Title
Simultaneous Equation Systems Involving Binary Choice Variables

Authors
van Wissen, Leo J.
Golob, Thomas F.

Publication Date
1990-05-01

Simultaneous Equation Systems Involving Binary Choice Variables

Leo J. van Wissen
Thomas F. Golob

May 1990
Reprint, No. 20

The University of California Transportation Center

University of California
Berkeley, CA 94720

The University of California Transportation Center

The University of California Transportation Center (UCTC) is one of ten regional units mandated by Congress and established in Fall 1988 to support research, education, and training in surface transportation. The UC Center serves federal Region IX and is supported by matching grants from the U.S. Department of Transportation, the California State Department of Transportation (Caltrans), and the University.

Based on the Berkeley Campus, UCTC draws upon existing capabilities and resources of the Institutes of Transportation Studies at Berkeley, Davis, and Irvine; the Institute of Urban and Regional Development at Berkeley; the Graduate School of Architecture and Urban Planning at Los Angeles; and several academic departments at the Berkeley, Davis, Irvine, and Los Angeles campuses. Faculty and students on other University of California campuses may participate in Center activities. Researchers at other universities within the region also have opportunities to collaborate on selected studies. Currently faculty at California State University, Long Beach, and at Arizona State University, Tempe, are active participants.

UCTC's educational and research programs are focused on strategic planning for improving metropolitan accessibility, with emphasis on the special conditions in Region IX. Particular attention is directed to strategies for using transportation as an instrument of economic development, while also accommodating to the region's persistent expansion and while maintaining and enhancing the quality of life there.

The Center distributes reports on its research in working papers, monographs, and in reprints of published articles. For a list of publications in print, write to the address below.

University of California
Transportation Center

108 Naval Architecture Building
Berkeley, California 94720
Tel: 415/643-7378
FAX: 415/643-5456

Authors of papers reporting on UCTC-sponsored research are solely responsible for their content. This research was supported by the U.S. Department of Transportation and the California State Department of Transportation, neither of which assumes liability for its content or use.

Simultaneous Equation Systems Involving Binary Choice Variables

Leo J. van Wissen
Thomas F. Golob

Institute of Transportation Studies
University of California at Irvine

Reprint, No. 20

Geographical Analysis
Vol. 22, pp. 224-243, 1990

The University of California Transportation Center
University of California at Berkeley

TABLE OF CONTENTS

1. INTRODUCTION

In this paper a simultaneous modeling system for dichotomous endogenous variables is developed and applied empirically to longitudinal travel demand data of modal choice. The reported research is motivated by three factors. First, the analysis of discrete data has become standard practice among geographers, sociologists, and economists. In the seventies a number of new tools were developed to handle multivariate discrete data (Bishop, et al., 1975; Fienberg, 1980; Goodman, 1972). However, while these methods are invaluable in studying empirical relationships among sets of discrete variables, they have a limited ability to reveal the underlying causal structure that generated the data.

Second, in travel demand analysis and housing market modeling, attention has been focused largely on single-equation models. It can be argued that this scope is too limited. Human decisions are usually not taken in isolation but in conjunction with other decisions and events. There may be complex feedback relations, recursive, sequential, and simultaneous decision structures that cannot be adequately described in a single equation. This has been a major motivation in the seventies in sociology for the development of a new modeling approach: linear structural equations with latent variables. Such models combine the classical simultaneous equation system model with a linear measurement model. Original developments, particularly the LISREL model

(Jöreskog, 1973, 1977), did not allow for discrete dependent variables. More recently, Muthén (1983, 1984, 1987) and others (e.g., Bentler, 1983, 1985) developed models that incorporate various types of non-normal endogenous variables, including censored/ truncated polytomous and dummy variables. This paper explores the possibilities of this method for simultaneous equation models in dynamic analysis of mobility.

A third motivation for the present research is the rapid growth of longitudinal data sets. In recent years many longitudinal surveys have become available for geographical, economic, and transportation analyses. In labor and housing market analysis the Panel Study of Income Dynamics (PSID, 1984) has played an important role (Heckman and Singer, 1985; Davies and Crouchley, 1984, 1985). In consumer behavior, the Cardiff Consumer Panel has been a major motivation for the development and testing of dynamic discrete choice models (Wrigley, et al., 1985; Wrigley and Dunn, 1984a, 1984b, 1984c, 1985; Dunn and Wrigley, 1985; Uncles, 1987). In the Netherlands a large general mobility panel has been conducted annually since 1984 (J. Golob, et al., 1985; van Wissen and Meurs, 1989). Here analyses have focused on discrete data on modal choice (T. Golob, et al., 1986), as well as on dynamic structural modeling (Golob and Meurs, 1987, 1988; Kitamura, 1987; Golob and van Wissen, 1988; Golob, 1988). The present paper is an extension of this line of research to incorporate dynamic structural models of modal choice, using data from the Dutch Mobility Panel.

This paper is organized as follows: In Section 2 the basic methodology is developed. In Section 3 the simultaneous equation system of dummy variables is compared with the conditional logistic model, which is derived from, and equivalent to, the familiar log-linear model. In the fourth section, both models are applied to a dynamic data set of train and bus usage. Some major conclusions regarding the above are drawn in the final section.

2. METHODOLOGY

2.1 Two Equation Systems

The basic model can be specified as a set of structural equations involving dummy endogenous variables. The dummy endogenous variables can be viewed as being generated by unobserved latent variables. Consider a latent variable y_i^* which is not observed, but the event $y_i^* > 0$ is observed, through the indicator $y_i = 1$:

$$y_i = 1 \quad \text{iff} \quad y_i^* > 0$$
$$y_i = 0 \quad \text{otherwise} \tag{2.1}$$

Next, suppose y^* is explained by the following relationship:

$$y_i^* = \gamma x_i + u_i \tag{2.2}$$

Then

$$\text{Prob}(y_i = 1) = \text{Prob}(u_i > -\gamma x_i) \tag{2.3}$$

Equations (2.1) through (2.3) define a binary choice model. If it is assumed that u is standard normal distributed then the probit model results:

$$\text{Prob}(y_i = 1) = \Phi(\gamma x_i) \tag{2.4}$$

where Φ is the standard normal cumulative distribution function.

For generalization to the multivariate case, suppose we have two latent variables, y_1^* and y_2^*, and two observed dummy indicators y_1 and y_2. The latent variables are assumed to be linear functions of the latent variables themselves as well as exogenous variables x. So, the following structural model is defined:

$$y_{1i}^* = \beta_1 y_{2i}^* + \gamma_1 x_{1i} + \zeta_{1i} \tag{2.5a}$$

$$y_{2i}^* = \beta_2 y_{1i}^* + \gamma_2 x_{2i} + \zeta_{2i} \tag{2.5b}$$

The β's denote the direct effects among the latent variables, and the γ's are the direct effects between the exogenous and endogenous variables. If the error terms ζ_1 and ζ_2 are assumed to be bivariate normally distributed with variance-covariance matrix Ψ then the bivariate probit model results.

In order to estimate the parameters in a structural equations system, a three-stage procedure is followed (see, e.g., Judge, et al., Ch. 14). First, the equations are written in reduced form. Second, least squares or maximum likelihood methods can be applied to the reduced-form equations to obtain consistent estimates of the reduced-form coefficients. Finally, the structural parameters are derived from the reduced-form parameters. A unique solution of the structural parameters in terms of the reduced-form estimates is not always possible, which is the problem of identification.

The three-stage procedure can be illustrated with the structural model given above. The reduced-form model is

$$y^*_{1i} = \pi_{11} x_{1i} + \pi_{12} x_{2i} + v_{1i} \tag{2.6a}$$

$$y^*_{2i} = \pi_{21} x_{1i} + \pi_{22} x_{2i} + v_{2i} \tag{2.6b}$$

where

$$\pi_{11} = \frac{\gamma_1}{1 - \beta_1 \beta_2}, \qquad \pi_{12} = \frac{\beta_1 \gamma_2}{1 - \beta_1 \beta_2}$$

$$\pi_{21} = \frac{\beta_2 \gamma_1}{1 - \beta_1 \beta_2}, \qquad \pi_{22} = \frac{\gamma_2}{1 - \beta_1 \beta_2}$$

$$v_1 = \frac{\beta_1 \zeta_2 + \zeta_1}{1 - \beta_1 \beta_2}, \qquad v_2 = \frac{\beta_2 \zeta_1 + \zeta_2}{1 - \beta_1 \beta_2} \tag{2.7}$$

The joint distribution of v_1 and v_2 is assumed bivariate normal according to:

$$E(v_{1i}) = E(v_{2i}) = 0,$$

$$E(v^2_{1i}) = \omega_{11} = 1, \qquad E(v^2_{2i}) = \omega_{22} = 1,$$

$$E(v_{1i} V_{2i}) = \omega_{12} \tag{2.8}$$

The variances ω_{11} and ω_{22} are set to 1 because the scale of the probit transformation is not identified (Maddala, 1983, p. 22).

