LA-UR- 10-07950

Approved for public release; distribution is unlimited.

Title:	TRAVEL DETERMINANTS AND MULTI-SCALE TRANSFERABILITY OF NATIONAL ACTIVITY PATTERNS TO LOCAL POPULATIONS
Author(s):	Kriste M Henson Konstadinos G. Goulias
Intended for:	90th Annual Meeting of the Transportation Research Board, January 23-27, 2011

Los Alamos
NATIONAL LABORATORY
EST. 1943

Form 836 (7/06)

TRAVEL DETERMINANTS AND MULTI-SCALE TRANSFERABILITY OF NATIONAL ACTIVITY PATTERNS TO LOCAL POPULATIONS

Kriste M. Henson
Los Alamos National Laboratory
Decision Applications Division
P.O. Box 1663, MS F609
Los Alamos, NM 87545
Tel (505) 667-0933
kriste@lanl.gov

and

Konstadinos G. Goulias
Department of Geography
University of California – Santa Barbara
3611 Ellison Hall, Santa Barbara, CA 93106
goulias@geog.ucsb.edu

Paper submitted to the 90th Annual Meeting of the Transportation Research Board, Washington D.C.
January 23-27, 2011

ABSTRACT

The ability to transfer national travel patterns to a local population is of interest when attempting to model megaregions or areas that exceed metropolitan planning organization (MPO) boundaries. At the core of this research are questions about the connection between travel behavior and land use, urban form, and accessibility. As a part of this process, a group of land use variables have been identified to define activity and travel patterns for individuals and households. The 2001 National Household Travel Survey (NHTS) participants are divided into categories comprised of a set of latent cluster models representing persons, travel, and land use. These are compared to two sets of cluster models constructed for two local travel surveys. Comparison of means statistical tests are used to assess differences among sociodemographic groups residing in localities with similar land uses. The results show that the NHTS and the local surveys share mean population activity and travel characteristics. However, these similarities mask behavioral heterogeneity that are shown when distributions of activity and travel behavior are examined. Therefore, data from a national household travel survey cannot be used to model local population travel characteristics if the goal to model the actual distributions and not mean travel behavior characteristics.

INTRODUCTION

When developing activity-based transportation models of megaregions or areas that exceed the boundary of a single metropolitan planning organization (MPO), non-typical issues must be addressed. At the core of activity-based models are "synthetic" schedules (computer produced time allocations of individuals and households to activity such as work, leisure, and travel) that are formulated and specified using data from activity/travel diaries. All activity-based simulations require activity schedules to depict a specific city's population movements. The data are collected using activity/travel diaries in which respondents report all their activities over the span of a few days. Confounding the problem of modeling large regions is that travel surveys may not be available for the entire region. In order to model cities or regions where non-overlapping or no data exists, it is important to establish whether other survey data, such as the 2001 National Household Travel Survey (NHTS), can be utilized to model the activity-travel behavior of persons at local levels by "transferring" a survey participant's daily travel schedules to in a different geographic location.

At the core of this research are questions about the connection between travel behavior and land use, urban form, and accessibility. This relationship is essential to establish when designing a process for the transferring activity-travel patterns. This research seeks to first identify a group of land use variables that are essential to define activity and travel patterns for individuals and households along with the well documented socioeconomic characteristics normally utilized in travel behavior research. To make the analysis tractable, persons were grouped by type, travel patterns, and a select combination of variables that best, within the available data resources, capture the nature of the areas where survey participants live. An experiment using NHTS data from the Midwest census region was conducted to compare them with survey data collected within the same region from the Minneapolis/St. Paul, MN metropolitan area. The NHTS data is also compared to a southern California survey data because of potential differences between regions.

REVIEW OF TRANSFERABILITY RESEARCH

It has been a long-held research question and practitioner-unfounded assumption in transportation that individuals with similar sociodemographics, such as occupation, education, income, gender, marital status, and stage in their life cycle, have similar travel patterns (1-6). For example, Oppenheim (1) believed that "if two travelers have the same personal characteristics and live in similar urban environments, they will have the same travel behavior". If Oppenheim's overlying assumption was true, theoretically it may be possible that a person's activity-travel patterns from one area could be applied to a person with similar demographics who has similar landform and accessibility opportunities. Several authors, including Reuscher et al. (7), attempted to define clusters of population based mainly on person and household characteristics. Although in his statement Oppenheim included urban form as a matching factor, it is only recently that researchers have focused more on landform and its effect on travel behavior. With land form, land use has also been increasingly recognized as another fundamental matching characteristic to be included in the process of modeling and simulating activity and travel patterns.

At the core of this discussion is the transfer of information from one source to another. The issue of transferability has been an important topic in transportation modeling for quite some time. In 1976, Atherton and Ben-Akiva published research on the transferability of trip generation models (9). Researchers continued to investigate into this area, both spatially and

temporally (9-12). Other researchers focused on transferring mode choice and trip distribution models both spatially and temporally (13-14) as well as estimating the number of commercial trips made in a day applying a national database to a regional level. Nationally, work is underway to update transferable parameters for simple planning and analysis – *Travel Demand Forecasting: Parameters and Techniques* – Project 8-61 (15).

The ability to utilize activity-travel surveys from a different area when generating synthetic households has been a more recent topic in the transferability arena. Greaves and Stopher (16) developed a simulation that applied a classification and regression tree (C&RT) to group the 1995 NPTS households into relatively homogeneous groups based on six trip purposes and then sampled from frequency distributions using Monte Carlo simulation to synthesize trip production models for Baton Rouge, LA. This methodology was further extended to model Adelaide and Sydney in Australia as well as experimenting with increasing accuracy of the data by including more local characteristics through Bayesian updating (17-19). Based on these studies, Mohammadian et al. (20) experimented with using neural networks to determine household memberships in clusters first determined in two-step cluster analysis routine before estimating average travel measures using Bayesian updating and Monte Carlo simulation. Lin and Long (21) used the NHTS to define neighborhoods at the tract levels with a two-step cluster method. This was then extended to build a two-step random coefficient model (22), which models the transferability of household auto work trip rates. It is worth noting the difference between transferring average characteristics (e.g., average number of trips per day), model parameters (e.g., the value of time in a discrete choice model), and the entire activity and travel schedule of an individual (or a distribution of membership in group of activity and travel patterns). The latter is more demanding in its requirements for closeness between transferable characteristics.

