candidate's policy positions, X_{1i} is a vector of demographic variables measuring information costs, X_{2i} is a vector of variables expressing the voter's exposure and depth of political information, X_{3iJ} is a vector of variables for policy-specific information about the candidate, X_{4iJ} are variables for non-policy information about the candidates, and u_{iJ} denotes the voter's utility for the candidate.

Above, the voter's decision rule in a situation of two candidate competition was developed, and it was shown that they voted for the candidate with the greatest expected utility. Two changes need to be made in the model to reflect this. First, the voter's uncertainty for the other candidate (G) must be modeled. This is accomplished by simply adding another equation to the model for v_{iG}. To express comparative candidate evaluations, however, I substitute a dichotomous indicator for whether the voter prefers candidate J to the other G, and add terms relevant to the voter's evaluation of candidate G to this equation.[10]

These two changes yield the model:

$$\begin{aligned} v_{iJ} &= \beta_1 + \beta_{11}X_{1i} + \beta_{12}X_{2i} + \tau_{11}u_i + \xi_{1i} && (14)\\ v_{iG} &= \beta_2 + \beta_{21}X_{1i} + \beta_{22}X_{2i} + \tau_{21}u_i + \xi_{2i} && (15)\\ u_i* &= \beta_3 + \beta_{31}X_{1i} + \beta_{33}X_{3iJ} + \beta_{34}X_{4iJ} + \\ &\quad \tau_{31}v_{iJ} + \beta_{35}X_{3iG} + \beta_{36}X_{4iG} + \tau_{32}v_{iG} + \xi_{3i} && (16) \end{aligned}$$

where v_{iJ} is the voter's uncertainty for candidate J, v_{iG} is their uncertainty of candidate G, X_{3iJ} and X_{3iG} are vector of variables for policy-specific information about each candidate, X_{4iJ} and X_{4iG} are variables for non-policy information about the candidates, and u_i* denotes the binary preference variable expressing whether i prefers candidate J to candidate G. Also, as before the β's are τ's are parameters to be estimated, and the ξ's are error terms.

Given that the structural model here is non-recursive, estimation of this uncertainty and evaluation model must be carefully considered. First, there must be enough information in the data matrix to obtain a unique set of parameter estimates — in other words, the structural model must be identified. Given that there are actually a number of restrictions in the structural model, the equations are overidentified, and given overi-

[10]In the empirical models below, this dichotomous indicator will always be preference for the Democratic candidate to the Republican candidate, where Democratic preference is coded one. Other measures of evaluation and preference are possible to operationalize from the survey data, but as I noted in Chapter 3, the use of the "thermometers" is suspect. Also, "thermometer" ratings were not available in the 1976 panel data.

dentification, instrumental variables procedures could be employed to produce consistent estimates of the parameters in the structural model.[11]

There is a further complication involved in the estimation of the structural model. Since the comparative utility variable is a dichotomy — representing whether the voter stated preferring one candidate to the other — straightforward application of two-stage least squares, the typical solution to the problem of simultaneous equations, is rendered inappropriate. Thus I estimate the binary choice equation (Equation 16) using both the 2SPLS and 2SCML approaches. The continuous variable equations (Equations 14 and 15) are estimated using the 2SLS approach.

4.2 Uncertainty and Voting — Estimation Results

In the substantive example presented here, I use data from the 1976 presidential election. This is a particularly interesting election to examine, for methodological and substantive reasons. Methodologically, this allows the use of the 1976 panel study of the presidential election campaign conducted by Thomas E. Patterson. This dataset, along with the accompanying data on the media coverage of this campaign, provides an excellent vehicle for studying uncertainty and voter decision making.

Substantively, recall that Carter had been a virtually invisible governor before the spring months of 1976; and before being appointed to serve the remainder of Nixon's term, Ford was a low-key House Minority Leader. Neither candidate was a long-established national figure, and neither was very well-known at the beginning of the campaign. Partly as the result of their relative obscurity, but also partly caused by lingering memories of Watergate, both the candidates and the mass media were largely preoccupied during the general election with the character of the candidates. Yet this does not imply that matters of "substance", like their positions on policy issues like abortion, foreign affairs, or the domestic economy were ignored. The statements of the candidates were covered widely in the press (Patterson 1980; Witcover 1977). But Ford and Carter were extremely moderate in most of their positions, and typically there was little which differentiated them on public policy.

Therefore, this could be an election in which voter uncertainty over the candidate's policy positions was exceptionally extreme, given two moderate and relatively poorly-known candidates. So it is important to understand the determinants of this uncertainty in 1976. Also, it is reasonable to expect that this uncertainty may have strongly influenced voter evaluations of the candidates. Given that the campaign was focused on character questions, it would also not be surprising to see these non-policy aspects of

[11] Below, when this structural model is operationalized for each election, verification of the identification status of each equation is apparent through checking the order condition, which is a necessary, but not sufficient condition, for identification of recursive systems (Hanushek and Jackson 1977). The order condition states that the number of exogenous variables excluded from each equation must be at least as great as the number of endogenous variables included in each equation.

candidate evaluation looming large in voter assessments of the candidates. This might have been further exacerbated by the difficulties associated with distinguishing between two relatively proximal candidates. But policy differences between the candidates might also have had an effect, since policy issues were discussed in the campaign, especially in the two fall debates.

In the next two sections, I present and discuss the two-stage results from the continuous two-stage equations (uncertainty) and then the binary voter choice two-stage equations. Full discussion of the operationalization of variables used in these models is in the Appendix, along with the reduced form estimates.

4.2.1 Determinants of Uncertainty

The uncertainty models from the 1976 campaign are given in Table 7, which gives the independent variables in the first column, the estimates and standard errors for the Carter uncertainty equation in the second column, and similar statistics for the Ford uncertainty equation in the third column. The data used to estimate these models came from the October wave of the 1976 Patterson panel study.

First, note that the models fit the data reasonably well. Both equations have similar R^2's and standard errors, which indicate that the two-stage models do account for a good deal of the variance in the voter's uncertainty of Carter and Ford in 1976. Additionally, almost all of the estimates are in the anticipated direction and have standard errors sufficiently small that the estimates are reliable indicators of the population parameters at reasonable levels of statistical significance.

