

# WINNOWING CHOICES: Political Choice Sets in Multi-Party Elections\*

Carole J. Wilson, *University of Texas at Dallas*  
Marco R. Steenbergen, *Universität Bern*

## Abstract

This paper reviews a class of choice models, known as choice set models, that have hitherto received scant attention in political science. Choice set models construe decision making as a two-stage process. In a first stage, the set of alternatives that are available to decision makers are pared down into a smaller set of alternatives that are attractive to a particular individual. The size and composition of this choice set does not have to be specified a priori, which offers great advantages over other sequential choice models such as McFadden's (1978) nested logit model. In a second stage, one alternative is selected from the pared-down choice set. We present one choice set model in great detail, namely Manski's (1977) model, and apply it to electoral data from the U.K. We conclude by discussing other potential applications of choice set models in political science.

---

\*Paper prepared for "Visions in Methodology: A Workshop for Women in Political Methodology".

## A Theory of Choice Sets

Scholars usually treat voting as a discrete choice process where voters are assumed to have certain levels of utility associated with each choice, and then to choose the candidate or party that maximizes that utility. In national elections, all voters are typically assumed to be presented with the same choice set because the alternatives listed on the ballot is the same for everyone. Research in consumer behavior should give us pause about such an assumption. Studies in this field have provided evidence in empirical and experimental studies to suggest that there are internal constraints on the choice sets decision makers face when choosing between alternatives. The electoral reality may be more complex than voters simply choosing from amongst all alternatives on the ballot—the choice set from which decisions makers actually choose may be smaller than the universal set of choices and may vary across voters. Research primarily in the field of marketing, but also in psychology and economics, has argued that choices are not necessarily based on evaluations of the universe of alternatives available to the decision maker; rather, the decision maker may choose from a subset of the universe (e.g. Hauser and Wernerfelt 1990; Nedungadi 1990; Punj and Brookes 2001; Roberts and Lattin 1991; Roberts and Nedungadi 1995; Roberts and Lattin 1997; Shocker *et al* 1991; Swait et al 2002). This smaller set is known as the *choice set*—the set of alternatives that reside in long-term memory that are “purposefully constructed and can be viewed as consisting of those goal-satisfying alternatives salient and *accessible* on a particular occasion. While an individual may have knowledge of a large number of alternatives, it is likely that only a few of these will ‘come to mind’ for a relevant use or purpose.” (Shocker, Ben-Akiva, Boccara, and Nedungadi 1991, p.183). The choice set is therefore the range of alternatives that a voter would reasonably investigate, evaluate, and potentially choose. Thus, the decision making process is a sequential one whereby choosers winnow alternatives in the universal set to a choice set from which they settle upon a single alternative.

Researchers in these fields have generally argued that modeling choice from choice sets is a more accurate representation of the decision making process than are models that assume decision-makers consider the universe of alternatives. These arguments are grounded in utility maximization theory (e.g. Hauser and Wernerfelt 1990), learning theory (e.g. Howard and Sheth 1969), and information processing theory (e.g. Alba and Chattopadhyay 1985; Shocker et al 1991). Empirical and experimental studies within the context of consumer brand choice have found strong evidence for the

theoretical construct of the choice set and its utility in improving the predictive power of choice models. Indeed, it has been shown that empirical models of choice are problematic if the choice set is misspecified.

Choice set theory in consumer behavior recognizes the notion that although consumers may have some utility for all brands, the decision to consume a particular brand is constrained at a given purchasing occasion. The composition of the choice set may be determined exogenously by limiting the alternatives presented to the decision maker (a consumer may consider only those labels of wine carried by the most convenient store), the set may be the result of a conscious heuristic used by the decision-maker to pare the number of alternatives to be evaluated (a consumer may choose to evaluate red wines, or only those that he has consumed in the past), or the set may be a function of the alternatives accessible and salient in memory during the evaluation processes (a consumer may choose only from those labels displayed at eye level-Roberts and Nedungadi 1995).

A logical extension of this consumer behavior model of decision making is to the political arena. The advantage of approaching voting behavior from the choice set perspective is that it allows us to incorporate a number of factors known to be important in political decision making into a single utility maximization model by recognizing the effects of memorial processes (e.g. Lodge and Stroh 1993; Wyer and Ottati 1993), attitude accessibility (e.g. Krosnick 1989; Wyer and Ottati 1993; Zaller 1992), and heuristics that voters as “cognitive misers” use to minimize information costs (e.g. Popkin 1991; Sniderman, Brody, and Tetlock 1991). Further, the sequential element of the choice set theory emphasizes the dynamic nature of political decisions. It seems quite a strong assumption that voters approach the voting booth (or even the pre-electoral survey) *de novo*—thoughtfully considering all alternatives. Rather, vote choice is more likely to have formed over time as voters winnow down the alternatives based on experience, information, campaigns, etc. (not dissimilar to Campbell et al’s (1960) classic “funnel of causality”).

The choice set theory implies that decisions are sequential processes whereby alternatives are winnowed to a subset of the universe of alternatives. These choice sets can, and do, vary across individuals. Models that do not explicitly take into account the heterogeneity of these choice sets can result in parameter estimates that are inconsistent and biased (Chaing et al 1999; Stopher 1980; Swait and Ben-Akiva 1986; and Williams and Ortuzar 1982). Unfortunately, the assumptions underlying the choice models typically used in political science are fundamentally at odds with the choice set theory.

## Current Approaches to Choice Set Modeling in Political Science

Over the past two decades, political scientists have drawn from a variety of choice models, the best known of which are the multinomial and conditional logit and multinomial probit models. Most of these models have been motivated in terms of random utility maximization (RUM), which has roots both in economic and psychological theory.<sup>1</sup> The key premises of the RUM are threefold: (1) decision makers are utility maximizers, (2) utility includes both a systematic and a random component, and (3) the systematic utility component includes both attributes of the alternatives and of the decision makers. The presence of a stochastic component in the utilities means that choice becomes probabilistic in nature. Thus, RUM models are sometimes also referred to as probabilistic choice models.

