Graduate Theses and Dissertations   Graduate School

4-22-2004

An Exploration of the Relationship between Mode Choice and Complexity of Trip Chaining Patterns

Xin Ye
University of South Florida

Follow this and additional works at: https://scholarcommons.usf.edu/etd

Part of the American Studies Commons

An Exploration of the Relationship between Mode Choice and

Complexity of Trip Chaining Patterns

by

Xin Ye

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Civil Engineering
Department of Civil and Environmental Engineering
College of Engineering
University of South Florida

Major Professor: Ram M. Pendyala, Ph.D.
Manjriker Gunaratne, Ph.D., P.E.
John Lu, Ph.D., P.E.

Date of Approval:
April 22, 2004

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

AN EXPLORATION OF THE RELATIONSHIP BETWEEN

MODE CHOICE AND COMPLEXITY OF TRIP CHAINING PATTERNS

Xin Ye

ABSTRACT

This thesis investigates the relationship between mode choice and the complexity of trip chaining patterns. An understanding of the causality between these two choice behaviors may aid in the development of tour-based travel demand modeling systems that attempt to incorporate models of trip chaining and mode choice. The relationship between these two aspects of travel behavior is represented in this thesis by considering three different causal structures: one structure in which the trip chaining pattern is determined first and influences mode choice, another structure in which mode choice is determined first and influences the complexity of the trip chaining pattern, and a third structure in which neither is predetermined but both are determined simultaneously. The first two structures are estimated within a recursive bivariate probit modeling framework that accommodates error covariance. The simultaneous logit model is estimated for the third structure that allows a bidirectional simultaneous causality. The analysis and model estimation are performed separately for work tour and non-work tour samples drawn from the 2000 Swiss Microcensus travel survey. Model estimation results show that the causal structure in which trip chaining precedes mode choice performs best for the non-work tour sample.

iv

For the work-tour sample, the findings were less conclusive because two causal structures, one in which trip chaining affects mode choice and the other in which both are determined simultaneously, gave virtually identical goodness-of-fit measures. But the structure in which mode choice precedes trip chaining pattern choice gave significantly inferior goodness-of-fit measures for the work tour sample. These findings should be reflected in the development of activity-based and tour-based modeling systems.

CHAPTER 1

INTRODUCTION

1.1 Background

Over the past few decades, there has been considerable research on people's trip chaining patterns, i.e., the propensity to link a series of activities into a multi-stop tour or journey. The analysis of trip chaining activity may lead to a better understanding of travel behavior and provide a more appropriate framework for examining various transportation policy issues (Strathman & Dueker, 1995). Indeed, the profession has seen tour-based models being developed and increasingly applied in the travel demand forecasting arena in place of the more traditional trip-based models that do not reflect trip chaining behavior and tour formation.

In this thesis, the terms trip chain and tour are used synonymously to refer to a sequence of trips that begins at home, involves visits one or more other places, and ends at home. Depending on the number of places visited within the tour or chain, the tour may be classified into two categories: simple and complex. A tour or chain with a single stop or activity outside the home location is defined as a "simple" tour, whereas a tour or chain with more than one stop outside the home location is defined as a "complex" tour.

Thus a tour or chain of the form: home → shop → home is considered a simple tour while a tour of the form: home → work → shop → home is considered a complex tour.

As people's activity patterns become increasingly complex and involve interactions with other household and non-household members and as time is a finite resource, it may be conjectured that trip chains are likely to be increasingly complex over time. The ability to chain multiple activities together in a single tour or chain may provide greater efficiency and convenience than a series of single-stop simple tours (Hensher et al., 2000). There are at least two reasons as to why this has significant traffic and policy implications. First, complex tours or chains may lead to an increase in automobile usage. If one needed to pursue complex tours or chains, then the flexibility afforded by the private automobile is desirable. The ability to pursue multiple activities in a single journey is rather limited when constrained by the schedules, routes, and uncertainty associated with public transportation. Thus, complex trip chaining may contribute to an increased auto dependency and consequently, automobile traffic. Second, in the case of workers (commuters), the formation of complex trip chains may entail the linking of non-work activities with the work trip (commute). Then, non-work trips that could have taken place outside the peak periods now occur in the peak periods simply because they are being tied together with the commute. Thus, complex trip chaining patterns may contribute to an increase in peak period travel demand.

1.2 Objective

The above discussion clearly points to the possible interdependency between trip chaining, auto usage, and trip timing. Strathman and Dueker (1995), in an analysis of the 1990 NPTS, found that complex trip chains may tend to be more auto-oriented. However, the nature of the causal relationship is not unilaterally evident because the availability of an automobile may provide the flexibility and convenience that contributes to the formation of complex trip chains. The flexibility of the automobile may stimulate the desire to undertake additional activities in one tour. For example, the lower travel times typically associated with the auto mode choice may relax time constraints and lead to more stop-making (Bhat, 1997). Moreover, shared rides, which constitute a portion of total auto mode share, are more likely to involve complex tours due to the variety of trip purposes and destinations between the driver and passengers. The central question that is being addressed in this thesis is: "Does mode choice influence the complexity of trip chaining patterns or does the complexity of the trip chaining patterns influence mode choice?". The ambiguity in the causal relationship between the complexity of trip chains and mode choice motivates this investigation. This research is aimed at understanding and quantifying the causal relationships between tour complexity and mode choice using econometric methods.

1.3 Outline

The rest of this thesis is organized as follows. Chapter 2 reviews literature concerning the causal analysis in transportation research. Chapter 3 presents the modeling methodology and formulation for the different causal structures considered in this thesis. Chapter 4

introduces the Swiss Travel Microcensus 2000 and the process by which the tour data set needed for model estimation was prepared. Model estimation results are discussed in Chapter 5. Chapter 6 presents the performance comparison across models representing three different causal structures. Conclusions are drawn and some recommendations for future research are provided in Chapter 7.

CHAPTER 2

LITERATURE REVIEW

2.1 Structural Equation Model: Causal Analysis among Continuous Variables

Structural Equations Modeling (SEM) is a powerful modeling methodology that takes a confirmatory hypothesis-testing approach to the causal analysis among continuous endogenous variables. A typical structural equations model (with 'G' continuous endogenous variables) is defined by a matrix equation system as shown in Equation 2.1.

$$\begin{bmatrix} Y_1 \\ . \\ . \\ . \\ Y_G \end{bmatrix} = \begin{bmatrix} Y & X \end{bmatrix} \begin{bmatrix} B \\ \Gamma \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ . \\ . \\ . \\ \varepsilon_G \end{bmatrix} \qquad 2.1$$

This equation can be rewritten as

$$Y = BY + \Gamma X + \varepsilon \qquad\qquad 2.2$$

$$(or) \, Y = (I - B)^{-1}(\Gamma X + \varepsilon) \qquad\qquad 2.3$$

Where Y: a column vector of endogenous variables,

B: a matrix of parameters associated with right-hand-side endogenous variables,

X: a column vector of exogenous variables,

Γ: a matrix of parameters associated with exogenous variables, and

ε: a column vector of error terms associated with the endogenous variables.

SEM specifies dependent variables Y as explanatory variables associated with the other exogenous explanatory variables and estimates parameter matrix B to capture inherent causal relationship among dependent variables Y. The correlations between Y and ε caused by simultaneous equations disables Ordinary Least Square (OLS) to consistently estimate the parameter matrix B associated with right-hand-side Y. Econometrician brought forward 2-Stage Least Square (2SLS) and 3-Stage Least Square (3SLS) to achieve consistent estimators based on Instrumental Variable (IV) approach. 3SLS estimator is more efficient than 2SLS estimator, since former accommodates unequal variance of ε in each single equation. In addition to Least Square (LS) approach, Maximum Likelihood (ML) method can be applied to consistently estimate the parameters in SEM as well. Limited-Information Maximum Likelihood (LIML) and Full-Information Maximum Likelihood (FIML) in ML estimators are exactly the counterparts of 2SLS and 3SLS in LS estimators. With normally distributed disturbances, FIML is efficient among all the estimators. The other advanced estimation approaches, such as Asymptotically-Distribution-Free (ADF), are also available and applied in literature.

Travel behavior investigators applied SEM in their research in order to analyze complex causal relationship among travel-related variables, such as trip frequency, travel time or travel distance, activity duration, etc.. Past research regarding SEM application in travel behavior is briefly reviewed here.

Kitamura, *et al.* (1992) and Golob, *et al* (1994) are the first known application of SEM to joint activity duration and travel time data.

Kitamura (1996) and Pas (1996) are two overviews that include discussions of the role of SEM in activity and time-use modeling.

Lu and Pas (1997) present an SEM of in home activities, out-of-home activities (by type), and travel (measured various ways), conditional on socioeconomic variables. Estimation is by normal theory- maximum likelihood, and the emphasis is on interpretation of the direct and indirect effects. The data are derived from the Greater Portland, Oregon metropolitan area.

Golob and McNally (1997) present an SEM of the interaction of household heads in activity and travel demand, with data from Portland. Activities are divided into three types, and SEM results are compared using maximum likelihood (ML) and generalized least squares (GLS) estimation methods. They conclude that GLS methods should be used to estimate SEM when it is applied to activity participation data.

Fujii and Kitamura (2000) studied the latent demand effects of the opening of new freeways. The authors used an SEM to determine the effects of commute duration and scheduling variables on after work discretionary activities and their trips. Data are for the Osaka-Kobe Region of Japan.

Golob (2000) estimated a joint model of work and non-work activity duration using Portland data.

Kuppam and Pendyala (2000) presented three SEMs estimated by GLS using data from Washington, DC. The models focused on relationships between: (1) activity duration and trip generation, (2) durations of in-home and out-of-home activities, and (3) activity frequency and trip chain generation.

Simma and Axhausen (2001) developed an SEM that captured relationships between male and female heads of household with regard to activity and travel demands. The dependent variables included car ownership, distances traveled by males and females, and male and female trips by two types of activities using data from the Upper Austria.

Meka and Pendyala (2002) investigated the interaction between two adults in one household in their travel and activity time allocation by SEM using Southeast Florida data. Interesting trade-off within non-work travel time and non-work activity time between two adults was quantified and interaction of travel decision between household members was verified by SEM.

Application of SEM in travel behavior research initiates the analysis of complex causal relationship among individual's travel decisions, however, its limitation is quite apparent. Existing SEM only allows continuous dependent variables, but most travel decisions, such as travel mode choice, destination choice or route choice, are discrete in nature. It is necessary to introduce discrete dependent variables into SEM framework in aim at more comprehensively analyzing the causal relationship with respect to individual's travel

decisions. Next subchapter will review the recent research in attempt to integrate discrete dependent variable into SEM framework.

2.2 Discrete-Continuous Econometric Modeling Framework: Causal Analysis between Continuous Variable and Discrete Variable

Pendyala and Bhat et. al. (2002) made an attempt to integrate discrete dependent variable into SEM-like modeling framework. Using this discrete-continuous econometric modeling framework, they analyzed the causal relationship between timing and duration of maintenance activities. Details of this methodology are presented as below.