The indicators of voter information costs in these models, the first four variables in the table (education, political information, gender and race) are all in the expected direction. That is, better educated and informed voters were less uncertain of the policy positions of Carter and Ford in 1976, while both women and racial minorities were more uncertain of the positions of these two candidates. And only the estimate for racial minorities fails to reach statistical significance in these models.

The rest of the variables in these two models, besides the relative candidate preference indicator, measure various dimensions of voter attachment to the political world, and their exposure to political information. Unlike the National Election Studies, however, the 1976 Patterson data contained a useful set of questions which allowed me to incorporate three additional variables into these two models, which indicated whether the respondent watched either or both of the televised debates, or recalled seeing some advertisement from the particular candidate's paid media campaign. So the estimates for watching the debates or campaign advertisements allows examination of two specific types of exposure to campaign information.[12]

Of all of these variables, only the estimates for partisan strength and the first debate indicators are incorrectly signed, but are not statistically significant. All of the rest have the correct sign, and most do reach reasonable levels of statistical significance. The first two indicators — media exposure and political efficacy — both show that the more exposed and efficacious voters are statistically more certain of the policy positions of the two candidates in this election. Also, voters who watched the first debate were more certain of Ford's policy positions, while those who watched the second debate were more certain of Carter's positions. What is fascinating about these results is that they comport with both the information made available by these two debates in 1976, as well as popular perceptions concerning which of the two candidates had more effectively presented themselves and their campaign positions in each debate. Most observers concluded that in the first debate, Ford had articulated his positions on unemployment and the economy quite forcefully, and had put Carter on the defensive. Then, while Carter began to respond, the debates were interrupted by technical difficulties which most believed damaged Carter's ability to get his arguments across (Witcover 1977). During the second debate, both candidates attacked their opponent's foreign policy positions, which might account for the negative effect watching this debate appears to have had on the policy uncertainty for both candidates in the models. But the second debate was marred by Ford's "no Soviet domination of Eastern Europe" comment, retracted within the next five days — which might account for the only marginal reduction in Ford uncertainty for voters who watched the second debate.

Last, the candidate preference indicator is correctly signed for both candidates, but is statistically significant in only the Carter equation. The signs of these estimates indicate that the greater the likelihood of Carter support, the higher a respondent's uncertainty about Ford's policy positions and the lower their uncertainty about Carter's positions, *ceteris paribus*. But the magnitude of this effect is interesting, since it is significant for only the challenging candidate. This suggests that while voters do engage in selective information processing about presidential candidates, such strategies may not be necessary for incumbent candidates: voters may have already obtained enough information about incumbents to make selective processing unnecessary.

4.2.2 Determinants of Choice

Not surprisingly, the expectations about the role of uncertainty and voter decision making are demonstrated in the choice models. The results are presented in Table 8. Recall that

[12]The debate and advertisement questions were included in the October wave of the Patterson study, just after the two presidential debates (September 23 and October 6). However, in the NES surveys, questions about viewing advertisements are not included, and the debate questions, when included in the survey instrument, are in the post-election instrument and hence are not very reliable measures of exposure to debate or advertising information. Consequently, I do not include these indicators in the 1980-1988 uncertainty models.

the dichotomous dependent variable here is coded so that support for the Democratic candidate is the high category, and for the Republican candidate is the low category. Thus the parameter estimates express the relative effect of the particular variable on the probability of Carter support. I expect that the closer a voter is to Carter on the policy issues, the greater the voter's support for Carter (negative sign); the closer a voter to Ford on the issues, the lesser their support for Carter and the greater their support for Ford (positive sign). The more uncertain a voter is about Carter, the less likely they are to support Carter while greater uncertainty about Ford should lead to a greater likelihood of Carter support: thus I expect the uncertainty parameters to be negative for Carter and positive for Ford. The "better" a voter's evaluation of Carter's personal and professional characteristics, the more likely they should be to support Carter (positive sign); and likewise for Ford, where a higher evaluation of his character leads to greater probabilities of Ford support (negative sign). Last, Democratic identifiers should support Carter, and Republican identifiers Ford (negative sign).

In Table 8, the first column gives the names of the independent variables in the model. The second column gives the parameter estimates and associated standard errors as estimated by the 2SCML model, and the third column gives estimates and standard errors from the 2SPLS model. Notice that both models fit the data very well. Each correctly predicts slightly over 95% of the cases in the sample, and the χ^2 statistic for each model shows that the models perform vastly better than a null, intercept-only model.[13] Furthermore, the variables of interest are all correctly signed and statistically significant at reasonable levels.

First, in both models, the effects of uncertainty are correctly signed, and are estimated relatively precisely. That is, the more uncertain a voter was of Carter, the lower the probability they would support Carter; the more uncertain a voter was of Ford, the lower the probability they would support Ford. This finding directly supports the first hypothesis advanced by the spatial model briefly discussed above.

[13]Neither of these goodness-of-fit statistics, however, are strong indicators of fit. The percent correctly predicted is strongly influenced by the distribution of high and low cases — when there are many high cases (say 80% of the cases are in the high category), the model will "correctly" predict better than when there are roughly similar frequencies of high and low cases. An alternative criterion is the "proportional reduction in error" (Hildebrand et al. 1976). Here, a simple PRE measure would be the model's prediction rate relative to a naive prediction rate, the percentage falling into the modal category. In the 1976 data, the modal category was Carter support (53.9%). By this criteria, the 2SCML model predicts 41.4% better and the 2SPLS model predicts 41.8% better than the naive model. The χ^2 statistic reported here is twice the difference between the log-likelihood of a naive model (with only an intercept included) and the log-likelihood of the actual models, often called the model's "deviance function" (McCullagh and Nelder 1991). The difference between log-likelihoods is often claimed to be distributed as a χ^2, with degrees-of-freedom equal to the number of restrictions. In the 2SCML model there are 11 degrees-of-freedom, and 9 in the 2SPLS model, making the χ^2 for each model highly significant. But the claim that the difference in log-likelihoods is approximated by a χ^2 is contingent upon the observations being distributed independently according to the binomial distribution, and upon the number of observations being small. Thus as n increases, the approximation breaks down, and large differences will be observed (McCullagh and Nelder 1991). With these caveats in mind, I report percent correctly predicted and χ^2 statistics.

