TECHNICAL REPORT STANDARD PAGE

1. Report No. FHWA/LA.03/356	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Simulating Household Travel Survey Data in Metropolitan Areas	5. Report Date: September 2003 6. Performing Organization Code	
7. Author(s) Peter R. Stopher Stephen P. Greaves	8. Performing Organization Report No.	
9. Performing Organization Name and Address: Department of Civil and Environmental Engineering Louisiana State University Baton Rouge, LA 70803-6405	10. Work Unit No. 11. Contract or Grant No. State Project No. 737-99-0462	
12. Sponsoring Agency Name and Address Louisiana Transportation Research Center 4101 Gourrier Avenue Baton Rouge, LA 70808	13. Type of Report and Period Covered: Final Report, 11/01/98 - 4/30/00 14. Sponsoring Agency Code	
15. Supplementary Notes		

16. Abstract

Census data provide a rich range of socioeconomic characteristics from which it is shown that trip characteristics can be simulated. This report summarizes research into the simulation of the trips and trip characteristics for a random sample of households drawn from census data. The simulation source is the 1990 PUMS data from the 1990 Decennial Census of the United States.

A set of categories is defined for the simulation that allows the development of significantly different statistical distributions of trip characteristics, using the 1995 NPTS data. Based on the census data, samples of households are obtained and their trip characteristics in terms of number of trips by purpose, mode, time of departure, and trip length are simulated, using a Monte Carlo type of simulation procedure. This is performed for three regions: Baton Rouge, Louisiana, Dallas-Fort Worth, Texas, and Salt Lake City, Utah. While there are found to be a number of statistically significant differences in the various trip characteristics between the simulation data and actual household travel surveys conducted in 1997 in Baton Rouge, 1996 in Dallas-Fort Worth, and 1993 in Salt Lake City, the numeric differences in many of the characteristics are actually quite small. It is found that the simulation, as currently defined, does not capture trip-length variations that may be attributable to city size, nor does it do as well as might be hoped in capturing effects resulting from differences in household size between cities such as Dallas and Salt Lake. Further refinement of the simulation procedure appears to be warranted.

In the case of Baton Rouge, comparisons are made on the trip rates by purpose with the existing trip generation models (which were borrowed in 1991 for the Baton Rouge area), with national default figures, and with new trip-generation models developed from the 1997 data. The simulation was found to perform much better than the borrowed trip-generation models and the national default figures. In comparison with new trip-generation models, the simulation was found to perform quite well, although the poorest results were obtained with home-based shopping trips.

It is concluded that simulation is a feasible procedure for creating synthetic household travel survey data, using the procedure outlined in this report. A number of new avenues for research are identified, which should enhance the results further.

17. Keywords Household/Personal Travel Surveys, Monte Carlo Simulation, PUMS, NPTS, Modeling		18. Distribution Statement Unrestricted. This document is available through NTIS, Springfield, VA 21161	
19. Security Classif. (of this report) Not Applicable	20. Security Classif. (of this page) Not Applicable	21. No. of Pages:	22. Price

SIMULATING HOUSEHOLD TRAVEL SURVEY DATA IN METROPOLITAN AREAS

by

Peter R. Stopher, Ph.D.

Department of Civil and Environmental Engineering
Louisiana State University
Baton Rouge, Louisiana 70803-6405

and

Stephen P. Greaves

Ph.D. Candidate
Department of Civil and Environmental Engineering
Louisiana State University
Baton Rouge, Louisiana 70803-6405

LTRC PROJECT NO. 99-2SS
STATE PROJECT NO. 737-99-0462

conducted for

LOUISIANA DEPARTMENT OF TRANSPORTATION AND DEVELOPMENT
LOUISIANA TRANSPORTATION RESEARCH CENTER

September 2003

ABSTRACT

Census data provide a rich range of socioeconomic characteristics, from which it is shown that trip characteristics can be simulated. This report summarizes research into the simulation of the trips and trip characteristics for a random sample of households drawn from census data. The simulation source is the 1990 PUMS data from the 1990 Decennial Census of the United States.

A set of categories is defined for the simulation that allows the development of significantly different statistical distributions of trip characteristics, using the 1995 NPTS data. Based on the census data, samples of households are obtained and their trip characteristics in terms of number of trips by purpose, mode, time of departure, and trip length are simulated, using a Monte Carlo type of simulation procedure. This is performed for three regions: Baton Rouge, Louisiana, Dallas-Fort Worth, Texas, and Salt Lake City, Utah. While there are found to be a number of statistically significant differences in the various trip characteristics between the simulation data and actual household travel surveys conducted in 1997 in Baton Rouge, 1996 in Dallas-Fort Worth, and 1993 in Salt Lake City, the numeric differences in many of the characteristics are actually quite small. It is found that the simulation, as currently defined, does not capture trip-length variations that may be attributable to city size, nor does it do as well as might be hoped in capturing effects resulting from differences in household size between cities such as Dallas and Salt Lake. Further refinement of the simulation procedure appears to be warranted.

In the case of Baton Rouge, comparisons are made on the trip rates by purpose with the existing trip generation models (which were borrowed in 1991 for the Baton Rouge area), with national default figures, and with new trip-generation models developed from the 1997 data. The simulation was found to perform much better than the borrowed trip-generation models the national default figures. In comparison with new trip-generation models, the simulation was found to perform quite well, although the poorest results were obtained with home-based shopping trips.

It is concluded that simulation is a feasible procedure for creating synthetic household travel survey data, using the procedure outlined in this report. A number of new avenues for research are identified, which should enhance the results further.

ACKNOWLEDGMENTS

The research reported in this document was conducted with funding from the Louisiana Transportation Research Center under the guidance of Mr. Harold R. Paul and Mr. Arthur P. Rogers of LTRC. Funding was provided for LTRC by the Federal Highway Administration Travel Model Improvement Program. The authors wish to acknowledge the help and support of Dr. Fred Ducca of FHWA in making this research funding possible.

The authors would also like to thank graduate students Sirisha Kothuri, Kanthi Kumar Adapi, and Anurag Dwivedi for their help in preparing data, conducting analyses, and applying procedures used in this research. Thanks are also given to staff of the Lafayette Metropolitan Planning Organization for their help to Stephen Greaves in the conduct of this research. Finally thanks are given to Philip Bullock of the Institute of Transport Studies, The University of Sydney, for help in finalizing the analyses reported herein.

IMPLEMENTATION STATEMENT

The results of this research offer an alternative low-cost method for a Metropolitan Planning Organization to develop data that could be used for model-building purposes and for model testing. While most larger MPOs and many medium-sized MPOs will most likely still desire to collect a full household/personal travel survey from time to time, the procedures developed herein provide opportunities for many of the medium and smaller MPOs to create a personal travel survey data set and to update it with a very small sample of households. This offers a potential savings of hundreds of thousands of dollars for many MPOs.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION / RATIONALE

Problems with Conducting Household Travel Surveys

Household travel data are used to estimate and update travel-demand models designed to analyze proposed transportation policy decisions as well as to provide information on regional travel characteristics. Typically, the data are obtained from a household-based survey in which a small sample of the population records their sociodemographic data and travel patterns over a given time period. These data are used to develop travel-demand models based on relationships between individual/household characteristics and observed travel patterns.

Household travel surveys (HTSs) have always been a problematic, high-cost activity for metropolitan planning organizations (MPOs). Surveying 1,300 to 2,000 households, which according to Cambridge Systematics, Inc. (1996), is the minimum number of households for satisfactory travel-demand model estimation in small MPOs, could cost between $156,000 and $350,000. This represents a substantial portion or even exceeds the annual planning budget of an MPO. Consequently, many MPOs have not conducted local surveys and it is unlikely they will in the future. An additional concern for all MPOs is the increasing difficulty of conducting high quality HTSs. The problems stem from the reluctance of people to participate in surveys, the general move away from face-to-face surveys, the increasing unreliability of using the telephone as a household recruiting/retrieval device, and the difficulty of catching people at home. While methodological and technological survey techniques continue to become increasingly refined, high unit costs and public resistance will plague future survey efforts. The implications are that *all* metropolitan areas, regardless of size and financial resources, may face problems of unavailable or inadequate travel data in the future.

