Title
Calibration and Test of a Discrete Choice Model with Endogenous Choice Sets

Journal
Geographical Analysis, 30

Author
Dai, Jian

Publication Date
1998

Peer reviewed

Jian Dai

Calibration and Test of a Discrete Choice Model with Endogenous Choice Sets

Conventional discrete choice models assume implicitly that the choice set is independent of the decisionmaker's preferences conditional on the explanatory variables of the models. This assumption is implausible in many choice situations where the decisionmaker selects his or her choice set. This paper estimates and tests a discrete choice model with endogenous choice sets based on Horowitz' theoretical work. To calibrate the model, a new probability simulator is introduced and a sequential estimation procedure is developed. The model and calibration methods are tested in an empirical application as well as Monte Carlo simulations. The empirical results are used to test the theory of endogenous choice sets and to examine the differences between the new model and a conventional choice model in parameter estimates and predicted choice probabilities. The empirical results strongly suggest that ignoring the endogeneity of choice sets in choice modeling can have serious consequences in applications.

1. INTRODUCTION

Explaining the outcomes of choices by individuals among sets of discrete alternatives has been a research objective of many fields including economics, geography, marketing research, psychology, and transportation systems analysis. In geography spatial choice modeling has been studied extensively [see Wrigley (1982, 1988); Golledge and Rushton (1984); Pitfield (1984); Wrigley and Longley (1984); Horowitz (1985); Golledge and Stimson (1987); Golledge and Timmermans (1988); Fischer, Nijkamp, and Papageorgiou (1990); Timmermans and Golledge (1990); and Thill (1992) for recent reviews and books. See also Rushton (1969a, 1969b, 1969c, 1971); Golledge and Rushton (1976) for early work]. In applications discrete choice random-utility models have been widely used to analyze choice behavior and to predict choices. These models

This research was in part supported by a scholarship from the Midwest Transportation Center at the University of Iowa and the Iowa State University. The author is deeply grateful to Joel L. Horowitz for insightful suggestions and many discussions. The author is also grateful to Marc Armstrong, Clair Pavlic, Gerard Rushton, Eugene Savin, and Frank Weirich for their support during the research.

Jian Dai is a researcher in the Graduate School of Management and the Center for Statistics in Science and Technology at the University of California at Davis.

Geographical Analysis, Vol. 30, No. 2 (April 1998) © 1998 Ohio State University Press
Submitted: 6/13/97. Revised version accepted: 9/23/97.

assume that an individual's choice is an expression of his or her preferences, that the individual's preferences among available alternatives can be described by a utility function that depends on the attributes of alternatives and the individual, and that the individual selects the alternative with the highest utility. The utility of an alternative is represented as the sum of two components: a deterministic component that accounts for systematic effects of observed attributes on choice, and a random component that accounts for the effects of unobserved variables that influence choice. The random-utility models then predict the probability that a randomly selected individual will choose a particular alternative, given the values of the observed variables and a joint probability distribution of the random variables (for example, Manski and McFadden 1981; Ben-Akiva and Lerman 1985; Ortuzar and Willumsen 1994).

In discrete choice modeling it is assumed that the decisionmaker makes a choice from the choice set which is usually a subset of the set of all alternatives. Many questions can be asked about the choice set, such as these: Where does the choice set come from? Is it known? Is it fixed? How is it selected or how might it be determined? How might it be modeled given the choice set generation process? Furthermore, how might the choice be modeled given the choice process? Choice set modeling is an important issue in discrete choice analysis because correct estimation of a choice model and correct prediction of choice probabilities are conditional on correct information about the choice set (Manski 1977; Williams and Ortuzar 1982). It is also a complex issue involving behavior theories, model structures, computational methods, and data availability. Researchers are increasingly concerned with the choice set issue and have been developing approaches to solving some of the problems.

In this paper we are concerned with a particular problem, namely, modeling choice and the choice set when both are the outcomes of a choice process in which the individual selects the choice set based on his or her preferences. The work is based on an approach developed by Horowitz (1991) to modeling choice with endogenous choice sets. This paper has two objectives. The first objective is to implement and estimate the choice model with endogenous choice sets proposed by Horowitz. Choice models involving endogenous choice sets are more complex than conventional choice models. This paper introduces new methods to compute choice probabilities and to estimate the parameters of the model. The second objective is to test the model. The model and estimation methods are first tested using simulation data. The model is then tested in an empirical application. The empirical test is carried out in a particular choice context, namely, college choice by high school seniors. College choice is an example of choice that depends on both spatial and nonspatial variables (for example, locational college attributes and academic quality) and, as will be explained later, it has characteristics that make it an ideal context for the purpose. Using the empirical results, the theory of endogenous choice sets is tested. Furthermore, the results of the new model are compared to those of a conventional choice model and the consequences of ignoring the dependence between choice set selection and final choice are explored. In this paper we consider the case where the choice set can be observed.

The rest of the paper is organized as follows. Section 2 provides a brief review of approaches to choice set modeling. Section 3 defines the choice context and presents the choice model with endogenous choice sets. Section 4 discusses the computational issues and calibration methods. Section 5 tests the models and calibration methods in Monte Carlo simulations. Section 6 presents the application and results of empirical testing. Section 7 gives concluding comments.

2. APPROACHES TO CHOICE SET MODELING

In most applications of random-utility models, the choice set of each individual is specified by the analyst a priori and is assumed to be fixed. The simplest approach to choice set specification consists of assuming that the choice set includes all conceivable alternatives and is the same for all individuals in the population. These assumptions are, however, often tenuous. For example, a commuter whose residence and work locations are not served by transit does not have the option to go to work by transit. An alternative approach is to use deterministic rules to decide the availability (or unavailability) of a particular alternative to the individual based on the knowledge and judgment of the analyst (for example, Ben-Akiva and Lerman 1974; Adler and Ben-Akiva 1976; Gautschi 1981). For example, in destination choice modeling, the individuals living in different geographic areas are often assumed to have different choice sets for the reason that they might be subject to different spatial constraints (for example, Southworth 1981; Miller and O'Kelly 1983). The specification of choice sets based on deterministic rules is simple to carry out but can result in incorrect choice sets. The prespecified choice set may include an alternative whose choice probability is zero or it may exclude an alternative whose choice probability is nonzero.

One possible solution to the problem of misspecifying choice sets is to model the choice set for each individual. In the probabilistic models of choice set generation, the choice set is considered to be random with a probability distribution that can be estimated from the data. Although the choice set may be fixed to the individual, it is random for the modeling purpose because the analyst does not have perfect information about the choice set generation process and therefore cannot predict it with certainty. A discrete choice model with random choice sets can be expressed as (Manski 1977)

$$P(i) = \sum_{C \subseteq G} P(i|C)^{*}P(C|G)$$

where G is the universal set of alternatives, $P(C|G)$ is the probability that an individual's choice set is $C, P(i|C)$ is the probability that alternative i is chosen conditional on choice set C. Similar ideas of decomposing a choice model into two submodels, a choice set submodel and a choice submodel, are also presented by Burnett (1980) and Burnett and Hanson (1979, 1980).

A high degree of complexity is implied in the choice models with choice set generation. The set of choice sets that can be formed from G is the power set of G, and it contains $(2^{m(G)} - 1)$ elements, where $m(G)$ is the number of alternatives in G. For example, 1,023 choice sets can be formed from a universal set with only ten alternatives. Thus, the number of choice sets is intractably large when G is large. To make a choice set model useful in applications, it is necessary to place a priori restrictions on the possible choice sets. Usually, $P(C|G)$ is restricted to a parametric family of distributions. For example, the dogit model (Gaudry and Dagenais 1979) assumes that an individual is either captive to an alternative or is free to choose from the full choice set. Random-constraints models (Swait and Ben-Akiva 1987a, 1987b) use probabilistic constraints to derive a family of parametric choice set models. See also Boccara (1989) and Thill and Horowitz (1997).

A search model is proposed by Richardson (1982) in which the choice set is the result of a sequential search and is not fully known to the individual until the choice has been made. Meyer (1980) develops a theoretical model of choice set generation that incorporates learning in the context of destination choice. Fotheringham (1983, 1988a, 1988b) suggests a competing-destination model in which the individual first selects a cluster of similar alternatives and then chooses from within this cluster. See also Fotheringham and Trew (1993) and Haynes and Fotheringham (1991). Miron and Lo (1997) propose a destination choice model that accounts for the selection bias due to correlation between error terms in the choice set model and the choice model. A more comprehensive review of choice set modeling, in the context of destination choice, can be found in Thill (1992).