Let i be an index for time of day of activity participation ($i = 1, 2, \ldots, I$) and let q be an index for observations ($q = 1, 2, \ldots, Q$). Consider the following equation system:

$$\begin{cases} u_{qi}^* = \beta_i' z_{qi} + \gamma_i' a_q + \varepsilon_{qi} \\ a_q = \theta' x_q + \delta' D_q + \omega_q \end{cases} \qquad 2.4$$

$\varepsilon_{qi} \sim$ IID Gumbel(0,1), $\omega_q \sim N(0,\sigma^2)$.

where u_{qi}^* is the indirect (latent) utility associated with the ith time of day for the qth observation, D_q is a vector of the time of day dummy variables of length I, δ is a vector of coefficients representing the effects of different times of the day of activity participation on activity duration, ε_{qi} is a standard extreme-value (Gumbel) distributed error term assumed to be independently and identically distributed across times of the day and observations. a_q is logarithm of activity duration and γ is its coefficient. The error term ω_q is assumed to be i.i.d. normally distributed across observations with a mean of zero

and variance of σ^2. In Equation 2.4, the time of day alternative i will be chosen (*i.e.*, D_{qi} =1) if the utility of that alternative is the maximum of the I alternatives. Defining

$$v_{qi} = \max_{j\in\{1,2,\dots I\},\ j\neq i}(u^{*}_{qj}) - \varepsilon_{qi}, \qquad\qquad 2.5$$

the utility maximizing condition for the choice of the ith alternative may be written as: 1 $D_{qi}=1$ if and only if $\beta_i' z_{qi} > v_{qi}$. Let $F_i(v_{qi})$ represents the marginal distribution function of v_{qi} implied by the assumed IID extreme value distribution for the error terms ε_{qi} (i = 1,2,…,I) and the relationship in Equation 2.5. Using the properties that the maximum over identically distributed extreme value random terms is extreme value distributed and the difference of two identically distributed extreme values terms is logistically distributed, the implied distribution for v_{qi} may be derived as:

$$\Pr(v_{qi} < y) = F_i(y) = \frac{\exp(y)}{\exp(y) + \sum_{j\neq i}\exp(\beta_j' z_{qj} + \gamma' a_q)}$$

The non-normal variable v_{qi} is transformed into a standard normal variate using the integral transform result: $v_{qi}^{*} = \Phi_1^{-1}[F_i(v_{qi})]$, where $\Phi(.)$ is the standard cumulative distribution function.

Equation system 2.4 may now be rewritten as:

$$\begin{cases} D^{*}_{qi} = \beta_i' z_{qi} + \gamma_i a_q - v^{*}_{qi} \\ a_q = \theta' x_q + \delta' D_q + \omega_q \end{cases} \qquad\qquad 2.6$$

A correlation ρ_i between the error terms v_{qi}^{*} and ω_q is allowed to accommodate common

unobserved factors influencing the time of day choice for activity participation and the duration of the participation. Since a_q is partially determined by ω_q and $v_{qi}*$ is correlated with ω_q if ρ_i is unequal to zero, a_q is apparently correlated with random error term $v_{qi}*$ in the first equation. Similarly, D_q is also correlated with random error term ω_q in the second equation. The endogenous nature of dependent variables D_q and a_q entails full-information maximum likelihood method to jointly estimate their corresponding parameters γ and δ. Limited-information maximum likelihood (sequential estimation) does not provide consistent estimators of the coefficients on endogenous variables. The full-information likelihood function for estimating parameters is:

$$LL = \prod_{q=1}^{Q} \left\{ \prod_{i=1}^{I} \left[\frac{1}{\sigma} \phi(l_q) \Phi(b_{qi}) \right]^{D_{qi}} \right\}$$ 2.7

where $\phi(.)$ is the standard normal density function, and l_q and b_{qi} are defined as follows:

$$l_q = \left[\frac{a_q - \theta' x_q - \delta' D_q}{\sigma} \right], b_{qi} = \left[\frac{\Phi^{-1} F_i(\beta_i' z_{qi} + \gamma a_q) - \rho_i l_q}{\sqrt{1 - \rho_i^2}} \right]$$ 2.8

However, they also showed that either γ or δ should be equal to zero for logical consistency. It leads to two recursive model system indicating two different causal relationships between discrete dependent variable and continuous dependent variable. If $\gamma = 0$ and $\delta \neq 0$, then discrete dependent variable, time of day participation, affects continuous dependent variable, the logarithm of activity duration (but not vice versa), inversely, if $\gamma \neq 0$ and $\delta = 0$, then the continuous dependent variable affects discrete

11

dependent variable (but not vice versa). By comparing the goodness-of-fit from the models in two recursive system, the causal relationship between them may be analyzed and identified.

The most creative part of this study is to specify the endogenous discrete and continuous variables as explanatory variables into mutual explanatory functions, then by taking advantage of logical consistency to identify two recursive model structures indicating two unidirectional causal relationships. In addition, discrete-continuous econometric modeling approach is appropriately adopted here to consistently estimate parameters on endogenous variables with accommodation of correlation between normal error term and Gumbel error terms.

2.3 Nested Logit Model: Sequential Discrete Choice Analysis

Hensher et. al. (2000) analyzed trip chaining as a barrier to the propensity to use public transportation using multinomial logit, nested logit and mixed logit model. The theme of that paper is fairly close to the current thesis. Following Strathman & Dueker (1995), they classified trip chains into 7 categories as below:

1) Simple work: h – w (– w –) – h

2) Complex to work: h – nw (– nw/w –) – w – h

3) Complex from work: h – w (– nw/w –) – nw – h

4) Complex to and from work: h – nw – (– nw/w –) – w – (– nw/w –) – nw – h

5) Complex at work: h – w – (– nw/w –) – nw – (– nw/w –) – w – h

6) Simple non-work: h – nw – h

12

7) Complex non-work: h – nw – (– nw –) – h

h: home, w: work, nw: non work. The bracketed terms represent additional trips that may be in the chain.

On the other side, only two alternatives: Car and Public Transport are specified in mode choice set. They presumed the sequential choice behavior in which trip chaining type choice is ahead of trip mode choice, and then imposed 1 on the inclusive value parameter coefficients associated with all the complex chain type choices and equalized the inclusive value parameter coefficients associated with all the rest simple chain type choices. The estimation results showed that inclusive value parameter (0.866) associated with simple chains appeared significant and fell into the interval from 0 to 1. This statistical result is consistent with utility maximization theory behind nested logit model and validates the presumed sequential choice behavior.

From this study, it can be realized that nested logit model (NL) can be applied in dealing with problems regarding the relationship between two discrete dependent variables. Based on the assumption of a sequential choice mechanism, nested logit models representing two alternative tree structures can be formed. By checking the reasonableness of the estimated inclusive value parameter coefficients and/or comparing measures of goodness-of-fit between models of two different structures, the more plausible structure that is supported by the data may be identified. Based on this identification, the causal relationship between two choice behaviors can be clarified. However, the nested logit model has restriction on inclusive value parameter coefficient.

Once estimated inclusive value parameter coefficients fall outside the allowable range from 0 to 1, the corresponding nested logit structure becomes invalid. Besides, unlike SEM, nested logit model does not provide a set of parameters directly measuring the impact of endogenous variables on the others.

2.4 Summary

Chapter 2 extensively reviews the existing methodologies in literature concerned with causal analysis between/among travel-related variables. These methodologies include Structural Equation Model, Discrete-Continuous Econometric Model and Nested Logit Model. As reviewed earlier, SEM only allows the causal analysis among multiple continuous endogenous variables; NL is potentially applicable in causal analysis evolved with discrete dependent variables but modeling structure is completely different from SEM and does not provide parameters directly measuring the causal effect. Discrete-Continuous Econometric Model initially introduced discrete variable into a SEM-like modeling framework. Following this good start, it seems to be straightforward to generalize this approach from discrete-continuous situation towards discrete-discrete situation. The details of modeling methodology for discrete-discrete causal analysis are brought forward in the next chapter.

CHAPTER 3

MODELING METHODOLOGY

3.1 Outline

As discussed in last chapter, the causal relationship between tour type choice and mode choice cannot be analyzed in SEM or Discrete-Continuous Econometric Model since dependent variables are both discrete in nature. In this thesis, two different econometric modeling methods are employed to analyze the relationship between two discrete variables. The first is the recursive simultaneous bivariate probit model, which allows the analysis of one-way causal relationships between two choice behaviors. In this formulation, the random error terms are assumed to follow the bivariate normal distribution. The bivariate normality assumption implies that two endogenous dummy variables may not coexist in mutual functional relations. The existence of an endogenous dummy variable in either function corresponds to two different causal structures (see Section 3.2 for details). Intuitively, this feature of the bivariate probit model provides an appropriate approach to distinguish the causality between tour complexity and auto mode choice. However, this approach also entails an underlying assumption that an explicit unidirectional causal relationship (or at least the tendency of such a unidirectional causal relationship) exists in the population being studied. A unidirectional causal relationship may exist in a specific tour, but the nature of the causal relationship may vary across individuals and across tours for the same individual. Macroscopically, the presence of a

bidirectional causality would possibly appear in the population if neither unidirectional causal relationship dominates the other.

In order to address the possibility of a simultaneous bidirectional causality, this thesis also uses the simultaneous logit model formulation presented by Schmidt and Strauss (1975) and initially introduced in the transportation context by Ouyang, et.al. (2002). This model formulation enables the modeling of bidirectional causality that might exist in tour complexity and mode choice. Essentially, the simultaneous logit model may be considered an extension of the more commonly known multinomial logit model, where two endogenous dummy variables can be incorporated into the mutual utility functions simultaneously. The only restriction on these two dummy variables is that their model coefficients must be identical for logical consistency (see Section 3.3 for details).

Thus, three different possible causal structures are considered in this thesis:

1) Mode choice $\rightarrow$ Trip chain complexity (recursive bivariate probit model)

2) Trip chain complexity $\rightarrow$ Mode choice (recursive bivariate probit model)

3) Trip chain complexity $\leftrightarrow$ Mode choice (simultaneous logit model)

Through a performance comparison of models across the three causal structures, it is envisaged that the relationship between tour complexity and mode choice may be discussed and clarified.

3.2 Recursive Simultaneous Bivariate Probit Model

If the tour's complexity/simplicity and auto/non-auto mode choice are treated as two binary choices, the bivariate probit model can be formulated at the tour level to simultaneously analyze their probabilities with accommodation of random error correlation. The general formulation is as follows:

$$\begin{cases} M_q^* = \gamma' z_q + \alpha T_q + \varepsilon_q \\ T_q^* = \beta' x_q + \eta M_q + \omega_q \end{cases} \qquad 3.1$$

where

q is an index for observations of tour ($q = 1, 2, \ldots, Q$)

M_q^* is a latent variable representing the mode choice for tour q

T_q^* is a latent variable representing the complexity of tour q

$M_q = 1$ if $M_q^* > 0$, $= 0$ otherwise

i.e., M_q is a dummy variable indicating whether tour q uses the auto mode

$T_q = 1$, if $T_q^* > 0$, $= 0$ otherwise

i.e., T_q is a dummy variable indicating whether tour q is complex

z_q is a vector of explanatory variables for M_q^*

x_q is a vector of explanatory variables for T_q^*

γ, β are two vectors of model coefficients associated with the explanatory variables z_q and x_q, respectively

α is a scalar coefficient for T_q to measure the impact of tour's complexity on mode choice

η is a scalar coefficient for M_q to measure the impact of mode choice on the choice of tour complexity

ε_q and ω_q are random error terms, which are standard bivariate normally distributed with zero means, unit variances, and correlation ρ, i.e. $\varepsilon_q, \omega_q \sim \phi_2(0,0,1,1,\rho)$.

Based on this normality assumption, one can derive the probability of each possible combination of binary choices for tour q:

$$prob(M = 0, T = 0) = \Phi_2[-\gamma'z, -\beta'x, \rho]$$ 3.2

$$prob(M = 1, T = 0) = \Phi_1[-(\beta'x + \eta)] - \Phi_2[-\gamma'z, -(\beta'x + \eta), \rho]$$ 3.3

$$prob(M = 0, T = 1) = \Phi_1[-(\gamma'z + \alpha)] - \Phi_2[-(\gamma'z + \alpha), -\beta'x, \rho]$$ 3.4

$$prob(M = 1, T = 1) = 1 - \Phi_1[-(\gamma'z + \alpha)] - \Phi_1[-(\beta'x + \eta)]$$

$$+ \Phi_2[-(\gamma'z + \alpha), -(\beta'x + \eta), \rho]$$ 3.5

where

$\Phi_1[\cdot]$ is the cumulative distribution function for standard univariate normal distribution,

$\Phi_2[\cdot]$ is the cumulative distribution function for standard bivariate normal distribution.