Worthy of additional notice, though, is the observation that the effects of policy uncertainty were greater for Carter uncertainty than for Ford uncertainty. In other words, the voter's uncertainty for Carter's positions had a greater effect on which candidate the voter supported than the voter's uncertainty of Ford's positions. These differential effects are probably the result of greater uncertainty about the positions of Carter, who was somewhat less visible before the general election began, and who did not have the tools of an incumbent president to make his policy positions known to the electorate.

Relatedly, the estimates for the effects of policy issues upon candidate support in 1976 support this argument. Both squared issue distance terms are correctly signed, yet the estimate for Carter issue distance is statistically significant at only the p=0.10 level, while the similar estimate for Ford issue distance is estimated more precisely (p=0.05). But the effect of issues is greater for Ford than for Carter, as witnessed by the relative sizes of the coefficients for Ford issue distance. With more uncertainty about Carter's positions, and with Carter uncertainty having more of an effect on candidate support, it is not surprising that Carter's positions on the issues had less of an influence on voter evaluations of the candidates in 1976.

The other two sets of parameters of interest — the non-policy dimensions of candidate evaluations — are all correctly signed and precisely estimated. That is, the partisan affiliation of the voter clearly influenced their evaluations of Ford and Carter. Also, their assessments of the characters of the candidates influenced their evaluations, with higher evaluations of either candidate's personal and professional characteristics leading to greater support for the candidate. But it is interesting to note that here again, the effect of candidate characteristics on candidate support is greater for Ford than for Carter. Perhaps the decision by Ford and his advisors to focus on the "character issue" and to employ the "Rose Garden strategy" had some effect on the electorate, leading to more positive assessments of his character than for Carter.

Lastly, there are the two parameters from the reduced-form regressions for both Ford and Carter uncertainty (presented in the appendix), presented for the 2SCML model. Recall the discussion of these two parameters from above: Rivers and Vuong (1988) demonstrate that these parameters give a robust test of exogeneity. The two parameters are larger than their standard errors, but the relative magnitudes of the parameters to their standard errors are not great enough for the parameters to be considered statistically significant. But, the deviance function for the 2SCML model versus a similar model without these two parameters yields a χ^2 of 10.88, which is is larger than the critical value of 5.99 at 2 degrees of freedom. Therefore, endogeneity between candidate evaluation and voter policy uncertainty is evident, and needs to be accounted for in these models.

However, what is the magnitude of the estimated effects of these variables upon candidate evaluations in 1976? As is usually the case in non-linear probit models, the parameters in the evaluation models cannot be interpreted directly, since the models are non-linear and the effect of any particular variable on the probability of supporting one candidate is dependent on the values of the other variables and parameters in the model.

To give a more intuitive feel for the magnitude of two of the effects in the models, of candidate uncertainty and squared policy issue distance upon candidate support, I utilize graphical methods (King 1989; McCullagh and Nelder 1991).[14] The results for the candidate policy uncertainty parameters are graphed in Figure 4.

In Figure 4 the voter's uncertainty about each candidate's policy positions is graphed along the x-axis, and the probability that the voter would support Carter on the y-axis. The dark line gives the effect of Ford policy uncertainty on probability of Carter support, while the light line gives the effect of Carter policy uncertainty on the probability the voter would support Carter (holding the other variables constant at their sample mean values). The strong effect of uncertainty on candidate evaluation is clear in this graph. Take two identical voters, with "mean" values on all the variables, but one who is very certain of Carter's position on the issues (1) and the other who is relatively uncertain of Carter's positions (5). The graph indicates that the former voter — the relatively certain voter — has a very high probability of supporting Carter, while the uncertain voter is slightly less than 50% likely to support Carter. Thus, by changing the relative uncertainty of the voter from very certain to relatively uncertain, the probability of supporting Carter in 1976 changes by over 50%.

To examine the relative impacts of these two variables, note that both Ford and Carter uncertainty had roughly similar distributions: the uncertainty means (3.96 for Ford, 4.08 for Carter) and standard deviations (3.03 for Ford, and 3.07 for Carter). Under the assumption that these two variables are normally distributed, we would expect that 65% of the sample would have had uncertainty for both candidates within one standard deviation of the mean: roughly ranging from 1 to 7.[15] Across this range, it is apparent that Carter uncertainty does have a marginally larger impact of candidate evaluation than Ford uncertainty: there is a 95% change in probability of Carter support across this range of Carter uncertainty, and a 92% change in the same range for Ford uncertainty.

5 Discussion

This substantive example should make clear the importance of non-recursive choice models in political science. The theoretical model connecting voter preferences for candidates with their uncertainty about the candidates shows the endogeneity of these two factors,

[14]Graphical methods involve simple simulations using the parameters of the model and some combination of values of the independent variables. Here, I set all but one of the independent variables to their mean values in the sample of voters used to estimate the model (the descriptive statistics are in the appendix). Then, I vary the one variable of interest across a range of values the variable takes in the actual data. This produces an estimate of the linear predictor for each value of the variable of interest, which is then transformed into a probability by the use of the appropriate link function. In each of the graphical interpretations here, I use the parameters from the 2SPLS model.

[15]As I will show in subsequent chapters, though, the distributions of uncertainty for the presidential candidates examined in this dissertation are not normally distributed. In fact, they are by definition bounded on the left by zero (a variance or standard deviation is, by definition, positive), with a long tail to the left. I only mention the normal distribution in this context for convenience.

a point which has not been taken into consideration in past research on either topic. The empirical estimation of this non-recursive model demonstrates the inter-relation of preferences and uncertainty, and in so doing, produces interesting insights into the politics of presidential campaigns (Alvarez 1992).

As a research methodology, then, these two-stage models of binary and continuous variables have great promise. But the Monte Carlo evidence presented earlier in this paper shows that these models do have some troubling properties; while the 2SCML model does estimate the true values of the parameters accurately, it tends to understate the actual uncertainty inherent in these estimates. The 2SPLS and 2SLS models both estimate the coefficients relatively well when error correlations are modest; but with greater error correlations each produces estimates further from the true values. These two-stage non-recursive choice models are quite easy to estimate. In fact, they take no longer to estimate than recursive specifications.

