제 26회 대한교통학회 학술발표회

ECONOMETRIC MODELING OF ALL-DAY ACTIVITY INVOLVEMENT AND DURATION WITH TRANSPORTATION PANEL DATA

(일기식 교통판넬 데이타를 이용한 전일활동의 계량모델 개발에 관한 연구)

1994년 10월

서울시정개발연구원

책임연구원 김 순관

I. INTRODUCTION

In the past, the application of activity-based travel demand models was limited to the static and cross-sectional sphere, with few exceptions, due to unestablished theory and lack of data. The introduction of dynamic features into travel demand models is a relatively recent phenomenon which reflects the shift in emphasis and widening of outlook that has taken place within transportation demand modeling in the past. The realization that peoples' trips result from activities, justifies the need to analyze activity behavior patterns as a basis for understanding the dynamic features of travel demand.

From this point of view, "dynamic" analysis of travel behavior follows the emphasis in the field of activity-based analysis on the time dimension and adaptation to changes in the travel environment (Kitamura, 1988). This emphasis is an outgrowth of the realization that activity modeling is the key to developing dynamic, comprehensive transportation models in which the all-important time dimension and individual diversity (i.e. heterogeneity) are explicitly accounted for. Having the capability to address the time dimension and heterogeneity can potentially provide suggestive information that could lead to the development and implementation of more effective transportation demand management policies (Mannering, Murakami, and Kim, 1994).

There has been a recent proliferation of activity-based studies that have sought to include time dimension and individual diversity. However, the absence of one universal theory that would be acceptable to everyone is a major reason for the lack of planning applications of activity-based analysis. Another reason for a low acceptance of activity-based approaches in the transportation field is its fragmental development. However, without doubt, existing research on activity-based travel modeling has provided important directions for achieving the goal of applied forecasting.

In this vein, two fundamental attempts are sought to further promote the case of activity-based travel behavior models:

- Fragmentary devotions to a sound methodological foundation focusing on state dependence and heterogeneity.
- Provide some empirical evidence relating to the temporal stability of activity-based models. To achieve this, the time stability of model parameters is tested by comparing two survey "waves" separated by one year.

From these view points, the objectives of the thesis research are threefold as follows:

- Provide a methodology which will help describe, explain, and forecast the allocation of all-day activities classified by workers/non-workers in the time dimension.
- Explore the role that state dependence and heterogeneity play in models of activity behavior by using transportation panel data; specifically models of home-stay duration. In the other models, no unobserved heterogeneity is assumed so that only the role of state dependence can be explored.
- Provide some empirical evidence relating to the temporal stability of activity-based models. To achieve this, the time stability of model parameters is tested by comparing two survey "waves" separated by one year.

II. ALL-DAY BEHAVIORAL MODEL FORMULATION

The modeling system is developed for a traveler who is making at least one trip each day in two consecutive days. There are three conceptual frameworks of trip generating activity involvement for nonwork, pre-work and post-work respectively. Figures 2.1-2.3 represents travelers' all-day activity patterns and models needed at each stage. All-day activity of workers can be splitted into pre-work, while-work, and post-work activity due to significant differences in activity patterns. However, while-work activities are out of the scope of this study and left for further research.