METHODOLOGY

Survey data (activity patterns as well as demographics) from the Midwest potion of the NHTS and demographic information from the 2000 Census SF1 and SF3 data sets were utilized along with accessibility and urban form measurements to develop three separate latent class cluster analysis (LCCA) models representing individual survey participants (*person*), activity and travel patterns (*travel*), and *land use*. These models are used to divide the survey population into specific groups that are a combination of membership in each of the three LCCA models. Finally, to investigate to what degree transferability is possible, the sets of latent class cluster models are replicated using two local surveys: the 2000 Travel Behavior Inventory (TBI) for the Minneapolis/St. Paul, MN metropolitan area and data and the Southern California Association of Governments (SCAG) 2001 Post Census Regional Household Travel Survey (SCAG survey).

National Household Travel Survey

The NHTS is a national travel survey designed to capture comprehensive data on transportation patterns in the U.S. in order to assist transportation planners and policy makers. The survey spanned March 2001 to May 2002. According to the FHWA (23), data were collected on daily trips taken over a 24-hour period with travel days spanning the entire week. These included trip purpose, type of transportation used, trip start and end times, travel time, and the time spent at the activity locations. Also included in the NHTS data were person and household characteristics for each traveler. Person characteristics included age, gender, whether or not survey participant was a driver or a worker, and level of education completed. Household information included the number of household members, number of adults, income level, number of vehicles, number of

drivers, and whether their home is owned or rented. Block group and census tract locations where households resided for the majority of the survey participants are available for research purposes. In order to compare with the TBI data, a subset of 5047 households from the Midwest census region were used.

2000 Travel Behavior Inventory

The TBI was conducted by NuStats for the Twin Cities Metropolitan Council. The Household Inventory Survey (HIS) portion of the TBI randomly selected households between April and September 2001. The HIS measures person trips by motorized and non-motorized modes of transportation within the seven-county metropolitan area and within the 13 counties in Minnesota and Wisconsin that surround the seven-county area. For 8,219 households, members completed a 24-hour weekday travel diary that includes origin and destination locations, travel mode, trip start and end times, and activities at trip destinations. Socioeconomic data, including household size, number of vehicles, household income, dwelling type, age, gender, and employment/school status were also collected. During the survey, the collection of households was controlled through sample management to ensure an adequate representation of all households with respect to household size (1, 2, 3, and 4+ persons) and vehicle ownership (0, 1, 2, and 3+ vehicles) (24).

Post-Census Regional Household Travel Survey

The SCAG survey, administered by NuStats, encompasses households living in the six counties within the SCAG area. Data were collected during Spring 2001, Fall 2001, and Spring 2002 from 16,939 households. Survey participants completed a 24-hour weekday travel diary or a 48-hour travel log that includes one weekend day. Data captured are origin and destination locations, travel mode, time of day, and trip purpose. Socioeconomic data, including household size, number of vehicles, household income range, dwelling type, age, gender, and employment/school status were also collected (25).

Latent Class Cluster Modeling

The modeling focused on creating three sets of LCCA models. These models are a type of latent class model (also known as finite mixture models). They allow the inclusion of one or more discrete unobserved variables, known as latent variables, to classify units of analysis. Latent class models do not depend on traditional modeling assumptions, including the data being linearly related and having a normal distribution. These models identify groupings of data that have similar values, characteristics, and/or behavior. It also includes a k-category latent variable, each category representing a cluster. The clusters produced by the models are latent variables, which could be interpreted to define the latent cluster category. Cases (data) are assigned to a cluster based on membership probabilities estimated directly by an underlying classification model (26).

Person LCCA Models

For each survey participant utilized to identify clusters, a selection of demographics from each survey's person and household tables was assembled. These include:

- Person: age category (0-14, 15-17, 18-22, 23-25, 26-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, and 65-88), gender, whether survey participant was a driver or a worker, and level of education completed;

- Household: number of household members, number of adults, number of children, number of workers, income level, number of vehicles, number of drivers, and whether their home was owned or rented.

Latent Gold 4.5 was used to construct a large number of models with the NHTS data, with clusters ranging in number from one to 17. Based on a set of classification statistics such as the log-likelihood (LL), Bayesian information criterion (BIC), and classification error, an 11-cluster model was selected to classify the person data. Then, using the same set of data and cluster number, LCCA models were generated for the TBI and SCAG surveys. The subset of the variables that most impacted cluster formation are: number of household members, number of adults, number of drivers, number of workers, and number of children. The person variables with the largest impact are the person's age category and whether the person works and if a person has a driver's license. In general, the generated clusters were one of three categories: adults in adults-only households, adults in a household with children, and children. The cluster descriptions and the percentage of persons that are assigned to each category are shown in Table 1.