To formalize the RUM premises, assume that an individual  $i$  chooses among  $m$  alternatives and has an underlying utility  $U_{ij}$  for each alternative  $j$ . The utility can be expressed as

$$\begin{aligned} U_{ij} &= V_{ij} + \epsilon_{ij} \\ &= \mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_i\boldsymbol{\gamma}_j + \epsilon_{ij} \end{aligned} \tag{1}$$

where  $V_{ij}$  is the systematic utility component, which is modeled out in terms of a vector of attributes of the alternatives as perceived by the decision maker ( $\mathbf{x}'_{ij}$ ) and a vector of attributes of the decision maker ( $\mathbf{z}'_i$ ), and  $\epsilon_{ij}$  is a stochastic component or error. The effect of  $\mathbf{x}'_{ij}$  is assumed to be constant across the alternatives, while the effect of  $\mathbf{z}'_i$  is allowed to vary across alternatives. The utility maximization premise implies that alternative  $j$  is chosen if  $U_{ij} > U_{ik} \forall k \neq j$ . Thus, the probability of choosing  $j$  can be expressed as

$$\begin{aligned} \pi_{ij} &= \Pr(V_{ij} + \epsilon_{ij} > V_{ik} + \epsilon_{ik}) \\ &= \Pr(\epsilon_{ik} < \epsilon_{ij} + V_{ij} - V_{ik}) \\ &= \Pr[\epsilon_{ik} < \epsilon_{ij} + (\mathbf{x}'_{ij} - \mathbf{x}'_{ik})\boldsymbol{\beta} + \mathbf{z}'_i(\boldsymbol{\gamma}_j - \boldsymbol{\gamma}_k)] \end{aligned} \tag{2}$$

$\forall k \neq j$ .

Equation (2) provides the basis for a large number of different choice models. What distinguishes these models from one another is the specific

---

<sup>1</sup>RUM is not the only way in which choice models can be motivated. In biostatistics, dose response theory is the more common framework to motivate binary and multinomial response models.

error distribution that is being assumed. Apart from the specific density function that is chosen, choice models differ in whether they allow the errors to be correlated across the alternatives. In practice, the different distributional assumptions have led to at least three distinct traditions in choice modeling, which we refer to as Lucean, Thurstonian, and Tverskian choice models.

## Lucean Models

Lucean models assume that the errors are homoskedastic and uncorrelated across the alternatives. Luce (1959), reasoning (correctly in many cases) that the ratio of the probabilities of choosing alternatives  $j$  and  $k$  should not change with the introduction of a third alternative, developed a logit model in which the multinomial choice probabilities were given by

$$\pi_{ij} = \frac{\exp(V_{ij})}{\sum_{k=1}^m \exp(V_{ik})}$$

This particular specification of the choice probabilities can be obtained by assuming that the errors are independent and follow a Type-I extreme value distribution. McFadden (1973, 1978) extended the model by assuming that the  $V_{ik}$  are a linear function of the attributes of the alternatives,  $\mathbf{x}'_{ik}$ . This model is known as the conditional logit model. When  $V_{ik}$  is modeled as a linear function of the attributes of the decision makers,  $\mathbf{z}'_i$ , then the multinomial logit model is obtained. It is also possible to incorporate both attributes of the decision makers and the alternatives into the Lucean model.

Lucean models have a long history in political science. They have frequently been used to model vote choice in multi-party elections (e.g. Dow and Endersby 2004; Whitten and Palmer 1996), especially in the guise of the multinomial logit model. Undoubtedly part of their attraction has been that these models are comparatively easy to estimate and interpret.

This simplicity comes at a price, however, namely that we impose the assumption of *independence of irrelevant alternatives* or IIA. IIA implies that the ratio of the choice probabilities for two alternatives depends only on those alternatives. This assumption is quite rigid and may not do justice to the actual behavioral calculus that decision makers such as voters use. A famous illustration of violations of IIA in the econometric literature is the red bus, blue bus example. Imagine that commuters are offered the choice between a red bus, a blue bus, and a car as transportation means. They treat both color buses as the same and are indifferent between bus travel and car travel. In a choice between the blue bus and car only, the

Luce model would estimate the probability of choosing the blue bus as .50, resulting in a 1:1 probability ratio. The introduction of the red bus to the choice set should result in a .25 probability of choosing the blue bus (and a .25 probability of choosing the red bus). The probability ratio of choosing the car to the blue bus is now 2:1. The Luce model, however, must maintain the original 1:1 ratio, resulting in an over-estimation in the joint probability of choosing substitutes.

An additional and related shortcoming of Luce-type models is that they consider choice as a non-sequential process. This means that all of the utilities of all presented alternatives are compared at the same time. However, frequently choice is a winnowing process whereby, say a voter, narrows down the field of candidates to a smaller subset from which he or she then selects an alternative. This sequential process requires a different modeling approach.

