ANALYSIS OF LONGITUDINAL DATA FROM THE PUGET SOUND TRANSPORTATION PANEL

Task F: Cross Section and Dynamic Analysis of Activity and Travel Patterns in PSTP

FINAL REPORT

Submitted to

Federal Highway Administration
and
Battelle Memorial Institute
Work Order BAT-94-016

June 1996

PENNSTATE

1855

Pennsylvania Transportation Institute

The Pennsylvania State University
Research Offiice Building
University Park, PA 16802
(814) 865-1891

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Analysis of Longitudinal Data from the Puget Sound Transportation Panel Task F: Cross Section and Dynamic Analysis of Activity and Travel Patterns in PSTP		5. Report Date 6. Performing Organization Code
7. Author(s) Konstadinos G. Goulias and Jun Ma		8. Performing Organization Report No. PTI 9623
9. Performing Organization Name and Address The Pennsylvania Transportation Institute Research Office Building The Pennsylvania State University University Park, PA 16802-4710		10. Work Unit No. (TRAIS) 11. Contract or Grant No. BAT-94-016
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Highway Administration 400 Seventh Street SW Washington, DC 20590	Battelle Memorial Institute 505 King Avenue Columbus, OH 43201-2693	13. Type of Report and Period Covered Final Report 9/1/94-8/31/96 14. Sponsoring Agency Code

15. Supplementary Notes
Federal Highway Administration COTR: Elaine Murakami

16. Abstract

This report presents analyses of activity and travel patterns that explicitly account for transportation level of service. The analyses are conducted using composite accessibility measures, land-use type, and density around the residences and workplaces of survey participants in the Seattle region. Using Puget Sound Transportation Panel (PSTP) data and cluster analysis, individuals' activities and travel patterns are first classified into a few relatively homogeneous behavioral groups. Then, cross classification is used to study the relationship between activity and travel patterns from one day to another and from one year to the next. Study results indicate that although there is a strong link between activity and travel within a day, the link becomes weaker as the time frame increases. To study the relationship between individuals' characteristics and their activity and travel pattern choices, the derived accessibility measures of individuals' residences and workplaces are included in four-level multilevel pattern selection models that include temporal, spatial, household, and person effects. Contextual analysis is conducted by separating employed from unemployed persons. Analysis results provide evidence that level of service and land use are strong determinants of activity participation and trip making, however, their effect depends on employment status.

17. Key Words	18. Distribution Statement
Panel surveys, accessibility, contextual effect, activity and travel patterns	No restrictions. This document is available from the National Technical Information Service, Springfield, VA 22161

19. Security Classif. (Of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	46	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

This report highlights findings of a simultaneous contextual analysis of day-to-day and year-to-year activity and travel patterns of residents in the Seattle region. Individuals' activities and travel patterns are classified into a few relatively homogeneous behavioral groups using Puget Sound Transportation Panel (PSTP) data and cluster analysis. The representative groups are intended to reduce the diversity of individuals' behavior patterns and are developed using activity and travel indicators (e.g., the number of trips made by each mode and each activity type, the duration of activities by each type of activity, etc.). Observed changes in activity and travel-pattern groups-in terms of transitions from one pattern to another at both day-to-day and year-to-year levels-are also examined. This is followed by identification of influential factors that may govern the selection mechanism of daily activity and travel patterns. These factors include temporal effects (using day and year indicators), spatial effects (county of residence, workplace, etc.), accessibility of residence and workplace, socio-demographics of the place of residence, household socio-economics, and person characteristics. The temporal, spatial, and household variables are the macro-contextual effects influencing person-based selection of travel and activity patterns.

The remainder of this report is organized as follows. First, a brief description of the data is provided. The activity and travel patterns obtained for the four four waves (i.e., two-day travel diary information in 1989, 1990, 1992, and 1993) of PSTP are presented next. Then, a few multilevel models are presented and discussed. This is followed by a section on the relationship between factors that summarize accessibility and area so&demographics and a multivariate analysis of activity and travel pattern choice. An analysis of the dynamics of activity and travel using the contingency table technique is also provided, as are analyses of observed activity and travel pattern transitions. Final conclusions are offered in the report summary.

2. DERIVATION AND MULTIVARIATE ANALYSIS OF ACTIVITY AND TRAVEL PATTERNS

Activity-based travel demand forecasting is increasingly becoming a mainstream approach to travel demand modeling. Activity-based modeling aims to depict and predict the relationship between activity and travel using a finite number of representative patterns. This step is needed to create algorithms for the prediction of activity patterns and resulting (or concomitant) travel patterns. In the past, this stream of work attempted to summarize human behavior in terms of representative behavior of several relatively homogeneous groups using statistical/mathematical pattern recognition techniques (Pas, 1982, 1983; Reeker and McNally, 1985, 1986). A subsequent step along this line was to study changes in activity and travel patterns within a day or a week (Pas and Koppelman, 1987; Pas, 1988). Due to the lack of panel observations, these studies have adopted a limited time frame (e.g., a day), providing only a snapshot of continuously changing behavior and failing to capture long-term dynamic aspects of activity and travel behavior such as habit. Consideration of this is greatly facilitated by the availability of longitudinal data that allow researchers to not only model individuals' selection mechanisms of activity types and durations within a day (Hammed and Mannering, 1993; Ettema et al., 1995), but to extend the analysis to longer time frames (e.g., a year). Panel data from activity/travel diaries on different days of different years allow researchers to examine variations in activity and travel patterns from one day to the next and from one year to another simultaneously. In this way, the dynamics of activity participation and travel behavior are depicted in a holistic way. To do this, we first need to examine the effects of time on activity and travel patterns, the interaction between activity and travel over time, and the determinants of their changing relationship over time. Then, we can begin formulating activity-based travel forecasting systems.

Past activity and travel patterns can be used to extrapolate individual future behavior, relying on a series of more fundamental elements that determine behavior. These elements include individual characteristics and descriptors of the environment within which each individual acts (e.g., household, neighborhood, peer group, etc.). In this report, these elements are called contextual effects. Operationally, context can be reflected in individual-based models through the inclusion of socioeconomic characteristics of persons and their corresponding households, residence- and workplace-related land-use patterns, neighborhood characteristics, and so on (e.g., see the recent travel/activity modeling system in McNally and Wang, 1994). An individual' s activity participation and travel from a contextual viewpoint, used in this report, is based on the following rationale emerging from sociology (DiPrete and Forristal, 1994):

> The essential feature of all contextual-effects models is an allowance for macro processes presumed to have an impact on the individual actor over and above the effect of any individual-level variables that may be operating.

Models emerging from a contextual analysis of multilevel data, which consist of information on multiple macro units (contexts) and multiple micro units within each macro unit, are called contextual or multilevel models. Using this idea in activity and travel analysis, a multilevel model can be used to specify, at an individual level, the effect of social context on behavioral variables (e.g. individuals' activities and travel patterns as a function of "higher" levels of aggregation such as household, neighborhood, etc.). Such a model specification is the simplest form of multilevel models where the "macro-level" variables are used as explanatory variables of a micro-level model.

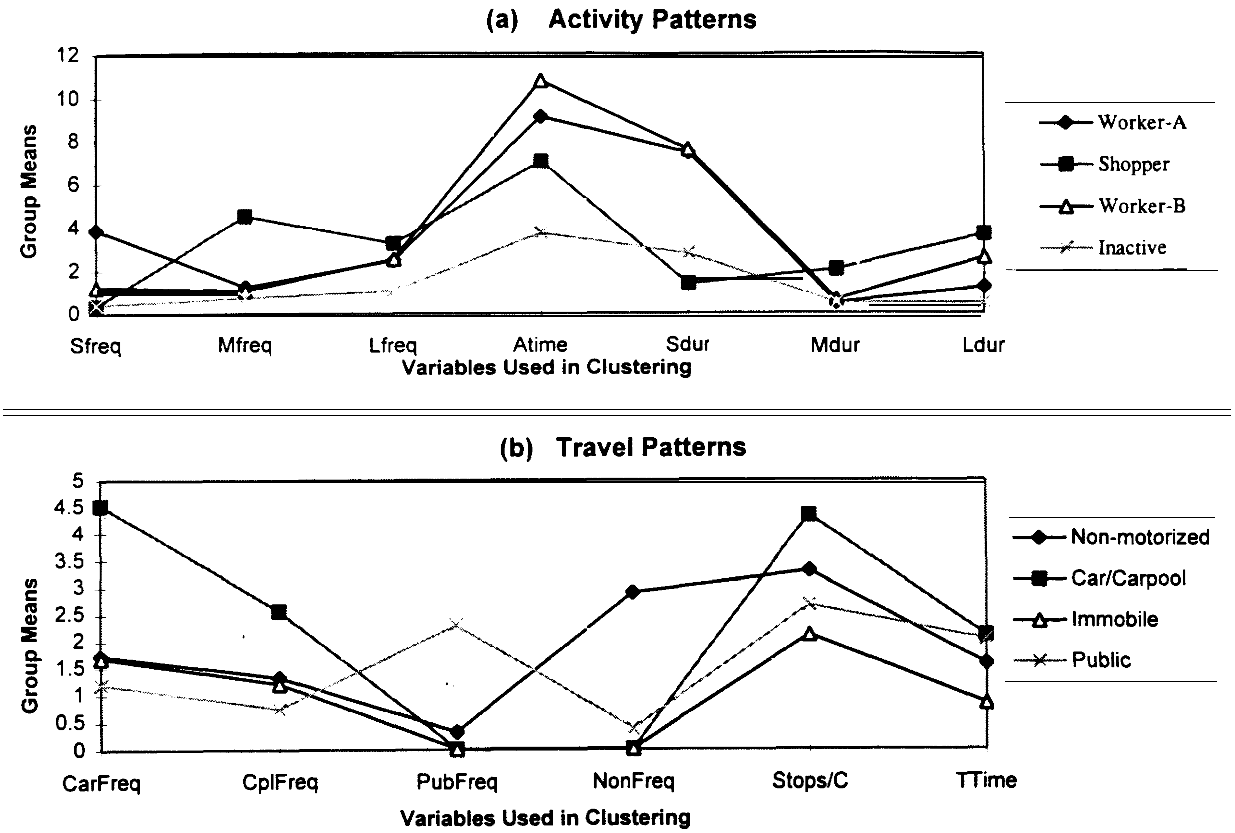
CLASSIFICATION OF ACTIVITY AND TRAVEL PATTERNS

PSTP, which started in 1989 in the Seattle metropolitan area (Murakami and Watterson, 1990), is a multi-wave panel that repeatedly observes the same individuals over time. The panel contains information of household and person socioeconomic characteristics and data from two-day travel diaries collected in 1989 (wave l), 1990 (wave 2), 1992 (wave 3), 1993 (wave 4), and 1994 (wave 5). Because the fifth wave was not available at the writing of this report, only the first four waves are used here. This includes 929 households, or 1,622 persons that participated in all four waves and had complete household, person, and 2-day activity and travel information. Although the data collection instrument is a travel diary, activity information can be derived from it. Activity types are the trip purposes provided by the respondents. The original nine activity types are aggregated into subsistence, maintenance, and leisure activities according to decreasing degrees of constraints and importance. This includes out-of-home or at-home activities before the evening return home. Similarly, the initial 16 travel modes are grouped into car, carpool, public transportation, and non-motorized/other modes. Adopting the same methodology defined in a previous report (Ma and Goulias, 1995), cluster analysis is issued to obtain activity and travel patterns.

Cluster analysis is a statistical technique used to search for homogeneous groups of observations (persons in this study) with respect to some variables of interest. Searching is based on distance (dissimilarity) or similarity in the values of a set of variables considered simultaneously. Distance measures how far apart two observations are, whereas similarity measures the closeness of two observations. The clustering algorithm, called nearest centroid sorting, is used here and can be applied to both situations with or without known initial cluster centers (i.e., representative values around which all observations in a given cluster are distributed). When the initial cluster centers are unknown (i.e., no information on characteristics of activity or travel groups is available before conducting cluster analysis, which is the case in the PSTP data set), we need to select the number of clusters externally. For a k-cluster solution, this algorithm first assigns the first k observations in the data set as temporary centers. As the combining process is going on, if the smallest distance from the subsequent observation to a center is greater than the distance between any two closest centers, this observation replaces a center that is the closest to it; otherwise, it is assigned into that cluster. The distance here is the commonly used Euclidian distance. It is very sensitive to the unit of the clustering variables. To avoid the effect of different variable sizes, all clustering variables are standardized into the form of the z-score, which is the ratio of the difference between the variable and its mean over its standard deviation.

Based on exploratory analyses using all the data in the four PSTP waves and the analysis in wave 1 and wave 2 (Ma and Goulias, EM), observations are classified into four groups (or clusters) for both activity and travel. Variables used in forming activity patterns are total amount of daily activity time (duration of out-of-home activities or for short at-home periods in a day, but not after returning home in the evening) and frequencies of subsistence, maintenance, and leisure activities. Average length of trip chains; total travel time; and frequencies of trips made by car, carpool, public transportation, and non-motorized/other modes are the variables used to obtain travel patterns. Figure l(a) shows a four-cluster solution of activity patterns. The first four variables (Sfreq, Mfreq, Lfreq, and Atime) on the x-axis are the variables used in cluster analysis. The last three (Sdur, Mdur, and Ldur) are provided for reference. The y-axis represents the absolute mean values of the variables reported on the x-axis. Frequencies are measured in number of events, whereas durations are measured in hours. These groups are named worker-A, shopper, worker-B, and inactive. The worker-A group not only has very high frequency and duration in subsistence activities, but also long total daily activity duration. People in worker-A and worker-B groups spend similar amounts of time in subsistence activities. The worker-B group, however, participates in fewer subsistence activities than the worker-A group. Both worker groups engage in activities longer than the other two groups. Except for the large differences in frequency of subsistence activities, the two worker groups behave similarly in maintenance and leisure activities. In contrast to the worker groups, the shopper group has the highest frequencies and longest durations in both maintenance and leisure activities. The inactive group, as the name suggests, has the lowest activity frequencies and total activity durations among the four pattern groups. The number of observations in each cluster is shown in figure 2(a). The worker-A and worker-B groups make up about 42 percent of the sample, of which only one sixth of the sample belongs to the worker-A group. In the inactive group, there are an equal number of people (42 percent) as in the two worker groups. The shoppers are 17 percent of the sample.

As with the activity patterns, a four-cluster solution is obtained for travel, as shown in figure l(b). The travel time variable is measured in hours, and the remaining variables are measured in number of events. These four groups are labeled non-motorized, car/carpool, immobile, and public groups. The non-motorized group is highly likely to make trips by non-motorized/other modes. This group is very efficient in trip chaining, as suggested by the second highest number of stops (or trips) per trip chain. People in the car/carpool group most likely travel by car and carpool, but less likely by public transportation and non-motorized/other modes. They have the highest total travel time among all four groups. The immobile group, on the other hand, is associated with people who have the shortest travel



Sfreq, Mfreq, Lfreq = frequencies of subsistence, maintenance, leisure activities

Sdur, Mdur, Ldur = durations of subsistence, maintenance, leisure activities

Atime = total out-of-home activity duration

Car-Freq, Cpl-Freq, Pub-Freq, Non-Freq = trip frequencies by car, carpool, public transportation, and non-motorized/other modes

Stops/C = number of stops per trip chain

TTime = total travel time

Figure 1. Activity and travel patterns.

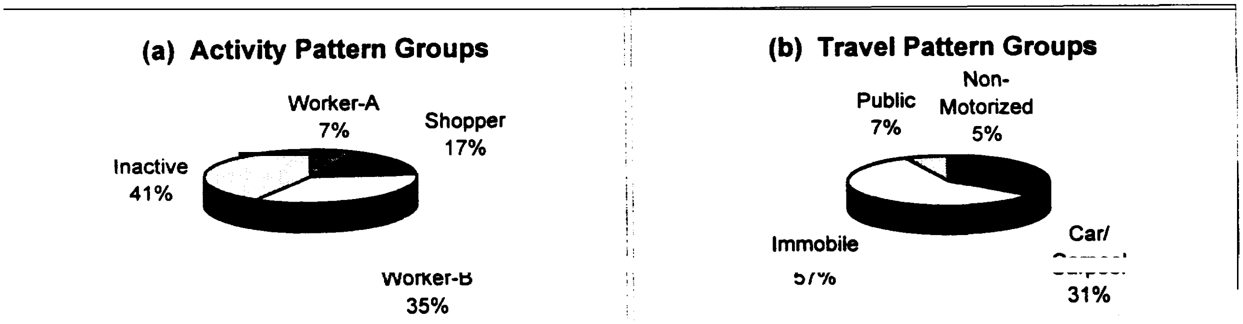


Figure 2. Compositions of activity and travel pattern groups.

time. This group also has very low trip frequencies by all the travel modes. The public group is characterized by, of course, the highest trip frequency by public transportation, but by the lowest number of trips by car and carpool. The group's total travel time is relatively high. Figure 2(b) shows the composition of each travel pattern group. More than half of the observations belong to the immobile group, while slightly less than one third are classified into the car/car-pool group. The remaining two groups account for 12 percent of the total observations, with a slightly higher percentage in the public group.

PERSON CHARACTERISTICS AND PATTERN GROUPS

The relationship between activity and travel patterns is examined in this section using the simplest form of multilevel models. This allows temporal effects on activity patterns and travel patterns to be examined while other influential factors in the data are being controlled The models explain micro-level outcomes in two ways. First, parameters of micro-level covariates in multilevel models are expressed as a function of context as macro-level variables. Second, this micro/macro relationship can be interpreted by the characteristics of the context. The definition of context is quite general and can include spatial references (counties, places), social groups (economic sectors, households, peer groups), and temporal contexts (different time points).

A distinction should be made here between theoretical multilevel analysis and pragmatic-level definitions for modeling purposes. Ideally, one should define levels based on micro-unit interactions within the level. For example, within a household individuals may subdivide tasks, and their activity participation is an expression of such a labor subdivision (e.g., grocery shopping). Likewise, a neighborhood functions as a social unit and interactions among the residents leads to specific activity and travel patterns (e.g., neighborhood association meetings). In addition, the availability and level of service offered by the activity and transportation system may inhibit or motivate activity and travel. Changes in all these macro-circumstances, surrounding an individual's activity and travel behavior, determine the change in activity and travel patterns. However, it is very difficult, if not impossible, to observe all the determinants of activity and travel behavior and to measure all the changes taking place. A more pragmatic approach, in which proxies to the macro-level effects for all contexts, is used in this report. In addition, instead of explicitly modeling transitions of context, a temporal context is adopted attempting to capture unmeasured changes in micro-level variables, changes in macro-level variables, systematic temporal trends in patterns, and noise.

A description of this idea is the four-level model structure shown in figure 3. The first macro level is the temporal context. Such a time effect is expressed in two different scales (which are analogous to the analysis in the previous section), year-to-year and day-to-day capturing long-term and short-term time effects, respectively. In this way the day and the wave effects can be examined in more detail. The second macrolevel is called the spatial effect. Ideally, such an effect should have been derived from spatial characteristics at the level of census tracks, block groups, blocks, traffic analysis zones, or even household dwelling unit location (currently there is an ongoing effort to supplement PSTP with these contextual variables by geocoding the information available). Because this information is unavailable, the spatial context is characterized by residential location and employment relocation status. Household residential location is measured at the county level, whereas the employment status is evaluated by changes in workplaces during the previous year. The household is the last macrolevel. Variables at the household level include income, household type and size, auto ownership, etc. These variables attempt to capture household interaction and related constraints in resources (e.g., income and cars). The forth level, a micro level, is made up of individuals' characteristics. Variables at this level contain specific individual effects such as age, education level, and so on.

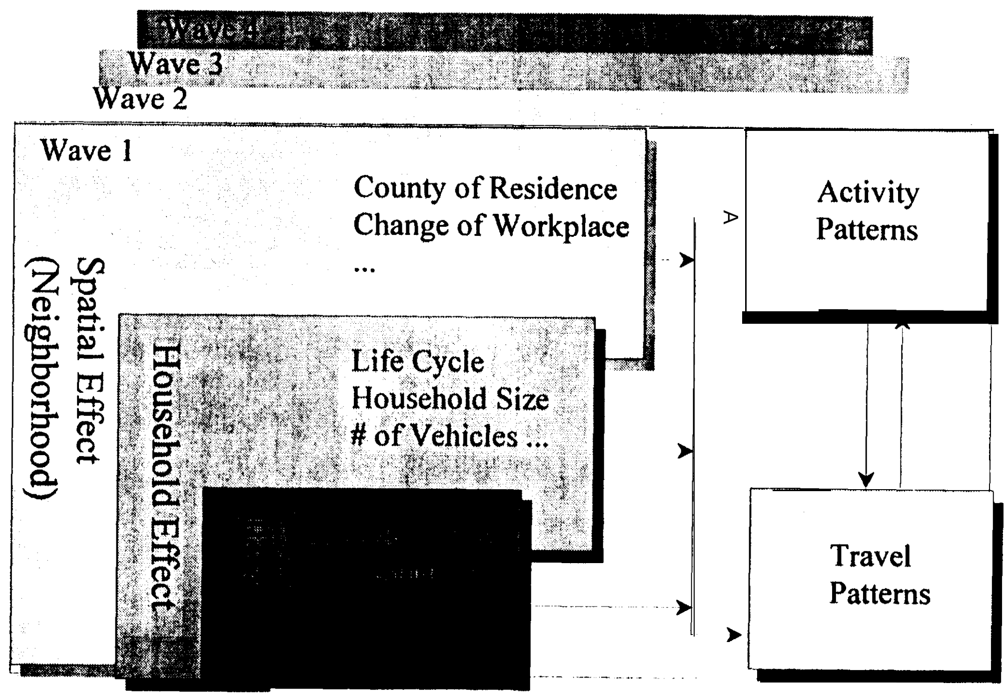


Figure 3. Multilevel analysis.

ACTIVITY MULTILEVEL MODELS

A multinomial logit model[1] that incorporates the four-level contexts is developed to describe the activity pattern membership (daily pattern choice). The multilevel model structure is implemented in the usual utility[2] function capturing the effects of time, space, household, and person levels. This model structure attempts to identify the linkage between the pattern groups and the individuals' characteristics augmented by household, neighborhood (spatial), and temporal effects. A model of activity patterns that includes all these effects is shown in table 1 (left-hand side) and is called the base model. The worker-A group is the reference group. Gender and driver's license holding are the variables capturing person effects in the model (the models here are a few examples of a large set of models estimated in a preliminary model specification exercise using a variety of person-based variables). Males seem less likely to be in the shopper and inactive groups. Household so&economics contribute greatly to determining individuals' activity patterns in various ways. People from large households are more likely to be in the shopper group. However, they are less likely to be in the inactive group if their household incomes are high. The number of employed people, an indirect measure of household disposable income, is insignificant for the inactive group but influential for the shopper and worker-B groups. Individuals from households with more employed members are less likely to be in the shopper group but more likely to belong to the worker-B group. It should be noted that income and the number of employed members are not significant in the same group. This may be due to the correlation between these variables. Households without a car are not likely to be in the shopper group. The household lifecycle variables (M35 and S35-65) show that people from young-adult-only and mid-aged-single households are less likely to be in the shopper and inactive groups. At the spatial contextual

---

'Given the preliminary nature of the study here, simple model forms are used to explore the strength of some hypotheses. Multilevel models can be formulated using a strong theoretical basis, which in turn may dictate the use of complex multi-tiered models for which estimation algorithms are not yet available.

[2]At best, the function used here can be called an indirect utility. The term utility is used in this report in a loose sense, and the logit model was not derived by considering the machinery of rational choice behavior. The models should be considered as nonlinear multivariate regressions that are used to make sense of what happens in the sample. Logit was chosen for convenience.

Table 1. Multinomial logit models of activity patterns.

	Excluding Travel Patterns (Base Model)						Including Travel Patterns					
	Shopper		Worker-B		Inactive		Shopper		Worker-B		Inactive	
	Coef.	t-test	Coef.	t-test	Coef.	t-test	Coef.	t-test	Coef.	t-test	Coef.	t-test
Constant	0.9760	4.39	1.9755	9.65	1.9768	9.54	-0.6456	-2.21	1.9513	8.18	1.3764	5.38
MALE	-0.8366	-9.80	-0.1317	-1.75	-0.3149	-4.14	-1.0590	-11.78	-0.2826	-3.52	-0.4848	-5.63
HHSIZE	0.2306	5.72	0.0248	0.67	-0.0460	-1.23	0.2405	5.65	-0.0133	-0.34	-0.1159	-2.72
HIGHINC	-0.0185	-0.19	-0.1330	-1.59	-0.2112	-2.47	-0.0201	-0.20	-0.1391	-1.56	-0.2260	-2.32
NUM_EMP	-0.1747	-3.00	0.1363	2.63	-0.0280	-0.53	-0.1961	-3.15	0.1216	2.16	-0.0384	-0.64
CAR0	-0.7686	-2.00	-0.5194	-1.50	0.1064	0.33	0.2681	0.66	-0.1952	-0.55	0.9207	2.69
M35	-0.9166	-3.32	-0.4155	-2.20	-0.3701	-1.89	-0.8263	-2.90	-0.3721	-1.84	-0.3201	-1.42
S35_65	-0.4824	-2.35	-0.0654	-0.41	-0.4206	-2.58	-0.4295	-2.00	-0.0202	-0.12	-0.3006	-1.64
DIST2B	0.1781	2.05	0.2427	3.12	0.1584	2.01	0.2478	2.70	0.2573	3.11	0.1812	2.05
KITSAP	-0.3654	-3.13	-0.6017	-5.78	-0.3036	-2.95	-0.2558	-2.02	-0.6225	-5.49	-0.3394	-2.83
YRHM1	0.1113	0.67	0.2985	2.07	0.1651	1.11	0.1902	1.10	0.3269	2.15	0.1874	1.14
YRHM20	0.3065	2.48	0.1005	0.87	0.3205	2.79	0.2030	1.55	-0.0319	-0.26	0.1354	1.06
YRCT1	-0.3976	-1.46	-0.5891	-2.56	-0.3394	-1.45	-0.3218	-1.13	-0.5902	-2.40	-0.3682	-1.38
WKPLACE	-0.0027	-16.27	-0.0008	-4.82	-0.0026	-16.08	-0.0028	-16.36	-0.0006	-3.76	-0.0023	-13.64
Day2Wave1	0.0355	0.18	0.0875	0.47	0.1531	0.80	0.0461	0.22	0.1195	0.62	0.1970	0.97
Day1Wave2	-1.0551	-6.05	-0.9486	-5.95	-0.7966	-4.89	-1.0776	-5.93	-1.0216	-6.12	-0.9111	-5.11
Day2Wave2	-1.0587	-5.97	-0.8136	-5.04	-0.6759	-4.10	-1.1320	-6.11	-0.8898	-5.27	-0.7855	-4.36
Day1Wave3	-0.6489	-3.59	-0.6997	-4.20	-0.2631	-1.56	-0.7562	-4.02	-0.8702	-4.99	-0.5328	-2.89
Day2Wave3	-0.7818	-4.34	-0.7483	-4.53	-0.3048	-1.82	-0.8531	-4.55	-0.8834	-5.11	-0.5295	-2.90
Day1Wave4	-0.9592	-5.32	-0.7649	-4.65	-0.4552	-2.72	-1.0510	-5.59	-0.9306	-5.40	-0.7333	-4.00
Day2Wave4	-1.2169	-6.80	-0.8936	-5.53	-0.5914	-3.60	-1.2982	-6.94	-1.0749	-6.33	-0.9005	-4.98
Immobile							0.5048	2.28	-0.9022	-5.91	-0.9817	-5.51
Car/Carpool							1.8390	9.91	-0.3803	-3.22	-1.0363	-7.40
Non-motorized							3.2942	13.08	3.0526	15.00	4.2684	20.19

Excluding Travel Patterns: $-2(\ln L(c) - \ln L(\beta)) = 3296$; Degrees of freedom = 60

Including Travel Patterns: $-2(\ln L(c) - \ln L(\beta)) = 8308$; Degrees of freedom = 69

Definitions of Variables

MALE	1 = Male
HHSIZE	Number of people in the household
HIGHINC	1 = Income > $50,000
NUM_EMP	Number of employed people
CAR0	1 = No car
M35	1 = Young multi-adult household
S35_65	1 = Mid-aged-single household
DIST2B	1 = Within two blocks to the nearest bus stop
KITSAP	1 = Living in Kitsap county
YRHM1	1 = Years in the current home < 1
YRHM20	1 = Years in the current home > 20
WKPLACE	1 = Changed workplace in the previous year
Day2Wave1	1 = The observation is from day 2 of wave 1

level, people who reside in Kitsap county or who changed workplaces during the previous year are more likely to be worker-A group members.

Regarding the temporal effects there is a systematic trend. In the model, day 1 of wave 1 is used as reference, and day 2 of wave 1 (1 day after the reference point), seems not to effect activity patterns (i.e., the distribution of activity patterns in day 1 is the same as in day 2 when we control for all other factors). Similar results are obtained for all four waves (the magnitudes of coefficients and t-tests of any 2 days in the same wave are very close). This may imply that the short-term time effects are not significant contributors in explaining activity pattern group membership. The long-term time effects tell a different story (e.g., compare the coefficients of Day-2 Wave-l to Day-2 Wave-2). This has been confirmed by stationary day-to-day and nonstationary year-to-year pattern switching in a parallel analysis (Goulias and Ma, 1995). The model also shows that all the variables have plausible signs and that the results do not contradict earlier studies. T-tests also suggest that the variables in the model are significant at a 90 percent confidence level. However, not every variable is significant in all the pattern groups, which may imply that each pattern group is governed by a different underlying selection mechanism and obviously selected by a different so&demographic population segment. The goodness-of-fit x^2 is 3,296 with 60 degrees of freedom, suggesting the choice of these explanatory variables is satisfactory.

To test the effects of day-of-week on the activity patterns, three additional model specifications are examined The first model specfication (not shown in this report) includes the same variables as in the base model plus three dummy variables representing the first 3 days of the week (i.e., Monday, Tuesday, and Wednesday). The estimated coefficients for all the variables have the same sign and are very close to those presented in table 1. Tuesday was the only significant variable in all three travel patterns. The x^2 increases to 3,320 with 69 degrees of freedom. Inclusion of these nine additional terms resulted in an increase in x^2 by 24, implying the significant effect of the day-of-week. The second model specification (shown in table 1) adds three variables to the base model representing three travel pattern groups (non-motorized, car/carpool, and immobile). This increases x^2 dramatically with an accompanying downward shift in the constant terms. Most of the variables preserve their sign, however, their magnitudes and t-statistics are different. The model shows that people who are in the non-motorized and car/carpool groups are likely to be in the shopper group but less likely to be in the worker-B and inactive groups. In addition, people in the immobile group are more likely to be in all the travel groups, except the worker-A group. The last model specification, combining both the day-of-week and the travel patterns with the base model, exhibits only a marginal increase in x^2 (not shown in table 1). This may be surprising and contradictory to the conventional belief that people follow weekly activity scheduling exercises. The results here may be due to an artifact emerging from the data collection scheme. In PSTP, an individual's behavior is recorded on the same days of the week in all waves. The day-of-week variable then captures differences among individuals and the days of a week. To isolate the effect of day-of-week, then, we would have needed a 7-day activity/travel diary.

TRAVEL MULTILEVEL MODELS

A parallel exercise to the activity patterns is done for travel pattern groups as well. The base model results are displayed in table 2 (left-hand side). At the micro level, gender and having a valid driver's license are determinants of the travel pattern group. At the household macro-level, household size, number of employed people, income category, car ownership, and household type are significantly influencing travel pattern group membership. Household perceptions, such as rating of living conditions of the county of residence and level of traffic congestion, are also influential. At the neighborhood macro-level, time lived in the current home, the county of residence, and workplace relocation are the significant variables in determining individuals' travel pattern groups. The degrees of their contributions, however, are different. For instance, living in the current home less than 1 year or more than 20 years significantly affects the car/car-pool group, whereas there is no noticeable impact on the public group. The temporal effects, unlike the activity models, do not exhibit any particular pattern and, except for the immobile group, are not significantly different from zero in most cases. The base model reveals that individuals who have valid driver's licenses, come from large households, are single seniors, and have lived in the current home for more than 20 years are likely to be in the car/car-pool and immobile groups. Individuals who changed workplaces in the previous year are also more likely to be in the car/car-pool group. The public group is

Table 2. Multinomial logit models of travel patterns.

	Excluding Activity Patterns (Base Model)						Including Activity Patterns					
	Car/Carpool		Immobile		Public		Car/Carpool		Immobile		Public	
	Coef.	t-test	Coef.	t-test	Coef.	t-test	Coef.	t-test	Coef.	t-test	Coef.	t-test
Constant	2.1653	7.52	2.5195	9.01	-0.0447	-0.13	1.6099	5.28	3.5052	11.66	0.2449	0.67
MALE	0.0942	1.03	0.2678	3.03	0.2222	2.02	0.2308	2.48	0.2893	3.11	0.1803	1.61
LICENSE	0.0019	5.52	0.0008	3.22	0.0003	1.04	0.0018	5.07	0.0011	3.97	0.0004	1.40
HHSIZE	0.1534	3.30	0.0912	2.01	0.0876	1.53	0.0955	2.03	0.1762	3.73	0.1354	2.36
HIGHINC	-0.0436	-0.40	-0.1184	-1.13	-0.2908	-2.17	-0.0515	-0.47	-0.0270	-0.25	-0.2539	-1.88
NUM_EMP	-0.1302	-2.09	-0.0629	-1.05	0.1036	1.37	-0.0860	-1.37	-0.0560	-0.90	0.0832	1.10
CAR0	-3.1136	-6.62	-1.6641	-6.38	1.2083	4.72	-3.0265	-6.34	-1.9756	-7.24	1.1034	4.26
S35_65	-0.5506	-3.02	-0.6053	-3.48	0.3046	1.48	-0.5616	-3.05	-0.5192	-2.84	0.3239	1.57
S65	0.7337	2.49	0.5569	1.99	0.3555	1.07	0.7083	2.37	0.5603	1.94	0.3848	1.15
LIVEGOOD	-0.2492	-1.30	-0.3534	-1.90	-0.0495	-0.21	-0.2254	-1.17	-0.3129	-1.63	-0.0342	-0.15
NONCONG	-0.3981	-3.74	-0.3209	-3.15	-0.6697	-4.91	-0.3977	-3.67	-0.3748	-3.49	-0.6861	-4.97
DIST2B	-0.3800	-4.14	-0.2959	-3.33	0.1248	1.11	-0.4116	-4.43	-0.3429	-3.70	0.1046	0.93
KITSAP	-0.3805	-2.79	-0.2677	-2.06	0.5293	3.38	-0.3755	-2.70	-0.1315	-0.95	0.6394	4.01
KING	-0.1578	-1.62	-0.2462	-2.61	0.0493	0.41	-0.1504	-1.53	-0.2177	-2.22	0.0600	0.50
YRHM1	-0.2381	-1.76	-0.1541	-1.19	-0.1201	-0.75	-0.2236	-1.62	-0.1608	-1.17	-0.1335	-0.82
YRHM20	0.5081	3.79	0.6095	4.69	0.2155	1.31	0.5016	3.69	0.5959	4.42	0.2431	1.47
WKPLACE	0.0003	2.84	-0.0002	-1.96	0.0006	4.10	0.0007	5.45	0.0002	1.30	0.0005	3.36
Day2Wave1	0.3374	1.99	0.3098	1.87	0.2321	1.15	0.3687	2.15	0.2592	1.51	0.2154	1.06
Day1Wave2	0.2949	1.74	0.3283	1.99	0.2807	1.40	0.3992	2.31	0.4685	2.71	0.3339	1.64
Day2Wave2	0.2577	1.52	0.3392	2.06	0.2005	0.99	0.3761	2.18	0.4231	2.46	0.2305	1.13
Day1Wave3	0.1159	0.69	0.3979	2.45	-0.2106	-1.02	0.1808	1.06	0.4375	2.58	-0.1748	-0.83
Day2Wave3	0.0790	0.48	0.3442	2.14	-0.1706	-0.84	0.1594	0.94	0.3529	2.10	-0.1364	-0.66
Day1Wave4	0.3220	1.84	0.5917	3.48	-0.0223	-0.10	0.4210	2.36	0.6712	3.78	0.0160	0.07
Day2Wave4	0.0824	0.49	0.4141	2.56	-0.2940	-1.40	0.2035	1.19	0.5232	3.07	-0.2542	-1.20
Worker-A							-0.0056	-0.04	-5.2747	-23.97	-1.0240	-5.74
Shopper							1.3533	9.42	-2.4507	-17.56	-1.5236	-7.62
Worker-B							0.5451	3.96	-1.3075	-10.27	-0.1077	-0.71

Excluding Activity Patterns: $2(\ln L(c) - \ln L(\beta)) = 1172$, Degrees of freedom = 69

Including Activity Patterns: $-2(\ln L(c) - \ln L(\beta)) = 6127$, Degrees of freedom = 78

Definitions of Variables

Variable	Definition	Variable	Definition	Variable	Definition
MALE	1 = Male	M35	1 = Young multi-adult household	KING	1 = Living in King county
LICENSE	1 = Have a valid driver license	S35_65	1 = Mid-aged-single household	YRHM1	1 = Years in the current home < 1
HHSIZE	Number of people in the household	LIVEGOOD	1 = Living condition is good	YRHM20	1 = Years in the current home > 20
HIGHINC	1 = Income > $50,000	NONCONG	1 = No congestion	WKPLACE	1 = Changed workplace
NUM_EMP	Number of employed people	DIST2B	1 = Two blocks to the nearest bus stop	Day2Wave1	1 = The obs. is from day 2 of wave 1
CAR0	1 = No cars	KITSAP	1 = Living in Kitsap county		

likely to include people who changed their workplace in the previous year and have no car. Males tend to fall into the immobile and public groups more often than their counterparts females.

Three additional model specifications are defined to examine the effect of day-of-week and the influence of activity patterns. Unlike activity patterns, day-of-week does not contribute to explaining travel pattern groups, as indicated by a seven-unit increase in x^2 when three dummy variables representing day-of-week are added (this causes a decrease in the degrees of freedom by nine). However, when activity patterns (expressed by three dummy variables) are included as the explanatory variables, the x^2 increases dramatically. The result is provided in table 2 (right-hand side). In addition, the signs and magnitudes of the explanatory variables change (the gender effect, for example, where male individuals are more likely to be in the car/carpool and immobile groups). As for the public group, gender does not play an important role. The car/carpool group tends to include people who are in the shopper and worker-B groups. The immobile group is less likely to be associated with the worker-A, shopper, and worker-B groups. People in the worker-A and shopper groups are less likely to belong to the public group.

3. ACTIVITY, TRAVEL, AND ACCESSIBILITY

In the 1970's and early 1980's there was a substantial amount of research examining the influence of accessibility on trip making (Leake and Huzayyin, 1979; Leake and Huzayyin, 1980; Richardson and Young, 1982). Accessibility has been variously interpreted as being the "nearness to places," the "nearness to activities," or "ease of participating in activities" (Richardson and Young, 1982). Early use of accessibility indicators provides some evidence that individuals with different levels of accessibility would exhibit distinctly different travel patterns. The reason for this is that individuals with high levels of accessibility can reach many places at relatively low cost and, thus, make more trips than people with similar socio-demographic characteristics but lower levels of access (Hanson and Schwab, 1987). This significant effect of accessibility on trip making also motivates efforts on how to measure accessibility (Vickerman, 1974; Pirie, 1979; Smith, 1980; Bach, 1981; Hanson, 1982; Weibull, 1982; Hanson and Schwab, 1987). Among a wide spectrum of accessibility measures, integral accessibility, which is an account for several possible destinations, is the most appropriate measure to assess easiness of travel for a given location (e.g., the residence of a household). Such accessibility can be measured by the ease to reach a place and/or the number of (weighted) opportunities (or activities) that can be reached from a given location. The former is a function of the travel cost to the place from all other places, whereas the latter considers a sort of unit cost for each opportunity. All these efforts emphasize destinations only, which are presumably motivated by attempts to distribute trips within a certain area or areas.

As the emphasis on trip making shifts from analyzing single trips to individuals' activity participation, there is a need to examine the effect of accessibility on activity engagement. In fact, as early as 1975 researchers noticed that accessibility is a key concept for characterizing a fundamental principle of human activity: maximum contacts through minimum activity (Karlqvist, 1975). However, there has not been much research in this area since then. Activity participation involves not only locations where an individual is pursuing the activity, but also his or her residence. For instance, consider two persons who have similar socio-demographic characteristics but live in two different places. One person is in a place that is close to many activity locations and has a congested transportation network, while the other lives in a low-density area without any traffic congestion. The two individuals may have quite different activity participation patterns. The first one may be sensitive to the network performance and, thus, select some activities or travel modes in order to avoid traffic congestion. The second person may not be constrained by any of these concerns. On one hand, this implies that accessibility is as important at residential locations as it is at destinations, although they may require different ways to evaluate the accessibility. On the other hand, accessibility at the same location may be perceived differently by different people. For example, a person who lives in an area may measure ease of access in a place quite differently from a person who works in the area. Again, this calls for different accessibility measures for the same location.

In spite of the complexity in measuring accessibility, the type of available information in this project adds another dimension of complexity in deriving accessibility measures. The data obtained from the PSRC include a variety of socio-demographic characteristics of the traffic analysis zones within which each household resides (e.g., population and household counts and densities, employment, and housing structure units). They also contain peak and off-peak highway and transit skims, which can further be broken down into more detailed measures such as car and carpool travel time and distance, total transit time and in-vehicle time for auto access and walk access, and boarding time and number of transfers for transit auto access and walk access, etc. How to compute accessibility and land-use measures that are suitable for activity participation and travel analyses using such a large amount of data has been one of the concerns in this project. The example offered here uses factor analysis to summarize land use/socio-demographics and access offered by the transportation system around the dwelling unit of a given household. In addition, splitting the sample into people for whom we can and whom we cannot derive workplace accessibility made it possible to analyze workplace accessibility and activity/travel patterns.