Development of Synthetic Household Travel Survey Data

The U.S. Bureau of the Census has collected a wealth of social and demographic data as part of the census for past decades. Due to their policy of confidentiality, the Bureau releases only summary statistics for the more disaggregate levels of census geography such as the census tract and the block group[1]. However, fully disaggregate data are provided at the level of the Public Use Micro-data Area (PUMA)[2] for either a one or a five percent sampling of households. PUMAs are large areas, containing approximately 100,000 people. The data available at this level, known as the Public Use Micro-data Sample (PUMS), are provided with complete disaggregation of the characteristics of households but with geographic specificity only at the level of the

---

1 For example, the Standard Tape Files released for public use by the U.S. Bureau of the Census and referenced as STF-1, STF-2, STF-3, etc.

2 Census of Population and Housing, 1990: Public Use Microdata Sample U.S. Technical Documentation. Prepared by the Bureau of the Census, Washington D.C.

PUMA, County, or State. Within the data available in PUMS are such variables as: household size, vehicles available, working status of household members, ages of household members, genders and relationships of household members, education levels of household members, race and ethnicity of household members, marital status of household members, and household income.

A typical household travel survey collects similar demographic information plus the travel and activities performed by the household most often for a 24-hour weekday period. Probably the only variable that transportation planners have traditionally used that is not in the PUMS is driver's license status. The PUMS provides very rich household data; each PUMA provides in the 5 percent sample about 5,000 households. The second resource on which this concept is based is the existence of a national transportation data set that collects similar travel and activity data to that of a typical HTS such as the 1995 Nationwide Personal Transportation Survey (NPTS).

Travel-demand models assume relationships between sociodemographic characteristics and travel characteristics. It seems plausible, therefore, that some parsimonious set of household and person characteristics could be chosen and, from data such as the NPTS, distributions of characteristics of travel could be developed that would differ based on the values of those demographic variables. This would provide a basis for simulating travel patterns. Specifically, if characteristics such as the number of trips by each trip purpose, the trip length, start time, and mode of travel were to be simulated from the distributions appropriate to a particular household, simulated data on trips for that household could be produced. For this to work, the characteristics of travel must be related to sociodemographic groupings, and actual households must be able to be sampled and described by the sociodemographic variables of interest. Using PUMS or any reliable source of local socio-demographic data, a sample of households can be drawn from an urban area for which a fairly exhaustive set of characteristics can be provided. With those characteristics, random drawings can be made from the distributions for each travel characteristic, thus defining a set of daily household trips with trip purpose, length, start time, and mode. The only characteristics that are undefined in this process are the specific geographic locations of the households and the locations of the trip ends or activities of the household.

OBJECTIVES

The objectives of this study are:

1. To develop and validate a method for MPOs to synthesize household travel survey data using local sociodemographic characteristics of Baton Rouge, Louisiana, in conjunction with a national source of simulated travel data.

2. To evaluate whether travel-demand models estimated with these synthetic data as applied to Baton Rouge, perform better than models or statistics developed in other contexts.

3. To evaluate the transferability of the approach. Here, the procedures will be run for two other urban areas with different socio-demographic characteristics from Baton Rouge.

SCOPE

The research reported here was aimed at developing a procedure for synthetically generating data that would typically be derived from a household travel survey. This includes trip rates by purpose, then, for each trip, the mode of travel, departure hour, and travel time. The only characteristics undefined in this process are the specific geographic locations of households and the activity/trip ends. The procedure was to be developed and tested in Baton Rouge to satisfy the first two objectives, and then extended to two other geographic areas apart from Louisiana and, preferably, representing markedly different urbanized areas.

METHODOLOGY

Setting Up the Simulation

The Classification and Regression Tree (C&RT)[3] method was used in the initial delineation of schemes to categorize the NPTS sample into homogeneous groups with respect to the travel attributes of interest. C&RT is a computationally intensive exploratory classification tool developed in the early 1960s, which has only recently become a realistic option for practical research. C&RT involves a binary recursive partitioning of the data with respect to a dependent variable. At each node, all predictor variables are evaluated to determine the best groupings based on the improve-ment score (reduction in the residual sum of squares). The independent variable with the largest score is selected for the split and this process continues until certain user-defined stopping criteria are met. For this research, the 30,400 NPTS households were partitioned into two equal parts, the models estimated, and the results were compared. C&RT is only an *exploratory* tool that identifies groupings and interactions of which the researcher may not be aware. Consequently, the analyst still must make informed subjective decisions about final groupings. Standard Analysis of Variance (ANOVA) procedures were used to assist in the validation of the final schemes.

Developing the Distributions

After establishing the categories, the second phase of the simulation procedure was to develop frequency distributions to use as samples for the travel characteristics of interest. This procedure is best illustrated with the following example. Consider the home-shop case where households have been classified into ten categories from the C&RT procedure. For each occurrence of a category in the NPTS database, the magnitude of home-work trips is recorded. The result is that ten discrete frequency distributions can be created.

These distributions can then be reconstructed as cumulative frequency distributions based on an arbitrarily large number of observations, e.g. 100,000. A random number generator[4] is then used to generate a value that falls within a particular probability range and hence indicates the appropriate number of trips assigned to that household. This is the same principle underlying Monte Carlo sampling procedures except that, in this case, empirically-derived probability function is used as the basis for sampling. The process is repeated for each travel characteristic of interest so that, in effect, a full travel survey data set is created for each sampled household.

---

[3] Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984) Classification and Regression Trees. Wadsworth International Group, Belmont, California.

[4] Rand Corporation. (1955) A Million Random Digits with 100,000 Normal Deviates. Available from http://www.rand.org/publications/classics/randomdigits/randomdata.html. Accessed June 12th, 1999.

Simulating A Synthetic Survey For Baton Rouge

If the concept of a synthetic household travel survey is valid, one should theoretically be able to reproduce (within an acceptable error range) the data collected by and build similar models to a real survey. To establish whether this is the case, the simulation procedure was tested using the Baton Rouge Personal Transportation Survey (BRPTS). The BRPTS was conducted from April-June of 1997 for the Baton Rouge MPO region as an additional sample to the NPTS. It used virtually identical methods and collected the same data plus fully geocoded location data. In all, 1,395 households were sampled of which 984 households provided weekday travel (approximately one-thirtieth of the NPTS sample used in this research).

Transferability of the Approach

To demonstrate the application of the simulation to other locations, the methodology was designed to be tested on two other metropolitan areas: the Dallas-Fort Worth Metroplex and the combined Wasatch Front and Mountainland MPO areas in Utah, covering Salt Lake City, Ogden, and Provo/Orem. These two locations were chosen for comparison to represent areas of differing population size and areas in which differing survey instruments were used in the most recent household travel survey. Some useful comparative statistics for the two locations are provided in Table 1, which shows that while the percent of females in the sample, the average vehicle ownership levels, and the average age of respondents appear very similar, the other statistics exhibit some significant differences between the Dallas-Fort Worth and Utah areas. As a result, the selection of these two regions should offer a good test of the simulation procedure.

Table 1: Summary Statistics from the Household Travel Surveys

Statistic	Wasatch Front/Mountainland	Dallas/Fort Worth
Average Age	34.3	34.8
Average Household Size	3.14	2.47
Percent in Single Family Dwellings	72.7%	78.1%
Percent from Non-Car-Owning Households	4.4%	5.2%
Average Vehicles per Household	1.97	1.84
Percent Females in Sample	52.5%	51.8%
Percent Home Owners	76.2%	67.5%

ANALYSIS

The Categorization Results

Person Trip Frequencies

Table 2 shows the final categorization schemes for each trip purpose together with the GLM results. Four schemes were used: (1) Home-Work and Work-Other, (2) Home-School, (3) Home-College, and (4) Home-Shop, Home-Other, and Other-Other.

Work trips were driven primarily by the number of household workers. While this is not surprising, it is interesting to observe the interactions between the number of workers, household vehicles, and the presence and age of children. For one-worker households, the critical issue is the availability of vehicles. For households with zero or one vehicle, the trip rates are significantly less than in households with two or more vehicles where the worker is likely to have his or her own vehicle. For two-worker households there is a suppression of trips caused by the presence of children, particularly those under age five. A logical conclusion would be that this is a result of workers dropping off children at daycare or school on their way to/from work. Although this involves a work trip, the data would record it as a home-other followed by a work-other trip (and vice versa for the trip home). Therefore, work-other trips should be higher for workers with children, which is generally the case.

