

UC Berkeley

Earlier Faculty Research

Title

A Joint Household Travel Distance Generation And Car Ownership Model

Permalink

<https://escholarship.org/uc/item/72h4k912>

Authors

Golob, Thomas F.
Van Wissen, Leo

Publication Date

1989-03-01



**A Joint Household
Travel Distance Generation
And Car Ownership Model**

Thomas F. Golob
Leo Van Wissen

March 1989
Reprint No. 8

**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.

**A Joint Household Travel Distance
Generation and Car Ownership Model**

Thomas F. Golob
Leo Van Wissen

Institute of Transportation Studies
University of California at Irvine

Reprint No. 8

Transportation Research Magazine
December 1989

The University of California Transportation Center
University of California at Berkeley

A JOINT HOUSEHOLD TRAVEL DISTANCE GENERATION AND CAR OWNERSHIP MODEL

THOMAS F. GOLOB

Institute of Transportation Studies, University of California Irvine, Irvine CA 92717, U.S.A

and

LEO VAN WISSEN

Institute of Transportation Studies, University of California, Irvine.
and Department of Spatial Economics, Free University, Amsterdam, The Netherlands

(Received 21 November 1988)

Abstract—The product of this research is a dynamic simultaneous equations model of car ownership and modal travel distances as a function of income. The data are from the Dutch National Mobility Panel (1984–1987); and four modes are encompassed: car driver, car passenger, train, and bus-tram-subway. A novel feature of the simultaneous equation system is the consistent treatment of the measurement scales of the variables: ordered probit functions for income and car ownership and tobit functions for distances. The dynamics are expressed in terms of pooled panel survey measurements of the variables at two points in time one year apart. This allows the identification of lagged responses and serial correlations over a one-year time-horizon. Results indicate that increased car ownership and car kilometers at time T_1 is influenced by heavy usage of other modes at time T_1 . This indicates there are significant noninstantaneous adjustments of car ownership and usage that represent modal substitutions.

1. OBJECTIVES AND SCOPE

This research is aimed at developing a dynamic simultaneous equation model of car ownership and travel distances by mode as a function of income. The dynamics involve a one-year time horizon, and travel distances represent car kilometers and passenger kilometers for one-week periods. If successful, such a model can provide a better understanding of the mutual interdependencies among these variables, providing insight into questions such as:

1. To what extent do changes in car ownership cause shifts in travel among modes?
2. Which modes are positive functions of income, and which are negative, both contemporaneously and with time lags?
3. Is car ownership a function of previous demand levels for either car travel, or competing modes, or both?
4. What are the relative strengths of the causal influences on car ownership of income and built-up travel demand, contemporaneously and with time lags?
5. Which modes are complements and which are supplements, contemporaneously and with time lags?

A purported strength of the simultaneous equation model formulated to address these and other questions is that its specification is consistent with the scales of the endogenous variables: car ownership and income (in four categories) are treated as ordinal variables estimated using ordered-response probit functions; and distances by mode are treated as continuous variables censored at zero estimated using tobit transformations. Each variable is treated the same in every equation in the system, regardless of whether it is dependent or independent in a given equation. This allows a great deal of flexibility in testing alternative hypotheses of cause and effect.

The modeling approach also has several weaknesses: First, the number of exogenous

variables is extremely limited at present; the effects on car ownership and travel demand of factors such as the employment status of household members, household structure, location, etc., are outside the scope of the model. The focus is on causal interactions among the endogenous variables, rather than explanation in terms of exogenous factors. Second, there is no accounting for individual-specific disturbances, as is possible with panel data. Third, dynamic effects are limited to those of one-year duration. Finally, there is no accounting for biases due to repeated measurement or sample attrition, and there is compensation for only certain serially correlated errors.

2. THE STRUCTURAL EQUATIONS METHODOLOGY

The model is based on an extension of linear structural equations to categorical and censored variables. (Categorical and censored variables are often, as a class, referred to as non-normal variables, to express the fact that they are not normally-distributed continuous variables.) The model equation system can be expressed in terms of two subsystems: (1) the linear structural equations, and (2) the nonlinear transformation of non-normal dependent variables.

The structural equations model defines a set of p dependent variables, organized in a vector y^* , in terms of each other and in terms of a set of m exogenous x variables:

$$y^* = By^* + \Gamma x + \zeta \quad (1)$$

where B is a $(p \times p)$ parameter matrix of the structural coefficients among the y^* variables, Γ is a $(p \times m)$ parameter matrix of structural coefficients relating the dependent and exogenous variables, and ζ is a $(p \times 1)$ vector of disturbances. The variance-covariance matrix of the ζ disturbance terms is defined as

$$\Psi = \zeta \zeta' \quad (2)$$

where Ψ is a $(p \times p)$ matrix.

Each element (i, j) in the B and Γ matrices represents the direct effect of variable j on variable i . (The main diagonal of B is specified to contain only zeros.) Consequently, there is a one-to-one correspondence between equation system (1) and a flow diagram in which there is a unidirectional arrow between each variable pair with nonzero elements in the B and Γ matrices.

Under multivariate normality assumptions, it is sufficient to consider only the first two moments of y^* , conditional on any exogenous variables x . These moments are given by

$$E(y^*|x) = (I - B)^{-1}\Gamma x \quad (3)$$

and

$$\sum (y^*|x) = (I - B)\Psi(I - B)^{-1}. \quad (4)$$

3. PROBIT AND TOBIT TRANSFORMATIONS

The second submodel converts categorical and otherwise non-normally distributed observed dependent variables to the required multivariate normal forms. There are two types of non-normal variables handled this way in the present research: ordered polytomous (ordinal) variables, and continuous variables censored from below at zero. The first type includes household income, measured in four categories, and number of cars owned by a household; the second type includes distances traveled per household per week by four motorized modes: car driver, car passenger, train, and bus-tram-subway.

The transformation from an ordered polytomous observed dependent variable y with c categories ($y = 0, 1, 2, \dots, c - 1$) to an unobserved (latent) continuous variable

y^* is given in terms of an unknown set of thresholds $k_1, k_2, \dots, k_{c-1}$ (Muthén, 1984; Golob, 1988a):

$$y = \begin{cases} c - 1 & \text{if } k_{c-1} < y^* \\ c - 2 & \text{if } k_{c-2} < y^* \leq k_{c-1} \\ \vdots & \\ 1 & \text{if } k_1 < y^* \leq k_2 \\ 0 & \text{if } y^* \leq k_1 \end{cases} \quad (5)$$

In the special case of a dichotomous y variable ($c = 2$), there is only a single threshold k_1 to be determined.

For the second type of non-normal variable, continuous variables censored from below at the bound d , the transformation to the unlimited continuous latent variable y^* is given by

$$y = \begin{cases} d & \text{if } y^* - k \leq d \\ y^* - k & \text{if } y^* - k > d \end{cases} \quad (6)$$

Transformation (5) corresponds to the conventional ordered probit model (or, ordered-response probit model) (Maddala, 1983), or the conventional binomial probit model in the special case of $c = 2$ categories. Equation (6) corresponds directly to conventional tobit estimation of limited dependent observed variables (Tobin, 1958; Amemiya, 1973).

4. THE ESTIMATION PROCEDURE

The complete model is specified by eqn (1) through (6). Equation (5) holds for the categorical income and car ownership variables; eqn (6) holds for the censored travel distance variables for four modes; and eqns (1) through (4) capture all variable relationships. The model parameters are the elements selected to be free in any or all of the three parameter matrices: B , Γ , and Ψ .

Estimation proceeds in two stages. In the first stage the sample statistics are estimated. It is useful to summarize these sample statistics in three parts: a mean/threshold/reduced-form regression intercept part, a reduced-form regression slope part, and a covariance/correlation part.

In the LISCOMP estimation procedure (Muthén, 1987) used in the present study, the unknown thresholds in (5) and (6) and regression coefficients in (3) are determined by maximizing the probabilities that the latent variables are normally distributed conditional on any exogenous x variables in the structural equation system. For an ordered polytomous observed variable (5), the estimation parameters are expressed in

$$\begin{aligned} P(y = j|x) &= P(k_j \leq y^* < k_{j+1}) \\ &= \Phi[(k_{j+1} - \pi'x) - (k_j - \pi'x)] \end{aligned} \quad (7)$$

where x is an $(m \times 1)$ vector of exogenous variables, π is a vector of reduced form regression coefficients (intercepts and slopes),

$$E(y^*|x) = \pi'x, \quad (8)$$

and Φ denotes the standard normal distribution function.

Equation (8) is the reduced form regression equation, and the first two parts of the vector of sample statistics can be estimated consistently using univariate probit/tobit regressions. Part three of the sample statistics, the covariance-correlation structure, can be solved conditionally on the parts one and two estimates. However, when multiple

categorical variables are involved, the full information approach using maximum likelihood leads to heavy computations. Therefore, different approaches have been attempted, and in the LISCOMP approach, Muthén (1981) uses a limited information technique: For p endogenous variables, $p(p - 1)/2$ correlation coefficients among the y^* variables (called "tetrachoric" correlations) are estimated using only bivariate sample information.