The sum of the probabilities for the four combinations of two binary choices should be equal to one, i.e.,

$$prob(M = 0, T = 0) + prob(M = 1, T = 0) + prob(M = 0, T = 1) + prob(M = 1, T = 1) = 1$$ 3.6

Substituting equations 3.2 through 3.5 into equation 3.6, it can be shown that

$$\Phi_2[-\gamma'z, -\beta'x, \rho] + \Phi_2[-(\gamma'z + \alpha), -(\beta'x + \eta), \rho]$$

$$= \Phi_2[-\gamma'z, -(\beta'x + \eta), \rho] + \Phi_2[-(\gamma'z + \alpha), -\beta'x, \rho]$$ 3.7

18

This equation does not hold unless either α or η is equal to zero. This requirement, known as the logical consistency condition similar to the situation in Discrete-Continuous Econometric Model, will lead to two different recursive simultaneous modeling structures (Maddala, 1983) suggesting two different causal relationships:

1) $\alpha = 0$, $\eta \neq 0$ (Mode Choice $\rightarrow$ Tour Complexity)

$$\begin{cases} M_q^* = \gamma' z_q + \varepsilon_q \\ T_q^* = \beta' x_q + \eta M_q + \omega_q \end{cases}$$

3.8

In this structure, mode choice is predetermined as per the first functional relationship. Then, the choice of mode is specified as a dummy variable in the second functional relationship for tour complexity to directly measure the impact of mode choice on the complexity of the trip chain or tour.

2) $\alpha \neq 0$, $\eta = 0$ (Tour Complexity $\rightarrow$ Mode Choice)

$$\begin{cases} M_q^* = \gamma' z_q + \alpha T_q + \varepsilon_q \\ T_q^* = \beta' x_q + \omega_q \end{cases}$$

3.9

Conversely, one may consider the alternative structure in which tour complexity is predetermined as per the second functional relationship. The complexity of the tour is specified as an explanatory variable influencing mode choice as per the first functional relationship.

Thus, the desirable feature of the bivariate probit model in which the coefficients of two endogenous dummy variables do not coexist in both functional relationships provides an

appropriate modeling framework to analyze the unidirectional causality between tour complexity and mode choice.

The endogenous nature of one of the dependent variables in the simultaneous equation system can be ignored in formulating the likelihood function. To facilitate formulating likelihood functions, equations 3.2 through 3.5 can be rewritten in a format including only the cumulative distribution function of the standard bivariate normal distribution.

$$prob(M = 0, T = 0) = \Phi_2[-\gamma'z, -\beta'x, -\rho] \qquad \qquad 3.10$$

$$prob(M = 1, T = 0) = \Phi_2[\gamma'z, -(\beta'x + \eta), -\rho] \qquad \qquad 3.11$$

$$prob(M = 0, T = 1) = \Phi_2[-(\gamma'z + \alpha), \beta'x, -\rho] \qquad \qquad 3.12$$

$$prob(M = 1, T = 1) = \Phi_2[\gamma'z + \alpha, \beta'x + \eta, \rho] \qquad \qquad 3.13$$

Equations 3.10 through 3.13 and the corresponding likelihood functions can be summarized by the following general formulations for the two different unidirectional causal structures (Greene, 2003):

1) $\alpha = 0$, $\eta \neq 0$ (Mode Choice $\rightarrow$ Tour Complexity)

$$prob_q = \Phi_2[\mu_q\gamma'z_q, \tau_q(\beta'x_q + \eta M_q), \mu_q\tau_q\rho] \qquad \qquad 3.14$$

$$L = \prod_{q=1}^{Q} \{\Phi_2[\mu_q\gamma'z_q, \tau_q(\beta'x_q + \eta M_q), \mu_q\tau_q\rho]\} \qquad \qquad 3.15$$

20

2) $\alpha \neq 0$, $\eta = 0$ (Tour Complexity $\rightarrow$ Mode Choice)

$$prob_q = \Phi_2[\mu_q(\gamma' z_q + \alpha T_q), \tau_q \beta' x_q, \mu_q \tau_q \rho] \qquad 3.16$$

$$L = \prod_{q=1}^{Q} \left\{ \Phi_2[\mu_q(\gamma' z_q + \alpha T_q), \tau_q \beta' x_q, \mu_q \tau_q \rho] \right\} \qquad 3.17$$

where $\mu_q = 2M_q - 1$ and $\tau_q = 2T_q - 1$.

As the likelihood functions of the recursive bivariate probit model and the common bivariate probit model are virtually identical, parameter estimation can be accomplished using readily available software such as LIMDEP 8.0.

3.3 Simultaneous Logit Model

One may also consider the possibility where neither of the two unidirectional causal structures is dominant within the population, i.e., both causal structures are prevalent in the population. In addition, one may consider the possibility where the choices regarding tour complexity and mode are made simultaneously. To accommodate such plausible bidirectional causality, the simultaneous logit model is applied in this thesis. The simultaneous logit model may be considered an extension of the multinomial logit model commonly used in transportation modeling practice. In the simultaneous logit model, the logarithm of the ratio of probabilities for two alternatives to be selected from one choice set is assumed to equal a linear combination of a set of explanatory variables. One dummy variable indicating the choice of tour complexity may be added into the set of explanatory variables for mode choice; similarly, one dummy variable indicating mode choice may be added into the set of explanatory variables for tour complexity. The

formulations may be written as follows (all of the symbols have the same meaning as in Section 3.2):

$$\log\left[\frac{P(M=1\,|\,T)}{P(M=0\,|\,T)}\right]_q = \gamma' z_q + \alpha T_q \qquad\qquad 3.18$$

$$\log\left[\frac{P(T=1\,|\,M)}{P(T=0\,|\,M)}\right]_q = \beta' x_q + \eta M_q \qquad\qquad 3.19$$

By rewriting equations 3.18 and 3.19 across two possible values that T_q and M_q can take, one gets:

$$\log\left[\frac{P(M=1,T=0)}{P(M=0,T=0)}\right]_q = \gamma' z_q \qquad\qquad 3.20$$

$$\log\left[\frac{P(M=1,T=1)}{P(M=0,T=1)}\right]_q = \gamma' z_q + \alpha \qquad\qquad 3.21$$

$$\log\left[\frac{P(T=1,M=0)}{P(T=0,M=0)}\right]_q = \beta' x_q \qquad\qquad 3.22$$

$$\log\left[\frac{P(T=1,M=1)}{P(T=0,M=1)}\right]_q = \beta' x_q + \eta \qquad\qquad 3.23$$

The sum of the probabilities for the four combinations of binary choices should be equal to one, i.e.,

$$P(M=0,T=0)_q + P(M=1,T=0)_q + P(M=0,T=1)_q + P(M=1,T=1)_q = 1 \qquad 3.24$$

22

By converting simultaneous equations 3.20 through 3.23, it can be shown that

$$P(M=1,T=1)_q = P(M=0,T=0)_q \exp(\gamma' z_q + \beta' x_q + \alpha)$$

$$= P(M=0,T=0)_q \exp(\gamma' z_q + \beta' x_q + \eta) \qquad 3.25$$

For logical consistency, α must be equal to η. Endogenous dummy variables T_q and M_q are allowed to coexist in the simultaneous equation system.

By replacing η with α and solving the simultaneous equations 3.20 through 3.24, the probability for each combination is formulated as follows:

$$P_{00q} = P(M=0,T=0)_q = 1/\Delta_q \qquad 3.26$$

$$P_{10q} = P(M=1,T=0)_q = \exp(\gamma' z_q)/\Delta_q \qquad 3.27$$

$$P_{01q} = P(M=0,T=1)_q = \exp(\beta' x_q)/\Delta_q \qquad 3.28$$

$$P_{11q} = P(M=1,T=1)_q = \exp(\gamma' z_q + \beta' x_q + \alpha)/\Delta_q \qquad 3.29$$

where

$$\Delta_q = 1 + \exp(\gamma' z_q) + \exp(\beta' x_q) + \exp(\gamma' z_q + \beta' x_q + \alpha). \qquad 3.30$$

Finally, the likelihood function may be formulated as follows:

$$L = \prod_{q=1}^{Q} (P_{00q})^{(1-M_q)(1-T_q)} (P_{10q})^{M_q(1-T_q)} (P_{01q})^{(1-M_q)T_q} (P_{11q})^{M_q T_q} \qquad 3.31$$

Model estimation is performed using the Gauss programming language (see Gauss codes in Appendix).

CHAPTER 4

DATA SET AND SAMPLE DESCRIPTION

4.1 Survey Description

The data set used for analysis and model estimation is extracted from the Swiss Travel Microcensus 2000. A very detailed description of the survey and the survey sample can be found in Ye and Pendyala (2003). Only a brief description of the survey sample is provided in this thesis. The survey respondent sample consists of 27,918 households from 26 cantons in Switzerland. The person sample was formed by randomly selecting one person over 6 years old from each household with less than 4 household members and two persons over 6 years old from each household with 4 or more members. As a result of this sampling scheme, the person respondent sample consisted of 29,407 persons. All of the persons in the person sample were asked to report their travel in a one-day trip diary. The resulting trip data set includes 103,376 trips reported by 29,407 interviewed persons (including the possibility of some respondents making zero trips on the survey day). The household and person characteristics of these samples are respectively shown in Tables 4.1 and 4.2. Data corresponding to respondents from the Canton of Zurich was extracted to reduce the data to a more manageable size and to control for possible area specific effects. Tables 4.1 and 4.2 also include summary statistics for the subsample of respondents from Zurich in addition to those of the overall

Swiss sample. The Zurich subsample includes 5,128 households from which 5,241 persons provided travel information.

4.2 Household Characteristics of Sample and Subsample

Table 4.1 Household Characteristics of Swiss Travel Microcensus 2000

Characteristic	Swiss Sample	Zurich Subsample
Sample Size	27918	5128
Household Size	2.43	2.20
1 person	27.5%	31.0%
2 persons	35.1%	37.0%
3 persons	14.0%	12.6%
≥ 4 persons	23.4%	19.5%
Monthly Income		
Low (<Fr 4K)	20.8%	17.5%
Medium (Fr 4K~Fr 8K)	35.9%	34.7%
High (>Fr 8K)	18.4%	21.5%
Vehicle Ownership	1.17	1.09
0 auto	19.8%	23.8%
1 auto	50.5%	49.5%
2 autos	24.5%	21.8%
≥ 3 autos	5.2%	4.9%
Family Type		
Single	27.2%	30.6%
Unmarried couple (no child)	27.9%	29.3%
Married	43.6%	37.8%
Presence of Children		
Child <6 years old	10.6%	8.7%
Child 6~17 years old	22.5%	18.9%
Household Location		
Major city	42.4%	55.1%
Surrounding areas of city	30.4%	35.2%
Rural	26.1%	8.9%

Table 4.1 lists sample size, average household size and its distribution, household income distribution, average household vehicle ownership and its distribution, family type distribution, proportion of households with children and household location distribution by total Swiss household sample and Zurich household subsample.

As expected, the proportion of households without automobiles in this Swiss sample is substantially higher than in a typical sample from the United States. This may be reflective of the higher level of public transport service in Switzerland that enables mobility and accessibility without the same level of auto dependence. As a result, one might expect the automobile to play a smaller role in the Swiss travel environment than in the US environment.

In general, the Zurich subsample exhibits characteristics rather similar to the overall Swiss sample but to some degree reflects the urban characteristics of Zurich from various angles. Average household size in Zurich subsample is smaller than the average in Swiss sample in accord with the common sense that urban residents tend to live more independently. Monthly income distribution apparently shows that Zurich has 3.1% more families falling into high-income category but 3.3% less families into low-income category than the nation level distribution, which also reflects Zurich's urban characteristics. Average vehicle ownership in Zurich subsample is somewhat smaller than in Swiss sample and there are 4.0% more households in Zurich without vehicle than nation-level distribution. It may be conjectured that advanced public transit system in Zurich enables more people live without vehicles in household. Statistics also shows that

Zurich has substantially less married families and less families with children, which may account for smaller average household size. Finally, the household location distribution in Zurich subsample shows less than 10% households are located in rural area and around 55% in major city, as expected.