School trips are almost totally dependent on the presence and number of children. On average, each additional child generates approximately 1.3 school trips. Again, the effect of trip linking on the suppression of relationships of, in this case, the number of children and the number of school trips is clear. Two problems concerning school trips were observed. First, several middle-aged adults had school trips recorded; these were apparently drop-off trips that had been miscoded. Second, because the NPTS collects data year round, non-school periods (summer in particular) must be eliminated from home-based school analysis.

College trips as well are dependent on the number of college-aged persons (aged 18-24). The presence of infants tends to suppress the propensity to make college trips. Again, this could be attributed to drop-off trips but is more likely a simple case that adults with very young children are less likely to pursue college courses.

Table 2: Categorization Scheme for Trip Purpose Simulation

Trip Purpose	Categorization Scheme	Mean	Std. Dev'n	GLM Results
Home-Work	0 Workers	0	0	F = 3228,
	1 Worker, 0-1 Vehicles	1.29	1.05	df = 9
	1 Worker, 2+ Vehicles	1.45	1.09	$r^2 = 0.489$
	2 Workers, 0 Children (0-4), 0 Children (5-17)	2.78	1.56	
	2 Workers, 0 Children (0-4), 1+ Children (5-17)	2.56	1.56	
	2 Workers, 1+ Children (0-4), 0 Children (5-17)	2.14	1.40	
	2 Workers, 1+ Children (0-4), 1+ Children (5-17)	2.32	1.39	
	3 Workers, 0 Children (5-17)	4.12	2.05	
	3 Workers, 1+ Children (5-17)	3.75	1.94	
	4 + Workers	5.56	2.41	
Work-Other	0 Workers	0	0	F = 702,
	1 Worker, 0-1 Vehicles	0.98	1.52	df = 9
	1 Worker, 2+ Vehicles	1.06	1.73	$r^2 = 0.172$
	2 Workers, 0 Children (0-4), 0 Children (5-17)	1.91	2.23	
	2 Workers, 0 Children (0-4), 1+ Children (5-17)	2.11	2.39	
	2 Workers, 1+ Children (0-4), 0 Children (5-17)	2.26	2.37	
	2 Workers, 1+ Children (0-4), 1+ Children (5-17)	2.25	2.61	
	3 Workers, 0 Children (5-17)	2.47	2.60	
	3 Workers, 1+ Children (5-17)	2.49	2.63	
	4 + Workers	3.02	2.99	
Home-School	0 Children (5-17)	0	0	F = 14039,
	1 Children (5-17)	1.30	0.90	df = 4
	2 Children (5-17)	2.73	1.50	$r^2 = 0.704$
	3 Children (5-17)	4.16	2.17	
	4+ Children (5-17)	5.46	3.02	
Home-Shop	1 Person, 0 Workers, 0 Vehicles	0.61	1.02	F = 166,
	1 Person, 0 Workers, 1+ Vehicles	0.85	1.14	df = 15
	1 Person, 1 Worker	0.50	0.83	$r^2 = 0.075$
	2 Persons, 0-1 Workers, 0 Vehicles	0.98	1.47	
	2 Persons, 0 Workers, 1+ Vehicles	1.81	1.93	
	2 Persons, 1 Worker, 1+ Vehicles	1.23	1.50	
	2 Persons, 2 Workers	0.96	1.28	
	3 Persons, 1-2 Children (0-4)	0.98	1.28	
	3 Persons, 0 Children (5-17), 0 Children (0-4)	1.69	1.88	
	3 Persons, 1-2 Children (5-17), 0 Children (0-4)	1.40	1.76	
	4 Persons, 0-1 Children (5-17), 0-1 Children (0-4)	1.62	1.93	
	4 Persons, 0-1 Children (5-17), 2-3 Children (0-4)	1.05	1.34	
	4 Persons, 2+ Children (5-17)	1.71	2.15	
	5+ Persons, 0-1 Children (5-17)	1.96	2.18	
	5+ Persons, 2 Children (5-17)	1.84	2.12	
	5+ Persons, 3 +Children (5-17)	2.13	2.53	

Shopping and Other trips are captured in a more complex scheme that illustrates the dynamics occurring in households of different sizes and structures. For one and two-person households, the critical determinants are the presence and number of workers and vehicle availability. For one-person households with no workers, considerably more trips were made if a vehicle was available. For two-person households, the presence of workers suppressed trips of this type because of time constraints and the tendency to push such discretionary and optional trips to the weekends. In three-person households, the critical factor was the suppression of trips caused by the presence of an infant. While this is still true in four-person households, the presence of school age children (ages 5-17) has a reverse effect, causing dramatic increases in trips recorded as "other." This results from parents picking-up/dropping-off children on the way to/from work (again suppressing the number of recorded work trips) and taking their (non-driving) children to after-school activities.

These findings imply that life-cycle factors (particularly worker status and the presence and age of children) are critical determinants of the number of person trips. Traditionally, strong indicators such as income and vehicle availability have less impact on trip frequencies. It is also notable that significant differences in trip making are captured in a relatively simple segmentation of the population, supporting the findings of past researchers[5]. In fact, the majority of the variation is captured in the initial split for all trip purposes.

Travel Mode

The next step of the categorization was to establish groups of households that, for a given trip purpose, chose the same travel mode. The intent here was to maintain the conditionality between the purpose and mode. Given that mode is a nominal dependent variable, the objective was to delineate categories such that the proportional use of each mode was relatively similar. The final 39-category scheme is shown in Table 3.

Work mode shares are captured in a scheme reflecting the presence and availability of automobiles to house-hold workers. The only significant use of transit is by zero-vehicle households. Even in this category, over half the trips are made by private vehicles. With respect to vehicle availability, the impact of competi-tion (reflected by more workers than vehicles) on the driver/passenger split is seen for one-vehicle households only. Where there are two or more vehicles, the effects are relatively minor because the worker will probably have access to his or her own vehicle.

School mode shares reflect a number of underlying dynamics. First, where the number of vehicles equals or exceeds the number of household members, auto driver trips are significantly more frequent. This case is most likely representative of teenagers with their own vehicles who drive to school. Second, as the number of household members increases in relation to the number of vehicles, the proportion of transit or bike/walk trips increases, while the proportion of auto

[5] Vaughn, K.M., Speckman, P. and Sun, D. (1999) Identifying Relevant Socio-Demographics for Distinguishing Household Activity-Travel Patterns: A Multivariate Regression Tree Approach. Paper prepared for The National Institute of Statistical Sciences (NISS), P.O. Box 14006 Research Triangle Park, NC 27709-4006.

passenger trips decreases. This is probably attributable to the likelihood that larger households will include children attending different schools.

College mode shares exhibit somewhat similar dynamics to home-work trips except that the issue becomes automobile availability per college-aged person. Again, significant transit usage only occurs for zero-vehicle households and when there is competition for vehicles.

Shopping and Other mode shares are driven by the dual factors of vehicle availability and the presence of school-age children. Again, zero-vehicle households are the only significant transit users, although, for this purpose, they make a greater proportion of bike/walk trips compared to others. Children are included in the segmentation scheme because of their impact on passenger shares, which is highly significant as shown in Table 3.

As with trip frequencies, household-level mode shares are captured through a relatively simple demographic segmentation. However, the goodness-of-fit measures suggest that most of the variation remains unexplained. This is not surprising, because one would expect mode choice to be significantly affected by local characteristics, such as transit service, congestion, and pricing policies. These factors (or surrogates) must be built into the simulation process to create a data set that is more sensitive to the particular locality. This is a recommendation for future research.