TABLE 1 Person class cluster definitions and survey person membership

Cluster Type	NHTS	TBI	SCAG
Adult in households with no children			
Single, working adult	4.5%	10.5%	8.0%
Single, retired adult	4.2%	3.6%	4.7%
Adult in married/cohabitating couples, no kids	16.5%	20.8%	16.4%
Adult in married/cohabitating, retired couples	11.6%	7.8%	9.0%
Children			
Child in 2-parent households with only young children	16.6%	17.3%	14.6%
Child in 2-3 adult (extended family) households	6.1%	3.7%	3.6%
Child in single-parent, lower income households	2.5%	2.3%	5.8%
Adult in households with children			
Adult in larger, 2-adult households	15.4%	15.8%	14.5%
Adult in larger, multiple-adult households with fewer children	14.1%	12.1%	13.4%
Adult in non-working adult in multiple-adult households	5.3%	3.1%	4.8%
Adult in smaller, 1-2 adult household with limited income	3.0%	3.1%	5.2%

Travel LCCA Models

To identify clusters of activity and travel behaviors, trip statistics were assembled from the three surveys for each survey participant. These are:

- number of trips by purpose;
- travel time by purpose;
- time spent at the activity locations by purpose;
- travel day of week;
- whether the person started the day at home;
- and whether the person was out of town on the travel day.

The trip purposes used in the modeling are Daycare, Home, Medical, Other, Retail, School, Serve Passenger, Service, Social, and Work. Weather data from the National Climate Data

Center (NCDC) were included to capture possible travel patterns variations due to weather. The travel cluster model has 13 clusters, each representing a unique combination of trip patterns. Each of the surveys has at least two trip patterns that were not generated by the other cluster models. The cluster descriptions and the percent memberships are in Tables 2 and 3. Two separate comparisons are needed, since the NHTS travel cluster categories were combined to match certain TBI and SCAG categories.

TABLE 2 Comparison of NHTS and TBI travel cluster categories

Cluster Type	NHTS	TBI
Retail, Social	15.7%	5.5%
Work, Retail, Social	13.3%	7.6%
No trips	11.5%	14.9%
Service (100%), Retail, Social, Work, etc.	9.4%	21.3%
Work	8.2%	13.7%
Social	7.8%	4.8%
School (100%), Retail, Social, Work, Serve Passenger, etc.	12.7%	6.2%
School (100%), etc.	5.8%	5.2%
Serve Passenger, Retail, Social	5.6%	NA
Work, Serve Passenger, Social, Retail	5.2%	7.7%
Medical (100%), Retail, Social, Serve Passenger	2.9%	NA
Daycare (100%), School, Social, Retail, Serve Passenger, etc.	1.9%	1.6%
Work, Social	NA	5.8%
Service (100%), Work	NA	5.8%

TABLE 3 Comparison of NHTS and SCAG survey travel cluster categories

Cluster Type	NHTS	SCAG
Retail, Social	23.5%	13.8%
Work, Retail, Social	13.3%	9.2%
No trips	11.5%	21.1%
Service (100%), Retail, Social, Work, etc.	9.4%	NA
Work	8.2%	12.6%
Other (100%), Retail, Social, Work	NA	7.3%
School (100%), Retail, Social, Work, Serve Passenger, etc.	7.4%	6.5%
School (100%), Social, etc.	5.3%	NA
School (100%), etc.	5.8%	6.9%
Serve Passenger, Retail, Social	5.6%	11.3%
Work, Serve Passenger, Social, Retail	5.2%	4.4%
Medical (100%), Retail, Social, Serve Passenger	2.9%	4.2%
Daycare (100%), School, Social, Retail, Serve Passenger, etc.	1.9%	NA
School (100%), Serve Passenger, Other, etc.	NA	2.8%

As demonstrated in the tables, the surveys did not all contain the same quantities of activities and trip types. The NHTS and TBI have only a small sample of type other trips while the TBI also contains a few medical trips. The SCAG survey includes only a small number of daycare trips.

Land use LCCA Models

The land use latent cluster models were constructed using a variety of data sets to characterize the areas where the survey households live. The data sets were selected to represent: access to the roadway and its connectivity; employment opportunities; shopping, health care, agriculture and other opportunities; road congestion; crime indicators; mass transit availability; income and housing costs; and climate. Depending on the type of data, each data set is included in the clustering model at multiple geographic resolutions. These are census tracts and block groups from the 2000 Census and travel distance polygons, estimated for 0-5 km, 5-10 km, and 10-50 km travel zones. Travel distances were derived using ESRI's ArcGIS Network Analysis extension. Including multiple geographic levels is done to avoid introducing errors into the analyses. In particular, Ecological Fallacy and the Modifiable Areal Unit Problem (MAUP) are of great concern. Guo and Bhat (27) showed while investigating the effects of MAUP on residential choice location that multinomial Logit models, which utilized parameters for a single scale (block, block group, census tract, or TAZ) suffered from conflicting results due to MAUP. For example, a tract-level model may suggest that households, in general, prefer lower housing values while a block group-level model indicates that high income households prefer high housing values. The inclusion of multiple scales in the clustering model mitigates (but does not solve) this problem. Another issue is data vintage. An attempt was made to gather data near the time frames in which the NHTS, TBI, and SCAG survey data were assembled (approximately March 2001 to May 2002). They are:

Business data: Dunn and Bradstreet (D&B) businesses (dated March 2004) were assigned an industry category (from the 2000 CTPP) based on their SIC: agriculture; retail; educational, health and social services; and other trips. The number of businesses for each industry was summed for the block groups, census tracts, and travel distance zones.

Travel Opportunities and Accessibility: NAVTEQ NAVSTREETS, a national road network circa Quarter 2 2003, is used to characterize potential travel opportunities across the U.S. In order to represent levels of accessibility for block groups, census tracts, and travel distance zones, the number of intersections and total link lengths by function class were summed.

Crime Indicators: It is possible that persons who live in areas of high crime have different travel patterns than those who live in safer areas. No national sources of crime data exist currently. So, a set of demographics that could be determinants of crime, from the 2000 Census SF 1 and SF 3, are identified that appear to be associated with areas of high crime (28). These include measures of low income and poverty, lack of education, areas with larger concentrations of foreign born and minorities, youth, and residential stability.