4.3 Person Characteristics of Sample and Subsample

The person characteristics are shown separately for commuters and non-commuters. The statistics shows personal daily trip rates is less than 4.0 for both commuters and non-commuters, whereas personal daily trip rates in US are typically more than 4.0. Because non-commuter is defined as those people who make work trips at least twice per week, work trips rates for non-commuter are not exactly equal to zero but fairly close to zero, as shown in Table 4.2. Auto mode share within work trips and non-work trips is less than 60% in accordance with the lower vehicle ownership in comparison with US.

Person characteristics show similarities between the overall Swiss sample and the Zurich subsample. As expected, non-commuters show a greater proportion of elderly (retired) and young persons than commuter samples. On average, commuters make about 1.4 trip chains per day where the trip chain is defined as a complete home-to-home tour. Non-commuters make, on average, about 1.2 trip chains per day. Commuters make nearly four trips per day while non-commuters make fewer trips at about three trips per day. The difference in trip making between commuters and non-commuters is primarily due to

Table 4.2 Person Characteristics of Swiss Travel Microcensus 2000

Characteristic	Swiss Sample		Zurich Subsample	
	Commuters	Non-Commuters	Commuters	Non-Commuters
Sample Size	13296	16111	2504	2737
Age (in years)	40.7 (Mean)	46.5 (Mean)	41.0 (Mean)	49.0 (Mean)
Young (6~29)	20.8%	31.8%	20.4%	28.1%
Middle (30~59)	73.5%	26.3%	73.1%	25.3%
Old (≥60)	5.7%	41.8%	6.5%	46.6%
Sex				
Male	54.8%	39.3%	56.4%	39.6%
Female	45.2%	60.7%	43.6%	60.4%
Employment Status				
Full time	71.8%	8.9%	72.8%	8.7%
Part time	24.3%	6.0%	23.0%	7.5%
Licensed	87.7%	50.7%	87.3%	51.4%
#Chains/day	1.43	1.24	1.32	1.20
#Trips/day	3.98	3.13	3.77	3.13
Work trips	0.91	0.08	0.87	0.08
Non-work trips	3.07	3.05	2.90	3.05
Work Trip Mode Share				
Auto	57.1%	54.8%	51.8%	46.6%
Non-Auto	42.9%	45.2%	48.2%	53.4%
Non-Work Trip Mode Share				
Auto	58.4%	41.1%	52.6%	37.0%
Non-Auto	41.6%	58.9%	47.4%	63.0%

Notes: Commuters are those who make work trips at least twice per week. All others are non-commuters. Incomplete chains, defined as a series of trips that does not end at home, are not included.

the work trip as non-work trip generation is virtually identical between commuters and non-commuters. Finally, it is seen that the Zurich subsamples are less dependent on the auto mode as they exhibit a smaller auto mode share compared to the overall Swiss sample. This is presumably due to the high level of transit service available in the Zurich area.

4.4 Trip Chain Analysis of Subsample

In this thesis, the unit of analysis and modeling is the tour or trip chain. A trip chain is defined in this thesis as a complete home-to-home journey where the origin of the first trip is home and the destination of the last trip is home. No intermediate home stop is present within the trip chain. Whenever the home location is reached, a chain is formed. A tour-level data set was formed by aggregating the trip data set to the tour level. All person and household characteristics were merged into the tour level data set. In most cases, a single mode was prevalent for the trip chain. In cases where multiple modes were prevalent within the same trip chain or tour, a single mode was assigned based on the whether or not the auto mode was used in the chain. If the auto mode was used for any segment in the trip chain, then the chain was assigned an auto mode and vice versa. Each tour was classified as a simple or complex tour depending on whether it had one intermediate stop or more than one intermediate stops within the chain.

In addition, tours were also classified as work-based tours and non-work-based tours. Any tour that included a work stop (regardless of the presence of other types of stops) was classified as a work-based tour while any tour that included only non-work stops was

classified as a non work-based tour. It was felt that the causal relationships governing work-based tours may be different from those governing non work-based tours. This is because the presence of a work stop may impose a certain amount of spatial and temporal rigidity on the activity/travel behavior of the individual in the context of that tour. The constraints associated with the work activity may lead to a different causal structure underlying trip chain formation and mode choice.

Table 4.3 Crosstabulation of Mode Choice and Tour Type for Non-work Tours

Mode Choice	Tour Type		Total
	Simple	Complex	
Frequency			
Non-auto	2685	661	3346
Auto	1030	525	1555
Total	3715	1186	4901
Column Percent			
Non-auto	72.3%	55.7%	68.3%
Auto	27.7%	44.3%	31.7%
Total	100.0%	100.0%	100.0%
Row Percent			
Non-auto	80.2%	19.8%	100.0%
Auto	66.2%	33.8%	100.0%
Total	75.8%	24.2%	100.0%

As the model estimation was performed only on the Zurich subsample, all further analysis presented in the thesis pertains only to this subsample. The Zurich subsample included 4,901 non-work tours and 1,711 work tours. Tables 4.3 and 4.4 offer simple cross-tabulations of tour complexity against mode choice. Table 4.3 examines the distribution of tour complexity by mode choice for non-work tours while Table 4.4 examines the distribution for work tours.

Table 4.4 Crosstabulation of Mode Choice and Tour Type for Work Tours

Mode Choice	Tour Type		Total
	Simple	Complex	
Frequency			
Non-auto	436	355	791
Auto	397	523	920
Total	833	878	1711
Column Percent			
Non-auto	52.3%	40.4%	46.2%
Auto	47.7%	59.6%	53.8%
Total	100.0%	100.0%	100.0%
Row Percent			
Non-auto	55.1%	44.9%	100.0%
Auto	43.2%	56.8%	100.0%
Total	48.7%	51.3%	100.0%

An examination of column-based percentages in Table 4.3 indicates that about 28 percent of simple non-work tours involve the use of the automobile as the primary mode of transportation. This value is considerably higher at 44 percent for complex non-work tours. Thus it appears that there is a correlation (at least) between mode choice and tour complexity. Clearly, the auto mode is utilized to a greater degree in the context of complex multi-stop trip chains. Similarly, examining the row-based percentages shows that 80 percent of non-work non-auto tours are simple in nature (involve only one stop). On the other hand, only 66 percent of non-work auto tours are simple in nature. Thus it appears that non-auto tours tend to be more simple than auto-based tours.

Table 4.4 offers similar indications, albeit the tendencies are not as strong as those seen in Table 4.3. In the case of work tours, it is found that a majority of simple tours are non auto-based (52 percent) while a majority of complex tours are auto-based (60 percent). Similarly, a majority of non auto-based work tours tend to be simple in nature (55

percent), while a majority of auto-based tours tend to be complex in nature (57 percent). Once again, a clear correlation between auto use and trip chain complexity is seen in these cross tabulations. Given the difference in the percent distributions between work and non-work tours, it was considered prudent to examine the causal relationship between tour complexity and mode choice for work and non-work tours separately.

CHAPTER 5

MODEL ESTIMATION RESULTS

5.1 Outline

This chapter presents estimation results for the models developed in this thesis. Tables 5.1 and 5.5 present a description of the variables used in model estimation for non-work and work tours respectively. The variables are listed in alphabetical order and mostly constitute dummy variable indicators that take a value of one if the condition is satisfied and zero otherwise. Estimation results for non-work tour models are provided in Tables 5.2 through 5.4 and estimation results for work tour models are provided in Tables 5.6 through 5.8.

5.2 Estimation Results for Non-Work Tours

Table 5.2 provides estimation results for the causal structure where tour complexity affects mode choice, Table 5.3 provides estimation results for the causal structure where mode choice affects tour complexity, and Table 5.4 provides estimation results for the simultaneous logit model that is intended to capture simultaneous causality between the two variables. In Table 5.2, the coefficient for tour complexity is statistically significant and positive in the mode choice model. This lends credence to the hypothesis that the need to make a complex tour is likely to increase dependency on the auto mode.

33

In addition, it was found that demographic and socio-economic characteristics, the tour's primary purpose, and time-of-day significantly influence mode choice and tour complexity. For example, larger household sizes, lower income, and school tours tend to be associated with a lower propensity to use the automobile mode.

Table 5.1 Description of Explanatory Variables Used in Non-work Tour Models
(N = 4901)

Variable Name	Variable Description	Mean	Std Dev
AD1_YOU1	Household is composed of one adult and one child who is over 6 but below 17 years old	0.01	0.09
AMPEAK	Tour starts in AM peak period (7:00~8:59)	0.17	0.37
AUTO	Auto mode choice for tour	0.32	0.47
CAR_GE2	Number of autos in household ≥ 2	0.29	0.45
COMPLEX	Tour is complex (multi-stop)	0.24	0.43
HHSIZE	Number of household members	2.72	1.46
HHSIZE1	Single person household	0.25	0.44
HHSIZE4	Number of household members ≥ 4	0.36	0.48
HIGH_INC	Monthly household income > Fr10000	0.14	0.34
LOW_INC	Monthly household income < Fr4000	0.18	0.38
OLD	Person > 60 years old	0.27	0.45
MALE	Person is male	0.46	0.50
PMPEAK	Tour starts in PM peak period (16:00~17:59)	0.10	0.30
SCHOOL	Primary purpose of the tour is school	0.11	0.32
SERVICE	Primary purpose of the tour is service	0.06	0.23
SHOPPING	Primary purpose of the tour is shopping	0.30	0.46
TOURDIST	Total distance that the tour covers (km)	26.32	86.15

Note: Primary purpose of a tour is defined as the trip purpose other than "return home" that accounts for the longest cumulative distance within the tour. If two different trip purposes account for equal distances within the tour, then the primary purpose is defined based on the following priority sequence: work>school>service>shopping>recreation>other

On the other hand, higher income, higher car ownership levels, males, and tours primarily involving service (serve passenger) stops increase the propensity to use the automobile. In the tour complexity model, it is found that individuals in single person households tend to make complex tours as opposed to individuals in larger households. This is a rather

surprising result as it was expected that individuals in larger households would have to

make complex tours to serve the needs of multiple household members.

Table 5.2 Non-work Tour Model
(Complex Tour → Auto Mode Choice)

Variable	Parameter	t-test
Auto Mode Choice Model		
Constant	-0.8035	-16.344
HHSIZE	-0.0790	-5.135
HIGH_INC	0.1075	2.195
LOW_INC	-0.3315	-6.842
CAR_GE2	0.5050	12.077
MALE	0.3169	9.153
SERVICE	0.3684	4.583
SCHOOL	-1.1255	-10.622
COMPLEX	1.6323	23.151
Complex Tour Choice Model		
Constant	-0.5231	-13.638
HHSIZE1	0.1099	2.467
AD1_YOU1	0.3843	1.998
HHSIZE4	-0.1314	-2.597
OLD	-0.1620	-3.737
YOUNG	-0.6122	-9.464
SERVICE	0.5172	6.279
SHOPPING	-0.2089	-5.230
AMPEAK	0.2879	5.933
PMPEAK	-0.2372	-3.568
ρ (Error Correlation)	-0.8083	-19.744
Sample Size	4901	
Number of parameters	20	
Log-likelihood		
At convergence	-5179.689	
At market share	-5734.170	
At zero	-6794.229	

Table 5.3 Non-work Tour Model
(Auto Mode Choice → Complex Tour)

Variable	Parameter	t-test
Auto Mode Choice Model		
Constant	-0.3797	-7.247
HHSIZE	-0.1502	-9.542
HIGH_INC	0.1710	2.941
LOW_INC	-0.3951	-6.891
CAR_GE2	0.6317	13.577
MALE	0.3630	9.079
TOURDIST	0.0005	3.361
SERVICE	0.8842	10.815
SCHOOL	-1.3619	-12.936
Complex Tour Choice Model		
Constant	-0.9301	-17.228
HHSIZE1	0.2534	5.142
HHSIZE4	-0.1784	-3.613
SERVICE	0.3623	3.866
SHOPPING	-0.2518	-5.531
AMPEAK	0.3032	5.621
PMPEAK	-0.3263	-4.399
AUTO	0.7593	6.168
ρ (Error Correlation)	-0.2473	-3.158
Sample Size	4901	
Number of parameters	18	
Log-likelihood		
At convergence	-5207.736	
At market share	-5734.170	
At zero	-6794.229	