Table 3: Categorization Scheme for Travel Mode Simulation

Trip Purpose	Mode Categories	Driver	Pass.	Transit	Bike/ Walk	Other
Home-Work	0 Vehicles	20%	31%	29%	16%	4%
	1 Vehicle, 1 Worker	87%	6%	2%	4%	1%
	1 Vehicle, 2+ Workers	57%	27%	7%	6%	3%
	2+ Vehicles, 1-2 Workers	92%	5%	1%	1%	1%
	2+ Vehicles, 3+ Workers	85%	10%	2%	3%	1%
Home-School	0 Vehicles	0%	22%	49%	28%	0%
	1 Vehicle, 1-3 Persons	1%	45%	44%	10%	0%
	1 Vehicle, 4+ Persons	1%	31%	44%	24%	0%
	2 Vehicles, 1-2 Persons	46%	12%	21%	21%	0%
	2 Vehicles, 3 Persons	2%	50%	38%	10%	0%
	2 Vehicles, 4 Persons	2%	47%	41%	11%	0%
	2 Vehicles, 5+ Persons	1%	37%	50%	12%	0%
	3+ Vehicles, 1-3 Persons	33%	38%	23%	6%	0%
	3+ Vehicles, 4+ Persons	9%	45%	36%	10%	0%
Home-College	0 Vehicles	18%	20%	25%	36%	1%
	1 Vehicle, 0-1 Persons (18-24)	67%	15%	9%	9%	1%
	1 Vehicle, 2+ Persons	35%	25%	19%	19%	3%
	2+ Vehicles, 2+ Persons (18-24)	94%	5%	0%	1%	0%
	2 Vehicles, 1-2 Persons (18-24)	60%	22%	8%	8%	1%
	2+ Vehicles, 3+ Persons (18-24)	48%	17%	4%	31%	0%
	3+ Vehicles	84%	6%	2%	8%	0%
Home-Shop	0 Vehicles	13%	30%	13%	43%	1%
	1+ Vehicle, 1 Person	90%	6%	0%	4%	0%
	1 Vehicle, 2+ Persons, 0 Children (5-17)	65%	27%	1%	7%	0%
	2+ Vehicles, 2+ Persons, 0 Children (5-17)	83%	14%	0%	2%	0%
	1+ Vehicle, 1+ Children (5-17)	50%	38%	3%	9%	0%
	2+ Vehicles, 1 Child (5-17)	71%	25%	1%	3%	0%
	1+ Vehicle, 2+ Children (5-17)	60%	35%	1%	4%	0%
Home-Other and Non-Home-Other	0 Vehicles	13%	43%	12%	30%	2%
	1+ Vehicle, 1 Person (18+), 0 Children (5-17)	85%	10%	0%	4%	0%
	1+ Vehicle, 2+ Persons (18+), 0 Children (5-17)	73%	22%	1%	4%	0%
	1 Vehicle, 1 Child (5-17)	52%	37%	3%	8%	0%
	2+ Vehicles, 1 Child (5-17)	63%	31%	2%	4%	0%
	1 Vehicle, 2+ Children (5-17)	37%	48%	4%	11%	0%
	2+ Vehicles, 2+ Children (5-17)	47%	44%	1%	8%	0%
Work-Other	0 Vehicles	26%	29%	10%	31%	3%
	1 Vehicle, 1 Worker	81%	9%	1%	8%	1%
	1 Vehicle, 2+ Workers	69%	19%	1%	10%	1%
	2+ Vehicles	84%	11%	0%	4%	0%

Chi-square = 148394, df = 152; Cramer's V = .356; Goodman and Kruskal's tau = .179

Departure Times

The third step of the categorization was to classify households with respect to the departure hour of the trip conditional on the trip purpose and mode. The departure time scheme comprised the 23 categories shown in Table 4. The eta value of 0.321 indicates that the scheme captures approximately 10 percent of the variation in departure times.

Table 4: Categorization Scheme for Departure Time Simulation

Categorization Scheme	6 a.m. - 9 a.m.	9 a.m. – 4 p.m.	4 p.m. - 7 p.m.	7 p.m. – 6 a.m.
Home-Work, POV, Transit	34%	25%	26%	15%
Home-Work, Bike	26%	34%	25%	15%
Home-School, POV, Transit	52%	42%	6%	1%
Home-School, Bike	45%	51%	4%	0%
Home-College, Persons (18-24) = 0	22%	33%	21%	23%
HBCOL, 18-24 >= 1	32%	47%	13%	9%
HBSHOP, 0 Workers	7%	70%	17%	6%
HBSHOP, 1 Worker, 1 Person	7%	34%	36%	24%
HBSHOP, 1 Worker, 2+ Persons	5%	51%	26%	17%
HBSHOP, >=2 Workers, non-passenger	7%	39%	32%	22%
HBSHOP, >=2 Workers, passenger	3%	29%	38%	30%
HBO, 0 Workers	11%	58%	19%	12%
HBO, 1 Worker, 1 Person	9%	33%	27%	31%
HBO, 1 Worker, 2+ Persons	12%	41%	26%	20%
HBO, >=2 Workers, non-passenger	15%	32%	28%	24%
HBO, >=2 Workers, passenger	9%	24%	35%	32%
NHBW, non-bike	15%	62%	18%	5%
NHBW, Bike	5%	84%	8%	3%
NHBO, 0 Workers	5%	77%	13%	5%
NHBO, 1 Worker, 1 Person	4%	51%	25%	21%
NHBO, 1 Worker, 2+ Persons	7%	61%	20%	13%
NHBO, >=2 Workers, non-passenger	8%	53%	23%	15%
NHBO, >=2 Workers, passenger	7%	44%	28%	22%
Total	16%	44%	23%	16%

Chi-square = 9756, df = 66, eta = .321

Work and school departure times are delineated only by whether the mode used was bike/walk as opposed to POV or transit. For home-work and home-school, the bike/walk shares are proportionally higher during the off-peak day-time period (9 a.m. – 4 p.m.) while the opposite trend is observed for the other modes.

Home-college departure times are dependent on the presence of college-age persons, which have been defined here as ages 18-24. For households with no college-age persons, college trips are fairly evenly spread throughout the day reflecting both older adults attending regular college courses and evening classes. For households with one or more college-age persons, the departure times are focused during the morning peak and off-peak day-time periods, reflecting regular student activities.

Home-shop departure times are primarily affected by the presence of household workers. For zero-worker households, the restraints of having to be at work are lifted; over two-thirds of shop trips for this category depart during the off-peak day-time period. In contrast, for one-person working households and two or more worker households, shop trips are shifted to the after-work hours.

Other trip departure times exhibit similar trends to home-shop trips. Again, one sees the impact of workers on the shift of departure times to the after-work hours although the relationships are weaker because of the multitude of reasons underlying these trips.

In summary, departure times are largely dictated by the trip purpose with the presence of workers having an impact on shopping and, to a lesser extent, other trips. While household characteristics have little direct impact on departure times, there are some indirect effects through the demographics that underlie mode choice. Again, one might anticipate that factors pertaining to the particular locality need to be incorporated to more comprehensively account for these differences.

Trip Length

The fourth categorization step was to classify households with respect to reported trip length (in minutes). Trip length should be affected significantly by the size and structure of the region (this issue is addressed in a later section). In addition, trip length should be affected by the results of the previous three steps of the simulation: trip purpose, mode, and departure time. Table 5 shows the final 23-category scheme, which explains 11 percent of the variation in trip lengths. As with mode, this proportion should rise when records are selected from regions of similar spatial characteristics.

POV home-work trip lengths are impacted by the departure time and number of vehicles. Income was a more powerful predictor, but the high incidence of missing income data led its rejection. Shorter trip lengths are associated with the off-peak hours and households with only one vehicle.

POV home-shop trip lengths are longer for households with no workers. This presumably reflects the greater discretionary time available for these households. For households with one or more workers, trip lengths are shorter during the a.m. peak reflecting stops on the way to work such as at a gas station or fast food outlet, both of which are classified as shop trips in the NPTS data base.

POV other trip lengths are impacted by the presence of children and the departure time. For households with one or more children, the average trip length is significantly shorter than for households with no children. These trips are also shorter during the a.m. peak than the rest of the day, reflecting drop-offs at schools and day-care centers. For households with no children, the situation is reversed with the longest trips during the a.m. peak. This presumably reflects trips with a stop (other than dropping a child off) on the way to work.

