Sequential, History-Dependent Approach to Trip-Chaining Behavior

RYUICHI KITAMURA

The characteristics of trip-purpose chains are examined, and a sequential model of trip chaining, which consists of history-dependent probabilities of activity choices, is developed. Statistical analyses of the study indicate that there is a consistent hierarchical order in sequencing activities in a chain where less-flexible activities tend to be pursued first. The analyses also indicate that the set of activities pursued in a chain tends to be homogeneous. Thus activity transitions are more organized and systematic than what a Markovian model would depict. Based on these findings, a sequential model of activity choice is formulated that, in spite of its simplified representation of the history of a chain, satisfactorily represents the observed behavior. Although the focus of the model is on direct linkages between activities, the model is capable of representing those characteristics associated with the entire set of activities in a chain. The results of the study strongly support the sequential modeling approach and indicate its practical usefulness in the analysis of trip-chaining behavior.

The importance of understanding trip-chaining behavior has long been recognized in connection with nonresidential trip generation (1) or with urban land use development (2). Underlying this is the dissatisfaction with the way tripmaking has been dealt with in the conventional transportation planning process or in location theory (3). As planning emphases in transportation shifted from infrastructure construction toward systems management and policy development, it was recognized that there was an increased need for a more fundamental understanding of travel behavior (4-6). The responses of urban residents to the recent oil crises (7,8) have made evident the importance of investigating trip-chaining behavior. Its importance is clearly seen when considering how the temporal and spatial distribution of trips in an urban area is affected by the way people organize their daily schedule of activities and combine trips. Statistical analyses have been accumulated to form a substantial body of empirical evidence [reviews of previous works on related subjects can be found in Hanson (3) and Damm (9)]. Yet many questions that have arisen in model-building efforts of trip-chaining behavior remain to be answered.

In this study one of the critical issues in trip chain modeling is addressed: representation of the decision structure involved in trip chaining. From the viewpoint that people plan and schedule beforehand a set of activities to be pursued in a trip chain, the decision process can be best represented as a simultaneous one that concerns the entire set of activities to be pursued and combined. Statistical analyses have been accumulated to form a substantial body of empirical evidence [reviews of previous works on related subjects can be found in Hanson (3) and Damm (9)]. Yet many questions that have arisen in model-building efforts of trip-chaining behavior remain to be answered.

In examining the adequacy of the sequential approach, two aspects are discussed: sequencing of activities in a trip chain, and tendencies or preferences in formation of the set of activities to be pursued in a trip chain. (This study is concerned with types and sequences of activities in a chain, but not with their spatial or temporal attributes.) A modeling effort that extends the present study into the temporal dimension can be found in a paper by Kitamura and Kermani shown elsewhere in this Record.) How these two aspects affect sequential probabilities of activity choice is demonstrated. Following this, empirical observations are made, and the nature of trip-chaining behavior is characterized.

BACKGROUND

The equivalence of the sequential and simultaneous approaches can be found in the following identity. By letting $X_n$ be the nth activity in a trip chain for the case of three activities,

$$Pr(X_1 = A, X_2 = B, X_3 = C) = Pr(X_3 = C | X_1 = A, X_2 = B) Pr(X_2 = B | X_1 = A) Pr(X_1 = A)$$

(1)

The probability that a given set of activities is chosen and pursued in a given order can be represented by a set of sequential and conditional probabilities. (The same identity has been used in relating simultaneous and sequential formulations of discrete choice.) When the conditionalities in Equation 1 is appropriately represented in sequential probabilities, then the sequential approach is equivalent to the simultaneous approach to trip chaining.

It may be argued that activity choice cannot be adequately described by probabilities that are conditioned only on the past; activity choice may also be dependent on future activities because a set of activities to be pursued may have been planned beforehand. Nevertheless, it can be seen that the backward dependency on the past implies forward dependency on the future as well. By using Bayes's rule,

$$Pr(X_1, X_2) = Pr(X_2 | X_1) Pr(X_1) / \left[ \sum_{X_1} Pr(X_2 | X_1) Pr(X_1) \right]$$

(2)
and so forth. A forward dependent probability can be always expressed as a function of backward dependent probabilities. That a choice is conditioned on the past implies that it is also conditioned on the future.

The preceding discussion indicates that the problems of previous sequential analyses, many of which used Markov chains, do not lie in their sequential structure, but rather they lie in their inadequate representation of the conditionality. In the following discussion it is assumed that there are patterns in sequencing activities in a set, and also that the choice probability of a given activity set is predetermined. The intensity of direct linkages between activities, or transition probabilities, which have been the main focus of previous studies, is viewed as a consequence of the patterns and preferences in choosing activity sets and sequencing activities. It is then shown that these patterns and preferences can be represented by the conditional transition probability, whereas the two Markovian assumptions—history dependence and stationarity (or time homogeneity)—are inadequate.

Suppose the number of activities in the set (denoted by k) is fixed and the individual is completely indifferent to the sequence of activities. Consider an activity set (w) and two activity types (A and B). Because the sequencing is completely random, all the sequences obtained by permutating the activities in w have the identical probability. Accordingly, for all w,

$$Pr(X_n = A, X_{n+1} = B | w) = Pr(X_n = A, X_{n+1} = B | w)$$

and

$$Pr(X_n = A | w) = Pr(X_n = A | w) \quad m, n = 1, 2, \ldots, k - 1$$

Then, if A is included in at least one activity set,

$$Pr(X_{n+1} = B | X_n = A) = \sum_{w} Pr(X_n = A, X_{n+1} = B | w) Pr(w)$$

$$= \sum_{w} \left[ \sum_{w} Pr(X_n = A | w) Pr(w) \right]$$

$$= Pr(X_{n+1} = B | X_n = A)$$

Namely, the pairwise activity transition probabilities are stationary. Note that this conclusion is not affected by the probability with which w is chosen [Pr(w)], i.e., it does not depend on the preferences in activity set choice.