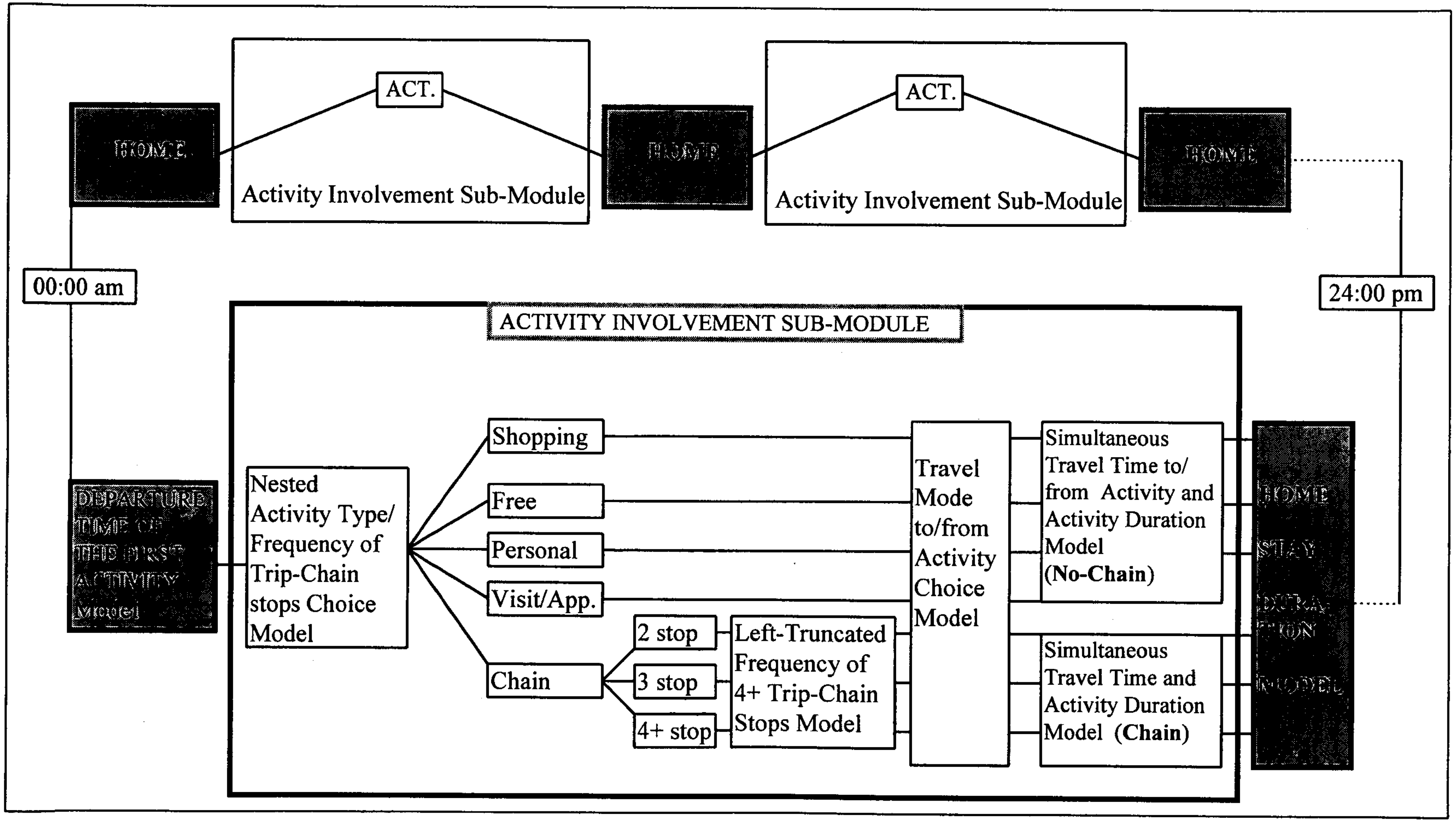

Fig. 2.1 Framework of all -day activity pattern of non-worker

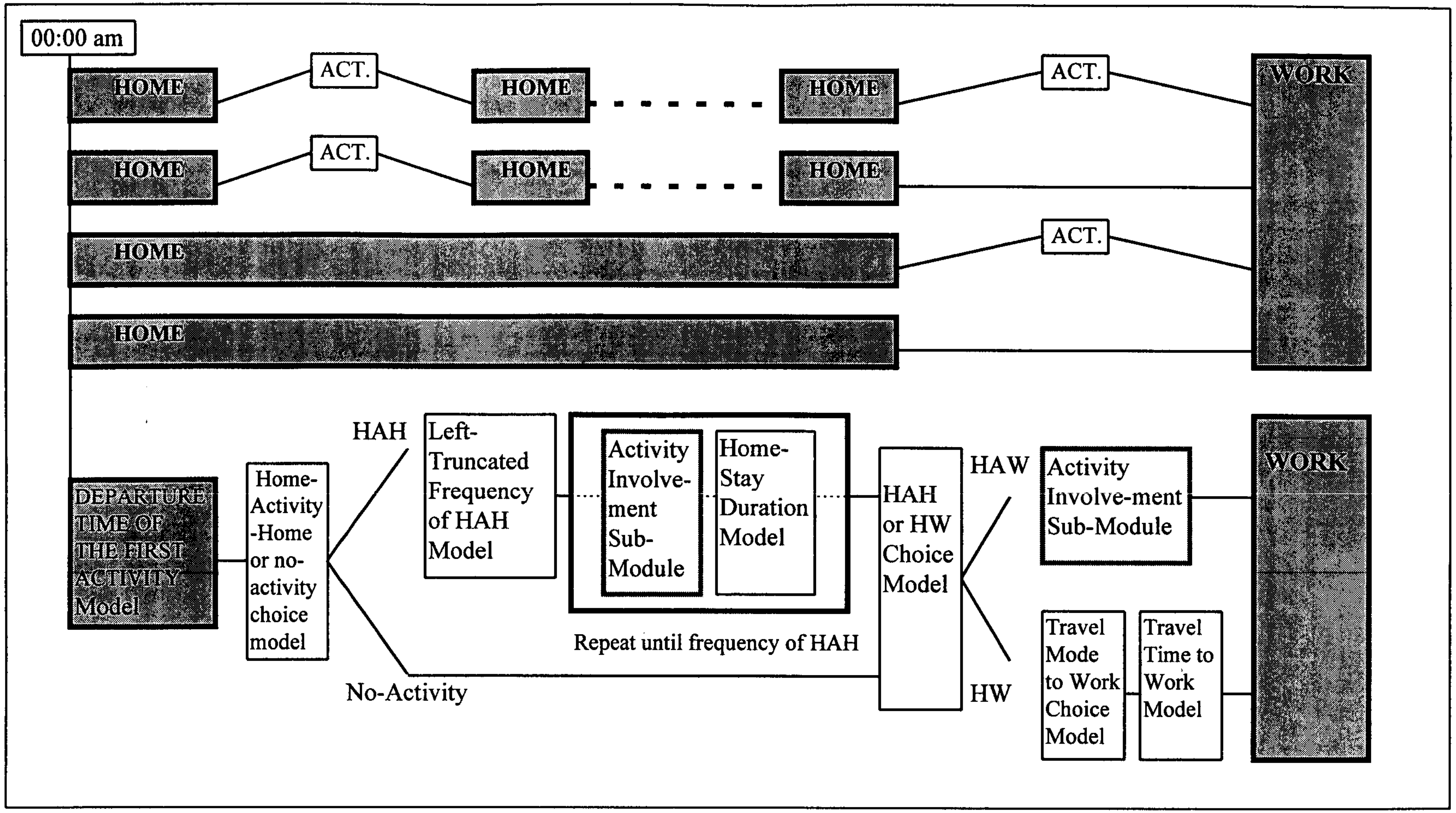

Fig. 2.2 Framework of pre-work activity pattern

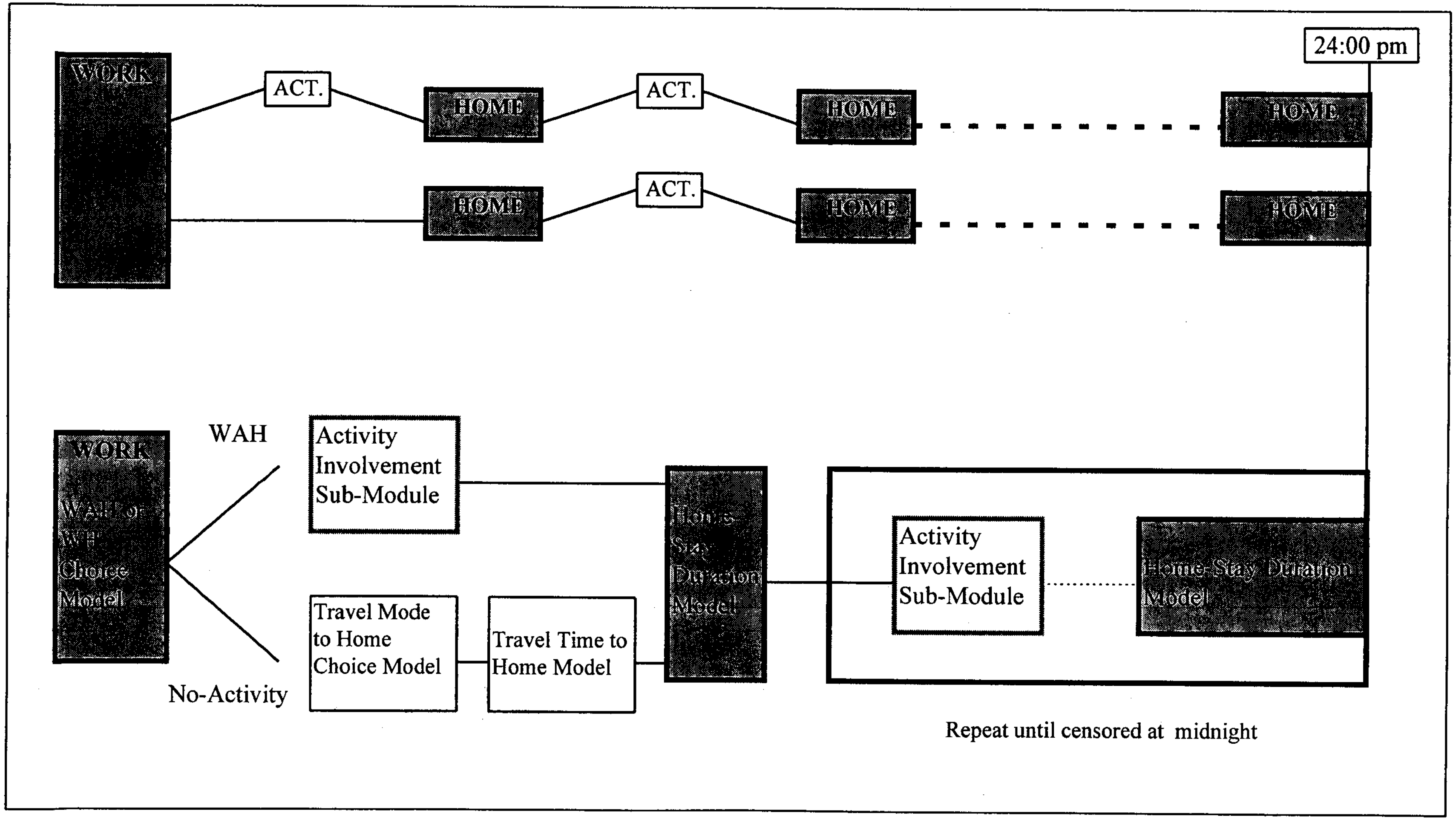

Fig. 2.3 Framework of post-work activity pattern

1. Activity or No-Activity Choice in Worker Model.

The traveler's choice of pursuing an activity or staying home or going directly home/to a work place will be modeled by the standard binary logit model. The linear indirect utility function provided by each alternative is,

$$\mathbf{V}_{kit} = \beta_{oi} + \mathbf{X}_k\beta_{1i} + \mathbf{NW}_{kt}\beta_{2i} + \mathbf{HB}_{kt}\beta_{3i} + \varepsilon_{kit} \tag{2.1}$$

where $\mathbf{V}_{kit}$ is the utility provided to traveler k by alternative i at time t, $\mathbf{X}_k$ is a vector of traveler k's socioeconomic variables, $\mathbf{NW}_{kt}$ is a vector of traffic congestion/network specific characteristics at time t, $\mathbf{HB}_{kt}$ is a vector of habitual behavior that includes previous activities pursued.

2. Nested Activity Type/Frequency of Trip-Chain Stops Choice Model

Modeling an individual's trip activity-type choice is also fundamental to dynamic transportation modeling because the type of activity chosen impacts the duration of activity involvement, destination choice, travel mode choice, and travel time and so on.

Using a standard utility maximizing approach, the mean indirect utility can be defined as in equation (2.1) and the generalized extreme value assumption of ε_{kit} produces a nested logit model structure (McFadden, 1981) with the lower nest being the choice of the number of trip-chain stops (two, three or four or more) and the upper nest being the activity type choice.

In equation form, the model structure is:

$$P_{kjt} = \frac{\exp(V_{kjt} + \phi_{AC}\ln \sum_{c \in S_j} \exp(V_{kct}))}{\sum_{l \in L} \exp(V_{klt} + \phi_{AC}\ln \sum_{d \in S_l} \exp(V_{kdt}))} \tag{2.2}$$

where P_{kjt} is the probability of individual k selecting alternative j (i.e. single-stop shopping, free-time, personal business, visit/appointment, or activity chain), V_{kjt} is the indirect utility for individual k derived from alternative j, ϕ_{AC} is an estimable coefficient defined only for the activity-chain alternative, L is the set of activity-type choices including j, S_l is the set of trip-chain stop choices available for activity-type l, and $\sum_{d \in S_l} \exp(V_{kdt})$ is the denominator (inclusive value) of the number of stops choice model specified from,

$$P_{knt} = \frac{\exp(V_{knt})}{\sum_{d \in S_l} \exp(V_{kdt})} \tag{2.3}$$

where P_{knt} is the probability of individual k, who is pursuing activity type l, choosing number of stops alternative n from the set of alternatives S_l, and V_{knt} is the indirect utility derived by individual k from choice n at time t.

3. Mode Choice Model

In this section, non-traditional mode choice models, which do not include alternative specific variables such as travel time and travel cost, are presented.