Population Density: This measure from the SF1 is included for both block groups and census tracts.

Transit: Very few sources of national geospatial bus or other mass transit data exist. Similar to crime statistics, the majority of the data are maintained by local authorities. Only three sources were identified: The Federal Transit Administration (FTA) Bus Route GIS Database was developed by the Moakley Center GeoGraphics Lab at Bridgewater State (28); commuter

rail stations from the NAVTEQ NAVSTREETS 2003 dataset; and the percentage of workers who use transit to travel to work, derived using the SF1.

Other land use data: A few other categories of data representing climate, congestion and connectivity were also used in the modeling process.

Thirteen land use clusters were generated for the NHTS. However, the TBI and SCAG survey do not generate models with the same categories. The areas where the TBI and SCAG survey participants live are both fairly homogenous regions, relative to the NHTS which contains persons from across the Midwest. For this reason fewer land use categories were needed to fully describe the TBI (using only eight clusters) and SCAG (using only six clusters) areas.

The latent clusters were most influenced by a combination of accessibility, transit, and population measures. Other types of measures including potential crime indicators, connectivity, congestion, commuter rail indicators, and climate have minimal impact on the land use clusters. The travel distance service areas have the most influence in cluster creation. So, each cluster is defined as a collection of overall accessibility for the zones. The overall accessibility is a unique combination of accessibility measures (whose definition was based on street length, number of intersections, and number of businesses), population densities, bus route lengths, and income ratios for. Table 4 contains the cluster types, which are a combination of the three accessibility zones (0-5 km, 5-10 km, and 10-50 km) described by a accessibility range (L=low, M=medium, and H=high). The land use ranges from L-L-L (a rural area) to H-H-H (a completely urban area). The last cluster combination, H-MH-ML, represents an area with a high neighborhood accessibility (0-5 km). But, the mid-range (5-10 km) and regional accessibilities (10-50 km) drop to medium-low.

TABLE 4 Comparison of the NHTS, TBI, and SCAG land use latent clusters

Cluster Type	NHTS	TBI	SCAG
L–L–L	14.2%	12.6%	10.8%
L–L–ML	11.2%	8.0%	
MH–MH–H	9.2%	17.1%	NA
ML-L-L	9.0%	NA	NA
ML-M-MH	9.0%	10.6%	NA
MH-MH-MH-H	7.9%	16.5%	15.5%
H-H-H	7.8%	16.8%	15.1%
ML-M-ML	6.1%	NA	18.3%
MH-MH-ML	5.3%	NA	
ML-L-ML	5.7%	6.7%	NA
H-H-H	5.4%	11.7%	13.7%
MH-M-L	4.7%	NA	26.6%
H-MH-ML	4.5%	NA	

The SCAG survey households live in areas that are dominated by high-density land use factors. Overall, the clusters primarily contain a combination of low and high accessibility ranges. Only a few areas had medium-ranged accessibilities. The TBI households live in slightly more diverse areas, as indicated by the larger number of clusters. However, the TBI land use clusters are still dominated by high levels of accessibility which would typically be found in cities. Only three of its eight clusters do not contain any high accessibility zones.

TRANSFERABILITY

To establish whether or not transferability of the NHTS is possible, it was necessary to determine whether people from the national and local travel surveys who share common socioeconomic demographics and live in similar areas have similar travel patterns. This is done by comparing person, travel, and land use cluster combinations for NHTS Midwest survey participants to people with the same cluster combinations in the TBI and SCAG survey. The NHTS travel and land use clusters were compared to their TBI and SCAG survey counterparts by comparing mean activity-travel characteristics for each group of persons cross-classified with each group of residence area. An example analysis and results are discussed below.

Comparison of Average Activity and Travel Behavior

To compare the NHTS to each of the local surveys, the mean activity-travel measures for each survey set (NHTS-TBI and NHTS-SCAG) are examined. As an example, results for one person cluster, persons living in a two adult both employed household with no children, are presented. The average travel statistics for the NHTS and TBI adults in this cluster are located in Table 5. The NHTS persons tend to make more retail and social trips in all land use categories. NHTS suburbanites travel ten minutes to retail locations on average per day and spend a total of 29 minutes at the locations. TBI persons have a daily average of six minute travel times and 12 minute average activity durations. Work and total trip statistics are more similar.

Table 5 also shows the t statistics for the NHTS and TBI clusters. For all persons in this cluster, only the averages for total activity durations, work travel times, and number of work trips had t values that were within the $\pm t_{critical}$ range. In general, subdividing the person clusters by land use type decreased the differences between the two surveys; this is reflected in the lower t statistic values and the greater number of person-land use cluster combinations that are calculated to have the same mean travel values. For example, the t statistic for the total number of trips for all cluster members is 3.04 ($t_{critical} = 1.96$) which is outside the $t_{critical}$ range. Once the cluster is divided by land use cluster memberships, the two sets of groups containing people living in urban areas are shown to statistically have the same mean total number of trips. The other two groups with adults living in rural and suburban areas still have t statistics outside the $\pm t_{critical}$ range. However, the calculated t statistics (rural = 2.46 and suburban = 2.97) are smaller than the overall person cluster t statistic for this measure. This indicates that the difference between the person-land use groups in the two surveys have less variability in the trip totals than the person-only clusters. This result is repeated for other travel characteristics, including the number of social and retail trips; retail, social, and total travel times; and retail, social, and work activity durations. Strikingly interesting is the pattern of the residents of urban environments that show remarkably similar average activity and travel characteristics with one exception.