The sample statistics are consistent estimates of the corresponding population vector σ . The elements in σ can be expressed in terms of the structural model parameters B , Γ , and Ψ following eqns (3) and (4). The second estimation stage involves finding optimal values of the free parameters in B , Γ , and Ψ that replicate the sample statistics s as closely as possible. In the LISCOMP approach, a weighted least-squares fitting function is used with a general, full-weight matrix W :

$$F = 1/2(s - \sigma)^T W^{-1}(s - \sigma). \quad (9)$$

If W is the (asymptotic) covariance matrix of the sample statistics s , then the generalized least-squares (GLS) estimator is obtained. It has been shown that the GLS estimator for this problem is asymptotically distribution free and thus appropriate for non-normal observed variables (Browne, 1982, 1984; Bentler, 1983a, 1983b; Muthén, 1983, 1984). Objective function (9) is minimized using a modified Fletcher-Power algorithm.

The statistic nF is chi-square distributed (n denoting the sample size) (Brown, 1974) with degrees-of-freedom equal to the number of elements in s (σ) minus the number of free parameters in B , Γ , and Ψ . This statistic provides a means of evaluating model goodness-of-fit by testing whether or not the model-replicated statistics are an accurate representation of the sample statistics (generally variance-covariances), but this test is subject to the usual sample-size problems associated with all chi-square tests (Bentler and Bonett, 1980). Another important use of the statistic, however, is in hypothesis testing: For nested or hierarchical models in which one model is a specialization (more restrictive) than another, the difference in model chi-square statistics is itself chi-square distributed with degrees-of-freedom equal to the difference in the number of free parameters in the two models. Such hypothesis testing is used frequently in the present research.

5. RELATED MODELS

Structural equation modeling with latent variables has been extensively applied in the fields of sociology and psychology, and more recently in marketing research (overviews being provided by Bentler, 1980; Fornell and Larcker, 1981; Hayduk, 1987; Jöreskog and Wold, 1982). Until recent developments in estimation methods for non-normal variables (Bentler, 1983a, 1983b; Browne, 1984; Maddala, 1983; Muthén, 1979, 1983, 1984; Winship and Mare, 1983), applications were generally based on assumptions of continuous dependent variables with multivariate normal distributions, and extensive use was made of two-stage and three-stage least-square estimations and normal-theory maximum likelihood estimation, the latter method being implemented in the LISREL program (Jöreskog and Sörbom, 1984, 1987). Early applications of structural equation models, including the special case of path analysis models, are provided in the field of travel demand modeling by Tardiff (1977), Dobson *et al.* (1978), den Boon (1980) and Golob and Zondag (1983).

The present research attempts to model both longitudinal and cross-sectional relationships among travel demand variables using panel data. It has been shown in other fields that structural equation models are particularly effective in capturing causal relationships when applied to panel data (Jöreskog and Sörbom, 1977; Jöreskog, 1979; Bentler, 1984; Arminger, 1987). This success is due partially to the ability to incorporate and test alternative cross-lagged effects: Does y_i at $t = 1$ influence y_j at $t = 2$, or does y_j at $t = 1$ influence y_i at $t = 2$?

In the field of travel demand analysis, Lyon (1981, 1984) developed a structural attitude-behavior model using panel data; the model accounted for non-normality in a single discrete choice variable through the use of instrumental variables and a serially-correlated error structure in a recursive estimation technique of the type reviewed in Maddala (1983, chapter 5). Golob and Meurs (1987, 1988) developed structural equation models of temporal changes in demand for five different models, using panel data with three waves six months apart, but applied only continuous-variable normal theory in estimating models for both trip rates and travel time expenditures. Golob (1988a) explored the uses of structural equation modeling with multiple non-normal variables in models of travel choice behavior but stopped short of producing a comprehensive model. In an application relevant to the present research, Kitamura (1987, 1988) applied a path model incorporating a probit response variable in analyzing the dynamic causal relationships between car ownership and overall mobility, measured in terms of total trips.

As an alternative to structural equation models with probit specifications of non-normal variables, it is also possible to estimate imposed causal relationships among multiple categorical endogenous variables using log-linear models applied to contingency tables (Goodman, 1972, 1973; Fienberg, 1980; Maddala, 1983). Such models represent systems of logit equations because log-linear and logit model specifications can be shown to be equivalent (Fienberg, 1980). Generally, these models are limited to recursive rather than simultaneous estimation (Maddala, 1983), and it is pointed out in Rosenthal (1980) and Winship and Mare (1983) that an inconsistency occurs in such models when a discrete variable appears as a logit transform of a probability in its dependent state and as a dummy variable in its independent state in the same system of equations. One of the basic problems is that there is no multivariate logistic distribution with logistic marginal distributions that have unconstrained correlation coefficients (Gumbel, 1961). Further limitations of log-linear/logit models with multiple endogenous variables involve difficulties in incorporating continuous variables (Fienberg, 1975) and problems with distinguishing structural association from purely statistical association through lack of free parameters (Heckman, 1978; van Wissen and Golob, (1988).

6. DATA

The data are from an ongoing national panel in the Netherlands instituted in 1984 with the goal of supporting studies of changes in the mobility of the Dutch population over time. The sample of approximately 1,800 households is stratified by life-cycle group, income, and community type, and is clustered in about twenty communities spread throughout the Netherlands. The survey structure and the general use of its data are described in Golob *et al.* (1985), and Meurs and van Wissen (1987).

The present study uses data from four waves of the Dutch panel conducted in the spring of each of the years 1984, 1985, 1986, and 1987. Each wave involved a household questionnaire and separate questionnaires and seven-day travel diaries for all household members over eleven years of age. (However, for the present study, the sample was restricted to persons eighteen years or older who are eligible to drive.) Total distances traveled during the seven diary days were calculated for each person for each of four modes: car driver, car passenger, train, and bus-tram-metro (subway).

A pooled wave-pair sample is used in the present study. The six variables are defined at each of two points in time for all persons in an adjacent pair of panel waves one year apart, and the wave pairs were pooled to form a sample of 7,238 person-wave pairs. This assumes that the same causal phenomena operate in each pair of adjacent waves. The pooled sample was then randomly divided into half to provide subsamples for estimation and testing.

The twelve variables, six state variables at each of the two points in time one year apart, are defined in Table 1.

Table 1. Variable definitions

<u>VARIABLE</u>	<u>AT TIME</u>	<u>VARIABLE TYPE</u>	<u>DESCRIPTION</u>
Income	T ₁	Ordinal - 4 categories	Household Income in four categories at Time 1
Car ownership	T ₁	Ordinal - 3 categories	Number of cars owned (0, 1, 2+) at Time 1
Distance by car driver	T ₁	Continuous, censored at 0	Distance traveled by car driver in 7 days at Time 1
Distance by train	T ₁	Continuous, censored at 0	Distance traveled by train in 7 days at Time 1
Distance by bus-tram-subway	T ₁	Continuous, censored at 0	Distance traveled by bus-tram-subway in 7 days at Time 1
Distance by car passenger	T ₁	Continuous, censored at 0	Distance traveled by car passenger in 7 days at Time 1
Income	T ₂	Ordinal - 4 categories	Household income in four categories at Time 2
Car ownership	T ₂	Ordinal - 3 categories	Number of cars owned (0, 1, 2+) at Time 2
Distance by car driver	T ₂	Continuous, censored at 0	Distance traveled by car driver in 7 days at Time 2
Distance by train	T ₂	Continuous, censored at 0	Distance traveled by train in 7 days at Time 2
Distance by bus-tram-subway	T ₂	Continuous, censored at 0	Distance traveled by bus-tram-subway in 7 days at Time 2
Distance by car passenger	T ₂	Continuous, censored at 0	Distance traveled by car passenger in 7 days at Time 2

7. MODEL SPECIFICATION

All of the twelve variables described in Table 1 were defined to be endogenous, reducing equation system (1) to

$$y^* = By^* + \zeta \quad (10)$$

where $p = 12$, so that y is a (12×1) vector, B is a (12×12) matrix of structural parameters (with zeros on the main diagonal), and ζ is a (12×1) vector of disturbance terms with a variance-covariance matrix $\Psi = \zeta\zeta'$. Because the twelve endogenous y^* variables represent six variables measured at two points in time, the y^* vector in (10) can be partitioned in half, the two halves being denoted y_1^* and y_2^* ($i = 1, \dots, 6$), and the B matrix can be correspondingly partitioned according to the scheme introduced in Golob and Meurs (1988):

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \quad (11)$$

where

- B_{11} = contemporaneous (synchronous) relationships among the y_i^1 variables ($i = 1, \dots, 6$) at time T_1 ;
- B_{21} = diachronic (lagged) relationships between variables y at time T_1 and variables Y_i^2 at time T_2 , one year later;
- B_{12} = reverse diachronic (leading) relationships between variables y_i^2 at time T_2 and variables y_i^1 at time T_1 , generally a null submatrix; and
- B_{22} = contemporaneous relationships among the y_i^2 variables at time T_2 .