Two different sets of choice alternatives are modeled for commuting modes without an activity involvement on the way and the travel mode to/from an activity involvement. Note that in the case of an activity involvement, alternatives of mode choice are defined not for single trips but for whole trips in an activity involvement and thus different modes to and from an activity include two possible cases: physically different modes such as bus and auto, and occupancy differences or driver/rider changes such as SOV and carpool driver/rider. Rarely is the first case observed in practice, so the latter case is defined as mix of SOV and car/vanpool alternatives (change of driver/rider is also rare).

1) Mode to/from Activity Choice Model

Since travel demand is a derived demand, the activity chosen impacts travel characteristics such as travel mode and travel time. However, it is also possible that the choice of travel mode is endogenous with the chosen activity, because the kinds of activities pursued are restricted by the set of available travel modes. Here, it is assumed that even the decision of travel mode is affected by the results of a long-term process, the decision can be contemporaneous with the decision of a current activity type. As such, activity indicator variables are replaced by the expected probabilities of activity type choices to overcome the possible endogeniety problem (see Dubin and MacFadden, 1984; Mannering and Winston, 1985; Hamed and Mannering, 1993).

The assumption here is that a traveler is likely to choose a travel mode providing the maximum utility. Consider a function that defines the linear indirect utility that each traveler, k, drives from the mode choice:

$$\mathbf{V}_{kmt} = \beta_{oi} + \mathbf{X}_k\beta_{1i} + \mathbf{NW}_{kt}\beta_{2i} + \mathbf{HB}_{kt}\beta_{3i} + E[\mathbf{AT}_{kt}]\beta_{4i} + E[DT_{kt}]\,\beta_{5i} + \zeta_{kmt} \qquad (2.4)$$

where $\mathbf{V}_{kmt}$ is the utility provided to traveler k by travel mode m at time t, $\mathbf{AT}_{kt}$ is a vector of indicator variables denoting the activity type for traveler k at time t and this is replaced by its expected values which is the vector of probabilities calculated from the activity type choice equation (2.2), DT_{kt} is the travel distance to/from an activity calculated from minimum-path network and this is also replaced by values instrumented by all exogenous values due to its endogenity. This gives rise to standard multinomial logit model as defined in equation (2.3).

2) Mode to Work and Mode from Work without Activity Choice Model

Commuting travel modes to work and from work, without an activity on the way, are defined as SOV, car/vanpool driver, car/vanpool rider, bus, and park and ride. Unlike activity trips, a considerable proportion of travelers use bus and park and ride modes (10% and 5% respectively) in the Puget Sound region.

The assumption here is the same as before in that travelers are likely to choose the travel mode providing the highest utility. Hence, the linear indirect utility function can be defined as equation (2.4) except for activity type indicator variables:

$$V^{*}_{kmt} = \beta_{oi} + \mathbf{X}_{k}\beta_{1i} + \mathbf{NW}_{kt}\beta_{2i} + \mathbf{HB}_{kt}\beta_{3i} + E[DIST_{kt}]\,\beta_{5i} + \zeta^{*}_{kmt} \qquad (2.5)$$

4. Frequency of Home-Activity-Home Trip in Pre-Work and Frequency of Four or More Trip-Chain Stops Model

One would expect the stops of a trip-chain to be some function of the utility that a individual traveler drives from chaining several trips as well as a random fluctuation in traffic congestion resulting from the occurrence of accidents and other disruptive incidents (Mannering and Hamed, 1990). With the independent randomness assumption of each trip-chain, the probability of traveler i having a n_i stops in trip-chain is,

$$P(n_i) = \frac{e^{-\lambda_i}\lambda_i^{n_i}}{n_i!} \qquad (2.6)$$

where λ_i is the Poisson parameter for traveler i. The Poisson parameter specification is,

$$\log(\lambda_i) = \mathbf{Z}_i\beta\,,\ \ \mathbf{Z}_i = (\mathbf{1},\ \mathbf{X}_i\,, \mathbf{NW}_i\,, \mathbf{HB}_i) \qquad (2.7)$$

where $\mathbf{Z}_i$ is a vector of regressors, in which each column vectors are defined as before. β is a vector of estimable parameters.

Since the model for the frequency of home-activity-home in pre-work is formulated with workers who make at least one home-based trip (HAH), the sample takes values strictly above zero. Thus the sample data is truncated at zero. Also, the model for the frequency of four or more stops in a chain has a truncation point of three.

With truncation at a value C, the distribution of n_i applies only to values above C. Thus,

$$\text{Prob}[n_i \mid n_i > C] = \frac{(e^{-\lambda_i}\lambda_i^{n_i})/n_i!}{\Pr ob[n_i > C]} \qquad \text{for } n_i = C+1, C+2, \ldots \tag{2.8}$$

5. Travel Time Model

2.6.1 Model of Travel Time From Home To Work and From Work To Home Without Activity

Based on the assumption of a traveler's ability controlling over travel time, a continuous linear travel time model (OLS) is defined as,

$$TT_{kt} = \beta_o + \mathbf{X}_{kt}\beta_1 + \mathbf{NW}_{kt}\beta_2 + \mathbf{HB}_{kt}\beta_3 + E[DT_{kt}]\beta_4 + \xi_{kt} \tag{2.9}$$

where TT_{kt} is the travel time from home to work or from work to home without an activity for traveler k at time t, ξ_{kt} is an error term assumed to be normally distributed, and the other variables and coefficients are defined as before.

When the selectivity bias is corrected (i.e. using the conditional value of travel time), Equation 2.9 becomes,

$$\begin{aligned} E[TT_{kt} \mid \text{No-activity}] = \beta_o &+ \mathbf{X}_{kt}\beta_1 + \mathbf{NW}_{kt}\beta_2 + \mathbf{HB}_{kt}\beta_3 + E[DT_{kt}]\beta_4 \\ &+ E[\mathbf{MD}_{kt}]\ \beta_5 + E[\xi_{kt} \mid \text{No-activity}] + \eta_{kt} \end{aligned} \tag{2.10}$$

where $E[TT_{kt} \mid \text{No-activity}]$ is the conditional expectation of travel time given the no-activity alternative, $\mathbf{MD}_{kt}$ is a vector of indicator variables denoting travel modes for traveler k at time t and this is replaced by its expected values which is the vector of probabilities calculated from the mode choice equation (2.5), $E[\xi_{kt} \mid \text{No-activity}]$ is the conditional expectation of the error term given the no-activity choice, and η_{kt} is a normally distributed error term.

As shown by Hay (1980) and Dubin and McFadden (1984), the closed form of the selectivity bias correction term for the no-activity choice is defined as,

$$E[\xi_{kt} \mid \text{No-activity}] = -(\sqrt{6\sigma^2} / \pi)\, \rho_{\text{no-activity}} [(1-P_{kt})\ln(1-P_{kt}) / P_{kt} + \ln P_{kt}] \qquad (2.11)$$

where σ^2 is the variance of ξ_{kt} in the entire population (not conditioned on the no-activity choice), $\rho_{\text{no-activity}}$ is the correlation of ξ_{kt} with the unobserved utility associated with the no-activity choice, and P_{kt} is the probability of traveler k selecting the no-activity choice at time t. Entering this into Equation 2.10 gives,

$$\begin{aligned} E[TT_{kt} \mid \text{No-activity}] = \beta_o + \mathbf{X}_{kt}\beta_1 + \mathbf{NW}_{kt}\beta_2 + \mathbf{HB}_{kt}\beta_3 + E[DT_{kt}]\beta_4 \\ + E[\mathbf{MD}_{kt}]\ \beta_5 + SB_{kt}\,\gamma_{\text{no-activity}} + \eta_{kt} \qquad (2.12) \end{aligned}$$

where SB_{kt} is the selectivity correction term calculated as $[(1-P_{kt})\ln(1-P_{kt}) / P_{kt} + \ln P_{kt}]$ and $\gamma_{\text{no-activity}}$ is the coefficient of the selectivity correction term, which equals $-(\sqrt{6\sigma^2} / \pi)\, \rho_{\text{no-activity}}$.