Conventional random-utility models assume implicitly that the choice set is independent of the decisionmaker's preferences conditional on the explanatory variables of the models. This assumption of independence is implausible in many choice situations where the decisionmaker chooses his or her choice set. Horowitz (1991) is concerned with this problem and develops an approach to modeling choice set selection and modeling choice conditional on endogenous choice sets. Horowitz's theoretical framework is presented in the next section.

3. THE CHOICE MODEL

3.1 The Choice Context

The choice context is the one in which the number of alternatives is large, collection of detailed information about all alternatives is costly, and sequential search is too time consuming to be feasible. There are many examples such as automobile choice, residential location choice, and college choice. In those choice situations, the number of possible alternatives can be very large. Therefore, it is very costly for an individual to acquire information about all alternatives. A strategy that the individual may apply consists of using easily available though incomplete information to choose a small subset of alternatives, acquiring detailed information about alternatives in the subset, and finally selecting a single alternative from this subset. Horowitz (1991) considered these choice situations and proposed a choice process that may consist of up to three consecutive stages:

(1) Using easily available but incomplete information, the individual selects a subset D of the set of all possible alternatives.
(2) An external process selects a subset C of D.
(3) The individual chooses a single alternative from C.

In this choice process, the choice set is selected by the individual based on his or her preferences. Thus, it is called an endogenous choice set. The possibility that choice may be a multistage process and that the choice set is the outcome of the choice process seriously complicates the development of choice models, even if the choice set is observed by the analyst. This is because, when the choice set is endogenous, the process of generating choice sets and the process of making the final choice may depend on the same attributes of the individual or alternatives, so that the two processes are not independent in general. For example, in modeling migration destination choice, the variable "relatives and friends living there" may be unobserved by the analyst but influence all stages of the choice process.

3.2 The Theoretical Model

In application of random-utility models, Horowitz (1991) develops a theoretical framework for modeling choice with endogenous choice sets, which encompasses the three-stage process. To minimize the complexity of the discussion and implementation, in this paper we collapse stages 1 and 2 of the process into a single stage called choice set generation. The notations are introduced as follows. Let G be the set of all possible alternatives and C be a subset of G from which the final choice is made. Thus, C is the choice set. For the example of automobile choice, G consists of all makes and models of automobiles available in the market, and C is the set of automobiles that the consumer has test-driven. Let z_i be a vector of attributes of alternative i that are both known to the individual when selecting C and observed by the analyst, and η_i be a random variable representing variables that are relevant to choice set selection but not observed by the analyst. Denote $U(x_i, \varepsilon_i)$ the utility of alternative i to the individual conditional on attributes that affect final choice, where x_i represents observed attributes and ε_i represents unobserved attributes. Let α and β be constant parameters. Following Horowitz, the probability that choice set C is chosen is given by

$$P(C) = P\left[\bigcap_{i\in C}(z_i\alpha + \eta_i \geq 0), \bigcap_{i\notin C}(z_i\alpha + \eta_i < 0)\right]. \tag{1}$$

The joint probability of final choice i and choice set C is

$$P(i,C) = P\left[\bigcap_{j\in C}(x_i\beta + \varepsilon_i > x_j\beta + \varepsilon_j, i \neq j), \bigcap_{j\in C}(z_j\alpha + \eta_j > 0), \bigcap_{j\notin C}(z_j\alpha + \eta_j < 0)\right]. \tag{2}$$

It follows that the probability of alternative i being chosen conditional on C is

$$P(i/C) = \frac{P(i,C)}{P(C)} = \frac{P(x_i\beta + \varepsilon_i > x_j\beta + \varepsilon_j, \forall j \in C; z_j\alpha + \eta_j \geq 0, \forall j \in C; z_j\alpha + \eta_j < 0, \forall j \notin C)}{P(z_j\alpha + \eta_j \geq 0, \forall j \in C; z_j\alpha + \eta_j < 0, \forall j \notin C)} \tag{3}$$

if $i \in C$; and $P(i|C) = 0$ if $i \notin C$. In the model of equations (3), the random component of the utility function ε_i is permitted to be correlated with η_i, the random component in the choice set generation function. In other words, the stochastic dependence of the random component of utility on C makes equation (3) differ from conventional choice models such as multinominal logit or probit models.

3.3 *The Operational Model*

Given a specification of the joint distribution of the random variables (ε_i, η_i) up to a set of constant parameters, α, β in equation (3) and any unknown parameters of the distribution of (ε, η) can be estimated in principle by the method of maximum likelihood subject to identification restrictions. The complexity of the specification that can be dealt with is, however, restricted by computational considerations. We want an operational model that preserves the essence of endogenous choice sets (η_i is correlated with ε_i) and that is as simple as possible in order to minimize computational complexities. The implementation is as follows. Assuming that the two-dimensional random variable (ε_i, η_i) is independently distributed across alternatives and ε_i is independently and identically distributed (IID), equation (3) yields

$$P(i/C) = \frac{\prod_{j \in C} P(x_i\beta + \varepsilon_i > x_j\beta + \varepsilon_j, j \neq i; \quad z_j\alpha + \eta_j \geq 0, \quad \forall j \in C)}{\prod_{j \in C} P(z_j\alpha + \eta_j \geq 0)} \tag{4}$$

In equation (4) the random variables ε_i and η_i for any alternative i may depend on common unobserved attributes and therefore be correlated. Further assume that

$$\eta_i = \gamma^2\varepsilon_i + \tau_i, \tag{5}$$

where τ_i is a random variable that is uncorrelated with ε_i, and γ is a scalar parameter. Note that, in equation (5), η_i is decomposed into two additive components: ε_i and τ_i. ε_i represents unobserved variables that affect both choice set selection and the final choice. The random variable τ_i represents unobserved attributes that affect choice set generation but not the final choice. For example, in college choice modeling, τ_i may represent unobserved variables such as expectations about financial aid that affect choice set generation but not the final choice. The parameter γ reflects differences in scale that may occur at different choice stages due to changes in the decisionmaker's information. Substitution of (5) into (4) yields

$$P(i/C) = \frac{\prod_{j \in C} P(x_i\beta + \varepsilon_i > x_j\beta + \varepsilon_j, j \neq i; z_j\alpha + \gamma^2\varepsilon_j + \tau_j \geq 0)}{\prod_{j \in C} P(z_j\alpha + \gamma^2\varepsilon_j + \tau_j \geq 0)} \tag{6}$$

The probability function of (6) can be evaluated by using either a probability simulator, or traditional numerical integration techniques. Let the random variables ε_i and η_i be IID normal and set the variances of ε and τ equal to one by normalization. For simplicity of notation, define $v_i = x_i\beta$, $w_i = z_i\alpha$, and $w_i^* = w_i/\sqrt{\gamma^4 + 1}$. Equation (6) then yields

$$\begin{aligned} P(i/C) = &\left[\int_{-\infty}^{+\infty} \Phi(w_i + \gamma^2\varepsilon) \prod_{j \neq i} \{[\Phi(v_i - v_j + \varepsilon) - 1]\Phi[w_j + \gamma^2(v_i - v_j + \varepsilon)]\right. \\ &\left. + \int_{-w_j - \gamma^2(v_i - v_j + \varepsilon)}^{+\infty} \Phi[(w_j + \tau)/\gamma^2]\phi(\tau)\,d\tau\}\phi(\varepsilon)\,d\varepsilon\right] \\ &\times \left[\prod_{j \neq i} \Phi(w_j^*)\right]^{-1} \end{aligned} \tag{7}$$

where $\Phi(\cdot)$ is the normal cumulative distribution function, $\phi(\cdot)$ is the normal probability distribution function. The derivation of equation (7) is given in the Appendix. Either equation (6) or equation (7) can be used as the operational model for applications.

mum likelihood estimation, is quite large. Thus, as a practical matter, the speed as well as accuracy of the probability calculation are important. In this section, the methods of calculating choice probabilities are discussed, the likelihood functions for estimating the model are defined, and a two-step sequential estimation procedure is proposed. The computationally efficient simulator to be introduced is also useful to evaluate other probability functions involving multiple integrals such as the multinominal probit model.