Although the pairwise activity transition probabilities are stationary, they are not history independent even in this simplified case of random activity sequencing. Suppose that the choice probabilities of sets that include activities A, B, and C are zero, while those of other sets are positive. Then

$$Pr(X_{n+1} = B, \ldots, X_0 = C, \ldots, X_n = A) > 0$$

and

$$Pr(X_{n+1} = B, \ldots, X_0 = D, \ldots, X_n = A) = 0$$

Therefore, $Pr(X_{n+1} = 1, X_2, \ldots, X_n) \neq Pr(X_{n+1} = X_n)$. For the activity transitions to be Markovian, the probabilities with which respective activity sets are chosen must conform with those depicted by the transition matrix of a Markov chain, a condition rather groundless to assume.

The pattern of sequencing activities in a trip chain is another source of history dependence, which also yields nonstationarity. Suppose activity A tends to be pursued before B, but the individual is indifferent to the sequencing of activity C. Then for w that involves A, B, and C, the probability $Pr(X_{m+1} = B \ldots, X_n = C | w)$ varies depending on whether A has been pursued before C. Now suppose both A and B tend to be pursued earlier in a chain, but again C is equally likely to be pursued in any order. Then, $Pr(X_{m+1} = A | X_n = C, w) > Pr(X_{m+1} = A | X_n = C, w)$ if m < n. The first example indicates that the sequencing patterns cause history dependence, and the latter indicates that pairwise transitions become nonstationary.

Any Markov chain exhibits certain patterns of activity set formation and sequencing. But the reversal is not always true; i.e., given patterns of set formation and sequencing cannot always be represented by a Markov chain. The discussion in this section also implies that sequencing and activity set formation can be represented when the conditional probabilities of activity transitions are appropriately specified. The failure of Markov chain models is caused by their invalid representation of the conditional probability. In this way, the characteristics of trip chaining are first observed, and then a sequential model is proposed.

**DATA SETS**

Empirical observations of this study are made by using the 1965 Detroit area transportation and land use study (TALUS) data set, the 1977 Baltimore travel demand data set, and published transition frequency matrices from Chicago, Buffalo, and Pittsburgh [reported by Hemmens (19)]. The TALUS data set is most extensively analyzed, whereas the other sets are used to examine the generality of the results obtained. A significant advantage of the TALUS data set—a conventional origin-destination survey result—is its ample sample size, which is crucial for the analysis carried out in this study.

The original TALUS data file, which contained records of 320,090 trips made by 82,050 individuals, was screened to exclude those individuals who did not have a closed series of trips that originated and terminated at home (which may include intermediate returns to home), who had no car available to them, or those who did not hold a driver's license, who were younger than 18 years old, who used travel modes other than car, and those who made work trips on the survey day (walk trips are not recorded in the TALUS data unless they were work trips). The last criterion is introduced because of the substantial differences in travel and time use patterns between those who worked and those who did not on the survey day (20, 21). As a result of this screening, the sample analyzed includes 76,025 trips and 27,901 trip chains made by 16,520 individuals (a geographical subsample of this was used in previous studies (16, 20, 21)).

Three screening criteria are also applied to the Baltimore data set, and a sample of 1,789 trips and 697 trip chains made by 435 individuals is obtained. The transition frequency matrices from the other three metropolitan areas include all observations without comparable screening. As is clear from the screening criteria, the internal homogeneity of the sample is emphasized in this analysis, whereas some aspects of travel behavior are placed out of its scope, such as the effect of travel mode on trip chaining. Individuals with transit trips are eliminated for this reason, and they are not analyzed because their sample size is too small for statistical analysis.

The 27,901 trip chains in the sample from the TALUS data set contain 48,124 sojourns with an aver...
The direct transitions between activities in trip chains in the TALOUS and Baltimore samples were first analyzed by using a transition matrix, with the assumption that trip chaining can be represented by a stationary and history-independent Markov chain. These two samples are different from those of other studies in that the individuals who make work trips are excluded. Nevertheless, this preliminary analysis of the pooled transition matrices indicated that the present samples share many of the trip-purpose linkage patterns reported in the literature (11, 19).

## Nonstationarity of Activity Transitions

Although traditional Markov chain analysis (which uses the pooled transition matrix) offers a convenient means of data summarization, the implicit stationarity assumption that the same transition matrix applies to all transitions in a trip chain is too restrictive for rigorous analysis of the behavior. In this section the nature of trip chaining is explored by using a nonstationary Markov chain, where each step of transition has its own transition matrix that is not necessarily identical to those of other steps (the first step of transition refers to the transition from the first purpose to the second, the second step of transition is the one from the second purpose to the third, and so forth).

### Nonstationarity in Trip-Purpose Chains

Nonstationarity in the observed trip-purpose transitions is statistically examined by applying the likelihood-ratio test (22). The results are summarized in Table 1. To eliminate empty cells in the frequency matrices for as many steps as possible, two trip purposes with fewer observed frequencies are merged with others, as indicated in the table. The test is conducted for the first nine transition matrices, and also for the eight matrices from steps 2-9. The null hypothesis is strongly rejected in both cases.