Thus, a when faced with a substantive problem which involves a non-recursive set of binary and continuous variable equations, an appropriate methodology would include estimating and presenting both the 2SPLS and 2SCML estimates for the binary variable equation(s). As the Monte Carlo results made quite evident, under some conditions the 2SPLS technique alone may yield incorrect results. With well-specified theoretical models and correctly measured RHS variables, thought, it is likely that the error term correlations will not be in the extreme ranges and both techniques can produce quite similar results. That this occurred in the empirical example presented above indicates that by verifying the estimated results from one technique against another we can be much more confident in our results.

With these considerations in mind, more work on similar two-stage techniques is needed. In particular, there has been no work on non-recursive models involving discrete choice variables — whether the discrete variables are ordered or unordered. Even though some researchers have employed two-stage techniques to estimate models involving such variables (Fiorina 1982; Franklin and Jackson 1983), almost nothing is known about the properties of models involving discrete choice variables. Expanding our knowledge of these models is important for empirical political science research.

6 Appendix

6.1 Operationalizations of Variables

The coding of the variables from the 1976 Patterson panel study is complicated by the fact that the ICPSR documentation for the study contains no variable numbers. Consequently, I assigned variable numbers to the ICPSR codebook (ICPSR study 7990, First Edition, 1982) sequentially (question 1, page 1, "location number" is V1, while the last codebook entry for "weight factor" on page 195 is the last variable, V1664).

In the uncertainty equation, the following variables were used. The variable for education was taken from V9, and was coded: 1 for those with a grade-school education or less; 2 for those with a high school education; 3 for those with some college or vocational education; 4 for those with college degrees. Political information is a ten-point scale where the respondent was given a point for each time both parties were placed and the Democratic party was placed to the "left" of the Republican party on the seven-point issue and ideology scales. Gender and Race are dummy indicators, where Gender is 1 for females and 0 for males (from V21), and race is 1 for minorities and 0 for whites (from V24). Partisan strength is the folded partisan identification scale (V1569). Media exposure was constructed as a factor scale from variables measuring the regularity with which the respondent was exposed to news coverage in newspapers (V1328), news magazines (V1339), television news (V1358), and conversations with others (V1348). The principal components factor analysis yielded one factor, eigenvalue 8.37. The political efficacy variable is an index of external political efficacy from questions concerning big interests and government (V1575), faith and confidence in government (V1577), public officials and people like me (V1579). A principal components factor analysis was used to make a factor scale; the eigenvalue of the only factor extracted from the data was 2.80. The indicators for the first and second debates, and for whether the voter saw a candidate's advertisement are dummy indicators, from V1455, V1456 (for the debates) and V1386, V1393 (for candidate advertisements. Nine seven-point issue scales are available in this survey: government provision of employment, involvement in the internal affairs of other nations, wage and price controls, defense spending, social welfare spending, tax cuts, legalized abortion, crime, and busing. The uncertainty variable was constructed by subtracting the respondent's placement of the candidate on the issue from the candidate's position, where the latter was measured by the mean position across all respondents placing the candidate on the issue. Respondents who did not place the candidate were assumed to be maximally uncertain about the candidate's position.

In the voting choice equation, the candidate traits ~~variables~~ were taken from questions in the Patterson study asking respondents to rate the attractiveness of the candidate's personality (V1426, V1427), their leadership abilities (V1431, V1432), their trustworthiness (V1436, V1437), and their ability or competence (V1441, V1442), for Ford and Carter respectively. Factor scales were constructed of these items for each candidate, with eigenvalues of 11.5 (Ford) and 8.28 (Carter). All of the available seven-point issue scales were used to calculate the uncertainty and squared issue distance terms (with the candidate means employed in the latter variable for the position of the candidates). Party identification came from the standard seven-point scale (V1569). The dichotomous candidate preference variable came from the post-election interview question as to whom the respondent had vote for (V1614), and was coded 1 for a Carter vote.

The measure of voter uncertainty about candidate policy positions comes from Alvarez (1992). It is based upon a statistical notion of uncertainty, where the "spread" of points around a central tendency is commonly defined for a mean as $\sigma^2 = \frac{1}{n}\sum_{n=1}^{N}(x-\overline{x})^2$, where x denotes the n points in the sample, and $\overline{x}$ represents the mean value, or the central tendency, in x. With this representation in mind, consider:

$$v_{iJ} = \frac{1}{k}\sum_{k=1}^{K}(P_{iJk} - T_{Jk})^2 \tag{17}$$

where v_{iJ} represents the voter i's uncertainty in their placement of J, P_{iJk} gives i's placement of J on each of the relevant k policy dimensions, and T_{Jk} indicates the position of candidate J on policy dimension k.

Less technically, this is a representation of the voter's uncertainty about the candidate's position across the policy space, in terms of the net dispersion of the voter's perception of the candidate's position and the candidate's true position. The greater the dispersion of their perceptions of the candidate's position from the candidate's true position, the more uncertain they are about the candidate's position on the policy issues; the tighter this dispersion of points, the less uncertain they are about the candidate's position.

This representation of voter uncertainty is appealing for three reasons. First, unlike the measures of uncertainty advanced in the literature, this representation directly operationalizes uncertainty from the survey data, and does not infer a uncertainty measure from ancillary information about respondents.[16] Second, this measure meshes closely with the theory of uncertainty discussed in the paper, which allows for rigorous tests of the hypothesis advanced about the role of uncertainty in presidential elections. Third, this measure can be applied to existing survey data, particularly the historical data from the National Election Studies, where there have been questions asking respondents to place candidates on policy scales since 1972.

However, it is also important to note what this particular measure cannot do. One, unless repeated questions about the same policy issue are posed to the respondent, this measure cannot gauge uncertainty *about specific issues.* Instead, it is intended to measure more generally the uncertainty the voter has *across issues.* Also, the accuracy of this measure will rely on the accuracy of the questions used to measure both the voter's own position on the issue, and even more importantly, the candidate's position on the issue.