## Thurstonian Models

An alternative specification of the error structure in the RUM can avoid the pitfalls of the IIA assumption. Thurstone (1927) described a “law of comparative judgment” in which an alternative  $j$  is perceived on the basis of the true utility,  $V_{ij}$  and a normally distributed error  $\epsilon_{ij}$ . The choice probability of binary choices is then given by  $\Phi(V_{ij} - V_{ik})$ , where  $j$  and  $k$  are two different alternatives. The extension of this idea to the polychotomous case is the multinomial probit model. Here the errors are jointly normally distributed rather than independently distributed and therefore are not saddled with the IIA assumption.<sup>2</sup>

In the multinomial probit model, the probability of choosing alternative  $j$  is given by

$$\pi_{ij} = \int_{-\infty}^{v_{ij1}} \int_{-\infty}^{v_{ij2}} \cdots \int_{-\infty}^{v_{ijm}} \phi_{m-1}(\eta_{i1j}, \eta_{i2j} \cdots \eta_{imj}) d\eta_{i1j} d\eta_{i2j} \cdots d\eta_{imj}$$

where  $v_{ijk} = v_{ij} - v_{ik} = (\mathbf{x}'_{ij} - \mathbf{x}'_{ik})\boldsymbol{\beta} + \mathbf{z}'_i(\boldsymbol{\gamma}_j - \boldsymbol{\gamma}_k)$ ,  $\eta_{ikj} = \epsilon_{ik} - \epsilon_{ij}$ , and  $\phi_{m-1}$  is the  $m - 1$ -variate normal density function with variance-covariance matrix  $\boldsymbol{\Sigma}_j$  defined over  $\boldsymbol{\eta}'_{ij} = (\eta_{i1j} \eta_{i2j} \cdots \eta_{imj})$ . For identification purposes, it is necessary to set  $\boldsymbol{\gamma}_k = \mathbf{0}$  for one of the alternatives, to fix the error variance to 1 for two of the alternatives, and the restrict the correlations to 0 for one of the alternatives (Keane 1992).

<sup>2</sup>Other efforts to avoid the IIA assumption include the DOGIT model (Gaudry and Dagenais 1979), mixed logit (see e.g. Glasgow 2001) and the nested choice models that will be considered in the next section.

While the multinomial probit model avoids the IIA assumption, its estimation is quite difficult because optimization of the log-likelihood function requires evaluation of a  $m - 1$ -variate integral. Considerable progress has been made in the estimation of these models using maximum simulated likelihood estimation (Geweke 1991; Hajivassiliou and McFadden 1998; Hajivassiliou, McFadden, and Ruud 1996; Keane 1992, 1994; Lerman and Manski 1981), the method of simulated moments (McFadden 1989), and Gibbs sampling (Dorfman 1996). Despite these developments, the multinomial probit model remains much more difficult to estimate and interpret than Lucean models.

Perhaps because of the greater computational complexity, there has been considerable debate in political science as to the need to estimate multinomial probit as opposed to logit models. Alvarez and his colleagues have been at the forefront of advocating the multinomial probit model, especially in studies of electoral choice in multi-party systems (Alvarez and Nagler 1995, 1998, 2001), and others have joined them (Schofield et al. 1998; Lacey and Burden 1999). But Quinn, Martin, and Whitford (1999) provide a more balanced assessment of the pros and cons of the multinomial probit model, and Dow and Endersby (2004) come down quite strongly on the side of multinomial logit, arguing that IIA is not frequently a problem and that the multinomial probit model is too susceptible to estimation problems.

Regardless of how one sides in the debate over the relative merits of multinomial probit versus multinomial logit, one should not overlook a major similarity between these two specifications, namely that choice is not considered sequential in either. That is, neither model is designed to address choice in multiple stages. We now turn to models that address the nature of sequential choice.

## **Tverskian Models**

Tverskian choice models explicitly assume a sequential or hierarchical decision structure. The most widely cited of these models is Tversky's (1979) elimination-by-aspects (EBA) model. The decision model assumes the basic random utility structure, but the process occurs across multiple decision stages whereby alternatives are eliminated based on undesired attributes. The pure EBA model has not been used in political science. In addition to the large number of parameters that the model requires and the lack of specialized software, it is difficult to incorporate continuous attributes in the analysis (see Manrai 1995; Rotondo 1986).

A more common approach to sequential choice is the nested logit model

of McFadden (1977, 1981). Using a generalized extreme value (GEV) distribution, the nested logit partitions choices into subsets of similar alternatives. The errors are correlated within but not across subsets. The probability of choosing alternative  $j$  from within subset  $C_l$  is given by

$$\pi_{ij} = \frac{\exp\left(\frac{V_{ij}}{\lambda_l}\right) \left[\sum_{k \in C_l} \exp\left(\frac{V_{ik}}{\lambda_l}\right)\right]^{\lambda_l - 1}}{\sum_s \left[\sum_{k \in C_s} \exp\left(\frac{V_{ik}}{\lambda_s}\right)\right]^{\lambda_s}}$$

Here  $\lambda \approx 1 - \rho$  is a measure of the correlation of the errors in a choice subset. If the subset consists of only one alternative,  $\rho = 0$  and  $\lambda = 1$ . For subsets with multiple alternatives,  $\lambda = 1$  indicates that the unobserved utility components of the alternatives are independent from one another.

Apart from accommodating sequential choice, the nested logit model also bypasses the IIA assumption, at least across subsets of alternatives. For two alternatives,  $j$  and  $k$ , in the same subset  $l$

$$\frac{\pi_{ij}}{\pi_{ik}} = \frac{e^{V_{ij}/\lambda_l}}{e^{V_{ik}/\lambda_l}}$$

Clearly, this quantity depends only on the two alternatives under consideration, thus reflecting the IIA property. For alternatives from different subsets,  $s$  and  $t$ , however, the ratio of probabilities depends on all of the alternatives in these subsets:

$$\frac{\pi_{ij}}{\pi_{ik}} = \frac{e^{V_{ij}/\lambda_s} \left(\sum_{l \in C_s} e^{V_{il}/\lambda_s}\right)^{\lambda_s - 1}}{e^{V_{ik}/\lambda_t} \left(\sum_{l \in C_t} e^{V_{il}/\lambda_t}\right)^{\lambda_t - 1}}$$

While the nested logit model has found limited application in political science (Born 1990; Kam 2006; Kamakura and Mazzon 2007), it has not penetrated the mainstream of the discipline like the multinomial logit model has. Among the limitations of the nested logit model is the requirement that one has to specify choice subsets ahead of time. This is frequently not possible or, in any case, there may be considerable variation in the size and composition of choice sets across decision makers.