2) Model of Simultaneous Travel Time To/From Activity and Activity Duration

Intuitively, the travel time to and from the selected activity and the duration of that activity are interrelated; that is, travelers may naturally be willing to accept longer travel times to activities requiring a longer duration and vice versa. In addition to the assumption of the interrelationship between travel time and activity duration, asymmetrical travel times to and from an activity are assumed. As a result, the simultaneous equations structure can be formulated as,

$$TT1_{kt} = \alpha_o + \alpha_1 TT2_{kt} + \alpha_2 ADUR_{kt} + \mathbf{Z}_{kt}\alpha_3 + \delta_{1kt} \tag{2.13}$$

$$TT2_{kt} = \beta_o + \beta_1 TT1_{kt} + \beta_2 ADUR_{kt} + \mathbf{Z}_{kt}\beta_3 + \delta_{2kt} \tag{2.14}$$

$$ADUR_{kt} = \gamma_o + \gamma_1 TT1_{kt} + \gamma_2 TT2_{kt} + \mathbf{Z}_{kt}\gamma_3 + \delta_{3kt} \tag{2.15}$$

$$\mathbf{Z}_{kt} = (\mathbf{X}_{kt}\ \mathbf{NW}_{kt}\ \mathbf{HB}_{kt}\ \mathbf{AT}_{kt}\ \mathbf{MD}_{kt})$$

where $TT1_{kt}$ and $TT2_{kt}$ are travel times to and from an activity by traveler k in minutes respectively, $ADUR_{kt}$ is activity duration in minutes and $\mathbf{Z}_{kt}$ is a vector of subvectors defined as before, and δ_{1kt}, δ_{2kt} and δ_{3kt} are error components allowing contemporaneous correlations .

As was the case with the travel time model from home to work or work to home, selectivity bias is present. It results from the facts that the travel time to and from the activity and the duration of the activity are only observed for the activity type and the travel mode that have been selected by the traveler. Thus, $\mathbf{AT}_{kt}$ and $\mathbf{MD}_{kt}$ are replaced by their expected probabilities (i.e. calculated from the activity type choice and the mode choice models respectively).

6. Home-Stay Duration Model

Home-stay duration is defined as the time spent at home between out-of-home trip generating activities. The duration of home-stay is extremely important in determining the timing of discretionary activities (e.g. shopping, free-time, etc.) which effectively end a home-stay. In this respect, n lengths of home-stays define the (n-1) frequencies of activity involvement.

In this study, the time-observation unit is one day (24 hrs), beginning at 12:01 am and ending at 12:00 midnight. Thus, all travelers who are home at midnight, will have left- or right-censored home-stay duration, and this censoring must be accounted for in any chosen econometric modeling procedure.

In terms of modeling home-stay duration, the use of hazard functions approach has been shown to be appropriate (Hamed and Mannering, 1993). This approach uses the hazard rate, (defined in this context as the rate at which home-stay durations are ending at some time t, given that the traveler has been at home until time t), as a basis for modeling home-stay duration. Mathematically, the hazard function is defined as,

$$h(t) = \frac{f(t)}{[1 - F(t)]} \tag{2.16}$$

where, t is a realization of continuous non-negative random variable T representing home-stay duration, and f(t) and F(t) are probability density and distribution functions of home-stay duration, respectively. Then the survival function S(t) is defined as,

$$S(t) = \text{Prob}\,[T>t] = 1 - F(t) \tag{2.17}$$

Using the hazard function as a basis, the proportional hazard formulation is,

$$h(t|Z) = h_o(t)\,\exp(\mathbf{Z}\beta) \tag{2.18}$$

where h(t|Z) is the hazard conditioned on covariate vector **Z**.

When home-stay duration T is a Weibull-distributed survival random variable with parameters $\lambda>0$ and $P>0$, a corresponding hazard function,

$$h_o(t) = \lambda P\,(\lambda t)^{P-1} \tag{2.19}$$

gives the proportional hazards model (from Equation 2.30),

$$h(t|Z) = \lambda P\,(\lambda t)^{P-1}\,\exp(-\beta Z) \tag{2.20}$$

The other types of state dependence, "occurrence dependence" and "lagged duration dependence" must be handled explicitly in the model by including appropriate variables in the covariate vector Z. Complications resulting from this inclusion will be discussed in a later section.

7. Departure Time Model

A non-worker is assumed to repeat home-stays and activity involvements until being censored at midnight and so the departure time of a non-worker is defined as the length of the first home-stay from the previous midnight. This length was modeled by OLS as such,

$$DT_k = \beta_o + \mathbf{X}_k\beta_1 + \mathbf{NW}_k\beta_2 + \mathbf{HB}_k\beta_3 + \zeta_k \qquad (2.21)$$

where DT_k is the departure time of the first activity involvement for non-worker k, $\mathbf{HB}_k$ includes lagged dependent values (i.e. previous-day departure time) and previous-day activity/travel habits, ζ_k is an error term assumed to be normally distributed, and the other variables and coefficients are defined as before.

The definition of pre-work departure time is the same as nonwork, but it is noted that applying the non-workers' model to the pre-work departure time may produce a operational problem, in which the scheduled work-arrival time calculated from the estimated model system can pass far over the scheduled work-start time. It is a practical concern. Hence, the pre-work departure time model will employ the Mannering, Abu-Eisheh, and Anadottir (1990)' approach which calculates the departure time back-ward from the work-start time. The model is,

$$DT_k = WST_k - SD_k - \sum TT_k - \sum ADUR_k - \sum HDUR_k \qquad (2.22)$$

where DT_k is the pre-work departure time for worker k, WST_k is the work start time for workers with fixed work-schedules and is the preferred arrival time for workers without fixed work-schedules, SD_k is the schedule delay, defined as the amount of time between scheduled work start time and actual arrival time, $\sum TT_k$ is the summation of travel times before work, $\sum ADUR_k$ is, if any, the summation of activity times undertaken before work, and $\sum HDUR_k$ is, if it is a home-based trip, the summation of home-stays before work.

III. STATE DEPENDENCE AND HETEROGENEITY

There is important information provided in panel data with regard to previous activity involvement and duration, that is potentially a good predictor of current behavior. In considering state dependence in duration modeling, three types of state dependence arise (see Heckman and Borjas (1980) for general discussions of state dependence types). The first type of state dependence is "duration dependence". This type of dependence focuses on the conditional probability of a duration ending given that it has lasted some known time. The second type, termed the "occurrence dependence", captures the effect that the number of previous involvements in certain behavior has on current behavior. The third type of state dependence is "lagged duration dependence", and accounts for the possibility that the duration of previous activities is a good indicator of the duration or occurrence of the current activity. This type of state dependence could uncover important habitual behavior. In the case of a logit model, only occurrence dependence and lagged duration dependence can be applied because the dependent variable is not a duration but a discrete choice.

Heterogeneity in activity behavior models is an outgrowth of the differences that remain among travelers' nuisance distributions after the effects of observable characteristics (e.g. socioeconomic characteristics, etc.) have been accounted for.

Heterogeneity produces the possibility that two travelers with identical observable characteristics may still have different activity responses. A natural solution to possible heterogeneity problems is to estimate separate models for each traveler, but this implies an unrealistic data burden. Thus some formal econometric correction is required.

With these general definitions of state dependence and heterogeneity, an appropriate econometric modeling structure can be formalized. A common correction for heterogeneity in duration models allows some parameters to be the same for all travelers but lets a single parameter vary across households to account for heterogeneity. If we define a random variable v to account for possible population heterogeneity, we can write a base-line survivor function (corresponding to $h_o(t)$),

$$S_o(t|v) = \exp[-v(\lambda t)^P] \tag{3.1}$$

It is assumed that v is Gamma distributed with parameters k>0 and c>0. To operationalize the Gamma heterogeneity component of the model, it is assumed that the model has a constant term and hence no generality is lost by constraining the mean of v to be 1. Therefore, the expected value of v is one (i.e. E[v]=c/k=1) implying k=c. Then it follows that,

$$h_o(t) = [S_o(t)]^\theta \ \lambda P(\lambda t)^{P-1} \tag{3.2}$$

Where $\lambda P(\lambda t)^{P-1}$ is the Weibull hazard. Since the variance of v is 1/k with k=c, θ=0 corresponds to the Weibull model with a homogeneous survival distribution. The further θ is from zero, the greater the effect of heterogeneity in the model.

The corresponding hazard, with covariates, is,

$$h(t|Z) = [S(t|Z)]^\theta \lambda P \ (\lambda t)^{P-1} \exp(-\beta Z) \tag{3.3}$$

Accounting for possible heterogeneity by assuming a Gamma distribution raises a number of important questions. First and foremost is whether or not the model (and the estimation process) will be able to distinguish heterogeneity from state dependence. Second, the Gamma heterogeneity model used herein handles pure heterogeneity, but does not explicitly account for possible state-dependent heterogeneity because it can not be as easily distinguished from the true state-dependent effects. To determine whether or not true state dependence is being captured, econometric tests must be conducted. The most common test is to instrument the state variable(s) (e.g. regressing it against variables known to be exogenous) and estimate the duration model with instrumented values. The significance of the coefficients associated with these instrumented values is an indication of the existence of true state dependence

IV. ESTIMATION RESULTS OF ALL-DAY BEHAVIORAL MODEL

The Puget Sound Transportation Panel (PSTP) was the source of the activity data used in this study. The PSTP is based on two-day travel diaries administered to all members of the sampled households, and includes the four counties of the Seattle-Tacoma metropolitan area. The panel has two waves; one wave administered in the fall of 1989 and consisted of 1,713 households; the second wave administered in the fall of 1990.