Table 5: Categorization Scheme for Trip Length Simulation

Categorization Scheme		Mean	Std. Dev'n
POV	Home-Work, 4 p.m. – 9 a.m., 1 Vehicle	20.9	20.6
	Home-Work, 4 p.m. – 9 a.m., 0, 2+ Vehicles	23.1	23.7
	Home-Work, 9 a.m. – 4 p.m., 1 Vehicle	17.0	18.0
	Home-Work, 9 a.m. – 4 p.m., 0, 2+ Vehicles	18.7	25.9
	Home-School	9.3	7.9
	Home-College	18.2	15.3
	Home-Shop, 0 Workers	13.2	12.7
	Home-Shop, 1+ Workers , 6 a.m. – 9 a.m.	11.2	29.7
	Home-Shop, 1+ Workers , 9 a.m. – 6 a.m.	11.7	13.8
	Home-Other/Other-Other, 0 Children, 6 a.m. – 9 a.m.	19.0	43.1
	Home-Other/Other-Other, 0 Children, 9a.m. – 6 a.m.	15.9	26.0
	Home-Other/Other-Other, 1+ Children, 6 a.m. – 9 a.m.	12.8	35.7
	Home-Other/Other-Other, 1+ Children, 9a.m. – 6 a.m.	13.3	20.0
	Work-Other, 4 p.m. – 9 a.m.	20.0	30.1
	Work-Other, 9 a.m. – 4 p.m.	15.0	23.5
Bus	Home-Work	40.9	28.8
	Home-School	22.6	13.8
	All Other Purposes	31.5	37.2
Bike/Walk	All Other Purposes	10.7	12.1
	Work-Other	7.3	7.2
Rail	Home-Work	49.5	23.7
	All Other Purposes	41.0	26.8
Other Modes	All Purposes	183.8	155.6
Total		16.2	24.6

F = 257, df = 22, r-squared = 0.116

Assessment of the Categorization Schemes

The results suggest that demographic factors alone can only partially explain differences in travel behavior particularly for mode, departure time, and trip length. Even with trip rates, most of the variation does not need to be explained. While this may be cause for concern, the critical issue is whether sufficient variation is captured for the simulation results to differentiate among households.

Comparisons with the BRPTS Survey Data

Trip Rate Comparisons

Table 6 shows the mean person trip rates per household from the simulation and the usage of BRPTS. The "p-value" of the z-test for equal population means shows the extent of the evidence against the null hypothesis (i.e., no significant difference in mean trip rates). The results are encouraging with no statistically significant differences except for home-shop trips, which the simulation overestimates.

Table 6: Comparison of BRPTS and Simulation Person Trip Rates per Household

Trip Purpose	BRPTS Data		Simulation Data[1]		
	Mean	Std. Dev'n	Mean	Std. Dev'n	p-value[2]
Home-Work	1.87	1.82	1.83	1.78	.51
Home-School	0.69	1.47	0.74	1.46	.33
Home-College	0.16	0.57	0.17	0.65	.62
Home-Shop	1.17	1.66	1.32	1.75	.01*
Home-Other	3.62	3.50	3.69	3.85	.64
Work-Other	1.46	2.14	1.34	2.06	.10
Other-Other	2.01	2.94	2.02	2.90	.92
All Purposes	11.00	7.85	11.11	7.57	.67

[1] *Average of 5 simulations*

[2] *p-value = probability of incorrectly rejecting the null hypothesis (i.e., making a Type I error)*

**Statistically significant difference in trip rates at the 95th percentile confidence level*

While it is important to show that the procedure produces aggregate trip rates that are comparable to observed trip rates, the procedure must also be validated at a disaggregate level for four reasons. First, one must be wary of potential aggregation bias, which can create misleading conclusions about whether the procedure is working correctly. Second, problems in this simulation step will be propagated through the remaining steps making it imperative to detect them early. Third, one can identify "problem" segments where the simulation appears to perform poorly. In this case, for instance, the home-shop discrepancies may be associated with particular segments of the population. Finally, these disaggregate relationships underpin the development of trip production models (based on cross-classifications of these relationships) such as those tested in the next section.

Again, it must be emphasized that the purpose of the data simulation is to produce a set of trips and their associated attributes that could have been derived from a survey. Therefore, it serves no purpose to compare simulation trip records with actual trip records on a household-by-household basis. Rather, the approach taken was to compare trip rates across segments of the population that were driving the simulation. These include household size, household workers, number of household vehicles, and number of school-age children.

Table 7 shows person trip rate comparisons by household size. Overall, the rates are comparable across the categories, except for home-shop trips for one or two-person households and other-other trips for four-person households. This could be attributed to erroneous or extreme values in the NPTS and/or BRPTS, genuine differences in behavior, or the failure to capture these differences in the categories underlying these cells.

With respect to the problem cells for shop trips, 65 percent of one and two-person households had recorded no shop trips in the BRPTS compared to 52 percent in the simulation. The anomaly for other-other trips was explained by one BRPTS household that had recorded 31 such trips. A review of this record showed it was a family with two workers and two school age children who, after work, had all made trips to a gas station, several stores, the park, and the school for an evening activity. The point here is to illustrate that one should not jump to the automatic conclusion that extreme values are erroneous. In actuality, the non-recording of trips is probably a far greater problem. However, one must be aware of how extreme values affect comparisons with a disaggregate analysis.

Table 7: Comparisons of Person Trip Rates per Household by Household Size

Trip Purpose	Data Source	Persons per Household				
		1	2	3	4	5+
Home-Work	BRPTS	0.75	1.64	2.19	2.52	3.00
	Simulation Data	0.71	1.67	2.21	2.40	2.73
	p-value	0.63	0.78	0.87	0.45	0.25
Home-School	BRPTS	0.00	0.05	0.66	1.57	2.80
	Simulation Data	0.00	0.07	0.66	1.73	2.96
	p-value	-----	0.24	0.96	0.27	0.58
Home-Shop	BRPTS	0.47	0.95	1.43	1.44	2.30
	Simulation Data	0.67	1.24	1.36	1.63	2.23
	p-value	0.01**	0.00**	0.57	0.24	0.81
Home-Other	BRPTS	1.46	2.84	3.58	5.35	7.57
	Simulation Data	1.32	2.75	3.83	5.84	7.41
	p-value	0.27	0.58	0.32	0.20	0.80
Other-Other	BRPTS	0.74	1.63	1.85	3.40	3.67
	Simulation Data	0.95	1.63	2.18	2.75	3.81
	p-value	0.11	0.99	0.14	0.03*	0.78
All Purposes	BRPTS	3.98	8.64	11.46	16.51	22.04
	Simulation Data	4.28	8.68	12.01	16.49	21.38
	p-value	0.20	0.86	0.19	0.96	0.52

**Statistically significant difference in trip rates at the 95th percentile confidence level*
***Statistically significant difference in trip rates at the 99th percentile confidence level*

Similar comparisons were made for demographic groups such as number of workers, number of children, and household vehicles. Other than a few “problem” cells the simulation was able to provide comparative rates to the actual survey data. The next step was to determine whether this comparison held true across cross-classifications of these descriptors. Of the various scenarios tested, a “life-cycle” scenario is reported here that incorporates elements pertaining to the presence/absence of workers, the number of adults, and the presence of children of pre-school and school-age. The scheme was based on life-cycle categories proposed for activity classification[6]. The major modification made for the current analyses was the delineation of single parent households because this group has somewhat unique travel characteristics. The final six-category scheme is shown in Table 8. In subsequent analysis, a seventh category was added, which is also shown. This category was not used in the Baton Rouge analysis.

---

[6] Vaderevu, R.V. and Stopher, P.R. (1996) Household Activities, Life Cycle and Role Allocation. *Paper presented at the 75th Annual Meeting of the Transportation Research Board*, Washington DC.