The reduced-form model parameters in equation (2.6) can be estimated in two steps. The π's can be estimated by means of probit regressions. Next, the correlation among the errors can be estimated using the theory of tetrachoric correlations in 2x2 tables (Tallis, 1962).

Given estimates $\hat{\pi}$ from equation (2.6), we can calculate the $\hat{y}^*$'s. The probabilities of the joint occurrence of the two events can now be expressed as:

$$P_{11}(i) = \text{Prob}(y_{1i} = 1 \text{ and } y_{2i} = 1) = \Phi_2(\hat{y}^*_{1i}, \quad \hat{y}^*_{2i}, \quad \rho) \qquad (2.9a)$$

$$P_{10}(i) = \text{Prob}(y_{1i} = 1 \text{ and } y_{2i} = 0) = \Phi_2(\hat{y}^*_{1i}, \quad -\hat{y}^*_{2i}, \quad -\rho) \qquad (2.9b)$$

$$P_{01}(i) = \text{Prob}(y_{1i} = 0 \text{ and } y_{2i} = 1) = \Phi_2(-\hat{y}^*_{1i}, \quad \hat{y}^*_{2i}, \quad -\rho) \qquad (2.9c)$$

$$P_{00}(i) = \text{Prob}(y_{1i} = 0 \text{ and } y_{2i} = 0) = \Phi_2(-\hat{y}^*_{1i}, \quad -\hat{y}^*_{2i}, \quad \rho) \qquad (2.9d)$$

where Φ_2 (.) is the standardized bivariate normal distribution. Maximum likelihood can be used to obtain a consistent estimator of ρ, the correlation coefficient.

With the estimated reduced-form coefficients, the $\hat{\pi}$'s, and $\hat{\rho}$'s, it is possible to calculate the structural parameters. From equation (2.7) we have:

$$\hat{\pi}_{12}/\hat{\pi}_{22} = \hat{\beta}_1, \qquad \hat{\pi}_{21}/\hat{\pi}_{11} = \hat{\beta}_2 .$$

Further, $\hat{\gamma}_1$ may be derived from the $\hat{\beta}$'s and $\hat{\pi}_{11}$ (or $\hat{\pi}_{21}$), and likewise $\hat{\gamma}_2$ from $\hat{\pi}_{22}$ (or $\hat{\pi}_{12}$). Finally, the variances and covariance of the structural model can be determined from the ω's. Since ω_{11} and ω_{22} are set to 1, these parameters are not uniquely identified. The equations relating the ω's to the (co)variances ψ are:

$$\psi_{11}^2 + 2\beta_1\psi_{12} + \beta_1\psi_{22}^2 = (1 - \beta_1\beta_2)\,\omega_{11} \qquad (2.10a)$$

$$\beta_1\psi_{11}^2 + (1 + \beta_1\beta_2)\,\psi_{12} + \beta_2\psi_{22}^2 = (1 - \beta_1\beta_2)\,\omega_{12} \qquad (2.10b)$$

$$\beta_2\psi_{11}^2 + 2\beta_2\psi_{12} + \psi_{22}^2 = (1 - \beta_1\beta_2)\,\omega_{22} \qquad (2.10c)$$

In general, the ψ's can be solved given the ω's and the β's.

2.2 q-Equation Systems

Extension of the model to more than two equations is straightforward. In matrix form, equation (2.5) becomes:

$$y^* = By^* + \Gamma x + \zeta \qquad (2.11)$$

where y* is a (q x 1) vector of latent endogenous variables, x is an (m x 1) vector of exogenous variables, B is a (q x q) parameter matrix of the structural coefficients among the y* variables, Γ is a (q x m) parameter matrix of structural coefficients relating the exogenous and endogenous variables, and ζ is a (q x 1) vector of disturbances. The variance-covariance matrix of the ζ's is defined as Ψ, with elements ψ_{kl}. Analogous to equation system (2.6), we obtain the reduced-form matrix equation:

$$\begin{aligned} y^* &= (I - B)^{-1} \Gamma x + (I - B)^{-1} \zeta \\ &= \Pi x + v \qquad (2.12) \end{aligned}$$

where $\Pi = (I - B)^{-1} \Gamma$ is the (q x m) matrix of reduced form regression coefficients, and v is a vector of random disturbances with covariance matrix Var(v) = Ω (q x q). A typical element of Ω is ω_{kl}. From equation (2.12) the expectation and variance of y* conditional on x can be derived:

$$E(y^* \mid x) = (I - B)^{-1} \Gamma x \qquad (2.13)$$

$$Var(y^* \mid x) = (I - B)^{-1} \Psi (I - B)^{-1^T} \qquad (2.14)$$

where, as before, Ψ denotes the variance-covariance matrix of the ζ disturbance terms. Analogous to the two-equations case, (2.12) is solved for Π and Ω. Given consistent estimates $\hat{\Pi}$ and $\hat{\Omega}$ the structural parameter matrices B, Γ and ψ are then estimated. In the two-equations case the Π parameters can be solved using univariate probit regressions. However, the covariance matrix Ω is in general much more difficult to estimate. In the two-equations case, maximum likelihood estimation involves the evaluation of the bivariate normal distribution function, but in the q-equation case this involves the evaluation of the multivariate normal distribution in q dimensions. There is no closed form solution for this integral and one has to rely on numerical solutions, which become computationally expensive with large numbers of variables. Consequently, various approximations have been developed. Daganzo (1979) developed an algorithm based on work by Clark (1961), in which the largest of a finite set of multivariate normally distributed variables is computed. Muthén (1983, 1984) developed a method where only bivariate information on sample distributions is used. This limited-information maximum likelihood approach, coupled with generalized least squares (GLS) estimation of the structural parameters, is implemented in the computer program LISCOMP (Muthén, 1987).

2.3 Limited-Information GLS Estimation

The modeling framework of LISCOMP implements structural equation models with latent endogenous variables that are not normally distributed. Endogenous variable types that can be handled in this way include dichotomous variables, ordinal variables, and censored or truncated continuous variables. A special case of this class of models is the multivariate probit model described here. An example of a structural equation model with mixed types of endogenous variables in the transportation context is provided in Golob and van Wissen (1988). Here, the focus is on the estimation of the multivariate probit model.

The distinction between reduced-form model parameters and structural model parameters is crucial in simultaneous equations modeling. The reduced-form coefficients may be called sample statistics. These are the regression coefficients Π (intercepts and slopes) and residual correlations Ω.

In the limited-information approach, the elements of Ω are estimated using only bivariate sample information. The estimation involves evaluation of equation system (2.9) for each observation for each pair of latent variables to obtain the corresponding residual correlations ω.

Estimation of the structural model parameters involves optimally replicating the sample statistics as close as possible in terms of the free model parameters in the B, Γ, and Ψ parameter matrices, using the generalized least squares (GLS) approach developed by Brown (1974, 1982, 1984). In the application of weights in the GLS approach, it is useful to distinguish between the regression statistics Π and the correlation statistics Ω. Consider the vector of sample statistics $S = (S_a, S_b)$ with the following elements:

$$S_a = \text{vec}\{\Pi\} \qquad (2.15a)$$

$$S_b = K\,\text{vec}\{\Omega\} \qquad (2.15b)$$

where Π and Ω are the sample statistics and K selects lower-triangular elements from the symmetric correlation matrix. The vec-operator strings out matrix elements row-wise in a column vector. Next, consider the population vector $\sigma = (\sigma_a , \sigma_b)$ where σ_a corresponds to the regression structure and σ_b to the correlation structure of the model. From equations (2.13) through (2.14) we have:

$$\sigma_a = \text{vec} \{ (1 - B)^{-1} \Gamma \} \tag{2.16a}$$

$$\sigma_b = K \, \text{vec} \{ (1 - B)^{-1)} \Psi (1 - B)^{-1T} \} \tag{2.16b}$$

The total number of parameters in σ_a is q x m and in σ_b is $^1/_2$ q (q-1). So the total number of free parameters in B, Γ, and Ψ, denoted as r, cannot exceed q (m + $^1/_2$ (q -1)). A generalized least-squares approach can then be used to obtain the structural equation parameters. The fitting function is

$$F = (S - \sigma)^T W^{-1} (S - \sigma) \tag{2.17}$$

where W is the estimator of the asymptotic covariance matrix of S. (For details on the computation of W, see Muthén, 1984, p. 119). F provides a large-scale chi-square test of model fit to the first and second order statistics. If r is the total number of free model parameters in B, Γ, and Ψ, then the appropriate degrees of freedom is q (m + $^1/_2$ (q - 1)) - r. See Bentler (1980), for interpretation of the chi-square statistic in large samples.

3. A COMPARISON WITH THE CONDITIONAL LOGISTIC MODEL

The multivariate probit model presented here can be compared with the log-linear model (LLM) and equivalent conditional logistic model (CLM). The LLM is one of the most frequently applied tools in applied multivariate categorical analysis (Bishop, et al., 1975). The CLM can be derived from the LLM and distinguishes between dependent and independent variables. The LLM-CLM model is well suited to test statistical associations among categorical variables, but there are difficulties in estimating structural parameters among the variables. In particular, there are two types of problems:

1. The CLM is not sufficiently rich in parameters to distinguish between statistical association and structural relations (Heckman, 1978; Maddala, 1983).