1. State dependence and Heterogeneity with home-stay duration model

As discussed earlier, the significance of occurrence dependence, which implies lagged duration dependence (because the higher number of home-stays means shorter average home-stay duration), in home-stay duration models may be not a result of true state dependence, but instead an outgrowth of persistent unobservables (see Heckman and Borjas (1980) for a detailed discussion). To test for this, all state variables are average home-stay duration), in home-stay duration models may be not a result of true state dependence, but instead an outgrowth of persistent unobservables (see Heckman and Borjas (1980) for a detailed discussion). To test for this, all state variables are instrumented by regressing actual values against exogenous variables and by using the regression-predicted values in the model estimation.

The results of this test are shown in Table 4.1 for non-workers and Table 4.2 for workers. From these two tables, it is shown that, except for the wave 1 non-worker model, in which their instrumented state effect is marginally significant along with slight shifts of coefficient magnitude, all other state effects are statistically insignificant, indicating that the state coefficients are not capturing true state dependence. This implies that true state effects may not exist or at least may be correlated with error components in home-stay duration models. This is because, although the Gamma heterogeneity specified in the models captures persistent unobservables, there may be unobservables which are not captured and correlated with state variables. That is, these state variables may be actually capturing a state dependent heterogeneity that is not accounted for by the Gamma heterogeneity variable θ.

2. Temporal Stability

Given the model structures proposed in the preceding chapter, the stability of coefficients over the two time periods (separated by one year) can be tested using a likelihood ratio (LR) test for models estimated by maximum likelihood method, and using a Chow test for OLS models.

First, for the maximum likelihood models, a chi-squared statistic that measures the probability of coefficient stability over time, where the null hypothesis is that two sets of coefficients are equal is,

$$\chi^2 = -2\,(\mathcal{L}_{POOL} - \mathcal{L}_{W1} - \mathcal{L}_{W2}) \tag{4.1}$$

Table 4.1. Test of state dependence in home-stay duration model by instrumented variables (non-worker), t-statistics in parentheses

Variable*	Wave 1 estimated coefficients	Wave2 estimated coefficients
Constant	12.412 (10.660)	11.221 (15.219)
Old age dummy (1 if aged > 57)	0.505 (2.570)	-
Number of children 0-17 years	-0.065 (-0.798)	-
Household size	-	-0.284 (-4.985)
Log of annual household income in $1K	-0.528 (-3.038)	-0.339 (-2.109)
Friday dummy (1 if Friday, 0 otherwise)	-0.404 (-2.394)	-
PM peak arrival dummy (1 if arrived at home 5-6 PM, 0 otherwise)	-1.054 (-4.033)	-
Time budget in hours	-0.243 (-7.875)	-0.273 (-9.267)
"Shopping" activity before getting home dummy (1 if shopping pursued before getting home, 0 otherwise)	-0.283 (-1.305)	-
"Personal" activity before getting home dummy (1 if personal pursued before getting home, 0 otherwise)	-0.426 (-2.105)	-
"Free" activity before getting home dummy (1 if free pursued before getting home, 0 otherwise)	-	0.429 (1.392)
"Chain" activity before getting home dummy (1 if chain pursued before getting home, 0 otherwise)	-	0.256 (1.612)
"Car/Vanpool ride" arrival mode dummy (1 if arrived at home by Car/Vanpool ride mode, 0 otherwise)	0.336 (1.485)	-
"Bus" arrival mode dummy (1 if arrived at home by Bus mode, 0 otherwise)	0.948 (1.074)	2.222 (2.576)
P (Weibull parameter)	0.836 (14.273)	0.831 (12.286)
θ (heterogeneity)	0.883 (3.041)	1.056 (3.147)
Instrumented previous day total number of home-stays	-0.457 (-1.483)	0.0008 (0.422)
Log-likelihood at zero	-1811.52	-1746.12
Log-likelihood at convergence	-1332.53	-1321.46
Number of observations	1159	1091
$\bar{\rho}^2$	0.261	0.240

* Dependent variable is log of home-stay duration in minutes.

Table 4.2. Test of state dependence in home-stay duration model by instrumented variables (worker), t-statistics in parentheses

Variable*	Wave 1 estimated coefficients	Wave 2 estimated coefficients
Constant	9.356 (8.016)	9.462 (21.749)
Young age dummy (1 if aged < 26)	0.001 (1.660)	-
Number of children 0-17 years	-0.178 (-2.294)	-0.180 (-2.683)
Friday dummy (1 if Friday, 0 otherwise)	-0.275 (-1.702)	-0.674 (-3.893)
PM peak arrival dummy (1 if arrived at home 4-6 PM, 0 otherwise)	-0.796 (-5.317)	-0.732 (-4.468)
Time budget after work in hours	-0.394(-8.388)	-0.368 (-8.520)
"Free" activity before getting home dummy (1 if free pursued before getting home, 0 otherwise)	0.520 (1.642)	-
"Car/Vanpool ride" arrival mode dummy (1 if arrived at home by Car/Vanpool ride mode, 0 otherwise)	0.417 (1.713)	-
"Bus" arrival mode dummy (1 if getting home by Bus mode, 0 otherwise)	-	1.232 (3.871)
Home-stay in pre-work dummy (1 if getting home after activity in pre-work, 0 otherwise)	-5.796 (-7.680)	-6.610 (-9.464)
Home-stay after work without activity dummy (1 if getting home after work without activity on the way, 0 otherwise)	-0.150 (-0.941)	-0.341 (-2.163)
P (Weibull parameter)	1.135 (12.770)	1.003 (12.717)
θ (heterogeneity)	4.754 (6.647)	4.411 (5.821)
Instrumented previous day total number of home-stays	-0.023 (-0.049)	0.0006 (0.506)
Log-likelihood at zero	-3514.36	-3507.63
Log-likelihood at convergence	-2281.00	-2242.45
Number of observations	2321	2242
$\bar{\rho}^2$	0.349	0.359

* Dependent variable is log of home-stay duration in minutes.

with the degrees of freedom = k (number of coefficients), where subscripting POOL denotes all observations, W1 denotes time period 1 observations, and W2 denotes time period 2 observations.

Second, for OLS models, the Chow test proceeds first by running a pooled regression and two individual regressions to obtain residual sum of squares (RSS) for each regression. With these results, then apply F-test as follows:

$$\mathrm{F} = \frac{(RSS_{POOL} - RSS_{W1} - RSS_{W2}) / k}{(RSS_{W1} + RSS_{W2}) / (N_{W1} + N_{W2} - k)} \tag{4.2}$$

with df = (k, $N_{w1}+N_{w2}-2k$), where N_{w1} denotes number of wave 1 observations, N_{w2} denotes number of wave 2 observations, and the other subscripts are defined as before. If the computed F exceeds the critical F, we reject the hypothesis that two regressions are the same (i.e. two models are temporally stable).

The results of LR-test and Chow-test are shown in Tables 4.3 and in Table 4.4 respectively. It can be concluded that, although the evidence of temporal stability /instability is not clear, activity-based models may be inherently unstable or, perhaps more likely, more extensive diary data is needed (e.g. seven or fourteen day diaries) to fully capture the cycle of human activity behavior. Clearly additional research on such temporal stability is warranted.

Other potential reasons for the severe instability in logit models are the unseparated habitual behavior (i.e. state effects) from unexplained heterogeneity in the population and the inclusion of unmatched households in both waves due to 19% attrition replacement in wave 2 (see Murakami and Watterson, 1990,1991). On the other hand, the practical application may not be impacted by the partial instability of the disaggregate behavioral models, since total outputs of models in different waves can be stable due to tradeoffs between positive and negative habitual changes (e.g. when applying an estimated logit model to different waves, choice probabilities may be similar in spite of different model specifications). This test is left for further research.

Table 4.3 Likelihood ratio tests for temporal stability of maximum likelihood estimates

Model	Wave 1 specification			Wave 2 specification		
	χ^2	d.f	p-value	χ^2	d.f	p-value
HAH/no-activity choice(pre-work)	5.54	9	0.785	4.72	7	0.694
HAW/no-activity choice(pre-work)	25.18	10	0.005*	25.06	10	0.005*
WAH/no-activity choice(post-work)	15.80	14	0.326	12.00	12	0.446
Frequency of activity-stops choice (non-work)	18.24	17	0.374	19.98	12	0.067
Frequency of activity-stops choice (pre-work)	23.37	15	0.077	20.90	12	0.052
Frequency of activity-stops choice (post-work)	38.18	14	0.000*	37.68	17	0.000*
Activity type choice (non-work)	16.80	20	0.666	34.20	22	0.047*
Activity type choice (pre-work)	44.64	20	0.001*	33.24	21	0.060
Activity type choice (post-work)	44.20	20	0.001*	43.40	22	0.004*
Mode to/from activity choice (non-work)	41.80	26	0.026*	9.80	27	0.999
Mode to/from activity choice (pre-work)	43.16	26	0.019*	34.76	25	0.093
Mode to/from activity choice (post-work)	38.00	30	0.150	67.20	39	0.003*
Mode to work without activity choice (pre-work)	23.60	25	0.543	17.60	23	0.779
Mode to work without activity choice (post-work)	42.40	30	0.066	22.32	27	0.721
Left-truncated Poisson for # HAH (pre-work)	14.24	7	0.047*	7.17	6	0.305
Left-truncated Poisson for 4+ stops in trip-chain (non-work)	10.43	9	0.317	11.20	7	0.130
Left-truncated Poisson for 4+ stops in trip-chain (work)	7.61	9	0.574	6.82	7	0.448
Home-stay duration (non-work)	24.50	14	0.009*	5.30	10	0.870
Home-stay duration (work)	15.43	13	0.281	10.26	11	0.507

* The null hypothesis that two models are temporally stable in coefficients, can be rejected at 5% significance level.