It is postulated that the contemporaneous relationships expressed in the B_{11} and B_{22} matrices be identical in terms of zero and nonzero effects (elements) at the two points in time. The consistent contemporaneous effects were determined from previous results from dynamic structural equation models of trip generation by mode (Golob and Meurs, 1987, 1988) and from cross-sectional models of the influences of income and car ownership on (household) trip generation by mode (Golob, 1988b; Kitamura, 1987, 1988). The postulated nonzero contemporaneous effects, together with the expected signs of the parameters, are depicted in the flow diagram of Fig. 1.

There are thirty possible direct effects in each of the two submatrices B_{11} and B_{22} (all ordered pairs among the six variables at each time point, excluding the main-diagonal elements representing a variable's effect on itself). It is postulated that only ten of these are nonzero, as shown by the ten arrows in Fig. 1, which can be interpreted as follows:

1. Income is contemporaneously exogenous.
2. Car ownership is a direct (+) function of income only.
3. Distance traveled by the car driver mode is a (+) direct function of car ownership only; it is an (+) indirect function of income through the intermediate car ownership variable.
4. Distance traveled by train is a (+) direct function of income and a (-) direct function of both car ownership and car-driver distance; it is consequently a (-) indirect function of income through the intermediate car ownership and car-driver distance variables. (The sign of the overall total effect of income on train distance depends on the relative magnitudes of the direct effects.)
5. Distance traveled by bus-tram-subway is a (+) direct function of train distance, but (-) direct functions of car ownership and car-driver distance; it is thus indirectly affected by income through various paths.
6. Finally, distance traveled by the car-passenger mode is a (+) direct function of car ownership but a (-) direct function of car-driver distance; it is thus an indirect function of income (the sign of which is determined by the relative magnitudes of the effects) and it is contemporaneously independent of train distance and bus-tram-subway distance.

These postulated contemporaneous relationships are based on the concept that both ownership and use of cars influences demands for other (nondriver) modes. They are also based on the common belief that train and bus-tram-subway are complementary modes, but, of the pair, only train use is a positive function of income, controlling for car ownership and use.

It is further postulated that the diachronic relationships in the B_{21} submatrix be restricted in the base model to the inertial relationships for each variable: y_i^1 to y_i^2 , for all $i = 1, \dots, 6$. These six relationships, expected to be relatively strong and positive, cover the variable autocorrelations and account for certain serially correlated errors as well.

The base model thus is specified with twenty-six parameters: ten identical synchronous effects at each of two points in time, and six inertial longitudinal effects. It is expected that several cross-lagged effects, representing noninstantaneous shifts in travel behavior, will be necessary to model fit if there are time lags in the causal relationships among

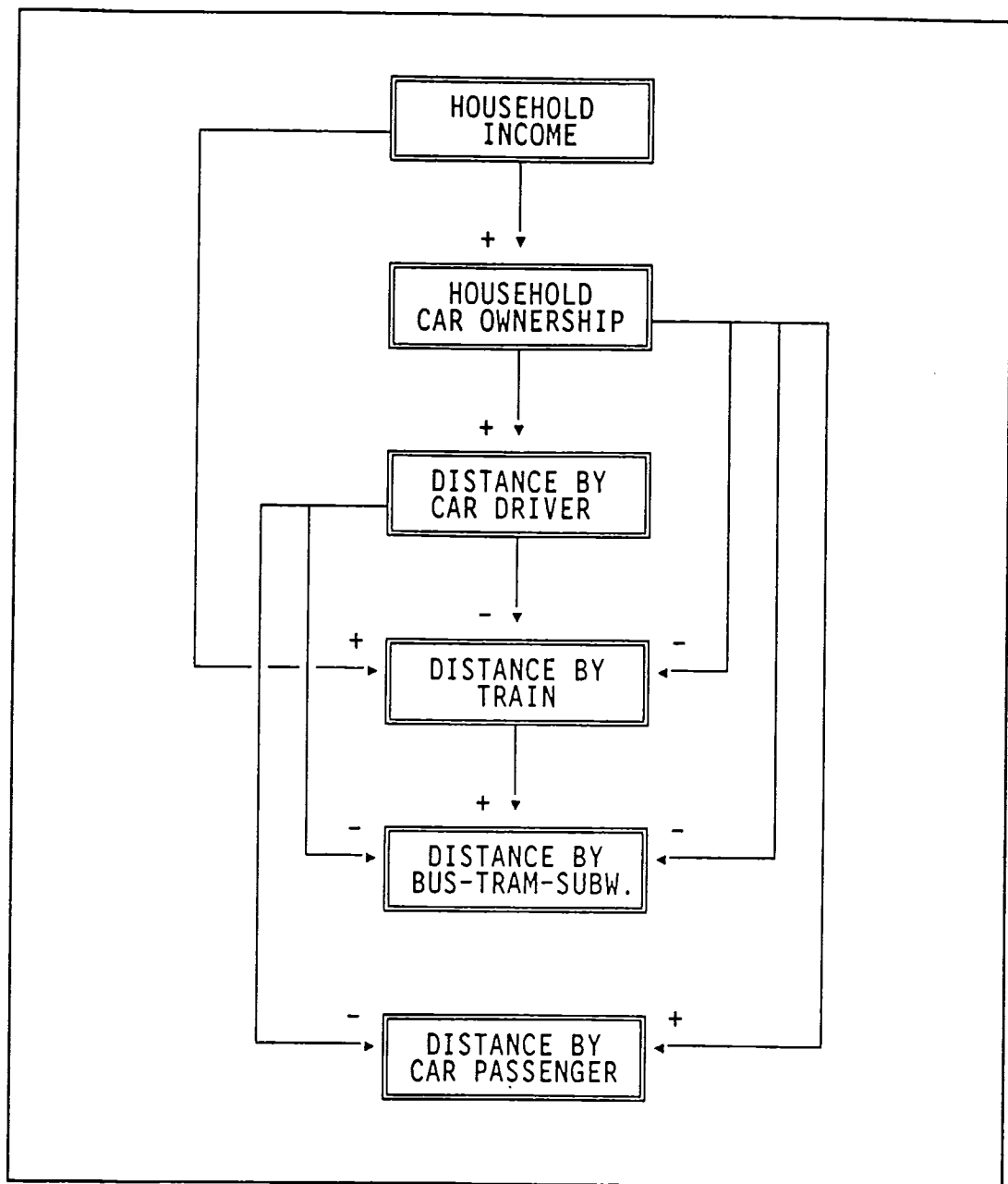


Fig. 1. Postulated synchronous relationships among the six endogenous variables at each point in time.

income, car ownership, and distance generation by mode. If such behavior is nonstationary (Kitamura, 1986), as argued by Goodwin (1987) and Golob (1988b), either cross-lagged effects will be required, or the cross-sectional structure will be different at the two points in time, or both lagged effects and restructuring of the synchronous effects will be needed. The need to add effects or modify the postulated structure was determined on the basis of hypothesis tests of nested models.

8. PARAMETER ESTIMATES

The final model involves thirty-seven parameters, eleven more than the twenty-six in the postulated base model. The synchronous and inertial structure of the base model is unchanged in the final model. The χ^2 value for the final model, determined as n (sample size) times the objective function (9), is 46.761 with 29 degrees of freedom (where degrees of freedom is given by the difference between 66 free elements in the covariance matrix of the 12 endogenous variables and the 37 free parameters in the

model). This corresponds to a probability value of $p = .0197$, indicating that the model cannot be rejected at the 99% confidence level. The root-mean-square residual fit between the sample and model covariance matrices (Golob and Meurs, 1988) is 0.027, indicating a good fit of the model-replicated and sample statistics.

All thirty-seven model parameters are significant at the 99% confidence level. The parameter estimates for the synchronous relationships at time T_1 (the B_{11} parameters) are listed in Table 2, and the estimates for the same synchronous relationships at time T_2 (the B_{22} parameters) are listed in Table 3. The two sets of synchronous parameters are different because of initial conditions (Heckman, 1981), the T_2 parameters being conditional on the simultaneously-estimated diachronal relationships.

The model is estimated using the correlation matrix form of the general variance-covariance structure, the variances of the ordinal dependent variables being standardized to one (Maddala, 1983), so the parameter estimates are standardized, allowing direct comparisons among them. The strongest of the T_2 relationships (Table 3) is the (positive) effect of train use on bus-tram-subway use, the (negative) effect of car driver use on bus-tram-subway use, the (positive) effect of car ownership on car use, and the (negative) effect of car ownership on train use.