FACTOR ANALYSIS

Factor analysis is a statistical technique that represents a set of variables in terms of a smaller number of derived variables (Kim and Mueller, 1986). It derives relationships that exist among. a set of observed variables by uncovering common dimensions or factors that link seemingly unrelated variables. Consequently, it provides insights into the underlying structure of the data. The basic assumption of factor analysis is that observed correlations between variables result from shared factors and, thus, these factors can be used to explain complex phenomena. The goal of factor analysis is to identify, based on a set of observed variables, the factors that are not directly observable. This is in agreement with the concept of accessibility. For instance, short total transit auto access time and short total waiting time are linked by a good accessibility indicator, although accessibility is not directly observable.

In factor analysis, a factor is a qualitative dimension or a coordinate axis. It defines the way in which entities differ and provides a dimensional structure for the data by indicating the important common qualities present in the data. In general, for a data set with p observed variables the mathematical model for the ith standardized variable can be written as:

$$X_i = A_{i1}F_1 + A_{i2}F_2 + \dots + A_{ik}F_k + U_i, \quad i=1, \dots, p \tag{1}$$

where:

X_i = **standardized observed variable (i.e., total car and carpool time and distance in highway skim)**

A_i = **coefficients used to combine the k factors called factor loadings**

F_i = **unobservable variables called common factors (i.e., the accessibility measures in this project)**

U_i = **an unobservable variable called a unique factor, which is equivalent to the unobserved variables that are related only to the observed variable X_i**

k = **the number of common factors**

The proportion of variance explained by the common factors is called the commonality of the variable. Therefore, the closer the commonality is to 1, the better the variable is represented by the common factors.

Because the common factors are unobservable, we need to determine the number of factors and the corresponding factor loadings. The jth factor, F_j, is expressed as a linear combination of a series of the observed variables:

$$F_j = \sum_{i=1}^{p} W_{ji} X_i \qquad (2)$$

where:

W_{ji} = factor score coefficients

W's can be estimated from principal components analysis or the maximum likelihood method. The factor scores give the projection of an observation on the common factors. They provide insights into the structure of the data by highlighting patterns of common variation and, thus, provide information on quantitative differences.

A factor can be obtained using the following four steps: variable identification, factor extraction, rotation, and score estimation (SPSS/PC+, 1992). The software first identifies variables that are not related to other variables from the correlation matrix for all variables and associated statistics. Then, it decides the number of factors necessary to represent the data using an appropriate method. An assessment of the model fit is also performed at this stage. Next, it transforms the factors to make them more interpretable by rotating and changing the correlations among the common factors (i.e., rotating the axis of the factors). Scores for each factor can then be computed for each observation, which can be used in a variety of other analyses.

RESULTS OF FACTOR ANALYSIS

Using the factor analysis technique, a set of factors are obtained for the socio-demographic characteristics, peak and off-peak transit skims, and peak and off-peak highway skims at the census tract level. The observed variables that are used to derive the corresponding factor(s) are provided in table 3. It should be noted that the transit and highway skims are aggregated from

Table 3. Variables used in factor analysis.

Type of Info	Name	Variable
Socio-demo info	DEMPL	No. of employed persons in civilian & armed forces per square mile
	DNOWRK	No. of persons not in labor force per square mile
	DUNEMP	No. of unemployed persons per square mile
	HDEN90	Household density in 1990
	INC	Median household income in 1990
	DMFHS	No. of multi-family housing structures per square mile
	PDEN90	Population density in 1990
	DOTHER	No. of housing structures other than single & multi-family per square mile
Highway (a.m.)	NCPLT	Non-carpool time
	CPLD	Carpool distance
	CPLT	Carpool time
Highway (daily)	NCPLT	Non-carpool time
	NCPLD	Non-carpool distance
	CPLT	Carpool time
Transit auto access (a.m. and daily)	TOTALT	Total transit time
	INVEH	In-vehicle time
	AUXI	Auxiliary transit (drive) time
	WAITT	Total wait time
	WAITB	Wait for bus time
	BOARD	Boarding time
	TRANSF	Number of transfers
Transit walk access (a.m. and daily)	TOTALT	Total transit time
	INVEH	In-vehicle time
	WAITT	Total wait time
	WAITB	Wait for bus time
	BOARD	Boarding time
	TRANSF	Number of transfers

traffic analysis zones (TAZs) to census tracts because the trip origins and destinations in the travel diaries are coded at the census tract level. The aggregate skims are then scaled by dividing each of them by the area of the corresponding census tract.

The summary statistics of factor analyses are provided in table 4. The Kaiser-Meyer-Olkin (KMO) measure of sample adequacy is an index for comparing the magnitudes of the observed correlation coefficients to the magnitudes of the partial correlation coefficients. Small values for KMO indicate that a factor analysis of the variables may not be a good idea. The KMO values in table 4 show that factor analysis is meritorious for peak transit walk access skim, but mediocre for peak and off-peak highway skims and off-peak transit auto access skim. The ranges of commonalities (minimum and maximum) for highway and transit skims suggest that there is a strong linear association among the variables for each skim. However, there might not be a strong linear relationship among the variables in the socio-demographic characteristics due to the low minimum communality. Nevertheless, the high percentage of explained variance for each set of variables shows that the extracted factors are able to capture the essential characteristics of the observed variables.

The factor score coefficients for socio-demographics and transit and highway skims are shown in tables 5, 6, and 7, respectively. For socio-demographics, two factors are obtained, each being a linear function of the observed variables:

SOCIOFl = 0.16629xDEMPL + 0.15157xDNOWRK + 0.13384xDOTHER + 0.14468xDUNEMP + 0.17176xHDEN90 + 0.00207xINC + 0.16871xPDEN90 + 0.15919xDMFHS

SOCIOF2 = 0.08702xDEMPL - 0.05294xDNOWRK + 0.01874xDOTHER - 0.112885xDUNEMP + 0.04729GIDEN90 + 0.974940X + 0.03257xPDEN90 - 0.0548xDMFHS

It is clear that SOCIOFl captures the effects of all the variables except INC, which is primarily captured by SOCIOF2 because the coefficient of INC for SOCIOFl is small but very large for SOCIOF2. A low value (which can be negative) in SOCIOFl means low employment,

Table 4. Summary statistics.

Statistics	Socio-Demo	Highway		Transit Auto Access		Transit Walk Access	
		A.M.	Daily	A.M.	Daily	A.M.	Daily
KMO	0.74	0.65	0.65	0.77	0.66	0.84	0.79
Minimum communality	0.59250	0.98588	0.99285	0.86154	0.92901	0.80534	0.43982
Maximum communality	0.99167	0.99706	0.99844	0.99890	0.99983	0.99592	0.93979
# of factors	2	1	1	2	2	1	1
% of variance explained	85.2	99.3	99.6	97.2	98.4	95.4	82.2
Correlation between factors	-0.28830	N/A	N/A	-0.15028	-0.14123	N/A	N/A

Table 5. Factor score coefficients of socio-demographic variables.

Factor	DEMPL	DNOWRK	DOTHER	DUNEMP	HDEN90	INC	PDEN90	DMFHS
SOCIOF1	0.16629	0.15157	0.13384	0.14468	0.17176	0.00207	0.16871	0.15919
SOCIOF2	0.08702	-0.05294	0.01874	-0.12885	0.04729	0.97494	0.03257	-0.0548

Table 6. Factor score coefficients of transit skims.

Skims	Time	Factor	TOTALT	INVEH	AUXI	WAITT	WAITB	BOARDT	TRANSF
Transit Auto Access	A.M.	AMAUF1	0.16682	0.18129	-0.10908	0.17889	0.15275	0.1772	0.1772
		AMAUF2	0.03861	-0.06192	1.0064	-0.04333	0.05063	-0.04362	-0.0436
	Daily	MDAUF1	0.16987	0.17603	-0.10363	0.17250	0.17014	0.16835	0.16835
		MDAUF2	0.00106	-0.04472	1.01312	-0.01876	-0.04597	0.00209	0.00209
Transit Walk Access	A.M.	AMWAF1	0.17444	0.17383	N/A	0.17417	0.15686	0.17206	0.17206
	Daily	MDWAF1	0.18690	0.19503	N/A	0.19654	0.13445	0.19114	0.19113

Table 7. Factor score coefficients of highway skims.

Time	Factor	NCPLT	CPLD	CPLT
A.M.	AMHYF1	0.32523	0.33334	0.335
Daily	MDHYF1	0.33427	0.33334	0.33419

low unemployment, and low residential population densities. SOCIOF2 is related to the median **income level.** Likewise, one factor is necessary to describe the effect of all the aspects of peak and off-peak highway skims and transit walk access skims, respectively. Low values in AMAUF1, MDAUD1, AMWAF1, and MDWAFl are associated with short travel times, while low values in AMHYFl and MDHYFl correspond to short travel times or short travel distances However, an additional factor is needed to depict the transit auto access skims. This additional factor captures the effect of auxiliary (drive) transit auto access time. Low values in AMAUF2 and MDAUF2 mean short driving times to park-and-ride lots. The reason for this may be that the auxiliary transit time has quite different characteristics than the other transit-related measures. For instance, a person may not equally weigh the travel time of driving his or her own car to a park-and-ride lot to the travel time of riding a bus.

ACTIVITY AND TRAVEL PATTERN SELECTION

As in chapter 1, individuals' activity and travel pattern choices can be described by the four-level multinomial logit models. These four levels are temporal, spacial, household, and person level effects. However, the spatial effect is enhanced by adding the accessibility measures. Since individuals' activity and travel patterns are affected by both residence and workplace, workers' and nonworkers' pattern selections are examined separately. For workers, accessibility measures at both residence and workplace are included as explanatory variables. For nonworkers, only the accessibility measures at residential locations are examined. Variables used in the models are described in table 8.

The model estimates for workers' and nonworkers' activity selections are provided in tables 9 and 10, respectively. For workers, the temporal indicators (i.e., day 2 in wave 1 and each of 2 days in each of the other three waves) have similar effects as when the accessibility measures are excluded. Specifically, the effect of day 2 in wave 1 is not significant, while the effects of the 2 days within the same wave are very similar, as evidenced by similar coefficients of 2-day temporal indicators for the last three waves. This may be due to the similar day effects in the same wave. Because the coefficient of day 1 in wave 1 is the reference, which is scaled down to zero, the coefficient of day 2 in wave 1 becomes trivial. As for the spatial effects, the

Table 8. Descriptions of variables in the models.

Variable	Descriptions
MALE	Indicator, 1 = male; 0 = female
LICENSE	Indicator, 1 = have a valid driver license; 0 = otherwise
WKMIS	Indicator, 1 = missing workplace; 0 = otherwise
NONWKOTH	Indicator, 1 = inconsistent info for nonworkers; 0 = otherwise
WORKCOM	Indicator, 1 = not work at home; 0 = otherwise
HHSIZE	Number of household members
TOT1_17	Number of persons who are younger than 17 (including 17)
TOTADULT	Number of persons who are older than 17 (excluding 17)
HIGHINC	Indicator, 1 = household income greater than $50,000; 0 = otherwise
NUM_EMP	Number of employed people in the household
CAR0	Indicator, 1 = no car; 0 = otherwise
NUMVEH	Number of vehicles in the household
S35_65	Indicator, 1 = mid-age single household, 0 = otherwise
S65	Indicator, 1 = senior single household; 0 = otherwise
M35	Indicator, 1 = young couple household; 0 = otherwise
LIVEGOOD	Indicator, 1 = living condition is good; 0 = otherwise
NONCONG	Indicator, 1 = live in a noncongested area; 0 = otherwise
DIST2B	Indicator, 1 = reside with two blocks to the nearest bus stop; 0 = otherwise
KITSAP	Indicator, 1 = live in Kitsap count, 0 = otherwise
KING	Indicator, 1 = l ive in King county; 0 = otherwise
YRHM1	Indicator, 1 = lived in the current home less than 1 year; 0 = otherwise
YRHM20	Indicator, 1 = lived in the current home more than 20 years; 0 = otherwise
YRCT1	Indicator, 1 = lived in the current county less than 1 year; 0 = otherwise
WKPLACE	Indicator, 1 = changed workplace; 0 = otherwise
P_R	Number of park & ride lots in the residence census tract
Day2Wave1	Indicator, 1 = Day2 in Wave 1; 0 = otherwise
Day1Wave2	Indicator, 1 = Day1 in Wave 2; 0 = otherwise
Day2Wave2	Indicator, 1 = Day2 in Wave 2; 0 = otherwise
Day1Wave3	Indicator, 1 = Day1 in Wave 3; 0 = otherwise
Day2Wave3	Indicator, 1 = Day2 in Wave 3; 0 = otherwise
Day1Wave4	Indicator, 1 = Day1 in Wave 4; 0 = otherwise
Day2Wave4	Indicator, 1 = Day2 in Wave 4; 0 = otherwise

NOTE—SOCIOF1, SOCIOF2, AMAUF1, AMAUF2, MDAUF1, MDAUF2, AMWAF1, MDWAF1, AMHYF1, MDHYF1 are the same as previously described. In the models, they are the accessibility measures at the residence census tract. When the prefix "K" is added, they refer to the accessibility measures of the workplace census tract.

Table 9. Model estimates of workers' activity patterns.

	Shopper		Worker-B		Inactive	
	Coef.	t-test	Coef.	t-test	Coef.	t-test
Constant	1.2633	3.61	1.1318	3.52	2.0756	6.81
Male	-1.1944	-11.67	-0.1444	-1.81	-0.1485	-1.80
WORKCOM	-0.4730	-1.89	1.1081	4.43	-0.1819	-0.82
NUM_EMP	-0.1265	-1.74	0.1525	2.66	0.0693	1.18
S35_65	-0.3796	-1.63	-0.0960	-0.57	-0.2950	-1.70
m35	-0.8300	-2.77	-0.4316	-2.18	-0.2897	-1.41
HHSIZE	0.2629	5.59	-0.0102	-0.26	-0.0813	-1.99
HIGHINC	0.0759	0.68	-0.1182	-1.34	-0.1984	-2.16
CAR0	-0.7169	-1.51	-0.2325	-0.69	0.1685	0.51
DIST2B	0.2061	1.88	0.2088	2.39	0.2282	2.52
KITSAP	-0.4843	-2.63	-0.6741	-4.66	-0.4885	-3.29
YRHM1	0.0289	0.16	0.2320	1.56	0.1539	0.99
YRHM20	-0.0091	-0.06	-0.1942	-1.57	0.1610	1.29
YRCT1	-0.3430	-1.12	-0.5468	-2.33	-0.4908	-2.00
WKPLACE	-0.1330	-0.94	-0.2315	-2.14	-0.1802	-1.60
SOCIOF1	0.0568	0.46	-0.1886	-2.40	-0.1755	-2.02
AMAUF1	1.1375	3.01	0.2261	0.96	0.1502	0.59
AMAUF2	0.4733	2.64	0.2756	1.91	0.2680	1.79
MDAUF1	-1.6735	-3.66	-0.2775	-1.05	-0.2569	-0.88
MDAUF2	-0.0818	-0.48	-0.3450	-2.20	-0.2665	-1.65
AMWAF1	0.0267	0.91	0.0054	0.23	0.0571	2.22
MDWAF1	0.0097	0.34	-0.0147	-0.64	-0.0543	-2.15
AMHYF1	-5.9734	-2.42	-1.1931	-0.60	-2.3157	-1.13
MDHYF1	5.5243	2.26	1.3639	0.69	2.3073	1.13
KSOCIOF1	0.1128	2.34	0.0212	0.51	-0.0295	-0.68
KAMAUF1	-0.3594	-2.34	-0.0868	-0.72	-0.1050	-0.86
KAMAUF2	0.0949	0.72	-0.1288	-1.19	-0.1714	-1.56
KMDAUF1	0.1992	2.14	-0.0283	-0.37	0.0604	0.75
KMDAUF2	0.0261	0.18	0.3163	2.54	0.3583	2.81
KAMWAF1	-0.0798	-1.38	-0.1237	-2.40	-0.1239	-2.39
KMDWAF1	0.1224	2.23	0.1202	2.46	0.1383	2.80
KAMHYF1	0.1403	0.09	-3.4329	-2.59	-2.2556	-1.66
KMDHYF1	-0.0457	-0.03	3.5197	2.58	2.3108	1.65
P_R	0.0565	0.72	-0.1147	-1.81	-0.0518	-0.79
KP_R	0.0798	1.11	0.0402	0.68	-0.0024	-0.04
Day2Wave1	0.0698	0.31	0.1141	0.58	0.1296	0.63
Day1Wave2	-1.0551	-5.25	-0.9621	-5.70	-0.8133	-4.62
Day2Wave2	-0.9498	-4.66	-0.7932	-4.63	-0.6757	-3.78
Day1Wave3	-0.8401	-4.05	-0.6988	-3.98	-0.2607	-1.45
Day2Wave3	-0.8186	-3.98	-0.7525	-4.31	-0.3272	-1.82
Day1Wave4	-1.1511	-5.41	-0.7486	-4.28	-0.5785	-3.18
Day2Wave4	-1.4473	-6.79	-0.9300	-5.44	-0.7510	-4.22

$-2(\ln L(c) - \ln L(\beta)) = 912$

Degrees of freedom = 123

Table 10. Model estimates of nonworkers' activity patterns.

	Shopper		Worker-B		Inactive	
	Coef.	t-test	Coef.	t-test	Coef.	t-test
Constant	2.3322	2.52	1.8772	2.03	4.5277	4.99
Male	-0.7725	-2.59	-0.3800	-1.26	-0.7371	-2.51
LICENSE	0.6839	1.20	-0.0616	-0.11	-0.7596	-1.38
NONWKOTH	-2.3708	-4.77	-1.0354	-2.11	-2.6407	-5.51
WKMIS	-2.9482	-7.60	-1.0319	-2.73	-2.7068	-7.33
NUM_EMP	-0.4897	-2.95	-0.3627	-2.15	-0.4950	-3.03
S35_65	-0.8583	-1.31	-0.8019	-1.20	-0.9444	-1.53
m35	-1.4708	-1.56	-0.8210	-0.99	-1.3728	-1.64
HHSIZE	0.8117	4.13	0.6039	3.05	0.5204	2.66
CAR0	1.6424	1.30	0.8540	0.66	1.7384	1.41
NUMVEH	-0.0605	-0.33	0.0167	0.09	-0.0095	-0.05
DIST2B	0.0247	0.07	0.0980	0.29	0.1737	0.53
KITSAP	-0.6113	-1.27	-1.0676	-2.15	-0.5873	-1.24
YRHM1	2.0866	1.72	2.1600	1.78	2.1180	1.75
YRHM20	1.0719	2.44	1.2832	2.89	1.0988	2.52
YRCT1	8.8741	0.06	8.2013	0.05	9.2331	0.06
SOCIOF1	0.3268	1.20	0.2366	0.89	0.3375	1.28
SOCIOF2	-0.0929	-0.49	-0.0525	-0.27	-0.0741	-0.40
AMAUF1	-0.2283	-0.24	-0.1449	-0.15	-0.7846	-0.85
AMAUF2	-0.7711	-0.99	-0.8186	-1.05	-0.9047	-1.17
MDAUF1	-0.1530	-0.20	-0.0889	-0.12	0.3779	0.52
MDAUF2	1.3810	1.14	1.2014	0.98	1.5125	1.25
AMWAF1	0.0428	0.33	-0.0889	-0.69	-0.0461	-0.36
MDWAF1	-0.0584	-0.44	0.1276	0.99	0.0829	0.65
AMHYF1	15.6310	1.88	14.5660	1.74	14.1520	1.71
MDHYF1	-15.7380	-1.94	-14.1340	-1.73	-14.0000	-1.74
P_R	0.8827	1.95	0.6926	1.52	0.8588	1.90
Day2Wave1	-0.0778	-0.13	-0.1368	-0.22	0.1092	0.18
Day1Wave2	-0.6958	-1.23	-0.4720	-0.82	-0.3857	-0.69
Day2Wave2	-1.1697	-2.15	-0.7357	-1.33	-0.6499	-1.21
Day1Wave3	0.3674	0.49	0.2950	0.39	0.5671	0.76
Day2Wave3	-0.1676	-0.24	0.0971	0.14	0.3178	0.46
Day1Wave4	-0.3180	-0.51	-0.2005	-0.32	0.2610	0.42
Day2Wave4	-0.0481	-0.07	0.3032	0.45	0.6542	1.00

$2(\ln L(c) - \ln L(\beta)) = 724$

Degrees of freedom = 99

accessibility measures have different influences on the pattern selection. Both residence and workplace accessibility measures are significant, although not all of them are significant determinants for all the activity patterns. The shopper group is likely to reside in census tracts characterized by bad peak transit auto access and off-peak highway accessibility measures, but good off-peak transit auto access and peak highway accessibility measures. The measure of socio-demographics and the accessibility measures in terms of off-peak transit auto access and highways adversely affects the choice of shopper pattern. For the worker-B group, accessibility measures at the workplace have a much stronger influence than those at residential locations (only SOCIOFl and MDAUF2 at residential locations are significant). For the inactive group, measures for transit walk access at residence and workplace contribute to explaining the pattern choice. However, the effects of these two measures are opposite, as indicated by the opposite signs of the coefficients. The other types of spatial factors include indicators of county of residence, the number of years in the county and in the current residence, and whether or not workplaces changed in the current year. If a person has been living in the current county for less than a year or changed his workplace, he is less likely to fall into either the worker-B group or the inactive group. Household socio-demographics and individual characteristics have various impacts on individuals' activity patterns. However, their influence is very similar as to when the accessibility measures are excluded.

The model estimates show that nonworkers have quite different activity pattern selection mechanisms than workers. First, almost all the temporal indicators are not significant. This may indicate that people who do not work at a workplace outside their residence may not be sensitive to the day effect because they are not restricted by any subsistence activities and they are likely to follow the same pattern group from year to year. Second, all the socio-demographic factors and accessibility measures for transit skims are not significant. However, the accessibility measures of highways are relatively significant in explaining activity pattern selection. A possible explanation may be that nonworkers are more dependent on SOV than workers.

Similar to the activity models, two multinomial logit models for workers' and nonworkers' travel patterns are estimated separately. The results are shown in tables 11 and 12. For workers, unlike the activity pattern selection, their travel pattern selection appears to be time homogenous because most temporal indicators are trivial. However, this may be misleading. It should be kept in mind that the accessibility measures are derived based on the information in 1990. If a person

Table 11. Model estimates of workers' travel patterns.

	Car/Carpool		Immobile		Public	
	Coef.	t-test	Coef.	t-test	Coef.	t-test
Constant	0.4352	0.75	1.3226	2.51	0.1774	0.28
Male	0.0288	0.25	0.3185	2.82	0.1173	0.86
WORKCOM	0.4762	1.45	0.5696	1.77	0.0120	0.03
LICENSE	2.2332	5.95	1.1074	3.67	-0.4673	-1.45
NUM_EMP	-0.2164	-2.27	-0.1514	-1.63	-0.1816	-1.61
S35_65	-0.7546	-3.29	-0.5651	-2.57	0.4150	1.65
S65	1.8877	1.72	1.7151	1.56	1.3801	1.18
TOT1_17	0.0656	1.01	-0.0154	-0.24	-0.0382	-0.48
TOTADULT	-0.2073	-1.64	-0.0572	-0.47	0.4281	2.97
HIGHINC	-0.1025	-0.78	-0.1731	-1.35	-0.3449	-2.16
CAR0	-1.0161	-2.03	-0.5592	-1.36	0.9431	2.27
NUMVEH	0.1406	2.17	0.1081	1.70	-0.1611	-2.00
DIST2B	0.0982	0.79	0.1316	1.08	0.5357	3.67
TRAFGOOD	-0.3874	-2.74	-0.3419	-2.49	-0.5202	-2.96
KITSAP	-1.8685	-9.09	-1.4985	-7.62	-0.1361	-0.54
YRHM20	0.6344	2.88	0.7161	3.31	0.3748	1.52
WKPLACE	0.5429	2.68	0.4773	2.40	0.2043	0.84
SOCIOF1	-0.3478	-3.41	-0.4394	-4.37	-0.7541	-4.40
SOCIOF2	0.1991	2.50	0.1721	2.21	0.1539	1.68
AMAUF1	-1.5859	-4.48	-1.1459	-3.28	-0.6109	-1.28
AMAUF2	0.2786	1.53	0.3171	1.80	0.7812	3.41
MDAUF1	1.4647	3.50	1.0679	2.59	0.6204	1.14
MDAUF2	-0.3589	-2.11	-0.4740	-2.80	-0.9880	-3.56
AMWAF1	0.1576	3.06	0.1758	3.44	0.2364	3.84
MDWAF1	-0.0565	-1.40	-0.0788	-1.96	-0.0771	-1.43
AMHYF1	-9.0162	-3.67	-7.3106	-3.14	1.0461	0.39
MDHYF1	9.4870	3.88	7.6491	3.29	-0.7334	-0.27
KSOCIOF1	0.0334	0.67	-0.0200	-0.40	-0.0585	-0.97
KSOCIOF2	0.5851	8.12	0.4485	6.36	-0.1368	-1.57
KAMAUF1	0.1851	0.95	0.0978	0.51	-0.2801	-0.85
KAMAUF2	0.5140	3.21	0.3712	2.38	0.3758	1.72
KMDAUF1	-0.1060	-1.09	-0.0510	-0.54	-0.0205	-0.19
KMDAUF2	-0.4903	-2.84	-0.2717	-1.62	-0.0218	-0.10
KAMWAF1	0.1895	2.18	0.1652	1.93	0.6253	3.64
KMDWAF1	-0.0844	-1.44	-0.0534	-0.93	-0.7939	-7.65
KAMHYF1	0.8932	0.46	-0.2513	-0.13	-7.5700	-3.28
KMDHYF1	-1.3107	-0.66	-0.1671	-0.09	7.1940	3.09
P_R	-0.1013	-1.13	-0.1977	-2.24	0.0623	0.59
KP_R	0.0962	1.00	0.0897	0.95	-0.1239	-0.97
Day2Wave1	0.3056	1.40	0.2005	0.94	0.2154	0.87
Day1Wave2	0.1551	0.72	0.0758	0.36	0.1001	0.40
Day2Wave2	0.1772	0.81	0.1914	0.89	0.1296	0.52
Day1Wave3	-0.0745	-0.35	0.2194	1.06	-0.3537	-1.41
Day2Wave3	-0.0629	-0.30	0.1454	0.71	-0.4977	-1.98
Day1Wave4	0.0431	0.19	0.2307	1.06	-0.0823	-0.32
Day2Wave4	-0.1621	-0.75	0.0869	0.41	-0.3942	-1.53

$-2(\ln L(c) - \ln L(\beta)) = 1872$

Degrees of freedom = 135

Table 12. Model estimates of nonworkers' travel patterns.

	Car/Carpool		Immobile		Public	
	Coef.	t-test	Coef.	t-test	Coef.	t-test
Constant	-1.115	2.52	0.9573	1.79	-2.3674	-2.99
Male	-0.1051	-2.59	-0.2188	-1.26	0.5110	2.10
LICENSE	2.0043	1.20	0.7946	3.82	-0.8222	-2.89
NONWKOTH	0.4092	-4.77	-0.0850	-0.18	-0.8656	-1.20
WKMIS	-0.9322	-7.60	-0.8984	-3.28	1.3232	3.90
NUM_EMP	-0.0063	-2.95	-0.0321	-0.27	0.0593	0.35
S35_65	0.4375	-1.31	-0.4243	-0.92	-0.7628	-1.23
S65	0.7107	-1.56	0.3617	1.09	0.5387	1.19
HHSIZE	0.0292	4.13	-0.0563	-0.60	-0.0866	-0.64
HIGHINC	0.1769	1.30	-0.0141	-0.05	-0.3264	-0.79
CAR0	-1.3253	-0.33	-0.6173	-1.64	1.2225	2.53
NUMVEH	0.5294	0.07	0.4796	3.96	0.4021	2.48
DIST2B	-0.3554	-1.27	-0.2720	-1.51	-0.1909	-0.74
TRAFGOOD	-0.4478	-2.26	-0.3295	-0.66	-0.6909	1.72
LIVEGOOD	-0.2633	-0.76	-0.2176	-1.77	0.8702	-2.49
KITSAP	0.9470	2.58	0.9346	2.65	1.9309	4.25
KING	0.5093	2.01	0.3564	1.48	1.0228	3.02
YRHM1	-0.2900	-0.85	0.2970	0.96	0.1335	0.30
YRHM20	0.4272	2.10	0.5017	2.58	0.1460	0.50
SOCIOF1	-0.2171	-0.99	-0.0753	-0.85	-0.0764	-0.50
SOCIOF2	-0.0305	-0.20	-0.0292	-0.26	-0.0656	-0.40
AMAUF1	-0.6239	1.14	-0.6150	-1.87	0.7507	1.04
AMAUF2	0.0255	0.33	0.0237	0.08	0.1825	0.46
MDAUF1	0.3179	-0.44	0.2496	1.13	-0.7871	-1.23
MDAUF2	0.1611	1.88	0.1824	0.51	0.0617	0.14
AMWAF1	0.1924	-1.94	0.1337	1.28	-0.3536	-0.66
MDWAF1	-0.1148	1.95	-0.0224	-0.26	0.1066	1.00
AMHYF1	-3.7621	-0.13	-4.1369	-1.50	-6.5278	-1.43
MDHYF1	3.3774	-1.23	3.8524	1.38	6.4330	1.41
P_R	-0.2807	-2.15	-0.1366	-0.94	-0.1636	-0.79
Day2Wave1	0.4897	0.49	0.6083	2.14	0.1135	0.28
Day1Wave2	0.4764	-0.24	0.7906	2.64	0.8963	2.24
Day2Wave2	0.3270	-0.51	0.6033	2.12	0.5075	1.28
Day1Wave3	0.2839	-0.07	0.6111	2.03	-0.3882	-0.79
Day2Wave3	0.0563	0.32	0.5691	1.89	0.6069	1.44
Day1Wave4	0.2484	0.32	0.8518	2.84	0.4720	1.06
Day2Wave4	-0.1016	0.30	0.5768	2.05	0.3459	0.82

$-2(\ln L(c) - \ln L(\beta)) = 882$

Degrees of freedom = 108

did not change residential location and workplace, his or her accessibility measures would remain the same. When a person changed residence or workplace from one census tract to another, the accessibility measure would change to reflect changes in accessibility, accordingly. Because the majority of the sample did not change residence and workplace, a good portion of the variables did not change across waves. This may also lead to stable temporal indicators. Therefore, there is not sufficient evidence to definitely conclude whether or not the travel pattern is time insensitive. As for the effects of the accessibility measures, the travel pattern selection is more dependent on various measures of accessibility than the activity pattern selection. This is presumably due to the fact that the derived accessibility measures are primarily based on various highway and transit skims, which directly reflect the travel conditions. It is worth noting that both measures of socio-demographics at residential locations and workplace are influential factors in determining the travel patterns. In other words, SOCIOFZ, the measure of median income at census tract, affects an individual's travel pattern but not his activity pattern.

Nonworkers' travel patterns are different from workers' travel patterns. First, nonworkers' travel pattern selection is not time homogenous because temporal indicators are significant. However, there is no obvious pattern in terms of how time affects pattern selection. Second, the accessibility measures are less important for nonworkers than workers because most of the accessibility variables are not significant.

4. DYNAMIC ANALYSIS

In this section, we examine day-to-day and year-to-year sequences of transitions among the four activity and the four travel pattern groups. This is done by examining the transitions of the same individual's activity or travel patterns during a given time frame (e.g., within wave 1 from day 1 to day 2). To differentiate between the two time scales, we call the yearly based time frame the long-term and the daily based time frame the short-term. Therefore, we can define four short-term periods (day 1 to day 2 in each of the four waves) and three long-term periods (wave 1 to wave 2, wave 2 to wave 3, and wave 3 to wave 4) for each of the 2 days. It should be noted, however, that the time span between wave 2 and wave 3 is longer (2 years) than those of the other consecutive waves (1 year). In the second part of this section, contingency table techniques are used to measure changes at two points in time for activity and travel pattern groups obtained with cluster analysis. The analysis here mirrors the two-wave analysis in Ma and Goulias, 1995.

SHORT-TERM DYNAMICS

Day-to-day transitions of activity pattern groups are obtained by cross tabulating the activity pattern memberships of the 2-day observations for each wave. Figure 4(a) shows the percentage of persons that maintained the same pattern group from day 1 to day 2 in each of the four waves (called the retention rate herein). The x-axis represents the pattern groups and the y-axis represents the short-term retention rate. First, we can see that short-term transitions for all the four waves present a very similar pattern with exception of the worker-A group in wave 1. The inactive group has the highest retention rate among the four groups, suggesting that people with this activity pattern are less likely to change their pattern from one day to the next. In addition, a systematic increase, in terms of the percentage of people who have the same pattern over time, is observed as the retention rate changes from 59 percent in wave 1 to 67 percent in wave 4. Both worker-A and worker-B groups have lower retention rates than that of the inactive group. People in the shopper group change their activity patterns most often among the four groups. In contrast to the inactive group, a gradual decrease for the shopper group is observed.

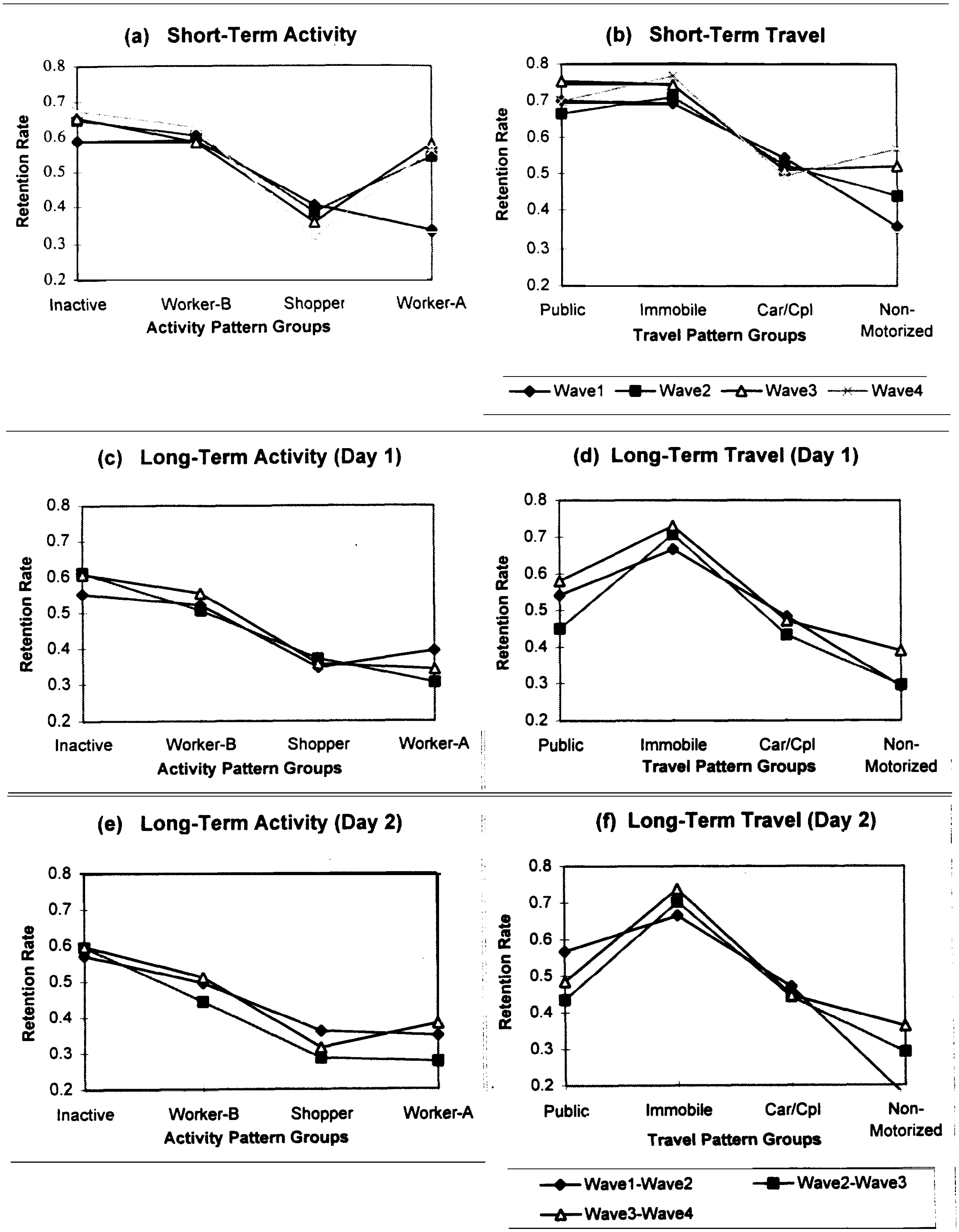


Figure 4. Short-term and long-term transitions of activity and travel pattern groups.

However, the other two pattern groups do not exhibit such a clear trend. Overall, approximately 57 percent of the people did not change patterns between the 2 days.

With the same approach, daily transitions of travel pattern groups are obtained (see figure 4[b]). Except for the public group, which does not follow any particular trend, the other three groups display a steady decrease or increase in the percentage of people who maintained the same travel pattern from wave 1 to wave 4. As indicated by a higher percentage of pattern turnover rates, ranging from 43 percent to 66 percent, public and immobile groups are likely to maintain the same patterns over time. Although many people in the car/car-pool group change their travel patterns, the average retention rate decreases slightly. In contrast, the non-motorized group shows a considerable increase in the retention rates from wave to wave. The dramatic change in retention rates, starting at 35 percent in wave 1 and ending at 57 percent in wave 4, suggests there are significant variations in individuals' travel behavior. It is also possible that such high transition rates are due to the small number of observations in this group (i.e., even if a few people switch their travel patterns from or to this group, it may result in a high percentage). The average retention rate of travel pattern groups is 65 percent.

From figures 4(a) and 4(b), we have seen that the activity pattern groups have a lower overall retention rate than the travel pattern groups. This implies that more people modify their activity patterns than their travel patterns in 2 days. In addition, people with different activity or travel pattern groups behave in different ways. For instance, compared to the other two groups, a higher percentage of people maintained the same patterns (and did so consistently over the 4-year period) in the inactive and shopper groups.

LONG-TERM DYNAMICS

It has been widely agreed that individuals' activity and travel behavior is influenced by various factors such as household socio-economics, transportation network characteristics, and other unobserved and unobservable factors. Changes in some of these factors may take place only within longer time frames (e.g., getting married and having children). Systematic changes in activity participation and travel behavior over longer time frames may be discovered by analyzing long-term dynamics (from year to year here) using cross-tabulation of two time points that are at least 1 year apart. Figures 3(c) and 3(e) provide the wave-to-wave retention rates for day 1 and day 2 of activity pattern groups, respectively. The total retention rate is 37 percent. Some notable differences in retention rates are observed from wave 2 to wave 3, especially on both days of the worker-A group and day 1 of the shopper and worker-B groups. This may be due to the longer time span between waves 2 and 3, which is 2 years. There are more changes during long-term transitions than during short-term transitions. For every activity pattern group, the corresponding retention rates are lower, ranging from 28 percent to 61 percent. These lower values show that more people changed their activity patterns from one year to the next. However, the marginal frequency (at each time point) of the pattern group membership is relatively homogeneous over time. This finding motivates a companion report that analyzes activity and travel transitions using models that are specifically designed for this purpose (Goulias and Ma, 1995).

The long-term transitions of travel pattern groups are provided in figures 4(d) and 4(f) for day 1 and day 2, respectively. Generally, there are more long-term than short-term transitions. The long-term retention rates are between 29 percent and 74 percent, while their short-term counterparts are between 35 percent and 75 percent. For the public group, there is a significant drop in the transition from wave 2 to wave 3 on both days. Again this may be the result of a longer time span between waves 2 and 3. However, such significant changes are not observed in other pattern groups. For the non-motorized group, a dramatic change from wave 1 to wave 2 on day 2 can also be seen. The total retention rate is 45 percent. As with short-term transitions, more people switch from activity group to activity group than from travel group to travel group.

Higher variations in activity patterns may be a direct result of individuals' weekly scheduling. Within a wave, PSTP data contain travel information from just 2 consecutive days and we are unable to clearly observe potential cyclic behavior (with a cycle shorter than 1 year) in activity scheduling. A higher variation in the year-to-year transitions confirms that changes requiring longer time frames (e.g., residential and/or workplace relocation, income increases or decreases, and lifestyle/lifecycle stage transitions) trigger more changes in activity and travel behavior than short-term changes related to individual and household scheduling of activities. This is reenforced when considering that PSTP was designed so that the same 2 days of the week are maintained for each individual. This point is taken one step further using multivariate models in this report.

RELATIONSHIP BETWEEN ACTIVITY AND TRAVEL PATTERNS

In this section, contingency table techniques are used to examine the “temporal interplay” between activity and travel patterns. This was measured by the contingency coefficient, the square root of the chi-square from cross-classification divided by the sum of the chi-square and the number of observations in the sample. The contingency coefficient is similar to the correlation coefficient. The closer the contingency coefficient is to 1, the stronger the relationship between the two variables is.

An earlier study on the same topic with data from the first two waves (Ma and Goulias, 1995) has shown a stronger relationship between activity and travel pattern groups in the same day than between different days. The same finding is obtained here with the time frame expanded to 5 years. Figure 5 shows a surface plot of the contingency coefficients for eight different time points, 2 days in each of the four waves. The peak of the surface, denoted by the diagonal area running from the left lower comer to the upper right comer, represents the contingency coefficients of travel and activity patterns within the same day. These highest values, ranging from 0.51 to 0.53, are a measure of the strong relationship between activity and travel within the same day. The off diagonal areas that are next to the peak of the surface, i.e., the middle left upper and middle right lower areas, represent the coefficients of different days within the same wave. They are consistently lower than the same-day contingency coefficients, ranging from 0.24 to 0.32, and higher than the remaining areas representing the coefficients among different days and waves. The smallest contingency coefficients correspond to activity and travel that are distant by the 2-year time span between waves 2 and 3 and by longer time spans (the least shaded areas in figure 5).