Table 4.4 Chow tests for temporal stability of OLS estimates

Model	Wave 1 specification			Wave 2 specification		
	F	d.f	$F_{0.05}$*	F	d.f	$F_{0.05}$*
Departure time of non-worker	1.301	(7,1426)	2.01	1.128	(8,1425)	1.94
Travel time from home to work without activity (pre-work)	0.705	(14,2937)	1.71	1.360	(15,2935)	1.67
Travel time from work to home without activity (post-work)	1.570	(12,2405)	1.75	1.169	(14,2401)	1.71

$F_{0.05}$* is the critical F-value at 5% significance level.

Based on the theoretical models of all-day travel behavior developed in preceding chapters, an operational system of models employing a simulation approach is developed mainly for an explanatory demonstration of activity/travel behavior process and for further policy evaluation relating to the travel demand. The system is divided into three sub-systems: non-work, pre-work, and post-work.

None of these works cover the state dependence (dependence on past experiences) along with empirical evidence at the system level. Hence, this study seeks to go one step further in the implementation of activity-based travel demand models by employing a simulation approach and by integrating a range of methodological issues advanced in the preceding chapters. In this study, only the conceptual framework for the application of the model system is presented and so empirical verifications (e.g. simulation) are left for further research.

Figures 5.1-5.3 illustrate operational structures generating the travel demand of non-workers and workers (in pre-work and post-work) by using the previously specified and estimated models.

As a practical concern, using the results of estimated models, each module computes the amount of time that each traveler has been on the road for traveling and then aggregates them to obtain total travelers on the road by time, which can be also classified by other characteristics such as travel mode, activity type, etc. This would contribute to the ability to develop and analyze demand management policies such as the flexible work hour effect.

VI. Conclusions and Future Research Directions

1. Conclusions

This study formulated a modeling system that accounted for all-day activity involvement and travel behavior, which are divided into three main categories of non-work, pre-work, and post-work. The system includes a set of discrete/continuous models, hazard-based duration models, and left-truncated poisson models. The role that state