[16] Other indirect measures of uncertainty have been estimates by Bartels (1986), Campbell (1983) and Franklin (1991). Additionally, direct survey-based measures of policy uncertainty have been studied by Aldrich et al. (1982) and Alvarez and Franklin (1994). Unfortunately, none of the direct survey-based measures are available for the historical NES presidential election studies, other than the 1980 election (Alvarez 1992).

Table 1: Two-stage Probit Coefficients
Coefficient Mean-Squared Errors

Error Corr.	γ	β_2	β_3
-.95	1.7789	0.5041	1.0201
-.70	0.2304	0.0676	0.1296
-.60	0.0625	0.0324	0.0441
-.45	0.0025	0.0001	0.0004
-.20	0.0196	0.0025	0.0121
-.10	0.0625	0.0100	0.0324
0	0.0625	0.0064	0.0289
.10	0.0576	0.0064	0.0289
.20	0.0144	0.0049	0.0081
.45	0.0025	0.0025	0.0025
.60	0.0676	0.0196	0.0400
.70	0.2401	0.0625	0.1296
.95	1.8225	0.4489	1.0201

Note: These are the mean-squared error estimates of the coefficients from 1000 trials of a Monte Carlo simulation. of the 2SPLS technique

Table 2: Two-stage CML Probit Coefficients
Coefficient Mean-Squared Errors

Error Corr.	γ	β_2	β_3
-.95	0.1156	0.0064	0.0484
-.70	0.0841	0.0100	0.0484
-.60	0.0961	0.0100	0.0400
-.45	0.0625	0.0225	0.0441
-.20	0.0729	0.0121	0.0441
-.10	0.1024	0.0196	0.0576
0	0.0900	0.0100	0.0441
.10	0.0961	0.0121	0.0484
.20	0.0576	0.0169	0.0324
.45	0.0784	0.0121	0.0361
.60	0.0729	0.0169	0.0400
.70	0.0729	0.0196	0.0441
.95	0.0961	0.0289	0.0625

Note: These are the mean-squared error estimates of the coefficients from 1000 trials of a Monte Carlo simulation. of the 2SCML technique

Table 3: Two-stage Least Squares Coefficients
Coefficient Mean-Squared Errors

Error Corr.	γ	β_2	β_3
-.95	21.996	0.0784	0.0000
-.70	0.7225	0.0064	0.0001
-.60	0.2304	0.0036	0.0000
-.45	0.0529	0.0036	0.0001
-.20	0.0025	0.0036	0.0000
-.10	0.0001	0.0064	0.0001
0	0.0016	0.0009	0.0000
.10	0.0004	0.0049	0.0001
.20	0.0009	0.0036	0.0001
.45	0.0484	0.0036	0.0000
.60	0.2304	0.0064	0.0001
.70	0.6724	0.0025	0.0001
.95	23.426	0.1764	0.0004

Note: These are the mean-squared error estimates of the coefficients from 1000 trials of a Monte Carlo simulation. of the 2PLS technique

Table 4: Two-stage Probit Standard Errors
Estimated Relative to Actual SE

Error Corr.	σ_γ	σ_{β_2}	σ_{β_3}
-.95	.86	.77	.85
-.70	.78	.77	.81
-.60	.81	.77	.75
-.45	.83	.79	.78
-.20	.83	.80	.82
-.10	.76	.76	.81
0	.77	.71	.79
.10	.77	.79	.77
.20	.84	.81	.84
.45	.85	.85	.82
.60	.88	.91	.89
.70	.96	.97	.95
.95	1.09	1.21	1.1

Note: These are the quotients of the estimated coefficient standard errors relative to the actual standard errors of the coefficients from 1000 trials of a Monte Carlo simulation of the 2SPLS technique

Table 5: Two-stage CML Standard Errors Estimated Relative to Actual SE

Error Corr.	σ_γ	σ_{β_2}	σ_{β_3}
-.95	.71	.53	.58
-.70	.73	.68	.74
-.60	.79	.69	.73
-.45	.84	.75	.78
-.20	.78	.78	.80
-.10	.75	.72	.77
0	.77	.82	.78
.10	.76	.78	.73
.20	.84	.81	.83
.45	.77	.79	.81
.60	.81	.84	.83
.70	.88	.89	.85
.95	.78	.82	.76

Note: These are the quotients of the estimated coefficient standard errors relative to the actual standard errors of the coefficients from 1000 trials of a Monte Carlo simulation of the 2SCML technique

Table 6: Two-stage LS Standard Errors
Estimated Relative to Actual SE

Error Corr.	σ_γ	σ_{β_2}	σ_{β_3}
-.95	.08	.09	.10
-.70	.15	.19	.19
-.60	.16	.20	.21
-.45	.17	.22	.23
-.20	.19	.24	.25
-.10	.19	.24	.26
0	.21	.27	.26
.10	.20	.25	.26
.20	.20	.25	.26
.45	.18	.25	.26
.60	.19	.24	.24
.70	.19	.23	.23
.95	.03	.03	.12

Note: These are the quotients of the estimated coefficient standard errors relative to the actual standard errors of the coefficients from 1000 trials of a Monte Carlo simulation of the 2PLS technique

Table 7: Two-Stage Uncertainty Results, 1976 Election
Carter and Ford Uncertainty

Independent Variables	Carter Uncertainty	Ford Uncertainty
Constant	5.5**	5.1**
	0.89	0.87
Education	-0.28**	-0.29**
	0.14	0.14
Political	-0.28**	-0.25**
Information	0.04	0.04
Gender	0.65**	0.78**
	0.24	0.24
Race	0.24	0.46
	0.49	0.49
Partisan	0.12	0.06
Strength	0.15	0.15
Media	-0.14**	-0.22**
Exposure	0.06	0.06
Political	-0.06**	-0.03*
Efficacy	0.03	0.03
First Debate	0.004	-0.25*
	0.17	0.17
Second Debate	-0.22*	-0.13
	0.16	0.16
Candidate	-0.17	-0.49*
Advertising	0.37	0.35
Candidate	-0.08**	0.02
Preference	0.03	0.03
Adjusted R^2	0.27	0.27
Model S.E.	2.6	2.7
Uncertainty mean	3.9	3.8
number of cases	464	464

Entries are two-stage least squares estimates, and their associated adjusted standard errors. * indicates a p=0.10 level of statistical significance, and ** a p=0.05 level, both one-tailed tests.