In a few instances the person cluster averages mask greater variability that is only revealed when these are subdivided by their land use cluster categories. For example, the total number of work trips for all persons in this cluster has a t statistic ($t = -0.53$) within the $t_{critical}$ range ($-1.96 < t < 1.96$). However, the average number of work trips for survey participants living in poor, urban areas have a $t = 2.08$, which is larger than $t_{critical}$. So, the average number of work trips taken by this subset of people is different for the TBI and NHTS surveys. Because each survey only contains a small subset of people in this person-land use group (67 out of 1492 in the NHTS and 212 out of 2270 in the TBI), this difference is masked in the LCCA person cluster t calculation.

TABLE 5 NHTS and TBI survey weekday average activity and travel statistics

NHTS Average Travel Values												
Land use	**Activity Duration**				**Travel time**				**Trips**			
	Retail	Social	Work	Total	Retail	Social	Work	Total	Retail	Social	Work	Total
Rural	29	46	299	459	21	12	38	107	0.99	0.55	1.17	4.61
Suburban	29	69	307	498	10	22	24	99	0.94	0.85	1.11	5.02
Urban	22	50	341	506	9	14	27	88	0.72	0.69	1.07	4.43
Urban – poor	22	41	315	465	14	13	20	92	0.97	0.66	0.87	4.61
TBI Average Travel Values												
Land use	**Activity Duration**				**Travel time**				**Trips**			
	Retail	Social	Work	Total	Retail	Social	Work	Total	Retail	Social	Work	Total
Rural	13	38	353	485	7	8	24	81	0.46	0.44	1.05	3.96
Suburban	12	44	366	494	6	9	29	84	0.45	0.48	1.15	4.07
Urban	16	48	345	517	8	10	24	85	0.61	0.61	1.19	4.88
Urban – poor	16	45	355	509	14	11	24	88	0.61	0.56	1.09	4.28
T-Statistics comparing NHTS to TBI												
Land use	**Activity Duration**				**Travel time**				**Trips**			
	Retail	Social	Work	Total	Retail	Social	Work	Total	Retail	Social	Work	Total
All	5.05 (1.96)	2.97 (1.96)	-2.99 (1.96)	-0.44 (1.96)	3.99 (1.96)	5.28 (1.96)	-0.01 (1.96)	2.37 (1.96)	7.83 (1.96)	5.91 (1.96)	-0.53 (1.96)	3.04 (1.96)
Rural	2.94 (1.98)	1.02 (1.98)	-2.27 (1.98)	-1.11 (1.98)	3.1 (1.98)	1.41 (1.98)	1.7 (1.98)	2.44 (1.98)	4.27 (1.98)	1.5 (1.98)	1.03 (1.98)	2.46 (1.98)
Suburban	2.73 (1.98)	2.09 (1.98)	-2.02 (1.98)	0.14 (1.98)	2.69 (1.98)	1.7 (1.98)	-1.05 (1.98)	1.32 (1.98)	3.8 (1.98)	3.62 (1.98)	-0.31 (1.98)	2.97 (1.98)
Urban	1.41 (1.98)	0.22 (1.98)	-0.15 (1.98)	-0.44 (1.98)	0.94 (1.98)	1.46 (1.98)	0.57 (1.98)	0.29 (1.98)	1.01 (1.98)	0.88 (1.98)	-1.09 (1.98)	-1.53 (1.98)
Urban – poor	1.04 (1.99)	-0.38 (1.98)	-1.18 (1.98)	-1.28 (1.98)	-0.15 (1.98)	0.54 (1.98)	-1.17 (1.98)	0.49 (1.98)	1.89 (1.99)	0.73 (1.99)	-2.08 (1.98)	0.8 (1.99)

The average travel statistics for the NHTS and SCAG survey adults in this cluster are located in Table 6. The average numbers of retail and social trips completed by NHTS cluster members are twice as large as their SCAG survey counterparts. NHTS average trips for rural, suburban, and urban areas range from 0.72 to 0.07 trips compared to 0.43 to 0.54 for SCAG survey participants. This is also reflected in higher travel times and activity durations. Work trip statistics are more similar.

The t statistics in Table 6 show that only the average work travel times for the person clusters are the same in the two surveys. Subdividing the person cluster into land use groups significantly increases the ability to match average travel characteristics for those people living in the suburban and both types of urban areas. These three person-land use categories have matching average values for retail, social, retail and total travel times. They also match most average activity durations, except for the average retail and social activity durations for suburbanites and total durations for urbanites. For average trip measures, only the average number of work trips for cluster members living in suburban and urban areas is the same. Although many of the person-land use groups do not have t values within the $\pm t_{critical}$ range, these are all smaller than the person-only t statistics. Table 6 shows the differences between the two surveys for these groups are heavily due to the rural population. The Midwest sample of the NHTS survey participants not only travel more than the Southern Californians but also participate in activities for longer periods in a day. Urbanites, however, appear to be similar in the NHTS sample in their time allocation but not their number of trips to retail and social purposes. This comparison also shows there are more differences between NHTS and SCAG than between NHTS and TBI, as expected. However, many cells show statistically no difference in their average activity-travel behavior.