Together with the overall chi-square values, the data in Table 1 present chi-square statistics for the row and column of each trip purpose, where the row total represents the nonstationarity in the transition probabilities from the trip purpose, and the column total represents that to the trip purpose. For the first case (steps 1, 2, . . . , 9), all rows and columns have significant statistics, except the column total for shopping, which indicates that shopping is pursued with a relatively stationary probability throughout a chain. The large chi-square value associated with the transitions to serve-passenger trips and that from social-recreation trips are also noted.

The second test excludes the transition matrix of the first step. The drastic reduction in the overall chi-square value from the first test indicates the extreme distinctiveness of the matrix for the first transition. Note that the first transition determines whether the individual pursues only one or more than one sojourn in a trip chain. The data in Table 1 also indicate that the variation in linkages with serve-passenger trips is a major source of nonstationarity in the second step and thereafter.

The pairwise distinctiveness of two successive transition matrices was also tested, and the first four matrices were found to be significantly different from each other (with chi-square values of 783.3 between the first and second steps, 92.9 between second and third, and 38.9 between third and fourth, all with df = 16). No significant difference was found after the fourth step. This is at least partly caused by the reduced sample size in the transition frequency matrices of later steps. At the same time, the implication of the result that the transition probabilities are stabilized in later steps of a trip chain is intuitively agreeable.

### Variations in Linkage Patterns

The nonstationarity in trip-purpose transitions implies that a pair of activities may have strengthening or weakening linkages with each other, and that some activities tend to be pursued earlier or later in a chain. The data in Table 2 indicate by step of transition those trip-purpose pairs for which more than expected transitions are observed in respective steps. Many of the diagonal cells are significant in all steps, which indicates that activities of the same type continue to have strong linkages among themselves throughout the chain. There are also several trip-purpose combinations that are significant only in the first few steps or in later steps.

### Table 1. Likelihood-ratio test of stationarity in trip-purpose transitions.

<table>
<thead>
<tr>
<th>Trip Purpose</th>
<th>For s = 1, . . . , 9</th>
<th>For s = 2, . . . , 9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Row Total(^a)</td>
<td>Column Total(^b)</td>
</tr>
<tr>
<td>Home</td>
<td>377.6(^e)</td>
<td>155.2(^e)</td>
</tr>
<tr>
<td>Personal business(^e)</td>
<td>377.6(^e)</td>
<td>155.2(^e)</td>
</tr>
<tr>
<td>Social-recreation(^b)</td>
<td>603.7(^e)</td>
<td>174.7(^e)</td>
</tr>
<tr>
<td>Shopping</td>
<td>374.1(^e)</td>
<td>36.5</td>
</tr>
<tr>
<td>Serve passengers</td>
<td>270.1(^e)</td>
<td>87.7(^e)</td>
</tr>
<tr>
<td>Total(^j)</td>
<td>1,585.4(^e)</td>
<td>1,585.4(^e)</td>
</tr>
</tbody>
</table>

Note: In places where degrees of freedom are indicated, the df for the column total cannot be defined in the conventional manner; therefore the ratio (total df) : (no. of columns) is presented here.

\(^a\) df = 22.

\(^b\) df = 22.4.

\(^c\) df = 22.4.

\(^d\) df = 22.4.

\(^e\) Significant at \(a = 0.005\).

\(^f\) Includes school.

\(^g\) Includes eat-meal trips.

\(^h\) A definition of the log-likelihood ratio statistic is given in Anderson and Goodman (7).

\(^i\) df = 128.

\(^j\) df = 112.

## Table 2. Salient trip-purpose linkages in nonstationary transition matrices for steps 1-4.

<table>
<thead>
<tr>
<th>Category</th>
<th>HOME</th>
<th>PBNS</th>
<th>SREC</th>
<th>MEAL</th>
<th>SHOP</th>
<th>SCHL</th>
<th>SVPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBNS</td>
<td>1,2,3,4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SREC</td>
<td>1</td>
<td>1,2,3,4</td>
<td>1,2,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEAL</td>
<td>1,2,3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHOP</td>
<td>1,2,3,4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCHL</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVPS</td>
<td>3</td>
<td>1,1,2,3,4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Steps 1 through 4 indicate the step of transition for which the cell has a chi-square value of 7.879 or greater with an expected frequency of 1 or greater. Abbreviations for trip-purpose categories are as follows: PBNS = personal business, SREC = social-recreation, MEAL = eat meal, SHOP = shopping, SCHL = school, and SVPS = serve passengers.
Especially noted is the transition from serve-passenger trips to home in the third step. The significance of this trip-purpose combination in this particular stage is caused by the dominance of the trip-purpose sequence: serve passengers to other activity to serve passengers to home (a later section indicates that this is a typical sequence when a trip chain involves serve-passenger trips). Thus the result suggests that the observed nonstationarity is partly caused by the sequencing by the tripmaker of the activities within a trip chain.

The variations in trip-purpose linkages were further characterized by evaluating for respective steps the mean first passage times (MFPTs); that is, the expected number of transitions from an origin state to each state visited for the first time (23). The result indicated that the linkages to personal business become weaker in later steps of a chain. On the other hand, the MFPTs to serve-passenger and social-recreation trips revealed strengthening linkages between these activities and others in later steps.

This analysis of nonstationarity in trip-purpose chains strongly suggests the existence of patterns in sequencing activities. An earlier section indicated that another possible source of the observed nonstationarity is the dependence of activity choice on the set of activities already pursued, which is closely related with the preferences in the choice of activity set. In the following sections these two aspects are discussed, and the reasons why such nonstationarity exists in trip-chaining behavior are illuminated.