Table 8: Two-Stage Voting Models, 1976 Election
Probability of Carter Support

Independent Variables	2SCML Estimates	2SPLS Estimates
Constant	1.58**	1.71**
	0.91	0.91
Ford Issue	0.18**	0.18**
Distance	0.08	0.08
Carter Issue	-0.10*	-0.11*
Distance	0.07	0.07
Ford	0.66*	0.68*
Uncertainty	0.51	0.51
Carter	-0.75*	-0.77*
Uncertainty	0.55	0.55
Ford	-0.40**	-0.39**
Traits	0.05	0.05
Carter	0.33**	0.33**
Traits	0.06	0.06
Party	-0.29**	-0.29**
Identification	0.07	0.07
Education	-0.40**	-0.39**
	0.18	0.18
Gender	-0.35	-0.34
	0.31	0.30
Ford Error	-0.61	
	0.53	
Carter Error	0.64	
	0.56	
% Correct	95.3	95.7
χ^2	699.8	697.6
number of cases	464	464

Note: Entries are maximum-likelihood estimates, and their associated asymptotic standard errors. * indicates a p=0.10 level of statistical significance, and ** a p=0.05 level, both one-tailed tests. 2SCML is the River-Vuong conditional- maximum likelihood model; 2SPLS is the limited-information probit and least squares model. Candidate uncertainty in the 2SPLS model are instruments from a reduced form regression; the Ford and Carter error indicators are the error terms from the same regression.

Table 9: Reduced Form Probit Model, 1976 Election

Probability of Carter Vote	
Independent Variables	
Constant	0.04
	1.1
Education	-0.43
	0.16
Partisan	0.51
Strength	0.16
Gender	-0.4
	0.28
Race	2.4
	2.5
Political	-0.01
Information	0.05
Political	-0.009
Efficacy	0.03
Media	0.01
Exposure	0.06
Ford Ad	-0.57
	0.39
Carter Ad	0.05
	0.39
First	-0.07
Debate	0.19
Second	-0.14
Debate	0.18
Ford Issue	0.29
Distance	0.08
Carter Issue	-0.26
Distance	0.07
Ford Traits	-0.46
	0.06
Carter Traits	0.41
	0.06
% Correct	95.7
χ^2	697.8

Entries are probit maximum-likelihood estimates and their associated standard errors.

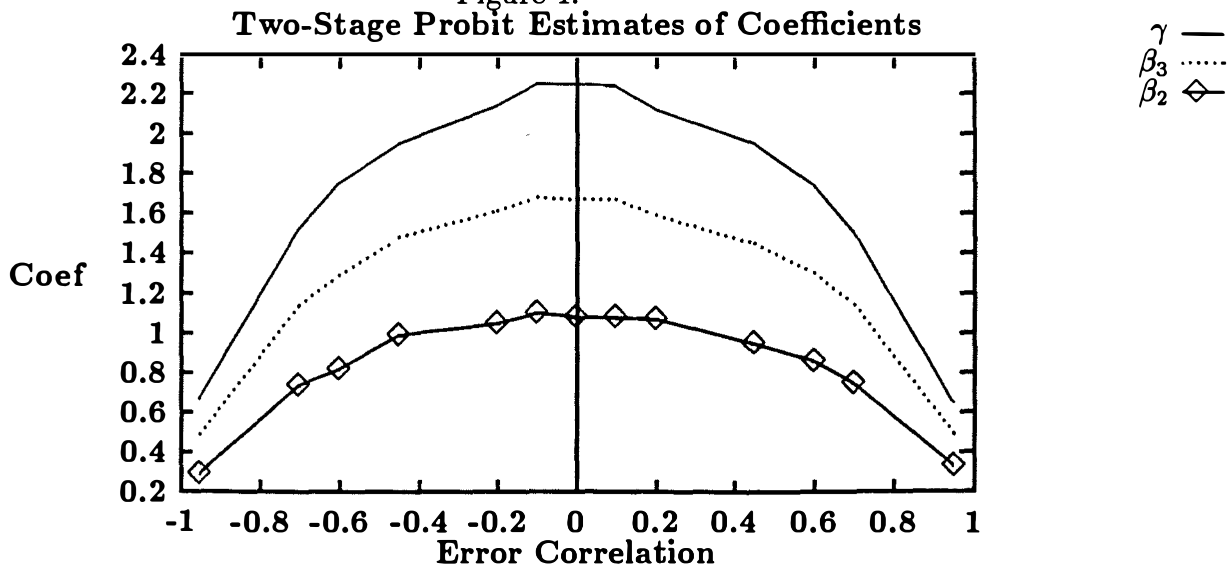
Table 10: Summary Statistics for Variables in 1976 Voting Models
Independent Variables

Variable	Mean	S. Dev.	Min.	Max
Ford Issue Distance	5.60	2.63	0.50	12.7
Carter Issue Distance	5.21	2.33	0.84	12.5
Ford Uncertainty	3.96	3.03	0.23	13.8
Carter Uncertainty	4.08	3.07	0.19	13.2
Ford Traits	12.0	5.46	3.51	24.6
Carter Traits	10.8	4.70	3.38	23.7
Party Identification	3.36	2.16	1.0	7.0
Education	2.62	0.87	1.0	4.0
Gender	0.54	0.50	0	1
Ford Error	-0.18	2.66	-7.45	8.53
Carter Error	-0.14	2.67	-7.19	8.06

Table 11: Reduced Form Models for 1976 Uncertainty Variables
Candidate Uncertainty