Table 8: Life-Cycle Categories

Category	Description	Occurrence in the BRPTS	
1	Single person household, person is employed	99	10.1%
2	Single parent household	58	5.9%
3	Multiple-adult households, at least one employed, no children	313	31.8%
4	Multiple-adult households, at least one employed, one or more children, none of school age.	60	6.8%
5	Multiple-adult households, at least one employed, one or more school-age children.	318	32.3%
6	One or more adults, none employed, no children	169	17.2%
7	One or more adults, none employed, one or more children present.	0	0%

Table 9 shows the trip rates across the six life-cycle categories. The simulation data again compare favorably except for a few anomalous cells. These data include category 6 and category 3 for home-shop trips and category 6 for all purposes. They follow the univariate trends for workers and children depicted previously. Overall, these results are an encouraging precursor to the development of trip-production models, which are built on cross-classifications of similar variables. In general, the simulation worked best for work, school, college, and home-other trips although there may be a possible aggregation bias within the "home-other" category because it encapsulates trips of many different purposes (i.e., serve passenger, social recreational, personal business). Home-shop trips performed the worst although the problems appear to be centered on small households and households with no workers.

Table 9: Comparisons of Person Trip Rates by Life-Cycle Categories

Trip Purpose	Data Source	Life-Cycle Category					
		1	2	3	4	5	6
Home-Work	BRPTS	1.39	1.03	2.51	1.93	2.68	0.00
	Simulation Data	1.32	1.30	2.49	2.12	2.48	0.00
	p-value	0.52	0.18	0.84	0.40	0.10	-----
Home-Shop	BRPTS	0.54	1.24	1.24	1.15	1.76	1.58
	Simulation Data	0.43	0.84	1.04	1.35	1.78	1.30
	p-value	0.29	0.07	0.04*	0.33	0.92	0.00**
Home-Other	BRPTS	1.19	3.95	2.72	2.95	6.12	2.73
	Simulation Data	1.12	4.01	2.80	3.14	6.23	2.68
	p-value	0.62	0.92	0.64	0.62	0.73	0.82
	BRPTS	4.61	11.21	10.00	9.58	18.49	4.92
All Purposes	Simulation Data	4.87	11.66	10.04	9.93	17.94	5.76
	p-value	0.41	0.65	0.89	0.63	0.30	0.03*

**Statistically significant difference in trip rates at the 95th percentile confidence level*

***Statistically significant difference in trip rates at the 99th percentile confidence level*

Mode Share Comparisons

Table 10 shows aggregate mode share comparisons for the BRPTS and the simulation data. Overall, auto driver shares are seriously under-estimated, while auto passengers, bike/walk, and transit shares are over-estimated. The situation is marginally improved by selecting records from

MSAs of a similar size to Baton Rouge (500,000 – 1,000,000). These records formed the basis of further comparisons of mode, departure times, and trip lengths.

Table 10: Comparison of BRPTS and Simulation Mode Shares

Trip Purpose	Mode	BRPTS	Simulation Trip Data	
			All NPTS Records	MSAs 500,000 – 1,000,000
All Purposes	Auto Driver	70.4%	64.6%**	66.1%**
	Auto Pass.	22.3%	23.2%	24.1%**
	Transit	3.5%	5.7%**	4.6%**
	Bike/ Walk	3.7%	6.5%**	5.1%**

**Statistically significant difference in mode shares at the 99^{th} percentile confidence level.

Table 11 determines whether these trends hold true across the seven trip purposes. Also shown are the partial simulation results. By way of recall, these are simulated mode shares based on actual trip production data from the BRPTS. The trend of under-estimating auto driver shares and over-estimating the other mode shares is apparent in most cases with the worse problems for home-shop and home-other trips. The partial simulation results show a general improvement, which suggests that some of the discrepancies may be due to problems in the prior (trip frequency) step of the simulation.

Table 11: Comparisons of BRPTS and Simulation Mode Shares by Trip Purpose

Trip Purpose	Mode	BRPTS	Full Simulation	Partial Simulation
Home-Work	Auto Driver	91.9%	89.0%**	89.5%**
	Auto Pass.	6.9%	7.5%	7.0%
	Transit	0.4%	1.8%**	1.8%**
	Bike/ Walk	0.8%	1.7%*	1.6%
Home-School	Auto Driver	4.4%	4.3%	3.9%
	Auto Pass.	42.0%	37.6%	37.8%
	Transit	46.0%	49.1%	48.4%
	Bike/ Walk	7.6%	9.0%	9.9%
Home-College	Auto Driver	76.0%	73.0%	74.8%
	Auto Pass.	10.1%	9.2%	12.2%
	Transit	4.7%	6.3%	4.6%
	Bike/ Walk	9.3%	11.2%	7.6%
Home-Shop	Auto Driver	76.2%	71.6%**	71.4%**
	Auto Pass.	20.6%	22.6%	22.0%
	Transit	0.1%	1.6%**	0.7%*
	Bike/ Walk	3.1%	4.2%	5.9%**
Home-Other	Auto Driver	65.6%	60.6%**	62.7%*
	Auto Pass.	28.0%	31.8%**	30.7%*
	Transit	1.1%	2.0%*	1.3%
	Bike/ Walk	5.4%	5.5%	5.2%
Work-Other	Auto Driver	86.7%	84.4%	85.1%
	Auto Pass.	10.2%	10.1%	9.6%
	Transit	0.3%	1.3%**	1.2%**
	Bike/ Walk	2.8%	4.2%*	4.0%
Other-Other	Auto Driver	64.8%	63.0%	61.3%*
	Auto Pass.	30.8%	30.1%	32.1%
	Transit	1.7%	1.8%	1.4%
	Bike/ Walk	2.6%	5.1%**	5.0%**

**Statistically significant difference in mode shares at the 95th percentile confidence level*

**Statistically significant difference in mode shares at the 99th percentile confidence level.

A disaggregate analysis was conducted to compare how the mode simulation operated for various segments of the population. For instance, Table 12 shows that the simulation data captures similar trends across the life-cycle groups to the actual data even if the actual magnitudes are dissimilar. The "1 person, 1 worker group" has the highest drive alone shares and is the most comparable category to the BRPTS. Single-parent households have the highest proportion of transit and bike/walk trips and the lowest proportion of drive-alone trips, while multiple-adult households with school-age children have the highest passenger shares, reflecting the presence of children. For multiple adult households without children, the private automobile dominates with 97 percent of trips made using this mode. This is the case whether workers are present or not.

Table 12: Mode Share Comparisons by Life-Cycle Groups

Life-Cycle Grouping	Data Source	Driver	Passenger	Transit	Bike/Walk
1 person, 1 worker	BRPTS	88.4%	7.5%	0.5%	3.6%
	Full Simulation	87.3%	8.0%	0.9%	3.8%
	Partial Simulation	86.5%	7.4%	1.2%	5.0%
1 adult, 1+ children	BRPTS	49.7%	25.1%	12.0%	13.1%
	Full Simulation	44.8%	33.5%**	12.7%	9.1%**
	Partial Simulation	45.2%	34.5%**	10.8%	9.5%**
2+ adults, 1 worker, 0 children	BRPTS	83.0%	14.5%	0.5%	1.9%
	Full Simulation	79.2%**	15.9%	1.3%**	3.5%**
	Partial Simulation	80.6%*	15.0%	1.1%*	3.2%**
2+ adults, 1 worker, 1+ children (0-4)	BRPTS	83.9%	12.2%	1.2%	2.8%
	Full Simulation	78.4%*	17.0%*	1.5%	3.1%
	Partial Simulation	78.7%*	17.5%*	1.3%	2.6%
2+ adults, 1 worker, 1+ children (5-17)	BRPTS	59.9%	30.7%	5.5%	3.8%
	Full Simulation	56.5%**	30.4%	7.7%**	5.3%**
	Partial Simulation	58.3%	30.3%	6.1%	5.3%**
1+ adults, 0 workers, 0 children	BRPTS	80.0%	17.2%	0.1%	2.7%
	Full Simulation	74.8%**	19.0%	1.7%**	4.4%*
	Partial Simulation	77.4%	17.2%	0.7%*	4.6%*

**Statistically significant difference in trip rates at the 95th percentile confidence level*
***Statistically significant difference in trip rates at the 99th percentile confidence level*

The results of this analysis indicate that demographic characteristics alone are only partially capable of capturing differences in mode shares. This is not surprising because elements of the transportation and spatial environment must be incorporated to fully capture the nuances between regions. For instance, in Baton Rouge, one would expect a higher than average household size, lower than average automobiles per worker, higher proportions of college-age persons, and relatively high percentages of non-car owning households to be reflected in lower proportions of auto driver trips and higher shares of the other modes. This is clearly not the case. Other factors are driving these differences such as ample supplies of free/low-cost parking, the lack of bicycle and pedestrian facilities, sparse transit coverage, and the dispersion of residential and commercial activities.