2. The use of endogenous dummy variables is inconsistent. If specified as a dependent variable, a dichotomous variable is treated as a probability; if entered as an explanatory variable, it is treated as a dummy variable (Winship and Mare, 1983).

To demonstrate these points, the LLM and the corresponding CLM are introduced and compared with the associated multivariate probit models. In Section 4 an empirical comparison among the models is given.

Consider a trinomial discrete variable problem defined by the dichotomous variables Y_1, Y_2, and Y_3, with values 0 and 1. The cross-classification of these variables is a three-way table with $2^3 = 8$ cells with frequencies f_{ijk}. The LLM models the joint distribution of the variables in terms of main effects and interaction effects. The saturated trinomial LLM is:

$$\begin{aligned}\log [P(Y_1, Y_2, Y_3)] = {} & \mu_0 + \mu_1 Y_1 + \mu_2 Y_2 + \mu_3 Y_3 + \\ & \mu_{12} Y_1 Y_2 + \mu_{13} Y_1 Y_3 + \mu_{23} Y_2 Y_3 + \\ & \mu_{123} Y_1 Y_2 Y_3 \end{aligned} \tag{3.1}$$

where the singly subscripted μ's are the main effects, μ_{12}, μ_{13}, and μ_{23} are second order interaction terms, and μ_{123} is the third order interaction term. The model extends easily to include more variables and more categories (see, e.g., Bishop, et al., 1975; Fienberg, 1980; Goodman, 1972).

Model (3.1) contains as many parameters as there are cell frequencies. By imposing constraints on the μ terms, a more parsimonious model results. The predicted cell frequencies F_{ijk} can be compared with the observed frequencies f_{ijk} to determine whether the hypothesis expressed in the model fits the data, using the χ^2 test or the log-likelihood ratio. Individual μ terms can also be tested using the conventional t-test.

The LLM models the joint distribution of all variables: No distinction between dependent and independent variables is made. In order to predict the outcome of one variable conditional on the outcomes of the other variables, the LLM can be transformed into the conditional logistic model (CLM), the equivalence between the log-linear model and linear logistic models being well known (McCullagh and Nelder, 1983, Section 6.4).

From equation (3.1):

$$L_{1|23} = \log \frac{P(Y_1 = 1 \mid Y_2, Y_3)}{P(Y_1 = 0 \mid Y_2, Y_3)}$$

$$= \mu_1 + \mu_{12}Y_2 + \mu_{13}Y_3 + \mu_{123}Y_2Y_3 \qquad (3.2)$$

All terms not involving Y_1 cancel out. Similarly we have:

$$L_{2|13} = \mu_2 + \mu_{12}Y_1 + \mu_{23}Y_3 + \mu_{123}Y_1Y_3 \qquad (3.3)$$

and

$$L_{3|12} = \mu_3 + \mu_{13}Y_1 + \mu_{23}Y_2 + \mu_{123}Y_1Y_2 \qquad (3.4)$$

Equations (3.2) through (3.4) correspond to the same LLM. The form of the three conditional equations is similar to that of a simultaneous equation system. However, there are a number of cross-equation parameter constraints that are usually not imposed in simultaneous equation systems: μ_{12} appears in both equations (3.2) and (3.3), μ_{13} appears in equations (3.2) and (3.4), and μ_{23} appears in equations (3.3) and (3.4). The term μ_{123} appears in all three conditional expressions. Consequently, they cannot be given a causal interpretation but should be considered as association-type parameters.

CLM's can be given a structural interpretation if the system is recursive (Maddala and Lee, 1976). Consider the joint probability structure:

$$P(Y_1, Y_2, Y_3) = P(Y_1) \cdot P(Y_2 \mid Y_1) \cdot P(Y_3 \mid Y_1, Y_2) \qquad (3.5)$$

From equation (3.5) we have:

$$L_1 = \log \frac{P\,(Y_1 = 1)}{P\,(Y_1 = 0)} = \mu_1^1 \qquad (3.6)$$

$$L_{2\,|\,1} = \log \frac{P\,(Y_2 = 1 \mid Y_1)}{P\,(Y_2 = 0 \mid Y_1)} = \mu_2^2 + \mu_{12}^2\, Y_1 \qquad (3.7)$$

$$L_{3\,|\,12} = \log \frac{P\,(Y_3 = 1 \mid Y_1\, ,\, Y_2)}{P\,(Y_3 = 0 \mid Y_1\, ,\, Y_1)}$$

$$= \mu_3^3 + \mu_{13}^3\, Y_1 + \mu_{23}^3\, Y_2 + \mu_{123}^3\, Y_{12} \qquad (3.8)$$

In these equations, the superscripts of the μ terms denote the dimensions of the table used in the estimation. So, equations (3.6) through (3.8) are estimated in three stages: First, μ_1^1 is estimated from the marginal distribution of Y_1. Next, μ_2^2 and μ_{12}^2 are derived from estimating a saturated LLM on the two-way table of Y_1 and Y_2. Finally, μ_3^3 , μ_{13}^3 , μ_{23}^3 , and μ_{123}^3 are estimated from the three-way table containing all three variables. The μ's of the recursive structure can be given a causal interpretation (Goodman, 1973).

If we compare the recursive model structure equations (3.6) through (3.8) with the types of models discussed in Section 2, an important difference emerges. In the multivariate probit model the dichotomous variables are treated as latent variables both as independent and as explanatory (intermediate) variables. This allows the models to be written in structural form and in reduced form. In the CLM formulation, dichotomous variables are treated as dummies if they are explanatory, but as log odds if they are dependent variables. A substantive interpretation can be given for both forms, depending on the type of problem being modeled. Heckman (1978) introduced a simultaneous equation system with dummy endogenous var[illegible] containi[illegible] the latent variable and the observed dummy indicator. Certain logical constraints [illegible] structural parameters apply, however. Maddala (1983, Section 5.7) gave a substantive interpretation to including latent variables or their observed dummy counterparts in structural equations. The latent variable y* can be interpreted as "intention" and the observed y as the actual action. The model

$$
\begin{aligned}
y_1^* &= \beta_1 y_2^* + \gamma_1 x_1 + \zeta_1 \\
y_2^* &= \beta_2 y_1^* + \gamma_2 x_2 + \zeta_2
\end{aligned}
\qquad (3.9)
$$

specifies that the intentions about y_1 and y_2 are determined jointly by the x's. On the other hand, the model

$$
\begin{aligned}
y_1^* &= \beta_1 y_2 + \gamma_1 x_1 + \zeta_1 \\
y_2^* &= \beta_2 y_1 + \gamma_2 x_2 + \zeta_2
\end{aligned}
\qquad (3.10)
$$

says that the intention for y_1 is determined by the actual outcome of y_2, but the intention for y_2 is also determined by the outcome of y_1. On the condition that intentions precede actions, such a model is not logically consistent. This can be shown formally (Maddala, 1983, p. 119). Let $F_1(.)$ and $F_2(.)$ be the distribution functions of ζ_1 and ζ_2, respectively. Then, (from 3.10):

$$\text{Prob}(y_1 = 1) = F_1(\beta_1 y_2 + \gamma_1 x_1) \tag{3.11}$$

and

$$\begin{aligned}
\text{Prob}(y_1 = 1 \text{ and } y_2 = 1) &= F_1(\beta_1 + \gamma_1 x_1)\ F_2(\beta_2 + \gamma_2 x_2) \\
\text{Prob}(y_1 = 1 \text{ and } y_2 = 0) &= F_1(\gamma_1 x_1)\ [1 - F_2(\beta_2 + \gamma_2 x_2) \\
\text{Prob}(y_1 = 0 \text{ and } y_2 = 1) &= [1 - F_1(\beta_1 + \gamma_1 x_1)]\ F_2(\gamma_2 x_2) \\
\text{Prob}(y_1 = 0 \text{ and } y_2 = 0) &= [1 - F_1(\gamma_1 x_1)]\ [1 - F_2(\gamma_2 x_2)]
\end{aligned} \tag{3.12}$$

The sum of these probabilities is equal to:

$$1 + [F_1(\beta_1 + \gamma x_1) - F_1(\gamma_1 x_1)]\ [F_2(\beta_2 + \gamma_2 x_2) - F_2(\gamma_2 x_2)] \tag{3.13}$$

It is clear that, in order for this expression to be 1, either β_1 or β_2 has to be zero. The key point to be made here is that the question of including the observed dummy indicator or its latent counterpart as explanatory variables in a simultaneous system of equations has important theoretical and substantive implications. By using conditional logistic models, one is restricted to using dummy indicators, which may not be appropriate for a given problem.