Table 5.4 Non-work Tour Model
Simultaneous Logit Model (Auto Mode Choice ↔ Complex Tour)

Variable	Parameter	t-test
Auto Mode Choice Model		
Constant	-0.8028	-8.814
HHSIZE	-0.2384	-8.476
HIGH_INC	0.2472	2.531
LOW_INC	-0.6879	-6.955
CAR_GE2	1.0678	13.417
MALE	0.6399	9.495
SERVICE	1.3552	9.600
SCHOOL	-2.6113	-10.818
Complex Tour Choice Model		
Constant	-1.2239	-15.641
HHSIZE1	0.3812	4.529
AD1_YOU1	0.5977	1.686
HHSIZE4	-0.3168	-3.224
OLD	-0.1727	-2.052
YOUNG	-0.4019	-3.454
SERVICE	0.7508	5.486
SHOPPING	-0.4656	-5.746
AMPEAK	0.5479	6.024
PMPEAK	-0.5577	-4.200
α (Joint Dependence)	0.6390	8.885
Sample Size	4901	
Number of parameters	19	
Log-likelihood		
At convergence	-5203.343	
At market share	-5734.170	
At zero	-6794.229	

On the other hand, one may conjecture that the possibility of task allocation present in a multi-person household may reduce the need to perform multi-stop trip chains (Strathman et al., 1994). Single parents, as expected, are more prone to engage in multi-stop trip chains. Compared to the middle age group, the older and younger individuals tend to be less prone to making multi-stop trip chains. This is possibly due to household and other obligations reaching their peak for many individuals during the middle age lifecycle stage. It is also rather surprising that tours undertaken in the AM peak show a greater propensity to involve multiple stops than those undertaken in the PM peak period. However, in the context of non-work tours, this may be a plausible result in that people combine a series of errands and school activities in the morning and complete their activities by mid-day. Another possible explanation is that time constraints towards the end of the day (PM period) limit the number of activities that an individual can pursue at that time. Another interesting finding is that gender does not significantly influence tour complexity in the case of non-work tours. Other studies have suggested that females tend to make more complex trip chains than males (McGuckin et al., 1999). The analysis in this thesis does not support that finding in the Swiss travel context. The tour's primary purpose appears to affect tour complexity. While service (serve passenger) tours tend to be complex in nature, shopping tours do not tend to be complex in nature. Thus it appears that the shopping activity may be more prone to being a stand-alone activity within a tour. The error correlation is found to be statistically significant and this is indicative of the validity of the assumption that non-work tour complexity and mode choice should be modeled in a simultaneous equations framework. The negative sign associated with the error correlation indicates that the unobserved factors influencing these two variables are

38

negatively correlated. Further analysis is warranted to fully explain the implications of the negative error correlation.

Table 5.3 provides estimation results for the causal structure where mode choice affects tour complexity for non-work tours. Interestingly, it is found that mode choice significantly affects tour complexity and that the choice of auto is positively associated with the formation of complex tours. Thus it appears from this model that the choice of the automobile mode for a tour contributes positively to the formation of multi-stop trip chains. In addition, the error correlation is significant and negative as in Table 5.2. All of the other indications provided by the model system are similar to those seen in Table 5.2. The tour length (distance in km) is found to significantly contribute to the choice of the auto mode.

Tables 5.2 and 5.3 appear to support the notion that there is a bidirectional causality between mode choice and tour complexity. In both models (representing two different causal structures), the coefficient associated with the endogenous variable on the right hand side is statistically significant and consistent with expectations and trends in the data set. In addition, both models offer significant and negative error correlation supporting the simultaneous equations formulation for representing the relationship between mode choice and tour complexity.

In light of these findings, Table 5.4 provides estimation results of the simultaneous logit model for non-work tours in which mode choice and tour complexity influence each

other simultaneously and bidirectional causality is allowed. The significantly positive joint dependence parameter, α, shows the presence of significant positive correlation between auto mode choice and tour complexity. The other explanatory variables provide similar indications as those in Table 5.2.

As all of the estimation results in Tables 5.2 through 5.4 offer plausible and similar interpretations, a more rigorous performance comparison must be conducted among the models to potentially identify the causal structure underlying the data set. This performance comparison is presented in Chapter 5 following the discussion of the estimation results for the work tour models.

5.3 Estimation Results for Work Tour Models

Estimation results for work tour models are provided in Tables 5.6 through 5.8. Table 5.6 provides estimation results for the causal structure where tour complexity affects mode choice, Table 5.7 provides estimation results for the causal structure where mode choice affects tour complexity, and Table 5.8 provides estimation results for the simultaneous logit model that is intended to capture simultaneous causality between the two variables.

In Table 5.6, it is found that tour complexity has a positive impact on auto mode choice. This is consistent with expectations, trends in the data, and the models of non-work tours. The coefficient associated with tour complexity variable in the mode choice model is positive and statistically significant.

Thus the model supports the notion that a complex tour or trip chaining pattern contributes to the choice of auto as the mode for the tour. In addition, the error correlation is negative and statistically significant, once again supporting the simultaneous equations formulation of the relationship between tour complexity and mode choice.

Table 5.5 Description of Explanatory Variables Used in Work Tour Models
(N = 1711)

Variable Name	Variable Description	Mean	Std Dev
AUTO	Auto mode choice for tour	0.54	0.50
BEG13_14	Tour starts in time period from13:00 to 14:59	0.11	0.31
BEG6_8	Tour starts in time period from 6:00 to 8:59	0.67	0.47
COMPLEX	Tour is complex (multi-stop)	0.51	0.50
COUNTRY	Residence is located in rural area	0.11	0.31
DIS_WORK	Distance between residence and work place (km)	11.02	15.02
END12	Tour ends in time period from 12:00 to 12:59	0.12	0.32
FREEPARK	Reserved parking lot at the work place is free	0.33	0.47
FULLTIME	Person is full-time employed	0.77	0.42
HHSIZE1	Single person household	0.30	0.46
HIGH_INC	Monthly household income > Fr10000	0.21	0.41
OWN_BUS	Person owns enterprise/business	0.14	0.35
OWN_RES	Person owns the residence	0.30	0.46
MALE	Person is male	0.61	0.49
SWISS	Person is of Swiss Nationality	0.85	0.36

With respect to other variables, it is found that free parking, longer commutes, full time employment, and rural residence are all positively influencing the choice of auto as the mode choice for work tours. All of these findings are consistent with expectations. In the tour complexity model, it is found that single persons tend to engage in multi-stop trip chains possibly due to the inability to share or allocate tasks among multiple household members.

Table 5.6 Work Tour Model
(Complex Tour → Auto Mode Choice)

Variable	Parameter	t-test
Auto Mode Choice Model		
Constant	-1.0013	-8.757
FREEPARK	1.2859	15.278
DIS_WORK	0.0062	2.673
COUNTRY	0.5123	4.469
FULLTIME	0.2458	3.094
COMPLEX	0.8106	3.723
Complex Tour Choice Model		
Constant	-0.4812	-4.844
HHSIZE1	0.1859	2.640
HIGH_INC	0.3173	4.000
OWN_BUS	0.3312	3.660
SWISS	0.2529	2.833
OWN_RES	0.1592	2.222
BEG6_8	0.3096	4.180
BEG13_14	-0.3678	-3.238
END12	-0.7236	-7.384
ρ (Error Correlation)	-0.3481	-2.404
Sample Size	1711	
Number of parameters	16	
Log-likelihood		
At convergence	-2076.249	
At market share	-2354.340	
At zero	-2371.950	

Table 5.7 Work Tour Model
(Auto Mode Choice → Complex Tour)

Variable	Parameter	t-test
Auto Mode Choice Model		
Constant	-0.6690	-8.596
MALE	0.1269	1.671
FREEPARK	1.3003	16.783
DIS_WORK	0.0081	3.546
COUNTRY	0.5346	4.541
FULLTIME	0.2229	2.519
Complex Tour Choice Model		
Constant	-0.5924	-4.765
HHSIZE1	0.1465	2.085
HIGH_INC	0.3003	3.693
OWN_BUS	0.2465	2.637
SWISS	0.2997	3.358
BEG6_8	0.3193	4.183
BEG13_14	-0.4198	-3.541
END12	-0.7380	-7.319
AUTO	0.2715	2.003
ρ (Error Correlation)	0.0050	0.054
Sample Size	1711	
Number of parameters	16	
Log-likelihood		
At convergence	-2078.843	
At market share	-2354.340	
At zero	-2371.950	

Table 5.8 Work Tour Model
Simultaneous Logit Model (Auto Mode Choice ↔ Complex Tour)

Variable	Parameter	t-test
Auto Mode Choice Model		
Constant	-1.2268	-9.351
MALE	0.2470	2.022
FREEPARK	2.1867	16.104
COUNTRY	0.9025	4.689
FULLTIME	0.3679	2.592
Complex Tour Choice Model		
Constant	-1.0245	-5.996
HHSIZE1	0.2917	2.498
HIGH_INC	0.4818	3.671
OWN_BUS	0.4424	2.967
SWISS	0.4495	3.076
OWN_RES	0.1666	1.401
BEG6_8	0.5174	4.184
BEG13_14	-0.6688	-3.498
END12	-1.2212	-7.286
α (Joint Dependence)	0.4859	4.906
Sample Size	1711	
Number of parameters	15	
Log-likelihood		
At convergence	-2079.840	
At market share	-2354.340	
At zero	-2371.950	

Individuals owning their business enterprise and place of residence are found to be more prone to engage in multi-stop trip chains, presumably for seeing multiple clients, visiting to office or supply stores, etc. A few variables were found to be statistically significant, but are potentially more difficult to interpret. For example, individuals of Swiss Nationality are more likely to engage in complex work tours. It is possible that these individuals have occupational characteristics that lead to the formation of complex trip chains. Another interesting finding is that time-of-day indicators play an important role in influencing tour complexity. Tours ending within the lunch hour are less prone to be complex possibly due to time constraints and the presence of a single lunch stop/destination. However, those beginning in the morning period of 6 to 9 AM are more prone to being multi-stop trip chains, possibly due to the linking of a non-work activity with the work activity in the overall tour. A more detailed time-of-day based analysis of trip chain formation is warranted to fully understand the relationship between trip chain complexity and time of day choice behavior. Within the context of this study, time of day choice is assumed exogenous to the model system. However, one may argue that time of day choice is endogenous to trip chain complexity and mode choice. The study of the simultaneous causal relationships among trip chain formation, mode choice, and time of day choice (three endogenous entities) remains a future research effort.

Table 5.7 gives estimation resuls for the model where mode choice affects work tour complexity. As before, the coefficient associated with the auto mode choice variable in the tour complexity equation is statistically significant and positive indicating that the choice of auto mode contributes positively to the formation of complex multi-stop trip

chains. However, unlike other models, the error correlation is statistically insignificant. Thus, this model suggests that tour complexity and mode choice can be modeled as two independent equations where mode choice affects tour complexity in a recursive unidirectional causal structure. While this may be possible, it is highly unlikely to be true given the findings suggested by the previous models where the error correlation is consistently significant and negative in value. Thus, this model is suggesting that the unobserved factors in the two equations are uncorrelated and challenging the assumption of simultaneity in the relationship between auto mode choice and tour complexity. Given that this model contains no additional explanatory variables or power than the other previous models, the authors feel that a rejection of the assumption of simultaneity is not warranted.   Despite the very different value of the error correlation between models presented in Tables 5.6 and 5.7, virtually all of the other variables show similar indications between the two model structures.

Table 5.8 furnishes estimation results of the simultaneous logit model for work tours. The joint dependence parameter, α, is found to be statistically significant and positive. This model supports the notion that there is a significant and positive bidirectional causal relationship between tour complexity and auto mode choice. All of the other explanatory variables are found to offer indications very similar to those seen in Tables 5.6 and 5.7.