## An Assessment

Of the models reviewed here, we believe that models of the Tverskian-type are particularly promising. In terms of the IIA assumption, they form a happy medium between the Lucean and Thurstonian models. Avoiding the

strict and often unrealistic IIA assumption of Luce-type models, Tverskian-type models avoid the accumulation of covariance parameters, and corresponding identification and estimation problems of Thurstonian models. Tverskian-type models also have the advantage of allowing for choice to be sequential. But herein lies also one of their main weaknesses, namely the need to specify the choice sequence a priori and to assume that it is uniform. What the discipline needs is a class of choice models that allow for choice to be sequential without researchers having to make an a priori guess of what that sequence is or having to assume that it is the same across decision makers. The choice set models to which we turn next satisfy this need.

## A Choice Set Model of Vote Choice

In this section we describe one version of a two-stage choice model that is promising for the development of vote models that are based on choice set theory. This model, which draws from Manski (1977), explicitly recognizes the electoral choice process as sequential, where voters develop a choice set in the first stage and choose from among the alternatives in this set in the second stage. The resulting model avoids the IIA assumption and allows for heterogeneous choice sets across voters. Moreover, it is relatively easy to estimate using maximum likelihood.<sup>3</sup>

### Derivation and Properties

The starting point for choice models that incorporate both a sequential decision making process and heterogeneous choice sets was Manski’s (1977) work on random utility models of choice, wherein he suggested that choice sets are probabilistic in nature and final choices are conditional on this choice set. He therefore argued for a model that captures both of these elements by representing the probability that  $i$  chooses  $j \in M$  as

$$\pi_{ij} = \sum_{C \in G} \pi_i(j|C)\pi_i(C) \quad (3)$$

Here  $M$  is the universal choice set, i.e. the choice set of all alternatives, and  $G$  is the set of all non-empty subsets ( $C$ ) of  $M$ , i.e.  $G$  is the powerset of

---

<sup>3</sup>Manski’s (1977) is not the only choice set model that is relevant for electoral behavior. For examples of other choice set models see e.g. van Nierop *et al.* (2000), Paap *et al.* (2005), and Rossi, Allenby, and McCulloch (2005). These models are more difficult to estimate and we intend to explore them in a future version of this paper.

$M$  excluding the null set. Further,  $\pi_i(C)$  is the probability that individual  $i$ 's choice set is  $C$  and  $\pi_i(j|C)$  is the probability that alternative  $j$  is chosen conditional on the choice set  $C$ .

Parameterizations of this model were first developed in transportation economics (e.g. Ben-Akiva and Boccara 1995; Basar and Bhat 2004) and applied in consumer behavior (e.g. Moe 2006). Following these, we let the inclusion of  $j$  in voter  $i$ 's choice set,  $C_i$ , be a linear function of a vector  $\mathbf{m}'_{ij}$ , which includes attributes of the alternative and/or the voter. If this function exceeds a latent threshold then  $j$  is included in  $C$ ; otherwise, it is not. Thus, we assume that the alternative needs to produce a minimum level of utility for  $i$  to be included in that voter's choice set. We further assume that the latent threshold follows a standard logistical distribution, so that the inclusion of  $j$  in  $C_i$  is

$$\omega_{ij} = \frac{1}{1 + \exp(-\mathbf{m}'_{ij}\boldsymbol{\gamma})} \quad (4)$$

where  $\boldsymbol{\gamma}$  is a vector of coefficients on  $\mathbf{m}'_{ij}$ .

Assuming that the thresholds are independent across alternatives, then the probability of choice set  $C_i$  is given by

$$\pi_i(C_i) = \frac{\prod_{j \in C_i} \omega_{ij} \prod_{j \notin C_i} (1 - \omega_{ij})}{1 - \prod_j (1 - \omega_{ij})} \quad (5)$$

where the denominator is a normalizing constant that eliminates the empty choice set. That is, consistent with Manski's (1977) formulation, the model assumes that the choice set contains at least one alternative. This is important for the next stage of the model, which pertains to the final vote choice of the voter.

In this stage, we use the Luce (1959) specification to model the probability of selecting  $j$  from  $C_i$ . Let  $\mathbf{x}'_{ij}$  denote a set of attributes of the alternative and/or voter that influence choice in the second stage. Then

$$\pi_i(j|C_i) = \begin{cases} \frac{\exp(\mathbf{x}'_{ij}\boldsymbol{\beta})}{\sum_{k \in C_i} \exp(\mathbf{x}'_{ik}\boldsymbol{\beta})} & \text{for } j \in C_i \\ 0 & \text{for } j \notin C_i \end{cases} \quad (6)$$

where  $\boldsymbol{\beta}$  is a vector of coefficients on  $\mathbf{x}'_{ij}$ . Note that  $\mathbf{x}'_{ij}$  and  $\mathbf{m}'_{ij}$  may contain different predictors.

The unconditional probability of choosing alternative  $j$  can be derived from (3). Note that  $G$  will include  $2^m - 1$  elements, where  $m$  is the number of elements in the universal choice set,  $M$ . Note that the estimation can get quite complex when  $m$  is large.

The Manski model, which is also known as the probabilistic choice multinomial logit (PCMNL) model, contains the conditional logit model as a special case. This case arises when  $\omega_{ij} = 1, \forall i, j$ . In this case, all alternatives are included in the choice set with probability 1, which is tantamount to saying that each voter is considering the universal choice set in his or her decision making process.