The parameter estimates for the inertial relationships for each of the six variables across time are listed in Table 4. Overall, this is the strongest set of relationships on the structural equations, indicating a relatively high degree of temporal stability, particularly in the income variable. Train and car driver demand also exhibit considerable stability. The least stable variable is distance by car passenger, confirming the results of the dynamic analyses of modal trip rates reported in Golob and Meurs (1987).

Finally, the parameter estimates for the cross-lagged relationships are listed in Table 5. These relationships can be interpreted as follows.

1. Income has a positive causal effect on future car ownership (i.e. car ownership one time period later).
2. Future income has a positive causal effect on present car ownership. The logical

Table 2. Parameter estimates for the synchronous relationships at time T_1 (elements of the B_{11} submatrix)

LINK FROM:		LINK TO:		PARAMETER	
Variable	At Time	Variable	At Time	Value	Z-Statistic
Income	T1	Car ownership	T1	0.407	21.107
Income	T1	Distance by train	T1	0.202	6.099
Car ownership	T1	Distance by car driver	T1	0.394	36.540
Car ownership	T1	Distance by train	T1	-.362	-13.509
Car ownership	T1	Distance by bus-tram-subway	T1	-.092	-5.297
Car ownership	T1	Distance by car passenger	T1	0.141	7.183
Distance by car driver	T1	Distance by train	T1	-.151	-5.584
Distance by car driver	T1	Distance by bus-tram-subway	T1	-.275	-16.328
Distance by car driver	T1	Distance by car passenger	T1	-.197	-12.395
Distance by train	T1	Distance by bus-tram-subway	T1	0.249	10.247

Table 3. Parameter estimates for the synchronous relationships at time T_2 (elements of the B_{22} submatrix)

LINK FROM:		LINK TO:		PARAMETER	
<u>Variable</u>	<u>At Time</u>	<u>Variable</u>	<u>At Time</u>	<u>Value</u>	<u>Z-Statistic</u>
Income	T2	Car ownership	T2	0.082	9.408
Income	T2	Distance by train	T2	0.047	3.004
Car ownership	T2	Distance by car driver	T2	0.130	15.952
Car ownership	T2	Distance by train	T2	-.110	-3.400
Car ownership	T2	Distance by bus-tram-subway	T2	-.038	-3.879
Car ownership	T2	Distance by car passenger	T2	0.045	3.228
Distance by car driver	T2	Distance by train	T2	-.051	-2.492
Distance by car driver	T2	Distance by bus-tram-subway	T2	-.145	-10.996
Distance by car driver	T2	Distance by car passenger	T2	-.085	-7.126
Distance by train	T2	Distance by bus-tram-subway	T2	0.174	9.555

explanation for this is that present car ownership decisions are based partially on anticipated future income.

3. Car ownership has a positive causal effect on future car vehicle kilometers, and a negative causal effect on future bus-tram-subway passenger kilometers. There are lagged effects from car ownership on both car use and use of the competing bus-tram-subway mode.

Table 4. Parameter estimates for the inertial relationships (diagonal elements of the B_{21} submatrix)

LINK FROM:		LINK TO:		PARAMETER	
<u>Variable</u>	<u>At Time</u>	<u>Variable</u>	<u>At Time</u>	<u>Value</u>	<u>Z-Statistic</u>
Income	T1	Income	T2	0.810	122.411
Car ownership	T1	Car ownership	T2	0.409	21.071
Distance by car driver	T1	Distance by car driver	T2	0.542	52.950
Distance by train	T1	Distance by train	T2	0.551	24.876
Distance by bus-tram-subway	T1	Distance by bus-tram-subway	T2	0.426	29.663
Distance by car passenger	T1	Distance by car passenger	T2	0.296	23.792

Table 5. Parameter estimates for the cross-lagged relationships (off-diagonal elements of the B_{12} and B_{21} submatrix)

LINK FROM:		LINK TO:		PARAMETER	
<u>Variable</u>	<u>At Time</u>	<u>Variable</u>	<u>At Time</u>	<u>Value</u>	<u>Z-Statistic</u>
Income	T1	Car ownership	T2	0.141	7.196
Income	T2	Car ownership	T1	0.083	7.574
Car ownership	T1	Distance by car driver	T2	0.074	9.002
Car ownership	T1	Distance by bus-tram-subway	T2	-.058	-3.649
Distance by car driver	T1	Car ownership	T2	0.178	15.638
Distance by car driver	T1	Distance by car passenger	T2	-.076	-4.891
Distance by train	T1	Car ownership	T2	-.109	-5.215
Distance by train	T1	Distance by car driver	T2	0.090	4.241
Distance by bus-tram-subway	T1	Car ownership	T2	-.040	-3.030
Distance by bus-tram-subway	T1	Distance by car driver	T2	-.039	-3.118
Distance by car passenger	T1	Car ownership	T2	0.062	4.476

4. Car usage (distance by both the car driver and car passenger modes) has a lagged positive effect on car ownership.

5. However, present car driver demand implies reduced future car passenger demand.

6. Present train passenger kilometers has a negative causal effect on future car ownership, but a positive effect on future car kilometers, implying train users adjust car ownership downward, but increase car mobility relative to that adjustment.

7. Finally, present bus-tram-subway passenger kilometers has a negative effect on both future car ownership and future car kilometers.

9. DIRECT EFFECTS

Each of the parameters in the structural equations model represents the direct effect of one variable on another. The results of Tables 2 through 5 can thus be interpreted by investigating each endogenous variable in terms of the direct effects to it from all other variables. As each effect corresponds to a unidirectional arrow in a flow diagram, it is instructional to visualize such direct effects in terms of the portions of a flow diagram that represent inputs to the variable in question. The focus here is limited to variables at the second point in time, T_2 .

Income at time T_2 is a direct function only of income at T_1 , and the magnitude of the effect is relatively strong, with a standardized beta parameter value of 0.810. (The corresponding flow diagram is not shown, as it consists of only a single arrow.)

In contrast to income, there are seven direct effects on car ownership at time T_2 , the most for any endogenous variable, and these are shown in Fig. 2. Only one of these

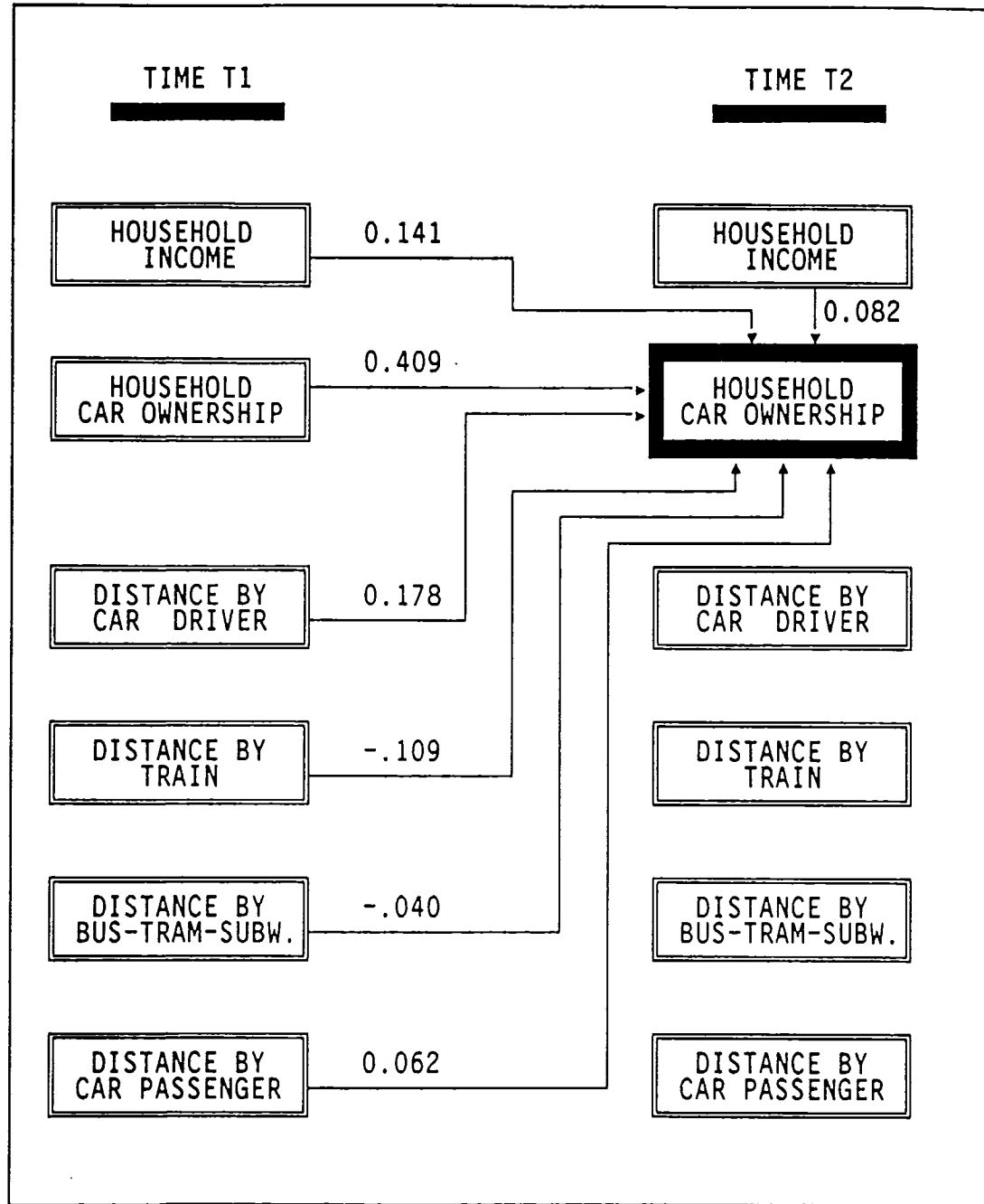


Fig. 2. Direct effects on household car ownership at time T_2 .