4. EMPIRICAL MODEL COMPARISON

The differences between the multivariate probit model introduced in Section 2 and the conditional logistic model in Section 3 are illustrated using data from an ongoing national mobility panel in the Netherlands that was initiated in 1984. The survey involves the yearly recording of one week of travel behavior of a sample of approximately 1,800 households. A stratified sampling scheme was used, based on life cycle, household income, and place of residence. All household members over eleven years of age were surveyed. For more information on the sampling scheme and the survey, see J. Golob, et al., (1986), or van Wissen and Meurs (1989). The present research uses data from four waves of the Dutch panel conducted in the spring of each of the years 1984 through 1987. The data used in this study is restricted to driver-aged household members over 18 years of age.

The relations among five dichotomous variables are analyzed. Four of the five variables pertain to the use of two public transport modes: train (T) and bus-tram-subway (B) at two points in time (denoted T_1, T_2 and B_1, B_2, respectively). The remaining variable relates to car ownership of the household (C). A pooled wave-pair sample was used. For each person, yearly interval records were constructed. The variables related to mode usage were defined for the beginning and the end of the intervals. Intervals in which a change in car ownership occurred were excluded from the analysis. Table 1 defines the variables used in this section. Since all variables are dichotomous, we can organize the data in tabular form. Table 2 shows the $2^5 = 32$ cell frequencies in the table.

We postulate the following hypothesis concerning the data:

$$P(B_1, T_1, B_2, T_2, C) = P(C) \cdot P(B_1, T_1 \mid C) \cdot P(B_2, T_2 \mid B_1, T_1, C) \tag{4.1}$$

TABLE 1

VARIABLE DEFINITIONS

VARIABLE NUMBER	VARIABLE	DESCRIPTION
1	C	Car ownership indicator of the household (0= no car owned, 1= 1 + cars owned).
2	T_1	Train-usage (0= no usage, 1= one or more trips made) in first 7-day observation period in one-year interval
3	B_1	Bus-train-subway usage (0= no usage, 1= one or more trips made) in first 7-day observation period in one-year interval
4	T_2	Train usage in second 7-day observation period
5	B_2	Bus usage in second 7-day observation period

This is a partially recursive structure, containing one marginal and two conditional probabilities. Bus usage and train usage at time $t = 1$ are jointly determined by car ownership. In addition, there is also a lagged effect: bus and train usage at time $t = 2$ are jointly determined by bus and train usage at time $t = 1$ and car ownership.

First, we estimate this model using a set of simultaneous conditional logistic equations:

TABLE 2

RESPONSE PATTERN TRAIN & BUS USAGE, AND CAR OWNERSHIP
(source: Dutch Mobility Panel)

CAR OWNERSHIP = 0

BUS TIME 1 = 0

		TRAIN TIME 1 = 0, TRAIN TIME 2: 0	TRAIN TIME 1 = 0, TRAIN TIME 2: 1	TRAIN TIME 1 = 1, TRAIN TIME 2: 0	TRAIN TIME 1 = 1, TRAIN TIME 2: 1
BUS TIME 2	0	402	20	26	22
	1	97	39	6	20

BUS TIME 1 = 1

		TRAIN TIME 1 = 0, TRAIN TIME 2: 0	TRAIN TIME 1 = 0, TRAIN TIME 2: 1	TRAIN TIME 1 = 1, TRAIN TIME 2: 0	TRAIN TIME 1 = 1, TRAIN TIME 2: 1
BUS TIME 2	0	113	7	46	17
	1	222	39	39	94

CAR OWNERSHIP = 1

BUS TIME 1 = 0

		TRAIN TIME 1 = 0, TRAIN TIME 2: 0	TRAIN TIME 1 = 0, TRAIN TIME 2: 1	TRAIN TIME 1 = 1, TRAIN TIME 2: 0	TRAIN TIME 1 = 1, TRAIN TIME 2: 1
BUS TIME 2	0	4065	83	79	48
	1	240	56	4	18

BUS TIME 1 = 1

		TRAIN TIME 1 = 0, TRAIN TIME 2: 0	TRAIN TIME 1 = 0, TRAIN TIME 2: 1	TRAIN TIME 1 = 1, TRAIN TIME 2: 0	TRAIN TIME 1 = 1, TRAIN TIME 2: 1
BUS TIME 2	0	276	12	64	14
	1	256	17	21	52

$$L_C = \mu_1^1 \tag{4.2a}$$

$$L_{T1} = \mu_2^3 + \mu_{12}^3 C + \mu_{23}^3 B_1 \tag{4.2b}$$

$$L_{B1} = \mu_3^3 + \mu_{13}^3 C + \mu_{23}^3 T_1 \tag{4.2c}$$

$$L_{T2} = \mu_4^5 + \mu_{14}^5 C + \mu_{24}^5 T_1 + \mu_{45}^5 B_2 \tag{4.2d}$$

$$L_{B2} = \mu_5^5 + \mu_{15}^5 C + \mu_{35}^5 B_1 + \mu_{45}^5 T_2 \tag{4.2e}$$

The model defined by equation system (4) is depicted in the flow diagram of Figure 1. The superscripts of the μ terms denote the dimension of the table used in the estimation of the parameters. Thus, μ_1^1 is estimated from the marginal distribution of C. Next, train and bus usage at time t = 1 are determined jointly from car ownership, using a three-dimensional table ($T_1 * B_1 * C$). The terms μ_{12}^3 and μ_{13}^3 relate to the causal effect of car ownership on train and bus usage, respectively. The μ_{23}^3 term appears in both equations and is the association among bus and train usage at time t = 1. No causal interpretation can be given to this particular parameter since there is no direction of the effect. The terms μ_{24}^5 and μ_{35}^5 measure the lagged effects of train and bus usage, respectively. It is assumed that there are no cross-lagged effects of B_1 to T_2 and T_1 to B_2 ($\mu_{34} = \mu_{25} = 0$).

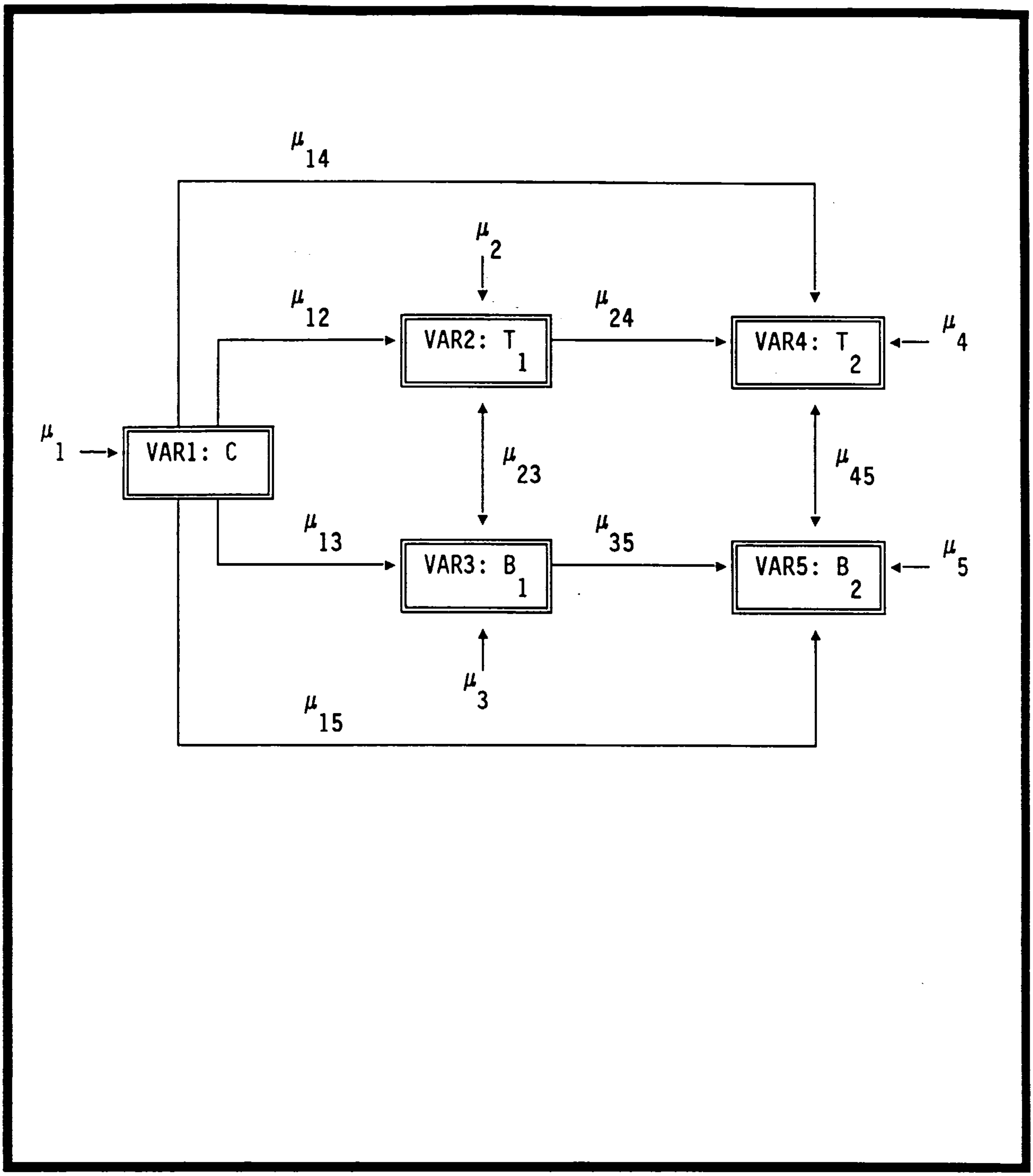

FIGURE 1

FLOW DIAGRAM
OF THE CONDITIONAL LOGISTIC MODEL

Table 3 contains the parameter estimates of the model. All estimated coefficients are highly significant. Regarding the effects of car ownership, there is the expected negative influence on both bus and train usage. The effect on bus usage is stronger than the effect on train usage in both time periods, which is also as expected because, for example, there is considerable train usage by higher income business travelers. Moreover, the coefficient values in t = 1 are larger than in t = 2, which could due to the absence of lagged effects for the t = 1 period. The lagged effects are highly significant. If we interpret these parameters as stability coefficients then train usage appears to be more stable than bus usage. Further, bus and train usage are highly complementary: the association between B_1 and T_1 is highly positive, as is the association between B_2 and T_2.