Other than this special case, however, the PCMNL model does not demonstrate the IIA property of the conditional logit model. This is demonstrated by Basar and Bhat (2004), who show that the cross-elasticities derived from the PCMNL formulation depend on all of the alternatives in the choice set, thus avoiding this restricting assumption. Specifically, let

$$B_j = \sum_{C \in G} \delta_j^C \pi(C) = \frac{\omega_j}{1 - \prod_k (1 - \omega_k)}$$

then the self- and cross-elasticities are given by

$$\begin{aligned} \eta_{z_{jk}}^{\pi_j} &= \frac{\partial \pi_j}{\partial z_{jk}} = \left[ (1 - B_j) \gamma_k + \frac{1}{\pi_j} \sum_{C \in G} \{ \pi(j|C) [1 - \pi(j|C)] \pi(C) \beta_k \} \right] z_{jk} \\ \eta_{z_{jk}}^{\pi_l} &= \frac{\partial \pi_l}{\partial z_{jk}} = \left\{ \frac{1}{\pi_l} \sum_{C \in G} \pi(l|C) \pi(C) \delta_{jl}^C - B_j \right\} \gamma_k + \\ &\quad \frac{1}{\pi_l} \sum_{C \in G} \{ -\pi(j|C) \pi(l|C) \pi(C) \beta_k \} z_{jk} \end{aligned}$$

respectively. Here  $z_{jk}$  is the  $k$ th attribute of the  $j$ th alternative, which may occur in  $\mathbf{m}'_j$ , in  $\mathbf{x}'_j$ , or in both. If  $z_{jk}$  occurs only in the first stage of the decision process, then all terms involving  $\beta_k$  will drop out of the expressions above. Similarly, if it only occurs in the second stage, then all terms involving  $\gamma_k$  will drop out. Note that the cross-elasticities are a function of  $\pi_j$  and of  $\pi(C)$  and, as such, entail all alternatives in the choice set.

## Estimation

Letting  $y_{ij}$  be an indicator variable that takes on the value 1 if alternative  $j$  is chosen and 0 otherwise, the log-likelihood function for the Manski model is given by

$$\begin{aligned} \ell &= y_{ij} \ln \pi_{ij} \\ &= y_{ij} [\ln \pi_i(j|C_i) + \ln \pi_i(C_i)] \end{aligned}$$

Note that the final choice stage is similar to the conditional logit model. If voter-specific attributes that do not vary across alternatives are included into the model, then it is customary to let the effect of those attributes vary across alternatives, designating one of these as the baseline category. For this category, the effect of the voter-specific attributes are constrained to zero, thus creating  $j - 1$  alternative specific constants. For the choice set stage, there are alternative specific constants for all alternatives in the universal choice set. This is a relatively straightforward log-likelihood function to optimize. For this paper, we performed the estimation in Stata.<sup>4</sup> We implement this method in the example below.

### **Example: Dimensions of Vote Choice in the United Kingdom**

The substantive question with which this analysis deals is the importance of party positions on the European integration dimension for inclusion in voters decision making in national elections. That is, to what extent do voters use this dimension in determining which parties they consider and ultimately vote for? Previous scholarship on this issue has been conflicted. The considerable research on the structure of attitudes toward EU membership has been quite successful in finding consistent European-wide determinants of opinions (e.g. Gabel and Palmer 1995, Gabel 1998a, Gabel 1998b, Carey 2002, McLaren 2004, Sánchez-Cuenca 2000, Kritzinger 2003). This has not been the case with the European integration issue in national voting models. Van der Eijk and Franklin (2004) refer to the European issue as a sleeping giant in European domestic politics—while the European integration issue has the potential to dramatically alter electoral politics in European countries, the issue has not been a consistently salient factor in national elections. Although recent studies of the United Kingdom by Evans (2002) and of Sweden, Finland, and Austria by Tillman (2004) find that attitudes toward European integration do influence vote choice in some national elections, there appears to be little agreement among scholars as to if and when the European issue affects domestic vote choice.

This analysis attempts to shed new light on the subject of the role of the issue of European integration in national elections by dividing the decision making process into two components—the choice-set stage and vote stage. Indeed, we illustrate that the importance of European integration is masked by modeling decision making in a single stage and in fact integration might be a more important factor in national elections than suggested by previous

---

<sup>4</sup>The Stata code is available upon request from the authors.

research.

In a simple model of political choice in the UK we explore the possibility that choice sets may be a function of two possible decision heuristics. Voters may use a proximity heuristic whereby they first choose a choice set of parties that are nearby on the Left-Right and/or the EU dimension. Alternatively, voters may use a directional heuristic and choose use a a set of parties that are on his or her side of the Left-Right and/or EU dimension. These hypotheses represent two competing theories of vote choice—Down’s (1957) spatial model and the Rabinowitz and MacDonald (1989) directional model, respectively. At the vote choice stage, voters may choose the closest party in the choice set on either or both dimensions. Or, extending the logic of the directional model, they may choose the furthest party on the same side of the dimension within the choice set.

We estimate the PCMNL model as described above using data from the 1999 European Election Study.<sup>5</sup> Proximity on the Left-Right and EU dimensions is measured as the distance of the voter’s self-placement from the voter’s party placement (both on 1-10 scales). Voter and parties are considered to be on the same side as a party on a given dimension if self placement and party placement are on the same side of the midpoint. The furthest party on the same side of the dimension is the party that has the largest distance from the voter and is on the same side of the midpoint as the voter.

In this model we examine only the three major parties in the United Kingdom. While it would be nice to include all of the parties in the analysis, especially given the importance we place on the defining the choice set, there are not enough people in the sample who voted for the minor parties to estimate their effects on choice.<sup>6</sup> The estimates of the PCMNL model and comparison conditional logit are presented in Table 1.

---

<sup>5</sup>The data utilized in this analysis were originally collected for the European Election Study Workgroup, consisting of Cees van der Eijk, Klaus Schoenbach, Hermann Schmitt, Holli Semetko, Wouter van der Brug, Mark Franklin, Sren Holmberg, Renato Mannheimer, Jacques Thomassen and Berhard Wessels. Fieldwork was carried out by a consortium of European survey organizations, co-ordinated by IPSOS. This study has been made possible with grants from the University of Amsterdam, the Dutch National Science Foundation, The Bundespresseamt, the CIS, the University of Mannheim, the ISPO Institute and Trinity College. Neither the original collectors of the data nor their sponsors bear any responsibility for the analyses or interpretations published here. The data are distributed by Steinmetz Archive, Amsterdam, the Netherlands, and associated data-archives.