effects is from a contemporaneous variable, income; one represents inertia from car ownership in the previous period; and the remaining five are cross-lagged effects, one from each of the other variables in the previous period. The conclusion is that car ownership is relatively stable over time, is sensitive to past, present, and future income, and is sensitive to previous, but not contemporaneous, demand for four motorized modes. Demands for car driver and car passenger travel increase the probability of higher levels of car ownership, with car driver demand being relatively more important. Alternatively, demands for train and bus-tram-subway decrease the probability of higher car ownership levels, with train demand being more important. Of the two pairs of direct effects, the lagged effects of car usage are more important than the lagged effects of public transport usage.

The direct effects on demand for the car driver mode at time T_2 are depicted in Fig. 3. In this case, there are three lagged effects in addition to the inertial effect and

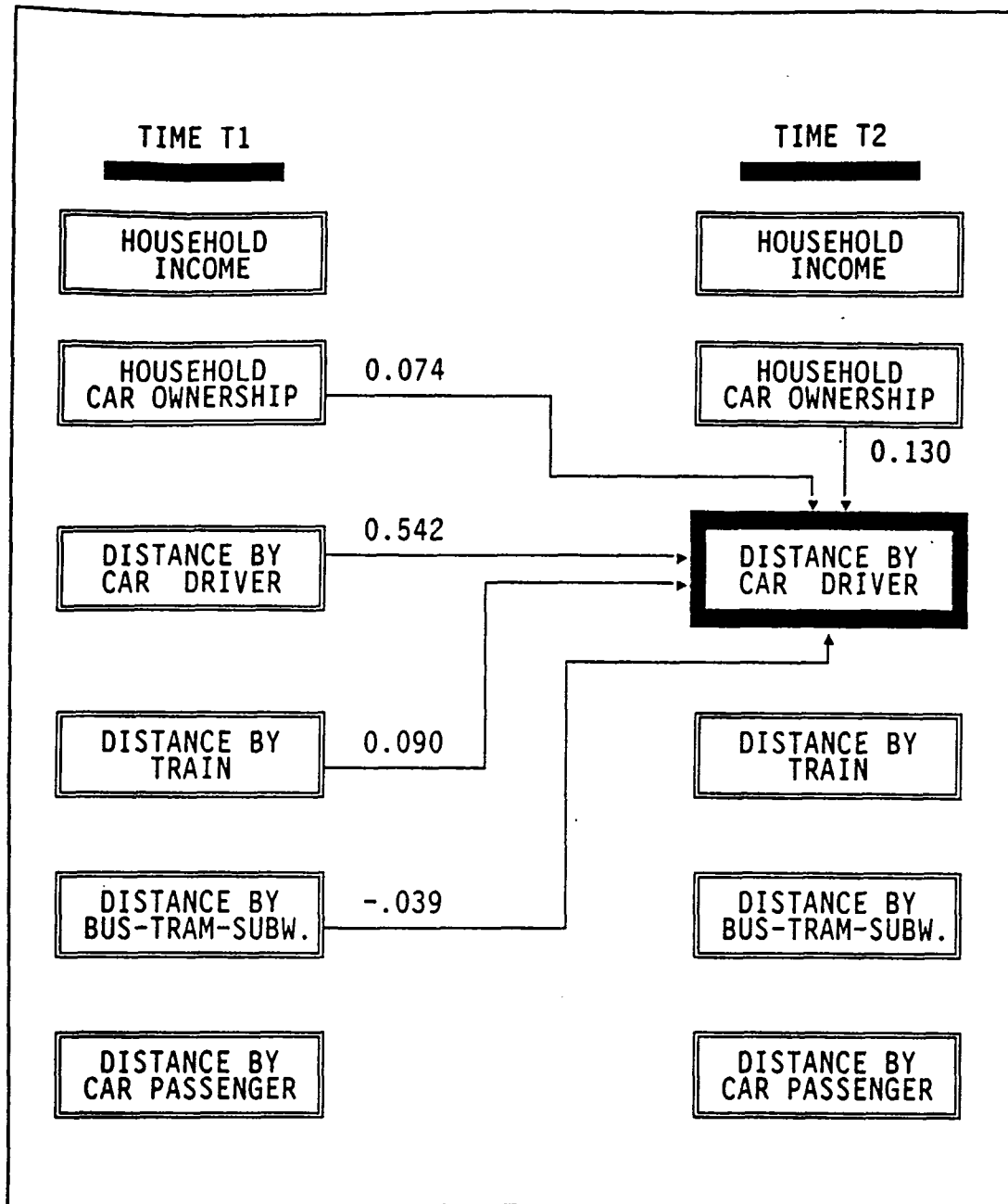


Fig. 3. Direct effects on distance traveled by car driver at time T_2 .

one contemporaneous car-ownership influence. The lagged effects are from previous car ownership, train usage, and bus-tram-subway usage. The positive causal effect from previous train demand is unexpected and indicates complementary mobility in terms of these modes.

In contrast to car ownership and car kilometers traveled, there are no cross-lagged direct effects on distance traveled by train (Fig. 4). In addition to the inertial effect, there are only contemporaneous direct influences on train demand from income (positive), car ownership, and car kilometers traveled (both negative). Adjustments in train demand are channeled through adjustments in car ownership and car driver demand.

There are five direct effects on the fifth variable, distance traveled by bus-tram-subway (Fig. 5). Three of these effects are contemporaneous, being from car ownership (negative), car kilometers (negative), and train passenger kilometers (positive). In addition there is a lagged effect from previous car ownership.

Finally, there are four direct effects on distance traveled by car passenger (Fig. 6). In addition to the relatively weak inertial effect, there is a contemporaneous positive