However, the model cannot answer the question of whether train usage implies bus usage or, conversely, whether there is indeed mere statistical association. Structural modeling provides the capability of answering this question through hypothesis testing and involving alternative model specifications. The first alternative model to be tested (designated as Model I) has the following form:

$$C^* = \zeta_1 \quad (4.3a)$$

$$T_1^* = \beta_{21} C^* + \zeta_2 \quad (4.3b)$$

$$B_1^* = \beta_{31} C^* + \zeta_3 \quad (4.3c)$$

$$T_2^* = \beta_{41} C^* + \beta_{42} T_1^* + \zeta_4 \quad (4.3d)$$

$$B_2^* = \beta_{51} C^* + \beta_{53} B_1^* + \zeta_5 \quad (4.3e)$$

with

TABLE 3

PARAMETER ESTIMATES OF THE
CONDITIONAL LOGISTIC MODEL
(Equation 4.2)

PARAMETER	EFFECT	COEFFICIENT	T-VALUE
μ_{12}^{3}	$C \rightarrow T_1$	-0.930	-9.22
μ_{23}^{3}	$B_1 \leftrightarrow T_1$	1.818	18.29
μ_{13}^{3}	$C \rightarrow B_1$	-1.577	-21.31
μ_{14}^{5}	$C \rightarrow T_2$	-0.489	- 4.30
μ_{24}^{5}	$T_1 \rightarrow T_2$	2.566	23.43
μ_{45}^{5}	$B_2 \leftrightarrow T_2$	1.593	15.48
μ_{15}^{5}	$C \rightarrow B_2$	-1.066	12.60
μ_{35}^{5}	$B_1 \rightarrow B_2$	2.191	28.13

χ^2 = 133.59
DF = 18
P = 0.000

$$\Psi = \text{cov}(\zeta) = \begin{bmatrix} 1 & & & & \\ 0 & 1 & & & \\ 0 & \psi_{32} & 1 & & \\ 0 & 0 & 0 & 1 & \\ 0 & 0 & 0 & \psi_{54} & 1 \end{bmatrix} \tag{4.3f}$$

Model I, specified in equation system 4.3 represents an implementation of the general structural equation formulation 2.11 with no exogenous x variables (Γ = null matrix) and

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \beta_{21} & 0 & 0 & 0 & 0 \\ \beta_{31} & 0 & 0 & 0 & 0 \\ \beta_{41} & \beta_{42} & 0 & 0 & 0 \\ \beta_{51} & 0 & \beta_{53} & 0 & 0 \end{bmatrix} \tag{4.4}$$

where the endogenous variables are in the same order as those in Table 1: C, T_1, B_1, T_2, B_2. The model is depicted in the flow diagram of Figure 2.

All variables are latent variables in this simultaneous equation system, so equation (2.4) holds for all y_i, i = 1,2,...,5:

$$\begin{aligned} \text{Prob}\ (y_i = 1) &= 1 - \Phi(k_i) \\ &= \Phi(-k_i) \end{aligned} \tag{4.5}$$

where the k_i unknown thresholds are estimated using the method described in Section 2.3.

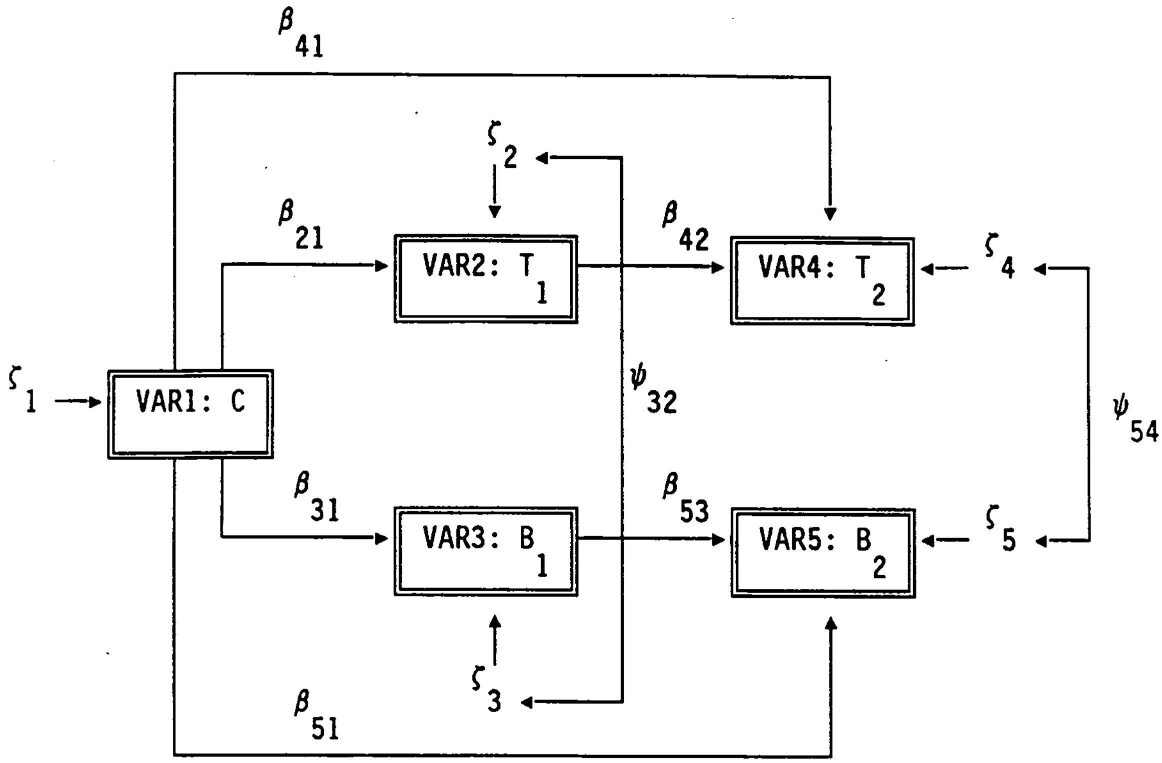

FIGURE 2

FLOW DIAGRAM
OF SIMULTANEOUS PROBIT MODEL I

Model I implies that there are no structural links between the bus and train choices at each point in time. These links are specified in terms of the correlations ψ_{32} and ψ_{54} between the respective residuals. These two parameters are analogous to the μ_{23} and μ_{45} parameters in the conditional logistic model (4.2). Further, the cross-lagged relations B_1--> T_2 and T_1--> B_2 are zero, so $\psi_{52} = \psi_{43} = 0$ and the effects from C and lagged relations B_1-->B_2 and T_1-->T_2 are modeled through the structural β parameters, which implies ($\psi_{21} = \psi_{31} = \psi_{41} = \psi_{51} = \psi_{42} = \psi_{53} = 0$). Thus, there are eight free parameters in the model. Table 4 displays the results of the estimation, using the limited information, GLS method of LISCOMP.