<sup>6</sup>To minimize the effect of these minor parties on the estimation of the choice set, we eliminate all subjects from the analysis that, when asked the likelihood of voting for one of the minor parties on a scale of 1-10, indicated 3 or higher.

Table 1: Effects of Left/Right and EU Dimensions on Vote Choice in the UK

Vote Stage Estimates		
	PCMNL	Cond. Logit
Left-Right Distance	-0.665* (0.293)	-0.411* (0.176)
EU Distance	-0.647* (0.235)	-0.288 (0.184)
Same Side L-R	-- --	0.968* (0.301)
Same Side EU	-- --	-0.487* (0.295)
Furthest Party on Same Side L-R	-0.554 (0.869)	-0.424 (0.438)
Furthest Party on Same Side EU	1.059 (0.665)	0.421 (0.399)
Choice-Set Stage		
Left-Right Distance	-0.742* (0.168)	
EU Distance	0.334* (0.225)	
Same Side EU	-1.866* (0.983)	
Same Side L-R	3.135* (0.719)	
Constant-Conservative Party	1.848* (0.036)	
Constant-Labour Party	3.969* (2.201)	
Constant-Lib/Dem Party	-2.931* (0.789)	
<i>N</i>	104	104

Note: Standard errors in parentheses. One tailed tests. \*p<0.05

The results of the PCMNL model indicate that voters select a choice set that includes parties nearby on the Left-Right dimension and on the same side. The parties that comprise the choice set are distant from the voter on the EU dimension and tend to be on the opposite side of the dimension. However, in the final vote stage, the EU distance variable is significant and in the expected direction, as is Left-Right distance, indicating that voters choose the party that minimizes the distance on both dimensions from among the choice set. In the conditional logit model, only Left-Right distance is significant and in the expected direction. In contrast, the two-stage model shows that the EU dimension is an important factor in national elections, but only in choosing from among parties that are acceptable parties on the Left-Right dimension. In terms of predictive ability, the PCMNL model predicts 71 percent of cases correctly compared to less than 54 percent for the conditional logit. A likely explanation for this is the ability of the PCMNL model to eliminate less preferred options in the choice set stage. In Table 2 we present the percentages of the most likely choice sets predicted by the PCMNL model.

Table 2: Predicted Choice Sets of UK Voters

Choice set	Percent
Conservative	0.96
Labour	23.08
Lib/Dem	0.00
Conservative—Labour	64.42
Conservative—Lib/Dem	0.00
Labor—Lib/Dem	3.85
All Parties	7.69

Note: Percent of predicted choice sets from PCMNL in sample.

Clearly, the universal set of alternatives is not the most likely choice set. More than 64 percent of the sample are predicted to have chosen between the Conservative and the Labour parties while nearly a quarter are captive to the Labour Party. Only about eight percent are predicted to have chosen from among all three parties.

As this example illustrates, the PCML model enables us to directly model the choice set from which voters choose. In addition to making more realistic assumptions about voters' decision making process this model has greater predictive power and can provide additional insight into the factors which affect vote choice. While the PCML model offers the advantages of allowing for sequential decision making, avoiding the often unrealistic IIA

assumption, and allowing for heterogeneous choice sets, there are some disadvantages. The most problematic aspect of the PCML model is that the probability of all possible choice sets must be estimated. While this is easy with only the three parties in the example explored here (seven possible choice sets), the possible number of choice sets increases dramatically such that vote choice in large multiparty systems may be impossible to estimate using this technique. We therefore intend to explore Bayesian approaches to choice set estimation in future work.

## Future Directions in Choice Set Models

This study brings a new theory of decision making to political science from the field of consumer behavior. This is a useful innovation for the study of voting behavior, because the use of choice sets is a more realistic model of decision making than current models that assume voters choose from the universal set of alternatives. Moreover, failure to incorporate the notion of heterogeneous consideration sets into empirical and theoretical models of voting has can lead to wrong conclusions about the decision making process.

While choice sets offer new insight into the decision making process of voters, the basis of this approach is grounded in two well-accepted theories of political behavior—utility maximization and information processing. Thus, while an innovation, the consideration set builds on existing theories to create more realistic assumptions about voters, and a more nuanced understanding of the choice process. As this study shows, incorporating consideration sets in vote choice models can lead to interesting substantive conclusions.

As with any new theoretical development, the study of consideration sets creates avenues for future research. One line of inquiry involves the composition of consideration sets. For example, the average number of parties contained in consideration sets in large party systems, the variation in size of consideration sets across levels of political sophistication, and the elements in consideration sets (similar or dissimilar parties) are all open questions. Another type of inquiry involves the effect of priming and cues on the consideration set. Still another avenue for future research is the implication of consideration sets for party positioning and campaign strategy. These are but a few substantive questions, the answers to which may provide us with a more complete understanding of political decision making and its effects on party systems.