4.1 Methods of Calculating Choice Probabilities

Historically, numerical integration has been the most frequently used method for calculating choice probabilities (Hausman and Wise 1978; Daganzo 1979; Ben-Akiva and Lerman 1985). It is highly accurate and, if the function involves a single integral, it is also fast. However, this method has severe limitations in evaluating multiple integrals because the computational effort increases exponentially with the dimensionality of the integral. Another method for computing choice probability is numerical approximation. The well-known Clark approximation (Clark 1961; Daganzo 1979) sequentially approximates the distribution of the maximum of two normal variables by a normal variable with the mean and variance of this maximum. This method is fast, the computation effort increasing only quadratically with the number of alternatives in the choice set. However, the accuracy of the approximation is highly variable. The method appears to overestimate small probabilities, and its estimation bias cannot be reduced by increasing sample size (Horowitz, Sparmann, and Daganzo 1982).

An alternative approach is to use Monte Carlo methods to evaluate multiple integrals in the probability function. This approach approximates the choice probability by sampling the random variables in the choice model (for example, Lerman and Manski 1981; Geweke 1989; McFadden 1989). For any given number of Monte Carlo draws, the number of operations in the simulation increases almost linearly with the dimension of the integral. Thus, the simulation method is much more efficient than the numerical integration method. Moreover, estimation errors in simulations can be reduced by increasing the number of draws of random variables. In other words, the simulators have good asymptotic properties. The smooth recursive conditioning (SRC) simulator (Börsch-Supan and Hajivassiliou 1993) is especially promising. Like the Stern (1992) simulator, the SRC simulator is unbiased, bounded between zero and unity, and smoothed. Moreover, it can simulate choice probabilities more accurately than the Stern simulator. The use of the SRC simulator in maximum likelihood estimation is easy, simply adding the simulator as a user subroutine to a standard maximum likelihood estimation software. The maximum likelihood estimation with smoothed probability simulator is called smooth simulated maximum likelihood (SSML) estimation.

4.2 The Conditional Likelihood Function

To estimate the parameters of the model using the method of maximum likelihood estimation (MLE), the log likelihood function must be specified. Let N be the sample size. The log likelihood function of the choice model conditional on endogenous choice sets is

$$\ln L \quad \sum_{n=1}^{N} \sum_{i \in C} y_{in} \ln P(i_n | C_n, z_{in}, x_{in}, \alpha, \beta, \gamma), \tag{8}$$

where y_{in} is a choice indicator, which is equal to 1, if individual n has chosen alternative i, and equal to 0 otherwise. One way to obtain the estimates $\hat{\alpha}_N$, $\hat{\beta}_N$, and $\hat{\gamma}_n$ is to maximize and solve equation (8). However, this approach to estimating the model does not work. The results of numerical experiments, which will be presented later, show that the estimates of $\hat{\alpha}_N$ and $\hat{\gamma}_N$ obtained from maximizing (8) are highly erroneous. There might be identification problems with the parameters if information about the alternatives not in the choice set is not utilized. Thus, we use an alternative approach to obtaining the parameter estimates, which is discussed below.

4.3 *The Joint Likelihood Function and Sequential Estimation*

The alternative approach is to maximize a joint likelihood function, in which the information about the alternatives not in the choice set is utilized. The joint likelihood function is given by

$$\begin{aligned}\ln L &= \sum_{n=1}^{N} \ln P(i_n, C_n | z_{in}, x_{in}, \alpha, \beta, \gamma) \\ &= \sum_{n=1}^{N} \ln[P(i_n | C_n, z_{in}, x_{in}, \alpha, \beta, \gamma) P(C_n | G, z_{in}, \alpha)] \\ &= \sum_{n=1}^{N} \sum_{i \in C_n} \ln P(i_n | C_n, z_{in}, x_{in}, \alpha, \beta, \gamma) + \sum_{n=1}^{N} \sum_{C_n \subseteq G} \ln P(C_n | G, z_{in}, \alpha). \qquad (9)\end{aligned}$$

Equation (9) involves two likelihood functions: the likelihood function of the choice model and the likelihood function of a choice set model. Since the choice set model uses alternatives in and out of the choice set, it helps avoid the identification problem. Two methods can be used to estimate the parameters of the choice model. Although full information maximum likelihood (FIML) estimation can maximize (9) in one step, it is computationally burdensome. For this reason, we employ an alternative estimation procedure, termed sequential estimation.

In sequential estimation, we first maximize the likelihood function of the choice set model to get $\hat{\alpha}$, the estimate of α; and then, insert $\hat{\alpha}$ into the likelihood function of the choice model and maximize the second function to obtain estimates of $\hat{\beta}$ and $\hat{\gamma}$. The log likelihood of the choice set model is

$$\ln L_1 = \sum_{n=1}^{N_1} \sum_{C_n \subseteq D_n} \ln P(C_n | D_n, z_{in}, \alpha), \qquad (10)$$

where N_1 is the size of the sample for estimating α, and $P(C_n | D_n, z_{in}, \alpha)$ is the probability that individual n chooses choice set C_n, conditional on set D_n which is a subset of the set of all alternatives. The reason for using only a subset is to make the choice set model computationally tractable. The choice set model (Dai 1995) is

$$P(C_n | D_n) = \frac{P(C_n)}{P(C_n) + P(C_n^*)} = \frac{\prod_{i \in C_n} \Phi(w_{in}) \prod_{j \in C_n^*} [1 - \Phi(w_{in}^*)]}{\prod_{i \in C_n} \Phi(w_{in}) \prod_{j \in C_n^*} [1 - \Phi(w_{in}^*)] + \prod_{i \in C_n^*} \Phi(w_{in}^*) \prod_{j \in C_n} [1 - \Phi(w_{in})]}, \qquad (11)$$

where C_n is the individual's choice set, C_n^* is a set of alternatives selected randomly from the set $\{G - C_n\}$, and $D_n = \{C_n, C_n^*\}$. This choice set model requires C_n and C_n^* to be the same size in order to have the uniform conditioning property (McFadden 1978) hold. The second log likelihood for maximizing over β and γ is given by

$$\ln L_2 = \sum_{n=1}^{N_2} \sum_{i \in C} y_{in} \ln P(i_n | C_n, x_{in}, w_{in}, \beta, \gamma), \tag{12}$$

where $w_{in} = z_{in}\hat{\alpha}$, and N_2 is the size of the sample used in the second-step estimation. Given equations (10), (11), and (12), the sequential estimation procedure consists of two steps:

(1) Estimate the parameter vector α by applying MLE to the choice set model, $P(C_n | D_n, z_n, \alpha)$. Denote the estimates by $\hat{\alpha}$.
(2) Compute $w_{in} = z_{in}\hat{\alpha}$, for all $i \in C_n$. Use w_{in} as a separate independent variable to estimate the parameter vector β and scalar γ.

The estimates obtained from this procedure are consistent and asymptotically normal. The proof of consistency of α estimated using equation (11) is given in Dai (1995). The proof of consistency and asymptotic normality for the rest of the parameters is standard (for example, Amemiya 1985; Judge et al., 1985). Note that the variance-covariance matrix of the estimates obtained in step 2 cannot be estimated consistently using the usual formulae involving second derivatives of the log-likelihood function or the outer-product of the gradient. This problem is similar to that of sequential estimation of nested logit models (Amemiya 1978; McFadden 1981). The corrected asymptotic variance-covariance matrix of the sequential estimators is given by Dai (1995).

5. NUMERICAL EXPERIMENTS

The objectives of the numerical experiments are twofold: (1) to assess the accuracy of the probability simulator and the simulated maximum likelihood estimators; and (2) to evaluate the performance of the models and the estimation procedure. We want to know if the SRC simulator can produce estimates of the choice probabilities as accurate as those by numerical integration. We would like to see if the models and estimation methods work and how well they perform.