The coefficient values for the simultaneous probit Model I are not directly comparable to those of the conditional logistic model for two reasons. First, the variance of the logistic distribution is $3^{1/2} / \pi$, while the scale of the probit is set to 1. Amemiya (1981) suggests multiplying the logit estimates by 0.625 to get comparable values. Second, the explanatory variables are dummies in the CLM, but latent variables in the multivariate probit model, which makes direct comparison difficult. Further, the χ^2 tests are different and cannot be compared directly. For the log-linear model the test measures the difference in observed and predicted cell frequencies. In the multivariate normal model, the test measures the differences in sample statistics and estimated correlations by the structural model. The CLM can be rejected as a fit to the data, according to the χ^2 value (critical χ^2 value with $\alpha = 0.01$ is 34.8), while the simultaneous probit model cannot be rejected, the χ^2 statistic indicating an excellent fit to the sample statistics. On the other hand, a more complicated CLM specification, involving higher-order interaction terms cannot be rejected, implying that bivariate information is not sufficient in this

TABLE 4

PARAMETER ESTIMATES OF THE
SIMULTANEOUS PROBIT MODEL I
(Equation 4.3)

PARAMETER	EFFECT	COEFFICIENT	T-VALUE
β_{21}	$C^* \rightarrow T_1^*$	-0.485	-14.85
β_{31}	$C^* \rightarrow B_1^*$	-0.553	-21.25
β_{41}	$C^* \rightarrow T_2^*$	-0.166	-4.34
β_{51}	$C^* \rightarrow B_2^*$	-0.363	-11.41
β_{42}	$T_1^* \rightarrow T_2^*$	0.568	18.69
β_{53}	$B_1^* \rightarrow B_2^*$	0.412	15.27
ψ_{32}	$B_1^* \leftrightarrow T_1^*$	0.343	10.34
ψ_{54}	$B_2^* \leftrightarrow T_2^*$	0.280	10.46

χ^2 = 1.195
DF = 2
P = 0.547

particular sample. Therefore, a full information estimation method might give better results for the structural model. Given the computational difficulties involved in estimating a five-dimensional normal variate, the limited information procedure was used.

Unlike the CLM, the simultaneous probit model allows testing of alternative causal hypotheses through alternative structural forms. The first set of hypotheses involves causal influences between the bus and train variables. Instead of a correlation among these variables, unidirectional links can be specified: train usage implies bus usage, as specified in Model II--Figure 3; or bus usage implies train usage, as specified in Model III--Figure 4. In Model II (Figure 3), there are no free (nonzero) off-diagonal parameters in the Ψ matrix, and there are two additional free parameters in the B matrix (4.4): β_{32} and β54. Model III (Figure 4) differs from Model II in that the causality between T and B is reversed, and β_{23} and β_{45} are freed, rather than β_{32} and β_{54}. Model estimation results for Models II and III, contrasted with the base Model I, are given in Table 5. The χ^2 goodness-of-fit measure is best for the base Model I (correlated train and bus choice residuals). Regarding the two causal hypotheses, Model II is preferable to Model III: the hypothesis that train usage leads to bus usage fits the data better than the competing hypothesis that bus usage implies train usage.

A second set of hypotheses that can be tested with simultaneous probit models involves the lagged relations $T_1 \dashrightarrow T_2$ and $B_1 \dashrightarrow B_2$: the question is whether these are indeed structural links, implying true state dependence, or whether the errors are correlated over time (serial correlation). The base Model I implies the former, while the latter hypothesis of serial correlation corresponds to Model IV, depicted in Figure 5. In Model IV, two free parameters (ψ_{42} and ψ_{53}) are added to the Ψ residual correlation matrix 4.3f, while the two structural parameters β_{42} and β_{53} in the B-matrix in equation 4.4