## Works Cited

- Alba, Joseph W., and Amitava Chattopadhyay. 1985. "Effects of Context and Post-Category Cues on Recall of Competing Brands." *Journal of Marketing Research*. 22(August):340-349.
- Alvarez, R. Michael, and Jonathan Nagler. 1995. "Economics, Issues and the Perot Candidacy: Voter Choice in the 1992 Presidential Election." *American Journal of Political Science*. 39:714-44.
- Alvarez, R. Michael, and Jonathan Nagler. 1998. "When Politics and Models Collide: Estimating Models of Multiparty Competition." *American Journal of Political Science*. 42:55-96.
- Alvarez, R. Michael and Jonathan Nagler. 2000. "A New Approach for Modelling Strategic Voting in Multiparty Elections." *British Journal of Political Science*. 30:57-75.
- Başar, Gözen, and Chandra Bhat. 2004. "A Parameterized Consideration Set Model for Airport Choice: An Application to the San Francisco Bay Area." *Transportation Research Part B*. 38(December):889-904.
- Born, Richard. 1990. "Surge and Decline: Negative Voting: and the Midterm Loss Phenomenon: A Simultaneous Choice Analysis." *American Journal of Political Science*. 34:615-645.
- Campbell, Angus and Philip E. Converse and Warren E. Miller and Donald E. Stokes. 1960. *The American Voter*. New York: Wiley & Sons.
- Carey, Sean. 2002. "Undivided Loyalties: Is National Identity an Obstacle to European Integration?" *European Union Politics*. 3(4): 387-413.
- Chaing, Jeongwen, Siddhartha Chib, and Chakravarthi Narasimhan. 1999. "Markov Chain Monte Carlo and Models of Consideration Set and Parameter Heterogeneity." *Journal of Econometrics*. 89:233-248.
- Dorfman, Jeffrey H. 1996. "Modeling Multiple Adoption Decisions in a Joint Framework." *American Journal of Agricultural Economics*. 78(August):547-557.
- Dow, Jay K. and James W. Endersby. 2004. "Multinomial Probit and Multinomial Logit: A Comparison of Choice Models for Voting Research." *Electoral Studies*. 23:107-122.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper & Row.
- Eijk, Cees van der and Mark Franklin. 2004. "Potential for Contestation on European Matters at National Elections in Europe." In Gary Marks and Marco Steenbergen (eds.) *European Integration and Political Conflict*. Cambridge: Cambridge University Press, 33-50.
- Evans, Geoffrey. 2002. "European Integration, Party Politics and Voting in

- the 2001 Election.” *British Elections and Parties Review*. 12:95-110.
- Gabel, Matthew. 1998a. *Interests and Integration: Market Liberalization, Public Opinion and European Union*. Ann Arbor, MI: University of Michigan Press.
- Gabel, Matthew. 1998b. “Economic Integration and Mass Politics: Market Liberalization and Public Attitudes in the European Union.” *American Journal of Political Science*. 42(3): 936-53.
- Gabel, Matthew C. and Harvey Palmer. 1995. “Understanding Variation in Public Support for European Integration.” *European Journal of Political Research*. 27(1): 3-19.
- Gaudry, Marc J. I. and Marcel G. Dagenais. 1979. “The Dogit Model.” *Transportation Research B*. 13B(June):105-111.
- Glasgow, Garrett. 2001. “Mixed Logit Models for Multiparty Elections.” *Political Analysis*. 9(2):116-136.
- Geweke, John. 1987. “Bayesian Inference in Econometric Models Using Monte-Carlo Integration.” *Econometrica* 57:13 1740.
- Geweke, John. 1991. “Efficient Simulation from the Multivariate Normal and Student-t distributions Subject to Linear Constraints.” Computer Science and Statistics: Proceedings of the Twenty-Third Symposium on the Interface. Alexandria, VA: American Statistical Association, 571-8.
- Hajivassiliou, Vassilis and Daniel McFadden. 1998. “The Method of Simulated Scores with Application to Models of External Debt Crises.” *Econometrica*. 66:273-286
- Hajivassiliou, Vassilis, Daniel McFadden, and Paul Ruud. 1996. “Simulation of multivariate normal rectangle probabilities and their derivatives: Theoretical and computational results.” *Journal of Econometrics*. 72:85-134.
- Hauser, John R., and Birger Wernerfelt. 1990. “An Evaluation Cost Model of Consideration Sets.” *Journal of Consumer Research*. 16(4):393-408.
- Howard, John A. and Jagdish N. Sheth. 1969. *The Theory of Buyer Behavior*. New York: John Wiley & Sons, Inc.
- Kam, Cindy. 2006. “Political Campaigns and Open-Minded Thinking.” *Journal of Politics*. 68(4):931-945.
- Kamakura, Wagner A. and José Afonso Mazzon. 2007. “Accounting for Voter Heterogeneity within and across Districts with a Factor-Analytic Voter-Choice Model.” *Political Analysis*. 15(1):67-84.
- Keane, M. P. 1992. “A note on identification in the multinomial probit model.” *Journal of Business and Economic Statistics* 10:193200.
- Kriztzing, Sylvia. 2003. “The Influence of Nation-State on Individual Support for the European Union.” *European Union Politics*. 4(2): 219-

- Krosnick, Jon A. 1989. "Attitude Importance and Attitude Accessibility." *Personality and Social Psychology Bulletin*. 15:297-308
- Lacy, Dean, and Barry C. Burden. 1999. "The Vote Stealing and Turnout Effects of Ross Perot in the 1992 U.S. Presidential Election." *American Journal of Political Science* 43(1, January):233-55.
- Lerman, S. and C. Manski. 1981. "On the Use of Simulated Frequencies to Approximate Choice Probabilities". In Charles F. Manski and Daniel McFadden, (eds.) *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: MIT Press.
- Lodge, Milton and Patrick Stroh. 1993. "Inside the Mental Voting Booth: An Impression-Driven Process Model of Candidate Evaluation." In *Explorations in Political Psychology*. Ed. Shanto Iyengar and William J. McGuire. Durham: Duke University Press.
- Luce, R. Duncan. 1959. *Individual Choice Behavior*. New York: John Wiley & Sons, Inc.
- Manrai, Ajay K. 1995. "Mathematical Models of Brand Choice Behavior." *European Journal of Operational Research*. 82:1-17.
- Manski, Charles F. 1977. "The Structure of Random Utility Models." *Theory and Decision*. 8:229-254.
- McFadden, Daniel. 1978. "Modelling the Choice of Residential Location." In A. Karlqvist, L. Lundqvist, S. Snikars, and J. Weibull (eds.), *Spatial Interaction Theory and Planning*. Amsterdam: North Holland. p. 75-96.
- McFadden, Daniel. 1973. "Conditional logit analysis of qualitative choice behavior." In P. Zarembka, (ed.) *Frontiers in Econometrics*. New York: Academic Press.
- McFadden, Daniel. 1998. "A Method of Simulated Moments for Estimation of Discrete Choice Models without Numerical Integration." *Econometrica*. 57:995-1026.
- McFadden, Daniel. 1981. "Economic Models of Probabilistic Choice." In Charles F. Manski and Daniel McFadden, (eds.) *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: MIT Press. p.198-272.
- McLaren, Lauren. 2004. "Opposition to European integration and fear of loss of national identity: Debunking a basic assumption regarding hostility to the integration project." *European Journal of Political Research*. 43: 895-911.
- Moe, Wendy (2006), "An Empirical Two-Stage Choice Model with Decision Rules Applied to Internet Clickstream Data." *Journal of Marketing*