Equation (6) is used for obtaining the simulated probabilities and equation (7) is used for numerical integration. To compute the choice probabilities, it is necessary to specify the values of the parameters and explanatory variables in the model. The simplest example that can be constructed consists of the three parameters $\theta = [\alpha, \beta, \gamma]'$ and two explanatory variables, z and x, observed attributes relevant to choice set selection and final choice, respectively. Let 5 be the size of the choice set, $\Delta v = \Delta x\beta = [1, 0.5, 0.5]'$ and $w = z\alpha = [-0.24, -0.24, -0.24, -0.24]'$.The choice probabilities are computed, and the results are presented in Table 1. The empirical distributions of the simulated probabilities are based on one hundred simulations. Within each simulation the random variables are sampled R times. Table 1 shows that, even with only twenty Monte Carlo draws, the simulated probabilities (mean = 0.44085 and S.D.= 0.03100) are close to that by numerical integration (0.43970). Furthermore, the accuracy of the simulated probability estimates can always be increased by using a larger number of draws. The SRC simulator produces good results.

TABLE 1
Distribution of the Probability Simulator

	$R = 10$	$R = 20$	$R = 50$	$R = 100$
Simulated Probability				
Mean	0.44304	0.44085	0.44125	0.44098
Standard Deviation	0.04090	0.03100	0.01996	0.01380
"True" probability = 0.43970				
$\Delta v = \Delta x\beta = [1, 0.5, 0.5]'$ and $w = [-0.24, -0.24, -0.24, -0.24]'$				

To examine the performance of the simulated maximum likelihood estimates, the same set up of three parameters and two explanatory variables is used. In addition, two functions are employed to generate the data sets:

- Choice set generation function:

$$i_C = 1(z_i\alpha + \gamma^2\varepsilon_i + \tau_i > 0, i \in G)$$

- Choice function

$$y_i = 1(x_i\beta + \varepsilon_i > x_j\beta + \varepsilon_j, \forall j \in C, i \neq j)$$

Using the two functions, the data are generated in three steps:

Step 1. Specify the true values of the parameter vector and generate variables. The specified parameter vector is $\theta_0 = [\alpha_0, \beta_0, \gamma_0]' = [-0.4, 1.0, 0.5]'$. These values are chosen to obtain choice sets of reasonable size and choice probabilities of reasonable magnitude. The explanatory variables are realized as IID standard normal random variables. For each sample, the random variables τ and ε are drawn from IID standard normal distributions.

Step 2. Use the choice set generation function to obtain the choice set for eachobservation. The size of the master choice set G is set to be 16. The size of choice set C_n is allowed to vary from 2 to 7 and is drawn from G. For sequential estimation, a second set C_n^* is drawn from the set $\{G - C_n\}$. The elements in C_n and C_n^* are mutually exclusive. The range of 2 to 7 is chosen because 2 is the minimum size of a choice set and the double of 7 is smaller than 16.

Step 3. Use the choice function to generate choice from choice set C_n for each observation.

The parameters are estimated using the smooth simulated maximum likelihood (SSML) estimation method. The choice probabilities are estimated by the SRC simulator with one hundred Monte Carlo draws for each observation. The experiments are repeated forty times for each of the likelihood functions. The estimation results are presented in Table 2. The first part of the table presents the means and standard deviations of the estimates computed using the conditional likelihood function defined in equation (8). The second part gives the results of sequential estimation using likelihood functions (10) and (12). The third part shows the true values.

The conditional likelihood function clearly produces biased estimates of α and γ. The results of sequential estimation using the likelihood functions in (10) and (12) are much better. The mean value of $\hat{\alpha}$ is -0.4124, which is close to the true value -0.4000. The standard deviation is small (0.0494). These suggest that the

TABLE 2
Distribution of the SSML Estimates

(1) SSML estimates based on the conditional likelihood function				
		$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$
	Mean	−0.1852	0.9348	0.6692
	S.D.	1.7637	0.1764	0.4215
(2) SSML estimates based on the joint likelihood function–Sequential estimation				
Step 1				
		$\hat{\alpha}$		
	Mean	−0.4124		
	S.D.	0.0494		
Step 2				
			$\hat{\beta}$	$\hat{\gamma}$
	Mean		1.0051	0.5500
	S.D.		0.1113	0.2373
(3) True values				
		α_0	β_0	γ_0
		−0.4000	1.0000	0.5000

choice set model works well. The estimates obtained from the second step of estimation are also good. The mean values of estimated β and γ are [1.0051 0.5500], close to the true values [1.0000 0.5000], and the standard deviations are smaller than those obtained from the conditional likelihood function. The results indicate that the joint log-likelihood function and the sequential estimation method work very well.

6. EMPIRICAL APPLICATION

6.1 The Choice Context—College Choice

The objective of the empirical work is to test the choice model with endogenous choice sets using real world data. The empirical study is conducted in a particular choice context, namely, college choice by college-bound high school seniors. College choice provides an ideal context for the empirical application. There are at least two reasons for this. First, college choice data are readily available. To investigate the external validity of choice models based on laboratory choice experiments, Horowitz and Louviere (1990) have collected data on college choice in cooperation with the American College Testing Program (ACT). We have access to the data sets at no cost. In contrast, acquisition of real world data with similar quality for other choices (for example, automobile choice, residential location choice) could cost large amounts of money if the data could be acquired at all. Second, college choice by high school seniors does not involve certain complications that can be present in models of other choices. For example, college choice is clearly a multistage choice process (that is, application, admission, and final choice) and the choice set (the set of colleges to which the student has been admitted) is readily observable.

6.2 Data and Variables

The college choice data sets were collected during 1987 to 1988. A survey of a random sample of 1,265 ACT-tested high school seniors was conducted by mail to determine the colleges to which each of the seniors had been admitted, the amount of financial aid each senior had received, and the college the senior had chosen to attend. The survey was conducted in the fall, after the students had enrolled in their chosen colleges. This survey plus an ACT sup-

TABLE 3
Definition of Variables

Symbol	Definition
TUI	Annual tuition in dollars
RMBD	Annual cost of room and board in dollars
FIN	Financial aid received in dollars
WS	Dummy variable equal to 1 if work study is available, 0 otherwise
DIST	Distance from student's home to college in miles
TLSTU	Total students enrolled in the college
POP	Population of the city or town in which the college is located
HSEL	Dummy variable equal to 1 if the average ACT composite score or SAT total score of entering freshmen exceeds 90% of the scores of other students in the nation, 0 otherwise
MSEL	Dummy variable equal to 1 if the average ACT composite score or SAT total score of entering freshmen exceeds 50% of the scores of other students in the nation but less than those of the top 10% of students, 0 otherwise
LSEL	Dummy variable equal to 1 if the average ACT composite score or SAT total score of entering freshmen is less than 50% of the scores of other students in the nation, 0 otherwise
FOUR	Dummy variable equal to 1 for four-year or more college and 0 otherwise
BA	Dummy variable equal to 1 if the highest degree offered is the B.A. or B.S., 0 otherwise
MA	Dummy variable equal to 1 if the highest degree offered is the M.A. or M.S., 0 otherwise
SPORT	Dummy variable equal to 1 if the college is in NCAA division 1 or 1A for football or basketball and 0 otherwise
PRIV	Dummy variable equal to 1 if the college is private, 0 otherwise
COLD	Dummy variable equal to 1 if the average daily maximum temperature in January is below 32F and 0 otherwise
MILD	Dummy variable equal to 1 if the average daily maximum temperature in January is between 32 and 45F
ACTS	The student's ACT composite score
FINC	The student's family income in 1987 dollars 0 = Less than \$6,000, 1 = \$6,000 to 11,999, 2 = \$12,000 to \$17,999, 3 = \$18,000 to 23,999, 4 = \$24,000 to \$29,999, 5 = \$30,000 to 35,999, 6 = \$36,000 to \$41,999, 7 = \$42,000 to 49,999, 8 = \$50,000 to \$59,999, 9 = \$60,000 and over.
ENW	Dummy variable equal to 1 if the student expected not to work during first year of college

plemental survey yielded 602 responses, of which 302 observations are usable for this study. The data include a number of attributes of the students and the colleges in their choice sets. These attributes represent factors appearing to be important in college choice, such as college cost and financial aid, selectivity on admission, students' academic aptitude, and family income. The attributes are coded into twenty variables, which are defined in Table 3. In choice modeling, the variables representing individual attributes are usually used in interaction terms with attributes of the alternatives. The interaction variables used in this study are defined in Table 4. A detailed discussion of the data collection process can be found in Horowitz and Louviere (1990). Discussions of variables affecting college choice are referred to by Chapman (1981); Manski and Wise (1983); Cook and Zallocco (1983); Erdman (1983); Hearn (1984, 1991); Hossler, Braxton, and Coopersmith (1989); Sanders (1990); Dixon and Martin (1991); Flint (1991, 1993); Rickman and Green (1993).