CHAPTER 6

MODEL PERFORMANCE COMPARISONS

6.1 Goodness-of-Fit Measure

The model estimation results presented in Chapter 5 generally offer plausible indications for alternative causal paradigms. The only model that may be rejected on qualitative grounds is that in Table 5.7 where the mode choice decision precedes the tour complexity decision. The statistically insignificant random error correlation which implies that there are no correlated unobserved factors between mode choice and tour complexity appears difficult to explain and defend in light of the simultaneity shown by the other models. This chapter presents a more rigorous comparison across models to see if it is possible to identify the most likely causal structure governing the relationship between mode choice and trip chaining.

A goodness-of-fit comparison among the models of different causal structures is conducted first. The adjusted likelihood ratio index as a goodness-of-fit measure can be used for testing and comparing non-nested relationships in discrete choice models. The indices are given as follows:

$$\bar{\rho}_0^2 = 1 - \frac{L(\beta) - K_0}{L(0)} \qquad\qquad 6.1$$

$$\bar{\rho}_c^2 = 1 - \frac{L(\beta) - K_c}{L(c)} \qquad\qquad 6.2$$

$\bar{\rho}_0^2$: Adjusted likelihood ratio index at zero

$\bar{\rho}_c^2$: Adjusted likelihood ratio index at market share

$L(\beta)$: Log-likelihood value at convergence

$L(0)$: Log-likelihood value at zero

$L(c)$: Log-likelihood value at market share (model including only the constant term)

K_0 and K_c: the number of parameters in the corresponding model.

The adjusted likelihood ratio indices for all of the models are presented in Tables 6.1 and 6.2.

Table 6.1 Likelihood Ratio Comparison in Non-work Tour Models
(N = 4901)

Causal Structure	Number of Parameters (K)	ρ_0^2	ρ_c^2	$\bar{\rho}_0^2$	$\bar{\rho}_c^2$
Complexity → Auto	20	0.238	0.097	0.235	0.093
Auto → Complexity	18	0.234	0.092	0.231	0.089
Complexity ↔ Auto	19	0.234	0.093	0.231	0.089

Table 6.2 Likelihood Ratio Comparison in Work Tour Models
(N = 1711)

Causal Structure	Number of Parameters (K)	ρ_0^2	ρ_c^2	$\bar{\rho}_0^2$	$\bar{\rho}_c^2$
Complexity → Auto	16	0.125	0.118	0.118	0.111
Auto → Complexity	16	0.124	0.117	0.117	0.111
Complexity ↔ Auto	15	0.123	0.117	0.117	0.110

6.2 Non-nested Test

Likelihood ratio test, a statistical test of the goodness-of-fit between two models, is widely applied for selecting more appropriate models estimated by maximum likelihood method. A relatively more complex model is compared to a simpler model to see if it fits the dataset significantly better. If so, the more complex model is considered as the better one. The likelihood ratio test is given as below.

$$LR = -2\,(\ln L_R - \ln L_U) \sim \chi^2(N) \qquad\qquad 6.3$$

under the null hypothesis that the restrictions on unrestricted model are jointly valid.

LR: Log-likelihood Ratio

$\ln L_R$: Log-likelihood function value for restricted model

$\ln L_u$: Log-likelihood function value for unrestricted model

N: number of restrictions imposed on the parameters in unrestricted model to achieve restricted model.

However, likelihood ratio test is only valid if it is used to compare hierarchically nested models. That is, the more complex model must differ from the simple model only by the

additional restrictions on the existing parameters in complex model. Restricted model should be developed by imposing restriction on the more general unrestricted model. In this thesis, the models in three different causal structures need to be compared one another, but no pair models of them are hierarchically nested since none model can be formed from either of the other two models with additional restriction.

The adjusted likelihood ratio index $\overline{\rho}_0^2$ as a goodness-of-fit measure can be used for testing non-nested hypothesis of discrete choice modes. To choose between two models (say, 1 and 2), Ben-akiva and Lerman (1985, p. 172) provide a test where under the null hypothesis that model 1 is the true specification, the following holds asymptotically:

$$\Pr(\; \overline{\rho}_2^2 - \overline{\rho}_1^2 > z\;) \le \Phi\;\{-[-2\,zL\,(0) + (K_2 - K_1)]^{1/2}\},\; z > 0 \qquad 6.4$$

where

$\overline{\rho}_i^2$: the adjusted likelihood ratio index at zero for model $i = 1, 2$

K_i : the number of parameters in model i

Φ : the standard normal cumulative distribution function

$L(0)$: log-likelihood value at zero;

        if all N observations in the sample have all J alternatives, $L(0) = N \ln(1/J)$.

The probability that the adjusted likelihood ratio index of model 2 is greater by some $z >$ 0 than that of model 1, given that the latter is the true model, is asymptotically bounded by the right-hand side of equation 6.4 above. If the model with the greater $\overline{\rho}^2$ is selected, then this bounds the probability of erroneously choosing the incorrect model over the true

specification. Using this procedure, models of alternative causal structures can be compared against one another.

6.3 Comparison Results

For non-work tour models, the difference in adjusted likelihood ratios is approximately 0.004 between the models in Table 5.2 and the models in Tables 5.3 and 5.4. According to equation 6.4, the calculated bounding probability on the right hand side of the expression is almost zero. Thus, it may be concluded that the model of Table 5.2 is more closely capturing the causal structure underlying the relationship between mode choice and tour complexity. The significantly better goodness-of-fit of the model in Table 5.2 suggests that the causal structure where the complexity of the tour affects mode choice (tour complexity → auto mode choice) is statistically, and possibly behaviorally, dominant in the population for non-work tours. One must be careful when drawing inferences regarding behavioral causality from statistical indicators.

For work tour models, however, the situation is not as clear. In comparing the models, the seemingly better model of Table 5.6 has an adjusted likelihood ratio index that is only 0.001 greater than those of the models in the other two causal structures. The bounding probabilities, as per the right hand side of equation 6.4, are calculated to be 0.036 and 0.067, respectively. The statistical test rejects the model of Table 5.7, i.e., the causal structure where auto mode choice drives the complexity of the work tour (auto mode choice → tour complexity). However, the test fails to reject the simultaneous logit model, i.e., bidirectional simultaneous causality, at the 0.05 level of significance. In addition to

this non-nested test, the insignificance of the random error correlation in the model of auto mode choice affecting tour complexity suggests that the assumed causal structure in that joint model may not be valid and may be indicative of the irrationality of causality in that direction. Thus, for work tours, two possible causal structures can not be rejected from this analysis. Either, the decision to make a complex work tour tends to result in the choice of the auto mode or both of these decisions are made contemporaneously.

From the viewpoint of activity-based travel behavior theory where travel choices are considered to be derived from activity patterns (and activity needs that are distributed in time and space), one may consider the findings of this thesis to be quite consistent with expectations. For non-work tours, the statistical model estimation results show that tour complexity (which is reflective of the activity pattern) drives mode choice. For work tours, the statistical model estimation results reject the notion that auto mode choice drives tour complexity. Once again, either tour complexity drives auto mode choice or both decisions are simultaneously related to one another.

CHAPTER 7

DISCUSSION AND CONCLUSION

7.1 Discussion

Mode choice behavior is a fundamental element of travel behavior that has significant implications for transportation planning. Estimates of public transit ridership and the use of alternative modes of transportation are largely based on studies of mode choice behavior and modal split models. Public transport agencies face increasing competition from the automobile as automobiles become increasingly affordable and the road infrastructure becomes increasingly ubiquitous. Undoubtedly, the automobile is considered to provide greater flexibility and convenience when compared with public transport modes that are generally constrained with respect to schedules and routes/destinations.

This study examines the inter-relationship between the complexity of people's activity-travel patterns and their mode choice. In order to conduct the analysis, this thesis examines mode choice behavior in the context of multi-stop (complex) vs single-stop (simple) trip chains. Through a series of econometric model formulations, this thesis presents a rigorous analysis of the most likely causal relationship between these two phenomena at the level of the individual trip chain or tour. It should be emphasized that

the analysis in this thesis does not attempt to replicate causality at the level of the individual traveler, but rather at the macroscopic level to identify the causal tendency that appears to be dominant in the population.

Using data derived from the 2000 Swiss Travel Microcensus, the thesis estimates bivariate probit models and simultaneous logit models that provide a rigorous analytical framework for analyzing and testing alternative causal structures. In the case of non-work tours (i.e., tours that do not involve any work stops), the analysis suggests that the causal structure where the complexity of the trip chaining pattern drives mode choice is the dominating causal trend in the population. In the case of work tours (i.e., tours that involve at least one work stop), the analysis suggests that the causal structure where the auto mode choice drives the complexity of the tour can be rejected statistically. Thus, either tour complexity drives mode choice or the two phenomena occur simultaneously.

7.2 Contribution

These findings have important implications for public transport service providers who are interested in attracting choice riders. If mode choice decisions precede activity pattern/agenda decisions, then it may be possible for public transport service providers to simply attract choice riders by improving amenities, schedule, route coverage, safety and security, and comfort. On the other hand, if the formation of the activity agenda precedes mode choice decisions, then the public transport industry has a greater challenge before it. Trip chaining and tour complexity serve as impediments to public transport usage. The analysis in this thesis suggests that this is the predominant relationship in the data set.

Then, not only do public transport service providers have to improve service amenities, but they also have to try and cater to a multi-tasking oriented complex activity agenda. This is extremely difficult to do with a fixed route fixed schedule system. As activity-travel patterns and tours become increasingly complex, it is likely that public transport agencies will have to develop new types of services to try and retain existing riders in addition to attracting new riders.

The analysis and findings of this thesis are also useful from an activity-based and tour-based model development standpoint. Most activity-based and tour-based travel demand model systems consist of hierarchical structures involving activity agenda or tour formation, mode choice, destination choice, and time of day choice. The development and application of these model systems calls for the ability to accurately represent causal relationships that are prevalent in the population. This thesis suggests that the activity agenda or tour formation step may precede the mode choice step for both non-work and work tours, although the hierarchy is less clear for the latter.

7.3 Future Research

Future research efforts should focus on analyzing whether these findings regarding causal relationships between tour complexity and mode choice hold in other data sets as well. In addition, the modeling framework can be extended to consider multinomial choice situations as opposed to pure binary choice variables considered in this thesis. Mode choice can be expanded to consider multiple modes including SOV, shared ride, public transit, and non-motorized modes. Similarly, tour complexity can be expanded to

consider different levels of tour complexity or different tour types such as that presented in Strathman and Dueker (1995). Another consideration that merits further investigation is the extent to which findings such as those presented in this thesis are sensitive to model specification. It is possible that statistical indicators of model performance will change depending on the model specification chosen. Such efforts would further aid in understanding important causal relationships underlying travel behavior.

Finally, it must be noted that causal relationships are being extracted and examined in this thesis from statistical relationships estimated on revealed outcome data. While such data provides insights into what people have done, it does not provide true insights into the decision mechanisms and behavioral processes underlying the revealed outcomes. In order to truly understand and identify causal relationships, data regarding processes and decision mechanisms are needed. Future research into the development of microsimulation models of activity and travel behavior should include attempts to collect and analyze such data.

REFERENCES

Aptech (1996) *GAUSS 3.2*, Aptech Systems, Maple Valley, Washington.

Ben-Akiva, M. and Lerman, S.R. (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*. The MIT Press, Cambridge, MA.

Bhat, C.R. (1997) Work Travel Mode Choice and Number of Non-work Commute Stops. *Transportation Research B* 31(1), pp. 41-54.

Dissanayake, D. and Morikawa, T. (2002) Household Travel Behavior in Developing Countries: Nested Logit Model of Vehicle Ownership, Mode Choice, and Trip Chaining, *Transportation Research Record* 1805, *Journal of the Transportation Research Board*, National Research Council, Washington, D.C., pp. 45-52.

Fujii, S., Kitamura, R., and Kishizawa, K. (1999) Analysis of Individuals' Joint Activity Engagement Using a Model System of Activity-travel Behavior and Time Use, *Transportation Research Record* 1676, pp. 11-19.