TABLE 6 NHTS and SCAG survey weekday average activity and travel statistics

NHTS Average Travel Values												
Land use	**Activity Duration**				**Travel time**				**Trips**			
	Retail	Social	Work	Total	Retail	Social	Work	Total	Retail	Social	Work	Total
Rural	25	51	329	487	17	19	30	104	0.85	0.64	1.12	4.46
Suburban	29	69	307	498	10	22	24	99	0.94	0.85	1.11	5.02
Urban	22	50	341	506	9	14	27	88	0.72	0.69	1.07	4.43
Urban – poor	22	41	315	465	14	13	20	92	0.97	0.66	0.87	4.61
SCAG Average Travel Values												
Land use	**Activity Duration**				**Travel time**				**Trips**			
	Retail	Social	Work	Total	Retail	Social	Work	Total	Retail	Social	Work	Total
Rural	16	21	290	391	8	8	25	81	0.52	0.34	0.86	3.33
Suburban	15	38	324	451	8	10	31	85	0.54	0.5	1.05	3.81
Urban	15	37	297	436	7	11	26	86	0.52	0.48	0.94	3.71
Urban – poor	15	23	312	441	9	7	26	83	0.43	0.33	0.84	3.2
t-Statistics comparing NHTS to SCAG survey												
Land use	**Activity Duration**				**Travel time**				**Trips**			
	Retail	Social	Work	Total	Retail	Social	Work	Total	Retail	Social	Work	Total
All	**5.07** (1.96)	**5.88** (1.96)	**2.4** (1.96)	**8.21** (1.96)	**4.99** (1.96)	**4.63** (1.96)	**-0.25** (1.96)	**3.28** (1.96)	**9.67** (1.96)	**9.09** (1.96)	**4.23** (1.96)	**12.58** (1.96)
Rural	**2.37** (1.97)	**5.01** (1.97)	**2.25** (1.97)	**5.7** (1.97)	**3.14** (1.98)	**3** (1.97)	**1.17** (1.97)	**2.86** (1.97)	**4.01** (1.97)	**5.56** (1.97)	**3.41** (1.97)	**6.29** (1.97)
Suburban	**2.36** (1.98)	**3.02** (1.98)	**-0.65** (1.98)	**1.83** (1.98)	**1.48** (1.98)	**1.62** (1.98)	**-1.49** (1.98)	**1.35** (1.98)	**3.27** (1.98)	**3.65** (1.98)	**0.48** (1.98)	**4.15** (1.98)
Urban	**1.69** (1.98)	**1.41** (1.98)	**1.96** (1.98)	**3.12** (1.98)	**1.37** (1.98)	**1.11** (1.98)	**0.2** (1.98)	**0.16** (1.98)	**2.03** (1.98)	**2.63** (1.98)	**1.43** (1.98)	**2.75** (1.98)
Urban – poor	**1.18** (1.99)	**1.69** (1.99)	**0.1** (1.99)	**0.72** (1.99)	**1.78** (1.98)	**1.98** (1.99)	**-1.42** (1.98)	**1.06** (1.98)	**2.89** (1.99)	**2.66** (1.99)	**0.24** (1.98)	**3.58** (1.99)

Tables 5 and 6 give the impression of feasible transferability for specific groups of household that live in urbanized environments. They even give the impression that data can be transferred from the Midwest to California. To verify this, the distribution of behaviors is examined for the travel clusters discussed earlier. Table 7 shows the membership percentage of adults in travel clusters. The clusters have been selected to focus on retail, social, and work trips, as well as including persons who did not travel. The NHTS contains a higher proportion of people who make retail and social trips. For example, 13.1% of NHTS cluster members complete retail and social trips on their travel day compared to only 5.3% of TBI and 8.7% SCAG survey participants. This same pattern is repeated for persons making retail, social and work trips (NHTS = 35.5%, TBI = 13.5%, and SCAG = 22.3%). SCAG survey members tended to have more work-only schedules and TBI persons had more people completing other types of trip combinations (represented by the remaining travel clusters combined together). Finally, the local surveys (particularly the SCAG survey) had persons who did not travel.

Recall from Table 5 that these urban dwellers have statistically no difference between NHTS and TBI in their average activity durations, daily time allocated to travel, and number of trips by purpose. Similarly, from Table 6 the comparison between NHTS and SCAG shows statistically the same average duration of activities but fewer trips for the SCAG urban residents. The travel clusters in Table 7 show the three surveys display different activity and travel behavior distributions even when their average behavioral characteristics are statistically the same for specific groups of people. Even when survey participants are cross-classified by important determinants of activity participation and travel, heterogeneity in behavior is masked. This leads to the incorrect conclusion that the transferability of average behavior is feasible. Average behavior is a blend of different activity and travel clusters in each survey and this undermines the possibility of transferring activity and travel scheduling from one region to another.

TABLE 7 NHTS, TBI, and SCAG survey land use and travel cluster distributions

Land use	Survey	Travel Clusters				
		Retail & Social	Retail, Social & Work	Work	No Trips	Other Clusters
All	NHTS	13.1%	35.5%	21.4%	4.4%	25.7%
	TBI	5.3%	13.5%	24.3%	7.4%	49.5%
	SCAG	8.7%	22.3%	28.5%	12.5%	28.0%
Rural	NHTS	14.5%	35.4%	22.5%	4.7%	23.0%
	TBI	5.4%	12.2%	26.4%	7.7%	48.3%
	SCAG	10.9%	15.7%	28.8%	15.3%	29.4%
Suburban	NHTS	15.7%	35.2%	15.7%	4.6%	28.7%
	TBI	4.7%	15.1%	27.5%	8.7%	44.0%
	SCAG	7.4%	24.1%	28.2%	11.2%	29.1%
Urban	NHTS	10.4%	37.3%	22.4%	6.0%	23.9%
	TBI	4.9%	12.7%	20.0%	7.6%	54.9%
	SCAG	9.9%	22.6%	25.4%	12.8%	29.4%
Poor Urban	NHTS	17.9%	29.9%	22.4%	3.0%	26.9%
	TBI	8.0%	18.4%	21.2%	5.7%	46.7%
	SCAG	6.1%	19.2%	34.1%	14.1%	26.5%

The actual number of survey participants and estimated population for these groups were compared using a χ^2 analysis. The results in Table 8 demonstrate that both the groups of urbanites (NHTS-TBI and NHTS-SCAG) have relatively similar numbers of survey participants and estimated population with social and retail trip patterns and work, social, and retail trip patterns. While this result reinforces the NHTS-TBI average travel behavior comparison in Table 5, the NHTS-SCAG average numbers of retail and social trips for urban dwellers in Table 6 are different. They are not represented in this trip cluster. Both tables do show that persons in rural areas from both sets of surveys had different average activity-travel behavior. The χ^2 distributions in Table 8 reflect these differences, except for those persons in the work-only NHTS-TBI travel cluster. In this case, both sets of survey participants and estimated population contain a smaller proportion of persons who make only work trips. These adults also statistically make the same average number of work trips.