Departure Time Comparisons

Table 13 shows the departure time comparisons for the BRPTS and simulation data. The Kolmogorov-Smirnov z-test indicates that, overall, the distributions of hourly departure times are generally comparable across the trip purposes apart from other-other trips. The z-test of proportions, however, shows that home-work trips leaving during the a.m. peak are overestimated, while home-shop and home-other trips leaving during the same period are underestimated. These trends are difficult to definitively explain, but they seem to be region-specific. This suggests that local information must be incorporated to capture the idiosyncrasies driving these differences.

Table 13: Comparisons of BRPTS and Simulation Departure Times by Trip Purpose

Trip Purpose	Departure Time-Period	BRPTS	Full Simulation		Partial Simulation	
			Proportion	K-S z-value[1]	Proportion	K-S z-value
Home-Work	6 a.m. – 9 a.m.	33%	36%*	.807	35%	.533
	9 a.m. – 4 p.m.	25%	24%		26%	
	4 p.m. – 7 p.m.	27%	24%*		24%*	
	7 p.m. – 6 a.m.	15%	15%		15%	
Home-School	6 a.m. – 9 a.m.	51%	50%	1.152	52%	1.363
	9 a.m. – 4 p.m.	44%	44%		41%	
	4 p.m. – 7 p.m.	4%	5%		6%	
	7 p.m. – 6 a.m.	1%	1%		1%	
Home-Other	6 a.m. – 9 a.m.	15%	12%**	1.312	13%	.956
	9 a.m. – 4 p.m.	34%	37%*		38%**	
	4 p.m. – 7 p.m.	28%	28%		28%	
	7 p.m. – 6 a.m.	23%	22%		21%*	
Other-Other	6 a.m. – 9 a.m.	9%	7%	1.322	8%	2.240**
	9 a.m. – 4 p.m.	51%	58%**		58%**	
	4 p.m. – 7 p.m.	24%	20%**		20%**	
	7 p.m. – 6 a.m.	16%	15%		15%	
All Purposes	6 a.m. – 9 a.m.	18%	17%	.90	18%	1.44*
	9 a.m. – 4 p.m.	42%	44%**		45%**	
	4 p.m. – 7 p.m.	24%	23%*		23%*	
	7 p.m. – 6 a.m.	16%	16%		15%	

**Statistically significant difference in departure times at the 95th percentile confidence level.*
***Statistically significant difference in departure times at the 99th percentile confidence level.*
[1]Kolmogorov-Smirnov z-value of hourly departure times.

Comparing the full and partial simulation results again, one does not see notable discrepancies between them. This reflects the fact that the demographics encapsulated in the categories have little direct effect on departure times as noted in the discussion in the section on departure time schemes.

Trip Length Comparisons

Table 14 provides trip length comparisons between the BRPTS and simulation data. While the mean trip lengths are closely replicated for all trip purposes, the Kolmogorov-Smirnov z-value statistic suggests problems in how well the distribution of values matches for home-work, home-school, and home-other trips. Analysis by time-interval suggests that the discrepancy is due to the over-estimation of short trips (less than 10 minutes) using the simulation data. Again, this suggests that local spatial information is required to fully capture these nuances that are not captured directly through the demographic categories.

Table 14: Comparison of BRPTS and Simulation Vehicle Trip Lengths (minutes)

Trip Purpose	BRPTS		Full Simulation			Partial Simulation		
	Mean	Std. Dev'n	Mean	Std. Dev'n	K-S z-value[1]	Mean	Std. Dev'n	K-S z-value
Home-Work	21.6	15.8	19.3	15.3	2.784**	20.1	16.3	2.06**
Home-School	16.8	11.7	8.7	7.0	3.611**	9.3	8.0	3.274**
Home-College	19.5	13.1	17.5	13.0	1.289	16.9	9.2	1.014
Home-Shop	10.5	8.3	11.4	10.3	1.073	11.3	9.5	.863
Home-Other	14.2	13.6	13.2	13.3	3.042**	13.5	14.9	3.199**
Work-Other	15.4	15.7	14.6	15.0	.600	15.9	17.8	.642
Other-Other	13.8	13.3	13.2	12.9	1.164	13.6	14.3	.872
All Purposes	15.4	14.0	14.1	13.6	4.133**	14.9	15.1	2.89**

***Statistically significant difference in trip lengths at the 99th percentile confidence level.*

[1]Kolmogorov-Smirnov z-value (asymptotic significance) of reported trip lengths.

Summary of Findings

The results are generally encouraging and suggest that the data simulation is capable of producing comparable person trip rates per household to those derived from an actual survey. This was shown to be true for sociodemographic segments of the population as well as for the population as a whole. However, the simulation is only partially able to produce comparable mode shares, departure times, and reported trip lengths. For these attributes, characteristics pertaining to the particular location must be intuitively incorporated within the simulation procedure itself. Two suggestions on how this could be achieved are proposed and tested in the final section of this analysis.

Trip Production Model Comparisons

The next research phase was to determine whether trip production models estimated with the simulation data offered improvements over borrowed trip production models, improvements over the use of national default values, and comparable results to models estimated from actual survey data. The measuring tool for establishing these improvements was the various models estimated with the BRPTS data.

Comparisons with Borrowed Trip Production Models

The first objective was to establish that trip production models estimated with the simulation data offered improvements over the use of borrowed models. Baton Rouge currently uses borrowed vehicle trip production models comprising three household size categories (1-2, 3-4, 5+) to predict three internal trip purposes (home-work, home-other and non-home-based). The models date from the early 1980s and other than some minor updates have retained their basic functional form.

Table 15 compares the vehicle trip rates and calculated trips for the current Baton Rouge models with the rates estimated from the BRPTS and the simulation data. The calculated trips are simply the trip rates multiplied by the number of households in that category. For this comparison, the BRPTS and simulation vehicle trip rates were derived by applying a calculated auto occupancy factor to the person trip rates (by vehicles) for each household size/trip purpose grouping.

The first point to note is that if one assumes the BRPTS to be the closest representation of reality, the current rates are seriously deficient. The lower rates for work trips, particularly for larger households, reflect the increase in multi-worker households since the models were originally estimated. The increase in home-other and non-home-based trips is probably attributable to declining auto-occupancy rates and increases in trip chaining. It is also probable that these discrepancies are exaggerated by differences in survey methodologies between the 1995 NPTS and the survey used to originally derive the rates used in the current Baton Rouge models. For instance, the NPTS was particularly effective at capturing short, discretionary trips compared to previous waves of the survey.

The second point to note is that the simulation data provides trip rates that are significantly closer to reality than what is currently used. Overall, trips are under-estimated by a rate approximately five trips per percentage point. This can be attributed to the under-estimation of auto driver trips in the mode share simulation and consequently the over-estimation of auto occupancy rates. The RMSE of .34 suggests some significant discrepancies although this is largely attributable to home-other trips for households with five or more persons.

Table 15: Comparisons with Current Borrowed Vehicle Trip Production Models

Trip Purpose	Household Size Category	Vehicle Trip Production Rates			Calculated Trips		
		BRPTS Data(1)	Current Models(2)	Synthetic Data(3)	BRPTS Data	Current Models	Synthetic Data
Home-Work	1-2	1.259	0.966	1.204	106,183	81,467	101,496
	3-4	2.108	1.402	2.049	112,724	74,962	109,571
	5+	2.699	1.863	2.362	47,901	33,063	41,921
	Total (% diff.)				266,807	189,491 (-28%)	252,988 (-4%)
Home-Other	1-2	2.553	1.956	2.467	215,271	164,957	208,040
	3-4	4.041	4.100	3.885	216,077	219,219	207,720
	5+	5.727	4.650	4.964	101,633	82,524	88,091
	Total (% diff.)				532,981	466,700 (-12%)	503,851 (-5%)
Non-Home-Based	1-2	1.945	1.179	1.839	164,053	99,430	155,076
	3-4	3.024	2.557	3.039	161,687	136,718	162,514
	5+	4.109	2.816	3.682	72,928	49,976	65,342
	Total (% diff.)				398,667	286,123 (-28%)	382,932 (-5%)
All Trips					1,198,456	942,314 (-21%)	1,139,771 (-5%)

Root Mean Square Error of trip rates: (1) versus (2) = .81; (1) versus (3) = .34; (2) versus (3) = .56.