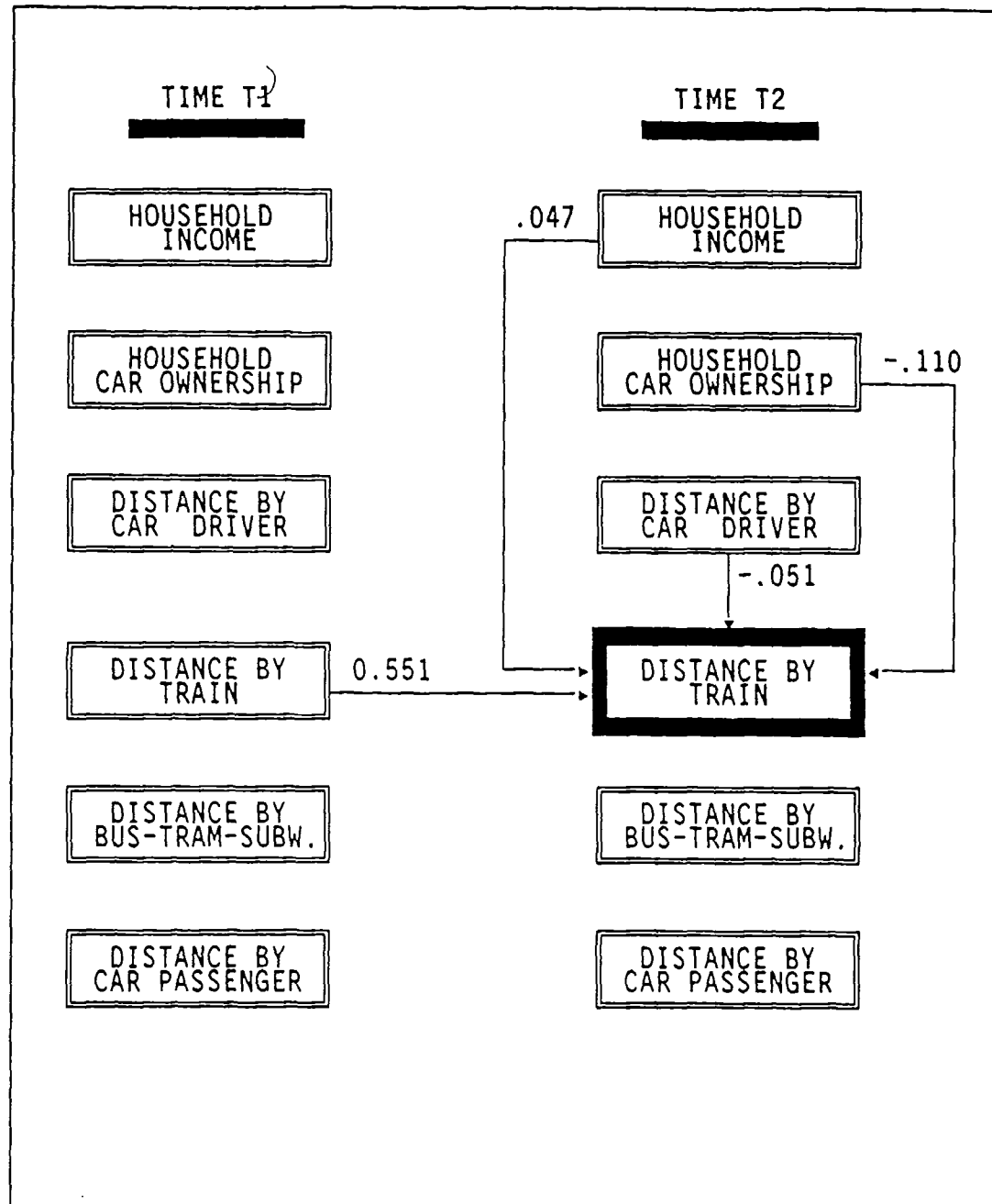


Fig. 4. Direct effects on distance traveled by train at time T_2 .

effect from car ownership and both contemporaneous and lagged negative effects from car driver demand.

10. TOTAL EFFECTS

The total effect of one variable on another variable is the sum total of any direct effect and all the indirect effects represented by paths through intermediate variables. For the structural equations system of (11), the matrix of total effects of every endogenous variable on every other endogenous variable is given by

$$A = (I - B)^{-1} - I \quad (12)$$

where A (Alpha) is a $(p \times p)$ matrix with elements α_{ij} measuring the total effect of variable y_j on variable y_i . These are the coefficients of the reduced-form equations.

The Alpha matrix of total effects for the present model is reproduced in Table 6, where only nonzero α_{ij} effects are listed. The following interpretations can be made of the effects on the variables at the second point in time, given by the entries in the bottom six rows of the matrix of Table 6.

1. Income at time T_2 is effected only by previous income (the total effect being equal to the direct effect). Income is relatively stable over yearly measurement intervals.

2. Car ownership at time T_2 is a function of all six of the endogenous variables at time T_1 plus income at time T_2 . In addition to the expected strong effects from income and previous car ownership, there are substantial lagged causal influences on car ownership from demand for (positive) car driver travel and (negative) train demand. There are also weaker lagged total effects from (positive) car passenger demand and (negative) bus-tram-subway demand. These important lagged causal influences on car ownership indicate that car ownership decisions are made in response to demand for car travel and in opposite (and weaker) response to public transport demand; and there is a time lag in these responses.

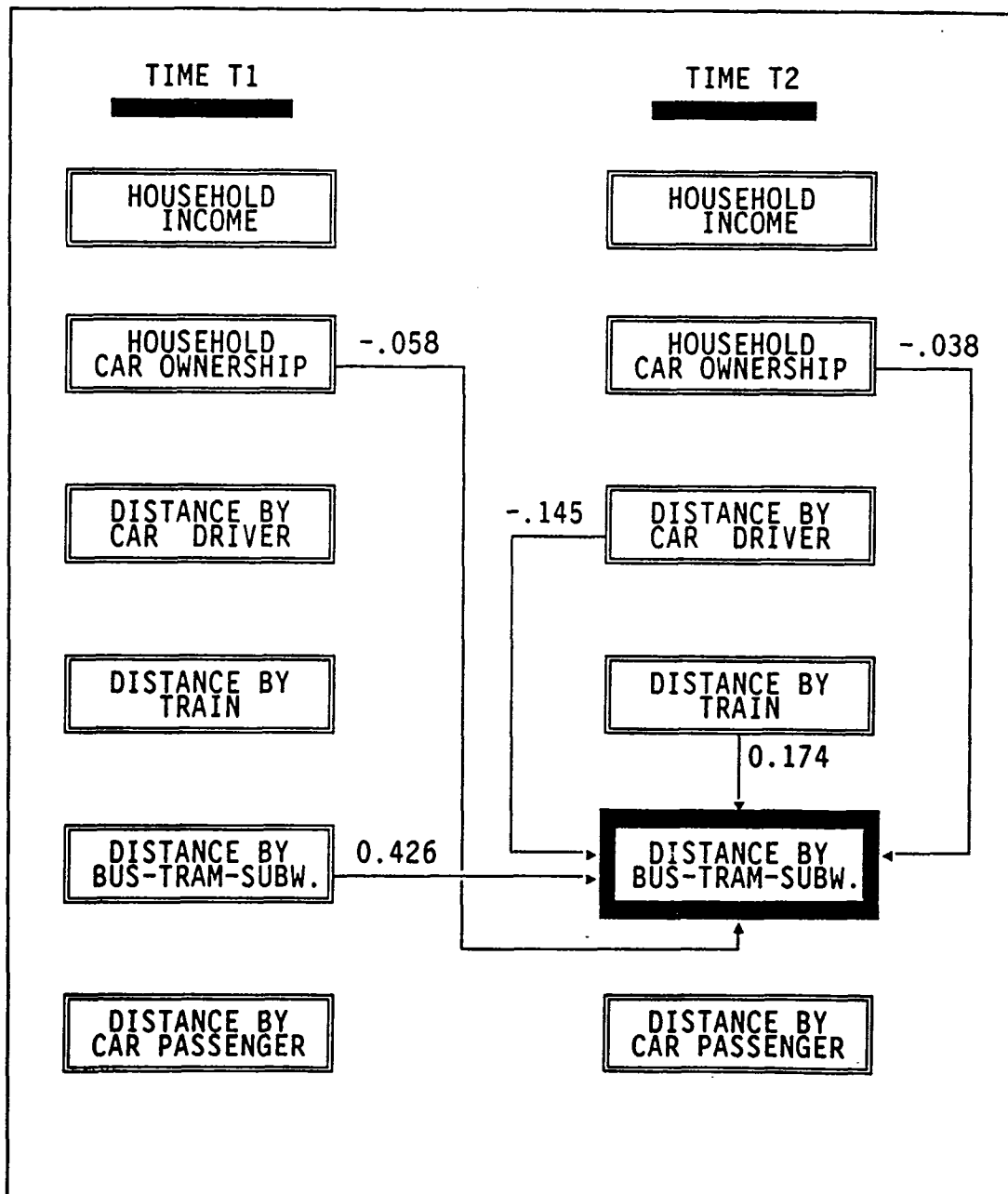


Fig. 5. Direct effects on distance traveled by bus-tram-subway time T_2 .