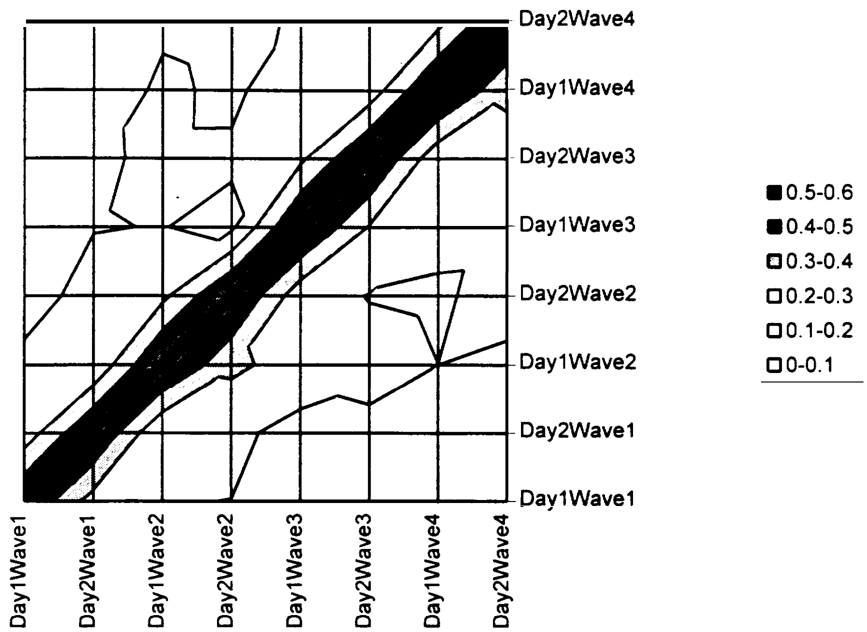


Figure 5. Contingency coefficients of activity and travel pattern groups.

5. SUMMARY AND CONCLUSIONS

In this report, the relationship between activity engagement and trip making behavior is explored by using cluster analysis to aggregate individuals into manageable and relatively homogeneous behavioral groups. This yields four activity groups and four travel groups based on representative variables. A four-level model specification based on contextual analysis provides results that indicate possible avenues for enhanced models of activity and travel. This analysis is then expanded using residence and workplace-related accessibility and socioeconomic factors. The analysis clearly shows that the level of service offered by the transportation system and land-use indicators significantly affects activity scheduling and trip making by individuals.

Cross-classification techniques are used to study the activity and travel pattern changes from one day to another and from one year to the next in the first four waves of PSTP. Studies show higher variations in day-to-day activity than in travel. Substantial year-to-year changes in activity and travel are also observed, with travel exhibiting more regularity than activity (confirming past research on habitual behavior Buff and Hanson, 19901). Changes in activity and travel occur so that the marginal frequencies remain time homogeneous. The link between activity and travel is strong within a day, but it gets weaker as the time span increases, thus, supporting the existence of short-term schedules.

Models of activity behavior cannot be specified without considering time explicitly. Indeed, the study of activity participation is about time allocation by individuals, and this was clearly reflected in this analysis. In addition, time needs to be considered in time scales such as short-term versus long-term. This separates components of behavioral variation that are due to day-to-day, week-to-week, month-to-month, and year-to-year scheduling and/or habits from structural effects (due to changes in lifestyle and activity/transportation system changes). Also related to this is the effect of day-of-the-week. In this study we were unable to examine this effect satisfactorily because of the data collection scheme used (the same individual was observed on the same days of the week in each panel wave).

Regarding model specification, the models here support the inclusion of four-level explanatory variables based on temporal, spatial (neighborhood, city, region), household, and person contexts to explain choice of daily activity and travel patterns. An attempt to include other contexts, especially for activity participation, such as peer contexts (e.g., employer and work-related groups, social groups, etc.) was successful using proxies such as the socio-demographic tract-level factors. Current data collection schemes exclude questions related to this type of information, in spite of their potential for enhanced activity participation model specification.

As for the reciprocal influence of activity and travel, different model specifications reveal that there is a dual directional influence between activity and travel pattern groups. This supports the conventional belief that the desire to participate in activity determines the demand for travel and that travel, in turn, enables activity engagement by releasing and imposing constraints in a complex fashion. Modeling this relationship requires considering activity participation and travel behavior simultaneously and explicitly accounting for the circumstances under which travel enables activity participation. This is a research area that has not been looked into carefully yet. From a policy analysis perspective, the study here shows one way to account for level of service in activity and travel models (i.e., the effects of accessibility on activity participation and travel).

The last implication of this study regards the unit of analysis for modeling purposes. In the past, most travel demand models used the household as the decision making unit, masking the influence of the household on an individual's behavior. Here we have used the person as the decision making unit and have included the household as a context within which the person acts. Person-based models operate at the most detailed level and can include all the other contexts within which a person operates as explanatory variables. Then, the effect of contextual changes (e.g., household type transitions) can be studied in a more explicit way to determine if these changes equally influence each household member (this is left as a future task) and to study how people interact over time.

The results presented in this report, however, should be considered preliminary for many reasons. The data used are from the 1,622 persons who participated in all four waves of PSTP, which is about 47.8 percent of the first wave participants. Sample selection in PSTP took place in a nonrandom way, and the conclusions drawn in this report are not representative of the entire Puget Sound population. Methods to account for this recently have been developed for the panel and are provided in Ma and Goulias, 1996. Extensions, beyond the Puget Sound, of the results in this report would also require considering other surveys that have been specifically designed for activity analysis.

6. REFERENCES

Bach, L. (198 1). "The Problem of Aggregation and Distance for Analyses of Accessibility and Access Opportunity in Location-Allocation Models," ***Environment and Planning A,*** vol. 13, pp. 955-978.

DiPrete, T. A., and J. D. Forristal (1994). "Multilevel Models: Methods and Substance." ***Annual Review Sociology,*** vol. 20, pp. 33 l-357.

Ettema, D., A. Borgers, and H. Timmermans. "A Competing Risk-Hazard Model of Activity Choice, Timing, Sequencing, and Duration." Preprint No. 951118. Report presented at the 74th Annual Meeting of the Transportation Research Board, Washington D.C., January 22-28, 1995.

Goulias, K. G., and J. Ma (July 1995). "Panel Analysis of Activity and Travel Patterns Using Generalized Markovian Models for Transitions." Working Paper.

Hamed, M. M., and F. L. Mannering (1993). "Modeling Travelers' Post-Work Activity Involvement: Toward a New Methodology." ***Transportation Science,*** vol. 27, no. 4, pp. 38 l-394.

Hanson, S. (1982). "The Determinants of Daily Travel-Activity Patterns: Relative Location and Socio-demographic Factors." ***Urban Geography,*** vol. 3, no. 3, pp. 179-202.

Hanson, S., and M. Schwab (1987). "Accessibility and Intraurban Travel." ***Environment and Planning A,*** vol. 19, pp. 73 5-745.

Huff, J., and S. Hanson (1990). "Measurement of Habitual Behaviour: Examining Systematic Variability in Repetitive Travel." ***In Developments in Dynamic and Activity-Based Approaches to Travel Analysis,*** edited by P. Jones. United Kingdom: Aldershot.

Karlqvist, A. (1975). "Some Theoretical Aspects of Accessibility-Based Location Models." In ***Dynamic Allocation*** of ***Urban*** Space, edited by A. Karlqvist, L. Lundqvist, and F. Snickars. Lexington: D. C. Heath.

Leake, G. R., and A. S. Huzayyin (December 1979). "Accessibility Measures and Their Suitability for Use in Trip Generation Models." ***Traffic Engineering & Control,*** pp. 566-572.

Leake, G. R., and A. S. Huzayyin (1980). "Importance of Accessibility Measures in Trip Production Models." ***Transportation Planning and Technology,*** vol. 6, pp. 9-20.

Ma, J., and K. G. Goulias. "A Dynamic Analysis of Activity and Travel Patterns Using Data from Puget Sound Transportation Panel." Preprint No. 950331. Paper presented at the 74th Annual Meeting of the Transportation Research Board, Washington, D.C., January 22-28, 1995.

Ma, J., and K. G. Goulias. "Sample Weights for the Puget Sound Transportation Panel Using Stratification Anchors in the Public Use Microdata Sample and Probabilistic Models for Self-Selection." Preprint No. 960562. Paper presented at the 75th Annual Meeting of the Transportation Research Board, Washington, D.C., January 7- 11, 1996.

McNally, M. G., and R. Wang. "Generation of Dynamic Trip Tables via an Activity-Based Microsimulation Model." Presented at the INFORMS Conference, Los Angeles, CA, 1994.

Murakami, E., and T. W. Watterson (1990). "Developing a Household Travel Panel Survey for the Puget Sound Region." ***Transportation Research Record,*** 1285, pp. 40-46.

Norusis, M. J. (1992). SPSS/PC+ Advanced Statistics, Version 5.0. SPSS Inc., pp. 121-136.

Pas, E. I. (1982). "Analytically Derived Classifications of Daily Travel-Activity Behavior: Description, Evaluation, and Interpretation." ***Transportation Research Record,*** 897, pp. 9- 15.

Pas, E. I. (1983). "A Flexible and Integrated Methodology for Analytical Classification of Daily Travel-Activity Behavior." ***Transportation Science,*** vol. 17, no. 4, pp. 405-429.

Pas, E. I. (1988). "Weekly Travel-Activity Behavior." ***Transportation,*** vol. 15, nos. 1-2, pp. 89-109.

Pas, E. I., and F. S. Koppelman (1987). "An Examination of the Determinants of Day-to-Day Variability in Individuals' Urban Travel Behavior." ***Transportation,*** vol. 14, no. 1, pp. 3-20.

Pirie, G. H. (1979). "Measuring Accessibility: A Review and Proposal." ***Environment and Planning A,*** vol. 11, pp. 299-3 12.

Reeker, W. W., and M. G. McNally (1985). "Travel/Activity Analysis: Pattern Recognition, Classification, and Interpretation." ***Transportation Research** A,* vol. 19A, no. 4, pp. 279-296.

Reeker, W. W., M. G. McNally, and G. S. Root (1986). "A Model of Complex Travel Behavior: Part I-Theoretical Development." ***Transportation Research A,*** vol. 2OA, no. 4, pp. 307-318.

Richardson, A. J., and W. Young (1982). "A Measure of Linked-Trip Accessibility." ***Transportation Planning** and Technology,* vol. 7, pp. 73-82.

Smith, T. E. (1980). "Additive Measures of Perceived Accessibility." *Environment and Planning* ***A,*** vol 12, pp. 829-841.

Vickerman, R. W. (1974). "Accessibility, Attraction, and Potential: A Review of Some Concepts and Their Use in Determining Mobility." Environment *and Planning A,* vol. 6, pp. 675-691.

Weibull, J. W. (1980). "On the Numerical Measurement of Accessibility." *Environment and Planning A,* vol. 12, pp. 53-67.

ANALYSIS OF LONGITUDINAL DATA FROM THE PUGET SOUND TRANSPORTATION PANEL

Task B: Integration of PSTP Databases and PSTP Codebook

FINAL REPORT

Submitted to

Federal Highway Administration
and
Battelle Memorial Institute
Work Order BAT-94-016

June 1996

PENNSTATE

1 8 5 5

Pennsylvania Transportation Institute

The Pennsylvania State University
Research Office Building
University Park, PA 16802
(814 865-1891

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Analysis of Longitudinal Data from the Puget Sound Transportation Panel Task B: Integration of PSTP Databases and PSTP Codebook		**5. Report Date** **6. Performing Organization Code**
7. Author(s) Konstadinos G. Goulias and Jun Ma		**8. Performing Organization Report No.** PTI 9619
9. Performing Organization Name and Address The Pennsylvania Transportation Institute Research Office Building The Pennsylvania State University University Park, PA 16802-4710		**10. Work Unit No. (TRAIS)** **11. Contract or Grant No.** BAT-94-016
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Highway Administration 400 Seventh Street SW Washington, DC 20590 Battelle Memorial Institute 505 King Avenue Columbus, OH 43201-2693		**13. Type of Report and Period Covered** Final Report 9/1/94-8/31/96 **14. Sponsoring Agency Code**
15. Supplementary Notes Federal Highway Administration COTR: Elaine Murakami		
16. Abstract		
17. Key Words Panel analysis, travel demand forecasting, longitudinal data, travel behavior		**18. Distribution Statement** No restrictions. This document is available from the National Technical Information Service, Springfield, VA 22161

16. Abstract

Each wave of the Puget Sound Transportation Panel contains trip-based, person-based, and household-based information. The trip-based information contains 2-day travel diary information that may lack consistency (e.g., missing trip records). In this report, we present the analysis of missing data and missing data augmentation that has been performed. A description of all the Puget Sound Transportation Panel databases—including integrated databases at the trip, person, and household levels—is also provided. In addition, a detailed "codebook" is included in this report for all the databases that were in usable format as of April 1996.

19. Security Classif. (Of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	124	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

FOREWORD

The Puget Sound Transportation Panel (PSTP) was established in 1989 in the Seattle metropolitan area by the Puget Sound Council of Governments (now called the Puget Sound Regional Council [PSRC]). It was initiated by Ms. Elaine Murakami (project manager) and Dr. Tim Watterson (director of technical services) with help from consultants and academic researchers. Through their efforts and prudent financial management, and with financial support from the Washington State Department of Transportation and other agencies, Ms. Murakami and Dr. Watterson, along with the PSRC staff, delivered the first general purpose transportation panel survey in the United States. Under their leadership and guidance, the panel continued until 1994.

Panel continuity, a common difficulty encountered by most panel surveys, is an indicator of Ms. Murakami's and Dr. Watterson's success. Continuity requires substantial effort in seeking financial sponsors, assuring panel participation by respondents, and providing continuous support for analysis. Because the PSTP is designed to be a panel survey, Ms. Murakami and Dr. Watterson concentrated on retaining as many continuous panel members as possible. These efforts included, and still include, sending holiday greeting cards and providing progress reports to the panel members. As a result, the PSTP attrition rates are lower than those reported by the Dutch National Mobility Panel. More important, its success provides valuable information for improving travel demand forecasting models and experience to other metropolitan regions that move toward panel surveys.

In 1993, Dr. Steve Fitzroy (director of technical services) took over leadership of the PSTP. Through his efforts and enthusiasm, two more surveys came out in 1993 and 1994, and plans for future data collection are under way. These efforts make PSTP a five-time point panel that provides tremendous potential for practice and research in travel behavior dynamics. Support for the panel was also provided by the Federal Highway Administration to evaluate the frost four-wave data and supplement PSTP with variables on the transportation level of service.

In addition to directing the data-collection efforts, extensive data-cleaning efforts were undertaken at PSRC as well. At PSRC, major contributors to PSTP data quality are Neil

Kilgren (who is also the person that has nurtured the panel since its conception), Robert Sicko, and Greig Lipton (now with the University of Washington). Such efforts greatly increase the credibility of the PSTP data analysis facilitate the data cleaning in this project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The Puget Sound Transportation Panel (PSTP) is the first general purpose, urban household panel survey in the United States (Murakami and Watterson, 1990) designed for transportation analysis. The major goals of the PSTP are to:

- Track changes in employment, work characteristics, household composition, and vehicle availability.
- Monitor changes in travel behavior and responses to changes in the transportation environment.
- Examine changes in attitudes and values of transit and non-transit users.

The PSTP was initiated in the Seattle metropolitan area in the fall of 1989 by the Puget Sound Regional Council (PSRC, then the Puget Sound Council of Governments) in partnership with transit agencies in the region. Funding for surveys in 1989 and 1990 was from a special transit data grant administered through the Washington State Department of Transportation. Unlike the traditional transportation cross-sectional surveys, the PSTP is a panel or longitudinal survey in which similar measurements (i.e., surveys) are made repeatedly on the same observations over time. A survey conducted at each point in time is called a wave. Each wave of the PSTP includes a travel survey, which may or may not include an attitudinal survey. The travel survey's three components are household demographics, person so&economics, and reported travel behavior, The attitudinal survey includes subjective travel attitudes. The PSTP contains data from five travel surveys (1989, 1990, 1992, 1993, and 1994-the fifth survey data cleaning and reporting is currently being performed at PSRC) and three attitudinal and values surveys (1990, 1991, and 1993). The data analyzed in this project (BAT-94-016) are from the first four travel surveys.

Research conducted between 1985 and 1988 in the Puget Sound region found that transit trips accounted for only 6 percent of all household trips. Therefore, to obtain several transit users, PSTP data were collected based on the stratification of travel-mode choice (drive alone, car pool; and transit).. The PSTP consists of three distinct household populations: households with regular transit users (defined as households in which at least one person makes at least four one-way trips by transiteach week), households with regular work-trip car poolers (defined as households in which at least one person makes at least four one-way trips in a car pool to work on a weekly basis), and households without regular transit users or car poolers. Three survey methods were used to recruit potential panel members in 1989. The telephone random digit dialing (RDD) technique was primarily for drive-alone and car pool households. People contacted through RDD were the representatives of their households. They were first asked some questions regarding household demographic information and then requested to indicate their willingness to be involved in the panel survey (for the whole household). Household representatives who agreed to participate in the panel were considered members of the potential panel, which consisted of all driving-age household members (15 years of age and older). In addition to RDD, an increase in the transit-user portion of the sample was achieved by “recontacting” Seattle Metro transit survey respondents who showed a willingness to participate in future research. Other recruits were obtained via distribution of letters to request volunteers on randomly selected bus routes. The panel was designed to track changes in mode choice, and the panel members were recruited according to their travel-mode use (this sampling scheme is called choice-based-i.e., data are collected based on the population characteristic that we attempt to explain, herein called endogenous travel modes).

In addition to the stratification on the endogenous travel modes, each sample of the three household populations is further stratified by county of residence (King, Kitsap, Pierce, and Snohomish), which is considered an exogenous variable to the study (this will have some operational implications when weights are created for the sample). The objective of this “double” sample stratification (i.e., by travel mode and county of residence) is to increase the sample size of transit users for acceptable statistical analysis. This may increase reliability of the statistical analyses performed.

After the potential panel members were recruited, mail-out mail-back travel diary forms were sent to them to collect their trip-making behavior. This instrument followed an “open” format in which potential panel members filled in blanks to identify the location, the reason (purpose), the mode, and the time traveled for each trip they made. Each driving-age household member (15 years and older) of the potential panel members was required to record every trip during two consecutive weekdays of the sampling period (i.e., Monday and Tuesday, Tuesday and Wednesday, Wednesday and Thursday, or Thursday and Friday). Each household maintained the same assignment days in the subsequent surveys. The option to declare no trips made in one or more of the survey days was also provided. Data collection in each wave took place in the fall of each year, excluding the Thanksgiving holiday weekend (between November 23 and November 25). However, not every potential panel member returned or completed the travel diary. Only potential panel members who completed the travel diaries became the panel members.

Because of the relocation of households outside the region analyzed (i.e., households leaving the region are not followed in this survey) and the panel fatigue from repeatedly flling out the travel diary, the number of observations remaining in the panel for all waves declined as the panel survey continued. The phenomenon of sampling units (in this case households) that refuse to participate in subsequent waves of the panel is called attrition. To maintain the same number of observations over time and to reflect changes in the population, new households of similar demographic and trip-making characteristics were recruited in each panel wave. Participants starting in the second and all subsequent waves are called sample refreshment (note that this is one way of replenishing a dynamic sample). Unlike the original sample-collection method, these replacement households were recruited only through the telephone random digit dialing technique. Random digit numbers ensured that the generated households were proportional to the existing sample in geographic locations, which were represented by telephone prefix. Household evolution is also traced by PSTP following their splitting process. For example, when a person leaves her/his original "nest" because of marriage or divorce, PSTP follows both the original nest and the newly formed household. When a household member moves out of the household and has a different residence in the panel, the formation of a new household is assumed and it is assigned a "sub-household" identification number. The identification number (HHID) of the original nest (core household) does not change from wave to wave.

Initially, 5,175 households' were contacted by telephone, including transit recontacts and bus-ride volunteers. Of these, 2,944 (56.9%) agreed to complete travel diaries. Out of the 2,944 participants, 1,712 (58.2%) households returned their diaries. For those that refused to participate in the panel survey, some of their so&demographics were included in the household file. Once a household became a panel member, its sample category, which was based upon the initial mode usage and recruitment method, was carried through the following surveys, regardless of the household's actual travel-mode category. The sample profile of trip records in each of the four waves is provided in table 1. The numbers in parentheses are the number of households or persons that had complete travel information but did not have either associated household or person information. These persons are listed in appendix C. The attrition rate (percent of households dropping out) from wave 1 to wave 2 is 15.2%. The overall retention rate (the opposite of attrition) during the four waves is 57.8% (i.e., 990 households, including split households, participated in all four waves). Figure 1 shows the number of stayers and replacements for each wave.

In table 2, households that appeared in the household data files in each wave are categorized by their primary travel modes. Because not all of the sociodemographic information of households that returned travel diaries is available, table 3 provides the number of households with both available household and person information for each travel-mode group. In a subsequent report on sample weighting we develop a different set of definitions for a variety of sample components. Participation in the panel is somewhat more complex than what is described in this report.

---

[1]Five records have duplicate HHIDs. The HHIDs of these households are 5223,5653,6889, and 7195. Household 5,653 has two duplicates, which have exactly the same fields.

	Initial Contacts	*Initial Respondents	Wave 1 **Participants	Wave 2 Participants	Wave 3 Participants	Wave 4 Participants
Stayers	5175	2944	1712	1452	1131	990
Split Households				41	33	48
Replacement 1				400	270	142
Replacement 2					362	282
Replacement 3						565
Others (w/o household or person info)				2	4	2
Total	5175	2944	1712	1895	1797	2030

*Initial Respondents: Households that agreed to receive travel diaries
**Participants: Households that returned travel diaries

Figure 1. Sample composition.

Table 1. Sample provide by wave.

Sample Profile	1989 (Wavel)	1990 (Wave2)	1992 (Wave3)	1993 (Wave4)
# of households[1]	1713 (1)	2157(264)	1696 (97)	2003 (71)
Original households (SUBID=0)	1713 (1)	2112(259)	1693 (97)	1972 (66)
Refreshment households	0	400	362	565
First split households (SUBID=1)	0	43 (4)	3	30 (5)
Second split households (SUBID=2)	0	2 (1)	0	1
# of households in all previous surveys	N/A	1711(259)	1146 (67)	970 (34)
# of persons	3391 (1)	4625(969)	3360(304)	4081(230)
# of refreshment persons	0	1023(263)	686 (16)	1063 (5)
# of persons in all previous surveys	N/A	2822 (85)	1942 (24)	1633 (8)

[1]In wave 2, five households with SUBID = 0 did not complete the travel diary. But their split households (SUBID = 1) returned the diary. These households' HHIDs are 402,2425,3023,3546, and 3,752. Therefore, the number of households in wave 2 are the sum of households with SUBlD=O and those five households with SUBID= 1 and without SUBlD=O. Similarly, in wave 3, households 128 and 3,023 with SUBlD=O did not return the travel diary, whereas their split households did. Thus, the number of households that returned the travel diary is the number of households with SUBlD=O plus two.

Table 2. Choice-based sample composition by wave (based on household files).

Primary Travel Mode	Number of Households (inc. split households)			
	Wave 1	Wave 2	Wave 3	Wave 4
Drive-alone	1138	1435	1132	1286
Bus	382	460	371	474
Car pool	193	259	192	241
Unknown	0	3[1]	1	2
Total	1712	2157	1696	2003

[1]2 out of 3 were coded as O's, which is not among the codes in the description file.

Table 3. Choice-based sample composition by wave (based on trip files).

Primary Travel Mode	Number of Households (inc. split households)			
	Wave 1	Wave 2	Wave 3	Wave 4
Drive-alone	1137	1252	1072	1240
Bus	382	391	346	464
Car pool	193	210	180	227
Unknown	0	38	1	1
Total	1712	1893	1599	1932

2. DATA CONSISTENCY EVALUATION AND RECTIFICATION (TASK B.l)

It is very important to have consistent data in the sample, particularly when **analyzing** trip-making behavior. The data records were examined for inconsistencies. The original trip, household, and person data records of each wave were examined first. To the greatest extent possible, inconsistencies were removed. A summary of the corrections is provided in this report.

DESCRIPTION OF ORIGINAL TRIP RECORD DATABASES

Most PSTP information was collected via trip diaries, except for trip distance (DISTANCE), which was derived from trip origins and destinations reported by respondents. This was computed by first converting trip ends into regional traffic analysis zones (TAZs) and then matching the trip with the calculated zone-to-zone distance using the regional transportation forecasting model. One way to analyze the PSTP data is to use the trip diary information as an activity diary and create measures of activity participation and segments of time allocated to each activity during the interview period (i.e., activity durations). This provides information for activity-based analyses (reflecting the rationale that people travel to participate in activities and that understanding travel behavior requires first understanding activity participation and its dynamics). Travel diaries are much simpler to administer and complete than activity diaries. Although travel diaries have some advantages over other approaches (e.g., simplicity, economy, etc.), it was found that respondents had difficulty selecting the type of trip in the given questionnaire categories (e.g., work-related, personal, recreation, etc.) Therefore, responses were recoded after the diaries were returned. This effort, “recoding” the travel diaries after their return, helped to increase data quality. However, some inconsistency in the trip data remained. There are mainly two types of inconsistencies in the trip data. The first type regards disagreements between data documentation and the electronic data set, and the second regards violation of "logic” within the electronic data set.

An example of the first type of inconsistency is a case where the trip purpose in the trip data file was coded as 15. The code 15 was not included and was not explained in the trip database description file. Another example is that person trips were coded twice within a day in the trip data file (i.e., the same trip appears twice in the database). The second type of inconsistency by violation of logical relations needs a somewhat more detailed analysis. For example, consider the visit to an activity location. The difference between departure time from this location and arrival time at this location, which is the time that an activity is pursued, should be positive. There are several instances where this value is negative. Details about the treatment of all of these cases are offered later in this report. Descriptive summaries of original trip data of the four waves are listed in tables 4, 5, 6, and 7. The detailed data "codebook," called the PSTP Codebook, is provided in appendix A.

TRIP RECORD DATA TREATMENT

A first pass of data cleaning[1] was made to generate consistent data records. However, inconsistencies persist. First, not all of the missing values have been imputed in the first pass because of lack of other supplementary information. For instance, when a person reported making only one trip within a day, a close examination was conducted based on the trip purpose, the trips of previous and following days, and the beginning and ending times of trips. A new trip was added if the returning trip was missing. This imputed trip has missing departure time, arrival time, origin, and destination. Second, origin/destination continuity was not investigated due to the intensive data manipulation needed, except when time inconsistency was encountered. Finally, if either starting time or ending time of a trip was missing, the associated travel time and activity time could not be computed. Therefore, they were coded as non-positive values (-9 for missing).

---

[1]This is after the extensive data cleaning and analyses done at PSRC.

Table 4. Descriptive variables of trip data in wave 1.

Original Trip Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	2279.93	1298.44	1	4209	31342
SUBID	.00	.00	0	0	31342
PERS	1.59	.74	1	6	31342
DIARY	1.49	.50	1	2	31342
DAYOFWK	2.89	1.26	1	5	31342
MDD	265.10	52.59	125	421	31342
TOTTRIP	6.21	3.07	0	18	31342
TRIPNUM	3.60	2.52	0	18	31342
PURPOSE	5.17	2.71	1	9	30845
TYPE	2.16	.74	1	3	30845
MODE	1.87	1.63	1	16	30845
BEGTIME	1372.49	452.96	100	2450	30841
ENDTIME	1401.02	452.51	100	2455	30840
MINUTES	23.20	39.85	1	615	31342
D_R	1.28	.56	1	3	30844
NUM	1.57	.91	1	9	30837
REL1	2.55	1.85	0	7	11853
REL2	3.39	1.91	0	7	1970
REL3	4.75	2.39	0	7	130
ORIGCT	44274.99	26775.39	100	99999	31341
DESTCT	44192.76	26784.24	100	99999	30844
DISTANCE	6.22	7.17	.30	82.00	30845

Rectified Trip Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	2280.51	1298.23	1	4209	31335
SUBID	.00	.00	0	0	31335
PERS	1.59	.74	1	6	31335
DIARY	1.49	.50	1	2	31335
DAYOFWK	2.89	1.26	1	5	31335
MDD	265.09	52.59	125	421	31335
TOTTRIP	6.21	3.07	0	18	31335
TRIPNUM	3.60	2.52	0	18	31335
PURPOSE	5.09	2.77	0	9	31335
TYPE	2.12	.78	0	3	31335
MODE	1.84	1.64	0	16	31335
BEGTIME	1374.44	453.22	1	2700	30833
ENDTIME	1404.17	452.48	15	2715	30831
MINUTES	18.62	17.94	1	540	30831
D_R	1.16	.41	0	2	29682
NUM	1.54	.92	0	9	31327
REL1	2.46	1.88	0	7	12283
REL2	2.72	2.18	0	7	2451
REL3	1.00	2.23	0	7	613
ORIGCT	44308.30	26769.59	100	99999	31308
DESTCT	44223.12	26779.09	100	99999	30813
DISTANCE	6.12	7.15	.00	82.00	30858
ATIME	125.94	173.11	0	1040	24548
FLAG	1.57	19.63	0	323	31335

Table 5. Descriptive variables of trip data in wave 2.

Original Trip Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	3573.35	2970.95	1	10356	33271
SUBID	.02	.13	0	2	33271
PERS	1.57	.72	0	6	33271
DIARY	1.50	.50	1	2	33269
DAYOFWK	2.92	1.24	1	7	33271
MDD	143.03	49.35	101	318	33271
TOTTRIP	6.25	3.22	0	18	33271
TRIPNUM	3.61	2.59	0	18	33271
PURPOSE	5.02	2.83	1	9	32640
TYPE	2.13	.75	1	3	32640
MODE	1.89	1.76	1	16	32640
BEGTIME	1350.02	456.89	0	2355	32639
ENDTIME	1371.82	461.49	0	2400	32637
MINUTES	21.24	32.17	1	540	32640
D_R	1.26	.52	1	3	32640
NUM	1.54	.95	1	35	32640
REL1	.96	1.69	0	7	32640
REL2	.22	1.01	0	7	32640
REL3	.03	.38	0	7	32638
ORIGCT	44611.07	26751.32	100	99999	33270
DESTCT	44467.79	26764.44	100	99999	32640
DISTANCE	6.47	7.50	.30	81.20	32639

Rectified Trip Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	3573.84	2970.93	1	10356	33284
SUBID	.02	.13	0	2	33284
PERS	1.57	.72	1	6	33284
DIARY	1.50	.50	1	2	33284
DAYOFWK	2.92	1.24	1	7	33284
MDD	143.04	49.35	101	318	33284
TOTTRIP	6.25	3.22	0	18	33284
TRIPNUM	3.61	2.59	0	18	33284
PURPOSE	4.92	2.89	0	9	33284
TYPE	2.09	.80	0	3	33284
MODE	1.86	1.76	0	16	33284
BEGTIME	1360.06	450.22	20	2645	32561
ENDTIME	1389.88	449.83	15	2705	32561
MINUTES	18.75	18.75	0	540	32554
D_R	1.16	.42	0	2	31976
NUM	1.52	.97	0	35	33284
REL1	2.49	1.88	0	7	12640
REL2	2.79	2.39	0	7	2580
REL3	1.10	2.16	0	7	823
ORIGCT	44615.37	26773.82	100	99999	33259
DESTCT	44467.12	26788.31	100	99999	32630
DISTANCE	6.36	7.49	.00	81.20	32914
ATIME	123.54	172.78	0	1135	25868
FLAG	3.23	28.99	0	401	33284

Table 6. Descriptive variables of trip data in wave 3.

Original Trip Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	5030.49	4057.29	4	13210	29510
SUBID	.01	.13	0	2	29510
PERS	1.60	.77	1	8	29510
DIARY	1.50	.50	1	2	29510
DAYOFWK	3.00	1.24	1	5	29510
MDD	197.49	46.21	122	508	29510
TOTTRIP	5.87	3.04	0	18	29510
TRIPNUM	3.42	2.44	0	18	29510
PURPOSE	5.12	2.81	1	15	28860
TYPE	2.09	.74	1	3	28860
MODE	1.88	1.71	1	16	28860
BEGTIME	1349.10	454.52	0	2430	28860
ENDTIME	1378.74	454.05	0	2450	28860
MINUTES	19.00	19.39	1	466	28860
D_R	1.18	.38	1	2	27619
NUM	1.58	1.23	1	43	28860
REL1	.91	1.59	0	7	28860
REL2	.26	1.10	0	7	28860
REL3	.04	.43	0	7	28860
ORIGCT	43241.14	26348.25	100	99999	29506
DESTCT	43081.50	26383.05	100	99999	28856
DISTANCE	7.02	9.09	.20	151.00	28764

Rectified Trip Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	5030.40	4056.96	4	13210	29518
SUBID	.01	.13	0	2	29518
PERS	1.60	.77		8	29518
DIARY	1.50	.50	1	2	29518
DAYOFWK	3.00	1.24	1	5	29518
MDD	197.50	46.20	122	508	29518
TOTTRIP	5.87	3.04	0	18	29518
TRIPNUM	3.42	2.44	0	18	29518
PURPOSE	5.01	2.87	0	9	29517
TYPE	2.05	.79	0	3	29518
MODE	1.83	1.71	0	16	29518
BEGTIME	1355.96	452.76	30	2700	28854
ENDTIME	1385.94	452.03	100	2710	28852
MINUTES	18.91	18.68	1	466	28852
D_R	1.15	.42	0	2	28275
NUM	1.55	1.24	0	43	29518
REL1	2.33	1.78	0	7	11316
REL2	2.88	2.40	0	7	2626
REL3	1.16	2.16	0	7	876
ORIGCT	43225.44	26348.23	100	99999	29470
DESTCT	43066.63	26383.72	100	99999	28822
DISTANCE	6.87	9.06	.00	151.00	29422
ATIME	130.78	175.37	0	1215	22623
FLAG	.06	4.10	0	311	29518

Table 7. Descriptive variables of trip data in wave 4.

Original Trip Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	7230.98	5429.94	4	14713	33172
SUBID	.01	.13	0	2	33172
PERS	1.61	.80	1	9	33172
DIARY	1.50	.50	1	2	33172
DAYOFWK	2.97	1.26	1	7	33172
MDD	171.99	62.10	113	412	33172
TOTTRIP	5.97	3.17	0	20	33172
TRIPNUM	3.47	2.51	0	20	33172
PURPOSE	5.17	2.81	1	9	32441
TYPE	2.09	.74	1	3	32441
MODE	1.93	1.76	1	16	32441
BEGTIME	1351.27	452.04	0	2540	32441
ENDTIME	1381.07	450.99	0	2555	32441
MINUTES	18.95	21.19	0	900	32441
D_R	1.19	.39	1	2	30789
NUM	1.58	1.14	0	40	32441
REL1	.94	1.64	0	7	32441
REL2	.24	1.02	0	7	32441
REL3	.04	.47	0	7	32440
ORIGCT	41804.24	26377.73	100	99999	33163
DESTCT	41613.83	26371.34	100	99999	32432
DISTANCE	6.89	9.37	.20	216.00	32314

Rectified Trip Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	7230.70	5430.11	4	14713	33179
SUBID	.01	.13	0	2	33179
PERS	1.61	.80	1	9	33179
DIARY	1.50	.50	1	2	33179
DAYOFWK	2.97	1.26	1	7	33179
MDD	171.98	62.10	113	412	33179
TOTTRIP	5.97	3.17	0	20	33179
TRIPNUM	3.47	2.51	0	20	33179
PURPOSE	5.05	2.88	0	9	33179
TYPE	2.05	.80	0	3	33179
MODE	1.89	1.76	0	16	33179
BEGTIME	1351.30	451.61	30	2730	32425
ENDTIME	1380.97	450.77	20	2805	32428
MINUTES	18.85	20.14	0	810	32425
D_R	1.16	.43	0	2	31527
NUM	1.54	1.15	0	40	33177
REL1	2.39	1.83	0	7	12695
REL2	2.68	2.27	0	7	2879
REL3	1.32	2.28	0	7	1052
ORIGCT	41785.58	26380.36	100	99999	33152
DESTCT	41594.20	26371.73	100	99999	32431
DISTANCE	6.74	9.32	.00	216.00	33052
ATIME	129.76	174.56	0	1025	25450
FLAG	.05	2.52	0	223	33179

Missing Values and Coding Errors

Missing data are not very frequent in the database. The number of missing values for each variable is shown in table 8. In wave 1, three underscores appeared in RELl and the corresponding number of people in the vehicle is 1. This implies that RELl (the relation to the vehicle occupant 1) does not apply because there are no passengers in the vehicle. It should be kept in mind that the record of people making no trips in a day (herein referred to as the “zero trip” people) is also included in the trip record files. When the total number of trips is 0 (TOTTRIP = 0), TRIPNUM, PURPOSE, TYPE, MODE, BEGTIME, ENDTIME, D-R, NUM, DESTCT, and DISTANCE (in waves 1 and 2) were coded as missing values in the original file. These zero trips, therefore, artificially inflate the missing data statistics. In order to obtain valid statistical analysis, all variables of zero trips were converted to zeros.

Similarly, some problems may arise when the number of total trips is equal to 1. A person usually makes at least two trips in a day, starting and ending at his/her residence (except for rare cases). In wave 1, there are 29 single trips (no return or no departure from the same location). Of these, 14 trips were found to return home from previous trips, which were made beyond the time frame of the survey, or to return home the next day. Eight out of the remaining 15 trips were made to the airport. They were assumed to leave the region and, thus, were considered to be reasonable in reporting single trips. The remaining seven trips were primarily work and leisure related trips with travel time ranging from 10 minutes to about 10 hours. For any trip longer than 4 hours or for a work trip made after 3:00 p.m., no returning home trip is considered necessary. Therefore, only two additional trips that were likely to be left out by respondents were added. Each added trip was assumed to have the same characteristics as the reported single trip. Because of lack of information of activity duration, the starting times and ending times of the added trips were coded as missing. In the same way, of all 61 single trips (in both waves 2 and 3), 13 and 8 trips were added in waves 2 and 3, respectively. Similarly, 12 out of 65 single trips were included in wave 4.

Table 8. Number of missing values in trip data.

Variables	Wave 1	Wave 2	Wave 3	Wave 4
Original # of observations	31342	33271	29510	33172
# of observations w/ 0 trip	496[1]	631	650	731
PERS		7 obs coded as 0		
DIARY		2		
BEGTIME	4	1		
ENDTIME	5	3		
D R	1		1241	1652
NUM	8			
ORIGCT	4		4	9
DESTCT	1		4	9
DISTANCE		1	96	127
# of obs. w/ single trip	29	61	61	65
# of duplicate obs.	1			4
# of missing trips		1		
# of obs. w/ an added trip	2	13	8	12
# of trips deleted	8			1
Final # of obs.	31335	33284	29518	33179

[1]With help from the PSRC staff, one of the zero-trips was confirmed to be incorrectly coded and, thus, was removed from the database.

Examination of trips for sequencing revealed some data problems. First, there were duplicate trip records in the trip file. In other cases, trip sequences were broken or exhibited mixed trip sequences. In waves 1, 2, and 4, there were several cases in which a person's total number of trips was not equal to the sum of the individual trips. In addition, occasionally both zero and non-zero trips were found for the same person on the same day. These were corrected by contacting staff of the PSRC. The final number of trip records is shown in table 8.

Data Errors from Observations

Data errors from observations are due to mistakes in time notations, numeric errors, and problems in trip sequencing. The data in the four waves presented various degrees of erratic records. Wave 3 is the “cleanest” of all the waves.

Time Notation

In daily life, a 24-hour day is usually denoted by two 12-hour periods, which are differentiated by a.m. and p.m. In order to facilitate the coding, time in the trip file used the military time notation (i.e., a 24-hour period in which 00:01 means 12:01 a.m. and 24:00 stands for 12:00 midnight). Most people are not used to such time notations in their daily lives. The miliary notation, however, reduces confusion in data coding and requires only numeric data fields. Therefore, it is more efficient in terms of data storage than the 12-hour time notation. In trip files, the 24-hour notation was used inconsistently. It seems that when an error was suspected, the time was coded incorrectly (there is no document to support the reasoning for this). An example from the trip record file is the following:

PURP.	MODE	BEGT.	ENDT.	MIN.	DIST.
6	1	715	845	90	19.20
6	1	900	920	20	5.50
6	1	*2130*	*2135*	5	1.00
6	1	940	955	15	5.90
6	1	1000	1010	10	8.40

The example above provides a sequence of trips (in terms of origins and destinations). It is obvious that 2130 and 2135 mean 9:30 a.m. and 9:35 a.m., respectively. However, this creates problems when one analyzes the data and attempts to provide statistics of departure and arrival times in a day. The strategy to correct this was to convert the time to the correct time notation, if changing time makes perfect sense. Therefore, the previous trip records were corrected as follows:

PURP.	MODE	BEGT.	ENDT.	MIN.	DIST.
6	1	715	845	90	19.20
6	1	900	920	20	5.50
6	1	*930*	*935*	5	1.00
6	1	940	955	15	5.90
6	1	1000	1010	10	8.40

After this rectification, it is possible that some other inconsistency exists. **For** example, few trips were coded as follows:

PURP.	MODE	BEGT.	ENDT.	MIN.	DIST.
1	2	1020	1045	25	13.30
1	2	1135	*10*	301[2]	17.70
5	2	1220	1310	50	36.00

Again, it is obvious that the second trip's ending time was not at 00: 10 a.m., instead it ended at 00:10 p.m., which was equivalent to 12: 10. In addition, DISTANCE suggested that 35 minutes would be reasonable for traveling 17.70 miles.

[2]301 was a code for incorrect travel time duration in the original dahbase. However, this was not documented in the original description files.

Numeric Errors

In the trip files some "numeric” errors were also found. This might have resulted from unclear writing by respondents and/or erroneous coding. Sometimes, 1 was mistakenly recorded as 2, 7 as 9, 8 as 6, 3 as 5, etc. Trip duration, distance, and purpose can be used jointly to assist correcting these errors. Examples include the following:

PURP.	MODE	BEGT.	ENDT.	MIN.	DIST.	ORIG.	DEST.
4	1	2000	2020	20	7.20	60902	63000
8	1	*2230*	*1250*	301	7.20	63000	60902

There is no doubt that 2250 was incorrectly coded as 1250. However, sometimes this is not as clear as in the example. Regarding starting or ending times one may observe:

PURP.	MODE	BEGT.	ENDT.	MIN.	DIST.	ORIG.	DEST.
6	1	1715	1719	4	4.10	92700	92400
8	1	*1931*	*1735*	301	4.10	92400	92400

It may be that 193 1 should be coded as 173 1. It is also possible that 1735 should be coded as 1935. In this case, trip purpose may be helpful to decide the correct choice, but this is not always the case. In this case, changing starting or ending time is "randomly” chosen to avoid bias.