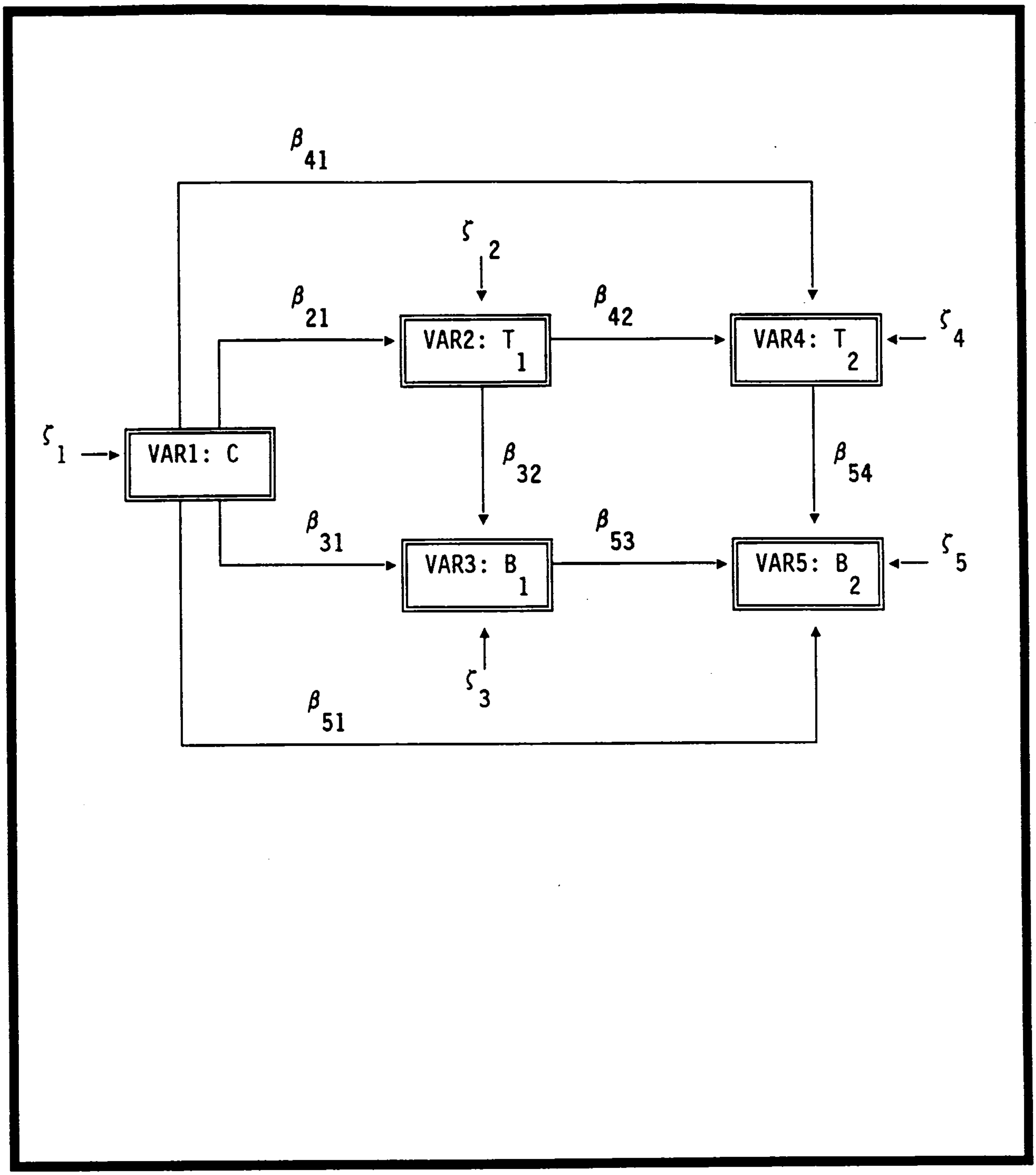

FIGURE 3

FLOW DIAGRAM
OF SIMULTANEOUS PROBIT MODEL II

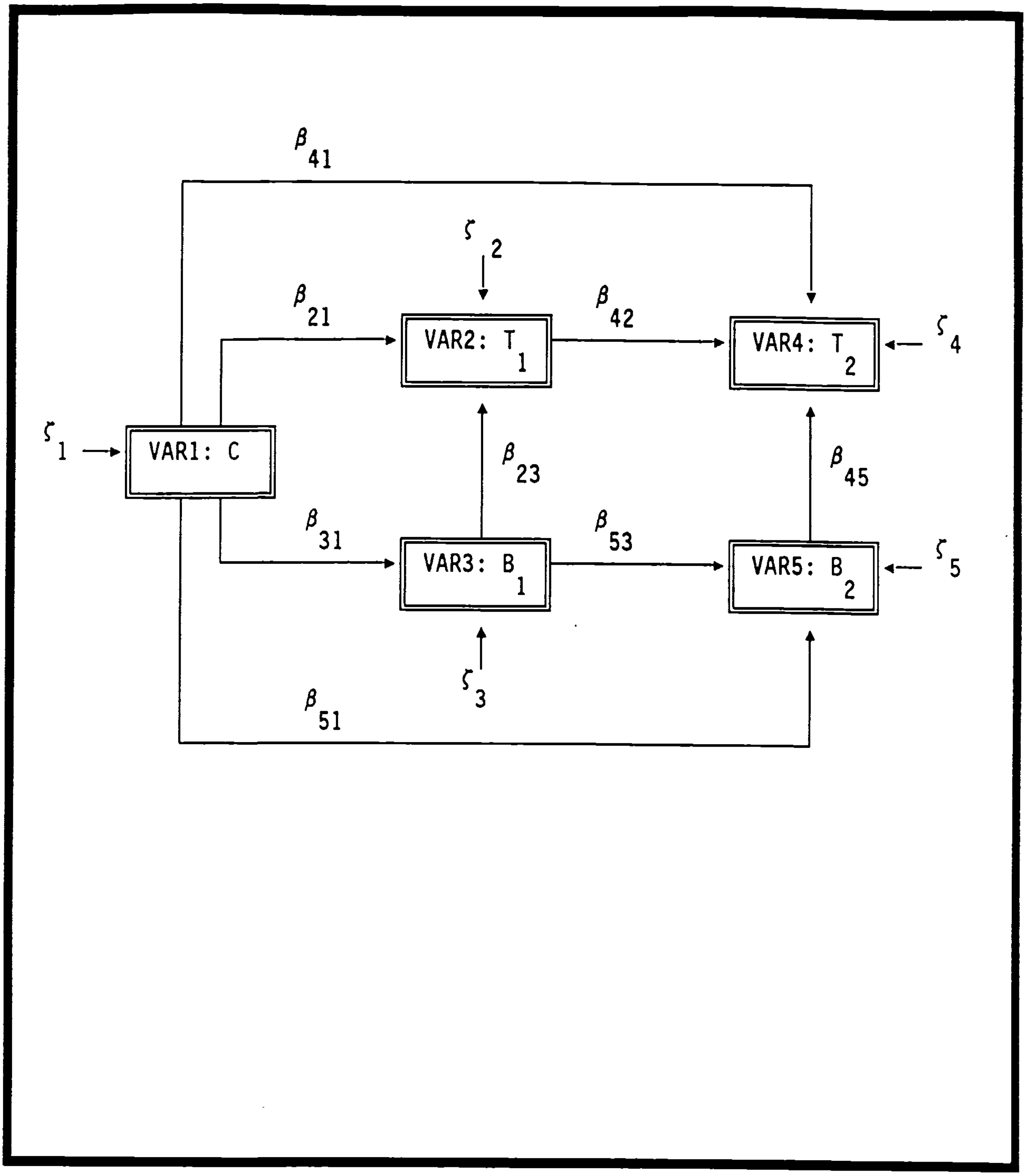

FIGURE 4

FLOW DIAGRAM
OF SIMULTANEOUS PROBIT MODEL III

TABLE 5

TESTS OF CAUSAL LINKS AMONG TRAIN AND BUS USAGE

PARAMETER	EFFECT	MODEL I		MODEL II		MODEL III	
		COEFF	(T-VAL)	COEFF	(T-VAL)	COEFF	(T-VAL)
β_{21}	$C^* \rightarrow T_1^*$	-.485	(-14.85)	-.478	(-14.68)	-.307	(-7.17)
β_{31}	$C^* \rightarrow B_1^*$	-.553	(-21.25)	-.394	(-12.03)	-.554	(-21.31)
β_{41}	$C^* \rightarrow T_2^*$	-.166	(-4.34)	-.171	(-4.52)	-.069	(-1.62)
β_{51}	$C^* \rightarrow B_2^*$	-.363	(-11.41)	-.313	(-9.79)	-.371	(-11.60)
β_{32}	$T_1^* \rightarrow B_1^*$			.351	(10.67)		
β_{42}	$T_1^* \rightarrow T_2^*$	.568	(18.69)	.594	(19.10)	.472	(17.98)
β_{23}	$B_1^* \rightarrow T_1^*$					.352	(10.47)
β_{53}	$B_1^* \rightarrow B_2^*$	.412	(15.27)	.323	(13.75)	.414	(15.15)
β_{54}	$T_2^* \rightarrow B_2^*$			.225	(10.26)		
β_{45}	$B_2^* \rightarrow T_2^*$					.243	(10.03)
ψ_{32}	$B_1^* \leftrightarrow T_1^*$	.343	(10.34)				
ψ_{54}	$B_2^* \leftrightarrow T_2^*$	.280	(10.46)				
χ^2 (DF)		1.195	(2)	7.535	(2)	10.141	(2)
P		0.5473		0.0226		0.0061	

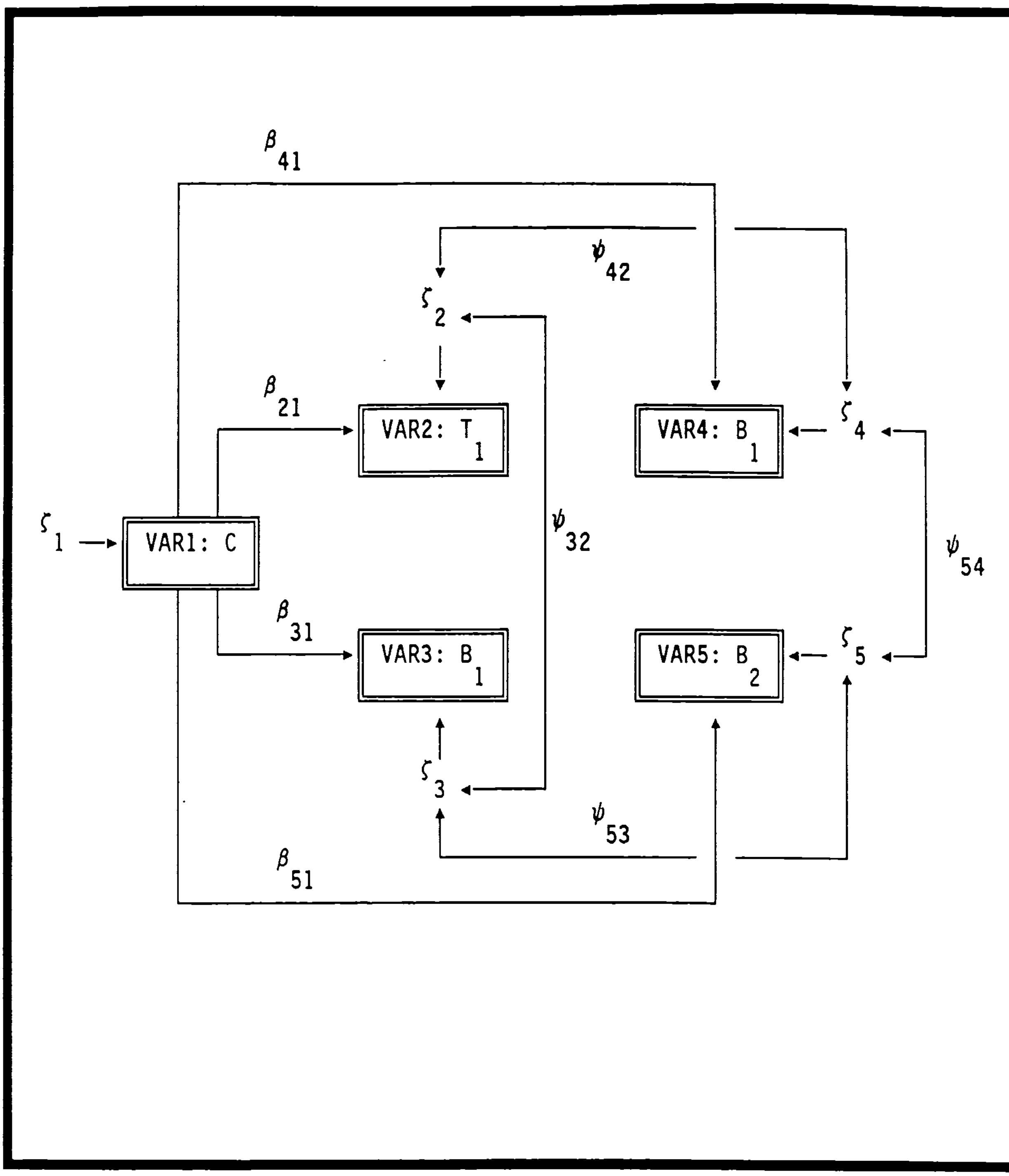

FIGURE 5

FLOW DIAGRAM
OF SIMULTANEOUS PROBIT MODEL IV

are restricted to zero. The estimation results for base Model I (with true state dependence) and Model IV (with serial correlation) are compared in Table 6. Clearly, the fit is much worse for Model IV, which implies true state dependence in public transport choice.

The aim here has been to compare the conditional logistic model with the structural simultaneous probit model. It was shown that the probit model allows testing of alternative causality in situations where the CLM only allows for correlation-type parameters.

5. CONCLUSIONS

In this paper, a simultaneous modeling system for dichotomous endogenous variables has been presented, based on the multivariate probit model. This model allows for causal hypotheses testing of sets of related discrete choice processes. One potentially fruitful application of this method is in the dynamic modeling of recurrent choices in time. The choice processes are usually linked through time lags, state dependencies, and serial correlation (heterogeneity). In principle, these dynamic relationships can be modeled in the framework presented here. A simple empirical example was given in Section 4. Public transport choice (train and bus) was modeled at two points in time, conditional on fixed car ownership levels. A number of hypotheses could be tested with this model. First, it was shown that train and bus choice was linked through correlation effects. Such an effect could be the result of mutual causation by excluded variables. Each of the competing hypotheses indicating that choice of one

TABLE 6

TEST OF STATE DEPENDANCE VERSUS SERIAL CORRELATION OF TRAIN AND BUS USAGE

PARAMETER	EFFECT	MODEL I		MODEL IV	
		COEFF	(T-VAL)	COEFF	(T-VAL)
β_{21}	$C^* \rightarrow T_1^*$	-.485	(-14.85)	-.546	(-17.01)
β_{31}	$C^* \rightarrow B_1^*$	-.553	(-21.25)	-.605	(-24.29)
β_{41}	$C^* \rightarrow T_2^*$	-.166	(-4.34)	-.509	(-15.36)
β_{51}	$C^* \rightarrow B_2^*$	-.363	(-11.41)	-.637	(-26.52)
β_{42}	$T_1^* \rightarrow T_2^*$	.568	(18.69)		
β_{53}	$B_1^* \rightarrow B_2^*$	.412	(15.27)		
ψ_{32}	$B_1^* \leftrightarrow T_1^*$	.343	(10.34)	.224	(7.96)
ψ_{42}	$T_1^* \leftrightarrow T_2^*$			.474	(16.10)
ψ_{53}	$B_1^* \rightarrow B_2^*$			.335	(12.96)
ψ_{54}	$B_2^* \leftrightarrow T_2^*$	.280	(10.46)	.254	(9.18)
χ^2 (DF)		1.195 (2)		35.087 (2)	
P		0.5473		0.0000	

mode causes the choice of the other mode showed a worse fit. Second, the nature of the time dependency of mode choice could be tested. The hypothesis of causally-related mode usage over time showed a much better fit than the rival hypothesis of serially-correlated errors, which is evidence for state dependency in train and bus choice.