ACTIVITY SEQUENCING IN A TRIP CHAIN

Consider the transition frequency matrix presented in Table 3, which gives direct transitions between activities in 10,555 multijourney trip chains in the TALUS sample. The matrix is obviously not symmetric, i.e. the frequency of (i;j) transitions is not always similar to that of (j;i) transitions. Examination of this asymmetric nature of the matrix leads to inferences as to the sequencing of activities within a trip chain. Suppose that three activities (A, B, and C) are to be pursued in a chain. If the tripmaker is completely indifferent to the sequence of these activities, all of the 3! possible sequences would have the equal likelihood of occurrence, and the occurrence of each one of the 6 (= 3! 2) possible direct transitions would have the identical probability. Accordingly, the observed transition frequency matrix must be asymmetric. The asymmetric matrix of Table 3, therefore, suggests that certain activities tend to precede others in multijourney chains.

<table>
<thead>
<tr>
<th>Category</th>
<th>PBNs</th>
<th>SREC</th>
<th>MEAL</th>
<th>SHOP</th>
<th>SCHL</th>
<th>SVPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBNs</td>
<td>1,522a</td>
<td>81b</td>
<td>212a</td>
<td>1,820b</td>
<td>27a</td>
<td>515a</td>
</tr>
<tr>
<td>SREC</td>
<td>663a</td>
<td>1,562a</td>
<td>285b</td>
<td>1,901a</td>
<td>16a</td>
<td>722</td>
</tr>
<tr>
<td>MEAL</td>
<td>114a</td>
<td>277b</td>
<td>16a</td>
<td>191b</td>
<td>17a</td>
<td>158</td>
</tr>
<tr>
<td>SHOP</td>
<td>844a</td>
<td>1,122a</td>
<td>188b</td>
<td>3,109a</td>
<td>4b</td>
<td>674a</td>
</tr>
<tr>
<td>SCHL</td>
<td>46a</td>
<td>43b</td>
<td>27a</td>
<td>59b</td>
<td>23a</td>
<td>59f</td>
</tr>
<tr>
<td>SVPS</td>
<td>618a</td>
<td>737b</td>
<td>140b</td>
<td>1,030a</td>
<td>93a</td>
<td>1,564a</td>
</tr>
</tbody>
</table>

Note: Abbreviations are defined in Table 2. The footnotes in the table, except a, give the significance of the asymmetry between (i,j) and (j,i) cells.

Tendencies in Activity Sequencing

Examination of the pooled transition frequency matrix of Table 3 indicates that the frequencies that are statistically most asymmetric involve personal business; for example, 815 transitions from personal business to social-recreation versus 462 transitions from social-recreation to personal business; 212 transitions from personal business to eating meal versus 114 from eating meal to personal business; and so forth (the differences are significant at α = 0.005). Obviously, personal business tends to be pursued in a chain before the other activities. School trips have a similar tendency, and they precede personal business trips more frequently, and serve-passenger trips have a tendency to precede school and personal business trips.

There are also several pairs of trip purposes of whose sequences the tripmaker is apparently indifferent; 285 transitions from social-recreation to eating meal versus 277 from eating meal to social-recreation; 1,091 from social-recreation to shopping versus 1,122 from shopping to social-recreation; and so forth. None of these differences is statistically significant at any appropriate level.

Nine of the 15 (= C2) pairs of different trip purposes have statistically significant asymmetry (α = 0.05). Based on these relationships, a hierarchy diagram is constructed to show the tendencies in sequencing activities within a trip chain (Figure 1a). The perfect consistency in the hierarchical relationship among the trip purposes is shown in the figure; for example, serving passengers, which precedes school, also precedes those trip purposes that school precedes. These consistent tendencies in the observed direct transitions are quite noteworthy.

Hierarchical relationships among activities are evaluated in the same manner by using transition frequency matrices from Chicago, Buffalo, and Pittsburgh; these are summarized in Figure 1b. The result is in satisfactory agreement with the TALUS result. This is also the case for Baltimore, but the sample size is insufficient to be conclusive.

Activity Sequencing and Uncertainty

The hierarchical order of activities presented in Figure 1a,b indicates that activities in the higher order tend to be accompanied with spatial or temporal fixity, or both. For example, serving a passenger quite often implies that a person must be chauffeured to a given location by a given time, personal business such as banking must be pursued at a predetermined location, and so forth. The result indicates that activities of less flexibility tend to be pursued in a trip chain before more flexible activities, such as social-recreation and shopping. Cullen and Godson (24) argued that an individual's itinerary for a day is formed by articulating activities with high spatial or temporal fixity or both, which act as pegs in daily activity scheduling. A previous study (29) revealed that serve-passenger trips largely prescribe an individual's daily travel pattern because of their fixity. The present study reveals another tendency in urban travel behavior: a relationship between sequencing of activities and their fixities.

The information available from the data set does not allow statistical determination of the reason why this sequencing pattern is observed. Nevertheless, the consistent observations from the four metropolitan areas offer the basis for constructing behavioral inferences on the subject. A rather
straightforward conjecture postulated here is that the sequencing pattern observed in this study is a result of individuals' consideration of uncertainty in activity planning.

Consider the case where an individual is combining both fixed and flexible activities into a chain. Quite typically, the exact amount of time required to accomplish an activity is not known to the individual beforehand. If a flexible activity is pursued first, and if it takes longer than initially thought, then the individual may not be able to attend the fixed activity in time. Note that an activity with spatial and temporal fixity by definition demands the individual to be at a certain location by a certain time. On the other hand, if the flexible activity takes less time, an unexpected block of time must be somehow spent. In either case, if the individual recognizes this uncertainty, it appears logical for him to pursue the fixed activity first. The observed activity-sequencing pattern thus suggests that uncertainty plays a significant role in the activity planning of an individual. The pattern is perhaps a result of an individual's effort to minimize risks because of the uncertainty and to pursue a set of activities efficiently in a trip chain.