6.3 Estimation and Testing

The coefficients in the choice model with endogenous choice sets are estimated using the two-step sequential estimation procedure discussed in section

TABLE 4
Definition of Interaction Variables

Symbol	Definition
HACT	Dummy variable equal to 1 if the student's composite ACT score is equal to or greater than 26, 0 otherwise
MACT	Dummy variable equal to 1 if the student's composite ACT score is between 19 and 25, 0 otherwise
	Dummy variable equal to 1 if the student's composite ACT score is equal to or less than 18, 0 otherwise
	Dummy variable equal to 1 if family income is equal to or higher than $42,000, 0 otherwise
MINC	Dummy variable equal to 1 if family income is between $24,000 and $41,999, 0 otherwise
LINC	Dummy variable equal to 1 if family income is less than $24,000, 0 otherwise
D30	Dummy variable equal to 1 if distance from student's home location to college location is greater than 30 miles, 0 otherwise
*HSEL*HACT*	Higher selectivity times higher ACT composite score
*HSEL*MACT*	Higher selectivity times medium ACT composite score
*MSEL*MACT*	Medium selectivity times medium ACT composite score
*LSEL*HACT*	Lower selectivity times higher ACT composite score
*PRIV*HACT*	Private college times higher ACT composite score
*PRIV*ENW*	Private college times expected not to work
*PRIV*HINC*	Private college times higher family income
*RBC*D30*	Room and board cost times distance greater than 30 miles

3.3. In this procedure the parameters affecting choice set generation are estimated first by sampling choice sets [see equations (10) and (11)]. Conditional on those estimates, the rest of the parameters in the model are estimated [equation (12)]. The choice probabilities are computed using the smooth recursive conditioning simulator with one hundred Monte Carlo draws for each observation. The results from the first estimation step are presented in Table 5, and the results from the second step are shown in Table 6. Reported in the tables are parameter estimates and their (asymptotic) standard errors and values of the *t* statistic.

All estimates have expected signs and most of them are statistically significant at the 0.05 significance level. Note that a positive value of an estimate indicates a positive relationship between the dependent variable and the corresponding explanatory variable, and a negative value vice versa. We first examine the estimated parameters that affect choice-set selection (see Table 5) and then those affecting final choice (see Table 6). The coefficients associated with *HSEL*HACT* and *LSEL*HACT* have opposite signs with the former being positive and the latter being negative. This suggests that students with higher ACT scores are more likely to apply for colleges with higher selectivity and unlikely to apply for those that admit almost everyone, other things being equal. The positive values with *MSEL*MACT* and *HSEL*MACT* indicate that students with medium ACT scores tend to choose to apply to moderately and highly selective colleges. It may also imply that the highly selective colleges are likely to admit students not only in the higher ACT score group but also in the medium ACT score groups. The coefficients of *PRIV*HACT* and *PRIV*ENW* suggest that private colleges are likely to be in the choice sets of those students who have higher ACT scores and who do not expect, at the time of application, to work during the first year of college. The signs of *FOUR* and *TUI* suggest that the students are more interested in four-year colleges than two-year ones, and that students are less likely to include colleges with higher tuition into their choice sets, other things being equal.

TABLE 5
Results of First-Stage Estimation

No.	Variable	Parameter Estimate	Asymptotic Std. err.	t-value
1	*HSEL*HACT*			
2	*LSEL*HACT*			
3	*MSEL*MACT*			
4	*HSEL*MACT*			
5	*PRIV*HACT*			
6	*PRIV*ENW*			
7	*FOUR*			
8	*TUI*			

First-Stage Summary
Number of observation = 287
Number of cases = 1,684
$L(a = 0) = -198.93$
$L(\hat{\alpha}) = -87.78$
$-2[L(a = 0) - L(a = \tilde{\alpha})] = 222.3$

TABLE 6
Results of Second-Stage Estimation Using Simulated Likelihood

No.	Variable	Parameter Estimates	Asymptotic Std. err.	t stat.
9	*AID*			
10	*WS*			
11	*TUI*			
12	*RBC*D30*			
13	*BA*			
14	*COLD*			
15	*PRIV*HINC*			
16	*LSEL*			
17	Gamma			

Second-stage summary
Number of observations = 226
Number of cases = 659
$L(\gamma = 0.5, \beta = 0, \alpha = \tilde{\alpha}) = -230.00$
$L^*(\gamma = \hat{\gamma}, \beta = \hat{\gamma}, \alpha = \tilde{\alpha}) = -191.80$
$-2[L - L^*] = 76.41$

Combined model summary
$L(\gamma = 0.5, \beta = 0, \alpha = 0) = -428.92$
$L^*(\gamma = \hat{\gamma}, \beta = \hat{\gamma}, \alpha = \tilde{\alpha}) = -279.58$

In the final stage of the choice process, a student selects a college to attend from the ones that have admitted him or her. The coefficients related to final choice are shown in Table 6. The coefficients of financial aid variables have expected positive signs and those of cost variables have expected negative signs. Specifically, the coefficient for the room and board cost variable has a significantly negative effect on an alternative's utility only when home-to-college distance is beyond a certain distance (thirty miles in this case). The signs of the remaining parameters suggest that, other things being equal, students from higher-income families are more likely to attend private colleges than those from lower-income families, students are less likely to choose a college that has low selectivity and offers no degrees higher than the baccalaureate, and students are also less likely to attend a college located in areas where winter is very cold. The results are consistent with the findings of previous research in college choice (for example, Manski and Wise 1983; Hearn 1991; Flint 1991). Other variables in the data are also tried in model specifications. Those variables include the college's involvement in major sports, college size, and the size of the city or town in which the college is located. It turns out that additional variables contribute little to increasing the likelihood values. Therefore, they are not included in the final model.

The coefficients in a random-utility model are measures of marginal utilities of the corresponding variables. A simple economic hypothesis is that the marginal utilities of financial aid variables are equal to the marginal disutilities of schooling cost variables. It is expected that a student would be indifferent in choosing a college to attend if the college gives the student an extra dollar in financial and increases the cost of attending that college by an extra dollar. In technical terms, the sum of coefficients of financial aid and cost variables should be approximately equal to zero. Thus, one way to check the estimation results is to test the null hypothesis that

$$(\beta_{AID} + \beta_{WS}) - (\beta_{TUI} + \beta_{RBC \cdot D30}) \quad 0,$$

where the coefficients for cost variables have negative values. The test statistic has chi-square distribution with one degree of freedom. The critical value for the test at the 0.05 significance level is 3.841 and the computed test value is 0.8181. Thus, the null hypothesis is not rejected, indicating that the estimation results are consistent with the expectation.

Using the estimation results, the hypothesis of endogenous choice sets can be tested easily. The null hypothesis is that choice set selection and final choice are independent in college choice. This is equivalent to testing that the coefficient of gamma is zero. The test statistic is the t test. Table 6 shows that the t value for gamma is 2.114. Thus, the null hypothesis is rejected at the 0.05 significance level. This provides empirical evidence that in college choice the choice set is indeed endogenous and that a model of college choice without considering the endogeneity of choice sets would have been misspecified.