Out-of-Order Trip Sequence

Occasionally trips are out of their sequential order. Origin and destination may provide information of the trip sequence. Most of the time, however, the origin/destination sequence is correct. If changing the trip sequence does not result in any inconsistency of the origin/destination, then trip sequence can be rearranged in a way that will produce a correct trip sequence in terms of time consistency. If switching trip records breaks the continuity of the origin/destination sequence, only the trip times are switched. Consider the following example:

PURP.	MODE	BEGT.	ENDT.	MIN.	DIST.	ORIG.	DEST.
1	1	1215	1225	10	4.80	90400	91201
6	1	1830	1850	20	2.70	91201	91300
8	2	*1900*	*1905*	5	2.00	91300	91300
7	1	*1850*	*1900*	10	2.00	91300	91300
8	1	2020	2030	10	2.00	91300	91300

The third and fourth trips are switched in the sequence because the resulting origin/destination sequencing makes sense.

The second type of error, in out-of-order trip sequence, is that either starting time or ending time was not coded correctly. For example:

PURP.	MODE	BEGT.	ENDT.	MIN.	DIST.	ORIG.	DEST.
1	1	814	830	16	11.25	70202	73300
8	1	1715	1845	90	11.25	73300	70202
5	2	1818	1827	9	2.97	70202	70202
8	2	1934	1939	5	2.97	70202	70202

It is reasonable to suspect that it would take 90 minutes to travel 11.25 miles, although this is only possible when the highway is highly congested. However, the fact that the ending time of the second trip was after the starting time of the third trip may lead to the conclusion that 1845 is wrongly coded and should be 1745.

The third type of error in sequencing is when two trips have exactly the same starting and/or ending time, The following illustrates such a case:

PURP.	MODE	BEGT.	ENDT.	MIN.	DIST.	ORIG.	DEST.
8	2	1810	1848	38	6.20	72105	72000
7	2	1818	1848	38	3.20	72000	91902

It is very difficult to identify the true ending time of the first trip and the starting time of the second trip. If the trip distance is available, the duration will be assigned a new value proportionally to the distance, and the associated starting/ending time will be calculated accordingly. If the distance is missing, the duration will be given equal value as the previous trip. There were some other cases where two trips had exactly the same ending times but different starting times. The time is corrected in a similar manner.

The last type of error in this category is when there is no information about the trip sequence. For example:

PURP.	MODE	BEGT.	ENDT.	MIN.	DIST.	ORIG.	DEST.
3	2	732	734	7	0.8	25700	25802
4	7	1615	1617	2	0.7	25802	25802
2	2	1445	1449	4	1.4	25802	25802
5	2	1505	1506	4	1.4	25802	25802
4	2	1615	1617	2	0.7	25802	25802

The second and the last trips were both visiting trips made at exactly the same time. The mode of the second trip seems unreasonable because it is impossible for a person to walk 0.7 miles within 2 minutes. Thus, this mode was changed to 2 because this person's trips were mainly made by car pool. It is likely that the second trip was the duplicate of the last trip, therefore, it was removed from the trip record file.

DESCRIPTION OF RECTIFIED TRIP RECORD DATABASE

To facilitate data analysis, an additional variable was created; that is, the duration of activity time (ATIME). Based on the maintained hypothesis that travel is derived from a person's desire to participate in activity, the purpose of a trip is to engage in a certain type of activity, and the trip destination is the place where the activity is pursued. The corresponding activity duration is, thus, the time difference from when a person leaves and when a person arrives at the destination. That is, each trip record, except for the last one a person makes at the end of the day, can be associated with an activity with a duration that equals the difference between the starting time of the next trip record and the ending time of the current trip record (i.e., $ATIME_i = BEGTIME_{i+1} - ENDTIME_i$).

The variables and corresponding summary statistics of the cleaned trip record database are shown in tables 4, 5, 6, and 7 for each of the four waves. The summary statistics are not significantly different from the original data because of the small number of incorrect records. However, the derived activity duration indicates that there have been noticeable changes. This derived measure is crucial in activity analysis. For this reason, each trip record contains a new variable indicative of the type of "corrections" performed, called FLAG. This allows researchers to analyze the original database.

DESCRIPTION OF ORIGINAL HOUSEHOLD RECORD DATABASES

Information included in waves 1 and 2 is the same, except that two additional variables that indicate panel and attrition status were added in wave 2. However, the format of waves 3 and 4 household data was slightly different from those of the previous two waves. For instance, in waves 1 and 2 income is recorded as a categorical variable. Waves 3 and 4 contain two separate indicators for income. One is categories (a categorical variable), the other is household income in dollars (a continuous variable). In addition, classification of the income categories is different in the first two and the last two waves. Although income in dollars could be used to recode the categorical income in waves 3 and 4, overwriting the categorical income using the definition of waves 1 and 2 would result in more missing values.

To maintain as much informtion as possible, the original categorical and continuous income variables in **waves** 3 and 4 are kept in the database. In addition to incompatibility in income, the date on which the travel diary was completed by each person (DATE1 and DATE2) was not included in waves 2, 3, and 4 data files. Similarly to the trip data files, each household data was accompanied by a description file that provided the name, length, beginning column, and meaning of the variables. The number of households obtained from household files is provided in table 9.

Table 9. Number of households.

	Number of households	Number of households by SUBID		
		Subid= 0	Subid = 1	Subid = 2
Wave 1	5175	5175	0	0
Wave 2	2157	2112	43	2
Wave 3	1696	1693	3	0
Wave 4	2003	1972	30	1

There are two types of inconsistencies in the household data. One is that the data file is inconsistent with its corresponding description file. For instance, the underscore coding ( ) existed in the household data file, but not in the corresponding description file. The other type of inconsistency is the violation of logical relations. For example, the size of any household equals the sum of the number of adults who are 18 years and older, the number of children whose ages are between 6 and 17, and the number of children who are 5 years and younger. However, in the household data file it was found that this relation did not always hold. The former type of inconsistency (data miscoding error) may be attributed to miscoding at the data coding stage, while the latter inconsistency (error from observations) to the errors made by the respondents.

Data consistency was checked in all four waves. No effort, except the day of the week of travel diaries (DAY1 and DAY2), was made in checking the data between waves during this initial pass. In-general, waves 3 and 4 contain fewer inconsistencies than waves 1 and 2. The detailed description of correcting inconsistent data is discussed in the following section. The summary statistics of each variable in each wave is listed in tables 10, 11, 12, and 13. The rectified files also include information on the zip code, Census Tract, and TAZ where each household resides at each wave.

HOUSEHOLD DATA TREATMENT

Missing Values and Coding Errors

Household data of waves 1 through 4 were investigated. The wave 1 database contains fewer errors than the wave 2 database. Missing information in waves 1 through 4 is shown in table 14.

The following is a sample of coding:

- In both waves 1 and 2, TOTADULT, TOT6-17, TOTl-5, AND TOT-LOG are coded as missing values '.' when HHSIZE= 1 or missing.
- In wave 2, sometimes household size (HHSIZE) was coded as star (*), which was not defined in the description file.
- In wave 2, column 37 is not defined, but some numbers appeared in this column.
- In wave 3, there were two cases where DAY1 and DAY2 were coded as 8, which was not defined in the descriptive file.
- In waves 2, 3, and 4, HHTYPE was incorrectly coded sometimes. For instance, HHTYPE 1 and 2 were associated with children and adult households. However, several type 2 households (HHRPE=2) were found to have no children at all.

To simplify the data retrieval by the software used, characters such as Y/y and N/n were converted to numeric values. These conversions are shown in table 15.

Table 10. Descriptive variables of household data in wave 1.

Original Household Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	3950.09	2176.21	1	7231	5175
SUBID	.00	.00	0	0	5175
COMPLETE	1.19	.79	0	2	5175
SAMPLE	1.45	.69	1	3	1713
INCOME	4.89	2.15	0	9	4746
HHTYPE	4.62	2.77	1	8	5175
HHSIZE	2.79	1.40	1	10	5162
TOTADULT	2.13	.79	1	9	4785
TOT6_17	.54	.91	0	7	4784
TOT1_5	.26	.59	0	4	4785
TOT_LOG	2.08	.82	1	10	2944
NUMVEH	2.23	1.30	0	9	5171
DAY1	2.39	1.15	1	4	1713
DATE1	262.99	52.08	125	420	1713
DAY2	3.39	1.15	1	5	1713
DATE2	266.28	52.53	126	421	1713

Rectified Household Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	3950.09	2176.21	1	7231	5175
SUBID	.00	.00	0	0	5175
COMPLETE	1.19	.79	0	2	5175
SAMPLE	1.45	.69	1	3	1713
INCOME	4.89	2.15	0	9	4746
HHTYPE	4.62	2.77	1	8	5175
HHSIZE	2.79	1.40	1	10	5162
TOTADULT	2.04	.82	1	9	5161
TOT6_17	.50	.89	0	7	5161
TOT1_5	.25	.58	0	4	5162
TOT_LOG	2.08	.82	1	10	2944
NUMVEH	2.23	1.31	0	9	5174
DAY1	2.39	1.15	1	4	1713
DATE1	262.99	52.08	125	420	1713
DAY2	3.39	1.15	1	5	1713
DATE2	266.28	52.53	126	421	1713
HHZIP	98197.52	152.89	98001	98580	1712
HHTRACT	461.85	260.80	1.00	999.82	1712
HHTAZ	493.67	233.63	1	849	1712
FLAG	.11	.60	0	10	5175

Table 11. Descriptive variables of household data in wave 2.

Original Household Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	3440.21	2871.84	1	10356	2158
SUBID	.02	.15	0	2	2158
SAMPLE	1.44	.71	0	3	2132
INCOME	5.35	2.10	0	9	1886
HHTYPE	4.39	2.59	0	8	1913
HHSIZE	2.62	1.27	1	8	1920
TOTADULT	1.93	.63	1	5	1915
TOT6_17	.51	.89	0	9	1752
TOT1_5	.27	.62	0	9	1686
TOT_LOG	2.02	.75	1	6	1922
NUMVEH	2.15	1.12	0	9	1921
DAY1	2.44	1.15	1	7	1999
DATE1	162.07	84.09	101	426	2001
DAY2	3.43	1.15	1	7	1999
DATE2	163.25	83.99	102	427	1999
PANEL	.84	.37	0	1	2154
STATUS	.36	.80	0	3	1608

Rectified Household Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	3441.32	2872.05	1	10356	2157
SUBID	.02	.15	0	2	2157
SAMPLE	1.45	.70	0	3	2156
INCOME	5.35	2.10	0	9	1886
HHTYPE	4.44	2.58	1	8	2127
HHSIZE	2.61	1.28	1	11	1921
TOTADULT	1.93	.63	1	5	1915
TOT6_17	.49	.86	0	6	1810
TOT1_5	.25	.57	0	3	1744
TOT_LOG	2.02	.75	1	6	1922
NUMVEH	2.15	1.12	0	9	1921
DAY1	2.44	1.15	1	7	1999
DAY2	3.43	1.15	1	7	1999
PANEL	1.16	.37	1	2	2157
STATUS	.36	.80	0	3	1608
HHZIP	98194.73	150.96	98001	98580	2035
HHTRACT	465.48	262.72	1.00	999.75	2035
HHTAZ	496.12	233.62	1	846	2035
FLAG	.04	.38	.00	9.00	2157

Table 12. Descriptive variables of household data in wave 3.

Original Household Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	4892.38	4076.62	4	13210	1696
SUBID	.00	.04	0	1	1696
SAMPLE	1.45	.69	1	3	1695
INCOME	44825.78	32326.29	-9	600000	1062
INCCAT	5.47	2.21	1	11	1660
HHTYPE	4.66	2.58	1	8	1696
HHSIZE	2.54	1.25	1	13	1696
TOTADULT	1.90	.62	1	6	1696
TOT6_17	.46	.85	0	6	1696
TOT1_5	.19	.52	0	6	1696
TOT_LOG	1.98	.71	1	6	1696
NUMVEH	2.23	1.19	0	15	1668
DAY1	2.47	1.16	1	7	1687
DAY2	3.46	1.16	1	7	1687
PANEL	.94	.23	0	1	1696

Rectified Household Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	4892.38	4076.62	4	13210	1696
SUBID	.00	.04	0	1	1696
SAMPLE	1.45	.69	1	3	1695
INCOME	45906.65	31945.72	5000	600000	1037
INCCAT	5.32	2.04	1	10	1616
HHTYPE	4.66	2.58	1	8	1696
HHSIZE	2.54	1.25	1	13	1696
TOTADULT	1.90	.62	1	6	1696
TOT6_17	.46	.85	0	6	1696
TOT1_5	.19	.52	0	6	1696
TOT_LOG	1.98	.71	1	6	1696
NUMVEH	2.23	1.19	0	15	1668
DAY1	2.47	1.16	1	7	1687
DAY2	3.46	1.16	1	7	1687
PANEL	.94	.23	0	1	1696
HHZIP	98190.23	149.70	98001	98580	1696
HHTRACT	457.18	259.79	1.00	928.00	1696
HHTAZ	490.71	231.45	1	831	1696
FLAG	.16	.96	.00	6.00	1696

Table 13. Descriptive variables of household data in wave 4.

Original Household Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	7410.83	5449.29	4	14713	2003
SUBID	.02	.13	0	2	2003
SAMPLE	1.48	.70	0	3	2003
INCOME	48671.49	49475.71	1000	900000	1166
INCCAT	5.42	2.18	1	11	2003
HHTYPE	4.59	2.53	1	8	2003
HHSIZE	2.50	1.26	1	9	2003
TOTADULT	1.87	.63	0	6	2003
TOT6_17	.44	.84	0	6	2003
TOT1_5	.19	.52	0	4	2003
TOT_LOG	1.96	.73	1	6	2003
NUMVEH	2.07	1.09	0	9	2003
DAY1	2.46	1.16	1	6	1932
DAY2	3.46	1.16	2	7	1932
PANEL	.96	.18	0	1	2003

Rectified Household Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	7410.83	5449.29	4	14713	2003
SUBID	.02	.13	0	2	2003
SAMPLE	1.48	.70	1	3	2001
INCOME	48671.49	49475.71	1000	900000	1166
INCCAT	5.34	2.10	1	10	1978
HHTYPE	4.59	2.53	1	8	2003
HHSIZE	2.50	1.26	1	9	2003
TOTADULT	1.87	.63	0	6	2003
TOT6_17	.44	.84	0	6	2003
TOT1_5	.19	.52	0	4	2003
TOT_LOG	1.96	.73	1	6	2003
NUMVEH	2.07	1.09	0	9	2003
DAY1	2.46	1.16	1	6	1932
DAY2	3.46	1.16	2	7	1932
PANEL	.96	.18	0	1	2003
HHZIP	98185.84	149.61	98001	98580	2001
HHTRACT	439.97	260.66	1.00	999.82	2001
HHTAZ	475.11	233.88	1	849	2001
FLAG	.08	.68	.00	6.00	2003

Table 14. Missing values in household data files.

Variables	Wave 1	Wave 2	Wave 3	Wave 4
Total Obs.	1712	2158	1696	2003
SAMPLE			1	
HHSIZE (missing) * 0	 1	237 1		
TOT LOG	1855			
HHTYPE=0		19		
TYPE		245		
TOTADULT	237	243		
DAY1		159	9	71
DATE1		158		
DAY2		159	9	71
DATE2		160		
NUMVEH			26	

Table 15. Code conversions of household data.

	Wave 1		Wave 2		Waves 3 & 4	
Variables	Original Code	Converted Code	Original Code	Converted Code	Original Code	Converted Code
COMPLETE	R	2				
	Y	1				
	N	0				
INCOME	9	-9				
	A/a	9				
	B/b	-9				
PANEL					Y	1
					N	0
					U	-9

In addition to checking trip records within a wave, DAY1 and DAY2 were examined across waves, According to the sample design, each household should have reported trips that were made during the same two consecutive days in all waves. However, several exceptions were found. Those households that did not have two consecutive days or did not maintain the same period across waves are provided in appendix D.

Data Errors From Observations

Logical relationships among the variables were used to scrutinize the data validity. The following "principles" were assumed to hold for each observation:

(a) Household size equals the sum of the number of adults who are 18 years or older, children whose ages are between 6 and 17, and children who are 5 years and younger.

(b) The total number of travel logs is greater than or equal to the number of adults.

The following three types of errors violate the above relations.

1. When household size is 1 (HHSIZE = l), TOTADULT, TOT6-17, TOTl-5, and TOT-LOG were coded as missing values in wave 1. This coding is incorrect because information about the number of adults and children is known. Therefore, when HHSIZE = 1, these four variables are recoded as follows:

   TOTADULT = 1
   TOT6-17 = 0
   TOT13 = 0
   TOT-LOG = 1

2. When relation (a) does not hold (i.e., HHSIZE + TOTADULT +TOTl-5 + TOT6-17): If HHTYPE = 8 (i.e., a household with two or more adults 65+ and without any child) and TOTADULT, TOTl-5, TOT6-17, and TOT-LOG were missing, then these variables were assigned the following values:

   TOTADULT = 2
   TOTl-5 = 0
   TOT6-17 = 0
   TOT-LOG = 2

When HHSIZE < TOTADULD+TOTI-5 +TOT6-17, there are two possible errors. One is that people may count a 15+ as an adult. Then, the total number of adults is calculated as the difference between the household size and the number of people age 17 or younger (i.e., TOTADULT = HHSIZE - TOTl-5 - TOT6-17). The other possible error may be due to miscalculation regarding household size (e.g., excluding the children under 5 years in the count). Then, household size would be the sum of the total number of adults and the number of people age 17 or younger (i.e., HHSIZE = TOTADULT+TOTl_S+TOT6J7). There may be instances when this cannot be identified. If the data verification does not lead to any acceptable conclusion, a missing value is entered.

3. When relation (b) does not hold (i.e., TOT-LOG > HHSIZE): There were four such cases. When checking with trip data and person data, no trip data were found for these cases. Thus, TOT-LOG was assigned a zero.

DESCRIPTION OF RECTIFIED HOUSEHOLD DATABASES

Three additional variables were added to the original database. The variables were zip code, Census Tract, and TAZ of household residential locations. These location variables were obtained by geocoding the household home addresses and matching them with the regional zip code, Census Tract, and TAZ maps by PSRC. Demographics and travel information at the Census Tract and sometimes the TAZ levels can be obtained and, thus, attached to the household files.

Descriptive summary statistics for the rectified household database are reported in tables 10 through 13.

PERSON RECORD DATABASES

Person information was collected for every member who was 15 years and older. The information included age, sex, employment status, occupation, typical mode to work/school, driver’s license, etc. Person data were collected simultaneously with the household data. However, some persons did not have any so&demographic information in the corresponding household files. Table 16 gives the number of persons and aggregated households that were in the person data file and not in the household data file. These persons' HHID, SUBID, and PERS are provided in appendix E.

Table 16. Number of households and persons in the person data file.

Wave	Number of persons	Number of persons not in household file	Number of households	Number of households not in household file
Wave 1 (1989)	3069	0	2944	0
Wave 2 (1990)	4631[1]	268	2285	128
Wave 3 (1992)	3360	2	1697	2
Wave 4 (1993)	4081	162	2101	98

[1]The following observations have duplicate HHID, SUBID, and PERS. Some of them have partially different or completely different information:

HHID	SUBID	PERS
197	0	2
5210	0	1
5210	0	2
215	0	1
6215	0	2
6219	0	1
6219	0	2

To simplify the data manipulation, person data coding containing characters was replaced by numeric values as follows:

Y = 1

N = O

u = -9

D = -9

' = -9 (Character period (.), which denoted missing values, was changed to -)

In addition, alphabetic values of SEX were converted to numeric values as follows:

F = O

M = 1

Like household data, some variables were coded with unexplained values. Stars (*) **were found in the field of WK-CITY in wave 2. Characters 'A' and 'B' were also used** for **work and school modes** in **wave 1.** Table 17 lists the number of records that have unknown values.

In some cases, it was found that variables were coded as both numbers and letters. The person-data description file explained that several variables are coded 1, 2, and 3, which indicated "Yes," "No," and "DK/refuses," respectively. However, LICENSE, BUSPASS, and STUDENT are coded as Y, N, and 3. BUS_FREQ of a person (HHID=2,519, SUBID=O, and **PERS** =2) is 200, which seems highly unlikely, although possible. BUS_FREQ is sometimes coded as 999, which is not explained in the description file. In this case, 999 and 200 are assumed to be missing codes. Complete code conversion is provided in table 18.

Data coding errors in SUBID, PERS, and OLDPER are also found in waves 2 and 3. For instance, in wave 2, two records have the same HHID=197, SUBXD=O, PERS=2, gender, and age. One has full information, while the other has several empty fields. It is, therefore, assumed that the latter is a duplicate of the former and, thus, is removed from the person file. Another example is that a record (HHID=10,218, SUBID=O, PERS=O) does not have a proper PERS code and that most fields have missing values. With the information from the corresponding household file, it turns out that this record is redundant and, thus, is deleted from the person file. However, due to limited information, not all errors are able to be corrected. For example, the following five records are found in wave 2:

HHID	SUBID	PERS	OLDPER	SEX	AGE
6215	0	1	-9	1	30
6215	0	1	-9	0	30
6215	0	2	-9	0	48
6215	0	2	-9	-9	30
6215	0	3	-9	1	25

The corresponding household file shows that the household has only two persons, which implies that the person records of both persons 1 and 2 have been duplicated and that person 3 should not be included. If the household information is correct, then the gender of

Table 17. Missing information in person data.

Variables	Type of Errors	Wave1	Wave2	Wave3	Wave4
OCC	13 17	4 1	1		
WK_MODE1	A 0 9	8 7 58	 6 1		
WK_MODE2	0 9	39 1	 28		
WK_MODE3	0	1	3		
SCH_CITY	***		1		
SCH_MOD1	A B 0 9	1 34 3 46	 1		
SCH MOD2	0	14	6		
SCH_MOD3	0	1			
WK_NUM	0		7		
SCH_NUM	0		1		
BUS_FREQ	200 999	1 35			

Table 18. Code conversions of person data.

Variables	Wave 1		Wave 2		Wave 3		Wave 4	
	Original Code	Converted Code	Original Code	Converted Code	Original Code	Converted Code	Original Code	Converted Code
EMPLOY	3	-9	3	-9				
WK_BUS WK_POOL CAR_FREQ CAR_CHLD STUDENT	3	-9					2 3	0 -9
AGE	99	-9	99	-9			0, 99	-9
AGE_GP	8	-9	8	-9			8	-9
OCC	13,14,17	-9	13,14	-9				
WK_MODE1	0,8,9	-9	0,9	-9				
WK_MODE2	0,8,9	-9	0	-9				
WK_MODE3	0	-9	0	-9				
WK_NUM SCH_NUM	5	-9	0,5	-9			5	-9
SCH_MOD1	0,8,9	-9	8,9	-9			14	-9
SCH_MOD2	0	-9	0	-9				
SCH_MOD3	0	-9						
BUS_FREQ	200,999	-9						
LICENSE BUSPASS	3	-9	3	-9				
OCCCHG			2	0			2	0
MDCHG PLCHG			2	0				

person 1 and the age of person 2 are indecisive. This is problematic because we cannot decide which record is a duplicate. However, it is also possible that the person file is correct after assigning each record a unique person ID. This would require a correction of HHSIZE in the household file.

Some data fields across waves are investigated to ensure data consistency from time to time. A few variable definitions were changed between waves. For instance, in waves 1 and 2, EMPLOY is defined as a binary variable indicating whether a person was employed outside the home. In waves 3 and 4, however, EMPLOY is defined as employment status such as employed, student, and neither employed or student. Mode to work (WK_MODE1 , . . . , WK_MODE4) and mode to school (SCH_MOD1 , . . . ,SCH_MOD4) are also coded differently in the first two waves than in the last two waves. AGE and AGE_GP are checked across waves. It was found that some people's ages were not consistent with time. Appendix F provides these records.

Some data fields were reexamined using information from the survey forms by PSRC staff in the data cleaning process. Two such variables that are related to work are added to the original database. One variable is the primary mode to work (MODE); which is equivalent to WK_MODEl. The other is whether or not a person worked at home, which is similar to EMPLOY. These two variables are more complete and accurate than the original variables. More important, they are coded consistently across the waves, which makes analysis or comparisons meaningful. However, WK_MODE1 and EMPLOY are included in the rectified database to maintain the integrity of the original database. Similar to the household files, zipcode and Census Tract of a person's workplace are attached to the person files. Summary statistics for the original and the corrected four waves are reported in tables 19 through 22.

ATTITUDE AND VALUES DATABASES

Three attitude and values surveys were conducted among the panel members who were employed (full or part time) or in school at the time of the survey. The first two surveys were collected in 1990 and 1991, shortly after the travel diary surveys in the fall of 1989 and 1990. The third survey was in 1993, but without any travel diary survey. Attitude and values

Table 19. Descriptive variables of persons in wave 1.

Original Person Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	2299.57	1317.14	1	4210	6096
SUBID	.00	.00	0	0	6096
PERS	1.69	.81	1	7	6096
OLDPER	Variable is missing for every case				
SEX	.48	.50	0	1	6096
AGE	41.61	18.00	15	99	6096
AGE_GP	4.04	1.60	1	8	6096
EMPLOY	.69	.47	0	3	6096
OCC	4.80	3.34	1	17	4160
WK_CITY	293.87	320.05	2	999	4158
WK_FREQ	4.91	.95	1	8	4158
WK_MODE1	1.44	1.32	0	9	4161
WK_MODE2	2.79	2.95	0	9	77
WK_MODE3	5.33	2.73	0	7	6
WK_MODE4	1.00		1	1	1
WK_NUM	1.32	.79	1	5	3422
WK_BUS	.03	.18	0	3	2842
WK_POOL	.11	.35	0	3	2842
CAR_REQD	.43	.53	0	3	3237
CAR_CHLD	.24	.45	0	3	3236
CAR_FREQ	3.79	1.55	1	8	748
STUDENT	.16	.39	0	3	6096
SCHOOL	306.36	300.46	1	999	922
SCH_CITY	300.94	302.49	2	999	922
SCH_MOD1	1.89	1.82	0	9	919
SCH_MOD2	2.86	3.04	0	7	28
SCH_MOD3	.00		0	0	1
SCH_MOD4	Variable is missing for every case.				
SCH_NUM	1.62	.99	1	5	577
BUS_FREQ	7.07	75.54	0	999	6090
BUSPASS	.53	.72	0	3	1084
LICENSE	.91	.30	0	3	6093

Rectified Person Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	2299.57	1317.14	1	4210	6096
SUBID	.00	.00	0	0	6096
PERS	1.69	.81	1	7	6096
OLDPER	Variable is missing for every case				
SEX	.48	.50	0	1	6096
AGE	40.06	15.54	15	95	5936
AGE_GP	4.02	1.58	1	7	6072
EMPLOY	.69	.46	0	1	6085
OCC	4.72	3.25	1	12	4125
WK_CITY	293.87	320.05	2	999	4158
WK_FREQ	4.91	.95	1	8	4158
WK_MODE1	1.33	.95	1	7	4092
WK_MODE2	5.57	.96	3	7	37
WK_MODE3	6.40	.89	5	7	5
WK_MODE4	1.00		1	1	1
WK_NUM	1.31	.76	1	4	3412
WK_BUS	.03	.17	0	1	2841
WK_POOL	.10	.30	0	1	2831
CAR_REQD	.41	.49	0	1	3216
CAR_CHLD	.23	.42	0	1	3225
CAR_FREQ	3.79	1.55	1	8	748
STUDENT	.15	.36	0	1	6080
SCHOOL	306.36	300.46	1	999	922
SCH_CITY	300.94	302.49	2	999	922
SCH_MOD1	1.84	1.59	1	7	868
SCH_MOD2	5.71	1.27	2	7	14
SCH_MOD3	Variable is missing for every case.				
SCH_MOD4	Variable is missing for every case.				
SCH_NUM	1.59	.95	1	4	572
BUS_FREQ	1.31	3.46	0	50	6054
BUSPASS	.42	.49	0	1	1036
LICENSE	.91	.29	0	1	6082
CHANGE	.00	.00	0	0	2240
HOME	.02	.13	0	1	2240
MODE	1.54	.98	1	6	2240
WKZIP	98182.26	139.82	98001	98597	2173
WKTRACT	372.23	273.99	-8.00	999.99	2224
FLAG	.04	.55	.00	7.00	6096

Table 20. Descriptive variables of persons in wave 2.

Original Person Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	3677.20	3046.67	1	10356	4633
SUBID	.02	.14	0	2	4633
PERS	1.67	.80	0	6	4633
OLDPER	.69	1.32	0	5	256
SEX	.48	.50	0	1	4131
AGE	42.78	16.43	0	99	4138
AGE_GP	4.52	1.84	1	8	4147
EMPLOY	.66	.47	0	1	4218
OCC	3.86	3.08	0	13	2701
WK_CITY	282.34	313.95	2	999	2686
WK_FREQ	4.84	.88	0	8	2687
WK_MODE1	1.31	.93	0	9	2730
WK_MODE2	2.41	2.35	0	6	98
WK_MODE3	2.89	2.80	0	6	9
WK_MODE4	6.00		6	6	1
WK_NUM	1.60	1.25	0	5	2598
WK_BUS	.11	.32	0	3	2437
WK_POOL	.17	.39	0	3	2439
CAR_REQD	.36	.49	0	3	2536
CAR_CHLD	.21	.42	0	3	2538
CAR_FREQ	.91	1.76	0	7	2148
STUDENT	.09	.29	0	1	4132
SCHOOL	333.16	300.05	1	789	380
SCH_CITY	327.24	305.46	2	999	381
SCH_MOD1	2.69	2.40	1	9	380
SCH_MOD2	3.39	2.74	0	7	38
SCH_MOD3	1.80	.45	1	2	5
SCH_MOD4	1.00	.00	1	1	2
SCH_NUM	2.76	1.63	0	5	340
BUS_FREQ	1.05	3.09	0	40	4012
BUSPASS	.06	.25	0	3	4134
LICENSE	.93	.28	0	3	4143
OCCCHG	1.84	.37	1	2	2010
MDCHG	1.00	.00	1	1	378
PLCHG	1.00	.00	1	1	324
CONT	1.00	.00	1	1	3414

Rectified Person Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	3676.54	3045.38	1	10356	4631
SUBID	.02	.14	0	2	4631
PERS	1.67	.80	1	6	4631
OLDPER	.69	1.32	0	5	256
SEX	.48	.50	0	1	4130
AGE	42.56	16.05	15	95	4119
AGE_GP	4.25	1.62	1	7	4137
EMPLOY	.66	.47	0	1	4216
OCC	3.86	3.08	0	12	2700
WK_CITY	282.34	313.95	2	999	2686
WK_FREQ	4.84	.88	0	8	2687
WK_MODE1	1.31	.92	1	6	2723
WK_MODE2	3.37	2.11	1	6	70
WK_MODE3	4.33	2.25	1	6	6
WK_MODE4	6.00		6	6	1
WK_NUM	1.31	.78	1	4	2384
WK_BUS	.11	.31	0	1	2435
WK_POOL	.17	.37	0	1	2436
CAR_REQD	.36	.48	0	1	2533
CAR_CHLD	.21	.40	0	1	2535
CAR_FREQ	.91	1.76	0	7	2148
STUDENT	.09	.29	0	1	4132
SCHOOL	333.16	300.05	1	789	380
SCH_CITY	327.24	305.46	2	999	381
SCH_MOD1	2.65	2.36	1	7	377
SCH_MOD2	4.03	2.51	1	7	32
SCH_MOD3	1.80	.45	1	2	5
SCH_MOD4	1.00	.00	1	1	2
SCH_NUM	1.96	1.06	1	4	248
BUS_FREQ	1.05	3.09	0	40	4012
BUSPASS	.06	.24	0	1	4132
LICENSE	.93	.26	0	1	4132
OCCCHG	.16	.37	0	1	2010
MDCHG	1.00	.00	1	1	377
PLCHG	1.00	.00	1	1	323
CONT	1.00	.00	1	1	3413
CHANGE	.84	.63	0	2	2601
HOME	.03	.17	0	1	2601
MODE	1.65	1.22	1	6	2601
WKZIP	98180.24	141.85	98001	98597	2525
WKTRACT	377.05	276.08	-8.00	999.99	2584
FLAG	.01	.29	.00	7.00	4631

Table 21. Descriptive variables of persons in wave 3.

Original Person Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	4831.64	4028.37	4	13210	3360
SUBID	.00	.05	0	1	3360
PERS	1.67	.85	1	9	3360
OLDPER	.01	.16	0	3	837
SEX	.47	.50	0	1	3356
AGE	45.16	16.14	15	93	3337
AGE_GP	4.49	1.62	1	7	3347
EMPLOY	1.56	.85	1	3	3353
OCC	3.82	3.01	1	12	2068
WK_CITY	277.03	306.77	2	999	2283
WK_FREQ	4.76	.92	0	7	2072
WK_MODE1	1.30	1.03	1	12	2058
WK_MODE2	4.92	3.63	1	13	134
WK_MODE3	5.14	4.00	1	11	14
WK_MODE4	Variable is missing for every case.				
WK_NUM	1.43	.74	1	5	1851
WK_BUS	.09	.28	0	1	2543
WK_POOL	.29	.45	0	1	2562
CAR_REQD	.26	.44	0	1	2483
CAR_CHLD	.16	.37	0	1	2480
CAR_FREQ	1.39	1.99	0	6	1022
SCHOOL	327.45	296.78	1	783	354
SCH_CITY	313.41	297.13	2	775	355
SCH_MOD1	2.31	2.31	1	11	259
SCH_MOD2	4.97	2.68	1	10	31
SCH_MOD3	5.75	2.06	3	8	4
SCH_MOD4	Variable is missing for every case.				
SCH_NUM	2.08	1.11	1	4	169
LICENSE	.94	.24	0	1	3291
BUSPASS	.11	.31	0	1	3289
BUS_FREQ	.91	2.81	0	30	3283
OCCCHG	.14	.35	0	1	1597
CONT	.94	.23	0	1	3360
PARKING	2.14	10.91	0	160	2482
DIARY	.91	.29	0	1	3360
DAY1	2.45	1.16	1	7	3339
DAY2	3.45	1.15	1	6	3339

Rectified Person Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	4831.64	4028.37	4	13210	3360
SUBID	.00	.04	0	1	3360
PERS	1.67	.85	1	9	3360
OLDPER	.02	.25	0	4	840
SEX	.47	.50	0	1	3356
AGE	45.16	16.14	15	93	3337
AGE_GP	4.49	1.62	1	7	3347
EMPLOY	.68	.47	0	1	3353
OCC	3.82	3.01	1	12	2068
WK_CITY	277.03	306.77	2	999	2283
WK_FREQ	4.76	.92	0	7	2072
WK_MODE1	1.30	1.03	1	12	2058
WK_MODE2	4.92	3.63	1	13	134
WK_MODE3	5.14	4.00	1	11	14
WK_MODE4	Variable is missing for every case.				
WK_NUM	1.42	.73	1	4	1850
WK_BUS	.09	.28	0	1	2543
WK_POOL	.29	.45	0	1	2562
CAR_REQD	.26	.44	0	1	2483
CAR_CHLD	.16	.37	0	1	2480
CAR_FREQ	1.39	1.99	0	6	1022
STUDENT	.08	.27	.00	1.00	3353
SCHOOL	327.45	296.78	1	783	354
SCH_CITY	313.41	297.13	2	775	355
SCH_MOD1	2.31	2.31	1	11	259
SCH_MOD2	4.97	2.68	1	10	31
SCH_MOD3	5.75	2.06	3	8	4
SCH_MOD4	Variable is missing for every case.				
SCH_NUM	2.08	1.11	1	4	169
BUS_FREQ	.91	2.81	0	30	3283
BUSPASS	.11	.31	0	1	3289
LICENSE	.94	.24	0	1	3291
OCCCHG	.14	.35	0	1	1597
CONT	.94	.23	0	1	3360
PARKING	2.14	10.91	0	160	2482
DIARY	.91	.29	0	1	3360
DAY1	2.45	1.16	1	7	3339
DAY2	3.45	1.15	1	6	3339
CHANGE	.84	.66	0	2	2355
HOME	.05	.22	0	1	2355
MODE	1.84	1.39	1	9	2355
WKZIP	98178.66	140.60	98001	98597	2295
WKTRACT	378.47	275.30	-8.00	999.99	2348
FLAG	.00	.12	.00	4.00	3360

Table 22. Descriptive variables of persons in wave **4.**

Original Person Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	7185.50	5404.21	4	14713	4081
SUBID	.02	.15	0	2	4081
PERS	1.67	.87	1	9	4081
OLDPER	.77	1.33	0	5	196
SEX	.47	.50	0	1	4081
AGE	43.64	16.41	15	94	4058
AGE_GP	4.35	1.64	1	7	4078
EMPLOY	1.69	.89	1	3	3920
OCC	3.76	3.03	1	12	2359
WK_CITY	281.48	308.00	2	999	2328
WK_FREQ	4.78	.96	0	7	2526
WK_MODE1	1.30	.94	1	11	2328
WK_MODE2	3.77	2.77	1	13	163
WK_MODE3	6.03	4.08	1	11	35
WK_MODE4	8.00	2.74	6	11	5
WK_NUM	1.34	.82	1	5	2118
WK_BUS	1.89	.32	1	3	2887
WK_POOL	1.80	.40	1	3	2901
CAR_REQD	1.75	.44	1	3	3081
CAR_CHLD	1.84	.37	1	3	3062
CAR_FREQ	1.24	1.93	0	7	1497
SCHOOL	384.89	339.80	1	849	412
SCH_CITY	299.61	304.46	1	800	412
SCH_MOD1	2.44	2.12	1	14	399
SCH_MOD2	4.76	2.36	1	7	71
SCH_MOD3	5.07	2.50	1	7	14
SCH_MOD4	8.00	2.65	6	11	3
SCH_NUM	1.78	.99	1	4	58
LICENSE	.97	.25	0	2	3904
BUSPASS	.58	.83	0	2	3896
BUS_FREQ	1.09	3.14	0	50	3859
OCCCHG	1.83	.37	1	2	1823
CONT	.91	.29	0	1	4081
PARKING	2.66	13.86	0	250	2453
DIARY	.94	.23	0	1	4081
DAY1	2.46	1.16	1	6	3855
DAY2	3.46	1.16	1	7	3855

Rectified Person Data

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	7185.50	5404.21	4	14713	4081
SUBID	.02	.15	0	2	4081
PERS	1.67	.87	1	9	4081
OLDPER	.77	1.33	0	5	196
SEX	.47	.50	0	1	4081
AGE	43.64	16.41	15	94	4058
AGE_GP	4.35	1.64	1	7	4078
EMPLOY	.60	.49	0	1	3920
OCC	3.76	3.03	1	12	2359
WK_CITY	281.48	308.00	2	999	2328
WK_FREQ	4.78	.96	0	7	2526
WK_MODE1	1.30	.94	1	11	2328
WK_MODE2	3.77	2.77	1	13	163
WK_MODE3	6.03	4.08	1	11	35
WK_MODE4	8.00	2.74	6	11	5
WK_NUM	1.33	.81	1	4	2115
WK_BUS	.11	.31	0	1	2883
WK_POOL	.20	.40	0	1	2899
CAR_REQD	.26	.44	0	1	3072
CAR_CHLD	.16	.37	0	1	3056
CAR_FREQ	1.24	1.93	0	7	1497
STUDENT	.10	.31	.00	1.00	3920
SCHOOL	384.89	339.80	1	849	412
SCH_CITY	299.61	304.46	1	800	412
SCH_MOD1	2.41	2.04	1	7	398
SCH_MOD2	4.76	2.36	1	7	71
SCH_MOD3	5.07	2.50	1	7	14
SCH_MOD4	8.00	2.65	6	11	3
SCH_NUM	1.78	.99	1	4	58
BUS_FREQ	1.09	3.14	0	50	3859
BUSPASS	.36	.48	0	1	1367
LICENSE	.98	.13	0	1	3728
OCCCHG	.17	.37	0	1	1823
CONT	.91	.29	0	1	4081
PARKING	2.66	13.86	0	250	2453
DIARY	.94	.23	0	1	4081
DAY1	2.46	1.16	1	6	3855
DAY2	3.46	1.16	1	7	3855
CHANGE	.81	.70	0	5	2662
HOME	.03	.18	0	1	2662
MODE	1.78	1.52	1	9	2659
WKZIP	98175.16	140.67	98001	99201	2591
WKTRACT	366.28	273.12	-8.00	999.99	2650
FLAG	.00	.00	.00	.00	4081

surveys asked panel members' opinions regarding their general travel behavior and some basic facts of work-related travel behavior. This information was grouped into four categories: (1) importance of performance measures that were related to mobility, accessibility, and comfort; (2) rating of these performance measures by different travel modes such as SOV, car pool, and bus; (3) agreement/disagreement with statements on mode usage; and (4) work commute behavior such as usual travel mode to work, availability of bus and car pool, and car requirement before, during, or after work.

In the 1990 attitude and values data file, information from an additional section and another survey was included, but without any descriptions attached to the description file. To retain as much information as possible, although not necessarily useful at this time, data without any descriptions were also included Attitude data files were not examined for any consistency concerns, except that agreement between data coding and the corresponding data description was checked. A few inconsistent codes were found in wave 1 and are shown in table 23. These codes were not described in the description files and, thus, converted to a missing code, which is -9. No miscodings were found in the wave 2 data file.

For the sake of consistency with household, person, and trip data files, data codes that represented unknown values were converted to missing values. In the 1991 data file, the variable LEAVE (time usually leaving home for work) was coded using a range from 00:00 to 1,259 followed by a.m. or p.m. To facilitate data manipulation (and maintain consistency in PSTP), such a notation was converted to a 24-hour notation, as described earlier in this chapter. Descriptions of the new data file are provided in the PSTP Codebook (appendix A). Summary statistics of the 1990 and 1991 attitude and values data are given in tables 24 and 25, respectively.