The simultaneous probit model was also tested against the conditional logistic model. This model is derived from, and equivalent to, the log-linear model. Although the conditional logistic model is highly valuable for determining empirical relationships, it has only limited capability to test causal relationships versus mere statistical association. If the model is not fully recursive, then it is not possible to determine the true underlying causal structure. Moreover, the conditional logistic model is restricted in the types of model specifications that it allows: endogenous dichotomous variables are treated as dummies whenever they appear as explanatory variables in the equation system. This may not always be the proper representation of the underlying theory. The simultaneous probit model does not have this limitation.

Despite the theoretical advantages, there are still a number of methodological problems in estimating the simultaneous probit model. First, the assumption of multinormality for observed outcomes may not be appropriate in many cases. Second, full maximum likelihood estimation is still not feasible with large numbers of variables, given the current state of computer technology. The reliance on limited information solutions simplifies the estimation procedure, but more work is necessary to study all the consequences of the simplifications invoked.

REFERENCES

Amemiya, T. (1978). "The estimation of a simultaneous equation generalized probit model." *Econometrica*, 46: 1193-1205.

Amemiya, T. (1981). "Qualitative response models: A survey." *Journal of Economic Literature*, 19: 488-505.

Bentler, P.M. (1980). "Multivariate analysis with latent variables: Causal modeling." *Annual Review of Psychology*, 31: 419-456.

Bishop, T., S. Fienberg, and P. Holland (1975). *Discrete Multivariate Analysis.* Cambridge, MA: M.I.T. Press.

Brown, M.W. (1974). Generalized least squares estimators in the analysis of covariance structures. *South African Statistical Journal*, 8:1-24. (Reprinted in D.J. Aigner and A.S. Goldberger (eds.), (1977), *Latent Variables in Socio-Economic Models*. Amsterdam: North-Holland.)

Brown, M.W. (1982). Covariance structures. In Douglas M. Hawkins (ed.), *Topics in Applied Multivariate Analysis*: 72-141. Cambridge University Press.

Brown, M.W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37; 62-83.

Clark, C. (1961). "The greatest of a finite set of random variables." *Operations Research*, 9: 145-162.

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

Davies, R.B., and R. Crouchley (1984). "Calibrating longitudinal models of residential mobility and migration: An assessment of a nonparametric marginal likelihood approach." *Regional Science and Urban Economics*, 14: 231-247.

Davies, R.B., and R. Crouchley (1985). "Control for omitted variables in the analysis of panel and other longitudinal data." *Geographical Analysis*, 17: 1-15.

Dunn, R., and N. Wrigley (1985). "Beta-logistic models of urban shopping centre choice." *Geographical Analysis*, 17: 95-113.

Fienberg, S.E. (1980). *The Analysis of Cross-Classified Data*, 2nd Ed. Cambridge, MA: M.I.T. Press.

Golob, J.M., L.J.M. Schreurs, and J.G. Smit (1985). "The design and policy applications of a panel for studying changes in mobility over time." In, *Behavioural Research for Transport Policy*: 81-95. Utrecht, VNU Press.

Golob, T.F. (1988). "The causal influence of income and car ownership on trip generation by mode." *Journal of Transport Economics and Policy* (forthcoming).

Golob, T.F., and L. van Wissen (1988). "A joint household travel distance generation and car ownership model." Working Paper UCI-ITS-WP-88-15. Institute of Transportation Studies, University of California, Irvine.

Golob, T.F., L. van Wissen, and H. Meurs (1986). "A dynamic model of travel demand." *Transportation Research*, 20A: 401-414.

Golob, T.F., and H. Meurs (1987). "A structural model of temporal change in multi-modal travel demand." *Transportation Research*, 21A: 391-400.

Golob, T.F., and H. Meurs (1988). "Development of structural equation models of the dynamics of passenger travel demand." *Environment and Planning*, 20A: 1197-1218.

Goodman, L.A. (1972). "A generalized model for the analysis of surveys." *American Journal of Sociology*, 77: 1035-1086.

Goodman, L.A. (1973). "The analysis of multidimensional contingency tables when some variables are posterior to others: A modified path analysis approach." *Biometrica*, 60: 179-192.

Heckman, J.J. (1978). "Dummy endogenous variables in a simultaneous equation system." *Econometrica*, 46: 939-959.

Heckman, J.J., and B. Singer (1985). *Longitudinal Analysis of Labor Market Data*. Cambridge, MA: Cambridge University Press.

Jöreskog, K.G. (1973). A general method for estimating a linear structural equation system." In, *Structural Equation Models in the Social Sciences*, edited by A.S. Goldberger and O.D. Duncan: 85-112. New York, NY: Seminar Press.

Jöreskog, K.G. (1977). "Structural equations models in the social sciences: Specification, estimation and testing." In, *Applications of Statistics*, edited by P.R. Krishnaiah: 265-287. Amsterdam: North Holland.

Judge, G.G., W.E. Griffith, R. Carter Hill, H. Lütkepohl, and T.C. Lee (1985). *The Theory and Practice of Econometrics*, 2nd Edition. New York, NY: Wiley Series in Probability and Mathematical Statistics.

Kitamura, R. (1987). "A panel analysis of household car ownership and mobility." *Proceedings of the Japan Society of Civil Engineers*, 383: 13-27.

Lee, L.F. (1978). "Unionism and wage rates: A simultaneous equation model with qualitative and limited dependent variables." *International Economic Review*, 19: 415-433.

Lee, L.F. (1981). "Fully recursive probability models and multivariate log-linear probability models for the analysis of qualitative data." *Journal of Econometrics*, 16: 51-69.

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

Maddala, G.S., and L.F. Lee (1976). "Recursive models with qualitative endogenous variables." *Annals of Economic and Social Measurement*, 5: 525-545.

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

Muthén, B. (1983). "Latent variable structural equation modeling with categorical data. *Journal of Econometrics*, 22: 43-65.

Muthén, B. (1984). "A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 49: 115-132.

Muthén, B. (1987). *LISCOMP: Analysis of Linear Structural Equations with a Comprehensive Measurement Model*. Mooresville, IN: Scientific Software, Inc.

PSID--Panel Study of Income Dynamics (1984). Users Guide. Survey Research Center, Institute for Social Research, The University of Michigan, Ann Arbor, Michigan.

Tallis, G.M. (1962). "The maximum likelihood estimation of correlation from contingency tables." *Biometrics*, 18: 342-353.

Uncles, M.D. (1987). "A beta-logistic model of mode choice: Goodness of fit and intertemporal dependence." *Transportation Research*, 21B: 195-205.

Winship, C., and R.D. Mare (1983). "Structural equations and path analysis for discrete data." *American Journal of Sociology*, 89: 54-110.

Wissen, L. van, and H.J. Meurs (1989). "The Dutch Mobility Panel: Experiences and evaluation." *Transportation* (forthcoming).

Wrigley, N., and R. Dunn (1984a). "Stochastic panel-data models of urban shopping behaviour: 1. Purchasing at individual stores in a single city." *Environment and Planning A*, 16: 629-650.

Wrigley, N., and R. Dunn (1984b). "Stochastic panel-data models of urban shopping behaviour: 2. Multistore purchasing patterns and the Dirichlet model." *Environment and Planning A*, 16: 759-778.

Wrigley, N., and R. Dunn (1984c). "Stochastic panel-data models of urban shopping behaviour: 3. The interaction of store choice and brand choice." *Environment and Planning A*, 16: 1221-1236.

Wrigley, N., and R. Dunn (1985). "Stochastic panel-data models of urban shopping behaviour: 4. Incorporation of independent variables into the NBD and Dirichlet models." *Environment and Planning A*, 17: (3) March.

Wrigley, N., G.M. Guy, R. Dunn, and L.G. O'Brien (1985). "The Cardiff Consumer Panel: Methodological aspects of the conduct of a long-term panel survey." *Transactions of Institute of British Geographers*, New Series 10: 63-76.