**HISTORY DEPENDENCE IN ACTIVITY CHOICE**

An earlier section indicated that preferences in activity set choice in general make activity transitions history dependent. The sequencing pattern observed in the previous section implies that activity choice depends on the series of activities already pursued in a trip chain. The strong direct linkages among activities of the same type also suggest history dependence. However, little exploration of the nature of history dependence in trip chaining has been made in the past, and most analyses were concerned only with direct linkages between pairs of activities. The analysis of this section, which focuses on the entire series of activities in trip chains, reveals additional characteristics of activity set formation and activity sequencing.

Although there are many possible ways of statistically examining the history dependence in trip chaining (e.g., triples used by Parkes and Wallies [25] and also see Anderson and Goodman [23]), most of them encounter problems with sample size because of the scarcity in the sample of chains with a large number of sojourns. Accordingly, this study takes on an approach of tabulating the frequency of chains by the trip-purpose sequence and directly examining the history-independence assumption by using a contingency table analysis technique.

**History Dependence of Three-Sojourn Chains**

Consider those trip chains with three sojourns, namely, \(X_1\), \(X_2\), and \(X_3\) ≠ home, and \(X_4\) = home. The history-independence assumption can be stated for these chains as

\[
\Pr(X_3 = k|X_1 = i, X_2 = j) = \Pr(X_3 = k|X_2 = j)
\]

for all \(i, j\), and \(k\) ≠ home. Namely, the conditional probability that the third activity is \(k\) given the second activity (= \(j\)) is independent of the first activity (= \(i\)). This null hypothesis can be tested by tabulating, for given \(X_2\), the frequencies of the third activity categories by the first categories, then by examining the independence of the resulting two-way contingency table. This contingency analysis is equivalent to applying a nonstationary Markov chain of the first order to test the history independence of three-sojourn chains. The results for 2,760 three-sojourn chains found in the TALUS sample are given in Table 4. To ensure a sufficient number of observations for each sequence of trip purposes, the original six trip-purpose categories are collapsed into four, as in Table 1.

In part A of Table 4 the results for those three-sojourn chains whose second trip purpose is personal business (including school) are presented. The row represents the first trip purpose, and the column represents the third trip purpose. If the history-independent assumption holds, then every row should have the same distribution of cell frequencies. The expected cell frequencies under this independence assumption are shown in parentheses.

As expected, the four contingency tables (parts A-D, Table 4) are all highly significant, which indicates that the conditional probability that a certain activity is pursued as the third activity, given the second one, does depend on the first activity pursued in the chain. Especially notable are the much higher-than-expected frequencies of the diagonal cells; tripmakers tend to repeat the same type of activity as the first and third activities in a trip chain. This recurrence of the same activity type is particularly noticeable for serve-passenger trips when the second purpose is not serving passengers (see parts A-C of Table 4). The diagonal cell for serving passengers alone accounts
Table 4. Frequencies of three-sojourn chains by sequence of trip purposes.

<table>
<thead>
<tr>
<th></th>
<th>( X_1 ) = PBNS</th>
<th>( X_2 ) = SREC</th>
<th>( X_3 ) = SHOP</th>
<th>( X_3 ) = SVPS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X_1 ) = PBNS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBNS</td>
<td>(66.2)</td>
<td>(48.3)</td>
<td>(58.4)</td>
<td>(51.7)</td>
<td>(56.7)</td>
</tr>
<tr>
<td>SREC</td>
<td>(20.3)</td>
<td>(13.0)</td>
<td>(25.9)</td>
<td>(15.8)</td>
<td>(21.3)</td>
</tr>
<tr>
<td>SHOP</td>
<td>(30.2)</td>
<td>(19.4)</td>
<td>(38.7)</td>
<td>(23.6)</td>
<td>(31.3)</td>
</tr>
<tr>
<td>SVPS</td>
<td>(48.3)</td>
<td>(31.4)</td>
<td>(61.8)</td>
<td>(37.4)</td>
<td>(46.2)</td>
</tr>
<tr>
<td>Total</td>
<td>(165)</td>
<td>(106)</td>
<td>(211)</td>
<td>(129)</td>
<td>(611)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( X_1 ) = PBNS</th>
<th>( X_2 ) = SREC</th>
<th>( X_3 ) = SHOP</th>
<th>( X_3 ) = SVPS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBNS</td>
<td>(17.6)</td>
<td>(65.8)</td>
<td>(35.9)</td>
<td>(67.7)</td>
</tr>
<tr>
<td>SREC</td>
<td>(15.0)</td>
<td>(109.2)</td>
<td>(39.8)</td>
<td>(79.6)</td>
</tr>
<tr>
<td>SHOP</td>
<td>(12.5)</td>
<td>(46.5)</td>
<td>(25.4)</td>
<td>(33.7)</td>
</tr>
<tr>
<td>SVPS</td>
<td>(23.3)</td>
<td>(80.3)</td>
<td>(63.9)</td>
<td>(38.2)</td>
</tr>
<tr>
<td>Total</td>
<td>(81)</td>
<td>(302)</td>
<td>(165)</td>
<td>(219)</td>
</tr>
</tbody>
</table>

Personal business (PBNS) includes school, and social-recreation (SREC) includes eating meal. For other abbreviations, see Table 2.