- Research*. 43(4):680-692.
- Nedungadi, Prakash. 1990. "Recall and Consumer Consideration Sets: Influencing Choice without Altering Brand Evaluations." *Journal of Consumer Research*. 17(3):263-276.
- Paap, Richard, Erjen van Nierop, Harald J. van Heerde, Michel Wedel, Philip Hans Franses, Karel Jan Alsem. 2005. "Consideration Sets, Intentions and the Inclusion of 'don't know' in a Two-Stage Model for Voter Choice". *International Journal of Forecasting*. 21:53-71.
- Popkin, Samuel L. 1991. *The Reasoning Voter*. Chicago: University of Chicago Press.
- Punj, Girish, and Richard Brookes. 2001. "Decision Constraints and Consideration-Set Formation in Consumer Durables." *Psychology and Marketing*. 18(8):843-863.
- Quinn, Kevin, Andrew Martin and Andrew Whitford. 1999. "Voter Choice in Multiparty Democracies: A Test of Competing Theories and Models." *American Journal of Political Science*. 43(4):1231.
- Rabinowitz, George, and Stuart E. MacDonald. 1989. "A Directional Theory of Issue Voting." *American Political Science Review*. 83(1):93-121.
- Roberts, John H., and James M. Lattin. 1991. "Development and Testing of a Model of Consideration Set Composition." *Journal of Marketing Research*. 28(4):429-440.
- Roberts, John H., and James M. Lattin. 1997. "Consideration: Review of Research and Prospects for Future Insights." *Journal of Marketing Research*. 34(3):406-410.
- Roberts, John H., and Prakash Nedungadi. 1995. "Studying Consideration in the Consumer Decision Process: Progress and Challenges." *International Journal of Research in Marketing*. 12:3-7.
- Rossi, Peter E., Greg M. Allenby and Robert McCulloch. 2005. *Bayesian Statistics and Marketing*. West Sussex: John Wiley & Sons.
- Rotondo, John. 1986. "Price as an Aspect of Choice in EBA." *Marketing Science*. 5(4):391-402.
- Sánchez-Cuenca. 2000. "The Political Basis of Support for European Integration." *European Union Politics*. 1(2): 147-171.
- Schofield, Norman, Andrew Martin, Kevin Quinn, and A. Whitford. 1998 "Multiparty Electoral Competition in the Netherlands and Germany: A Model Based on Multinomial Probit", *Public Choice*. 97(2):247-93
- Shocker, Allan D., Moshe Ben-Akiva, Bruno Boccara, and Prakash Nedungadi. 1990. "Consideration Set Influences on Consumer Decision-Making and Choice: Issues, Models, and Suggestions." *Marketing Letters*. 2(3):181-197.

- Sniderman, Paul M., Richard A. Brody, and Philip E. Tetlock. 1991. *Reasoning and Choice: Explorations in Political Psychology*. New York: Cambridge University Press.
- Stopher, P. 1980. "Captivity and Choice in Travel-Behavior Models." *Transportation Journal of ASCE*. 69:99-118.
- Swait, Joffre, and Moshe Ben-Akiva. 1986. "Incorporating Random Coefficients in Discrete Models of Choice Set Generation." *Transportation Research, B*. 21:92-102.
- Swait, Joffre, Wiktor Adamowicz, Michael Handemann, Adele Diederich, John Krosnick, David Layton, William Provencher, David Schkade, and Roger Tourangeau. 2002. "Context Dependence and Aggregation in Disaggregate Choice Analysis." *Marketing Letters*. 13(3):195-205.
- Tillman, Erik R. 2004. "The European Union at the Ballot Box? European Integration and Voting Behavior in the New Member States." *Comparative Political Studies*. 37(5):590-610.
- Thurstone, L. 1927. "A Law of Comparative Judgement." *Psychological Review*. 34:273-86.
- Tversky, A. 1972. "Elimination by Aspects: A Theory of Choice." *Psychological Review*. 79(4):281-299.
- van Nierop, Erjen, Richard Paap, Bart J. Bronnenberg, Philip H. Franses and Michel Wedel. 2000. "Modeling Unobserved Consideration Sets for Household Panel Data." *ERIM Report Series Research in Management ERS-2000-42-MKT*. Erasmus University Rotterdam.
- 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(1):231-260.
- Williams, H. and J.D. Ortuzar. 1982. "Behavior Theories of Dispersion and the Mis-specification of Travel Demand Models." *Transportation Research, B*. 16(June):167-219.
- Wyer, Robert S., and Victor C. Ottati. 1993. "Political Information Processing." In *Explorations in Political Psychology*, ed. Shanto Iyengar and William J. McGuire. Durham: Duke University Press.
- Zaller, John. R. 1992. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.