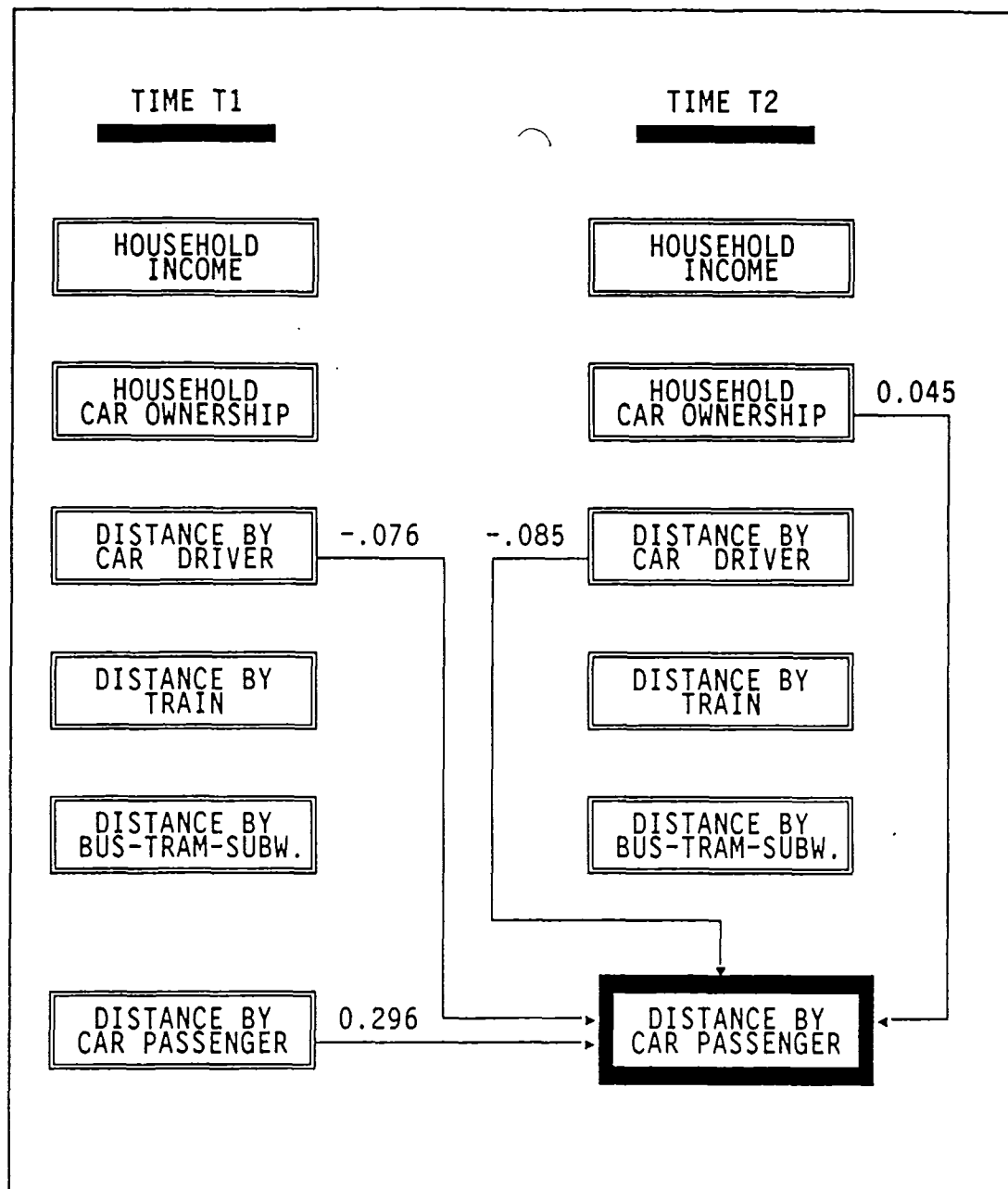


Fig. 6. Direct effects on distance traveled by car passenger at time T_2 .

3. Car driver demand at time T_2 is primarily a function of prior and current car ownership and prior income, in addition to its own inertia (temporal stability). However, it is also a positive function of lagged train demand, indicating a lagged substitution effect: Train users tend to become drivers to some degree.

4. Train demand at time T_2 , which has a temporal stability similar to car driver demand, is primarily a (negative) function of prior and contemporaneous car ownership and prior and contemporaneous car driver demand. It is approximately independent of income due to compensatory (positive) direct and (negative) indirect income effects.

5. Bus-tram-subway demand, which is less stable than car driver and train demand, is a moderately strong negative function of all variables with the exception of prior and contemporaneous train demand, of which it is a moderately strong positive function, and car passenger demand, of which it is approximately independent.

6. Finally, car passenger demand exhibits the lowest level of temporal stability and

Table 6. Total effects of matrix A , where α_{ij} = total effect of variable j on variable i (nonzero elements only)

		TIME T1					TIME T2						
		Distance					Distance						
		Income	Car Owner- ship	by Car Driver	by Train	by Bus- Tram- Subway	by Car Pass- enger	Income	Car Owner- ship	by Car Driver	by Train	by Bus- Tram- Subway	by Car Pass- enger
	Income												
T	Car Ownership	0.474						0.083					
I	Distance by Car Driver	0.187	0.394					0.033					
M													
E	Distance by Train	0.002	-.421	-.151				-.035					
T	Distance by Bus- Tram- Subway	-.094	-.305	-.315	0.249			-.025					
1													
	Distance by Car Passenger	0.030	0.063	-.197				0.005					
	Income	0.810											
T	Car Ownership	0.440	0.541	0.195	-.119	-.040	0.062	0.127					
I	Distance by Car Driver	0.197	0.332	0.556	0.065	-.044	0.008	0.038	0.130				
M													
E	Distance by Train	-.019	-.309	-.133	0.561	0.007	-.007	0.012	-.117	-.051			
T	Distance by Bus- Tram- Subway	-.116	-.310	-.246	0.199	0.435	-.005	-.024	-.077	-.154	0.174		
2													
	Distance by Car Passenger	-.002	-.015	-.174	-.011	0.002	0.298	0.002	0.034	-.085			

is moderately well explained in terms of only two variables: prior and contemporaneous car driver demand, which influence car passenger demand negatively.

11. HYPOTHESIS TESTING

The base model, restricted to only synchronous and inertial effects, and the final model, with the same synchronous and lagged effects and eleven additional cross-lagged effects, are nested models. The hypothesis that the final model represents a statistically significant improvement over the base model can be tested in terms of the difference in model χ^2 values with eleven degrees of freedom. The χ^2 values for the base model and the final model are 688.048 and 46.761, respectively, and the difference of 641.287 indicates that the final model is a highly significant improvement over the base model.

The level of significance of the improvement in model fit associated with each of the eleven cross-lagged effects that were added to the base model in creating the final model can be assessed by applying the test of nested models to the final model and a

less restrictive model in which one of the links is deleted. This was done for all eleven cross-lagged effects, with the results shown in Table 7. A χ^2 difference of 6.637 or greater indicates that the parameter leads to a significant improvement in model fit (at the $p = .01$ level), and all parameters pass this test.

The χ^2 difference values in Table 7 also provide a relative indicator of the importance of each of the parameters to overall model fit. Measured in this way, the link from distance traveled by car driver in time T_1 to car ownership in time T_2 is the most important of the eleven cross-lagged effects, followed by the effects on car ownership of lagged and anticipated income. The least important effects are from distance by bus-tram-subway in time T_1 to car ownership and car driver distance in T_2 . However, even if the least important effect is deleted, the model fit declines to the point where it can be rejected at the $p = .01$ level ($\chi^2 = 55.850$ with 30 degrees of freedom, which corresponds to a probability value of $p = .0028$); all of the cross-lagged effects are required for model fit.

12. CONCLUSIONS AND DIRECTIONS FOR FURTHER RESEARCH

It can be concluded that the dynamic interrelationships among income, car ownership level, and demand for four motorized modes, measured in terms of distances traveled per week, can be successfully modeled using a simultaneous equation structure.

Table 7. Nested model χ^2 tests of significance of each cross-lagged effect (* = parameter significant at $P = .01$ Level)

EFFECT DELETED					
LINK FROM:		LINK TO:		TEST STATISTIC	
Variable	At Time	Variable	At Time	χ^2	χ^2 Difference
Income	T1	Car ownership	T2	106.157	59.396*
Income	T2	Car ownership	T1	103.798	57.037*
Car ownership	T1	Distance by car driver	T2	140.093	93.332*
Car ownership	T1	Distance by bus-tram-subway	T2	59.988	13.227*
Distance by car driver	T1	Car ownership	T2	262.154	215.393*
Distance by car driver	T1	Distance by car passenger	T2	83.619	36.858*
Distance by train	T1	Car ownership	T2	74.094	27.333*
Distance by train	T1	Distance by car driver	T2	64.913	18.152*
Distance by bus-tram-subway	T1	Car ownership	T2	55.850	9.089*
Distance by bus-tram-subway	T1	Distance by car driver	T2	56.484	9.723*
Distance by car passenger	T1	Car ownership	T2	66.822	20.061*

The dynamics are expressed in terms of panel survey measurements of the six variables at two points in time one year apart. Incomes and car ownership levels at the two time periods are measured in terms of ordered categories and treated in the structural equations as ordered probit response variables; the four modal distance variables are treated in terms of censored continuous variables (censored at zero) subject to tobit transformations.

Results indicate that the synchronous relationships among the variables are similar at the two points in time, but there are significant causal effects that have a one-year time lag. Notable among these lagged effects are the causal influences on car ownership and car kilometers at time T_2 of demand for other modes at time T_1 ; there appears to be substantial noninstantaneous adjustment of car ownership and use that represents modal substitutions. Also detected was an intriguing effect of (anticipated) future income on present car ownership. These results indicate that cross-sectional models are approximations of decision processes with time lags.

Further research is envisioned along three directions: First, additional explanatory variables can be added to the model structure in the form of exogenous variables (whose effects are captured in the Γ matrix in equation system (1)). Such variables might include life cycle family structure, employment status, and age—many of which would be dummy variables.

Second, the model (with or without additional exogenous variables) can be estimated for separate population segments, and statistical tests can be conducted of the equivalences of model parameters among the segments. Such group estimation and testing is a feature of most available computer programs for structural equation modeling.

Third, and finally, group analyses can be conducted for each pair of years of the Dutch National Mobility Panel (1984–1985, 1985–1986, 1986–1987, and 1987–1988) to determine whether or not the modeled travel behavior is systematically changing over the 1984–1988 period in the Netherlands. This can be useful for policy planning and evaluation.