Fujii, S. and Kitamura, R. (2000) Evaluation of Trip-inducing Effects of New Freeways Using a Structural Equations Model System of Commuters' Time use and Travel, *Transportation Research B* 34, pp. 339-354.

Golob, T.F. and E.R. Zondag (1984). *A causal model of mobility*. Presented at the Tenth Transportation Planning Research Colloquium. Zandvoort, the Netherlands.

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

Golob, T.F. and L.J. van Wissen (1989). A joint household travel distance generation and car ownership model. *Transportation Research B* 23: 471-491.

Golob, T.F., R. Kitamura and G. Occhiuzzo (1991). An attitude-behavioral intention model of the market potential for alternative-fuel vehicles: the "green" segment and the role of information. Presented at *Annual Meeting of Transportation search Board,* Washington, D.C.

Golob, T.F. and D. Brownstone (1992). A dynamic attitude-behavior model of consumer reactions to incentives for restricting solo driving trips for environmental reasons. Presented at *6th World Conference on Transport Research*, Lyon, France, June 29-July 3.

Golob, T.F. R. Kitamura and C. Lula (1994). Modeling the Effects of Commuting Time on Activity Duration and Non-Work Travel. Presented at *Annual Meeting of Transportation Research Board*, Washington, D.C.

Golob, T.F. and McNally, M.G. (1997) A Model of Activity Participation and Travel Interactions Between Household Heads, *Transportation Research B* 31(3), pp. 177-194.

Golob, T.F (2000). A Simultaneous Model of Household Activity Participation and Trip Chain Generation. *Transportation Research B* 34, pp. 355-376.

Golob T.F. (2000) A Simultaneous Model of Household Activity Participation and Trip Chain Generation, *Transportation Research B* 34, pp. 355-376.

Greene, W.H. (2003) *Econometric Analysis,* Fifth Edition*,* Pearson Education, Inc., NJ.

Greene, W.H. (2002) *LIMDEP Version 8.0: User's Manual*, Econometric Software, Inc., 2002, Plainview, NY.

Hensher, D.A. and Reyes A.J. (2000) Trip Chaining as a Barrier to the Propensity to Use Public Transport. *Transportation* 27, pp. 341-361.

Kitamura, R. (1988). An evaluation of activity-based travel analysis. *Transportation* 15, pp. 9-34.

Kitamura, R., J.P. Robinson, T.F. Golob, M.A. Bradley, J. Leonard and T. van der Hoorn (1992). A comparative analysis of time use data in the Netherlands and California. In Proceedings of the 20th PTRC Summer Annual Meeting: *Transportation Planning Methods*, pp. 127-138.

Kitamura, R. (1996). Activity-based travel demand forecasting and policy analysis. Presented at *the TMIP Conference on Activity-Based Travel Forecasting*, New Orleans, LA, June 2-5.

Kitamura, R., C. Chen, and R.M.Pendyala (1997). Generation of Synthetic Daily Activity-Travel Patterns. *Transportation Research Record* 1607, pp. 154-162.

Kitamura, R., Fujii, S. and Pas, E.I. (1997) Time-use Data, Analysis and Modeling: Toward the Next Generation of Transportation Planning Methodologies. *Transport Policy* 4(4), pp. 225-235.

Kitamura, R., C. Chen, and R. Narayanan (1998). The Effects of Time of Day, Activity Duration and Home Location on Travelers' Destination Choice Behavior. Paper Presented at *the 77th Annual Meeting of Transportation Research Board*, Washington, D.C.

Koppelman, F.S. and Wen, C. (1998) Alternative Nested Logit Models: Structure, Properties and Estimation, *Transportation Research B* 32(5), pp. 289-298.

Kuppam, A.R. (1999). A Limited Dependent Variable Structural Equations Model of Travel Behavior for Transportation Policy Evaluation. PhD Dissertation, Department of Civil and Environmental Engineering, University of South Florida, Tampa, FL.

Kuppam, A.R. and R.M. Pendyala (2000). An exploratory analysis of commuters' activity and travel patterns. Presented at the *79th Annual Meeting of the Transportation research Board*, Jan. 9-13, Washington, DC.

Lu, X. and E.I. Pas (1997). A structural equations model of the relationships mong socio-demographics, activity participation and travel behavior. Presented at *76th Annual Meeting of the Transportation Research Board*, January 12-16, Washington, DC.

Lu, X. and Pas, E.I. (1999). Socio-demographics, Activity Participation and Travel Behaviour, *Transportation Research A* 33, pp. 1-18.

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

McGuckin, N. and Murakami, E. (1999) Examining Trip-Chaining Behavior: A Comparison of Travel by Men and Women, *Transportation Research Record* 1693, *Journal of the Transportation Research Board,* National Research Council, Washington, D.C., pp. 79-85.

Meka, S. (2003). A Structural Equations Analysis of Within-Household Activity and Time Allocation Between Two Adults, Master Thesis, Department of Civil and Environmental Engineering, University of South Florida, Tampa, FL.

Misra, R., Bhat, C. and Srinivasan, S. (2003) A Continuous Time Representation and Modeling Framework for the Analysis of Nonworker Activity-Travel Patterns: Tour and Episode Attributes. *CD-ROM of the 82nd Annual Meeting of the Transportation Research Board*, National Research Council, Washington, D.C.

Ouyang, Y., Shankar, V. and Yamamoto, T. (2002) Modeling the Simultaneity in Injury Causation in Multi-vehicle Collisions. *CD-ROM of the 81st Annual Meeting of the Transportation Research Board,* National Research Council, Washington, D.C.

Pas, E.I. (1996). Recent advances in activity-based travel demand modeling. Presented at *the TMIP Conference on Activity-Based Travel Forecasting*, New Orleans, LA, June 2-5.

Pas, E. & Harvey. A.S. (1997). *Understanding travel behavior in an era of change*. (Eds.) Peter R. Stopher.and Martin Lee-Gosselin. New York: Pergamon Press.

Pendyala, R.M., Bhat, C., Parashar, A. and Muthyalagari, G.R. (2001) An Exploration of the Relationship between Timing and Duration of Maintenance Activities. *CD-ROM of the 81st Annual Meeting of the Transportation Research Board,* National Research Council, Washington, D.C.

RDC, Inc. (1995). Activity-Based Modeling System for Travel Demand Forecasting, TMIP.

Schmidt, P. and Strauss, R.P. (1975) Estimation of Models with Jointly Dependent Qualitative Variables: A Simultaneous Logit Approach. *Econometrica* 43(4), pp. 745-755.

Shiftan, Y. (1998) Practical Approach to Model Trip Chaining, *Transportation Research Record* 1645, Transportation Research Board, National Research Council, Washington, D.C., pp. 17-23.

Simma, A. and Axhausen, K.W. (2001). Within-Household Allocation of Travel: The Case of Upper Austria. Presented at *the 80th Annual Meeting of the Transportation Research Board*, Washington, D.C.

Strathman J.G. and Dueker K.J. (1994) Effect of Household Structure and Selected Characteristics on Trip Chaining. *Transportation* 21, pp. 23-45.

Strathman, J.G. and Dueker, K.J. (1995) Understanding Trip Chaining, *Special Reports on Trip and Vehicle Attributes*, 1990 NPTS Reports Series, Publication No. FHWA-PL-95-033, U.S. Department of Transportation, pp. 1-1 – 1-27.

Ye, X. and Pendyala, R.M. (2003) Description of the Switzerland Microcensus 2000 Travel Survey Sample. Research Report prepared for Jenni + Gottardi AG, Department of Civil and Environmental Engineering, University of South Florida, Tampa, FL.

APPENDICES

Appendix A: Gauss Codes for Simultaneous Logit Model of Non-work Tours

library maxlik;
maxset;

load data[4901,131] = "C:\\swiss\\multichoice\\non-work_tours_zurich_4901.dat";
one=ones(4901,1);
intnr   =   data[., 1   ];
tournum   =   data[., 2   ];
hhnr   =   data[., 3   ];
tripdist=   data[., 4   ];
tripdur =   data[., 5   ];
beg_time   =   data[., 6   ];
end_time   =   data[., 7   ];
nsegment   =   data[., 8   ];
trip_pur   =   data[., 9   ];
mode_dur   =   data[., 10   ];
mode_dis   =   data[., 11   ];
person =   data[., 12   ];
targetpn   =   data[., 13   ];
intdur =   data[., 14   ];
weekday   =   data[., 15   ];
day   =   data[., 16   ];
season =   data[., 17   ];
age   =   data[., 18   ];
sex   =   data[., 19   ];
employed   =   data[., 20   ];
study   =   data[., 21   ];
auto_lic   =   data[., 22   ];
motr_lic   =   data[., 23   ];
live_st =   data[., 24   ];
national   =   data[., 25   ];
study16   =   data[., 26   ];
edu_lev   =   data[., 27   ];
emp_sit   =   data[., 28   ];
school =   data[., 29   ];
emp_reg   =   data[., 30   ];
worktime   =   data[., 31   ];
occ_posi   =   data[., 32   ];
dis_work   =   data[., 33   ];
pklot_wk   =   data[., 34   ];
n_trp_wk   =   data[., 35   ];

Appendix A: (Continued)

```
dis_schl     =      data[., 36      ];
pklot_sc     =      data[., 37      ];
n_tr_sch     =      data[., 38      ];
av_bike      =      data[., 39      ];
av_egbik     =      data[., 40      ];
av_motor     =      data[., 41      ];
av_auto      =      data[., 42      ];
weather      =      data[., 43      ];
workday      =      data[., 44      ];
leavehom     =      data[., 45      ];
why_nlv      =      data[., 46      ];
when_lv      =      data[., 47      ];
comp_pwk     =      data[., 48      ];
kilo_pwk     =      data[., 49      ];
y_ljny_r     =      data[., 50      ];
n_jny_3m     =      data[., 51      ];
m_ljny_r     =      data[., 52      ];
d_ljny_r     =      data[., 53      ];
n_nights     =      data[., 54      ];
means_r      =      data[., 55      ];
ds_jny_i     =      data[., 56      ];
ds_jny_e     =      data[., 57      ];
air_12m      =      data[., 58      ];
air_5y =     data[., 59      ];
n_air_5y     =      data[., 60      ];
y_l_air=     data[., 61      ];
air_pur      =      data[., 62      ];
hol_trp=     data[., 63      ];
pck_tour     =      data[., 64      ];
mn_airpt     =      data[., 65      ];
y_tkoff =    data[., 66      ];
m_tkoff=     data[., 67      ];
d_tkoff =    data[., 68      ];
nair_12m     =      data[., 69      ];
n_ck_12m     =      data[., 70      ];
hhld_w=      data[., 71      ];
canton =     data[., 72      ];
hhdate =     data[., 73      ];
language     =      data[., 74      ];
city_typ     =      data[., 75      ];
city_rur     =      data[., 76      ];
hhsize  =    data[., 77      ];
```

Appendix A: (Continued)

```
hh6plus      =      data[., 78     ];
n_targtp     =      data[., 79     ];
hhincome     =      data[., 80     ];
perm_add     =      data[., 81     ];
add_swit     =      data[., 82     ];
y_in_add     =      data[., 83     ];
faml_typ     =      data[., 84     ];
rent_own     =      data[., 85     ];
n_apartm     =      data[., 86     ];
n_2ndhom     =      data[., 87     ];
own_park     =      data[., 88     ];
n_park =      data[., 89     ];
n_auto =      data[., 90     ];
n_motor      =      data[., 91     ];
n_smotor     =      data[., 92     ];
n_engbik     =      data[., 93     ];
n_bike =      data[., 94     ];
n_bik_lc     =      data[., 95     ];
bad_ch_1     =      data[., 96     ];
nwk_ch_1     =      data[., 97     ];
work_c_1     =      data[., 98     ];
sim_comx     =      data[., 99     ];
pr_mode4     =      data[., 100    ];
uni_mult     =      data[., 101    ];
kid    =      data[., 102    ];
kid5   =      data[., 103    ];
kid6_17      =      data[., 104    ];
dri_lic =     data[., 105    ];
age0_5=      data[., 106    ];
age6_17      =      data[., 107    ];
age18_24     =      data[., 108    ];
age25_34     =      data[., 109    ];
age35_44     =      data[., 110    ];
age45_54     =      data[., 111    ];
age55_64     =      data[., 112    ];
age65_74     =      data[., 113    ];
age75  =      data[., 114    ];
male   =      data[., 115    ];
female =      data[., 116    ];
pr_pur =      data[., 117    ];
nwf_firs     =      data[., 118    ];
nwf_mid      =      data[., 119    ];
```