TABLE 8 Scaled χ^2 distributions for weekday survey pairs based on trip and land use category combinations

Trip Type	Land use Type	NHTS-TBI χ^2		NHTS-SCAG χ^2	
		Survey Participants	Estimated Population	Survey Participants	Estimated Population
Social and Retail	Rural	18.3	13.2	21.3	14.6
	Suburban	2.6	2.9	0.1	2.3
	Urban	0.2	0.8	4.1	3.8
	Poor Urban	0.0	0.0	0.9	0.1
	Other	46.4	39.2	8.2	4.5
Work, Social and Retail	Rural	50.5	29.7	78.0	43.9
	Suburban	1.2	0.5	5.1	0.0
	Urban	5.1	12.7	1.2	0.5
	Poor Urban	0.8	0.5	0.4	0.6
	Other	155.0	116.1	32.0	16.3
No Trips	Rural	3.0	0.4	0.6	0.7
	Suburban	6.9	5.6	16.0	3.2
	Urban	4.3	2.4	16.8	17.0
	Poor Urban	3.6	1.2	12.5	11.3
	Other	0.4	1.3	22.7	26.6
Work	Rural	1.3	0.7	14.6	15.3
	Suburban	20.1	17.2	32.9	5.2
	Urban	4.8	1.6	15.5	14.6
	Poor Urban	5.1	6.0	13.3	16.3
	Other	6.6	6.0	2.0	2.3
Other Trip Types	Rural	36.8	21.4	15.1	17.0
	Suburban	27.3	33.2	16.9	2.3
	Urban	64.5	57.2	20.5	27.5
	Poor Urban	27.8	22.2	4.5	4.9
	Other	6.6	4.7	1.6	0.9
Total χ^2		499.1	396.6	356.8	251.6

Similar results are obtained for other sociodemographic groups such as adults in households with children and adults in "older" (e.g., retired) households.

COMPARISON TO PREVIOUS TRANSFERABILITY RESULTS

Mohammadian et al. (20) have completed some of the most extensive studies in transferring the NHTS to local populations. Their approach (based on work in 16-19) uses Bayesian updating to influence the distributions of selected variables. The methodology utilizes a small sample of survey households to calculate the mean and standard deviations of modeled variables (e.g., trip rates by type). The updated distributions are used to generate a new set of travel data using Monte Carlo simulation. Their studies use national portions of the NHTS, updated using an NHTS add-on. Eleven clusters were generated using a combination of household characteristics, land use attributes, and travel behavior. The averages for twenty different variables generated using the updated distributions were compared to NHTS averages using z tests, similar to the t statistic. Out of twenty variables, nine attributes were deemed to be transferable, because their means were similar. These variables include number of auto trips, trip length, commute distance, number of trips per tour, number of tours, number of mandatory trips, number of maintenance trips, and number of discretionary trips. Similar to Mohammadian et al., this study's methodology of clustering persons, activity and travel information, and land use separately was able to match average activity-travel measures for person-land use groups. The actual distributions of survey persons, travel patterns, and residential land use and accessibility measures were not the same between either of the two local surveys and the NHTS. It is unclear if the Mohammadian et al.'s simulated data suffers from the same problem, although their approach does have the added benefit of applying local data to ensure the average distributions will be more similar.

IS TRANSFERABILITY POSSIBLE?

The results of this research show that in fact people with different geographic residence locations do travel differently, even if they share the same sociodemographic characteristics. However, the inclusion of land use characteristics does improve the ability to match average activity-travel behavior characteristics. But, in order to transfer the data from one region to another, the distributions must be similar between the local and national surveys. The primary reason that distributions between the NHTS and a local survey must be analogous is that the travel patterns from the national survey will be used to generate travel schedules for a local synthetic population. Activity-based transportation models rely on the range of activity-travel characteristics to represent behaviorally appropriate travel across a street network. Because the entire distribution of travel behavior is included in these models (and not just averages), activity-based transportation models are suited to answer policy questions, typically not adequately addressed or able to be represented by a typical four-step transportation model. Transferring average travel measures or incorrectly applying survey patterns to a new area could lead to erroneous conclusions. It should be noted the comparison presented in this paper is a sample of a larger analysis that includes a more detailed description of the data assembly process, cluster/group identification, comparisons among additional social and demographic groups as well as the detailed comparison for both weekday and weekend activity and travel patterns (30).

ACKNOWLEDGEMENTS

This research was supported by the University of California Lab Fees Research program, the National Highway Institute of the Federal Highway Administration's Dwight David Eisenhower Graduate Transportation Fellowship, the University of California Transportation Center's UCTC Doctoral Dissertation Research Grant, and the National Science Foundation's Doctoral Dissertation Research Grant in Geography and Regional Science.