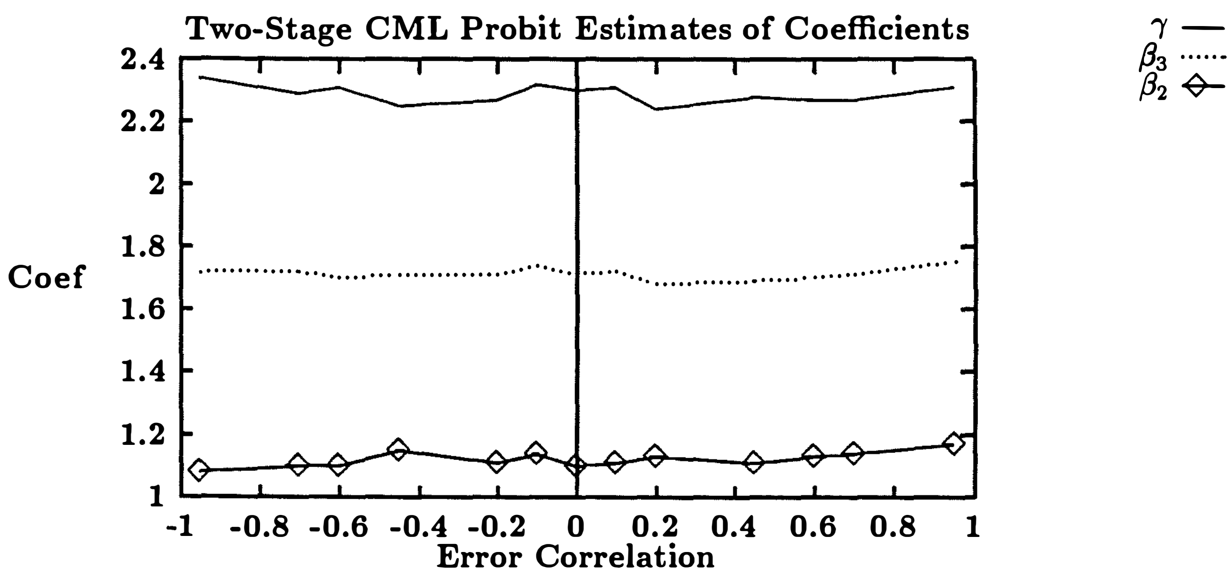
Independent Variables	Ford Uncertainty	Carter Uncertainty
Constant	5.02	5.34
	0.88	0.87
Education	-0.29	-0.32
	0.14	0.13
Party Strength	0.008	0.02
	0.13	0.12
Gender	0.72	0.56
	0.22	0.22
Political	-0.32	-0.33
Information	0.04	0.04
Media	-0.25	-0.17
Exposure	0.05	0.05
Political	-0.02	-0.04
Efficacy	0.02	0.02
Race	0.80	0.25
	0.44	0.44
Candidate	-0.50	-0.12
Advertisement	0.29	0.31
First	-0.17	0.02
Debate	0.16	0.15
Second	-0.14	-0.20
Debate	0.15	0.15
Ford Traits	-0.01	0.02
	0.02	0.02
Carter Traits	-0.01	-0.08
	0.03	0.02
Adjusted R^2	0.27	0.26
Model S.E.	2.77	2.75
Uncertainty Mean	4.17	4.14

Figure 1:



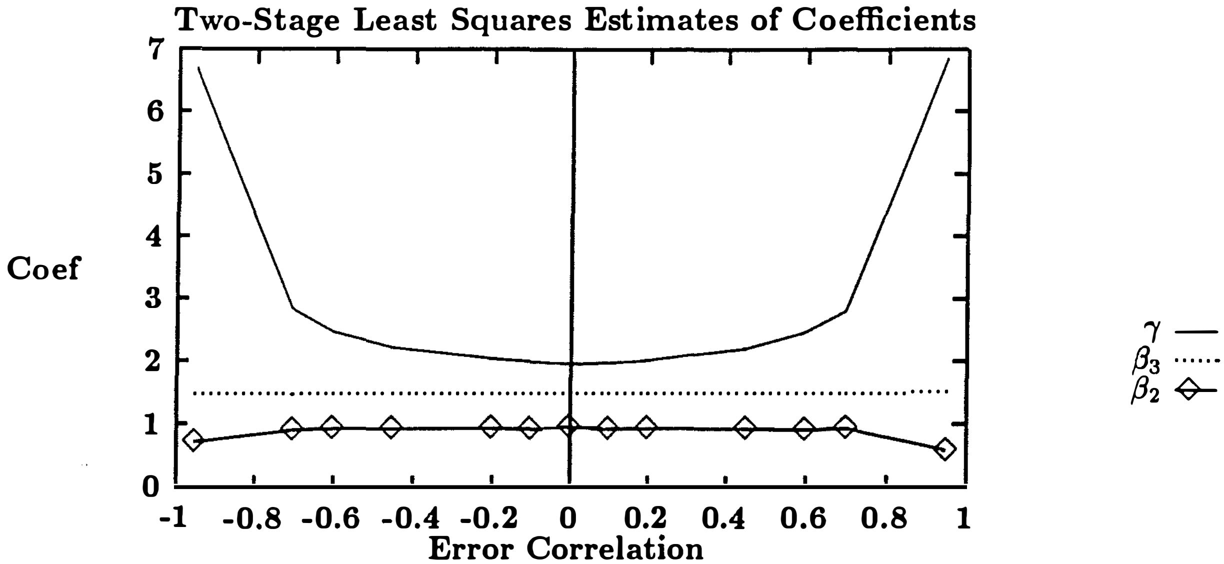
Note: These are the average estimated coefficients from 1000 trials of a Monte Carlo simulation of the 2SPLS. technique

Figure 2:



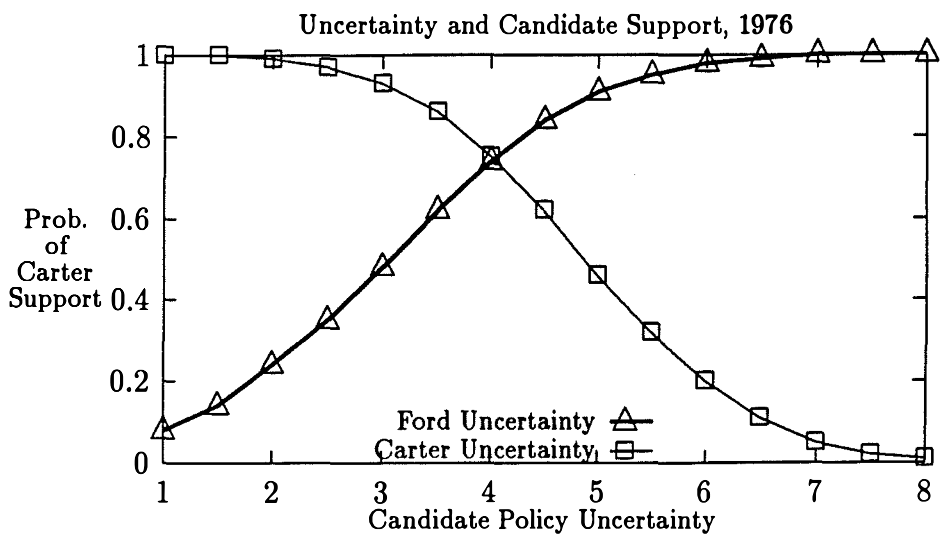
Note: These are the average estimated coefficients from 1000 trials of a Monte Carlo simulation of the 2SCML. technique

Figure 3:



Note: These are the average estimated coefficients from 1000 trials of a Monte Carlo simulation of the 2SLS tech-. nique

Figure 4:



Note: This gives the simulated probabilities of candidate support across the ranges of uncertainty about each candidate in the 1976 election. The coefficients these calculations are based upon are in Table 8.

7 References

Achen, C. H. *The Statistical Analysis of Quasi- Experiments.* Berkeley: University of California Press, 1986.

Aldrich, J. H., R. G. Niemi, G. Rabinowitz, and D. W. Rohde. "The Measurement of Public Opinion about Public Policy: A Report on Some New Issue Question Formats." *American Journal of Political Science* 1982(26): 391-414.

Aldrich, J. H., and F. D. Nelson. *Linear Probability, Logit, and Probit Models.* Beverly Hills: Sage Publications, Inc., 1984.

Alvarez, R. M. *Issues and Information in Presidential Elections.* Duke University, Ph.D. Dissertation, 1992.

Amemiya, T. "The Estimation of a Simultaneous Equation Generalized Probit Model." *Econometrica*, 1978(46): 1193- 1205.

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

Banks, J. S. and D. R. Kiewiet. "Explaining Patters of Candidate Competition in Congressional Elections." *American Journal of Political Science* 1989(33): 997-1015.

Bartels, L. M. "Issue Voting Under Uncertainty: An Empirical Test." *American Journal of Political Science*, 1986(30): 709-728.

Campbell, J. E. "Ambiguity in the Issue Positions of Presidential Candidates: A Causal Analysis." *American Journal of Political Science* 1983(27): 284-293.

Canon, D. T. *Actors, Athletes and Astronauts* Chicago: University of Chicago Press, 1990.

Downs, A. *An Economy Theory of Democracy.* New York: Harper and Row, 1957.

Dubin, J. A. and R. D. Rivers. *Statistical Software Tools*, version 2.0. Pasadena: Dubin-Rivers Research, 1989.

Enelow, J. M. and M. J. Hinich. *The Spatial Theory of Voting.* New York: Cambridge University Press, 1984.

Fiorina, M. P. *Retrospective Voting in American National Elections.* New Haven: Yale University Press, 1981.

Fishman, G. S. *Principles of Discrete Event Simulation.* New York: John Wiley and Sons, 1978.

Fiske, S. T. and M. A. Pavelchak. "Category-Based versus Piecemeal-Based Affective Responses: Developments in Schema-Triggered Affect." In R. M. Sorrentino and E. T. Higgins, *The Handbook of Motivation and Cognition.* New York: Guilford Press, 1986.

Franklin, C. H. "Eschewing Obfuscation? Campaigns and the Perceptions of U.S. Senate Incumbents." *American Political Science Review* 1991(85): 1193-1214.

Franklin, C. H. and J. E. Jackson. "The Dynamics of Party Identification." *American Political Science Review*, 1983(77): 957-973.

Graber, D. A. *Processing the News.* White Plains, New York: Longman, Inc., 1988.

Hanushek, E. A. and J. E. Jackson. *Statistical Methods for Social Scientists.* New York: Academic Press, 1977.

Hildebrand, D. K., J. D. Laing, H. Rosenthal. "Prediction Analysis in Political Research." *American Political Science Review*, 1976(70): 509-535.

Jacobson, G. C. and S. Kernell. *Strategy and Choice in Congressional Elections.* New Haven: Yale University Press, 1981.

Judge, G. G., R. C. Hill, W. E. Griffiths, H. Lutkepohl, and T.-C. Lee. *Introduction to the Theory and Practice of Econometrics*, second edition. New York: John Wiley and Sons, 1988.

King, G. *Unifying Political Methodology.* Cambridge: Cambridge University Press, 1989.

Lazarsfeld, P., B. Berelson, and H. Gaudet. *The People's Choice.* New York: Columbia University Press, 1944.

Maddala, G. S. *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge: Cambridge University Press, 1983.

Markus, G. B., and P. E. Converse. "A Dynamic Simultaneous Equation Model of Electoral Choice." *American Political Science Review*, 1979(73): 1055-1070.

McCullagh, P. and J. A. Nelder. *Generalized Linear Models*, second edition. London: Chapman and Hall, 1983.

Mitrani, I. *Simulation Techniques for Discrete Event Systems.* Cambridge: Cambridge University Press, 1982.

Page, B. I. and C. Jones. "Reciprocal Effects of Policy Preferences, Party Loyalties and the Vote." *American Political Science Review*, 1979(73): 1071-89.

Patterson, T. C. *The Mass Media Election.* New York: Praeger, 1980.

Rahn, W. M. "Perception and Evaluation of Political Candidates: A Social-Cognitive Perspective." Ph.D. Dissertation, University of Minnesota, 1990.

Rivers, D. and Q. H. Vuong. "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." *Journal of Econometrics*, 1988(39): 347-366.

Schlesinger, J. *Ambition and Politics.* Chicago: Rand McNally, 1966.

Verba, S. and N. H. Nie. *Participation in America* New York: Harper and Row, 1972.

Witcover, J. *Marathon* New York: Viking Press, 1977.

Wolfinger, R. E. and S. J. Rosenstone. *Who Votes?* New Haven: Yale University Press, 1980.