Table 23. Missing codings in 1990 attitude and values data file.

Variables	Original Codes
SECB14	8 (1)
SECB18	9 (1)
SECB22, SECB25, SECB29,SECB36, SECB38, SECB39, SECB49	0
DAYS	0 (16), 9 (1)
MODE	21 (1), 22 (2), 23 (1), 30(1), 72(1)
MOST	0 (4)
POOL	8 (2)
HHLDCOMM	5 (1), 7 (1), 8(2)

Table 24. Descriptive variables of 1990 attitude and values data.

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	2302.45	1289.52	1	4209	2477
PERS	1.54	.66	1	5	2477
STATUS	2.14	1.89	1	8	2439
SECA1	6.49	1.06	1	7	1985
SECA2	5.63	1.66	1	7	1974
SECA3	5.79	1.32	1	7	1983
SECA4	5.58	1.42	1	7	1979
SECA5	5.25	1.66	1	7	1974
SECA6	5.42	1.67	1	7	1972
SECA7	5.48	1.76	1	7	1964
SECA8	5.67	1.58	1	7	1954
SECA9	3.01	2.03	1	7	1965
SECA10	6.32	1.15	1	7	1975
SECA11	5.85	1.35	1	7	1975
SECA12	5.23	1.54	1	7	1975
SECA13	5.69	1.69	1	7	1973
SECA14	6.27	1.25	1	7	1979
SECA15	6.34	1.30	1	7	1974
SECA16	5.84	1.37	1	7	1982
SECA17	5.72	1.47	1	7	1983
SECB1	6.46	1.02	1	7	1946
SECB2	4.67	1.68	1	7	1908
SECB3	5.17	1.42	1	7	1897
SECB4	6.73	.83	1	7	1944
SECB5	3.92	1.93	1	7	1903
SECB6	5.41	1.59	1	7	1890
SECB7	4.51	1.86	1	7	1943
SECB8	5.41	1.56	1	7	1901
SECB9	4.72	1.56	1	7	1889
SECB10	5.75	1.48	1	7	1940
SECB11	3.65	1.77	1	7	1896
SECB12	4.87	1.42	1	7	1887
SECB13	4.64	1.74	1	7	1937
SECB14	5.17	1.44	1	7	1894
SECB15	5.19	1.38	1	7	1888
SECB16	6.39	1.14	1	7	1936
SECB17	4.44	1.68	1	7	1900
SECB18	5.86	1.32	1	7	1887
SECB19	6.76	.89	1	7	1924

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
SECB20	4.85	1.47	1	7	1903
SECB21	6.35	1.17	1		1889
SECB22	6.62	1.04	1	7	1922
SECB23	4.00	1.63	1	7	1900
SECB24	5.35	1.33	1	7	1886
SECB25	1.41	1.20	1	7	1894
SECB26	5.59	1.71	1	7	1902
SECB27	4.05	1.90	1	7	1889
SECB28	6.77	.81	1	7	1939
SECB29	3.57	1.74	1	7	1900
SECB30	3.73	1.84	1	7	1890
SECB31	6.70	.89	1	7	1937
SECB32	3.34	1.75	1	7	1901
SECB33	3.17	1.85	1	7	1888
SECB34	6.48	1.10	1	7	1935
SECB35	3.22	1.77	1	7	1899
SECB36	4.69	1.56	1	7	1886
SECB37	5.38	1.81	1	7	1936
SECB38	6.03	1.59	1	7	1850
SECB39	5.43	1.55	1	7	1878
SECB40	5.18	1.62	1	7	1935
SECB41	5.93	1.21	1	7	1894
SECB42	4.91	1.49	1	7	1880
SECB43	6.09	1.25	1	7	1934
SECB44	4.65	1.64	1	7	1900
SECB45	5.81	1.31	1	7	1888
SECB46	4.64	1.79	1	7	1938
SECB47	5.15	1.53	1	7	1904
SECB48	4.81	1.44	1	7	1890
SECB49	3.70	1.92	1	7	1931
SECB50	5.30	1.60	1	7	1898
SECB51	4.75	1.49	1	7	1889
SECC1	4.50	1.48	1	7	1973
SECC2	3.24	1.76	1	7	1971
SECC3	5.25	1.81	1	7	1968
SECC4	2.99	1.76	1	7	1965
SECC5	4.61	1.92	1	7	1960
SECC6	4.60	2.03	1	7	1956
SECC7	3.85	1.78	1	7	1966

Table 24. Descriptive variables of 1990 attitude and values data (cont.).

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
SECC8	5.26	1.62	1	7	1973
SECC9	3.70	2.03	1	7	1967
SECC10	4.28	1.31	1	7	1967
SECC11	2.95	1.61	1	7	1967
SECC12	5.81	1.28	1	7	1972
SECC13	5.38	1.54	1	7	1972
SECC14	3.91	1.67	1	7	1972
SECC15	2.27	1.63	1	7	1971
SECC16	4.76	1.56	1	7	1966
SECC17	4.72	1.68	1	7	1968
SECC18	5.98	1.37	1	7	1968
SECC19	4.04	1.71	1	7	1968
SECC20	4.08	1.51	1	7	1972
SECC21	3.71	1.54	1	7	1973
SECC22	4.28	1.64	1	7	1973
SECC23	3.61	1.56	1	7	1972

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
DAYS	4.87	.89	1	7	1944
MODE	2.10	2.36	1	12	1954
CARREQ	.35	.48	0	1	1950
CAREMPL	.37	.48	0	1	791
ZIP	192.05	153.26	1	993	1694
MOST	2.46	.78	1	3	1906
LEAST	1.43	.69	1	3	1867
BUSSTOP	1.71	.86	1	3	1825
DIRECT	.19	.39	0	1	1036
CONSID	.38	.48	0	1	1019
POOL	.09	.28	0	1	1963
HHLDCOMM	.09	.28	0	1	1845
NBHDCOMM	.10	.30	0	1	1587
WORKCOMM	.26	.44	0	1	1709
FREQCHLD	3.94	1.58	1	5	1903
FREQERRD	2.06	1.12	1	5	1937
FREQPERS	2.86	1.55	1	5	1952
FILLER	This is a String (Alphanumeric) variable.				

Table 25. Descriptive variables of 1991 attitude and values data.

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
HHID	3592.00	2955.89	4	10355	2395
SUBID	.00	.07	0	2	2395
PERS	1.58	.70	1	6	2395
STATUS	2.31	2.01	1	8	2352
SECA1	6.51	.99	1	7	1824
SECA2	5.66	1.60	1	7	1816
SECA3	5.72	1.30	1	7	1822
SECA4	5.66	1.33	1	7	1819
SECA5	5.18	1.62	1	7	1812
SECA6	5.57	1.55	1	7	1819
SECA7	5.62	1.65	1	7	1812
SECA8	5.83	1.44	1	7	1812
SECA9	2.97	2.01	1	7	1817
SECA10	6.35	1.05	1	7	1818
SECA11	5.97	1.26	1	7	1813
SECA12	5.37	1.45	1	7	1813
SECA13	5.66	1.64	1	7	1815
SECA14	6.29	1.20	1	7	1818
SECA15	6.40	1.23	1	7	1818
SECA16	5.83	1.31	1	7	1820
SECA17	5.68	1.40	1	7	1819
SECB1	6.51	.96	1	7	1789
SECB2	4.55	1.67	1	7	1759
SECB3	5.07	1.44	1	7	1749
SECB4	6.77	.79	1	7	1786
SECB5	3.87	1.86	1	7	1750
SECB6	5.42	1.59	1	7	1742
SECB7	4.45	1.86	1	7	1789
SECB8	5.48	1.45	1	7	1757
SECB9	4.73	1.46	1	7	1746
SECB10	5.75	1.44	1	7	1786
SECB11	3.64	1.72	1	7	1751
SECB12	4.78	1.38	1	7	1742
SECB13	4.51	1.73	1	7	1783
SECB14	5.21	1.47	1	7	1752
SECB15	5.25	1.28	1	7	1742
SECB16	6.45	1.11	1	7	1782
SECB17	4.38	1.68	1	7	1755
SECB18	5.85	1.22	1	7	1745
SECB19	6.79	.86	1	7	1777

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
SECB20	4.84	1.45	1	7	1751
SECB21	6.36	1.11	1	7	1743
SECB22	6.67	.98	1	7	1767
SECB23	3.91	1.60	1	7	1748
SECB24	5.31	1.31	1	7	1742
SECB25	1.35	1.10	1	7	1757
SECB26	5.67	1.67	1	7	1745
SECB27	4.16	1.88	1	7	1735
SECB28	6.80	.73	1	7	1780
SECB29	3.49	1.70	1	7	1749
SECB30	3.61	1.82	1	7	1740
SECB31	6.75	.78	1	7	1782
SECB32	3.30	1.72	1	7	1747
SECB33	3.11	1.82	1	7	1741
SECB34	6.57	1.00	1	7	1775
SECB35	3.19	1.72	1	7	1749
SECB36	4.61	1.61	1	7	1742
SECB37	5.46	1.78	1	7	1782
SECB38	6.00	1.66	1	7	1698
SECB39	5.50	1.50	1	7	1733
SECB40	5.09	1.59	1	7	1780
SECB41	5.85	1.27	1	7	1747
SECB42	4.86	1.41	1	7	1739
SECB43	6.12	1.22	1	7	1779
SECB44	4.60	1.61	1	7	1748
SECB45	5.82	1.25	1	7	1742
SECB46	4.64	1.77	1	7	1783
SECB47	5.04	1.56	1	7	1750
SECB48	4.83	1.37	1	7	1742
SECB49	3.58	1.88	1	7	1781
SECB50	5.39	1.60	1	7	1750
SECB51	4.80	1.42	1	7	1741
SECC1	4.32	1.51	1	7	1822
SECC2	3.25	1.75	1	7	1820
SECC3	5.31	1.76	1	7	1821
SECC4	3.00	1.70	1	7	1814
SECC5	4.67	1.91	1	7	1807
SECC6	4.59	2.00	1	7	1816
SECC7	3.83	1.75	1	7	1820
SECC8	5.21	1.61	1	7	1819

Table 25. Descriptive variables of 1991 attitude and values data (cont.).

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
SECC9	3.96	1.99	1	7	1821
SECC10	4.24	1.35	1	7	1818
SECC11	2.96	1.64	1	7	1820
SECC12	5.85	1.26	1	7	1819
SECC13	5.39	1.50	1	7	1819
SECC14	3.82	1.63	1	7	1820
SECC15	2.21	1.58	1	7	1818
SECC16	4.79	1.53	1	7	1815
SECC17	4.75	1.63	1	7	1808
SECC18	5.96	1.33	1	7	1815
SECC19	4.10	1.69	1	7	1815
SECC20	4.02	1.49	1	7	1812
SECC21	3.65	1.53	1	7	1814
SECC22	4.31	1.68	1	7	1814
SECC23	3.70	1.57	1	7	1820
ADDR	This is a String (Alphanumeric) variable.				
CITY	This is a String (Alphanumeric) variable.				
ZIP	97818.53	5643.67	98	98580	1266
DAYS	4.78	.89	1	7	1802
MODE	2.21	2.55	1	12	1806

Variable	Mean	Std Dev	Minimum	Maximum	#Obs.
LEAVE	890.16	313.54	100	2400	2395
TRAVTIME	26.24	20.54	1	300	1756
CARREQ	.35	.48	0	1	1805
CAREMPL	.37	.48	0	1	762
PAYPARK	.20	.40	0	1	1719
FREEPARK	1.49	.84	1	3	1711
MOST	2.53	.75	1	3	1789
LEAST	1.47	.70	1	3	1758
BUSSTOP	1.79	.88	1	3	1654
DIRECT	.20	.40	0	1	846
CONSID	.38	.49	0	1	819
BUSRIDES	1.09	2.94	0	30	1580
PASS	.09	.28	0	1	1710
POOL	.09	.28	0	1	1710
POOLRIDE	2.40	2.02	1	21	1793
HHLDCOMM	.09	.29	0	1	1682
NBHDCOMM	.10	.30	0	1	1444
WORKCOMM	.23	.42	0	1	1541
FREQCHLD	3.99	1.54	1	5	1799
FREQERRD	2.12	1.14	1	5	1812
FREQPERS	2.81	1.55	1	5	1804

3. INTEGRATION OF DATABASES

The trip-based information can be integrated with person and household information for each wave using statistical database software (e.g., SPSS). The key variables to match the three files are HHID, SUBID, and PERS. The resultant database may include household so&demographic information, characteristics of the household members, and out-of-home activity information. The format of this integrated trip-based database is provided in the PSTP Codebook.

In addition to individual trips, a person or a household is often the analytical unit of interest. From the trip-based database one can "aggregate" the data to obtain person- and/or household-based information. Aggregated person and household trip information can be obtained by aggregating individuals' trip-based records based on person and household identification numbers, which is not provided here. At the person or household level, information of each trip (such as trip purpose, mode, beginning/ending time, distance, and origin/destination) will not be maintained. Instead, only 2-day summary statistics can be obtained. The person-based trip database consists of person characteristics, associated household so&demographic information, and aggregated trip information. This includes trip frequencies, durations, and travel time of each trip purpose by each travel mode. Household-based trip information can be obtained in a similar way, either from the integrated trip-based database or from the aggregated person-based database. For instance, a household database could include all the so&demographics and the aggregated household-based trip information, which may be the total number of daily trips, the average distance per trip, the average activity time (duration) per activity, the average travel time per trip, and so on. Tables 26 and 27 list the aggregated trip information at person and household levels, respectively.

Table 26. Aggregated trip data at the person level.

	Wave 1	Wave 2	Wave 3	Wave 4
Avg. # of trips	4.98	4.87	4.59	4.63
Avg. trip length	6.56	6.71	7.08	7.09
Avg. activity duration	112.2	108.6	110.8	109.9
Avg. travel time	19.5	19.6	19.6	19.6

Table 27. Aggregated trip data at the household level.

	Wave 1	Wave 2	Wave 3	Wave 4
Avg. # of trips	9.88	9.41	8.78	8.79
Avg. trip length	12.98	12.95	13.55	13.47
Avg. activity duration	222.3	209.7	212.1	208.7
Avg. travel time	38.6	37.9	37.6	37.2

4. CONCLUSIONS AND RECOMMENDATIONS

Household, person, and trip data of the first four waves of PSTP were examined during this project. Data inconsistencies were found, however, the inconsistencies are not introducing substantial bias into the PSTP data sets. Two types of errors were found: namely, data coding errors and errors from observations. Data coding errors are those emerging from the data coding procedure. This includes unexpected codes that were not defined in the description files, incorrect numeric coding, and duplicate records. The incorrect numeric coding may be partially due to unclear records from the observations. Errors from observations are the result of inconsistent information provided by the respondents. This category includes missing variable values, discontinuity in origin/destination sequences, time inconsistency, numeric errors, and out-of-order trip sequences.

Based on the problems encountered during this data cleaning process, it is recommended that numeric data be used instead of characters. If characters are used, then numeric numbers and characters should not be used for the same variable because it is much easier to deal with numeric variables than character variables for most of the statistical/business software. When data are missing, indifferently of the data types, they should be denoted by blanks or some unique numeric values. In the data files, a period (.) is used for missing numeric values, while a blank (no symbol) is used for missing character values. This reenforces the suggestion above to avoid “mixing“ character data with numeric data.

During the data cleaning process, some issues regarding data coding arise. These issues include the following.

How should should starting/ending times from observations be co&d?

In the trip data files, it seems that there were two ways to mark the errors when starting/ending time or duration of a trip was not correct. One way was to record the time by adding or subtracting 12 hours, depending on the situation. For instance, if 2,190 was the correct time but not for sure, it would be coded as 9:00, or the other way around. The other way was to denote trip duration as 301 when time was not consistent. This coding method is a good way to represent the original data, but it causes substantial confusion in data analysis First, it is' redundant because two variables (starting/ending time and trip duration) are used to mark errors. No information would be lost if duration is coded as 301 while trip starting/ending time is corrected to reflect the might-be situation. Second, if one wants to use information of trip starting/ending time, one has to correct the miscoding information to have correct input for data analysis. Correcting trip times is time consuming. Therefore, it is recommended that starting/ending time be corrected based on recorders' knowledge. If there is doubt about the credibility of trip time or errors in trip times, trip duration (MINUTES) marked by a special code (e.g., 301) to indicate the uncertainty of trip time and the starting and ending times should be correctly coded.

What is the time span considered for trip recording?

The survey form of the PSTP travel diary defined 1:00 a.m. as the cutoff point of trip records without explicitly defining this to the respondents. This unclear survey period also causes confusion. Indeed, in the trip data files the period covered exceeds 48 hours. Some trips started several hours before the target day, and some ended a few hours after 190 a.m. the following day. Consider the following example:

PURP.	MODE	BEGT.	ENDT.	MIN.	DIST.
1	1	2215	2230	15	8.40
2	1	730	800	30	10.40
8	1	830	845	15	1.70

It seems that this person worked at night, went shopping the next day, and then went home. In a 24-hour interval with a cutoff point at 190 a.m., the first trip should be discarded. An additional trip might be added if the person's typical work schedule started at night. Neither of the two ways makes any difference in the traditional travel demand forecasting. However, it does make a difference when activity patterns are the subject of the interest. If this person is the one who usually works at night, his/her activity pattern would be shopping, home, and work, instead of work, shopping, and home. For duration analysis, recording the beginning/ending time of an activity eliminates left and right censoring. **'Therefore,** it' is crucial to explicitly define the time span of trip records and discard trips made beyond the time span. Alternatively, it could allow for longer time spans to capture starting and ending times that go beyond the 48-hour window for all respondents.

According to the implicitly defined rule of coding time, which was discussed previously, this trip sequence might be converted to the following trip sequence:

PURP.	MODE	BEGT.	ENDT.	MIN.	DIST.
1	1	1015	1030	15	8.40
2	1	1930	2000	30	10.40
8	1	2030	2045	15	1.70

Such time conversion is possible, but highly unlikely. Nevertheless, it is very confusing (this may be another reason to justify the argument made in the previous question).

Is there any missing information on activity time?

Frequently, in the data set the following sequence of trips appears:

PURPOSE	BEGTIME	ENDTIME	DISTANCE
6	1005	1020	3.7
7	1020	1025	0.8
5	1320	1330	1.6
8	1330	1340	2.2

This **sequence** implies that there is no time for engaging in the activity. A possible situation that might cause this problem is that the duration of the activity is so short that the activity time **cannot** be recorded (less than 1 minute) or that it might be ignored by the person. For instance, it may take a person less than a minute to stay in a drive-through bank. Here, the duration of activity can be assigned the value of 1 minute. However, it is also possible that people ignore short periods when reporting their trip sequences, even if the period may be longer than a minute. Short durations of activity participation may not be relevant for tradition& transportation planning models. Exactly the opposite is true when the data are used for activity analysis where short activity durations are equally important as long activity durations.

5. REFERENCES

Fitzroy, S. S. "Puget Sound Transportation Panel Four Waves." Presented at Travel **Model** Improvement Program Workshop, Fort Worth, Texas, August 14-17, 1994.

Murakami, E., and C. Ulberg (1992). ***Current Status*** *of the* ***Puget Sound Transportation*** Panel. Unpublished manuscript.

Murakami, **E.,** and **W. T.** Watterson (1990). "Developing a Household Travel Panel Survey for the Puget Sound Region." ***Transportation Research Record 1285, pp. 40-46.***

Murakami, E., and W. T. Watterson (1993). "The Puget Sound Transportation Panel after Two Waves." ***Transportation Research Record 27A, pp. 447-492.***

APPENDIX A

PSTP Codebook[1]

A.1 Wave 1 Household Data for All Households Contacted

This file contains all households originally contacted for the study, including those who declined to participate in the panel and those who did not complete travel diaries.

NUMBER OF OBSERVATIONS: 5 175
NUMBER OF VARIABLES: 20
RECORD LENGTH: 51

# VARIABLE	LENGTH	BEGIN COLUMN	LABEL
1 HHID	5	1	Household ID [1]
2 SUBID	1	6	Household split ID
3 COMPLETE	1	7	Completed travel diaries
4 SAMPLE	2	8	Sample category
5 INCOME	2	10	Household income category
6 HHTYPE	2	12	Life cycle
7 HHSIZE	2	14	Household size
8 TOTADULT	2	16	Number of adults (18+)
9 TOT6_17	2	18	Number of children (6-17)
10 TOT1_5	2	20	Number of children (<6)
11 TOT_LOG	2	22	Number in household 15+ for travel log
12 NUMVEH	2	24	Number of household vehicles
13 DAY1	2	26	Day-of-week of first day dairy
14 DATE1	3	28	Date (MDD) of first day diary
15 DAY2	2	31	Day-of-week of second day dairy
16 DATE2	3	33	Date (MDD) of second day diary
17 HHZIP	5	36	Zip code
18 HHTRACT	6 (F6.2)	41	Census tract
19 HHTAZ	3	47	Traffic analysis zone
20 FLAG	2	50	Correction indicator

NOTES:

[1] The participating households (whether or not they completed diaries) were assigned ID numbers from 1 through the 4,000s; the non-participating households (value of "2" in COMPLETE field) were assigned numbers beginning at 5000. These last household ID numbers should not be confused with the ID numbers assigned to the wave 2 replacement households, which also begin at 5000.

[1] -9 is the missing code.

COMPLETE

0 = The participating household did not return wave 1 travel diaries
1 = The participating household returned wave 1 travel diaries
2 = The contacted household declined to participate in the panel study

SAMPLE

1 = SOV
2 = Bus
3 = Carp001

INCOME

1 = < $7,500
2 = $7,500-15,000
3 = $15-25,000
4 = $25-30,000
5 = $30-35,000
6 = $35-50,000
7 = $50-70,000
8 = $70,000 +
0 = < $30,000
9 = $30,000+

HHTYPE (lifecycle)

1 = Any child < 6
2 = All children 6-17
3 = 1 adult, < 35
4 = 1 adult, 35-64
5 = 1 adult, 65 +
6 = 2+ adults, < 35
7 = 2 + adults, 35-64
8 = 2+ adults, 65 +

DAY l-DAY2

1 = Monday
2 = Tuesday
3 = Wednesday
4 = Thursday
5 = Friday
6 = Saturday
7 = Sunday

DATEl-DATE2

125-129 = September
202-231 = October
301-330 = November
401421 = December

FLAG

- 0 = No change has been made
- 1 = Assign TOT6_17 = 0 and TOT1_5=0 when HHSIZE = 1, TOTADULT = 1, and both TOT6_17 and TOT1_5 are missing
- 6 = HHTYPE has been changed according to TOTADULT, TOT6_17, and TOT1_5
- 7 = HHSIZE has been changed
- 9 = TOT6_17 has been changed
- 10 = TOT1_5 have been changed

A.2 Wave 2 Household Data

This file contains all households who continued with the survey from wave 1 (1989) as well as replacement and new recruit households.

NUMBER OF OBSERVATIONS: 2157
NUMBER OF VARIABLES: 21
RECORD LENGTH: 48

#VARIABLE	LENGTH	BEGIN COLUMN	LABEL
1 HHID	5	1	Household ID
2 SUBID	1	6	Household split ID[l]
3 SAMPLE	2	7	Sample category
4 INCOME	2	9	Household income category
5 HHTYPE	2	11	Lifecycle
6 HHSIZE	2	13	Household size
7 TOTADULT	2	15	Number of adults (18 +)
8 TOT6 17	2	17	Number of children (6-17)
9 TOTl-5	2	19	Number of children (< 6)
10 TOT -LOG	2	21	Number in household 15 + for travel log
11 NUMVEH	2	23	Number of household vehicles
12 DAY1	2	25	Day-of-week of first day dairy
13 DAY2	2	27	Day-of-week of second day, dairy
14 PANEL	2	29	Panel status [2]
15 STATUS	2	31	Household attributes status
16 HHZIP	5	33	Zip code
17 HHTIWCT	6 (F6.2)	38	Census tract
18 HHTAZ	3	44	Traffic analysis zone
19 FLAG	2	47	Correction indicator

NOTES:

[1] SUBID indicates that a member/members of a core household (either original or replacement) who has/have a different residence in current wave. The first split will have a SUBID = 1. If another member of the original household moves out (though remains in the panel), that new household will have a SUBID = 2, and so on.

[2] Used to determine whether or not a household can be considered a continuation household for wave 3, i.e., there are household person data and some or all of the panel members within the household return travel diaries.. Replacement or new panel households can be identified as those with HHIDs greater than 5,000.

SAMPLE
- 1 = SOV
- 2 = Bus
- 3 = Carp001

INCOME

- 1 = c $7,508
- 2 = $7,500-15,000
- 3 = $15-25,000
- 4 = $25-30,000
- 5 = $30-35,000
- 6 = $35-50,ooo
- 7 = $50-70,ooo
- 8 = $7O,ooo+
- 0 = < $30,000
- 9 = $30,000+

HHTYPE (lifecycle)

- 1 = Any child < 6
- 2 = All children 6- 17
- 3 = 1 adult, < 35
- 4 = 1 adult, 35-64
- 5 = 1 adult, 65 +
- 6 = 2+ adults, < 35
- 7 = 2 + adults, 35-64
- 8 = 2+ adults, 65+

DAYl-DAY2

- 1 = Monday
- 2 = Tuesday
- 3 = Wednesday
- 4 = Thursday
- 5 = Friday
- 6 = Saturday
- 7 = Sunday

DATEl-DATE2

- 101-131 = October
- 201-229 = November
- 303-318 = December

FLAG

- 0 = No change has been made
- 1 = Assign TOT6-17 = 0 and TOTl-5 =0 when HHSIZE = 1, TOTADULT = 1, and both TOT6-17 and TOTl-5 are missing
- 6 = HHTYPE has been changed according to TOTADULT, TOT6-17, and TOTl-5
- 7 = HHSIZE has been changed
- 9 = TOT6-17 has been changed
- 10 = TOTl-5 have been changed

A.3 Wave3 Household Data

This file contains all households who continued with the survey from wave 2 (1990) as well as replacement and new 'recruit households.

NUMBER OF OBSERVATIONS: 1696
NUMBER OF VARIABLES: 19
RECORD LENGTH: 52

#VARIABLE	LENGTH	BEGIN COLUMN	LABEL
1 HHID	5	1	Household ID
2 SUBID	1	6	Household split ID[l]
3 SAMPLE	2	7	Sample category
4 INCOME	6	9	Household income
5 INCCAT	2	15	Household income category
6 HHTYPE	2	17	Lifecycle
7 HHSIZE	2	19	Household size
8 TOTADULT	2	21	Number of adults (18 +)
9 TOT6 17	2	23	Number of children (6-17)
10 TOT1-5	2	25	Number of children (< 6)
11TOTLOG	2	27	Number in household 15 + for travel log
12NUMVEH	2	29	Number of household vehicles
13 DAY1	2	31	Day-of-week of first day dairy
14 DAY2	2	33	Day-of-week of second day dairy
15 PANEL	2	35	Panel status [2]
16 HHZIP	5	37	Zip code
17 HHTRACT	6 (F6.2)	42	Census tract
18 HHTAZ	3	48	Traffic analysis zone
19 FLAG	2	51	Correction indicator

NOTES:

[1] SUBID indicates that a member/members of a core household (either original or replacement) who has/have a different residence in current wave. The first split will have a SUBID = 1. If another member of the original household moves out (though remains in the panel), that new household will have a SUBID=2, and so on.

[2] Used to determine whether or not a household can be considered a continuation household for wave 4, i.e., there are household person data and some or all of the panel members within the household return travel diaries. Replacement or new panel households can be identified as those with HHIDs greater than 11,000.

SAMPLE

1 = SOV
2 =. Bus
3 = Carp001

INCCAT

1 = < $10,000
2 = $l0,000-15,000
3 = $15-25,000
4 = $25-35,000
5 = $35-45,000
6 = $45-55,000
7 = $55-75,000
8 = $75,000+
9 = < $35,000
10 = $35,000+

HHTYPE (life cycle)

1 = Any child < 6
2 = All children 6-17
3 = 1 adult, < 35
4 = 1 adult, 35-64
5 = 1 adult, 65 +
6 = 2+ adults, < 35
7 = 2+ adults, 35-64
8 = 2+ adults, 65 +

DAYl-DAY2

1 = Monday
2 = Tuesday
3 = Wednesday
4 = Thursday
5 = Friday
6 = Saturday
7 = Sunday

FLAG

0 = No change has been made
1 = Assign TOT6-17 = 0 and TOTl-5 =0 when HHSIZE = 1, TOTADULT = 1, and both TOT6-17 and TOTl-5 are missing
6 = HHTYPE has been changed according to TOTADULT, TOT6-17, and TOTl-5
7 = HHSIZE has been changed
9 = TOT6-17 has been changed
10 = TOTl-5 have been changed

A.4 Wave 4 Household Data

This file contains all households who continued with the survey from wave 3 (1992) as well as replacement and new recruit households.

NUMBER OF OBSERVATIONS: 2003
NUMBER OF VARIABLES: 19
RECORD LENGTH: 52

#VARIABLE	LENGTH	BEGIN COLUMN	LABEL
1 HHID	5	**1**	Household ID
2 SUBID	1	6	Household split ID[l]
3 SAMPLE	2	7	Sample category
4 INCOME	6	9	Household income
5 INCCAT	2	15	Household income category
6 HHTYPE	2	17	Lifecycle
7 HHSIZE	2	19	Household size
8 TOTADULT	2	21	Number of adults (18 +)
9 TOT6 17	2	23	Number of children (6-17)
10 TOTl- 5	2	25	Number of children (< 6)
11 TOT LOG	2	27	Number in household 15+ for travel log
12 NUMVEH	2	29	Number of household vehicles
13 DAY1	2	31	Day-of-week of first day dairy
14 DAY2	2	33	Day-of-week of second day dairy
15 PANEL	2	35	Panel status [2]
16 HHZIP	5	37	Zip code
17 HHTRACT	6 (F6.2)	42	Census tract
18 HHTAZ	3	48	Traffic analysis zone
19 FLAG	2	51	Correction indicator

NOTES:

[1] SUBID indicates that a member/members of a core household (either original or replacement) who has/have a different residence in current wave. The first split will have a SUBID= 1. If another member of the original household moves out (though remains in the panel), that new household will have a SUBID =2, and so on.

[2] Used to determine whether or not a household can be considered a continuation household for wave 5, i.e., there are household person data and some or all of the panel members within the household return travel diaries. Replacement or new panel households can be identified as those with HHIDs greater than 13300.

SAMPLE

1 = SOV
2 = Bus
3=Carpool

INCCAT

1 = < $l0,000
2 = $lO,OOO-15,000
3 = $15-25,000
4 = $25-35,000
5 = $35-45,000
6 = $45-55,000
7 = $55-75,000
8 = $75,000+
9 = < $35,000
10 = $35,000+

HHTYPE (lifecycle)

1 = Any child < 6
2 = All children 6-17
3 = 1 adult, < 35
4 = 1 adult, 35-64
5 = 1 adult, 65 +
6 = 2+ adults, < 35
7 = 2+ adults, 35-64
8 = 2+ adults, 65+

DAY l-DAY2

1 = Monday
2 = Tuesday
3 = Wednesday
4 = Thursday
5 = Friday
6 = Saturday
7 = Sunday

FLAG

0 = No change has been made
1 = Assign TOT6-17 = 0 and TOTl-5 =0 when HHSIZE = 1, TOTADULT = 1, and both TOT6-17 and TOTl-5 are missing
6 = HHTYPE has been changed according to TOTADULT, TOT6-17, and TOT13
7 = HHSIZE has been changed
9 = TOT6-17 has been changed
10 = TOTl-5 have been changed

A.5 Wave 1 Person Data for Households Participating in the Panel

This file contains members 15 and older of only those households that agreed to take part in the panel., whether or. not they completed travel diaries.

NUMBER OF OBSERVATIONS: 6096
NUMBER OF VARIABLES: 38
RECORD LENGTH: 87

#VARIABLE	LENGTH	BEGIN COLUMN	LABEL
1 HHID	5	1	Household ID
2 SUBID	1	6	Household split ID (always 0 in wave 1)
3 PERS	1	7	Person number
4 OLDPER	2	8	Applies to wave 2 and later data only [l]
5 SEX	2	10	Sex of respondent
6 AGE	2	12	Age of respondent
7 AGE GP	2	14	Age group of respondent
8 EMPLOY	2	16	Employed outside the home
9 OCC	2	18	Occupation of respondent
10 WK_CITY	3	20	City code for work location
11 WK FREQ	2	23	Number of days/week respondent works
12 WK-MODE1	2	25	Travel mode to/from work [2]
13 WK-MODE2	2	27	Travel mode to/from work
14 WK-MODE3	2	29	Travel mode to/from work
15 WK-MODE4	2	31	Travel mode to/from work
16 WK-NUM	2	33	Drive to work alone or with others
17 WK-BUS	2	35	Regularly take bus in past 6 months
18 W-K-POOL	2	37	Regularly pooled in past 6 months
19 Car REQD	2	39	Car required at work
20 CAR-CHLD	2	41	Car required to pick up children
21 CAR-FREQ	2	43	Frequency children are picked up
22 STUDENT	2	45	Currently attend school
23 SCHOOL	3	47	School code
24 SCH CITY	3	50	City code where school is located
25 SCH-MOD1	2	53	Travel mode to/from school [2]
26 SCH-MOD2	2	55	Travel mode to/from school
27 SCH-MOD3	2	57	Travel mode to/from school
28 SCH-MOD4	2	59	Travel mode to/from school
29 SCH-NUM	2	61	Drive to school alone or with others
30 BUS-FREQ	2	63	Frequency using bus per week
31 BUSPASS	2	65	Have transit pass
32 LICENSE	2	67	Have valid driver license
33 CHANGE	2	69	Workplace change code
34 HOME	2	71	Worked at home code
35 MODE	2	73	Primary mode to work
36 WKZIP	5	75	Zip code of workplace
37 WKTRACT	6 (F6.2)	80	Census tract of workplace
38 FLAG	2	86	Correction indicator

NOTES:

[1] OLDPER contains the original person ID number and is applicable only to members who have been split off into a new household and assigned new ID numbers in wave 2 and later:

[2] Travel mode to work (WK_MODElWK_MODE4) and travel mode to school (SCH_MODl-SCH_MOD4): If applicable and only one mode is used to travel to work or school, then only the first variable in each series has a value. If more than one mode is used, subsequent mode(s) are put into as many of the other three variables as needed.

SEX

1 = Male
0 = Female

AGE-GP

1 = 15-17
2 = 18-24
3 = 25-34
4 = 35-44
5 = 45-54
6 = 55-64
7 = 65-98

EMPLOY

1 = Employed outside the home
0 = Employed at the home

OCC

1 = Professional/Technical
2 = Manager/Administration/Business owner
3 = Secretary/Clerical
4 = Retail sales
5 = Other sales
6 = Shop or production worker
7 = Craftsman or foreman
8 = Equipment/Vehicle operator
9 = Service workers
10 = General laborer
11 = Military
12 = Other

WK_MODE1 - WK_MODE4, SCH_MODl-SCH_MOD4

1 = Car only
2 = Bus
3 = Car/bus combination
4 = Motorcycle
5 = Bicycle
6 = Walk
7 = Other

WK_NUM, SCH_NUM

1 = Drive alone
2 = Drive but with others
3 = Ride with others
4 = Take turns

WK_POOL

1 = Regularly pooled in past six months
0 = NOT regularly pooled in past six months

CAR REQD

1 = Car required to pick up children
0 = Car NOT required to pick up children

STUDENT

1 = Student
0 = NOT student

BUSPASS

1 = Have transit pass
0 = NOT have transit pass

LICENSE

1 = Have valid driver license
0 = NOT have valid driver license

CHANGE

0 = New panel member or no previous employment status
1 = Same workplace as previous wave
2 = Changed workplace from previous wave
3 = Different workplace than previous wave, same workplace as wave before that
4 = No record in previous wave, different workplace than wave before that
5 = No record in previous wave, same workplace as wave before that

HOME

0 = Worked outside the home
1 = Worked at home or from the home (work may be performed in the home or in various locations out side the home

MODE

1 = SOV
2 = Transit
3 = Carpool
4 = Walk
5 = Other
6 = Car/bus combination
7 = Ferry

WKTRACT

-8 = Workplace varies, or more than one current workplace

FLAG

0 = No change has been made
2 = SUBID has been changed
3 = PERS has been changed
7 = AGE_GP has been changed according to AGE

A.6 Wave 2 Person Data for Households Participating in the Panel

This file contains members 15 and older of only those households that participated in wave 1, along with replacement households (HHID > 05000).

NUMBER OF OBSERVATIONS: 4631
NUMBER OF VARIABLES: 42
RECORD LENGTH: 95

#	VARIABLE	LENGTH	BEGIN COLUMN	LABEL
1	HHID	5	1	Household ID
2	SUBID	1	6	Household split ID [1]
3	PERS	1	7	Person number
4	OLDPER	2	8	Applies to wave 2 and later data only [2]
5	SEX	2	10	Sex of respondent
6	AGE	2	12	Age of respondent
7	AGE_GP	2	14	Age group of respondent
8	EMPLOY	2	16	Employed outside the home
9	OCC	2	18	Occupation of respondent
10	WK_CITY	3	20	City code for work location
11	WK_FREQ	2	23	Number of days/week respondent works
12	WK_MODE1	2	25	Travel mode to/from work [3]
13	WK_MODE2	2	27	Travel mode to/from work
14	WK_MODE3	2	29	Travel mode to/from work
15	WK_MODE4	2	31	Travel mode to/from work
16	WK_NUM	2	33	Drive to work alone or with others
17	WK_BUS	2	35	Regularly take bus in past 6 months
18	WK_POOL	2	37	Regularly pooled in past 6 months
19	CAR_REQD	2	39	Car required at work
20	CAR_CHLD	2	41	Car required to pick up children
21	CAR_FREQ	2	43	Frequency children are picked up
22	STUDENT	2	45	Currently attend school
23	SCHOOL	3	47	School code
24	SCH_CITY	3	50	City code where school is located
25	SCH_MOD1	2	53	Travel mode to/from school [3]
26	SCH_MOD2	2	55	Travel mode to/from school
27	SCH_MOD3	2	57	Travel mode to/from school
28	SCH_MOD4	2	59	Travel mode to/from school
29	SCH_NUM	2	61	Drive to school alone or with others
30	BUS_FREQ	2	63	Frequency using bus per week
31	BUSPASS	2	65	Have transit pass
32	LICENSE	2	67	Have valid driver license
33	OCCCHG	2	69	Has occupation changed since last year
34	MDCHG	2	71	Changed travel mode to work
35	PLCHG	2	73	Changed workplace
36	CONT	2	75	Continuation member [4]
37	CHANGE	2	77	Workplace change code
38	HOME	2	79	Worked at home code

39 MODE	2	81	Primary mode to work
40 WKZIP	5	83	Zip code of workplace
41 WKTRACT	6 (F6.2)	88	Census tract of workplace
42 FLAG	2	94	Correction indicator

NOTES:

[1] SUBID indicates that a member/members of a core household (either original or replacement) who has/have a different residence in current wave. The first split will have a SUBID = 1. If another member of the original household moves out (though remains in the panel), that new household will have a SUBID=2, and so on.

[2] OLDPER is a variable applicable to wave 2 and later. It contains the member's original person number (assigned at the time the household or individual joined the panel) and is applicable only to members who have been split off into a new household and assigned new ID numbers.

[3] Travel mode to work (WK-MODEl-WK-MODE4) and travel mode to school (SCH MODl-SCH MOD4): If applicable and only one mode is used to travel to work or school, then only the first variable in each series has a value. If more than one mode is used, subsequent mode(s) are put into as many of the other three variables as needed,

[4] Continuation households are households that participated completely in the current wave (the household has person, travel and household data for that wave.) Replacement households replace dropouts from the previous wave and have the following HHIDs: 05001-10358

SEX

1 = Male
0 = Female

AGE_GP

1 = 15-17
2 = 18-24
3 = 25-34
4 = 35-44
5 = 45-54
6 = 55-64
7 = 65-98

EMPLOY

1 = Employed outside the home
0 = Employed at the home

OCC

1 = Professional/Technical
2 = Manager/Administration/Business owner
3 = Secretary/Clerical
4 = Retail sales
5 = Other sales
6 = Shop or production worker
7 = Craftsman or foreman
8 = Equipment/Vehicle operator
9 = Service workers
10 = General laborer
11 = Military
12 = Other

WK_MODE1 - WK_MODE4, SCH_MOD1-SCH_MOD4

1 = Car only
2 = Bus
3 = Car/bus combination
4 = Motorcycle
5 = Bicycle
6 = Walk
7 = Other

WK_NUM, SCH_NUM

1 = Drive alone
2 = Drive but with others
3 = Ride with others
4 = Take turns

WK_POOL

1 = Regularly pooled in past six months
0 = NOT regularly pooled in past six months

CAR_REQD

1 = Car required to pick up children
0 = Car NOT required to pick up children

STUDENT

1 = Student
0 = NOT student

BUSPASS

1 = Have transit pass
0 = NOT have transit pass

LICENSE

1 = Have valid driver license
0 = NOT have valid driver license

OCCCHG, MDCHG, PLCHG

1 = Yes
0 = No

CHANGE

0 = New panel member or no previous employment status
1 = Same workplace as previous wave
2 = Changed workplace from previous wave
3 = Different workplace than previous wave, same workplace as wave before that
4 = No record in previous wave, different workplace than wave before that
5 = No record in previous wave, same workplace as wave before that

HOME

0 = Worked outside the home
1 = Worked at home or from the home (work may be performed in the home or in various locations out side the home

MODE

1 = SOV
2 = Transit
3 = Carpool
4 = Walk
5 = Other
6 = Car/bus combination
7 = Ferry

WKTRACT

-8 = Workplace varies, or more than one current workplace

FLAG

0 = No change has been made
2 = SUBID has been changed
3 = PERS has been changed
7 = AGE_GP has been changed according to AGE

A.7 Wave 3 Person Data for Households Participating in the Panel

This file contains members 15 and older of only those households that participated in wave 2, along with replacement households (HHID ≥ 11000).