\( X^2 \) : Expected cell frequency
\( X^3 \) : Row, column, or grand total of chi-square values.

for 40.0 percent of the total chi-square value of part A where the second purpose is personal business. The corresponding values are 38.1 percent for part B (\( X_2 \) = social-recreation), and 48.8 percent for part C (\( X_2 \) = shopping). The sequence of serve passengers to other activity to serve passengers is observed much more frequently than the expectation under the history-independence assumption, and it is found in 12.2 percent of the all three-sojourn chains, or in 36.6 percent of those three-sojourn chains that involve serve-passenger trips at all. The corresponding statistics from the Baltimore sample are 14 and 36 percent, respectively. This sequence pattern is obviously caused by the typical requirement that a person chauffeured and dropped off at a place has to be picked up later. The examination of individual cells of parts A-C also indicates that the probability that the third purpose is serving passengers is significantly smaller than the expectation when the first and second purposes are not serving passengers.

The data in Table 4 also indicate that the activities pursued in a chain quite often all fall within one trip-purpose category. For example, the sequences shopping to shopping to shopping and social-recreation to social-recreation to social-recreation are the most frequently observed sequences. This, together with the recurring tendency previously discussed, indicates that the activities pursued in a chain tend to be homogeneous. Of the 2,760 chains, 61.9 percent involve only one trip-purpose category, 22.3 percent involve two, and only 14.5 percent involve three different trip-purpose categories as defined here. These observations differ substantially from the expected values obtained by assuming complete independence in trip-purpose transitions (i.e., Markov chain of the 0th order): 6.9, 56.8, and 35.3 percent, respectively. In the Baltimore sample 72 percent of chains with three or more sojourns involve only one or two trip-purpose categories.

The sequences of activities in these three-sojourn chains showed exactly the same hierarchical order as in Figure 1. Note that this analysis takes into consideration the sequences of indirectly linked activities. This can be seen in part by examining the asymmetry of the matrices presented in Table 4.

Similar tabulations and analyses were done for 1,164 four-sojourn chains in the TALUS sample with the same classification of trip purposes into four categories. However, of the 256 (+ 4) possible sequences of trip purposes, 150 had observed frequencies of 3 or less, which warranted only limited statistical examination of these chains. Even a data set of 76,025 trip records appears insufficient for rigorous statistical investigation of history dependence in trip chains. Nevertheless, available statistics indicate that the inferences made for the three-sojourn chains are likely to apply to the four-sojourn chains. For example, 547 (47 percent) of the all four-sojourn chains involved only one or two trip-purpose categories. Again, tripmakers tend to pursue only a few types of activities in a chain. Of the 292 chains that contain two or three serve-passenger trips, 220 (75.3 percent) involve the sequences of serve passengers to other activity to serve passengers, or serve passengers to other activity to other activity to serve passengers.

Possible Explanation of Homogeneity

Obviously, the temporal and spatial distributions of opportunities are among the factors that contribute to the homogeneity of activity types pursued in a trip chain. For example, pursuing personal business is not likely in the evening because businesses or shops are typically closed, and chains made in the evening tend to be social-recreation oriented. Commercial corridor development provides many shopping opportunities in close proximity, thus making shopping trip chains convenient and economical.

It may also be hypothesized that the individual has clear perception as to the compatibility of dif-
different types of activities in a chain (closely related is the viewpoint that individuals perceive different time periods of a day as suitable for different types of activities (15, 24)). In some extreme cases a set of activities may be viewed as a full course of activities and pursued as such; for example, movie to late dinner to home, or visit a friend to bowling lane to a pizza house to drive the friend home to home. The latter forms a four-sojourn chain that involves social-recreation and serve-passenger trips. If a bundle of activities is pursued in an integrated activity, introduction of a heterogeneous activity into the chain may not be acceptable to the individual. Even the way the individual is dressed may affect trip chaining. If the individual leaves home to pursue a homogeneous set of activities and is dressed suitably for this set, he may feel uncomfortable to visit locations for heterogeneous activities where he may be overdressed or underdressed. This may be especially the case in the TALUS sample because of its survey date and the conservative nature of the region.

SEQUENTIAL MODEL OF TRIP CHAINING

Based on the results of the previous sections, which examined the nature of trip-chaining behavior from various viewpoints, a simple model of the behavior is developed and tested. The key issue in the modeling effort is how to represent the past history of a chain in a simple and practical form.

Model Framework

The model development effort is based on the premise that the observed characteristics of trip chaining can be adequately represented by making the transition probabilities dependent on the past history of the chain. The findings that activities in a chain tend to be homogeneous and that inflexible activities tend to be pursued first suggest a rather simple and systematic structure of history dependence. The probability of a given activity transition strongly depends on the types of activities already pursued, but may not depend on the number of times these activities were engaged in or on the exact order in which they were pursued. For example, the sequencing pattern implies that once flexible activities have been pursued, the probability of an inflexible activity is small, but the number of the previous flexible activities and their sequence may have only a negligible effect on the probability. The exact representation of the history, as shown in Equation 1, may not be necessary, and a simpler representation may be adequate.

The conditional probability of activity choice is formulated as follows:

\[
\text{Pr}(X_{n+1}|X_1, X_2, \ldots, X_n) = \text{Pr}(X_{n+1}|X_n, D_{1n}, D_{2n}, \ldots, D_{Kn})
\]

where \(D_{jn}\) is a binary (0-1) variable, which indicates whether activity type \(j\) has been pursued in the chain by the \(n\)th transition, and \(K\) is the number of activity categories used to represent the history of the chain. The model assumes that activity transition probabilities depend on direct linkages; thus the probability of the next activity \((X'_{n+1})\) is conditioned on the current activity \((X_n)\). History of a chain, however, is not represented by the entire series of activities pursued, but by a set of binary variables \((D_{1n}, \ldots, D_{Kn})\).