6.4 Comparison with a Conventional Choice Model

We now examine the differences in estimated coefficients and in prediction between the new model and a conventional choice model. It can be shown (see Appendix) that the new model reduces to an identity probit model when the choice sets are exogenous. Thus, the identity probit model is chosen as the alternative model for comparison. Using the same data and same specification of the systematic component of the utility function for the new model, an identity probit model of college choice is estimated and the results are presented in Table 7. The differences between the two models in estimated parameter values are shown in Table 8. The probit model has fewer parameters than the new model since it ignores choice set selection. Of the parameters that are common to both models, the estimates have same signs but different values. It seems that the probit model consistently underestimates the values of parameters with a positive sign and overestimates those with a negative sign. The fractional differences between pairs of estimates from the two models range from 24 to 67 percent. Recall that the hypothesis in equation (15) is not rejected using the results of the new model. However, the hypothesis is now rejected using the results of the probit model. The test value computed using the probit model is

TABLE 7
Estimation Results of the Identity Probit Model

No.	Variable	Parameter Estimates	Asymptotic Std. err.	*t* stat.
1	*AID*			
2	*WS*			
3	*TUI*			
4	*RBC*D30*			
5	*BA*			
6	*COLD*			
7	*PRIV*HINC*			
8	*LSEL*			

Model summary
Number of observations = 226
Number of cases = 659
$L(b = 0) = -230.00$
$L^*(b = \hat{\beta}) = -192.47$
$-2[L - L^*] = 75.06$

TABLE 8
Differences in Estimated Parameter Values

Variable Name	Estimated Parameters (New Model) $\beta 1$	(Probit) $\beta 2$	Difference $\beta 2 - \beta 1$	$100^*(\beta 2 - \beta 1)/\beta 1$
	2.9106	2.1198		
	0.6227	0.4462		
	−1.0874	−0.5544		
	−1.3127	−1.0030		
	−0.4324	−0.3268		
	−0.7039	−0.5039		
	0.8693	0.6128		
	−0.2594	−0.0845		

7.6254, which is larger than the critical value (3.841) at the 0.05 significance level.

The differences in estimated coefficients do matter in choice analysis and in prediction because different coefficient values mean different marginal utilities of the corresponding variables in the models. The ratio of one coefficient to another measures the marginal rate of substitution (MRS) between one college attribute and another. For example, the ratio of marginal utility of low selectivity (LSEL) to the marginal utility of tuition (TUI) indicates the average student's trade-off between these two variables. The new model suggests that, on average, about a \$2,400 deduction in tuition is needed to compensate a student for going to a school with low selectivity, while the number estimated from the probit model is only about \$1,500. Similarly, the new model indicates that about a \$2,000 decrease in room and board cost is needed to compensate a student for attending a lower-selective college, whereas the value given by the probit model is about \$800. These examples show that students' willingness to pay inferred from the two models with and without endogenous choice sets are quite different.

Finally, the differences between the two models in prediction are assessed. Conventionally, aggregate demand for each alternative is predicted using competing models and then differences in aggregate demand are examined. However, aggregate demand for each alternative cannot be computed in this case because in the data set the number of alternatives (colleges) is larger than the number of students and few students share a given choice set. Thus, we can only compare predicted choice probabilities at the individual student level. Define

$$d_n = \sum_{i \in C_n} \frac{|P^*_{in} - P_{in}|}{P_{in}}$$

$$e_n = \frac{d_n}{I_n},$$

where P^*_{in} and P_{in} are the choice probabilities of student n choosing alternative i, computed by the probit model and the new model, respectively. d_n is the fractional difference between the two predicted probabilities summed over all alternatives in the choice set C_n. I_n is the size of C_n, and e_n is the average fractional difference in predicting the probability that a particular alternative is chosen. We call d_n the cumulative fractional difference and e_n the average fractional difference.

Three scenarios are developed for evaluating the distributions of d and e. In the first scenario, no values of the explanatory variables are changed. Thus, the predictions are based on existing conditions. In the second scenario, it is assumed that the amount of financial aid from the student's chosen college is increased by $2,000. In the third scenario, it is supposed that the chosen college reduces tuition by 50 percent as well as increases the amount of financial aid by $2,000. The purpose of the last two scenarios is to examine the differences between predictions by the two models when one or more of the explanatory variables change in values. Notice that in the new model the change in college attributes will not affect C_n and/or I_n. What will change are $P(C_n)$, the probability that C_n is chosen, and of course, P_{in}, the probability that alternative i is chosen conditional on C_n.

The distributions of the cumulative fractional differences and average fractional differences in predicted choice probabilities computed from the two models are shown in Table 9. The distributions are based on 226 observations. In the table, $d1$, $d2$, $d3$, and $e1$, $e2$, $e3$ represent the distributions of the differences based on scenarios 1, 2, and 3, respectively. In scenario 1, the medians of the two distributions $(d1, e1)$ are 0.3702 and 0.1358. It indicates that, on average, the

TABLE 9
Distribution of Fractional Difference in Prediction

(1) Cumulative Fractional Difference			
Percentile	$d1$	$d2$	$d3$
25	0.1772	0.3094	0.6359
Median	0.3702	0.6123	1.1510
75	0.7371	1.2423	2.2593
(2) Average Fractional Difference			
Percentile	$e1$	$e2$	$e3$
25	0.0634	0.1298	0.2637
Median	0.1358	0.2262	0.4159
75	0.2453	0.4002	0.6973

fractional difference is about 37 percent per observation, and approximately 14 percent per alternative. The differences are increased noticeably in scenarios 2 and 3, where predictions are made for changes in the financial aid and tuition variables. In scenario 2, the median values of d and e are 0.6123 and 0.2262, suggesting that the differences in prediction increase to 61 percent per observation and 23 percent per alternative, on average. In scenario 3, the average differences in d and e are about 115 percent and 42 percent, as a result of assuming an increase in the amount of financial aid and a decrease in tuition by the college the students chose to attend. Those results show that predictions by models with and without endogenous choice sets can be very different, especially when one or more of the explanatory variables have changed values.

7. CONCLUSIONS

In this paper a new choice model with endogenous choice sets has been estimated and tested. Choice models with endogenous choice sets are more complex both analytically and computationally than are conventional choice models. To estimate the model new calibration methods are introduced. In particular, the paper shows that the SRC simulator and the SSML estimation are promising: they are fast, accurate, efficient, and easy to use. These new methods are useful not only for calibrating this model but also other choice models with complex structures such as the multinominal probit model. Furthermore, a joint likelihood function and a two-step sequential estimation procedure are developed to estimate the parameters of the model. The results of numerical experiments show that the model, calibration methods, and estimation procedure work well.

The model is further tested in an empirical application, namely, college choice by high school seniors. The model fits the college choice data well. The empirical results are used to test the theory of endogenous choice sets. The test results show that, in college choice, choice set selection and final choice are not independent. The college choice model would have been misspecified if the endogeneity of choice sets were ignored in the modeling. Furthermore, the new model is compared with a conventional choice model in which choice sets are assumed to be exogenous. The results show that students' willingness to pay inferred from the two models is different and choice probabilities predicted by the two models are very different. The empirical evidence strongly suggests that ignoring the endogeneity of choice sets in choice modeling can have serious consequences in applications.

In implementing the theoretical model, this paper has made several simplifying assumptions. Some of the restrictive assumptions can be relaxed without making the model computationally intractable. For example, it is not difficult to extend the operational model to a choice process with two-stage choice set generation, thanks to the probability simulator. Some other assumptions, for instance, that the random variables influencing choice set selection are independent across alternatives, are very difficult to remove. In some applications the choice set may be difficult to identify or observe. These problems have not been dealt with in this paper and are treated as topics for further research. This paper is but a first step in implementing and applying this new choice model with endogenous choice sets. We intend to explore the possibility for the model to become an important element of several urban theories. For example, the model can be used in the study of urban housing markets and residential location choice, where it is well established that the two/three-stage choice process is common practice.

APPENDIX

Derivation of the Choice Model in Equation (7)

The probability that alternative i is chosen conditional on choice set C is given by

$$P(i|C) = \frac{P(v_i + \varepsilon_i > v_j + \varepsilon_j, \forall j \in C, j \neq i; w_j + \gamma^2 \varepsilon_j + \tau_j > 0, \forall j \in C)}{P(w_j + \gamma^2 \varepsilon_j + \tau_j > 0, \forall j \in C)}$$

$$\equiv \frac{P_i}{P_C}$$

We derive P_i first. Let P_i be conditional on ε_i and τ_j for all $j \in C$ and $i \neq j$ to obtain

$$P_i = P(\varepsilon_j < v_i - v_j + \varepsilon_i; \gamma^2 \varepsilon_j > -\tau_j - w_j; \tau_i > -w_i - \gamma^2 \varepsilon_i)$$

$$= P(\tau_i > -w_i - \gamma^2 \varepsilon_i) \prod_{j \neq i} P\left[-\frac{\tau_j + w_j}{\gamma^2} < \varepsilon_j < v_i - v_j + \varepsilon_i\right].$$