NUMBER OF OBSERVATIONS: 3360
NUMBER OF VARIABLES: 44
RECORD LENGTH: 100

#	VARIABLE	LENGTH	BEGIN COLUMN	LABEL
1	HHID	5	1	Household ID
2	SUBID	1	6	Household split ID (always 0 in wave 1)
3	PERS	1	7	Person number
4	OLDPER	2	8	Applies to wave 2 and later data only [1]
5	SEX	2	10	Sex of respondent
6	AGE	2	12	Age of respondent
7	AGE_GP	2	14	Age group of respondent
8	EMPLOY	2	16	Employment/student status
9	OCC	2	18	Occupation of respondent
10	WK_CITY	3	20	City code for work location
11	WK_FREQ	2	23	Number of days/week respondent works
12	WK_MODE1	2	25	Travel mode to/from work [2]
13	WK_MODE2	2	27	Travel mode to/from work
14	WK_MODE3	2	29	Travel mode to/from work
15	WK_MODE4	2	31	Travel mode to/from work
16	WK_NUM	2	33	Drive to work alone or with others
17	WK_BUS	2	35	Regularly take bus in past 6 months
18	WK_POOL	2	37	Regularly pooled in past 6 months
19	CAR_REQD	2	39	Car required at work
20	CAR_CHLD	2	41	Car required to pick up children
21	CAR_FREQ	2	43	Frequency children are picked up
22	STUDENT	2	45	Currently attend school
23	SCHOOL	3	47	School code
24	SCH_CITY	3	50	City code where school is located
25	SCH_MOD1	2	53	Travel mode to/from school [2]
26	SCH_MOD2	2	55	Travel mode to/from school
27	SCH_MOD3	2	57	Travel mode to/from school
28	SCH_MOD4	2	59	Travel mode to/from school
29	SCH_NUM	2	61	Drive to school alone or with others
30	BUS_FREQ	2	63	Frequency using bus per week
31	BUSPASS	2	65	Have transit pass
32	LICENSE	2	67	Have valid driver license
33	OCCCHG	2	69	Has occupation changed since last year
34	CONT	2	71	Continuation member
35	PARKING	3	74	Parking cost
36	DIARY	2	76	Returned travel diary
37	DAY1	2	78	First diary day
38	DAY2	2	80	Second diary day

39 CHANGE	2	82	Workplace change code
40 HOME	2	84	Worked at home code
41 MODE	2	86	Primary mode to work
42 WKZIP	5	88	Zip code of workplace
43 WKTRACT	6 (F6.2)	93	Census tract of workplace
44 FLAG	2	99	Correction indicator

NOTES:

[1] SUBID indicates that a member/members of a core household (either original or replacement) who has/have a different residence in current wave. The first split will have a SUBID = 1. If another member of the original household moves out (though remains in the panel), that new household will have a SUBID = 2, and so on.

[2] OLDPER is a variable applicable to wave 2 and later. It contains the member's original person number (assigned at the time the household or individual joined the panel) and is applicable only to members who have been split off into a new household and assigned new ID numbers.

[3] Travel mode to work (WK-MODEl-WK-MODE4) and travel mode to school (SCH MODl-SCH MOD4): If applicable and only one mode is used to travel to work or school, then only the first variable in each series has a value. If more than one mode is used, subsequent mode(s) are put into as many of the other three variables as needed.

[4] Continuation households are households that participated completed completely in the current wave (the household has person, travel and household data for that wave). Replacement households replace dropouts from the previous wave and have the following HHIDs: 11001-13210

SEX

1 = Male
0 = Female

AGE_GP

1 = 15-17
2 = 18-24
3 = 25-34
4 = 35-44
5 = 45-54
6 = 55-64
7 = 65-98

EMPLOY

1 = Employed outside the home
0 = Employed at the home

OCC

1 = Professional/Technical
2 = Manager/Administration/Business owner
3 = Secretary/Clerical
4 = Retail sales/sales clerk
5 = Other sales/stockbroker
6 = Shop or production worker
7 = Craftsman or foreman
8 = Equipment/Vehicle operator
9 = Service workers
10 = General laborer
11 = Military
12 = Other

WK_MODE1 - WK_MODE4, SCH_MOD1 - SCH_MOD4

1 = Car/carpool/vanpool
2 = Bus
3 = Car/bus combination
4 = Motorcycle
5 = Bicycle
6 = Walk
7 = School bus
8 = Walk to school/work
9 = Metro vanpool
10 = Ferry/walk
11 = Ferry
12 = Carpool
13 = Ferry/car

WK_NUM, SCH_NUM

1 = Drive alone
2 = Drive but with others
3 = Ride with others
4 = Take turns

WK_POOL

1 = Regularly pooled in past six months
0 = NOT regularly pooled in past six months

CAR_REQD

1 = Car required to pick up children
0 = Car NOT required to pick up children

STUDENT

1 = Student
0 = NOT student

BUSPASS

1 = Have transit pass
0 = NOT have transit pass

LICENSE

1 = Have valid driver license
0 = NOT have valid driver license

DAY1-DAY2

1 = Monday
2 = Tuesday
3 = Wednesday
4 = Thursday
5 = Friday
6 = Saturday
7 = Sunday

CHANGE

0 = New panel member or no previous employment status
1 = Same workplace as previous wave
2 = Changed workplace from previous wave
3 = Different workplace than previous wave, same workplace as wave before that
4 = No record in previous wave, different workplace than wave before that
5 = No record in previous wave, same workplace as wave before that

HOME

0 = Worked outside the home
1 = Worked at home or from the home (work may be performed in the home or in various locations out side the home

MODE

1 = SOV
2 = Transit
3 = Carpool
4 = Walk
5 = Other
6 = Car/bus combination
7 = Ferry

WKTRACT

-8 = Workplace varies, or more than one current workplace

FLAG

0 = No change has been made
2 = SUBID has been changed
3 = PERS has been changed
7 = AGE_GP has been changed according to AGE

A.8 Wave 4 Person Data for Households Participating in the Panel

This file contains members 15 and older of only those households that participated in wave 3, along with replacement. households (HHID > 13,300).

NUMBER OF OBSERVATIONS: 4081
NUMBER OF VARIABLES: 44
RECORD LENGTH: 100

# VARIABLE	LENGTH	BEGIN COLUMN	LABEL
1 HHID	5	1	Household ID
2 SUBID	1	6	Household split ID [1]
3 PERS	1	7	Person number
4 OLDPER	2	8	Applies to wave 2 and later data only [2]
5 SEX	2	10	Sex of respondent
6 AGE	2	12	Age of respondent
7 AGE_GP	2	14	Age group of respondent
8 EMPLOY	2	16	Employment/student status
9 OCC	2	18	Occupation of respondent
10 WK_CITY	3	20	City code for work location
11 WK_FREQ	2	23	Number of days/week respondent works
12 WK_MODE1	2	25	Travel mode to/from work [3]
13 WK_MODE2	2	27	Travel mode to/from work
14 WK_MODE3	2	29	Travel mode to/from work
15 WK_MODE4	2	31	Travel mode to/from work
16 WK_NUM	2	33	Drive to work alone or with others
17 WK_BUS	2	35	Regularly take bus in past 6 months
18 WK_POOL	2	37	Regularly pooled in past 6 months
19 CAR_REQD	2	39	Car required at work
20 CAR_CHLD	2	41	Car required to pick up children
21 CAR_FREQ	2	43	Frequency children are picked up
22 STUDENT	2	45	Currently attend school
23 SCHOOL	3	47	School code
24 SCH_CITY	3	50	City code where school is located
25 SCH_MOD1	2	53	Travel mode to/from school [3]
26 SCH_MOD2	2	55	Travel mode to/from school
27 SCH_MOD3	2	57	Travel mode to/from school
28 SCH_MOD4	2	59	Travel mode to/from school
29 SCH_NUM	2	61	Drive to school alone or with others
30 BUS_FREQ	2	63	Frequency using bus per week
31 BUSPASS	2	65	Have transit pass
32 LICENSE	2	67	Have valid driver license
33 OCCCHG	2	69	Has occupation changed since last year
34 CONT	2	71	Continuation member [4]
35 PARKING	3	73	Parking cost
36 DIARY	2	76	Returned travel diary
37 DAY1	2	78	First diary day

38 DAY2	2	80	Second diary day
39 CHANGE	2	82	Workplace change code
40 HOME	2	84	Worked at home code
41 MODE	2	86	Primary mode to work
42 WKZIP	5	88	Zip code of workplace
43 WKTRACT	6 (F6.2)	93	Census tract of workplace
44 FLAG	2	99	Correction indicator

NOTES:

[1] SUBID indicates that a member/members of a core household (either original or replacement) who has/have a different residence in current wave. The first split will have a SUBID = 1. If another member of the original household moves out (though remains in the panel), that new household will have a SUBID=2, and so on.

[2] OLDPER is a variable applicable to wave 2 and later. It contains the member's original person number (assigned at the time the household or individual joined the panel) and is applicable only to members who have been split off into a new household and assigned new ID numbers.

[3] Travel mode to work (WK-MODEl-WK-MODE4) and travel mode to school (SCH MODl-SCH MOD4): If applicable and only one mode is used to travel to work or school, then only the first variable in each series has a value. If more than one mode is used, subsequent mode(s) are put into as many of the other three variables as needed.

[4] Continuation households are households that participated completely in the current wave (the household has person, travel and household data for that wave.) Replacement households replace dropouts from the previous wave and have the following HHIDs: 13301-14713

SEX

1 = Male
0 = Female

AGE_GP

1 = 15-17
2 = 18-24
3 = 25-34
4 = 35-44
5 = 45-54
6 = 55-64
7 = 65-98

EMPLOY

1 = Employed outside the home
0 = Employed at the home

OCC

1 = Professional/Technical
2 = Manager/Administration/Business owner
3 = Secretary/Clerical
4 = Retail sales/sales clerk
5 = Other sales/stockbroker
6 = Shop or production worker
7 = Craftsman or foreman
8 = Equipment/Vehicle operator
9 = Service workers
10 = General laborer
11 = Military
12 = Other

WK_MODE1 - WK_MODE4, SCH_MOD1-SCH_MOD4

1 = Car/carpool/vanpool
2 = Bus
3 = Car/bus combination
4 = Motorcycle
5 = Bicycle
6 = Walk
7 = School bus
8 = Walk to school/work
9 = Metro vanpool
10 = Ferry/walk
11 = Ferry
12 = Carpool
13 = Ferry/car

WK_NUM, SCH_NUM

1 = Drive alone
2 = Drive but with others
3 = Ride with others
4 = Take turns

WK_POOL

1 = Regularly pooled in past six months
0 = NOT Regularly pooled in past six months

CAR_REQD

1 = Car required to pick up children
0 = Car NOT required to pick up children

STUDENT

1 = Student
0 = NOT student

BUSPASS

1 = Have transit pass
0 = NOT have transit pass

LICENSE
1 = Have valid driver license
0 = NOT have valid driver license

DAY1-DAY2
1 = Monday
2 = Tuesday
3 = Wednesday
4 = Thursday
5 = Friday
6 = Saturday
7 = Sunday

CHANGE
0 = New panel member or no previous employment status
1 = Same workplace as previous wave
2 = Changed workplace from previous wave
3 = Different workplace than previous wave, same workplace as wave before that
4 = No record in previous wave, different workplace than wave before that
5 = No record in previous wave, same workplace as wave before that

HOME
0 = Worked outside the home
1 = Worked at home or from the home (work may be performed in the home or in various locations out side the home

MODE
1 = SOV
2 = Transit
3 = Carpool
4 = Walk
5 = Other
6 = Car/bus combination
7 = Ferry

WKTRACT
-8 = Workplace varies, or more than one current workplace

FLAG
0 = No change has been made
2 = SUBID has been changed
3 = PERS has been changed
7 = AGE_GP has been changed according to AGE

A.9 Travel Data

NUMBER OF OBSERVATIONS: 31335 (1989)
33284 (1990)
29518 (1992)
33179 (1993)
NUMBER OF VARIABLES: 25
RECORD LENGTH: 66

# VARIABLE	LENGTH	BEGIN COLUMN	LABEL
1 HHID	5	1	Household ID
2 SUBID	1	6	Household split ID (always 0 in wave 1)
3 PERS	1	7	Person ID
4 DIARY	1	8	Day number
5 DAYOFWK	1	9	Day of week
6 MDD	3	10	Month/date
7 TOTTRIP	2	13	Total trips
8 TRIPNUM	2	15	Trip number
9 PURPOSE	2	17	Trip purpose
10 TYPE	2	19	Trip type
11 MODE	2	21	Trip mode
12 BEGTIME	4	23	Trip starting time
13 ENDTIME	4	27	Trip ending time
14 MINUTES	3	31	Duration of trip
15 D_R	2	34	Driver/Rider
16 NUM	2	36	Number in vehicle
17 REL1	2	38	Relation of other vehicle occupant 1
18 REL2	2	40	Relation of other vehicle occupant 2
19 REL3	2	42	Relation of other vehicle occupant 3
20 ORIGCT	5	44	Trip origin census tract
21 DESTCT	5	49	Trip destination census tract
22 DISTANCE	6 (F6.2)	54	Travel distance in miles
23 ATIME	4	60	Activity duration
24 FLAG	3	64	Correction indicator

DIARY
1 = First day of the travel diary
2 = Second day of the travel diary

MDD (The first character of the variable indicates month of year. Its meaning is given below. The last two characters represent date of month)
Wave 1:
1 = September
2 = October
3 = November
4 = December

Wave 2:
1 = October
2 = November
3 = December
4 = January

Wave 3:
1 = September
2 = October
3 = November
4 = December
5 = January

Wave 4:
1 = October
2 = November
3 = December
4 = January

DAYOFWK
1 = Monday
2 = Tuesday
3 = Wednesday
4 = Thursday
5 = Friday
6 = Saturday
7 = Sunday

PURPOSE
0 = No trip was made
1 = Work
2 = Shopping
3 = School
4 = Visiting
5 = Free-time
6 = Personal
7 = Appointments
8 = Home
9 = College

MODE
0 = No trip was made
1 = Car
2 = Carpool
3 = Vanpool
4 = Bus
5 = Para-transit
6 = Taxi
7 = Walk
8 = Bike
9 = Motorcycle
10 = School Bus

11 = Ferry/car
12 = Ferry/foot
13 = Monorail
14 = Boat
15 = Train
16 = Airplane

TYPE

0 = No trip was made
1 = Home-based work
2 = Home-based other
3 = Non-home-based

REL1, REL2, and REL3

0 = No trip was made
1 = Spouse/partner
2 = Child household member (< 15)
3 = Adult household member (15+)
4 = Non-household family member
5 = Co-worker/client
6 = Adult friend(s)
7 = Child friend(s)

D_R

0 = No trip was made
1 = Driver
2 = Rider

FLAG

0 = No change has been changed
1 = Adding a returning home trip
2 = Change order of a trip sequence
111 = BEGTIME and ENDTIME have been changed from 24xx to __xx, and MINUTES is the same
112 = BEGTIME and ENDTIME have been changed from 12xx to __xx, or vice versa, and MINUTES is the same
113 = BEGTIME and ENDTIME have been changed not in the same ways as in 111 and 112, but MINUTES is the same
121 = BEGTIME and ENDTIME have been changed from 24xx to __xx, and MINUTES has been changed accordingly
122 = BEGTIME and ENDTIME have been changed from 12xx to __xx, or vice versa, and MINUTES has been changed accordingly
123 = BEGTIME and ENDTIME have been changed not in the same ways as in 111 and 112, and MINUTES has been changed accordingly
211 = BEGTIME has been changed from 24xx to __xx, and MINUTES is the same
212 = BEGTIME has been changed from 12xx to __xx, or vice versa, and MINUTES is the same
213 = BEGTIME has been changed not in the same ways as in 111 and 112, but MINUTES is the same

221= BEGTIME has been changed from 24xx to __xx, and MINUTES has been changed accordingly
222= BEGTIME has been changed from 12xx to __xx, or vice versa, and MINUTES has been changed accordingly
223= BEGTIME has been changed not in the same ways as in 111 and 112, and MINUTES has been changed accordingly
311= ENDTIME has been changed from 24xx to __xx, and MINUTES is the same
312= ENDTIME has been changed from 12xx to __xx, or vice versa, and MINUTES is the same
313= ENDTIME has been changed not in the same ways as in 111 and 112, but MINUTES is the same
321= ENDTIME has been changed from 24xx to xx, and MINUTES has been changed accordingly
322= ENDTIME has been changed from 12xx to __xx, or vice versa, and MINUTES has been changed accordingly
323= ENDTIME has been changed not in the same ways as in 111 and 112, and MINUTES has been changed accordingly
401= Switch BEGTIME and ENDTIME

A.10 The 1990 Attitude and Values Survey

This file contains data from a survey asking the panel members about factors affecting their daily travel choices. Only those who were employed (full or part time) or in school (1 < STATUS < 4) at the time of the survey were asked to complete the survey.

NUMBER OF OBSERVATIONS:	2478
NUMBER OF FIELDS:	113
NUMBER OF ROWS PER RECORD:	4

#	VARIABLE	LENGTH	ROW	BEGIN COLUMN	DESCRIPTION
1	HHID	5	1	1	Household ID
2	SUBID	1	1	6	Household split ID [1]
3	PERS	1	1	7	Person ID
4	STATUS	2	1	8	Current work/school status
	SECTION A: IMPORTANCE RATING				
5	SECA1	2	1	10	Ability to arrive on time
6	SECA2	2	1	12	Ability to travel without changing vehicle
7	SECA3	2	1	14	Not having to deal with traffic congestion
8	SECA4	2	1	16	Short travel time
9	SECA5	2	1	18	Day-to-day costs like gas, parking, bus
10	SECA6	2	1	20	Protection from weather
11	SECA7	2	1	22	Having a seat
12	SECA8	2	1	24	Short wait time
13	SECA9	2	1	26	Ability to read while traveling
14	SECA10	2	1	28	Ability to travel when desired
15	SECA11	2	1	30	Flexibility to change plans
16	SECA12	2	1	32	Making few stops
17	SECA13	2	1	34	Parking availability
18	SECA14	2	1	36	Safety of vehicle from accidents
19	SECA15	2	1	38	Freedom from threats to personal safety
20	SECA16	2	1	40	Avoiding stress
21	SECA17	2	1	42	Minimizing pollution
	SECTION B: PERFORMANCE RATING				
22	SECB1	2	1	44	SOV: Ability to arrive on time
23	SECB2	2	1	46	Bus: Ability to arrive on time
24	SECB3	2	1	48	Pool: Ability to arrive on time
25	SECB4	2	1	50	SOV: Ability to travel without changing vehicle
26	SECB5	2	1	52	Bus: Ability to travel without changing vehicle
27	SECB6	2	1	54	Pool: Ability to travel without changing vehicle
28	SECB7	2	1	56	SOV: Not having to deal with traffic congestion
29	SECB8	2	1	58	Bus: Not having to deal with traffic congestion

30 SECB9	2	1	60	Pool: Not having to deal with traffic congestion
31 SECB10	2	1	62	SOV: Short travel time
32 SECB11	2	1	64	Bus: Short travel time
33 SECB12	2	1	66	Pool: Short travel time
34 SECB13	2	1	68	SOV: Day-to-day costs like gas, parking, bus
35 SECB14	2	1	70	Bus: Day-to-day costs like gas, parking, bus
36 SECB15	2	1	72	Pool: Day-to-day costs like gas, parking, bus
37 SECB16	2	1	74	SOV: Protection from weather
38 SECB17	2	1	76	Bus: Protection from weather
39 SECB18	2	1	78	Pool: Protection from weather
40 SECB19	2	2	1	SOV: Having a seat
41 SECB20	2	2	3	Bus: Having a seat
42 SECB21	2	2	5	Pool: Having a seat
43 SECB22	2	2	7	SOV: Short wait time
44 SECB23	2	2	9	Bus: Short wait time
45 SECB24	2	2	11	Pool: Short wait time
46 SECB25	2	2	13	SOV: Ability to read while traveling
47 SECB26	2	2	15	Bus: Ability to read while traveling
48 SECB27	2	2	17	Pool: Ability to read while traveling
49 SECB28	2	2	19	SOV: Ability to travel when desired
50 SECB29	2	2	21	Bus: Ability to travel when desired
51 SECB30	2	2	23	Pool: Ability to travel when desired
52 SECB31	2	2	25	SOV: Flexibility to change plans
53 SECB32	2	2	27	Bus: Flexibility to change plans
54 SECB33	2	2	29	Pool: Flexibility to change plans
55 SECB34	2	2	31	SOV: Making few stops
56 SECB35	2	2	33	Bus: Making few stops
57 SECB36	2	2	35	Pool: Making few stops
58 SECB37	2	2	37	SOV: Parking availability
59 SECB38	2	2	39	Bus: Parking availability
60 SECB39	2	2	41	Pool: Parking availability
61 SECB40	2	2	43	SOV: Safety of vehicle from accidents
62 SECB41	2	2	45	Bus: Safety of vehicle from accidents
63 SECB42	2	2	47	Pool: Safety of vehicle from accidents
64 SECB43	2	2	49	SOV: Freedom from threats to personal safety
65 SECB44	2	2	51	Bus: Freedom from threats to personal safety
66 SECB45	2	2	53	Pool: Freedom from threats to personal safety
67 SECB46	2	2	55	SOV: Avoiding stress
68 SECB47	2	2	57	Bus: Avoiding stress
69 SECB48	2	2	59	Pool: Avoiding stress
70 SECB49	2	2	61	SOV: Minimizing pollution
71 SECB50	2	2	63	Bus: Minimizing pollution
72 SECB51	2	2	65	Pool: Minimizing pollution

SECTION C: AGREEMENT/DISAGREEMENT STATEMENTS

73 SECC1	2	2	67	Riding a bus is relaxing way to commute
74 SECC2	2	2	69	I enjoy driving my car even in heavy traffic
75 SECC3	2	2	71	My schedule is too erratic to be in a carpool
76 SECC4	2	2	73	If gas prices go up, less likely to drive to work

77 SECC5	2	2	75	I don't know anyone to carpool with
78 SECC6	2	2	77	Taking the bus doesn't fit my lifestyle
79 SECC7	2	2	79	Willing to pay higher taxes to improve bus service
80 SECC8	2	3	1	I hate the idea of transferring buses
81 SECC9	2	3	3	SOV should pay more for parking than poolers
82 SECC10	2	3	5	Carpooling is an enjoyable way to travel
83 SECC11	2	3	7	It's easy to find someone I can carpool with
84 SECC12	2	3	9	Riding the bus helps reduce traffic congestion
85 SECC13	2	3	11	Won't rely on another to get to work on time
86 SECC14	2	3	13	Driving a car is a relaxing way to commute
87 SECC15	2	3	15	Not fair having special HOV/transit lanes
88 SECC16	2	3	17	Getting bus schedule information is easy
89 SECC17	2	3	19	It's a hassle to take the bus
90 SECC18	2	3	21	I like the freedom of driving my own car
91 SECC19	2	3	23	Would carpool with someone unknown if convenient
92 SECC20	2	3	25	Taking the bus is an enjoyable way to travel
93 SECC21	2	3	27	People only ride the bus to work if they have to
94 SECC22	2	3	29	More freeway carpool lanes should be built
95 SECC23	2	3	31	I'd rather drive with others than by myself

SECTION D: WORK COMMUTE

96 DAYS	2	3	33	How many days a week do you work
97 MODE	2	3	35	Usual travel mode to work
98 CARREQ	2	3	37	Does job require car?
99 CAREMPL	2	3	39	Does company have cars available for this purpose
100 ZIP	3	3	41	ZIP code
101 MOST	2	3	44	Most preferred way to work (Bus/Pool/SOV)
102 LEAST	2	3	46	Least preferred way to work (Bus/Pool/SOV)

SECTION D: BUS AVAILABILITY & USE

103 BUSSTOP	2	3	48	How far away is nearest bus stop
104 DIRECT	2	3	50	If >3 blocks, does it go directly to work/school
105 CONSID	2	3	52	If >3 blocks, have you considered riding the bus

SECTION D: CO-RIDER AVAILABILITY

106 POOL	2	3	54	Do you commute by carpool or vanpool
107 HHLDCOMM	2	3	56	Anyone in household with similar commute pattern
108 NBHDCOMM	2	3	58	Anyone in neighborhood with similar commute pattern
109 WORKCOMM	2	3	60	Anyone at work with similar commute pattern

SECTION D: NEED CAR BEFORE OR AFTER WORK

110 FREQCHLD	2	3	62	Freq. needing car to drop off/pick up children
111 FREQERRD	2	3	64	Freq. needing car for other personal errands

SECTION D: NEED CAR DURING WORK

112 FREQPERS	2	3	66	Freq. needing car for personal trips during day
113 filler	21	4	1	[2]

NOTES:

[1] In this file SUBID is always 0.

[2] Columns 1-4 are from Section E. Columns 5-21 are from the VISION 2020 questionnaire attached to the end of the survey and is not part of the panel.

STATUS

1 = Work, 35 hours/week or more
2 = Work, less than 35 hours/week
3 = Student, full-time
4 = Student, part-time
5 = Homemaker
6 = Retired
7 = Unemployed
8 = Other

SECTION A (Importance rating):

1 = Not important at all
2 = .
3 = .
4 = .
5 = .
6 = \./
7 = Extremely important

SECTION B (Performance rating):

1 = Extremely poorly
2 = \./
3 = Somewhat poorly
4 = \./
5 = Somewhat well
6 = \./
7 = Extremely well

SECTION C (Agreement/Disagree Statements):

1 = Very strongly disagree
2 = Strongly disagree
3 = Somewhat disagree
4 = Neutral
5 = Somewhat agree
6 = Strongly agree
7 = Very strongly agree

SECTION D:

MODE

1 = Drive alone
2 = Drive, but with others
3 = Ride with others
4 = Take turns drive/ride
5 = Bus only
6 = Bus/car combo: Catch bus at Park & Ride
7 = Bus/car combo: Catch bus other than at Park & Ride
8 = Motorcycle
9 = Bicycle
10 = Walk
11 = Ferry
12 = Other

BUSSTOP

1 = 3 block or less (less than 1/4 mile)
2 = 4 to 6 blocks (1/4 to ½ mile)
3 = 7 blocks or more (more than ½ mile)

FREQCHLD, FREQERRD, and FREQPERS

1 = 3 or more days a week
2 = 1 or 2 days a week
3 = 2 or 3 times a month
4 = Once a month or less
5 = Never

CARREQ, CAREMPL, DIRECT, CONSID, HHLDCOMM, NBHDCOMM, and WORKCOMM

1 = Yes
0 = No

A.11 The 1991 Attitude and Values Survey

This file contains data from a survey asking the panel members about factors affecting their daily travel choices. Only those who were employed (full or part time) or in school (1 ≤STATUS ≤ 4) at the time of the survey were asked to complete the survey.

NUMBER OF OBSERVATIONS: 2395
NUMBER OF FIELDS: 121
NUMBER OF ROWS PER RECORD: 4

#	VARIABLE	LENGTH	ROW	BEGIN COLUMN	DESCRIPTION
1	HHID	5	1	1	Household ID
2	SUBID	1	1	6	Household split ID [1]
3	PERS	1	1	7	Person ID
4	STATUS	2	1	8	Current work/school status
					SECTION A: IMPORTANCE RATING
5	SECA1	2	1	10	Ability to arrive on time
6	SECA2	2	1	12	Ability to travel without changing vehicle
7	SECA3	2	1	14	Not having to deal with traffic congestion
8	SECA4	2	1	16	Short travel time
9	SECA5	2	1	18	Day-to-day costs like gas, parking, bus
10	SECA6	2	1	20	Protection from weather
11	SECA7	2	1	22	Having a seat
12	SECA8	2	1	24	Short wait time
13	SECA9	2	1	26	Ability to read while traveling
14	SECA10	2	1	28	Ability to travel when desired
15	SECA11	2	1	30	Flexibility to change plans
16	SECA12	2	1	32	Making few stops
17	SECA13	2	1	34	Parking availability
18	SECA14	2	1	36	Safety of vehicle from accidents
19	SECA15	2	1	38	Freedom from threats to personal safety
20	SECA16	2	1	40	Avoiding stress
21	SECA17	2	1	42	Minimizing pollution
					SECTION B: PERFORMANCE RATING
22	SECB1	2	1	44	SOV: Ability to arrive on time
23	SECB2	2	1	46	Bus: Ability to arrive on time
24	SECB3	2	1	48	Pool: Ability to arrive on time
25	SECB4	2	1	50	SOV: Ability to travel without changing vehicle
26	SECB5	2	1	52	Bus: Ability to travel without changing vehicle
27	SECB6	2	1	54	Pool: Ability to travel without changing vehicle
28	SECB7	2	1	56	SOV: Not having to deal with traffic congestion
29	SECB8	2	1	58	Bus: Not having to deal with traffic congestion
30	SECB9	2	1	60	Pool: Not having to deal with traffic congestion

31 SECB10	2	1	62	SOV: Short travel time
32 SECB11	2	1	64	Bus: Short travel time
33 SECB12	2	1	66	Pool: Short travel time
34 SECB13	2	1	68	SOV: Day-to-day costs like gas, parking, bus
35 SECB14	2	1	70	Bus: Day-to-day costs like gas, parking, bus
36 SECB15	2	1	72	Pool: Day-to-day costs like gas, parking, bus
37 SECB16	2	1	74	SOV: Protection from weather
38 SECB17	2	1	76	Bus: Protection from weather
39 SECB18	2	1	78	Pool: Protection from weather
40 SECB19	2	2	1	SOV: Having a seat
41 SECB20	2	2	3	Bus: Having a seat
42 SECB21	2	2	5	Pool: Having a seat
43 SECB22	2	2	7	SOV: Short wait time
44 SECB23	2	2	9	Bus: Short wait time
45 SECB24	2	2	11	Pool: Short wait time
46 SECB25	2	2	13	SOV: Ability to read while traveling
47 SECB26	2	2	15	Bus: Ability to read while traveling
48 SECB27	2	2	17	Pool: Ability to read while traveling
49 SECB28	2	2	19	SOV: Ability to travel when desired
50 SECB29	2	2	21	Bus: Ability to travel when desired
51 SECB30	2	2	23	Pool: Ability to travel when desired
52 SECB31	2	2	25	SOV: Flexibility to change plans
53 SECB32	2	2	27	Bus: Flexibility to change plans
54 SECB33	2	2	29	Pool: Flexibility to change plans
55 SECB34	2	2	31	SOV: Making few stops
56 SECB35	2	2	33	Bus: Making few stops
57 SECB36	2	2	35	Pool: Making few stops
58 SECB37	2	2	37	SOV: Parking availability
59 SECB38	2	2	39	Bus: Parking availability
60 SECB39	2	2	41	Pool: Parking availability
61 SECB40	2	2	43	SOV: Safety of vehicle from accidents
62 SECB41	2	2	45	Bus: Safety of vehicle from accidents
63 SECB42	2	2	47	Pool: Safety of vehicle from accidents
64 SECB43	2	2	49	SOV: Freedom from threats to personal safety
65 SECB44	2	2	51	Bus: Freedom from threats to personal safety
66 SECB45	2	2	53	Pool: Freedom from threats to personal safety
67 SECB46	2	2	55	SOV: Avoiding stress
68 SECB47	2	2	57	Bus: Avoiding stress
69 SECB48	2	2	59	Pool: Avoiding stress
70 SECB49	2	2	61	SOV: Minimizing pollution
71 SECB50	2	2	63	Bus: Minimizing pollution
72 SECB51	2	2	65	Pool: Minimizing pollution

SECTION C: AGREEMENT/DISAGREEMENT STATEMENTS

73 SECC1	2	2	67	Riding a bus is relaxing way to commute
74 SECC2	2	2	69	I enjoy driving my car even in heavy traffic
75 SECC3	2	2	71	My schedule is too erratic to be in a carpool
76 SECC4	2	2	73	If gas prices go up, less likely to drive to work
77 SECC5	2	2	75	I don't know anyone to carpool with

78 SECC6	2	2	77	Taking the bus doesn't fit my lifestyle
79 SECC7	2	2	79	Willing to pay higher taxes to improve bus service
80 SECC8	2	3	1	I hate the idea of transferring buses
81 SECC9	2	3	3	SOV should pay more for parking than poolers
82 SECC10	2	3	5	Carpooling is an enjoyable way to travel
83 SECC11	2	3	7	It's easy to find someone I can carpool with
84 SECC12	2	3	9	Riding the bus helps reduce traffic congestion
85 SECC13	2	3	11	Won't rely on another to get to work on time
86 SECC14	2	3	13	Driving a car is a relaxing way to commute
87 SECC15	2	3	15	Not fair having special HOV/transit lanes
88 SECC16	2	3	17	Getting bus schedule information is easy
89 SECC17	2	3	19	It's a hassle to take the bus
90 SECC18	2	3	21	I like the freedom of driving my own car
91 SECC19	2	3	23	Would carpool with someone I don't know if convenient
92 SECC20	2	3	25	Taking the bus is an enjoyable way to travel
93 SECC21	2	3	27	People only ride the bus to work if they have to
94 SECC22	2	3	29	More freeway carpool lanes should be built
95 SECC23	2	3	31	I'd rather drive with others than by myself

SECTION D: WORK COMMUTE

96 ADDR	30	3	33	Work address
97 CITY	20	4	1	City
98 ZIP	5	4	21	ZIP code
99 DAYS	2	4	26	How many days a week do you work
100 MODE	2	4	28	Usual travel mode to work
101 LEAVE	4	4	30	Time usually leaving home for work
102TRAVTIME	3	4	34	Minutes to get to work
103 CARREQ	2	4	37	Does job require car?
104 CAREMPL	2	4	39	Does company have cars available for this purpose
105 PAYPARK	2	4	41	If driving, do you personally pay for parking
106 FREEPARK	2	4	43	Does employer provide free or reduced fee parking
107 MOST	2	4	45	Most preferred way to work (Bus/Pool/SOV)
108 LEAST	2	4	47	Least preferred way to work (Bus/Pool/SOV)

SECTION D: BUS AVAILABILITY & USE

109 BUSSTOP	2	4	49	How far away is nearest bus stop
110 DIRECT	2	4	51	If >3 blocks, does it go directly to work/school
111 CONSID	2	4	53	If >3 blocks, have you considered riding the bus
112 BUSRIDES	2	4	55	Number of times riding bus to and from work
113 PASS	2	4	57	Do you have a transit pass

SECTION D: CO-RIDER AVAILABILITY

114 POOL	2	4	59	Do you commute by carpool or vanpool
115 POOLRIDE	2	4	61	Number of times carpooling to and from work
116 HHLDCOMM	2	4	63	Anyone in household with similar commute pattern

117	NBHDCOMM	2	4	65	Anyone in neighborhood with similar commute pattern
118	WORKCOMM	2	4	67	Anyone at work with similar commute pattern

SECTION D: NEED CAR BEFORE OR AFTER WORK

119	FREQCHLD	2	4	69	Freq. needing car to drop off/pick up children
120	FREQERRD	2	4	71	Freq. needing car for other personal errands

SECTION D: NEED CAR DURING WORK

121	FREQPERS	2	4	73	Freq. needing car for personal trips during day

NOTES:

[1] In a split household (SUBID > 0) a person ID number may change in wave 2 (1990). To match person data across waves, use the PERS and OLDPER ID variables in wave 2 person data file to change to wave 2 person ID to the wave 1 ID, and visa versa.

STATUS

1 = Work, 35 hours/week or more
2 = Work, less than 35 hours/week
3 = Student, full-time
4 = Student, part-time
5 = Homemaker
6 = Retired
7 = Unemployed
8 = Other

SECTION A (Importance rating):

1 = Not important at all
2 = .
3 = .
4 = .
5 = .
6 = \./
7 = Extremely important

SECTION B (Performance rating):

1 = Extremely poorly
2 = \./
3 = Somewhat poorly
4 = \./
5 = Somewhat well
6 = \./
7 = Extremely well

SECTION C (Agreement/Disagree Statements):

1 = Very strongly disagree
2 = Strongly disagree
3 = Somewhat disagree
4 = Neutral
5 = Somewhat agree
6 = Strongly agree
7 = Very strongly agree

SECTION D:

MODE

1 = Drive alone
2 = Drive, but with others
3 = Ride with others
4 = Take turns drive/ride
5 = Bus only
6 = Bus/car combo: Catch bus at Park & Ride
7 = Bus/car combo: Catch bus other than at Park & Ride
8 = Motorcycle
9 = Bicycle
10 = Walk
11 = Ferry
12 = Other

FREEPARK

1 = Free
2 = Reduced
3 = Neither

BUSSTOP

1 = 3 block or less (less than 1/4 mile)
2 = 4 to 6 blocks (1/4 to ½ mile)
3 = 7 blocks or more (more than ½ mile)

FREQCHLD, FREQERRD, and FREQPERS

1 = 3 or more days a week
2 = 1 or 2 days a week
3 = 2 or 3 times a month
4 = Once a month or less
5 = Never

CARREQ, CAREMPL, DIRECT, CONSID, HHLDCOMM, NBHDCOMM, and WORKCOMM

1 = Yes
0 = No

A.12 Integrated Trip-Person-Household Database

This file contains trip-based integrated trip, person, and household information. Persons or households without trip records are excluded.

NUMBER OF OBSERVATIONS: 31335 (1989)
33284 (1990)
29518 (1992)
33170 (1993)
NUMBER OF FIELDS: 70
LENGTH 180

# VARIABLE	LENGTH	BEGIN COLUMN	DESCRIPTION
1 HHID	5	1	Household ID
2 SUBID	1	6	Household split ID [1]
3 PERS	1	7	Person ID
4 DIARY	1	8	Day number
5 DAYOFWK	1	9	Day of week
6 MDD	3	10	Month/date
7 TOTTRIP	2	13	Total trips
8 TRIPNUM	2	15	Trip number
9 PURPOSE	2	17	Trip purpose
10 TYPE	2	19	Trip type
11 MODE	2	21	Trip mode
12 BEGTIME	4	23	Trip starting time
13 ENDTIME	4	27	Trip ending time
14 MINUTES	3	31	Duration of trip
15 D_R	2	34	Driver/Rider
16 NUM	2	36	Number in vehicle
17 REL1	2	38	Relation of other vehicle occupant 1
18 REL2	2	40	Relation of other vehicle occupant 2
19 REL3	2	42	Relation of other vehicle occupant 3
20 ORIGCT	5	44	Trip origin census tract
21 DESTCT	5	49	Trip destination census tract
22 DISTANCE	6 (F6.2)	54	Travel distance in miles
23 ATIME	4	60	Activity duration
24 OLDPER	2	64	Applies to wave 2 and later data only [2]
25 SEX	2	66	Sex of respondent
26 AGE	2	68	Age of respondent
27 AGE_GP	2	70	Age group of respondent
28 HOME	2	72	Worked at home code
29 OCC	2	74	Occupation of respondent
30 WK_CITY	3	76	City code for work location
31 WK_FREQ	2	79	Number of days/week respondent works
32 WKMODE	2	81	Primary travel mode to/from work [3]
33 WK_NUM	2	83	Drive to work alone or with others
34 WK_BUS	2	85	Regularly take bus in past 6 months
35 WK_POOL	2	87	Regularly pooled in past 6 months

36 CAR_REQD	2	89	Car required at work
37 CAR_CHLD	2	91	Car required to pick up children
38 CAR_FREQ	2	93	Frequency children are picked up
39 STUDENT	2	95	Currently attend school
40 SCHOOL	3	97	School code
41 SCH_CITY	3	100	City code where school is located
42 SCH_MODE	2	103	Primary travel mode to/from school [4]
43 SCH_NUM	2	105	Drive to school alone or with others
44 BUS_FREQ	2	107	Frequency using bus per week
45 BUSPASS	2	109	Have transit pass
46 LICENSE	2	111	Have valid driver license
47 OCCCHG	2	113	Has occupation changed since last year
48 MDCHG	2	115	Changed travel mode to work
49 CONT	2	117	Continuation member [5]
50 PARKING	3	119	Parking cost
51 CHANGE	2	122	Workplace change code
52 WKZIP	5	124	Zip code of workplace
53 WKTRACT	6 (F6.2)	129	Census tract of workplace
54 SAMPLE	2	135	Sample category
55 INCOME	6	137	Household income
56 INCCAT	2	143	Household income category
57 HHTYPE	2	145	Life cycle
58 HHSIZE	2	147	Household size
59 TOTADULT	2	149	Number of adults (18+)
60 TOT6_17	2	151	Number of children (6-17)
61 TOT1_5	2	153	Number of children (<6)
62 TOT_LOG	2	155	Number in household 15+ for travel log
63 NUMVEH	2	157	Number of household vehicles
64 PANEL	2	159	Panel status [6]
65 HHZIP	5	161	Zip code
66 HHTRACT	6 (F6.2)	166	Census tract
67 HHTAZ	3	172	Traffic analysis zone
68 TFLAG	2	175	Correction indicator for trip record
69 PFLAG	2	177	Correction indicator for person record
70 HFLAG	2	179	Correction indicator for household record

NOTES:

[1] SUBID indicates that a member/members of a core household (either original or replacement) who has/have a different residence in current wave. The first split will have **a** SUBID = 1. If another member of the original household moves out (though remains in the panel), that new household will have a SUBID=2, and so on.

[2] **OLDPER** is a variable applicable to wave 2 and later. It contains the member's original person number (assigned **at** the time the household or individual joined the panel) and is applicable only to members who have been split off into a new household and assigned new ID numbers.

[3] It was derived from the person file and the travel diaries by PSRC. It is equivalent to WK-MODE1 in the person file.

[4] It is equivalent to SCH-MOD 1 in the person file.

[5] Continuation households are households that participated completely in the current wave (the household has person, travel and household data for that wave). Replacement households replace dropouts from the previous wave and have the following HHIDs:

Wave 2: 05001-10358
Wave 3: 11001-13210
Wave 4: 13301-14713

[6] Used to determine whether or not a household can be considered a continuation household for the subsequent wave, i.e., there are household person data and some or all of the panel members within the household return travel diaries.