For example, suppose that activities are classified into four categories (serve passengers, personal business, social-recreation, and shopping), and the \(D_{jn}\)'s are defined for \(j = 1, 2, 3\) as

- \(D_{1n} = 1\) if serving passenger has been pursued by \(n, 0\) otherwise;
- \(D_{2n} = 1\) if personal business has been pursued by \(n, 0\) otherwise; and
- \(D_{3n} = 1\) if social-recreation or shopping has been pursued by \(n, 0\) otherwise.

Two activity categories (social-recreation and shopping) are grouped together in defining the \(D_{jn}\)'s (note they are tied in the activity sequencing in Figure 1a). There are eight possible values of vector \(D_n = (D_{1n}, D_{2n}, D_{3n})\). Recall that some transitions are likely to occur with the sequencing pattern found earlier, whereas others are less likely to occur. Naturally, the probabilities of the first group will be larger than those of the second group for given \(X_{n+1}\) and \(X_n\). The activity choice probability of the model captures this variation because it is conditioned on \(D_n\) and thus replicates the sequencing pattern. The preferences in activity set choice are represented by the conditional probability in a similar manner.

The process \((X_1, X_2, \ldots)\) depicted by the history-dependent probability \(\text{Pr}(X_{n+1}|X_n, D_{1n}, D_{2n}, D_{3n})\) can be represented as a stationary, history-independent Markov chain process if the states are redefined and the state space is expanded appropriately. The set of states and the structure of the transition matrix for this example are shown in Figure 2. The states are now defined in terms of both activity type and past history of the chain (1.e., vector \(D_n\)). Accordingly, the past history, as expressed by \(D_n\), is automatically specified when the state of the process is designated, which implies that

\[
\text{Pr}(X_{n+1} = j|X_n = i; D_{1n}, D_{2n}, D_{3n}) = \text{Pr}(X_{n+1} = j|X_n = i) = P_{ij}
\]

where \(X_n^*\) is the redefined \(n\)th state. Namely, the process \((X_1^*, X_2^*, \ldots)\) is a Markov chain pro-
cess. This expansion makes the analysis quite straightforward, and statistical evaluation of the model can be done as in a standard Markov chain analysis.

**Estimation Result**

Sequential models of activity linkage are estimated by using the 27,901 trip chains in the TALUS sample with the classification of activities into four types, as in the preceding example. Five models with different transition structures are examined:

1. Stationary, history-independent model;
2. Nonstationary, history-independent model;
3. History-dependent model with three elements in $D_n$;
4. History-dependent model with four elements in $D_n$; and
5. History-dependent model, a hybrid of models 3 and 4.

Models 1 and 2 are studied here as references against which the history-dependent models are compared. Based on the results presented earlier, model 2 assumes a stationary transition matrix after the fourth transition. Model 3 is the one described in the previous example. The history indicator $D_n$ of model 4 is defined for the four activity types without grouping shopping and social-recreation together, as in model 3. Model 5 uses the same $D_n$ as model 4. However, no further difference is assumed in model 5 as to the history dependence of activity transitions after serving passengers, personal business, and either one of social-recreation or shopping are all pursued in a chain.

Transition probabilities of each model are estimated by the maximum likelihood method. The goodness of fit in terms of the log-likelihood value and square sum of errors in predicting the frequency of each activity sequence is given in Table 5. The latter statistic excludes chains with five or more sojourns (about 4 percent of the entire sample) for computational reasons. The improving goodness of fit of the model found in the table as the number of parameters increases is not surprising. More important, however, is that systematic prediction errors diminish as the thorough treatment of history dependence is made. The agreement between the observed and predicted frequencies of respective activity sequences is shown in Figure 3 for models 1 and 5.

Model 1 (Figure 3a), a standard Markov chain model, significantly underestimates the frequencies of single-sojourn chains, overestimates most of two-sojourn sequences, and makes extremely large errors in evaluating the frequencies of chains that involve recurrence of activities, especially those involving the following sequence: serve passengers to other activities to serve passengers. The nonstationary model (model 2) almost perfectly replicates the distribution of chain lengths. Nevertheless, chains starting with shopping are mostly underestimated, and sequences that involve serve-passenger trips are estimated with large errors.

<table>
<thead>
<tr>
<th>No.</th>
<th>Model</th>
<th>No. of Parameters</th>
<th>$-2(L)$</th>
<th>Chi-Square</th>
<th>df</th>
<th>SSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Stationary, history-independent</td>
<td>20</td>
<td>-58,403</td>
<td>-</td>
<td>-</td>
<td>723,648</td>
</tr>
<tr>
<td>2</td>
<td>Nonstationary, history-independent</td>
<td>100</td>
<td>-57,668</td>
<td>1,471.2</td>
<td>80</td>
<td>117,969</td>
</tr>
<tr>
<td>3</td>
<td>History dependent, three-element</td>
<td>100</td>
<td>-56,642</td>
<td>3,522.0</td>
<td>80</td>
<td>34,330</td>
</tr>
<tr>
<td>4</td>
<td>History dependent, four-element</td>
<td>180</td>
<td>-56,336</td>
<td>4,133.4</td>
<td>160</td>
<td>8,689</td>
</tr>
<tr>
<td>5</td>
<td>History dependent, hybrid</td>
<td>150</td>
<td>-56,362</td>
<td>4,083.0</td>
<td>130</td>
<td>8,737</td>
</tr>
</tbody>
</table>

Note: $L = \log$ likelihood; $-2(\Delta L) = -2(L$ of model 1) - (L of the model$); and SSE = square sum of errors.