REFERENCES

1. Oppenheim, N. (1975). A typological approach to individual urban travel behavior prediction. *Environment and Planning A*, **7**, pp. 141-152.
2. Hanson, S. and P. Hanson (1981). The Travel-Activity Patterns of Urban Residents: Dimensions and Relationships to Sociodemographic Characteristics. *Economic Geography*, **57**(4), pp. 332-347.
3. Hanson, S. and J. Huff (1986). Classification Issues in the Analysis of Complex Travel Behavior. *Transportation*, **13**, pp. 271-293.
4. Bowman, J. and M. Ben-Akiva (1996). Activity-Based Travel Forecasting, *Activity-Based Travel Forecasting Conference Proceedings*. June 2-5, 1996, http://tmip.fhwa.dot.gov/clearinghouse/docs/abtf/, accessed September 25, 2006.
5. Stopher, P. R. and H. M. A. Metcalf (1999). Household Activities, Life Cycle, and Role Allocation Tests on Data Sets from Boston and Salt Lake City. *Transportation Research Record*, No.1676, pp. 95-102.
6. Pribyl, O. and K. G. Goulias (2005). Simulation of Daily Activity Patterns Incorporating Interactions Within Households: Algorithm Overview and Performance. *Transportation Research Record*, No.1926, pp. 135-141.
7. Reuscher, T. R., R. L. Schmoyer, Jr., and P. S. Hu (2002). Transferability of Nationwide Personal Transportation Survey Data to Regional and Local Scales. *Transportation Research Record*, No.1817, pp. 25-32.
8. Atherton, T. J. and M. E. Ben-Akiva (1976). Transferability and Updating of Disaggregate Travel Demand Models. *Transportation Research Record*, No.610, pp. 12-18.
9. Elmi, A. M., D. A. Badoe, and E. J. Miller (1999). Transferability Analysis of Work-Trip-Distribution Models. *Transportation Research Record*, No.1676, pp. 169-176.
10. Wilmot, C. G. and P. R. Stopher (2001). Transferability of Transportation Planning Data. *Transportation Research Record*, No.1768, pp. 36-43.
11. Valenin, C. A., J. N. Prashker, and Y. Shiftan (2003). Analysis of Trip Generation Characteristics in Israel for the Years 1984, 1996/7 and Spatial & Temporal Transferability of Trip Generation Demand Models. *2003 Transportation Research Board Annual Meeting CD-ROM*, Washington D.C., January 2003.
12. Horner, J., J. R. Stone, and L. Huntsinger (2006). The Impact on Travel Behavior of Proximity to Major Urban Centers. *2006 Transportation Research Board Annual Meeting CD-ROM*, Washington D.C., January 22-26, 2006.
13. Karasmaa, N. (2001). The Spatial Transferability of the Helsinki Metropolitan Area Mode Choice Models. Presented at the 5th Workshop of the Nordic Research Network on Modelling Transport, Land-Use and the Environment, Nynäshamn, September 28-30, 2001.

14. Karasmaa, N. and M. Pursula (1997). Emperical Studies of Transferability of Helsinki Metropolitan Area Travel Forecasting Models. *Transportation Research Record*, No.1607, pp. 38-44.
15. Transportation Research Board (TRB) (2010). *Travel Demand Forecasting: Parameters and Techniques*. http://rip.trb.org/browse/dproject.asp?n=11814/, accessed July 25, 2010.
16. Greaves, S. P. and P. R. Stopher (2000). Creating a Synthetic Household Travel and Activity Survey: Rational and Feasibility Analysis. *Transportation Research Record*, No.1706, pp. 82-91.
17. Stopher, P. R. (2003). Simulating Household Travel Survey Data in Australia: Adelaide Case Study. *Road and Transport Research*, **12**(3), pp. 29-44.
18. Stopher, P. R., S. P. Greaves, and M. Xu (2004). Using Nationwide Household Travel Data for Simulating Metropolitan Area Household Travel Data. Presented at the National Household Travel Survey Conference: Understanding Our Nation's Travel Conference, Washington, D.C., November 1-2, 2004.
19. Pointer, G., P. R. Stopher, and P. Bullock (2004). Monte Carlo Simulation of Household Travel Survey Data for Sydney, Australia: Bayesian Updating Using Different Local Sample Sizes. *Transportation Research Record*, No.1870, pp. 102-108.
20. Mohammadian, A., M. Javanmardi, and Y. Zhang (2010). Synthetic household travel survey data simulation. *Transportation Research Part C*. **18**, pp. 869-878.
21. Lin, J., L. Long (2008). What Neighborhood Are You In? Empirical Findings of Relationships between Household Travel and Neighborhood Characteristics. *Transportation*, **35**(6), pp. 739-758.
22. Lin, J., L. Long (2008b). Model-based Approach to Synthesize Household Travel Characteristics across Neighborhood Types and Geographic Areas. *The ASCE Journal of Transportation Engineering*, **134**(12), pp. 493-503.
23. Federal Highway Administration (FHWA) (2001). National Household Travel Survey (NHTS) Users Guide. http://nhts.ornl.gov/2001/usersguide/index.shtml, accessed August 3, 2006.
24. Metropolitan Council (2003). 2000 Travel Behavior Inventory Home Interview Survey: Data and Methodology. Report No. 35-03-032, http://www.metrocouncil.org/planning/transportation/TBI_2000/TBI_Methodology.pdf, accessed June 2, 2010.
25. NuStats (2003). Year 2000 Post-Census Regional Travel Survey: Final Report of Survey Results. http://www.scag.ca.gov/travelsurvey/pdf/MainSurveyResults.pdf, accessed July 21, 2010.
26. Magidson, J. and J. K. Vermunt (2002). A Nontechnical Introduction to Latent Class Models. Statistical Innovations White Paper #1. http://www.statisticalinnovations.com/articles/lcmodels2.pdf, accessed July 21, 2010.
27. Guo, J. Y. and C. R. Bhat (2004). Modifiable Areal Units: Problem or Perception in Modeling of Residential Location Choice? *Transportation Research Record*, No.1898, pp. 138-147.
28. Sampson, R. J., S. W. Raudenbush, and F. Earls (1997). Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy. *Science*, **277**, pp. 918-924.
29. GeoGraphics Laboratory (1999). FTA Bus Route GIS Database. http://www.geolab.bridgew.edu/ftpserver/, accessed December 7, 2006.

30. Henson K. (2010) Travel Determinants and Multi-scale Transferability of National Activity Patterns to Local Populations, Unpubblished Ph.D. dissertation, Department of Geography, University of California Santa Barbara, Santa Barbara, CA.