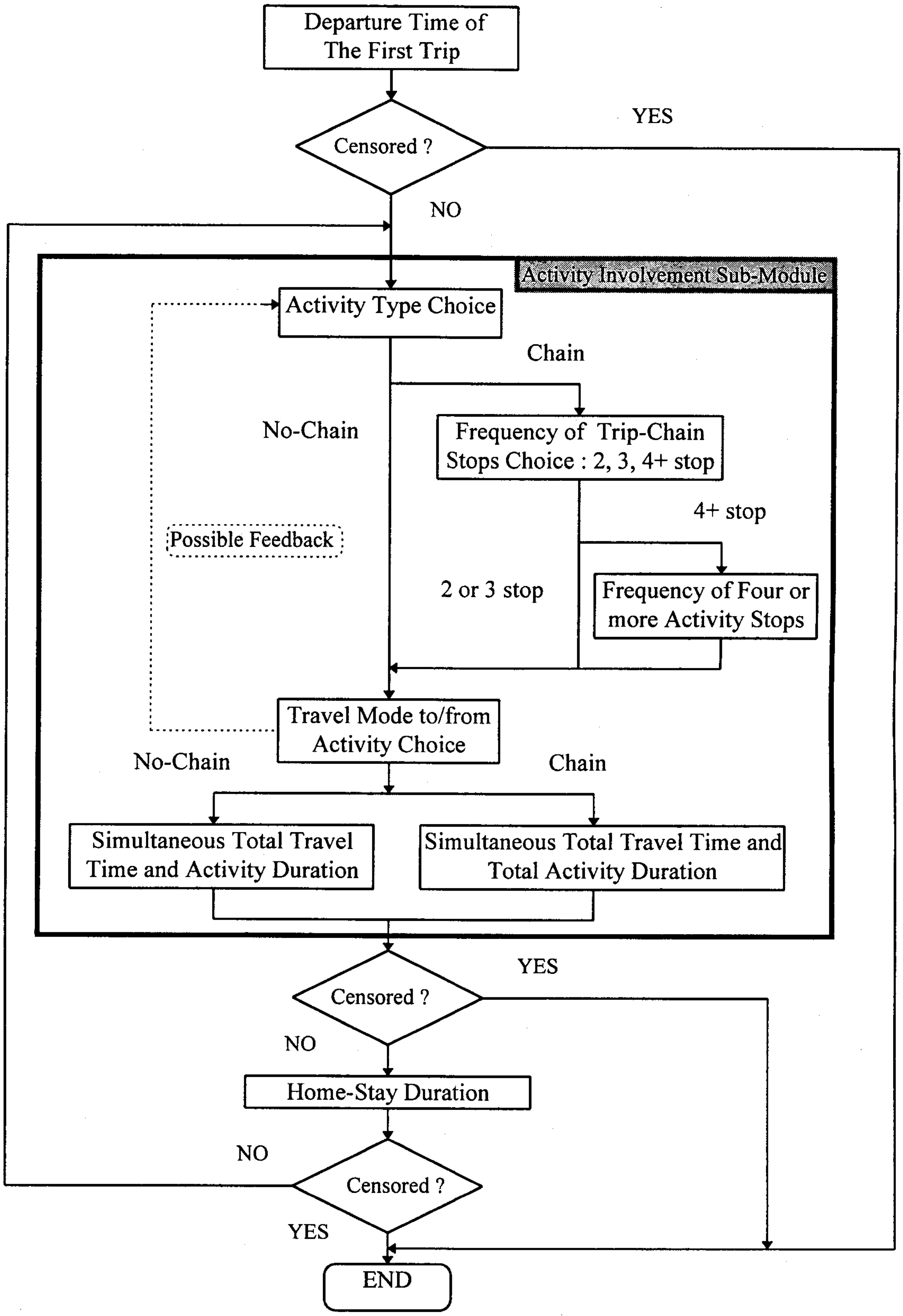

Figure 5.1. Flow chart of non-work model system

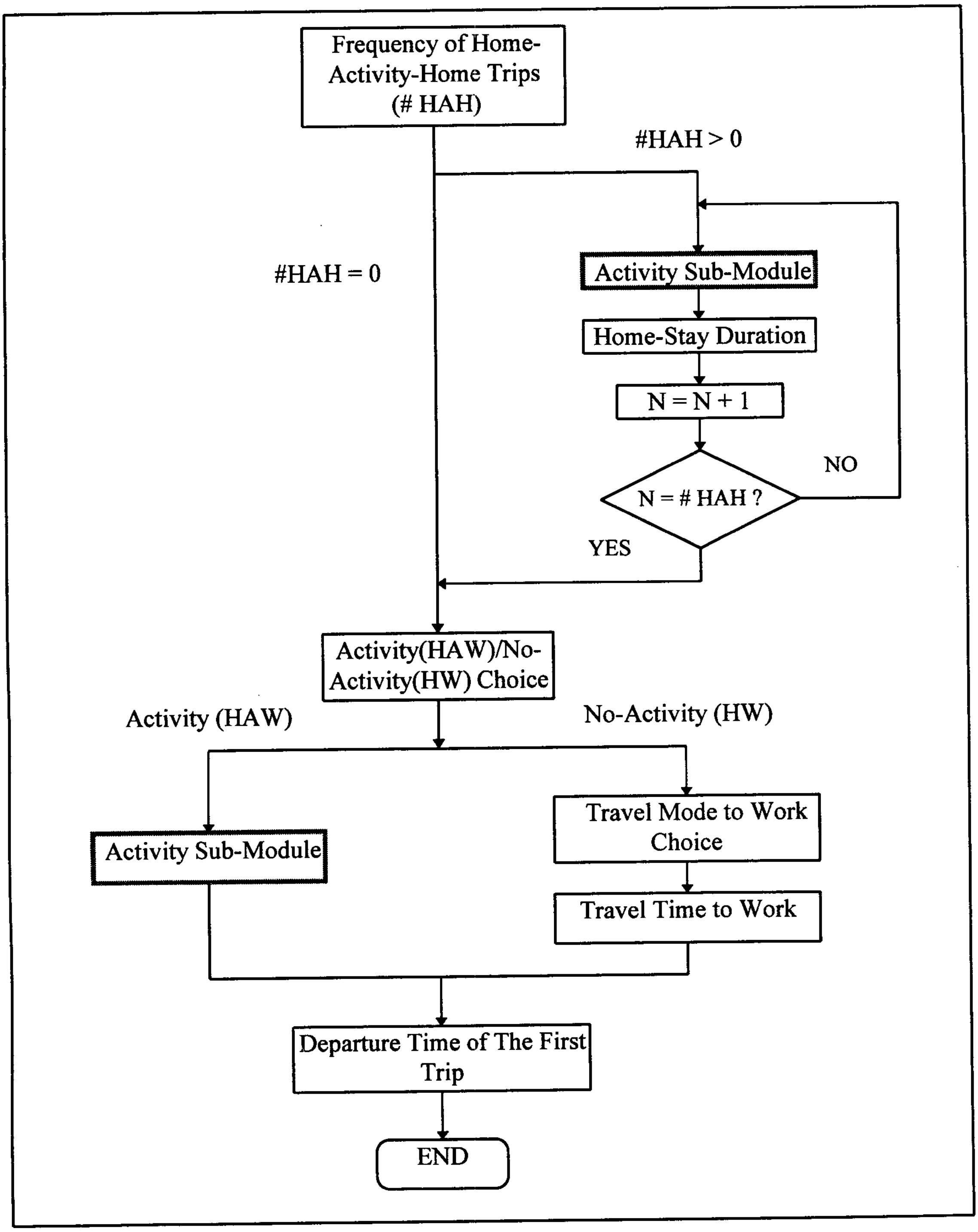

Figure 5.2. Flow chart of pre-work model system

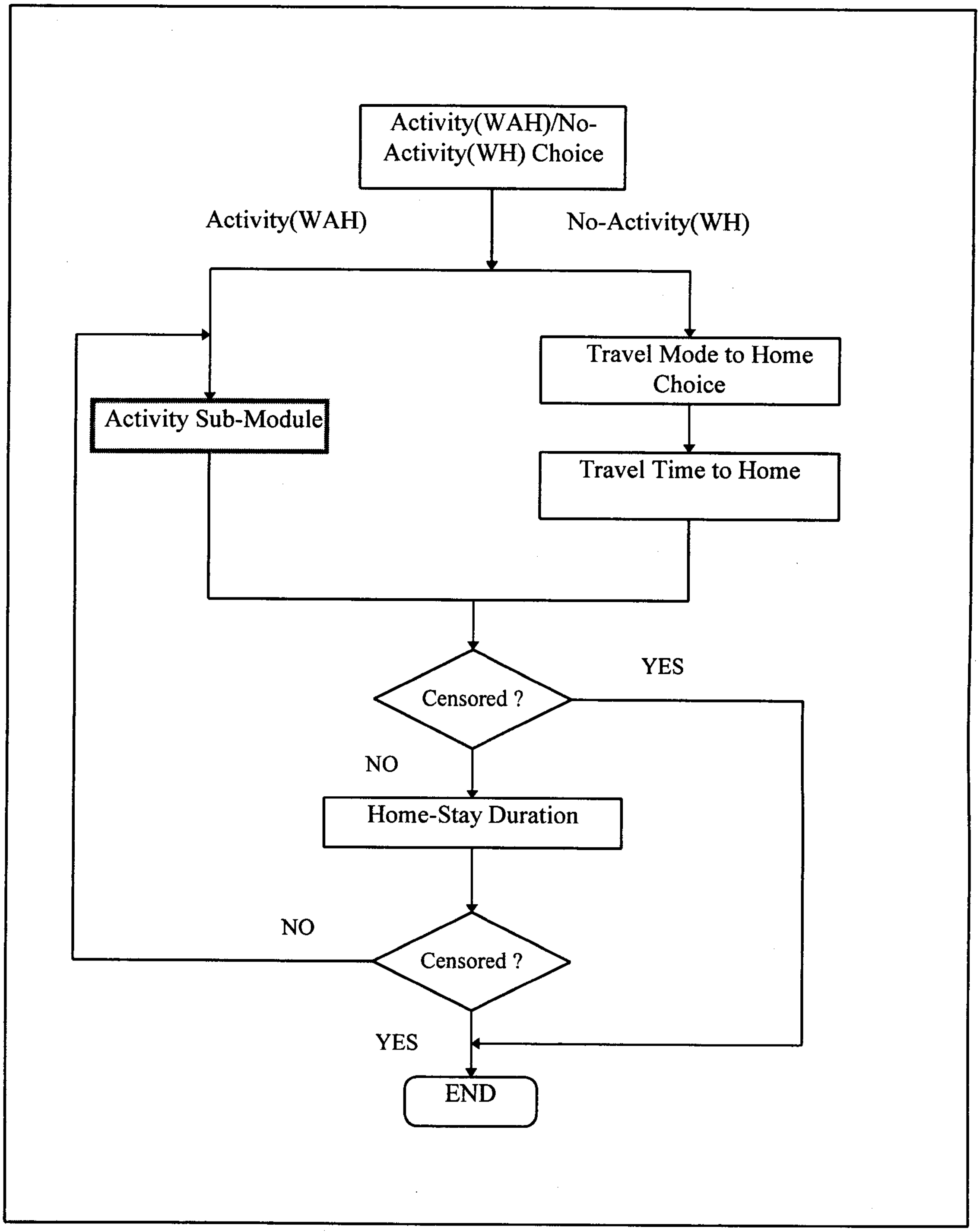

Figure 5.3. Flow chart of post-work model system

dependence and heterogeneity play in models of activity/travel behavior was explored in the context of home-stay duration models. Using data from two-day travel diaries, collected in 1989 and again in 1990, separate econometric models were estimated for 1989 and 1990, and the temporal stability of these models was statistically evaluated. Finally, three conceptual operational modeling systems were developed for possible policy evaluation and further travel demand forecasting.

- The empirical results show the importance of proper econometric specification in estimating activity-based travel behavior models.

- Habitual behavior was confirmed with highly significant and temporally stable coefficient estimates in most of the models. This underscores the need for multi-day travel data. In addition, current-day state effects (i.e. current-day experience negatively impacts on the same decision) showed consistently stable results.

- The Poisson model along with the nested activity-type/frequency of trip-chain stops choice model has an important implication in that these models can be a cornerstone for overcoming the current difficulty in identifying trip-chain patterns. This notion was supported by the fact that lagged variables were identified and found to be consistently significant.

- The model of travel times without activities, showed that the selectivity-bias correction term was significantly different from zero suggesting that the selectivity-bias of the activity/no-activity choice is clearly present in these models. The positive sign of selectivity-bias correction term means a negative correlation of the error term in travel time equation with the utility associated with no-activity choice.

- In the simultaneous travel time and activity duration models, consistent relationships between dependent and right-hand side dependent variables were found. In the non-work single-stop model, only the positive one-way impacts of activity duration on travel time were identified. The pre-work single-stop model also had activity duration positively impacting travel time to/from activity. Conversely, the travel time from the activity decreased the activity duration. The relationship in the post-work single-stop model is the same as the pre-work single-stop model, except for the positive impact of expected travel time from the activity on activity duration. This is because post-work activities are less constrained by scheduling. In the case of trip-chaining, the sum of travel times and the sum of activity durations are positively inter-correlated.

- In terms of the role that state dependence and heterogeneity play in models of activity behavior, true state effects may not exist or at least may be correlated with error components in home-stay duration models. This is because, although the Gamma heterogeneity specified in these models captures some portion of persistent unobservables, there may remain unobservables which are not captured and correlated with state variables.

- For temporal stability, likelihood ratio tests were employed for maximum likelihood estimates and Chow tests for OLS models. At the 5% level of confidence, the results show that the multinomial logit models for activity-type choice and for activity involvement mode choice were unstable. In contrast, all of OLS models for departure time and commute travel time without activity were stable over time. The other models were inconclusive. Generally, socio-economic variables showed the most unstable results, traffic congestion/network variables moderately stable results, and state dependent variables the most stable results.

- The proposed three conceptual frameworks, to operationalize the model system, demonstrate the decision process of activity/travel behavior classified by non-work, pre-work and post-work respectively. Using the previous estimation results, each framework can compute travel-related discrete/continuous decisions and thus the amount of time that each traveler is on the transportation network.

2. Future Research Directions

The empirical findings and conceptual framework of the activity-based travel model system presented in this study should provide a valuable methodological starting point for future activity-based travel-modeling research. In this regard, there are a number of important directions to follow.

First, the model system developed in this study includes only the time domain, which may contribute to the overall instability of the estimation results due to omission of geographic characteristics. With this said, destination choice models, which produce a dynamic origin/destination table, and route choice models, are warranted. In this case, zonal attributes can be included in the models.