Assume ε_js and τ_i are iid normal, and let $1(A)$ be an indicator function which takes value 1, if A occurs, and 0 otherwise. Then,

$$P_i = \Phi\left(\frac{w_i + \gamma^2 \varepsilon_i}{\sigma_\tau}\right) \prod_{j \neq i} \left\{ \left[\Phi\left(\frac{v_i - v_j + \varepsilon_i}{\sigma_\varepsilon}\right) - \Phi\left(-\frac{\tau_j + w_j}{\gamma^2 \sigma_\varepsilon}\right)\right] \right.$$

$$\left. \times 1\left[\frac{\tau_j + w_j}{\gamma^2} < \frac{v_i - v_j + \varepsilon_i}{\sigma_\varepsilon}\right]\right\}$$

$$= \Phi\left(\frac{w_i + \gamma^2 \varepsilon_i}{\sigma_\tau}\right) \prod_{j \neq i} \left\{ \left[\Phi\left(\frac{v_i - v_j + \varepsilon_i}{\sigma_\varepsilon}\right) - \Phi\left(-\frac{\tau_j + w_j}{\gamma^2 \sigma_\varepsilon}\right)\right] \right.$$

$$\left. \times 1\left[\tau_j > -w_j - \frac{\gamma^2}{\sigma_\varepsilon}(v_i - v_j + \varepsilon_i)\right]\right\}$$

where $\Phi(\,)$ is the normal cumulative distribution function. We first integrate over the distributions of the τ_js $(j \neq i)$ to obtain

$$P_i = \Phi[(w_i + \gamma^2 \varepsilon_i)/\sigma_\tau) \prod_{j \neq i} \int_{-w_j(\gamma^2/\sigma_\varepsilon)(v_i - v_j + \varepsilon_i)}^{\infty} \{\Phi[(v_i - v_j + \varepsilon_i)/\sigma_\varepsilon]$$

$$- \Phi[-(\tau_j + w_j)/\gamma^2 \sigma_\varepsilon]\} \times (1/\sigma_\tau)\phi(\tau/\sigma_\tau)\, d\tau,$$

where $\phi(\,)$ is the normal probability density function. Notice that σ_ε and σ_τ are unidentifiable in (A4). Setting $\sigma_\varepsilon = \sigma_\tau = 1$ by normalization and finishing the tegration

$$P_i \quad \Phi(\quad + \quad) \prod \quad \Phi \quad v_j \quad)\Phi[w_j + \quad v_j +$$

$$\Phi[w_j \quad v_j + \quad + \int_{+\varepsilon_i}^{\infty} \Phi \quad + w_j \quad \tau_j \, d\tau \qquad \text{A5}$$

Finally tegrate ε_i to get

$$\int^{+\infty} \Phi \quad + \quad \prod_{j} \Phi(\quad v_j + \varepsilon_i) \quad \Phi[w_j + {}^2(\quad v_j + \varepsilon_i$$

$$+ \int_{\gamma^2}^{\infty} \Phi \quad + w_j) \gamma^2]\phi \quad) \quad \varepsilon_i) \, d\varepsilon_i$$

Now we deri in A.

$$P(w_j + \quad \varepsilon_j + \tau_j \quad \forall j \quad C$$

$$P(\gamma^2 \varepsilon_j + \tau_j \quad w_j \ \forall j \quad C$$

$$\prod_{j \in C} \Phi \left(\frac{w_j}{\sqrt{(\gamma^2 \sigma_\varepsilon)^2 + \sigma_\tau^2}} \right.$$

$$\prod_{\in C} \Phi \quad -\frac{w_j}{\sqrt{\gamma^4 + 1}}$$

F mp city fine w_j^* $w_j / \sqrt{\gamma^4 + 1}$ Equatio A7) be itt

$$\prod_{\in C} \Phi(w_j^*$$

Co ining (A6) an A the probability that alternative ch conditional th dogenou choice C b

$$\int^{+\infty} \Phi(\quad + \quad \prod \quad [\Phi(\quad v_j \quad) \quad]\Phi[w_j + \quad (\quad v_j +$$

$$+ \int^{\infty} \Phi \quad + w_j \quad {}^2]\phi(\quad) \quad \phi(\quad d\varepsilon$$

$$\prod_{j \in C} \Phi(w_j^*)$$

Horowitz (1991) points out that the choice model with endogenous choice sets reduces to a conventional choice model if choice set generation is independent of final choice. Thus, the model in (A9) should reduce to an identity probit model if γ approaches 0. This is one way to verify the model. As $\gamma \rightarrow 0$, the term in the inner integral of equation (A9) becomes

$$\Phi\left(\frac{\tau + w_j}{\gamma^2}\right) \Rightarrow 1(\tau + w_j > 0)$$

$$\Rightarrow \int_{-w_j}^{\infty} 1(\tau + w_j > 0)\phi(\tau)\,d\tau$$

$$\int_{-w_j}^{\infty} \phi(\tau)\,d\tau \quad \Phi(w_j) \tag{A10}$$

Using (A10) and setting γ 0, equation (A9) reduces to

$$P(i|C) \quad \frac{\prod_{j\in C} \Phi(w_j) \int_{-\infty}^{\infty} \prod_{j\neq i}(v_i - v_j + \varepsilon)\Big]}{\prod_{j\in C} \Phi(w_j)}$$

$$\int_{-\infty}^{\infty} \left[\prod_{j\neq i}(v_i - v_j + \varepsilon)\right]\phi(\varepsilon)\,d\varepsilon, \qquad \text{(A11}$$

which is, indeed, an identity probit model.

LITERATURE CITED

Adler, T., and M. Ben-Akiva (1976). "Joint Choice Model for Frequency, Destination, and Travel Mode for Shopping Trips." *Transportation Research Record* 569, 136–50.

Amemiya, T. (1978). "On a Two-step Estimation of a Multivariate Logit Model." *Journal of Econometrics* 8, 13–21.

———. (1985). *Advanced Econometrics*. Cambridge: The Harvard University Press.

Ben-Akiva, M., and S. Lerman (1974). "Some Estimation Results of a Simultaneous Model of Auto Ownership and Mode Choice to Work." *Transportation* 4, 357–67.

———. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge: MIT Press.

Boccara, B. (1989). "Modeling Choice Set Formation in Discrete Choice Models." Ph.D. dissertation, Department of Civil Engineering, M.I.T.

Börsch-Supan, A., and V. A. Hajivassiliou (1993). "Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models." *Journal of Econometrics* 58, 347–68.

Burnett, K. P. (1980). "Spatial Constraints-Oriented Modeling as an Alternative Approach to Movement: Microeconomic Theory and Urban Policy." *Urban Geography* 1, 53–67.

Burnett, K. P., and S. Hanson (1979). "Rationale for an Alternative Mathematical Approach to Movement as Complex Human Behavior." *Transportation Research Record* 723, 11–24.

———. (1982). "The Analysis of Travel as an Example of Complex Human Behavior in Spatially Constrained Situations: Definitions and Measurement Issues." *Transportation Research* A 16, 87–102.

Chapman, D. W. (1981). "A Model of Student College Choice." *Journal of Higher Education* 52, 490–505.

Clark, C. E. (1961). "The Greatest of a Finite Set of Random Variables." *Operations Research* 9, 145–62.

Cook, R. W., and R. L. Zallocco (1983). "Predicting University Preferences and Attendance: Applied Marketing in Higher Education Administration." *Research in Higher Education* 19, 197–211.

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

Dai, J. (1995). "Modeling Discrete Choice with Endogenous Choice Sets." Ph.D. Dissertation, Department of Geography, University of Iowa, Iowa City, Iowa.

Dixon, P. N., and N. K. Martin (1991). "Measuring Factors That Influence College Choice." *NASPA Journal* 29(1), 31–36.

Erdman, D. G. (1983). "An Examination of Factors Influencing Student Choice in the College Selection Process." *The Journal of College Admissions* 100, 3–6.

Fischer, M. M., P. Nijkamp, and Y. Y. Papageorgiou, eds. (1990). *Spatial Choices and Processes*. North-Holland.

Flint, T. A. (1991). "Does Financial Aid Make Students Consider Colleges with a Wider Cost Range?" *Journal of Student Financial Aid* 21(2), 21–32.

———. (1993). "Early Awareness of College Financial Aid: Does It Expand Choice?" *The Review of Higher Education* 16(3), 309–27.

Fotheringham, A. S. (1983). "A New Set of Spatial Interaction Models: The Theory of Competing Destinations." *Environment and Planning* A 15, 15–36.