DIARY

1 = First day of the travel diary
2 = Second day of the travel diary

MDD (First character of the variable indicates month of year. Its meaning is given below. The last two characters represent date of month)

Wave 1:

1 = September
2 = October
3 = November
4 = December

Wave 2:

1 = October
2 = November
3 = December
4 = January

Wave 3:

1 = September
2 = October
3 = November
4 = December
5 = January

Wave 4:

1 = October
2 = November
3 = December
4 = January

DAYOFWK

1 = Monday
2 = Tuesday
3 = Wednesday
4 = Thursday
5 = Friday
6 = Saturday
7 = Sunday

PURPOSE

0 = No trip was made
1 = Work
2 = Shopping
3 = School
4 = Visiting
5 = Free-time
6 = Personal
7 = Appointments
8 = Home
9 = College

MODE

0 = No trip was made
1 = Car
2 = Carpool
3 = Vanpool
4 = Bus
5 = Para-transit
6 = Taxi
7 = Walk
8 = Bike
9 = Motorcycle
10 = School Bus
11 = Ferry/car
12 = Ferry/foot
13 = Monorail
14 = Boat
15 = Train
16 = Airplane

TYPE

0 = No trip was made
1 = Home-based work
2 = Home-based other
3 = Non-home-based

REL1, REL2, and REL3

0 = No trip was made
1 = Spouse/partner
2 = Child household member (< 15)
3 = Adult household member (15+)
4 = Non-household family member
5 = Co-worker/client
6 = Adult friend(s)
7 = Child friend(s)

D_R

0 = No trip was made
1 = Driver
2 = Rider

SEX

1 = Male
0 = Female

AGE_GP

1 = 15-17
2 = 18-24
3 = 25-34
4 = 35-44
5 = 45-54
6 = 55-64
7 = 65-98

HOME

0 = Worked outside the home
1 = Worked at home or from the home (work may be performed in the home or in various locations out side the home)

OCC

1 = Professional/Technical
2 = Manager/Administration/Business owner
3 = Secretary/Clerical
4 = Retail sales/sales clerk
5 = Other sales/stockbroker
6 = Shop or production worker
7 = Craftsman or foreman
8 = Equipment/Vehicle operator
9 = Service workers
10 = General laborer
11 = Military
12 = Other

WKMODE

1 = SOV
2 = Transit
3 = Carpool
4 = Walk
5 = Other
6 = Car/bus combination
7 = Ferry

WK_NUM, SCH_NUM

1 = Drive alone
2 = Drive but with others
3 = Ride with others
4 = Take turns

WK_POOL

1 = Regularly pooled in past six months
0 = NOT regularly pooled in past six months

CAR_REQD

1 = Car required to pick up children
0 = Car NOT required to pick up children

STUDENT

1 = Student
0 = NOT student

SCH_MODE

Wave 1 and wave 2:

1 = SOV
2 = Transit
3 = Carpool
4 = Walk
5 = Other
6 = Car/bus combination
7 = Ferry

Wave 3 and wave 4:

1 = Car/carpool/vanpool
2 = Bus
3 = Car/bus combination
4 = Motorcycle
5 = Bicycle
6 = Walk
7 = School bus
8 = Walk to school/work
9 = Metro vanpool
10 = Ferry/walk
11 = Ferry
12 = Carpool
13 = Ferry/car

BUSPASS

1 = Have transit pass
0 = NOT have transit pass

LICENSE

1 = Have valid driver license
0 = NOT have valid driver license

CHANGE

0 = New panel member or no previous employment status
1 = Same workplace as previous wave
2 = Changed workplace from previous wave
3 = Different workplace than previous wave, same workplace as wave before that
4 = No record in previous wave, different workplace than wave before that
5 = No record in previous wave, same workplace as wave before that

WKTRACT

-8 = Workplace varies, or more than one current workplace

SAMPLE

1 = SOV
2 = Bus
3 = Carpool

INCCAT

Wave 1 and wave 2:

1 = < $7,500
2 = $7,500-15,000
3 = $15-25,000
4 = $25-30,000
5 = $30-35,000
6 = $35-50,000
7 = $50-70,000
8 = $70,000+
0 = < $30,000
9 = $30,000+

Wave 3 and wave 4:

1 = < $10,000
2 = $10,000-15,000
3 = $15-25,000
4 = $25-35,000
5 = $35-45,000
6 = $45-55,000
7 = $55-75,000
8 = $75,000+
9 = < $35,000
10 = $35,000+

HHTYPE (Life cycle)

1 = Any child < 6
2 = All children 6-17
3 = 1 adult, < 35
4 = 1 adult, 35-64
5 = 1 adult, 65+
6 = 2+ adults, < 35
7 = 2+ adults, 35-64
8 = 2+ adults, 65+

TFLAG

0 = No change has been changed
1 = Adding a returning home trip
2 = Change order of a trip sequence
111 = BEGTIME and ENDTIME have been changed from 24xx to __xx, and MINUTES is the same
112 = BEGTIME and ENDTIME have been changed from 12xx to __xx, or vice versa, and MINUTES is the same
113 = BEGTIME and ENDTIME have been changed not in the same ways as in 111 and 112, but MINUTES is the same
121 = BEGTIME and ENDTIME have been changed from 24xx to __xx, and MINUTES has been changed accordingly
122 = BEGTIME and ENDTIME have been changed from 12xx to __xx, or vice versa, and MINUTES has been changed accordingly
123 = BEGTIME and ENDTIME have been changed not in the same ways as in 111 and 112, and MINUTES has been changed accordingly
211 = BEGTIME has been changed from 24xx to __xx, and MINUTES is the same
212 = BEGTIME has been changed from 12xx to __xx, or vice versa, and MINUTES is the same
213 = BEGTIME has been changed not in the same ways as in 111 and 112, but MINUTES is the same
221 = BEGTIME has been changed from 24xx to __xx, and MINUTES has been changed accordingly
222 = BEGTIME has been changed from 12xx to __xx, or vice versa, and MINUTES has been changed accordingly
223 = BEGTIME has been changed not in the same ways as in 111 and 112, and MINUTES has been changed accordingly
311 = ENDTIME has been changed from 24xx to __xx, and MINUTES is the same
312 = ENDTIME has been changed from 12xx to __xx, or vice versa, and MINUTES is the same
313 = ENDTIME has been changed not in the same ways as in 111 and 112, but MINUTES is the same
321 = ENDTIME has been changed from 24xx to __xx, and MINUTES has been changed accordingly
322 = ENDTIME has been changed from 12xx to __xx, or vice versa, and MINUTES has been changed accordingly
323 = ENDTIME has been changed not in the same ways as in 111 and 112, and MINUTES has been changed accordingly
401 = Switch BEGTIME and ENDTIME

PFLAG

0 = No change has been made
2 = SUBID has been changed
3 = PERS has been changed
7 = AGE_GP has been changed according to AGE.

HFLAG

0 = No change has been made
1 = Assign TOT6_17 = 0 and TOT1_5=0 when HHSIZE = 1, TOTADULT = 1, and both TOT6_17 and TOT1_5 are missing
6 = HHTYPE has been changed according to TOTADULT, TOT6_17, and TOT1_5
7 = HHSIZE has been changed
9 = TOT6_17 has been changed
10 = TOT1_5 have been changed

A.13 1990 Highway and Transit Skims

A.13.1 Highway skim

This file contains highway skims extracted from the travel demand forecast model of PSRC.

NUMBER OF OBSERVATIONS:	832×832
NUMBER OF FIELDS:	8
LENGTH	55

# VARIABLE	LENGTH	BEGIN COLUMN	DESCRIPTION
1 FM_TAZ	3	1	From-taz
2 TO_TAZ	3	4	To-taz
3 ANCPTI	8	7	AM non-car pool time
4 ACPLDI	8	15	AM car pool distance
5 ACPLTI	8	23	AM car pool time
6 DNCPTI	8	31	Daily non-car pool time
7 DNCPLDI	8	39	Daily non-car pool distance
8 DCPLTI	8	47	Daily car pool time

A.13.2 AM transit auto access skim

This file contains a.m. (peak) transit auto access skims extracted from the travel demand forecast model of PSRC.

NUMBER OF OBSERVATIONS:	832×832
NUMBER OF FIELDS:	9
LENGTH:	63

# VARIABLE	LENGTH	BEGIN COLUMN	DESCRIPTION
1 FM_TAZ	3	1	From-taz
2 TO_TAZ	3	4	To-taz
3 ATRTAU	8	7	AM total transit auto access
4 AIVTAU	8	15	AM in vehicle time auto access
5 AAUXAU	8	23	AM auxiliary transit (drive) time auto access
6 ATWTAU	8	31	AM total wait time auto access
7 AWTLAU	8	39	AM wait for bus time auto access
8 ABRDAU	8	47	AM boarding time auto access
9 ANBRAU	8	55	AM number of boarding (transfer) auto access

A.13.3 AM transit walk access skim

This file contains a.m. (peak) transit walk access skims extracted from the travel demand forecast model of PSRC.

NUMBER OF OBSERVATIONS:	832×832
NUMBER OF FIELDS:	9
LENGTH:	63

# VARIABLE	LENGTH	BEGIN COLUMN	DESCRIPTION
1 FM_TAZ	3	1	From-taz
2 TO_TAZ	3	4	To-taz
3 ATRTWA	8	7	AM total transit walk access
4 AIVTWA	8	15	AM in vehicle time walk access
5 AAUXWA	8	23	AM auxiliary transit (drive) time walk access
6 ATWTWA	8	31	AM total wait time walk access
7 AWTLWA	8	39	AM wait for bus time walk access
8 ABRDWA	8	47	AM boarding time walk access
9 ANBRWA	8	55	AM number of boarding (transfer) walk access

A.13.4 Daily transit auto access skim

This file contains daily (off-peak) transit auto access skims extracted from the travel demand forecast model of PSRC.

NUMBER OF OBSERVATIONS:	832×832
NUMBER OF FIELDS:	9
LENGTH:	65

# VARIABLE	LENGTH	BEGIN COLUMN	DESCRIPTION
1 FM_TAZ	3	1	From-taz
2 TO_TAZ	3	4	To-taz
3 DTRTAU	8	7	Daily total transit auto access
4 DAIVTAU	8	15	Daily in vehicle time auto access
5 DAUXAU	8	23	Daily auxiliary transit (drive) time auto access
6 DTWTAU	8	31	Daily total wait time auto access
7 DWTLAU	8	39	Daily wait for bus time auto access
8 DBRDAU	8	47	Daily boarding time auto access
9 DNBRAU	8	55	Daily number of boarding (transfer) auto access

A.13.5 Daily transit walk access skim

This file contains daily (off-peak) transit walk access skims extracted from the travel demand forecast model of PSRC.

NUMBER OF OBSERVATIONS:	832×832
NUMBER OF FIELDS:	9
LENGTH:	63

# VARIABLE	LENGTH	BEGIN COLUMN	DESCRIPTION
1 FM_TAZ	3	1	From-taz
2 TO_TAZ	3	4	To-taz
3 DTRTWA	8	7	Daily total transit walk access
4 DIVTWA	8	15	Daily in vehicle time walk access
5 DAUXWA	8	23	Daily auxiliary transit (drive) time walk access
6 DTWTWA	8	31	Daily total wait time walk access
7 DWTLWA	8	39	Daily wait for bus time walk access
8 DBRDWA	8	47	Daily boarding time walk access
9 DNBRWA	8	55	Daily number of boarding (transfer) walk access

A.13 Census Demographics

This file contains the following information:

- 1990 population and households extracted from tables P1 and P3 in STF1
- 1992-94 population and households taken from PSRC population and housing estimate reports
- 1990 employment status extracted from table P70 in STF3
- Wage and salary workers taken from PSRC subcounty employment estimates report

NUMBER OF OBSERVATIONS: 564
NUMBER OF FIELDS: 34
LENGTH 191

#	VARIABLE	LENGTH	BEGIN COLUMN	DESCRIPTION
1	TRACT	6 (F6.2)	1	Census tract
2	AREA	7 (F7.3)	6	Area in square miles
3	HH90	5	13	1990 households
4	HDEN90	7 (F7.1)	18	1990 household density
5	POP90	5	25	1990 population
6	PDEN90	7 (F7.1)	30	1990 population density
7	HH92	5	37	1990 households
8	HDEN92	7 (F7.1)	42	1992 household density
9	POP92	5	49	1992 population
10	PDEN92	7 (F7.1)	54	1992 population density
11	HH93	5	61	1993 households
12	HDEN93	7 (F7.1)	66	1993 household density
13	POP93	5	73	1993 population
14	PDEN93	7 (F7.1)	78	1993 population density
15	HH94	5	85	1994 households
16	HDEN94	7 (F7.1)	90	1994 household density
17	POP94	5	97	1994 population
18	PDEN94	7 (F7.1)	102	1994 population density
19	WKRS1990	5	109	Wage & salary workers in 1990, by workplace
20	WKRS1994	5	114	Wage & salary workers in 1994, by workplace
21	EMPLOYED	5	119	Civilian & armed forces employed
22	UNEMPL	5	124	Civilian unemployed
23	NONWKG	5	129	Not in labor force
24	INCOME	6	135	Median household income
25	SF1	5	141	1, detached
26	SF2	5	146	1, attached
27	MF1	5	151	2 units
28	MF2	5	156	3 or 4 units
29	MF3	5	161	5-9 units
30	MF4	5	166	10-19 units
31	MF5	5	171	20-49 units
32	MF6	5	176	50 or more units
33	MH	5	181	Mobile home or trailer
34	OTHER	5	186	Other

APPENDIX B

1990 Travel Diary of PSTP[2]

Mary Smith
00123400
02

MONDAY

OCTOBER 1

Puget Sound Transportation Panel

1990 Travel Diary

Address or Nearest Intersection to Where You Started This Day:

IF YOU DID NOT GO ANYWHERE ON THIS DAY PLEASE CHECK HERE: ☐

THEN WHERE? (Address or Intersection)	WHY? (Work/school/work/dinner/visit/etc.)	HOW? (Car/truck/van/walk/bus/taxi/etc.)	Driver or Rider ?	TIME STARTED	TIME ARRIVED	How many total in your group?	WHO? (Relationship of people with you)	How many family under 15?
1.			D R	AM PM	AM PM			
2.			D R	AM PM	AM PM			
3.			D R	AM PM	AM PM			
4.			D R	AM PM	AM PM			
5.			D R	AM PM	AM PM			
6.			D R	AM PM	AM PM			
7.			D R	AM PM	AM PM			
8.			D R	AM PM	AM PM			
9.			D R	AM PM	AM PM			
10.			D R	AM PM	AM PM			
11.			D R	AM PM	AM PM			
12.			D R	AM PM	AM PM			

Use the back for additional trips.

THINGS TO REMEMBER

- Fill out the diaries for the days indicated on the labels.
- Use a separate line for each stop.
- Record each return trip to home or work.
- The last entry should be your home, or where you were at 1 am.
- When household members travel together, each should record the trip.

[2] Source: Murakami, E., and C. Ulberg (1992).

APPENDIX C

Persons with Trip Records But without Person or Household Information

The variables in the following records are HHID, SUBID, and PERS

Wave 1

3705	0	3

Wave 2

19	0	1
19	0	2
19	0	3
3546	1	1

Wave 3

16	0	1
69	0	1
69	0	2
72	1	1
99	1	1
99	2	1
99	2	2
121	0	1
121	0	2
128	1	1
149	0	1
165	0	1
165	0	2
177	0	1
197	1	1
212	0	3
214	0	1
215	1	1
215	1	2
224	2	1
284	0	1
284	0	2
284	0	3
289	1	1
299	0	1
299	0	2
342	0	4
371	0	1
395	0	1
395	0	2
435	0	1
435	1	1
435	1	2
442	1	1
442	1	2
450	0	1
450	0	2
453	0	1
453	0	2
464	1	2
505	0	1
505	0	2
535	0	2
548	0	2
1039	0	3
1054	1	1
1054	1	2
1088	0	1
1109	0	1
1109	0	2
1129	0	1
1134	0	1
1134	0	2
1138	0	3
1169	0	3
1174	0	1
1174	0	2
1200	0	3
1200	2	1
2011	1	1
2011	1	2
2066	1	1
2139	1	1
2181	0	1
2181	0	2
2229	0	1
2308	0	3
2321	0	1
2344	1	1
2376	0	1
2376	0	2
2468	1	1
2504	1	1
2504	1	2
2536	2	1
2536	2	2
2546	0	1
2546	0	2
2546	0	3
2547	0	1
2547	0	2
2575	0	1
2583	0	4
2589	0	4
2613	1	1
2613	1	2
2643	1	1
2673	0	1
2734	0	2
3023	1	1
3043	0	1
3055	0	1
3055	0	2
3061	0	4
3082	1	1
3101	0	1
3121	0	3
3141	0	1
3141	0	2
3159	0	1
3173	0	3
3206	0	1
3229	1	1
3229	1	2
3287	0	3
3342	0	2
3407	0	1
3470	0	2
3470	0	3
3489	0	1
3489	0	2
3523	0	1
3523	0	2
3526	0	1
3526	0	2
3526	0	3
3526	0	4
3571	0	1
3589	0	4
3589	1	1
3617	0	2
3626	0	1
3626	0	2
3629	0	1
3629	0	2
3631	0	1
3631	0	2
3631	0	3
3639	0	1
3639	0	2
3771	0	1
3771	0	2
3771	0	3
3771	0	4
3771	0	5
3776	0	1
3776	0	2
3776	0	3
3782	0	2
3829	0	1
3829	0	2
3834	0	1
3834	0	2
3847	0	1
3847	0	2
3849	0	1
3871	0	1
3871	0	2
3910	0	1
3910	0	2
3967	0	1
3967	0	2
4029	0	4
4037	0	1
4037	0	2
4037	0	3
4124	0	1
4209	0	1
5041	0	1
5123	0	1
5125	0	1
5125	0	2
5137	0	1
5137	0	2
5178	1	1
5197	0	1
5197	0	2
5197	0	3
5204	0	1
5204	0	2
5231	0	1
5231	0	3
5249	0	1
5249	0	2
5257	0	1
5257	0	2
5268	0	1
5273	0	1
5273	0	2
5283	0	1

5413	0	1
5413	0	2
5505	0	1
5505	0	2
5601	0	1
5601	0	2
5701	0	1
5701	1	1
5701	1	2
5704	0	1
5704	0	2
5706	0	1
5706	0	2
5715	1	1
5715	1	2
5801	0	1
5801	0	2
5907	0	1
5907	0	2
5909	0	1
6019	0	1
6019	0	2
6023	0	1
6023	0	2
6023	0	3
6036	0	1
6105	0	1
6105	0	2
6110	0	1
6111	0	1
6111	0	2
6203	0	1
6203	0	2
6210	0	1
6210	0	2
6224	0	1
6224	0	2
6226	0	1
6226	0	2
6347	0	1
6349	0	1
6349	0	2
6509	0	1
6511	0	1
6511	0	2
6541	0	1
10005	1	1
10005	1	2
10010	0	1
10011	0	1
10011	0	2
10018	0	1
10018	0	2
10022	0	1
10022	0	2
10022	0	3
10038	0	1
10038	0	2
10038	0	3
10041	0	1
10041	0	2

10042	0	1
10042	0	2
10043	0	1
10043	0	2
10044	0	1
10047	0	1
10047	0	2
10048	0	1
10048	0	2
10049	0	1
10049	0	2
10051	0	1
10052	0	1
10054	0	1
10062	0	3
10077	0	1
10077	0	2
10079	0	1
10079	0	2
10088	0	1
10091	0	1
10091	0	2
10092	0	1
10092	0	2
10098	0	1
10098	0	2
10099	0	1
10099	0	2
10107	0	1
10107	0	2
10113	0	1
10113	0	2
10114	0	1
10114	0	2
10116	0	1
10116	0	2
10117	0	1
10120	0	1
10120	0	2
10121	0	1
10121	0	2
10123	0	1
10123	0	2
10124	0	1
10125	0	1
10125	0	2
10125	0	3
10125	0	4
10126	0	1
10127	0	1
10127	0	2
10128	0	1
10128	0	2
10129	0	1
10129	0	2
10130	0	1
10130	0	2
10131	0	1
10133	0	1
10140	0	1
10140	0	2

10142	0	1
10142	0	2
10143	0	1
10149	0	1
10149	0	2
10152	0	1
10159	0	1
10161	0	1
10161	0	2
10172	0	1
10172	0	2
10172	0	3
10177	0	2
10180	0	1
10180	0	2
10193	0	1
10193	0	2
10217	0	1
10217	0	2
10220	0	1
10220	0	2
10234	0	1
10234	0	2
10244	0	1
10249	0	3
10252	0	1
10252	0	2
10254	0	1
10254	0	2
10257	0	1
10259	0	1
10259	0	2
10259	0	3
10259	0	4
10260	0	1
10261	0	1
10261	0	2
10263	0	1
10263	0	2
10264	0	1
10264	0	2
10266	0	1
10266	0	3
10268	0	1
10269	0	1
10270	0	1
10271	0	1
10271	0	2
10272	0	1
10274	0	1
10274	0	2
10276	0	1
10276	0	2
10276	0	4
10279	0	1
10279	0	2
10282	0	1
10282	0	2
10282	0	3
10283	0	1
10283	0	2

10288	0	1
10293	0	1
10293	0	2
10296	0	1
10296	0	2
10302	0	1
10302	0	2
10310	0	1
10310	0	2
10313	0	1
10313	0	2
10334	0	1
10334	0	2
10338	0	1
10338	0	2
10338	0	3
10339	0	4
10339	0	5
10340	0	2
11156	0	2
11396	0	3
12026	0	2

Wave 4

43	0	1
43	0	2
50	0	1
50	0	2
102	0	1
102	0	2
123	1	1
168	0	1
168	0	2
264	1	1
265	0	1
265	0	2
412	0	1
412	0	2
456	1	1
532	1	1
557	0	1
557	0	2
568	0	1
568	0	2
1068	0	1
1068	1	2
1068	1	3
1068	1	4
1080	2	1
1146	0	1
1185	0	1
1185	0	2
1196	0	1
1196	0	2
1197	0	1
1197	0	2
1264	0	1
1264	0	2
1269	0	1
1293	0	1
1293	0	2

2009	1	1	2665	0	1	3708	0	2	11118	0	2
2041	0	1	2700	0	1	3710	0	1	11151	1	3
2139	1	1	2700	0	2	3717	1	1	11202	0	1
2148	0	1	3057	1	1	3754	0	1	11202	0	2
2148	0	2	3084	0	1	3757	0	1	11203	0	1
2155	0	1	3084	1	1	3822	0	1	11203	0	2
2155	0	2	3146	0	1	3822	0	2	11290	0	1
2238	0	1	3146	0	2	3887	0	2	11290	0	2
2238	0	2	3167	0	1	3890	0	1	11290	0	3
2253	1	1	3167	0	2	3890	0	2	11310	0	1
2267	0	1	3195	1	1	3916	0	1	11310	0	2
2267	0	2	3246	0	1	3916	0	2	11315	0	1
2279	1	1	3246	0	2	4018	0	1	11315	0	2
2347	0	1	3275	0	1	4018	0	2	11383	0	1
2347	0	2	3275	0	2	4040	0	1	11383	0	2
2395	0	1	3299	0	1	4040	0	2	11420	0	1
2395	0	2	3299	0	4	5008	0	1	11420	0	2
2395	0	3	3361	0	1	5008	0	3	12202	0	1
2454	0	1	3361	0	2	5113	1	1	12207	1	1
2468	0	1	3365	0	1	10181	0	1	12436	0	1
2468	0	2	3399	0	1	10181	0	2	12436	0	2
2553	1	1	3399	0	2	10186	0	1	12607	0	1
2555	0	1	3454	0	1	10186	0	2	12607	0	2
2578	0	2	3454	0	2	10215	0	1	12617	0	1
2637	0	1	3519	0	1	10215	0	2	12617	0	2
2637	0	2	3519	0	2	10299	0	1	13208	0	1
2657	0	1	3560	1	1	10299	0	2	13208	0	2
2657	0	2	3594	0	1	10300	1	1	13418	0	1
2657	0	3	3594	0	2	10354	1	1	13418	0	2
2657	0	4	3618	1	1	11036	0	1	13418	0	3
2659	0	1	3689	0	1	11039	0	1	14698	0	1
2659	0	2	3708	0	1	11118	0	1	14698	0	2

APPENDIX D

Households with Inconsistent Survey Days

D.l Households that did not have two consecutive days.

Variables in the following records are HHID, SUBID, DAY1, and DAY2.

Wave 1

422	0	2	2
560	0	2	2
616	0	4	4
1122	0	1	1
2337	0	3	3
3030	0	3	2
3729	0	1	1

Wave 2

159	0	4	4
1188	0	4	6
5317	0	7	1
5806	0	2	4

Wave 3

159	0	4	4
1188	0	4	6
5317	0	7	1
5806	0	2	4

Wave 4

433	0	3	3
10035	0	2	2

D.2 Households that did not have the same survey days

Variables in the following records are HHID, SUBID, and days of travel diary in waves 1, 2, 3, and 4, respectively.

Day1

4	0	3	4	4	4
376	0	4	6	6	4
433	0	2	2	2	3
524	0	4	4	4	6
1096	0	3	3	3	4
2151	0	2	2	2	1
2170	0	4	5	5	4
2187	0	4	4	4	1
2377	0	3	3	3	1
2589	0	1	2	2	2
2594	0	1	2	2	2
2595	0	1	2	2	2
2605	0	1	2	2	2
2613	0	1	2	2	2
2617	0	1	2	2	2
2758	0	2	2	2	1
3139	0	2	2	2	3
3181	0	3	4	4	4
3258	0	4	4	4	2
3288	0	4	5	5	4
3292	0	4	5	5	4
3604	0	1	1	1	4
3614	0	2	2	2	4

Day2

4	0	4	5	5	5
159	0	5	4	4	5
376	0	5	7	7	5
524	0	5	5	5	7
560	0	2	3	3	3
1096	0	4	4	4	5
1188	0	5	6	6	5
2151	0	3	3	3	2
2170	0	5	6	6	5
2187	0	5	5	5	2
2337	0	3	4	4	4
2377	0	4	4	4	2
2589	0	2	3	3	3
2594	0	2	3	3	3
2595	0	2	3	3	3
2605	0	2	3	3	3
2613	0	2	3	3	3
2617	0	2	3	3	3
2758	0	3	3	3	2
3139	0	3	3	3	4
3181	0	4	5	5	5
3258	0	5	5	5	3
3288	0	5	6	6	5
3292	0	5	6	6	5
3604	0	2	2	2	5
3614	0	3	3	3	5

APPENDIX E

Persons without Household Sociodemographics

Variables in the following records are HID, SUBID, and PERS.

```
Wave 2
  383 1 1
 2327 1 1
 2327 1 2
 2327 1 3
 2742 1 1
 3546 1 1
 3810 1 1
 3810 1 2
 3922 1 1
 3922 1 2
 3933 1 1
 3933 1 2
 5005 0 1
 5005 0 2
 5010 0 1
 5010 0 2
 5010 0 3
 5015 0 1
 5015 0 2
 5105 0 1
 5105 0 2
 5111 0 1
 5111 0 2
 5116 0 1
 5116 0 2
 5116 0 3
 5116 0 4
 5121 0 1
 5121 0 2
 5124 0 1
 5133 0 1
 5133 0 2
 5145 0 1
 5145 0 2
 5155 0 1
 5155 0 2
 5159 0 1
 5163 0 1
 5163 0 2
 5188 0 1
 5188 0 2
 5208 0 1
 5208 0 2
 5225 0 1
 5225 0 2
 5225 0 3
 5225 0 4
 5230 0 1
 5240 0 1
 5240 0 2
 5243 0 1
 5243 0 2
 5260 0 1
 5260 0 2
 5260 0 3
 5260 0 4
 5270 0 1
 5271 0 1
 5271 0 2
 5303 0 1
 5303 0 2
 5307 0 1
 5307 0 2
 5309 0 1
 5309 0 2
 5318 0 1
 5319 0 1
 5419 0 1
 5419 0 2
 5511 0 1
 5511 0 2
 5615 0 1
 5719 0 1
 5803 0 1
 5803 0 2
 6025 0 1
 6025 0 2
 6028 0 1
 6117 0 1
 6137 0 1
 6137 0 2
 6202 0 1
 6202 0 2
 6202 0 3
 6219 0 1
 6219 0 1
 6219 0 2
 6219 0 2
 6230 0 1
 6230 0 2
 6321 0 1
 6401 0 1
 6401 0 2
 6403 0 1
 6403 0 2
10007 0 1
10007 0 2
10012 0 1
10012 0 2
10013 0 1
10013 0 2
10015 0 1
10017 0 1
10017 0 2
10017 0 3
10021 0 1
10021 0 2
10023 0 1
10023 0 2
10024 0 1
10040 0 1
10040 0 2
10055 0 1
10055 0 2
10057 0 1
10057 0 2
10068 0 1
10068 0 2
10078 0 1
10078 0 2
10081 0 1
10081 0 2
10081 0 3
10083 0 1
10083 0 2
10087 0 1
10095 0 1
10095 0 2
10096 0 1
10096 0 2
10096 0 3
10096 0 4
10096 0 5
10101 0 1
10101 0 2
10103 0 1
10103 0 2
10103 0 3
10108 0 1
10109 0 1
10109 0 2
10109 0 3
10110 0 1
10110 0 2
10139 0 1
10139 0 2
10139 0 3
10154 0 1
10154 0 2
10166 0 1
10166 0 2
10167 0 1
10171 0 1
10171 0 2
10179 0 1
10179 0 2
10179 0 3
10185 0 1
10185 0 2
10189 0 1
10189 0 2
10189 0 3
10189 0 4
10191 0 1
10191 0 2
10194 0 1
10194 0 2
10194 0 3
10194 0 4
10195 0 1
10201 0 1
10201 0 2
10202 0 1
10202 0 2
10202 0 3
10203 0 1
10203 0 2
10205 0 1
10207 0 1
10207 0 2
10207 0 3
10211 0 1
10211 0 2
10211 0 3
10212 0 1
10212 0 2
10214 0 1
10214 0 2
10222 0 1
10222 0 2
10225 0 1
10231 0 1
10231 0 2
10236 0 1
10236 0 2
10239 0 1
10239 0 2
10241 0 1
10241 0 2
```

10246 0 1
10247 0 1
10256 0 1
10256 0 2
10265 0 1
10267 0 1
10267 0 2
10273 0 1
10273 0 2
10280 0 1
10280 0 2
10285 0 1
10285 0 2
10287 0 1
10287 0 2
10292 0 1
10295 0 1
10295 0 2
10295 0 3
10295 0 4
10297 0 1
10297 0 2
10301 0 1
10301 0 2
10304 0 1
10304 0 2
10306 0 1
10306 0 2
10306 0 3
10306 0 4
10308 0 1
10311 0 1
10311 0 2
10316 0 1
10316 0 2
10319 0 1
10319 0 2
10320 0 1
10320 0 2
10321 0 1
10321 0 2
10322 0 1
10322 0 2
10322 0 3
10324 0 1
10324 0 2
10325 0 1
10325 0 2
10325 0 3
10331 0 1
10331 0 2
10331 0 3
10332 0 1
10332 0 2
10333 0 1
10333 0 2
10336 0 1
10336 0 2
10341 0 1
10341 0 2
10341 0 3
10351 0 1
10351 0 2
10351 0 3
10351 0 4

Wave 3
2344 1 2
2468 1 1
5178 1 1

Wave 4
43 0 1
43 0 2
50 0 1
50 0 2
102 0 1
102 0 2
123 1 1
168 0 1
168 0 2
264 1 1
265 0 1
265 0 2
412 0 1
412 0 2
456 1 1
532 1 1
557 0 1
557 0 2
568 0 1
568 0 2
1068 1 1
1080 2 1
1146 0 1
1185 0 1
1185 0 2
1196 0 1
1196 0 2
1197 0 1
1197 0 2
1264 0 1
1264 0 2
1269 0 1
1293 0 1
1293 0 2
2009 1 1
2041 0 1
2139 1 1
2148 0 1
2148 0 2
2155 0 1
2155 0 2
2238 0 1
2238 0 2
2253 1 1
2267 0 1
2267 0 2
2279 1 1
2347 0 1
2347 0 2
2395 0 1
2395 0 2
2395 0 3
2454 0 1
2468 0 1
2468 0 2
2553 1 1
2555 0 1
2578 0 2
2637 0 1
2637 0 2
2657 0 1
2657 0 2
2657 0 3
2657 0 4
2659 0 1
2659 0 2
2665 0 1
2700 0 1
2700 0 2
3057 1 1
3084 0 1
3084 1 1
3146 0 1
3146 0 2
3167 0 1
3167 0 2
3195 1 1
3246 0 1
3246 0 2
3275 0 1
3275 0 2
3299 0 1
3299 0 4
3361 0 1
3361 0 2
3365 0 1
3399 0 1
3399 0 2
3454 0 1
3454 0 2
3519 0 1
3519 0 2
3560 1 1
3594 0 1
3594 0 2
3618 1 1
3689 0 1
3708 0 1
3708 0 2
3710 0 1
3717 1 1
3754 0 1
3757 0 1
3822 0 1
3822 0 2
3887 0 2
3890 0 1
3890 0 2
3916 0 1
3916 0 2
4018 0 1
4018 0 2
4040 0 1
4040 0 2
5008 0 1
5008 0 3
5113 1 1
10181 0 1
10181 0 2
10186 0 1
10186 0 2
10215 0 1
10215 0 2
10299 0 1
10299 0 2
10300 1 1
10354 1 1
11036 0 1
11039 0 1
11118 0 1
11118 0 2
11151 1 3
11202 0 1
11202 0 2
11203 0 1
11203 0 2
11290 0 1
11290 0 2
11290 0 3
11310 0 1
11310 0 2
11315 0 1
11315 0 2
11383 0 1
11383 0 2
11420 0 1
11420 0 2
12202 0 1
12207 1 1
12436 0 1
12436 0 2
12607 0 1
12607 0 2
12617 0 1
12617 0 2
13208 0 1
13208 0 2
13418 0 1
13418 0 2
13418 0 3
14698 0 1
14698 0 2

APPENDIX F

Person&with Inconsistent AGE and AGE GP

F.l Persons with inconsistent age

Variables in the following records are **HHID** SUBID, PERS, and AGE in wave 1 through wave 4.

132	0	2	87	86	88	89
143	0	1	49	50	47	48
330	0	2	65	66	71	70
440	0	2	55	53	55	56
472	0	2	55	54	55	57
499	0	1	42	43	47	46
499	0	2	42	43	47	46
585	0	1	67	68	67	70
1021	0	2	80	87	82	84
1022	0	2	36	27	39	40
1108	0	1	53	53	59	57
1169	0	1	50	48	50	51
1169	0	2	52	50	52	53
1188	0	1	62	63	55	56
1256	0	2	81	83	80	85
2112	0	1	27	25	26	27
2112	0	2	36	35	36	37
2221	0	1	33	53	36	37
2264	0	2	49	50	47	48
2339	0	1	44	43	45	46
2378	0	1	40	39	43	44
2449	0	4	55	56	59	56
2564	0	2	35	36	34	39
2600	0	2	53	51	53	54
2643	0	1	47	45	47	48
3009	0	2	67	64	66	64
3148	0	2	32	48	45	45
3192	0	2	38	35	36	37
3288	0	1	52	53	55	54
3356	0	2	49	50	49	53
3358	0	2	38	34	42	43
3371	0	1	69	67	73	73
3424	0	1	54	53	56	57
3424	0	2	54	53	57	58
3609	0	2	46	44	46	47
3609	0	4	17	16	18	19
3708	0	1	24	23	25	26
3748	0	2	58	57	59	60
3830	0	1	57	56	60	61
3830	0	2	63	62	65	67
3922	0	1	33	34	26	37
4002	0	1	36	37	49	40
4204	0	2	51	22	24	25

F.2 Persons with inconsistent AGE-GP

Variables in the following records are HHID, SUBID, PERS, AGE-GP in wave 1 through wave 4.

440	0	2	6	5	6	6
472	0	2	6	5	6	6
1022	0	2	4	3	4	4
1159	0	2	6	5	5	5
2221	0	1	3	5	4	4
2304	0	1	5	4	4	5
2316	0	1	5	4	4	5
2564	0	2	4	4	3	4
2690	0	1	7	6	7	7
3009	0	2	7	6	7	6
3282	0	2	6	5	5	5
3288	0	1	5	5	6	5
3358	0	2	4	3	4	4
3609	0	2	5	4	5	5
3716	0	2	6	3	3	3
3740	0	2	3	1	1	2
4002	0	1	4	4	5	4
4204	0	2	5	2	2	3

Received June 20, 1996

Travel Trends Using the Puget Sound Panel Survey: A Generalized Estimating Equations Approach

Julie L. Yee[1] and Debbie A. Niemeier[2]

[1] Doctoral Candidate, Division of Statistics, University of California, Davis
[2] Asst. Professor, Dept. Of Civil and Environmental Engineering, University of California, Davis, 95616

ABSTRACT. This paper examines longitudinal mode use trends using four waves of the Puget Sound Transportation Panel (PSTP). The analysis is conducted using generalized estimating equations (GEE) for model estimation. In addition to examining mode use frequencies over time, we also consider mode use trends conditioning on household income and lifecycle stage. As expected, results indicate an overall increase in the number of worktrips made between 1989 and 1993 and these trips were marked by increasing use of single occupancy vehicles. The full parameters of the model were also to estimate the rate of increase in terms of percentage increase (PI) and their confidence intervals. Results indicate that the mean number of worktrips made by driving alone significantly increased from wave 1 to wave 4; with a 95% C.I. the rate of percent increase was estimated between 8.2% and 24.5%. The ranges for rates of change in HOV-pool and non-motor worktrip frequencies overlap with the range for SOV rate of change, and thus, it cannot be said that rate of change for these modes was significantly different from the SOV rates of change. The rate of change in the mean frequency for HOV-transit is not only below the range for SOV trips but also suggests, with 95% confidence, the rate of percent decrease was between 2.88% and 44.0%.

1.0 INTRODUCTION

Recent summary statistics, based on the Nationwide Personal Transportation Study (NPTS), suggest that single occupancy vehicle trips in U.S. have increased in the past decade while public transport trips have declined. In 1983, single occupancy vehicles accounted for 71.1% of all journey to work trips, increasing to 83.0% in 1990. Concomitantly, between 1983 and 1990, public transport journey to work trips decreased from 4.5% to 4.0% (Hu and Young, 1993). Likewise, U.S. census data suggests that private vehicle use for the journey to work trip increased by approximately 27% between 1960 and 1990 while public transit use declined by as much as 60% during the same period (Rossetti and Eversole, 1993).

Much of the mode use literature associated with trends analysis relies on cross-section data to describe the various longitudinal travel patterns. The potential problems associated with travel behavior **analysis** using cross-section data have been well documented (e.g., Kitamura, 1990) and include, for example, lack of temporal insight and omitted and confounding variables (Golob, 1990). More recently, researchers have explored travel behavior with panel data, thus overcoming some of the difficulties associated with cross-section data. However, much of this research has been used to examine the effects of specific transportation policies on travel behavior.

For example, Kitamura et al (1990) examined how travel patterns change when telecommuting is considered using panel data. Others have examined the effects of staggered work hours on travel behavior (Golob and Guiliano, 1989) and changes in mobility (van Wissen and Meurs, 1989) using panel data. Perhaps the most notable panel data is the Dutch National Mobility Panel which includes weekly travel diaries and household and personal demographic data. Golob (1990) used the Dutch Panel to examine the relationship between travel time expenditures and car ownership while Meurs (1990) examined the characteristics of trip generation.

Absent from the travel behavior literature is a U.S. based analysis of mode use trends using panel survey data. The purpose of this paper is to examine mode use trends using four waves of the Puget Sound Transportation Panel (PSTP). The analysis is conducted using generalized estimating equations (GEE) for model estimation. In addition to examining mode use, we also consider mode use trends conditioning on exogenous variables such as household income and lifecycle stage. The paper begins with a brief description of generalized estimating equations. In Section 3, the empirical setting is discussed followed by the model specification. In Section 4, the results are described and finally, conclusions are presented.

2.0 GENERALIZED ESTIMATING EQUATIONS

The analysis of longitudinal count data requires specific statistical techniques (e.g., Kitamura, 1990; Maddala, 1987). The standard linear model $\mu_k = \mathbf{X}'_k \beta$ is limited in two basic respects. First, the range of $\mu_k = E(Y,)$ is not restricted; this hinders practical use in modeling, for example, count data or proportions. Second, the standard linear model assumes independent normally distributed errors with error variances independent of μ_k. Each individual *k* contributes correlated observations to the full likelihood of the panel data. Without understanding the nature of the correlation between these observations, the contribution to the likelihood by each subject cannot be known, much less used to computationally fit a model.

Generalized linear models (GLM) addresses these limitations by specifying a monotone differential link function $g(\mu_k)$ which is equated to $\mathbf{X}_k'\beta$, where $\mathbf{X}_k$ is a vector of covariate measures and β is a vector of coefficients relating $\mathbf{X}_k$ to $g(\mu_k)$. A general discussion of the model foundation may be found in McCullagh & Nelder (1989). $\mathbf{Y}_k$ is assumed to have a known distribution in the exponential family with a variance that is a known function of μ_k. For example, proportions may be modeled as a logit model with link function, $g(\mu_k) = \log(\frac{\mu_k}{1-\mu_k}) = \mathbf{X}_k'\beta$, thus restricting each component of μ_k to the range [0,1].

Suppose the interest lies in evaluating how Y_{ij}, the frequency of trips made during a specified wave *j* and of mode type *i*, changes over time, where $i=1,\dots,I$ and $j=1,\dots,J$. We begin by specifying a distribution for the Y_{ij}'s. If we assume a Poisson distribution, the canonical link is the natural log function with mean and variance $\exp(\mathbf{X}_{ij}'\beta)$. A general wave-mode GLM may now be formulated as:

$$\log(\mu_{ij}) = \beta_0 + \beta_1 wave_j + \sum_{s=1}^{I} \beta_{2,s} I\{s=i\} + \sum_{s=1}^{I} \beta_{3,s} wave_j * I\{s=i\} \qquad (1)$$

where μ_{ij} is the mean of Y_{ij} and $1\{\cdot\}$ is the mode indicator function. The GLM estimate for the vector β can be obtained by using the estimating equation,

$$\mathbf{D}'\mathbf{H}^{-1}(\mathbf{Y} - \mu) = 0 \qquad (2)$$

where $\mu = g^{-1}(\mathbf{X}'\beta)$, **D** represents the matrix $\frac{\partial \mu}{\partial \beta}$ and, assuming independence between observations, **H** is a diagonal covariance matrix. With a panel data model specified as in equation 1, and assuming *I* modes and *J* waves, there are *IJ* correlated observations and, without modifications, the standard GLM should not be applied.