**Figure 3.** Observed and expected frequencies of trip-purpose sequences.

a. **STATIONARY, HISTORY INDEPENDENT MODEL (MODEL 1)**

b. **HISTORY DEPENDENT MODEL (MODEL 5)**
The history-dependent models (models 3–5) largely improve these defects. Model 3, however, still shows significant errors for chains that involve social-recreational or shopping trips, which suggests that grouping these two activity types when representing the history of a chain is not adequate. Examination of the log-likelihood value between this model and models 4 and 5 also indicates this. The performances of models 4 and 5 (Figure 3b) are satisfactory, and very few activity sequences are predicted with significant errors. Note that model 5 performs almost as well as model 4, even though it has 30 less parameters. The satisfactory agreement between the observed and predicted frequencies implies that the patterns in activity sequencing and activity set formation are well represented by the model, and also that the model adequately captures the history of a chain. A simple representation of the history of a trip chain by means of a set of binary variables makes possible a satisfactory replication of trip-chaining behavior.

CONCLUSIONS

The statistical analysis of this study found that there is a consistent hierarchical order in sequencing activities where less-flexible activities tend to be pursued first. It was also found that the set of activities pursued in a trip chain tends to be homogeneous. Thus, activity transitions are more organized and systematic than what a Markovian process would depict. The homogeneity of activity types, patterns in sequencing activities, history dependence, and nonstationarity in activity transitions are all closely interrelated. Accordingly, it was possible to develop a sequential, history-dependent model of activity transition that, in spite of its simplified representation of the history of a chain, well replicated the observation. Although the focus of the model was on direct transitions of activities, the model was capable of representing those characteristics found for the entire chain (e.g., homogeneity and recurrence of activities and patterns in indirect transitions). The result strongly supports the sequential modeling approach adopted in this study. The usefulness of the model can be enhanced when the history-dependent probabilities are related to exogenous factors. This is another step that must be taken before the sequential model can be applied to practical problems.

Although the focus of this study was on the basic characteristics of trip chaining and its representation by sequential probabilities, the study results have some practical implications. The strong regularity implied by the homogeneity of trip chains suggests that people's responses to changes in the travel environment may be limited, as far as trip chaining is concerned. People organize their trip chains while considering the types of activities, but they may not necessarily minimize travel distance or cost. The importance of uncertainty in activity scheduling suggested by the observed sequencing pattern also implies this. Thus, travel patterns may be less sensitive to travel cost than was expected. The rather surprising result that the post-energy crisis Baltimore sample has a mean chain length that is 10 percent shorter than the 1965 Detroit sample also supports this claim. This conjecture, however, is subject to further investigation. Additional subjects that can be suggested for future investigation include examination of hierarchical relationships in time allocation and spatial choices for the trip chain, extension of the analysis to incorporate temporal and spatial aspects and verifying the present findings in that context, and investigation of the characteristics of all trip chains made by an individual within the study period and of the independence among these chains.

ACKNOWLEDGMENT

This study was partly supported by a grant from the National Science Foundation to the University of California, Davis.

REFERENCES

Identifying Time and History Dependencies of Activity Choice

RYUICHI KITAMURA AND MOHAMMAD KERMANSHAH

In this study a sequential model of activity patterns is formulated that consists of time- and history-dependent models of activity choice. This analytical framework is used to identify time-of-day and history-dependent characteristics of activity choice by statistically testing a series of hypotheses. The results indicate that the simplest expression of the history of activity engagements is an adequate descriptor, and also that non-home-based activity choice is conditionally independent of the activities in the previous day, given the activities pursued in the current trip chain. Interdependencies of activity types across trip chains are also characterized by estimated model coefficients. The results of the study indicate that the decisions associated with the entire activity pattern can be decomposed into interrelated activity choices whose conditional dependencies can be statistically evaluated.

The way individuals schedule their daily activities and organize their itineraries has immediate impacts on the spatial and temporal distribution of trips, or needs for trips, in an urban area. Therefore, representing how the choice and scheduling of activities are done and how travel patterns are formed are critical elements in travel-demand forecasting as well as in basic travel-behavior research (1-3). This is especially so when attempting to forecast the impact on novel changes in the travel environment or when seeking a transportation policy that will accomplish given objectives most effectively.

The mechanism by which trips as induced demand are generated is complex. Even when only scheduling is considered (i.e., when and in what order a given set of locations is visited and how these visits are arranged into trip chains), there are numerous scheduling possibilities. Choice of activities and their locations further complicates the problem. Constraints that govern the behavior are not limited to monetary and time budgets as in the classical utility maximization framework in economics, but include spatial and temporal fixity constraints associated with the respective activities (4), interpersonal linkage constraints (5), and other types of constraints that portray the travel environment of each individual (6). The interrelated activity choices underlying an activity-travel pattern are dependent on the time of day, as many previous studies on time use have indicated (7,8). Previous empirical evidence (9), and paper by Kitamura elsewhere in this Record) at the same time indicates that the choices are dependent on history, i.e., the set of activities already pursued on that day.

These aspects of daily activity and travel behavior are all of particular importance for the understanding and forecasting of the behavior. In particular, the time-of-day and history dependencies of activity choice may be viewed as the most fundamental elements, whose adequate representation will lead to representation of other important aspects of the behavior as well. For example, the preferences in forming a set of activities in a trip chain can be described by sequential probabilities of activity choice when their history dependencies are appropriately incorporated (see paper by Kitamura elsewhere in this Record). By specifying the structure of the time-of-day and history dependencies and estimating the model statistically, an important objective can be accomplished: characterization of activity and travel patterns along the time dimension. When the model includes exogenous factors that are related to changes in the travel environment or in the population characteristics, then the model serves as a tool for forecasting possible changes in activity and travel behavior.