Second, although the state effects in the home-stay duration model looked spurious, the use of longer travel diaries and longer panel data could reveal important state effects. Also, the state effects of the other should be tested with appropriate testing methods which can account for heterogeneity.

Third, the findings of temporal instability do not suggest that activity-based travel models are inherently unstable. They do, however, underscore the need for more extensive data both in terms of the length of the travel observation period (i.e. seven to fourteen day diaries) and the number of panels waves.

Fourth, another drawback in terms of application of the model system is in the area of "while-work" activities which is assumed to be constant. To complete an all-day behavior modeling system, while-work activities should be accounted for.

Fifth, in terms of practical application, the test of whether the partial temporal instability of the disaggregate behavioral models produce the same system output or not, is instructive. This could be important because total outputs of models in different waves can be stable due to tradeoffs between positive and negative habitual changes.

Finally, empirical testing of the three proposed conceptual model systems and integration of the model system into a complete travel demand forecasting model is needed.

BIBLIOGRAPHY

Abu-Eisheh, S., Mannering, F. (1987) A discrete/continuous analysis of commuters' route and departure time choice. *Transportation Research Record* **1138**, 27-34 .

Adiv, A. (1983) The structure of work-trip based on analysis of trip diaries in the San Francisco Bay Area. In: S. Carpenter and P. Jones (eds.) *Recent Advances in Travel Demand Analysis*. Gower, Adershot, England.

Arnot, R., Palma, A. and Lindsey, R. (1988) Schedule delay and departure time decisions with heterogeneous commuters. *Transportation Research Record* **1197**, 56-67.

Arnot, R., Palma, A. and Lindsey, R. (1990) Departure time and route choice for the morning commute. *Transportation Research* **24B**(3), 209-228.

Atherton, T. and Ben-Akiva, M. (1976) Transferability and updating of disaggregate travel demand models. *Transportation Research Record* **610**, 12-18.

Ben Akiva, M. and Lerman, S. (1985) *Discrete Choice Analysis*. MIT Press, Cambridge, MA.

Bhat, C. and Koppelman, F. (1993) A conceptual framework of individual activity program generation. *Transportation Research* **27A**(6), 443-446.

Damm, D. (1983) Theory and Empirical Results: A Comparison of Recent Activity-Based Research. In Carpenter and Jones (Eds.) *Recent Advances in Travel Demand Analysis*, Gower, Aldershot, England, 3-33.

Elbers, C. and Ridder, G. (1982) True and spurious duration dependence: The Identifiability of the Proportional Hazard Model. *Review of Economic Studies*, **44**, 403-409.

Fleming, T. and Harrington, D. (1990) *Counting processes and survival analysis*. John Wiley and Sons, New York, NY.

Gupta, S. (1991) Stochastic models of interpurchase time with time-dependent covariates. *Journal of Marketing Research*, 18, 1-15.

Hamed, M. (1990) Discrete/continuous econometric analysis of commuters' post-work activity involvement and duration. Ph.D. dissertation. Department of Civil Engineering, University of Washington, Seattle, WA.

Hamed, M. and Mannering, F. (1993) Modeling travelers' post-work activity involvement: toward a new methodology. *Transportation Science* **27(4),** 381-394.

Hausman, J., Hall, B. and Griliches, Z. (1984) Econometric models for count data with an application to the patents-R&D relationship. *Econometrica* **52,** 909-938.

Hay, J. (1980) Occupational choice and occupational earnings: selectivity bias in a simultaneous logit-OLS model. Unpublished PhD dissertation, Yale University, New Haven, CN.

Heckman, J. and Borjas, G. (1980) Does unemployment cause future unemployment? definitions, questions and answers from a continuous time model of heterogeneity and state dependence. *Econometrica*, **47**, 247-283.

Heckman, J. and Singer, B. (1984) A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, **52**(2), 271-320.

Hendrickson, C. and Plank, E. (1984) The flexibility of departure times for work trips. *Transportation Research* **18A**(1), 25-36.

Hensher, D. (1981) A practical concern about the relevance of alternative-constants for new alternatives in simple logit models. *Transportation Research* **15B**(6), 407-410.

Kalbfleisch, J. and Prentice, R. (1980) *The statistical analysis of failure time data*, John Wiley & Sons, New York, New York.

Kawakami, S. and Isobe, T. (1990) Development of a one-day travel-activity scheduling model for workers. In: P. Jones, eds., *Developments in Dynamic and Activity-Based Approaches to Travel Analysis*. Gower, Adershot, England.

Kiefer, N. (1988) Economic duration data and hazard functions. *Journal of Economic Literature*, **26**, 646-679.

Kim, S.-G., Hamed, M. and Mannering, F. (1993) A note on modeling travelers' home-stay duration and efficiency of proportional hazards model. Working paper, Department of Civil Engineering, University of Washington.

Kim, S.-G. and Mannering, F. (1994) Panel data and activity duration models: econometric alternatives and applications. Forthcoming in T. Golob (eds.) *Panels for Transportation Planning: Methods and Applications*. Kluwer Publishing, CA.

Kitamura, R. and Kermanshah, M. (1983) Identifying time and history dependencies of activity choice. *Transportation Research Record* **944**, 22-30.

Kitamura, R. (1984) A model of daily time allocation to discretionary out-of-home activities and trips. *Transportation Research* **18B**, 255-266.

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

Kitamura, R. (1990) Panel analysis in transportation planning: an overview. *Transportation Research* **24A**, 401-415.

Kitamura, R. and Bunch, D. (1990) Heterogeneity and state dependence in household car ownership: a panel analysis using ordered-response probit models with error components, *Transportation and Traffic Theory*, Yokohama, Japan.

Lancaster, T. (1985) Generalized residuals and heterogeneous duration models with application to the Weibull model. *Journal of Econometrics*, **28**, 155-169.

Mahmassani, H. and Herman, R. (1984) Dynamic user equilibrium departure time and route choice on idealized traffic arterials. *Transportation Science* **18**(4), 362-384.

Mahmassani, H. and Chang, G.-L. (1986) Experiments with departure time choice dynamics of urban commuters. *Transportation Research* **20B**(4), 297-320.

Mannering, F. and Winston, C. (1985) A dynamic empirical analysis of household vehicle ownership and utilization. *Rand Journal of Economics* **16**(2), 215-236.

Mannering, F. and Hensher, D. (1987) Discrete/continuous econometric models and their application to transport analysis. *Transport Reviews* **7**(3), 227-244.

Mannering, F. (1989) Poisson analysis of commuter flexibility in changing routes and departure times. *Transportation Research* **23B**(1), 53-60.

Mannering, F., Abu-Eisheh, S. and Arnadottir, (1990) A dynamic traffic equilibrium with discrete/continuous econometric models. *Transportation Science* **24**(2), 105-116 .

Mannering, F. and Hamed, M. (1990) Occurrence, frequency, and duration of commuters' work-to-home departure delay. *Transportation Research* **24B**(2), 99-109.

Mannering, F. and Winston, C. (1991) Brand loyalty and the decline of American automobile firms. *Brookings Papers on Economic Activity: Microeconomics*, 67-114.

Mannering, F., Kim, S.-G., Barfield, W. and Ng, L. (1994) Statistical analysis of commuters' route, mode and departure time. *Transportation Research* **2C**(1), 35-47.

Mannering, F., Murakami, E. and Kim, S.-G. (1994) Temporal stability of travelers' activity choice and home-stay duration: some empirical evidence. Forthcoming in *Transportation.*

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

Murakami, E. and Watterson, W. (1990) Developing a household travel panel survey for the Puget Sound region. *Transportation Research Record* **1285**, 40-46.

Murakami, E. and Watterson, W. (1991) Attrition and replacement issues in the Puget Sound transportation panel. Presented at the 70th Annual Meeting of the Transportation Research Board, Washington, D.C.

Pas, E. and Koppelman, F. (1986) An examination of the determinations of day-to-day variability in individuals' urban travel behavior. *Transportation* **14**, 3-20.

Recker, W., McNally, M. and Root, G. (1986a) A model of complex travel behavior: part I - theoretical development . *Transportation Research* **20A**(4), 307-318.

Recker, W., McNally, M. and Root, G. (1986b) A model of complex travel behavior: part II - an operational model. *Transportation Research* **20A**(4), 319-330.

Train, K. (1986) *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand.* MIT Press, Cambridge, MA.

Van der Hoorn, T. (1983) Development of an activity model using a one-week activity-diary data base. In: S. Carpenter and P. Jones (eds.) *Recent Advances in Travel Demand Analysis*. Gower, Adershot, England.

Vilcassim, N.J. and Jain, D.C. (1991) Modeling purchase-timing and brand-switching behavior incorporating explanatory variables and unobserved heterogeneity. *Journal of Marketing Research*, **28**, 29-41.