———. (1988a). "Consumer Store Choice and Choice Set Definition." *Marketing Science* 7, 299–310.

———. (1988b). "Market Share Analysis Techniques: A Review and Illustration of Current US Practice." In *Store Choice, Store Location, and Market Analysis*, edited by N. Wrigley. London: Routledge.

Fotheringham, A. S., and S. Trew (1993). "Chain Image and Store Choice Modeling: The Effects of Income and Race." *Environment and Planning* A 25, 179–96.

Gaudry, M., and M. G. Dagenais (1979). "The Dogit Model." *Transportation Research B* 13, 105–11.

Gautschi, D. A. (1981). "Specification of Patronage Models for Retail Center Choice." *Journal of Marketing Research* 18, 162–74.

Geweke, J. (1989). "Efficient Simulation from the Multivariate Normal Distribution Subject to Linear Inequality Constraints and the Evaluation of Constraint Probabilities." Mimeo, Duke University, Durham, N.C.

Golledge, R. G., and G. Rushton, eds. (1976). *Spatial Choice and Spatial Behavior*. Columbus: Ohio State University Press.

———. (1984). "A Review of Analytic Behavioral Research in Geography." In *Geography and Urban Environment*, edited by D. T. Herbert and R. J. Johnston, pp. 1–43. Englewood Cliffs: John Wiley and Sons.

Golledge, R. G., and R. J. Stimson (1987). *Analytical Behavioral Geography*. London: Croom Helm.

Golledge, R. G., and H. J. P. Timmermans, eds. (1988). *Behavioral Modeling in Geography and Planning*. London: Croom Helm.

Hansman, J. A., and D. A. Wise (1978). "A Conditional Probit model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferenees." *Econometrica* 46, 403–26.

Haynes, K. E., and A. S. Fotheringham (1991). "The Impact of Space on the Application of Discrete Choice Models." *Review of Regional Science* 20(2), 39–49.

Hearn, J. C. (1984). "The Relative Roles of Academic, Ascribed, and Socioeconomic Characteristics in College Destination." *Sociology of Education* 57, 22–30.

———. (1991). "Academic and Nonacademic Influences on the College Destinations of 1980 High School Graduates." *Sociology of Education* 64, 158–71.

Horowitz, J. L. (1985). "Travel and Location Behavior: State of the Art and Research Opportunities." *Transportation Research* 19A, 441–53.

———. (1991). "Modeling the Choice of Choice Set in Discrete-Choice Random-Utility Models." *Environment and Planning* A 23, 1237–46.

Horowitz, J. L., and J. J. Louviere (1990). "The External Validity of Choice Models Based on Laboratory Choice Experiments." In *Spatial Choices and Processes*, edited by M. M. Fischer, P. Nijkamp, and Y. Y. Papageorgiou, pp. 247–64. Amsterdam: North-Holland.

Horowitz, J. L., J. M. Sparmann, and C. F. Daganzo (1982). "An Investigation of the Accuracy of the Clark Approximation for the Multinomial Probit Model." *Transportation Science* 16, 382–400.

Hossler, D., J. Braxton, and G. Coopersmith (1989). "Understanding Student College Choice." In *Higher Education: Handbook of Theory*, vol. 5, edited by J. C. Smith. New York: Agathon Press.

John, E. P. (1991). "The Impact of Student Financial Aid: A Review of Recent Research." *Journal of Student Financial Aid* 21(1), 18–32.

Judge, G. G., W. E. Griffiths, R. C. Hill, H. Lutkepohl, and T. C. Lee (1985). *The Theory and Practice of Econometrics*. New York: John Wiley & Sons.

Lerman, S., and C. Manski (1981). "On the Use of Simulated Frequencies to Approximate Choice Probabilities." In *Structural Analysis of Discrete Data with Econometric Applications*, edited by C. Manski and D. McFadden, pp. 305–19. Cambridge, Mass.: MIT Press.

Lowe, J. C., and A. Viterio (1989). "Differential Spatial Attraction of Private Colleges and Universities in the United States." *Economic Geography* 65(3), 208–15.

Manski, C. (1977). "The Structure of Random Utility Models." *Theory and Decision* 8, 229–54.

Manski, C., and D. McFadden, eds. (1981). *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, Mass.: MIT Press.

Manski, C. F., and D. Wise (1983). *College Choice in America*. Cambridge, Mass.: Harvard University Press.

McFadden, D. (1978). "Modeling the Choice of Residential Location." In *Spatial Interaction Theory and Residential Location*, edited by A. Karlquist et al., 75–96. Amsterdam: North Holland.

———. (1981). "Econometric Models of Probabilistic Choice." In *Structural Analysis of Discrete Data with Econometric Applications*, edited by C. Manski and D. McFadden, pp. 198–272. Cambridge: MIT Press.

———. (1989). "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration." *Econometrica* 57, 943–52.

Meyer, R. (1980). "Theory of Destination Choice-Set Formation under Information Constraints." *Transportation Research Record* 673, 1–7.

Miller, E. J., and M. E. O'Kelly (1983). "Estimating Shopping Destination Models from Travel Diary Data." *Professional Geographer* 35, 440–49.

Miron, J. R., and L. Lo (1997). "A Selection Bias Approach to Destination Choice." *Geographical Analysis* 29(2), 145–68.

Ortuzar, J. de D., and L. G. Willumsen (1994). *Modelling Transport*. New York: John Wiley & Sons Ltd.

Pitfield, D. E. (1984). *Discrete Choice Models in Regional Science*. London Papers in Regional Science 14. London: Pion Limited.

Richardson, A. (1982). "Search Models and Choice Set Generation." *Transportation Research A* 16, 403–19.

Rickman, C. A., and G. Green (1993). "Market Segmentation Differences Using Factors of the College Selection Process." *College and University* 68(1), 32–37.

Rushton, G. (1969a) "Analysis of Spatial Behavior by Revealed Space Preference." *Annals of the Association of American Geographers* 59, 391–400.

———. (1969b). "The Scaling of Locational Preferences." In *Behavioral Problems in Geography*, edited by K. Cox and R. Gollege, Studies in Geography 17, 197–227. Northwestern University, IL.

———. (1969c). "Temporal Changes in Space Preference Structures." *Proceedings of the Association of American Geographers* 1, 129–32.

———. (1971). "Preference and Choice in Different Environments." *Proceedings of the Association of American Geographers* 3, 146–50.

Sanders, N. F. (1990). "Understanding Seniors' College Choices." *Journal of College Admission* 126, 3–8.

Southworth, F. (1981). "Calibration of Multinomial Logit Models of Mode and Destination Choice." *Transportation Research A* 15, 315–25.

Stern, S. (1992). "A Method for Smoothing Simulated Moments of Discrete Probabilities in Multinomial Probit Models." *Econometrica* 60, 943–52.

Swait, J., and M. Ben-Akiva (1987a). "Incorporating Random Constraints in Discrete Models of Choice Set Generation." *Transportation Research B* 21, 92–102.

———. (1987b). "Empirical Test of a Constrained Discrete Choice Model: Mode Choice in Sao Paulo, Brazil." *Transportation Research B* 21, 103–15.

Thill, J.-C. (1992). "Choice Set Formation for Destination Choice Modeling." *Progress in Human Geography* 16, 361–82.

Thill, J.-C., and J. L. Horowitz (1997). "Travel-Time Constraints on Destination-Choice Sets." *Geographical Analysis* 29(2), 108–23.

Timmermans, H. J. P., and R. G. Golledge (1990). "Applications of Behavioral Research on Spatial Problems. II: Preference and Choice." *Progress in Human Geography* 14, 311–54.

Williams, H. C. W. L., and J. D. Ortuzar (1982). "Behavioral Theories of Dispersion and Mis-specification of Travel Demand Models." *Transportation Research B* 16, 167–219.

Wrigley, N. (1982). "Quantitative Methods: Developments in Discrete Choice Modeling." *Progress in Human Geography* 6, 547–62.

———. (1988). *Store Choice, Store Location, and Market Analysis*. New York: Routledge.

Wrigley, N., and P. A. Longley (1984). "Discrete Choice Modeling in Urban Analysis." In *Geography and the urban Environment*, edited by D. T. Herbert and R. J. Johnston, pp. 45–94. New York: John Wiley & Sons Ltd.