Appendix A: (Continued)

```
nwf_last     =       data[., 120    ];
chain_tp     =       data[., 121    ];
beg_c   =    data[., 122    ];
end_c   =    data[., 123    ];
pmode5       =       data[., 124    ];
mode3   =    data[., 125    ];
auto    =    data[., 126    ];
complx  =    data[., 127    ];
tourmode     =       data[., 128    ];
m_sov   =    data[., 129    ];
m_hov   =    data[., 130    ];
m_other      =       data[., 131    ];
/* dummy variables definition */

shopping =    (pr_pur .==4) ;
leisure =     (pr_pur .==5) ;
service =     (pr_pur .==3) ;
pschool =     (pr_pur .==2) ;
business =    (pr_pur .==1) ;
pmale =       (sex .==1)    ;
old =   (age.>60)      ;
young =       (age.<18 .and age.>0)        ;
high_inc =    (hhincome.>=6)       ;
car_0 =       (n_auto .==0) ;
car_1 =       (n_auto .==1) ;
car_ge2 =     (n_auto.>=2) ;
low_inc =     (hhincome .<3 .and hhincome.>0)    ;
with_kid =    ((hhsize-hh6plus).>0) ;
hhsize1 =     (hhsize .==1) ;
hhsize2 =     (hhsize .==2) ;
hhsize3 =     (hhsize .==3) ;
hhsize4 =     (hhsize.>3)    ;
high_edu =    (EDU_LEV .==6 .or EDU_LEV .==7)        ;
with_kid =    (kid5.>0)      ;
with_you =    (kid6_17.>0)  ;
weekend =     (weekday .==5)       ;
bad_weat =    (weather.>=4 .and weather.<=7)     ;
rain = (weather .==5)       ;
working =     (workday .==1)       ;
ad1_you1= (kid6_17.>0 .and hhsize.==2)  ;
ampeak=(beg_c.==7 .or beg_c.==8);
pmpeak=(beg_c.==16 .or beg_c.==17);
```

Appendix A: (Continued)

```
proc lpr(b,z);

local x,y,ux,uy,alpha,delta,p00,p01,p10,p11;

x = one~hhsize~high_inc~low_inc~car_ge2~pmale~service~pschool;
y = one~hhsize1~ad1_you1~hhsize4~old~young~service~shopping~ampeak~pmpeak;

ux=x*b[1:8,.];     /*  utility for auto */
uy=y*b[9:18,.];   /*  utility for complex chain */

alpha=b[19,.];     /* joint dependence */

delta= 1 + exp(ux) + exp(uy) + exp(ux + uy + alpha) ;

p00=one./delta;           /*  probability of non-auto & simple   */
p01=exp(uy)./delta;     /*  probability of non-auto & complex   */
p10=exp(ux)./delta;     /*  probability of  auto & simple   */
p11= exp(ux+uy+alpha)./delta;       /*  probability of auto & complex   */

retp (     (1-auto) .* (1-complx).*ln( p00 )
       + (1-auto) .*  complx.*ln( p01 )
       +     auto .*  (1-complx).*ln( p10 )
       +     auto .*  complx.*ln( p11 )          ) ;

endp;

b0={0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0};

{b,f,g,cov,ret}=maxlik(data,0,&lpr,b0);

call maxprt(b,f,g,cov,ret);
```

Appendix B: Gauss Codes for Simultaneous Logit Model of Work Tours

```
library maxlik;
maxset;
load data[1711,131] = "C:\\swiss\\multichoice\\work_tours_zurich_1711.dat";
one=ones(1711,1);
intnr    =       data[., 1      ];
tournum          =       data[., 2      ];
hhnr     =       data[., 3      ];
tripdist=        data[., 4      ];
tripdur =        data[., 5      ];
beg_time         =       data[., 6      ];
end_time         =       data[., 7      ];
nsegment         =       data[., 8      ];
trip_pur         =       data[., 9      ];
mode_dur         =       data[., 10     ];
mode_dis         =       data[., 11     ];
person =         data[., 12     ];
targetpn         =       data[., 13     ];
intdur  =        data[., 14     ];
weekday          =       data[., 15     ];
day      =       data[., 16     ];
season  =        data[., 17     ];
age      =       data[., 18     ];
sex      =       data[., 19     ];
employed         =       data[., 20     ];
study    =       data[., 21     ];
auto_lic         =       data[., 22     ];
motr_lic         =       data[., 23     ];
live_st =        data[., 24     ];
national         =       data[., 25     ];
study16          =       data[., 26     ];
edu_lev          =       data[., 27     ];
emp_sit          =       data[., 28     ];
school  =        data[., 29     ];
emp_reg          =       data[., 30     ];
worktime         =       data[., 31     ];
occ_posi         =       data[., 32     ];
dis_work         =       data[., 33     ];
pklot_wk         =       data[., 34     ];
n_trp_wk         =       data[., 35     ];
dis_schl         =       data[., 36     ];
pklot_sc         =       data[., 37     ];
n_tr_sch         =       data[., 38     ];
```

Appendix B: (Continued)

```
av_bike      =      data[., 39      ];
av_egbik     =      data[., 40      ];
av_motor     =      data[., 41      ];
av_auto      =      data[., 42      ];
weather      =      data[., 43      ];
workday      =      data[., 44      ];
leavehom     =      data[., 45      ];
why_nlv      =      data[., 46      ];
when_lv      =      data[., 47      ];
comp_pwk     =      data[., 48      ];
kilo_pwk     =      data[., 49      ];
y_ljny_r     =      data[., 50      ];
n_jny_3m     =      data[., 51      ];
m_ljny_r     =      data[., 52      ];
d_ljny_r     =      data[., 53      ];
n_nights     =      data[., 54      ];
means_r      =      data[., 55      ];
ds_jny_i     =      data[., 56      ];
ds_jny_e     =      data[., 57      ];
air_12m      =      data[., 58      ];
air_5y =     data[., 59      ];
n_air_5y     =      data[., 60      ];
y_l_air=     data[., 61      ];
air_pur      =      data[., 62      ];
hol_trp=     data[., 63      ];
pck_tour     =      data[., 64      ];
mn_airpt     =      data[., 65      ];
y_tkoff =    data[., 66      ];
m_tkoff=     data[., 67      ];
d_tkoff =    data[., 68      ];
nair_12m     =      data[., 69      ];
n_ck_12m     =      data[., 70      ];
hhld_w=      data[., 71      ];
canton =     data[., 72      ];
hhdate =     data[., 73      ];
language     =      data[., 74      ];
city_typ     =      data[., 75      ];
city_rur     =      data[., 76      ];
hhsize  =    data[., 77      ];
hh6plus      =      data[., 78      ];
n_targtp     =      data[., 79      ];
hhincome     =      data[., 80      ];
```

Appendix B: (Continued)

```
perm_add    =       data[., 81      ];
add_swit    =       data[., 82      ];
y_in_add    =       data[., 83      ];
faml_typ    =       data[., 84      ];
rent_own    =       data[., 85      ];
n_apartm    =       data[., 86      ];
n_2ndhom    =       data[., 87      ];
own_park    =       data[., 88      ];
n_park =    data[., 89      ];
n_auto =    data[., 90      ];
n_motor     =       data[., 91      ];
n_smotor    =       data[., 92      ];
n_engbik    =       data[., 93      ];
n_bike =    data[., 94      ];
n_bik_lc    =       data[., 95      ];
bad_ch_1    =       data[., 96      ];
nwk_ch_1    =       data[., 97      ];
work_c_1    =       data[., 98      ];
sim_comx    =       data[., 99      ];
pr_mode4    =       data[., 100     ];
uni_mult    =       data[., 101     ];
kid     =   data[., 102     ];
kid5    =   data[., 103     ];
kid6_17     =       data[., 104     ];
dri_lic =   data[., 105     ];
age0_5=     data[., 106     ];
age6_17     =       data[., 107     ];
age18_24    =       data[., 108     ];
age25_34    =       data[., 109     ];
age35_44    =       data[., 110     ];
age45_54    =       data[., 111     ];
age55_64    =       data[., 112     ];
age65_74    =       data[., 113     ];
age75  =    data[., 114     ];
male   =    data[., 115     ];
female =    data[., 116     ];
pr_pur =    data[., 117     ];
nwf_firs    =       data[., 118     ];
nwf_mid     =       data[., 119     ];
nwf_last    =       data[., 120     ];
chain_tp    =       data[., 121     ];
beg_c  =    data[., 122     ];
```

```
end_c   =       data[., 123    ];
pmode5      =       data[., 124    ];
mode3 =     data[., 125    ];
auto    =   data[., 126    ];
complx =    data[., 127    ];
tourmode    =       data[., 128    ];
m_sov =     data[., 129    ];
m_hov =     data[., 130    ];
m_other     =       data[., 131    ];

/* dummy variables definition */
pmale =     (sex .==1)     ;
old =   (age.>60)      ;
young =     (age.<18 .and age.>0)        ;
high_inc =  (hhincome.>=6)       ;
car_0 =     (n_auto .==0) ;
car_ge2 =   (n_auto.>=2) ;
low_inc =   (hhincome .<3 .and hhincome.>0)    ;
with_kid =  ((hhsize-hh6plus).>0) ;
hhsize1 =   (hhsize .==1) ;
hhsize2 =   (hhsize .==2) ;
hhsize3 =   (hhsize .==3) ;
hhsize4 =   (hhsize.>3)   ;
high_edu =  (EDU_LEV .==6 .or EDU_LEV .==7)      ;
with_kid =  (kid5.>0)      ;
with_you =  (kid6_17.>0)  ;
weekend =   (weekday .==5)        ;
bad_weat =  (weather.>=4 .and weather.<=7)   ;
rain =  (weather .==5)       ;
working =   (workday .==1)        ;

freepark= (pklot_wk .==1);
country= (city_rur.==2);
fulltime=(emp_sit .==1);
independ= (occ_posi .==1);
swiss= (national .==1);
beg6_8=(beg_c .==6 .or  beg_c .==7 .or beg_c .==8);
beg13_14=(beg_c .==13 .or  beg_c .==14);
end12 = (end_c .==12) ;
owner= (rent_own.==3);
```

Appendix B: (Continued)

```
proc lpr(b,z);

local x,y,ux,uy,alpha,delta,p00,p01,p10,p11;

x=one~pmale~freepark~country~fulltime;
y = one~hhsize1~high_inc~independ~swiss~owner~beg6_8~beg13_14~end12;

ux=x*b[1:5,.];      /*  utility for auto */
uy=y*b[6:14,.];  /*  utility for complex chain */
alpha=b[15,.];  /* joint dependence */

delta= 1 + exp(ux) + exp(uy) + exp(ux + uy + alpha) ;

p00=one./delta;           /*  probability of non-auto & simple   */
p01=exp(uy)./delta;    /*  probability of non-auto & complex   */
p10=exp(ux)./delta;     /*  probability of  auto & simple   */
p11= exp(ux+uy+alpha)./delta;       /*  probability of auto & complex   */

retp (      (1-auto) .* (1-complx).*ln( p00 )
       + (1-auto) .*  complx.*ln( p01 )
       +     auto .*  (1-complx).*ln( p10 )
       +     auto .*   complx.*ln( p11 )           ) ;
endp;

b0={0,0,0,0,0,  0,0,0,0,0,   0,0,0,0,0};
{b,f,g,cov,ret}=maxlik(data,0,&lpr,b0);

call maxprt(b,f,g,cov,ret);
```