Liang and Zeger (1986a, 1986b) extend the GLM estimating equations to account for correlated measurements using quasi-likelihood. The approach, known as Generalized Estimating Equations (GEE), allows the distribution of $\mathbf{Y}_k$ to be known only to the extent that the mean can be reasonably expressed by $\mu_k = g^{-1}(\mathbf{X}'_k\beta)$ with the variance expressed as a function of the mean, $\mathbf{H}(\mu_k)/\phi$, where ϕ is a scale parameter. Further, in the GEE approach, an empirical variance function is estimated so that model parameters and their variances are consistently estimated even if the correlation structure is misspecified.

The basic form of the quasi-likelihood estimating functions for a correlated response vector is defined for the *IJ*-vector $\mathbf{Y}$ measured from a single individual. The coefficients may be estimated by implementing the score function given by,

$$\mathbf{U}(\beta) = \mathbf{D}'\phi\mathbf{H}^{-1}(\mathbf{Y} - \mu) \qquad (4)$$

where **D** is a $IJ \times p$ matrix where the $(s,t)^{th}$ element is given by $\partial\mu_s / \partial\beta_t$, i.e., the partial derivative of the s^{th} component of the *IJ*-dimensional vector μ with respect to the t^{th} component of the *p*-dimensional vector β. The function **U** has properties which are similar to the derivative of the log-likelihood function. For this reason, the estimating equations,

$$\mathbf{U}(\beta) = 0$$

are known as quasi-likelihood estimating equations. The covariance matrix for U is

$$i_\beta = -\mathrm{E}\left(\frac{\partial}{\partial\beta}\mathbf{U}\right)\Bigg|_\beta = \mathbf{D}'\mathbf{H}^{-1}\mathbf{D}\phi$$

which is analogous to the Fisher information derived for the usual ordinary likelihood functions. The covariance matrix for the estimates 6 is given by the inverse

$$\mathrm{Cov}(\hat{\beta}) = i_\beta^{-1} = \left(\mathbf{D}'\mathbf{H}^{-1}\mathbf{D}\right)^{-1}/\phi\,.$$

The quasi-likelihood can be extended to a *IJK* dimensional model vector encompassing the measurements from K individuals. Assuming that respondents are independent, then equation (4) represents the contribution of each individual to the quasi-likelihood. The estimating equations are then,

$$U(\beta) = \sum_{k=1}^{K} \mathbf{D}_k{}'\phi\, \mathbf{H}_k{}^{-1}(\mathbf{Y}_k - \mu_k) = 0 \qquad (5)$$

with solution denoted by $\hat{\beta}$. The corresponding covariance matrix for $\hat{\beta}$ is

$$\left[-\mathrm{E}\left(\frac{\partial}{\partial\beta}\mathbf{U}\right)\right]^{-1}\Bigg|_\beta = \left(\sum_{k=1}^{K}\mathbf{D}_k{}'\mathbf{H}_k{}^{-1}\mathbf{D}_k\right)^{-1}\phi\,.$$

The computational solution of the GEE, equation 5, requires iteration using an initial estimate of $\hat{p}^{o}$ and the recursive assignment,

$$\hat{\beta}^{l+1} = \hat{\beta}^{l} + \left[\sum_{k=1}^{K} \hat{\mathbf{D}}_k \hat{\mathbf{H}}_k^{-1} \hat{\mathbf{D}}_k\right]^{-1} \left[\sum_{k=1}^{K} \hat{\mathbf{D}}_k{}' \hat{\mathbf{H}}_k^{-1} (\mathbf{Y}_k - \hat{\mu}_k)\right],$$

where $\hat{\mathbf{D}}_k$, $\hat{\mathbf{H}}_k$, and $\hat{\mu}_k$ are functions of β and updated at every iteration using the current estimator $\hat{\beta}^{l}$. If $\hat{\beta}^{o}$ is sufficiently close to $\hat{\beta}$, then the sequence of $\hat{\beta}^{l}$ converges to $\hat{\beta}$. The estimated covariance matrix for $\hat{\beta}$ is given by

$$\hat{i}_\beta^{-1} = \left[-\mathrm{E}\left(\frac{\partial}{\partial\beta}\mathbf{U}\right)\right]^{-1}\Bigg|_{\beta=\hat{\beta}} = \left(\sum_{k=1}^{K} \mathbf{D}_k{}' \mathbf{H}_k^{-1} \mathbf{D}_k\right)^{-1} \hat{\phi}, \qquad (6)$$

where $\hat{\phi}$ is the estimate for the scale parameter and defined by

$$\hat{\phi}^{-1} = \frac{1}{IJK - p} \sum_{k=1}^{K} \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(Y_{ijk} - \hat{\mu}_{ijk})^2}{\hat{\mu}_{ijk}} = \frac{X^2}{IJK - p}$$

where X^2 is the generalized Pearson statistic (McCullagh & Nelder 1989).

In specifying the covariance matrix functional **H,** the generalized estimating equations (GEE) can be used to treat longitudinal data where H is not only a function of μ but also of an additional set of parameters. Specifying **H** in the form suggested by Liang and Zeger (1989b) leads to

$$\mathbf{H}_k = \mathbf{A}_k^{1/2} \mathbf{R} \mathbf{A}_k^{1/2} / \phi$$

where **R** is a suitable correlation matrix for the outcomes, $\mathbf{Y}_k$, and $\mathbf{A}_k$ is the IJ x IJ diagonal matrix with diagonal elements $\phi\,\mathrm{Var}(Y_{ijk})$, $1 \le i \le I$, $1 \le j \le J$. Using equation (6) to estimate the variances of the parameters implies that the correlation matrix **R** has been correctly specified. In practice, it is difficult to ascertain that the true correlation has in fact been specified. Liang and Zeger propose a "working" robust empirical estimate of the correlation structure to protect against misspecification. The alternate covariance matrix estimator is of the form:

$$\mathbf{V} = \left(\sum_{k=1}^{K} \mathbf{D}_k{}'\mathbf{H}_k^{-1}\mathbf{D}_k\right)^{-1} \left(\sum_{k=1}^{K} \mathbf{D}_k{}'\mathbf{H}_k^{-1}\,\mathrm{Cov}(\mathbf{Y}_k)\mathbf{H}_k^{-1}\mathbf{D}_k\right)\left(\sum_{k=1}^{K} \mathbf{D}_k\mathbf{H}_k^{-1}\mathbf{D}_k\right)^{-1} \phi\,. \quad (7)$$

Liang and Zeger also noted that, under mild regularity conditions, $\hat{\beta}$ and V provide consistent estimates even when **H** is misspecified. Furthermore, it can be shown that

$$\sqrt{n}\left(\hat{\beta} - \beta\right) \xrightarrow{L} W \sim \mathrm{Normal}\left(0, \lim_{n\to\infty} n\mathbf{V}\right).$$

This justifies the construction of t-ratios, which for large sample sizes, may be referred to a normal probability table. With data from a finite sample, $\mathrm{Cov}(\mathbf{Y}_k)$ need only be replaced by the *IJ*-dimensional square matrix $(\mathbf{Y}_k - \hat{\mu}_k)(\mathbf{Y}_k - \hat{\mu}_k)'$. In this framework, $\mathbf{H}_k(\mu_k)/\phi$ is known as the "working" covariance matrix. Although $\mathbf{Y}_k$ is believed to contain correlated elements, $\mathbf{H}_k$ can still be specified as if the structure was entirely independent. This would correspond to an independent working correlation matrix and is similar to the solution by ordinary quasi-likelihood with the exception that **V** differs from the covariance estimator in equation (6) by a correction factor,

$$\left(\sum_{k=1}^{K} \hat{\mathbf{D}}_k'\hat{\mathbf{H}}_k^{-1}\,\mathrm{Cov}(\mathbf{Y}_k)\hat{\mathbf{H}}_k^{-1}\hat{\mathbf{D}}_k\right)\left(\sum_{k=1}^{K} \hat{\mathbf{D}}_k'\hat{\mathbf{H}}_k^{-1}\hat{\mathbf{D}}_k\right)^{-1},$$

that makes it robust to departures from the assumption that H is correctly specified. With the PSTP data, this is advantageous in that it relaxes the necessity of understanding the nature of correlation of any two frequencies observed for the same person.

3.0 THE EMPIRICAL SETTING

The PSTP consists of four waves of travel data collected during the years 1989-1993 (see Murakami and Watterson, 1989). Each wave is organized into three data files: one each containing household, person, and trip diary information for every household member of driving age. Each household is represented by a single record in the household files indexed by a household identification number and carrying attribute information such as household income and lifecycle. Similarly, each person is represented by a single record in the person file and is indexed by a household and person identification number. The person files contain profiles of the individual participants with information such as age, sex and occupation. Finally, the trip diary file includes trip attributes for every trip taken during the two day travel period. Each individual trip is described in terms of the trip purpose, mode, and other related attributes; this file is indexed by the household and person identification and trip number.

Using the trip diary files for each wave, trip modes were categorized into four mutually exclusive categories:

- Single occupant vehicle;
- HOV-pool (carpool, vanpool, and taxi);
- HOV-transit (bus and paratransit); and
- non-motor (walk and bike).

Eight additional modes represented in the data were excluded from this analysis. These include motorcycle, school bus, ferry/car, ferry/foot, monorail, boat, train, and airplane and constitute only a small portion of the sample. The analysis is conducted on the subset of respondents participating in all four waves, giving a by-wave sample size of 519 subjects, each with 16 observations (four modes and four waves), for a total of 8304 observations. There are 222 records in which the subject -is associated with the same household as another subject. We will assume that individuals from the same household behave independently of each other.

For the analysis, the data is restricted to work-related trips for those individuals with work trip information in all four waves. Also, in the interest of looking at the subgroups defined by income and lifecycle, the data was further restricted to include only those individuals with reported income and household information in all four waves. The first two waves contained a few households whose incomes were categorized under an alternate scheme indicating only whether they made less than or greater than $30,000 as opposed to the $35,000 cutoff point used in our categorizing scheme. These subjects represented a small proportion of the sample and were also omitted from the analysis. Table 1 presents basic trip summary statistics for each wave.

Table 1. Journey to Work Trip Summary Statistics

	Wave 1	Wave 2	Wave 3	Wave 4
Mean No. of Trips (s.e.)				
SOV	1.92 (0.05)	2.34 (0.10)	2.22 (0.08)	2.33 (0.08)
HOV-Pool	0.34 (0.03)	0.37 (0.04)	0.37 (0.04)	0.41 (0.05)
HOV-Transit	0.23 (0.03)	0.22 (0.03)	0.16 (0.02)	0.18 (0.03)
Non-Motor	0.12 (0.02)	0.17 (0.03)	0.12 (0.02)	0.15 (0.03)
Percent of Total Trips				
SOV	73.8	75.5	77.2	75.8
HOV-Pool	12.9	12.0	12.9	13.5
HOV-Transit	8.6	7.0	5.7	5.8
Non-Motor	4.7	5.5	4.2	4.9

Finally, it is important to also note the limitations of this study. First, survey participants were randomly selected using a stratified sampling protocol based on mode use proportions derived from previous research. Respondents were recruited using three methods: random telephone digit dialing, contacting prior participants in the Seattle Metro transit surveys, and solicitation of volunteers on randomly selected bus routes. The random telephone digit dialing method was the primary way of collecting participants who drive alone or carpool. The latter two methods target transit users. The sample groups were obtained separately and controlling for the proportion of transit users, therefore, it is not appropriate to use this data to compute regionwide mode proportions. Moreover, to conclude that the results of the analysis applies to the general population, the probability of returning travel diaries in all four waves must be assumed independent of travel behavior.

4.0 MODEL SPECIFICATION

The model given by equation (1) may be specified for the PSRC panel data in the following form:

$$\log(\mu_{ijk}) = \beta_0 + \beta_1 j + \beta_{22} 1\{i=2\} + \beta_{23} 1\{i=3\} + \beta_{24} 1\{i=4\} + \beta_{32} j 1\{i=2\} + \beta_{33} j 1\{i=3\} + \beta_{34} j 1\{i=4\} \quad (8)$$

This model reflects the four modes and four waves represented in the panel data and results in a 16K x 9 data matrix, where the rows are made up of the 16 responses from each of the K individuals. The columns contain the response Y , 7 covariates and one identifying variable used to match rows of information corresponding to the same individual. The number of patkneters is *p=8,* but only 7 covariates need to be specified since the usual intercept term, $ß_o$, in the design matrix is automatically included.

Using the mode-wave model specified in *(1),* the response vector is multivariate $\mathbf{Y} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_K)'$, where $\mathbf{Y}_k = (Y_{11k}, Y_{12k}, Y_{13k}, Y_{14k}, Y_{21k}, \ldots, Y_{24k}, Y_{31k}, \ldots, Y_{44k})'$, with mean $\mu_k = E(\mathbf{Y}_k) = \exp(\mathbf{X}_k'\beta)$ where β is a 8 x 1 vector of coefficients to be estimated and $\mathbf{X}_k$ is a 16 x 8 matrix and each row of $\mathbf{X}_k$ is given as

$$X_{ijk} = (1, j, 1\{i=2\}, 1\{i=3\}, 1\{i=4\}, j1\{i=2\}, j1\{i=3\}, j1\{i=4\})', \quad (9)$$

for $i = 1,...,4$ and $j = 1,\ldots,4$, with $1\{\cdot\}$ denoting the indicator function. To avoid over-parameterizing the model, the coefficients β_{21} and β_{31}, corresponding to the effects associated with mode 1, the reference, are not included. The variance-covariance matrix of Y is given by $Cov(Y) = H(\mu) / ø$, where, for the PSTP model, $H(\mu) = Diag(\mu)$.

The analyses can be performed using Splus™ (Statsci, 1995) with programming extensions by Carey and McDermott (1995). To solve GEE problems approximately 50 megabytes of temporary virtual memory is required. The output consists of estimates of the coefficients, and the robust estimate of the variance of the estimated coefficients used by Liang and Zeger. Since the method is semi-parametric, there are no likelihood functions and goodness-of-fit tests are not available. However, this is not expected to be a problem since the model constraints are relaxed by typical standards.

A continuous wave variable is defined to indicate each wave of travel data. Although the wave variable may be treated as a categorical variable with four levels, it is better used as a continuous variable representing the time factor between the four periods. In this way, trends with time can be described by the effects of the wave variable in the statistical model. The model coefficients $ß_0$ and ß, can be interpreted as the intercept and slope for the linear relation between wave and log(μII,) when i =l, i.e., the mode of travel is by car. In particular, the slope ß , quantifies the effect on log(μ_{li}) for increasing wave numbers. Since the effect of wave on log(μ_{li}) is additive, then the effect of wave on μ_{li} is multiplicative. Consider the relationship between the two models corresponding to two consecutive wave numbers j and j +l. Using equations (1) and (3),

$$\begin{aligned}\mu_{1,j+1} &= \exp\{\beta_0 + \beta_1(j+1)\} \\ &= \exp\{\beta_0 + \beta_1 j\}\exp\{\beta_1\} \\ &= \mu_{1j}\exp\{\beta_1\}.\end{aligned}$$

In other words, the average frequency of SOV worktrips changed between any two consecutive waves by a factor of exp{ $ß_1$ } . When i = 2, 3, or 4, the intercept and slope for the relation between wave and $\log(\mu_{ij})$ are adjusted by the coefficients β_{2i} and β_{3i} respectively. So, for modes i = 2, 3, and 4, the intercept and slope for the linear relation between wave and $\log(\mu_{ij})$ are $\beta_0 + \beta_{2i}$ and $\beta_1 + \beta_{3i}$ respectively; the average frequency of trips made by mode i, ($i \neq 1$), changed between any two consecutive waves by a factor of $\exp\{\beta_1 + \beta_{3i}\}$.

As coefficients for wave-related terms, β_1 and $\beta_1 + \beta_{3i}$ are the primary parameters describing the rates of changes occurring in mode frequencies over wave time. The coefficient β_1 is negative if the mean frequency of mode 1 (SOV) worktrips is decreasing, positive if increasing, and zero if there were no changes. Similarly, $\beta_1 + \beta_{3i}$ is negative if the mean frequency of worktrips for mode i (i = 2, 3, or 4) is decreasing, positive if increasing and zero if there is no change. Each of the single parameters β_{3i} may be thought of as a measure for the rate of change over time of mode i frequency relative to that of driving alone.

Since the vector of estimates $\hat{\beta}$ is approximately normal with an approximate covariance matrix given by V, then confidence intervals may be calculated using the covariances supplied by **V**. For $se(\hat{\beta}_1) = \sqrt{\hat{var}(\hat{\beta}_1)}$, where $\hat{var}(\hat{\beta}_1)$ is the diagonal element of **V** which estimates $var(\hat{\beta}_1)$, then a (1-α)100% confidence interval for β_1 is $\left(\hat{\beta}_1 - z_{\alpha/2}se(\hat{\beta}_1), \hat{\beta}_1 + z_{\alpha/2}se(\hat{\beta}_1)\right)$. Thus, a (1-$\alpha$)100% confidence interval for $\exp\{\beta_1\}$, the factor by which μ_{1j} changes between consecutive waves, is

$$\left(\exp\left\{\hat{\beta}_1 - z_{\alpha/2}se(\hat{\beta}_1)\right\}, \exp\left\{\hat{\beta}_1 + z_{\alpha/2}se(\hat{\beta}_1)\right\}\right).$$

Similar confidence intervals for $\exp\{\beta_1 + \beta_{3i}\}$ are

$$\left(\exp\left\{\hat{\beta}_1 + \hat{\beta}_{3i} - z_{\alpha/2}\mathrm{se}(\hat{\beta}_1 + \hat{\beta}_{3i})\right\}, \exp\left\{\hat{\beta}_1 + \hat{\beta}_{3i} + z_{\alpha/2}\mathrm{se}(\hat{\beta}_1 + \hat{\beta}_{3i})\right\}\right),$$

where,

$$\mathrm{se}(\hat{\beta}_1 + \hat{\beta}_{3i}) = \sqrt{\hat{\mathrm{var}}(\hat{\beta}_1) + \hat{\mathrm{var}}(\hat{\beta}_{3i}) + 2\,\hat{\mathrm{cov}}(\hat{\beta}_1, \hat{\beta}_{3i})}\,.$$

The change in SOV mean worktrip frequency between the fourth and first waves may be expressed as $\exp\{(4\text{-}1)\beta_1\}$, and $\exp\{(4\text{-}1)\beta_1 + \beta_{3i}\}$ for modes 2, 3, or 4. The respective confidence intervals are

$$\left(\exp\left\{3\hat{\beta}_1 - z_{\alpha/2}\mathrm{se}(3\hat{\beta}_1)\right\}, \exp\left\{3\hat{\beta}_1 + z_{\alpha/2}\mathrm{se}(3\hat{\beta}_1)\right\}\right)$$

and

$$\left(\exp\left\{3(\hat{\beta}_1 + \hat{\beta}_{3i}) - z_{\alpha/2}\mathrm{se}(3(\hat{\beta}_1 + \hat{\beta}_{3i}))\right\}, \exp\left\{3(\hat{\beta}_1 + \hat{\beta}_{3i}) + z_{\alpha/2}\mathrm{se}(3(\hat{\beta}_1 + \hat{\beta}_{3i}))\right\}\right)$$

where

$$\mathrm{se}\left(3\hat{\beta}_1\right) = \sqrt{3}\,\mathrm{se}\left(\hat{\beta}_1\right)$$

and

$$\mathrm{se}\left(3\left(\hat{\beta}_1 + \hat{\beta}_{3i}\right)\right) = \sqrt{3}\,\mathrm{se}\left(\hat{\beta}_1 + \hat{\beta}_{3i}\right).$$

5.0 RESULTS

Table 2 presents the results for the model specified in equation (8). Notice that the robust standard errors from **V** are generally smaller than those obtained from naively assuming an independence correlation structure, allowing detection of more significant effects. Both the wave and the wave*mode3 terms (corresponding to the coefficients β_1 and β_{33}) are significant while the wave*mode2 and wave*mode4 terms (corresponding to the coefficients β_{32} and β_{34}) are not. These results imply that: (1) the mean frequency of single occupancy vehicle (SOV) worktrips is increasing significantly between waves; (2) the rates at which the mean frequencies of HOV-pools and non-motor worktrips change over time are not statistically different from the rate of increase in the mean frequency of SOV worktrips; and (3) conversely, the rate at which the mean frequency of the HOV-transit worktrips changed over time is statistically different than the rate of change in the frequency of SOV worktrips ($t=-3.0$, $p \leq 0.05$). Moreover, this rate of change $\beta_1+\beta_{33}$ is estimated as -0.1, indicating an overall decrease in the mean frequency over time.

Table 2. Model Results

Variable	*Coefficient*	*Naive s.e.*	*Robust s.e.*	*t-ratio*
intercept	0.67	0.05	0.04	18.75*
wave	0.05	0.02	0.01	4.16*
mode2	-1.81	0.14	0.13	-13.61*
mode3	-2.05	0.17	0.16	-13.15*
mode4	-2.67	0.21	0.20	-13.29*
wave*mode2	0.01	0.05	0.05	0.27
wave*mode3	-0.15	0.07	0.05	-3.03*
wave*mode4	-0.03	0.08	0.07	-0.42

Note: * $p \leq 0.05$, N=8304

Generally, wave effects are reflected by the t-ratios for the wave-related coefficients, β_1, β_{32}, β_{33} and β_{34}, however, examination of these coefficients alone gives only a partial indication of the overall travel trends. More usefully, the full parameters of the model can be used to estimate mean trip frequencies. Table 3 shows the estimated mean frequencies for the four modes for waves 1 and 4 along with estimates of the rate of increase in terms of percentage increase (PI) and their confidence intervals.

The results are consistent with the coefficients and their t-ratios given in Table 2, but provide additional insight. With 95% confidence, the mean number of worktrips made by driving alone significantly increased from wave 1 to wave 4 (1989-1993); the 95% C.I. indicates that the rate of percent increase was between 8.2% and 24.5%. The ranges for rates of change in HOV-pool and non-motor worktrip frequencies overlap with the range for SOV rate of change, corresponding to the insignificant t-ratios for the wave*mode coefficients for these modes; thus, it cannot be said that rate of change for these modes was significantly different from the SOV rates of change. The rate of change in the mean frequency for HOV-transit is not only below the range for SOV trips (as is consistent with Table 2) but also suggests an estimated overall negative rate of change; with 95% confidence, the rate of percent decrease was between 2.88% and 44.0%.

Table 3. Fitted Trip Frequencies[1]

	Wave 1 frequency	*Wave 4 frequency*	*Increase factor*	*Percent increase*	*PI: 95% (lower)*	*PI: 95% (upper)*
SOV	2.04	2.37	1.16	16.07	8.20	24.51
HOV-pools	0.34	0.41	1.21	20.79	-7.72	58.10
HOV-transit	0.23	0.17	0.74	-26.26	-44.00	-2.88
Non-motor	0.14	0.15	1.07	6.75	-27.11	56.35

1. The fitted mean frequencies of worktrips per two consecutive weekdays for wave 1 and wave 4, and the percent increase between the two waves.

The generalized linear model used above for the GEE naturally extends to include additional covariates. The design matrix X and coefficient vector ß are both augmented to include the effects represented by the additional terms. Analogous to ordinary regression modeling, when the interest lies in determining whether or not the values of other covariates are associated with changes in mode frequencies by wave, interaction terms are constructed and included in X and ß.

Income Effects

To test the effects of income, the subjects were identified as having household incomes either greater than or less than $35,000 (the median King County income) and a categorical covariate, Income, was added to the model to assess mode frequencies between the two income groups. This model was fitted with all possible interactions between the three factors: ***Wave, Mode*** and ***Income.*** Table 4 presents the coefficients and their t-ratios.

Table 4. Income Effects

Variable	*Coefficient*	*Naive s.e.*	*Robust s.e.*	*t-ratio*
intercept	0.50	0.10	0.07	7.20*
wave	0.09	0.04	0.03	3.39*
mode2	-1.77	0.28	0.27	-6.56*
mode3	-1.63	0.27	0.24	-6.94*
mode4	-2.66	0.40	0.45	-5.92*
income	0.24	0.12	0.08	2.85*
wave*mode2	-0.07	0.11	0.10	-0.67
wave*mode3	-0.08	0.11	0.08	-1.03
wave*mode4	-0.02	0.16	0.19	-0.08
wave*income	-0.06	0.04	0.03	-1.87
mode2*income	-0.03	0.32	0.31	-0.09
mode3*income	-0.71	0.35	0.33	-2.18*
mode4*income	-0.02	0.47	0.51	-0.04
wave*mode2*income	0.09	0.13	0.11	0.86
wave*mode3*income	-0.04	0.14	0.11	-0.37
wave*mode4*income	-0.01	0.18	0.20	-0.07

* $p \leq 0.05$, N=8304

One striking result is that the two-way interaction between *Wave* and *Mode3* is no longer significant; however, the *Wave* term is still significant. With all other covariates held **fixed,** the rate at which the mean frequency of HOV-transit worktrips changed over time is not statistically different than the rate of change in the mean frequency of SOV worktrips. In the earlier analysis, without the income variable, the interaction effect reveals an overall trend in the PSTP sample for a tendency to make proportionally more SOV worktrips and fewer HOV-transit worktrips. In this analysis, the effect of the two-way interaction between wave and mode is interpreted in the presence of income. For households *with unchanging income status, there* is an increasing number of SOV worktrips over time (the *Wave* term is positive and significant) while the rates of change in worktrip frequencies over time **observed** for the remaining modes are not significantly different from the increase observed for SOV. This can be observed in Tables 5 and 6, where the incomes are examined separately', by noticing that the confidence intervals for

[1] To examine incomes separately, the sample is divided according to income classifications for analyses. For those subjects whose income classification changed during the four waves, the portion of their data vector corresponding to the period in which they made less than $35,000 was included in the analysis for incomes less than $35,000, and the remaining portion was included in the over $35,000.

percentage increase for mode 1 are greater than zero but also overlap with each of the confidence intervals of the other modes, which are not statistically different from zero. The tables also suggest that respondents with higher incomes generally make more worktrips by all modes except for HOV-transit, in which they make considerably fewer trips.

The statistical significance of the two-way interaction between mode 3 and income hints at the difference between the two analyses with regard to the changes in HOV-transit use over time. The analysis indicates that HOV-transit mode use frequencies differ between the income groups; as might be expected, respondents with household incomes greater than $35,000 make fewer work-trips by HOV-transit. Since, by definition, we are looking at the same households over each of the four waves, this implies a possible increase in, the number of households earning greater than $35,000 vice versa. Goulias and Ma (1996) have shown that PSTP household incomes increased between 1989- 1993. As respondents moved from one income group to the next, their use of both HOV-transit and SOV modes for worktrips changed. For respondents remaining in the same income group for each of the four waves, only SOV mode use increased significantly.

Table 5. Fitted Mean Frequencies (<$35,000), N=1952

	Wave 1 frequency	*Wave 4 frequency*	*Increase factor*	*Percent increase*	*PI: 95% (lower)*	*PI: 95% (upper)*
SOV	1.79	2.34	1.30	30.44	11.87	52.10
HOV-pools	0.28	0.30	1.06	5.68	-41.23	90.06
HOV-transit	0.32	0.33	1.01	1.47	-33.10	53.91
Non-motor	0.12	0.15	1.25	24.63	-57.07	261.79

Table 6. Fitted Mean Frequencies (>$35,000), N=6352

	Wave 1 frequency	*Wave 4 frequency*	*Increase factor*	*Percent increase*	*PI: 95% (lower)*	*PI: 95% (upper)*
SOV	2.16	2.37	1.10	9.80	0.89	19.51
HOV-pools	0.36	0.43	1.18	18.04	-12.00	58.35
HOV-transit	0.18	0.14	0.75	-24.81	-49.86	12.76
Non-motor	0.14	0.14	1.01	0.67	-33.17	51.64

Including income and all possible interactions in one model effectively results in a separate model for each of the *two* groups (i.e., each of *the Intercept, Wave, Mode,* and *Wave *Mode* terms are all adjusted by the indicator variable, *Income).* Alternatively, the analysis was performed using separate income groups, with two models estimated. Both analyses are similar in that they estimate the parameters of the same pairs of models. However, in the former (aggregated) analysis, the covariance is assumed homogeneous over the entire dataset. In the latter (separated) analysis, this assumption is not necessary since it fits separate covariances for each income group.

Lifecyle Type Effects

The analysis was repeated with the covariate household type (lifecycle group). The eight group lifecycle types are listed in Table 7.

Table 7. Lifecycle Types

Lifecycle Group	Sample size
(1) Any child, less than 6 years	1284
(2) All children, between 6-17 years	2516
(3) 1 adult, less than 35 years	156
(4) 1 adult, 35-64 years	704
(5) 1 adult, 65 years or greater	52
(6) 2 or more adults, less than 35 years	204
(7) 2 or more adults, 35-64 years	3192
(8) 2 or more adults, 65 years or greater	196

To study the effect of *Wave* for each group in an aggregated analysis involves a model with 64 coefficients and creates computational difficulties. As an alternative, the analysis was run separately for the eight household lifecycle groups. The total sample was divided into 8 sub-samples by lifecycle group using the same analysis technique as noted for income. Results for household types 5 and 8, consisting of adults older than 65 years, failed to converge. These households tend to make fewer worktrips and accordingly, have estimated means close to zero; Poisson data close to zero have variances approaching zero and consequently, not infrequently a near-singular covariance matrix.

The remaining lifecycle groups reveal heterogeneous travel patterns. For lifecycle 1 (households with a child less than 6 years old), there was a significant increase only in the mean HOV-pool use (Table 9), increasing at a rate between 5.19% to 225.93%. The large variances in this subgroup might be attributed to within group heterogeneity. There appears to have been an increase in use of HOV-transit as well (-53%), but this increase is not statistically significant at the type I error level of 0.05.

Table 8. Estimated coefficients for households with any child 6 years or younger.

variables	*coefficients*	*naive s.e.*	*robust s.e.*	*t-ratio*
intercept	0.75	0.12	0.09	8.76*
wave	0.05	0.04	0.04	1.10
mode2	-2.25	0.33	0.34	-6.58*
mode3	-2.88	0.47	0.46	-6.28*
mode4	-2.57	0.48	0.49	-5.21*
wave*mode2	0.16	0.12	0.11	1.41
wave*mode3	0.09	0.17	0.13	0.70
wave*mode4	-0.11	0.19	0.20	-0.55

*p≤0.05

Table 9. Fitted mean frequencies, any child less than 6 years

	wave 1 frequency	*wave 4 frequency*	*increase factor*	*percent increase*	*PI: 95% lower bound*	*PI: 95% upper bound*
SOV	2.22	2.57	1.15	15.47	-10.70	49.32
HOV-pools	0.27	0.51	1.85	85.16	5.19	225.93
HOV-transit	0.14	0.21	1.53	53.27	-25.14	213.79
non-motor	0.15	0.13	0.83	-17.15	-72.76	152.03

For lifecycle 2, households with all children between the ages of 6 and 17, the only significant change was an increase in SOV worktrips of a rate between 6.23% and 39.26% (Table 11). Mean worktrip frequencies for the other modes did not significantly change.

Table 10. Estimated coefficients for households all children between 6 and 17.

variables	*coefficients*	*naive s.e.*	*robust s.e.*	*t-ratio*
intercept	0.62	0.01	0.07	8.68*
wave	0.07	0.03	0.02	2.84*
mode2	-1.58	0.25	0.25	-6.41*
mode3	-2.20	0.33	0.33	-6.67*
mode4	-2.91	0.41	0.54	-5.36*
wave*mode2	-0.09	0.09	0.09	-0.99
wave*mode3	-0.12	0.12	0.11	-1.01
wave*mode4	0.04	0.14	0.19	0.20

*p≤0.05

Table 11. Fitted mean frequencies, children between 6 and 17

	wave 1 frequency	*wave 4 frequency*	*increase factor*	*percent increase*	*PI: 95% lower bound*	*PI: 95% upper bound*
car (alone)	1.98	2.41	1.22	21.63	6.23	39.26
HOV-pools	0.37	0.35	0.93	-7.50	-44.75	54.86
HOV-transit	0.20	0.17	0.86	-14.16	-54.27	61.11
non-motor	0.11	0.15	1.36	-36.23	-53.01	295.97

Households with one adult less than 35 years old, lifecycle type 3, had a tendency towards reduced use of non-motor vehicles for worktrips with a decrease of 87.97% to 95.41%. There were no significant changes with the other mode frequencies although there was a marginally significant reduction in HOV-pools (a 90% CI for the percentage of decrease of worktrips by HOV-pools would be 17.1% to 88.8%).

Table 12. Estimated coefficients for households, 1 adult less than 35 years.

variables	*coefficients*	*naive s.e.*	*robust s.e.*	*t-ratio*
intercept	0.67	0.29	0.17	3.94*
wave	0.00	0.11	0.06	0.05
mode2	-1.57	0.94	0.76	-2.06*
mode3	-2.19	1.11	1.65	-1.33
mode4	-2.04	1.72	0.99	-2.07*
wave*mode2	-0.40	0.44	0.23	-1.71
wave*mode3	-0.26	0.49	0.53	-0.48
wave*mode4	-0.87	1.01	0.09	-9.38*

*p≤0.05

Table 13. Fitted mean frequencies, 1 adult less than 35 years

	wave 1 frequency	*wave 4 frequency*	*increase factor*	*percent increase*	*PI: 95% lower bound*	*PI: 95% upper bound*
car (alone)	1.97	1.98	1.01	0.94	-29.48	44.48
HOV-pools	0.27	0.08	0.30	-69.53	-90.73	0.15
HOV-transit	0.17	0.08	0.47	-53.09	-97.68	850.12
non-motor	0.11	0.01	0.07	-92.57	-95.41	-87.97

For lifecycle group 4, households with 1 adult between 35 and 64 years old, no significant wave effects were found. This was a moderately sized sample at 704, much larger than the groups of 156 and 204 represented by two other groups in this study, both of which had a sufficiently large enough sample size to detect changes between waves. Since the standard errors are relatively large, this suggests the group was very heterogeneous in their travel behavior (Tables 14 and 15).

Table 14. Estimated coefficients, 1 adult between 35 and 64 years.

variables	***coefficients***	***naive s.e.***	***robust s.e.***	***t-ratio***
intercept	0.59	0.18	0.13	4.45*
wave	0.05	0.06	0.04	1.36
mode2	-2.18	0.58	0.45	-4.87*
mode3	-1.11	0.40	0.44	-2.51*
mode4	-3.03	0.70	0.69	-4.39*
wave*mode2	-0.03	0.20	0.16	-0.18
wave*mode3	-0.17	0.15	0.14	-1.21
wave*mode4	0.24	0.22	0.20	1.16

*$p \leq 0.05$

Table 15. Fitted mean frequencies, 1 adult between 35 and 64 years

	wave 1 frequency	***wave 4 frequency***	***increase factor***	***percent increase***	***PI: 95% l. bound***	***PI: 95% u. bound***
car (alone)	1.90	2.23	1.18	17.84	-7.05	49.40
HOV-pools	0.21	0.22	1.08	7.70	-56.34	165.70
HOV-transit	0.53	0.38	0.72	-28.45	-62.93	38.11
non-motor	0.12	0.28	2.41	140.78	-26.73	691.20

For lifecycle group 6, households with more than two or more adults less than 35 years, there was an increased use worktrips made by driving alone with the percentage increase estimated to be between 30.34% and 181.36% (Table 16). The other modes did not have significantly different worktrip frequencies over the waves.

Table 16. Estimated coefficients, 2 or more adults less than 35 years.

variables	*coefficients*	*naive s.e.*	*robust s.e.*	*t-ratio*
intercept	0.42	0.24	0.19	2.18*
wave	0.22	0.10	0.07	3.31*
mode2	-1.56	0.60	0.58	-2.71*
mode3	-2.37	1.15	0.88	-2.69*
mode4	-2.06	0.92	0.77	-2.66*
wave*mode2	-0.07	0.25	0.26	-0.27
wave*mode3	-0.41	0.56	0.18	-2.32*
wave*mode4	-0.33	0.43	0.19	-1.68

*p≤0.05

Table 17. Fitted mean frequencies, 2 or more adults less than 35 years.

	wave 1 frequency	*wave 4 frequency*	*increase factor*	*percent increase*	*PI: 95% lower bound*	*PI: 95% upper bound*
car (alone)	1.89	3.61	1.92	91.50	30.34	181.36
HOV-pools	0.37	0.58	1.56	56.99	-56.07	453.93
HOV-transit	0.12	0.06	0.56	-44.30	-78.11	41.69
non-motor	0.17	0.12	0.72	-28.12	-75.60	111.77

For lifecycle group 7, households with more than one adult between 35 and 64 years old, there was a significant decrease, between 26.48% and 72.76%, in use of HOV-transit. The other mode frequencies show no statistically significant change (Table 19).

Table 18. Estimated coefficients, 2 or more adults between 35 and 64 years.

variables	*coefficients*	*naive s.e.*	*robust s.e.*	*t-ratio*
intercept	0.67	0.09	0.07	9.93*
wave	0.03	0.03	0.02	1.42
mode2	-1.70	0.22	0.20	-8.41*
mode3	-1.74	0.28	0.26	-6.72*
mode4	-2.42	0.33	0.30	-8.10*
wave*mode2	0.03	0.08	0.08	0.43
wave*mode3	-0.30	0.11	0.09	-3.29*
wave*mode4	-0.09	0.12	0.11	-0.82

*p≤0.05

Table 19. Fitted mean frequencies, 2 or more adults between 35 and 64 years.

	wave 1 frequency	*wave 4 frequency*	*increase factor*	*percent increase*	*PI: 95% lower bound*	*PI: 95% upper bound*
car (alone)	2.03	2.24	1.11	10.75	-3.80	27.51
HOV-pools	0.38	0.47	1.23	22.66	-19.65	87.26
HOV-transit	0.26	0.12	0.45	-55.25	-72.76	-26.48
non-motor	0.16	0.14	0.85	-15.21	-53.82	55.67

6.0 CONCLUSIONS

In conclusion, there was an increase in the number of worktrips made between 1989 and 1993 and these were marked by increasing use of single occupancy vehicles. HOV-transit mode use simultaneously declined, an effect apparently associated with the increasing income of the population. The two income groups defined by those making greater or less than $35,000 exhibited comparable mode use behavior with the exception of HOV-transit. Particularly, people in households with the higher income bracket made fewer worktrips by HOV-transit. When considering the factor by lifecycle group, it was found that different household types varied. Results indicated that households having all children between 6 and 17 or households having two or more adults less than age 35 made increasingly more worktrips by car; households with a child less than 6 years old made increasingly more trips by HOV-pool; households with more than one adult between 35 and 64 years made fewer worktrips by HOV-transit; and households with one adult less than 35 years made fewer non-motor worktrips.

The full parameters of the model were also used to estimate the rate of increase in terms of percentage increase (PI) and their confidence intervals. These results suggest that the mean number of SOV worktrips significantly increased from wave 1 to wave 4 with a 95% C.I. for the rate of percent increase of between 8.2% and 24.5%. The rates of change in HOV-pool and non-motor worktrip frequencies overlap with the range for SOV rate of change, and thus, it cannot be said that rate of change for these modes was significantly different from the SOV rates of change. The rate of change in the mean frequency for HOV-transit is not only below the range for SOV trips but also suggests, with 95% confidence, the rate of percent decrease was between 2.88% and 44.0%.

REFERENCES

Carey, V. & McDermott, A., 1995, Harvard Medical School

Golob, T. and Guiliano, G., "A simultaneous equation model of employee attitudes to a staggered work hours demonstration project," Working Paper UCI-ITS-WP-89- 1, Institute of Transportation Studies, University of California, Irvine, CA.

Golob, T., "The dynamics of household travel time expenditures and car ownership decisions," Transportation Research, Part A, 1990, 24A, No. 6: pp. 443-464.

Goulias, KG. & Ma, J., "Analysis of Longitudinal Data from the Puget Sound Transportation Panel -- Task Report," Pennsylvania Transportation Institute, Pennsylvania State University

Hu, P. & Young, .I., 1990 NPTS Databook, Nationwide Personal Transportation Survey, FHWA-PL-94-O 1 OA, November 1993, Office of Highway Information Management, Federal Highway Administration, HPM-40, Washington, DC.

Kitamura, R., "Panel Analysis in Transportation Planning: An Overview," Transportation Research, Part A, 1990,24A, No. 6: pp. 401-5 15.

Kitamura, R., Goulias, IS., and Pendyala, R., "Telecommuting and travel demand: An impact assessment for the State of California Telecommute Pilot Project Participants," Research Report UCD-TRG-RR-90-8, Institute of Transportation Studies, University of California, Davis, CA.

Liang, KY. & Zeger, S.L., "Longitudinal data analysis using generalized linear models," Biometrika, 1986a, 73,l pp. 13-22.

MaddaIa, G., "Recent developments in the econometrics of panel data analysis," Transportation Research, Part A, 1987, 21A, No. 4/5 pp. 303-326.

McCullagh, P. & Nelder, J.A., Generalized Linear Models, 2nd edition, Chapman & Hall 1989.

Meurs, H., "Dynamic analysis of trip generation," Transportation Research, Part A, 1990, 24A, No. 6: pp. 427-442.

Murakami, E. and T, Watterson, "Developing a household travel panel survey for the Puget Sound Region", *Preprint*, 73rd Annual Meeting, Transportation Research Board, Washington, D.C., (1991).

Rossetti, M & Eversole, B., Journey to work trends in the United States and its major metropolitan areas, HW362, November 1993, Office of Highway Information Management, Federal Highway Administration, HPM-40, Washington, DC

Statsci, S-PLUS Programmers, (1993), Mathsoft, Inc., Seattle, WA.

Van Wissen, L. and Meurs H., "The Dutch mobility panel: Experiences and evaluation," Transportation, 16: pp. 99-l 19.

Zeger, S.L., & Liang, KY. "Longitudinal data analysis for Discrete and Continuous Outcomes,,' Biometrics 42, March 1986b, pp. 121-l 30.