Acknowledgments—The research of Dr. van Wissen has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences.

The authors wish to thank the Project Bureau for Integrated Transport and Traffic Studies and the Office of the Director General for Transport of the Dutch Ministry of Transport and Public Works, and Bureau Goudappel Coffeng (Deventer, the Netherlands), for providing the data and encouragement necessary for the research reported here. However, the conclusions expressed are those of the authors alone, who are also solely responsible for any errors.

REFERENCES

- Amemiya, T. (1973) Regression analysis when the dependent variable is truncated normal. *Econometrica* **41**, 997–1016.
- Arminger, G. (1987) Misspecification, asymptotic stability, and ordinal variables in the analysis of panel data. *Soc. Methods Rsch.* **15**, 336–348.
- Bentler, P. M. (1980) Multivariate analysis with latent variables: Causal modeling. *Ann. Rev. Psy.* **31**, 419–456.
- Bentler, P. M. (1983a) Some contributions to efficient statistics in structural models: Specification and estimation of moment structures. *Psychometrika* **48**, 493–517.
- Bentler, P. M. (1983b) Simultaneous equation systems as moment structure models. *J. Econ.* **22**, 13–42.
- Bentler, P. M. (1984) Structural equation models in longitudinal research. In: *Handbook of Longitudinal Research*, Vol. 1, 88–105. (Edited by S. A. Mednick, M. Harway, and K. M. Finello) Praeger, New York.
- Bentler, P. M. and Bonett, D. G. (1980) Significance tests and goodness of fit in the analysis of covariance structures. *Psych. Bull.* **88**, 588–606.
- Browne, 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.)
- Browne, M. W. (1982) Covariance structures. In: *Topics in Applied Multivariate Analysis*, 72–141. (Edited by Douglas M. Hawkins) Cambridge University Press, Cambridge.
- Browne, M. W. (1984) Asymptotically distribution-free methods for the analysis of covariance structures. *Br J Math Stat Psychol* **37**, 62–83.
- den Boon, A. K. (1980) *Opvattingen over autogenbruik en milieuvervuiling*. Baschwitz Instituut, Universiteit van Amsterdam, Nederland.

- Dobson, R., Dunbar F., Smith C. J., Reibstein D., and Lovelock C. (1978) Structural models for the analysis of traveler attitude-behavior relationships. *Transportation* 7, 351-363.
- Fienberg, S. E. (1975) Comment. *J Am Stat Assoc* 70, 521-524.
- Fienberg, S. E. (1980) *The Analysis of Cross-Classified Categorical Data*, 2nd Ed. MIT Press, Cambridge, MA.
- Fornell, C., and Larcker D. F. (1981) Structural equation models with unobserved variables and measurement error: Algebra and statistics. *J Marketing Resch*, 382-388.
- Golob, J. M., Schreurs L. J. M., and Smit J. G. (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, The Netherlands.
- Golob, T. F. (1988a) Structural equation modeling of the dynamics of travel choice behavior. Presented at 1988 Oxford Conference on Travel and Transportation, July 13-15, Oxford, England.
- Golob, T. F. (1988b) The causal influences of income and car ownership on trip generation by mode. *J Transport Econ Policy*, forthcoming.
- Golob, T. F., and Meurs H. (1987) A structural model of temporal change in multi-modal travel demand. *Trans Rsch* 21A, 391-400.
- Golob, T. F., and Meurs H. (1988) Development of structural equations models of the dynamics of passenger travel demand. *Environ Planning A*, 20, 1197-1218.
- Golob, T. F., and Zondag E. R. (1983) A causal model of mobility. Presented at the Tenth Transportation Planning Research Colloquium, Zandvoort, The Netherlands. 14-16 December 1983.
- Goodman, L. A. (1972) A generalized model for the analysis of surveys. *Am J Sociol*, 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. *Biometrika*, 60, 179-192.
- Goodwin, P. B. (1987) Family changes and public transport use. Presented at Round Table Conference on the Analysis of Panel Data, The Hague, May 14-15, 1987.
- Gumbel, E. J. (1961) Bivariate logit distributions. *J Am Stat Assoc* 56, 335-349.
- Heckman, J. J. (1978) Dummy endogenous variables in a simultaneous equation system. *Econometrica*, 46, 931-959.
- Hayduk, L. A. (1987) *Structural Equation Modeling with LISREL*. Johns Hopkins University Press, Baltimore.
- Heckman, J. J. (1981) The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. In: *Structural Analysis of Discrete Data with Econometric Applications*, 114-178, (C. F. Manski and D. McFadden editors.) Cambridge, MA: M.I.T. Press.
- Jöreskog, K. G. (1979) Statistical estimation of structural models in longitudinal-developmental investigations. In: *Longitudinal Research in the Study of Behavior and Development*, 303-351. (Edited by Nesselroade J and Baltes P.) New York, Academic Press.
- Jöreskog, K., and Sörbom D. (1977). Statistical models and methods for analysis of longitudinal data. In: *Latent Variables in Socioeconomic Models*. 285-325. Aigner D. and Goldberger A. S. Amsterdam, North-Holland.
- Jöreskog, K., and Sörbom D. (1984) *LISREL VI: An Analysis of Linear Structural Relationships by Maximum Likelihood, Instrumental Variables, and Least Squares Methods*. Mooresville, IN: Scientific Software, Inc.
- Jöreskog, K., and Sörbom D. (1987) New developments in LISREL. *Data Analyst*, 4, 1-22.
- Jöreskog, K., and Wold H. (eds.) (1982) *System under indirect Observation: Causality, Structure, Prediction*, Part I. New York: North-Holland.
- Kitamura, R. (1986) Linear panel analysis of travel behavior. Netherlands Institute of Transport, Rijswijk, the Netherlands.
- Kitamura, R. (1987) A panel analysis of household car ownership and mobility. *Proceedings of the Japan Society of Civil Engineers*, 383, 13-27.
- Kitamura, R. (1988) A dynamic model system of household car ownership, trip generation and modal split: Model development and simulation experiment. *Proceedings of the 14th Australian Road Research Board Conference, Part 3*: 96-111.
- Lyon, P. K. (1981) Time-dependent structural equations modelling of the relationship between attitudes and discrete choice behavior of transportation consumers. Unpublished PhD dissertation, Department of Industrial Engineering/Management Sciences, Northwestern University, Evanston, IL.
- Lyon, P. K. (1984) Time-dependent structural equations modeling: A methodology for analyzing the dynamic attitude-behavior relationship. *Transportation Science*, 18, 395-414.
- Maddala, G. S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.
- Meurs, H., and van Wissen L. (1987) Analysis of panel data. Presented at Round Table Conference on the Analysis of Panel Data, The Hague, May 14-15, 1987.
- Muthén, B. (1979) A structural probit model with latent variables. *J Am Stat Assoc* 74, 807-811.
- Muthén, B. (1983) Latent variable structural equation modeling with categorical data. *J 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.
- Rosenthal, H. (1980) The limitations of log-linear analysis. *Contemp Soc*, 9, 207-212.
- Tardiff, T. J. (1977) Causal influences involving transportation attitudes and behavior. *Transportation Research* 11, 397-404.
- Tobin, J. (1958) Estimation of relationships for limited dependent variables. *Econometrica*, 26, 24-36.
- van Wissen, L., and Golob T. F. (1988) A simultaneous equation system of the dynamics in modal choice. Working Paper WP-88-16, Institute of Transportation Studies, University of California, Irvine.
- Winship, C., and Mare R. D. (1983) Structural equations and path analysis for discrete data. *Am J Soc*, 89, 54-110.

- I = the identify matrix
- m = number of exogenous variables
- n = sample size
- p = number of endogenous variables
- S = the $(p \times p)$ variance-covariance matrix of the y^* endogenous variables
- x = an $(m \times 1)$ column vector of x_i exogenous variables
- y = a $(p \times 1)$ column vector of (categorical or censored) y_i endogenous variables
- y^* = a $(p \times 1)$ column vector of y_i^* latent variables transformed from the y_i variables
- A = a $(p \times p)$ matrix of total effects α_{ij} of variables y_i^* on variable y_j^*
- B = a $(p \times p)$ parameter matrix of structural coefficients β_{ij} among the y^* latent variables.
- Γ = a $(p \times m)$ parameter matrix of regression coefficients γ_{ij} of the y_i^* latent constructs on the exogenous x_i variables
- ζ = a $(p \times 1)$ vector of disturbances ζ_i in the model specifications of the y_i^* variables
- Φ = the cumulative normal distribution function
- Σ = the $(p \times p)$ variance-covariance matrix σ_{ij} of the endogenous y_i^* variables generated by the model
- Ψ = the $(m \times m)$ variance-covariance matrix of the ζ disturbance terms of the estimations of the y_i^* variables