Comparisons with National Default Trip Rates

Given the age of the current models, the results of the previous test are not surprising. The second objective was to establish that the simulation data offered improvements over the use of national default statistics for trip-production model estimation. As noted previously, national default statistics are provided in NCHRP 365. These techniques are appealing because of their simplicity, low data requirements, and their recent incorporation in computer travel-demand modeling packages. However, it is suspected that (potentially significant) differences between regions may not be detected by using national averages differentiated on four categories of urban area size only. However it is suspected that aggregate national trip production statistics differentiated on four categories of urban area size may only hide (potentially significant) differences between regions.

For the purposes of this comparison, average daily person trips per household by household size (1-5+) were used for urbanized areas of 200,000 – 499,999. NCHRP 365 provides percentage breakdowns for three purposes (home-based work, home-based non-work, and non-home-based) for each household size category. Table 16 shows the comparisons with the person trip rates per household calculated from the BRPTS data and the simulation data. Substantial differences are apparent between the NCHRP 365 rates and the survey rates, which primarily affects the prediction of home-other and non-home-based trips. Again, this is probably indicative of the fact that many more of these types of trips were recorded in the 1995 NPTS and 1997 BRPTS surveys.

Table 16: Comparisons with NCHRP 365 Person Trip Production Rates

Scenario	Household Size	Average Daily Person Trips per Household	% Average Daily Person Trips by Purpose			Calculated Person Trips		
			HBW	HBO	NHB	HBW	HBO	NHB
BRPTS Data	1	4.0	0.19	0.51	0.30	28770	77345	45656
	2	8.6	0.19	0.46	0.35	75908	183432	139550
	3	11.5	0.19	0.51	0.30	61832	165646	96531
	4	16.5	0.15	0.52	0.33	63651	216052	136374
	5+	22.0	0.14	0.58	0.28	53241	228698	109226
	Totals	11.00				283402	871173	527337
NCHRP 365	1	3.7	0.20	0.56	0.24	28232	79049	33878
	2	7.1	0.23	0.53	0.24	75417	173787	78696
	3	10.8	0.22	0.54	0.24	67179	164894	73286
	4	13.4	0.18	0.61	0.21	60768	205936	70896
	5+	15.9	0.19	0.59	0.22	53614	166485	62079
	Totals	9.2				285210 (+1%)	790151 (-10%)	318835 (-65%)
Simulation Data	1	4.3	0.17	0.49	0.35	27269	79346	56622
	2	8.7	0.19	0.48	0.32	77135	194405	129558
	3	12.0	0.18	0.50	0.31	62431	170415	106784
	4	16.5	0.15	0.57	0.28	60435	237367	117552
	5+	21.4	0.13	0.60	0.27	48521	228625	102311
	Totals	11.11				275791 (-3%)	910158 (+4%)	512827 (-3%)

The simulation data compares more favorably with the BRPTS data supporting the contention that this approach is preferable to the use of national default statistics. However, this and the previous comparison were made without knowledge of the data underlying these borrowed rates and information on how they were calculated, which makes definitive conclusions difficult to reach.

Comparisons with New Trip Production Models

Having established that the simulation data offered improvements over existing modeling techniques (in terms of trip rates), the third objective was to determine whether these data were capable of estimating new trip production models that were comparable to the same models using actual travel survey data (the BRPTS in this case). Five new person trip production models were estimated: home-work, home-school/college, home-shop, home-other, and non-home-based. All models were specified using a cross-classification scheme that comprised five household (1-5+) and four vehicle ownership (0-3+) categories, a scheme that might be encountered in a "typical" regional modeling effort. Five of the categories were merged with neighboring cells because of sparse sample sizes. The final categories and their occurrence within the BRPTS are shown in Table 17.

Table 17: The New Baton Rouge Trip Production Model Scheme

Household	Household Size					
Vehicles	1	2	3	4	5+	Totals
0	1 = 39			5 = 27		66
1		3 = 90	6 = 25	9 = 26		301
2	2 = 160	4 = 233	7 = 94	10 = 81	12 = 47	455
3+			8 = 70	11 = 56	13 = 36	162
Totals	199	323	216	163	83	984

Trip production models were estimated for each purpose and calibrated using multiple classification analysis (MCA). Table 18 through Table 20 provide comparisons of the new models estimated with the BRPTS data and the simulation data together with chi-square and RMSE statistics. The "p-value" is provided to enable the reader to assess the weight of evidence against the null hypothesis (i.e., no significant differences in mean trip rates across the cells). An overall assessment suggests that home-work and home-school/college trips are well estimated, home-other and non-home-based trips are acceptably estimated, and home-shop trips are marginally well estimated.

While the aggregate statistical measures indicate an overall sense of the "success" of the models estimated from the simulation data, it is critical to look at the differences between each cell in the trip production model. Invariably, one or two cells contribute disproportionately to the finding of a significant difference. The problem in these anomalous cells could be attributable to spurious values in either of the two data sets that could inflate or deflate a final value. Clearly, this suspicion must be examined further.

Table 18: Home-Work Person Trips per Household

Model Category	Person Trips per Household		Calculated Household Trips		
	BRPTS Data	Simulation Data	BRPTS Data	Simulation Data	% Difference
1	0.57	0.53	4,381	4,032	-9%
2	0.82	0.79	27,144	26,354	-3%
3	1.11	1.20	12,592	13,608	7%
4	1.90	1.90	60,896	60,838	0%
5	0.98	1.10	4,537	5,047	10%
6	1.23	1.36	5,847	6,481	10%
7	2.02	2.06	27,920	28,496	2%
8	2.97	2.88	24,542	23,785	-3%
9	1.58	1.62	9,705	9,906	2%
10	2.38	2.32	29,969	29,234	-3%
11	3.33	3.14	24,283	22,885	-6%
12	2.84	2.52	23,384	20,764	-13%
13	3.79	3.34	20,758	18,288	-14%
Total	-----	-----	275,956	269,718	-2%

Chi-Square = 6.7, df = 12, p-value = 0.9; RMSE = 0.18

Table 19: Home-Other Person Trips per Household

Model Category	Household Trip Rates		Calculated Household Trips		
	BRPTS Data	Simulation Data	BRPTS Data	Simulation Data	% Difference
1	1.83	1.92	13,971	14,630	5%
2	1.47	1.23	48,752	40,860	-19%
3	2.77	3.03	31,549	34,494	9%
4	2.89	2.71	92,703	86,826	-7%
5	3.91	4.91	18,008	22,620	20%
6	3.54	4.22	16,891	20,122	16%
7	3.66	3.90	50,625	53,882	6%
8	3.73	3.83	30,780	31,665	3%
9	5.25	6.12	32,154	37,514	14%
10	5.36	5.80	67,669	73,125	7%
11	5.43	5.74	39,617	41,836	5%
12	7.96	7.62	65,542	62,726	-4%
13	8.03	7.56	43,925	41,352	-6%
Total			552,187	561,652	2%

Chi-Square = 30.7, df = 12, p = 0.00; RMSE = 0.53

Table 20: Non-Home-Based Person Trips per Household

Model Category	Household Trip Rates		Calculated Household Trips		
	BRPTS Data	Simulation Data	BRPTS Data	Simulation Data	% Difference
1	0.68	1.25	5,178	9,535	46%
2	1.38	1.57	46,046	52,055	12%
3	2.65	2.78	30,162	31,628	5%
4	3.25	2.89	104,124	92,671	-12%
5	2.06	3.31	9,472	15,259	38%
6	2.76	3.63	13,168	17,294	24%
7	3.36	3.74	46,439	51,685	10%
8	3.81	4.11	31,486	33,917	7%
9	4.92	4.51	30,153	27,668	-9%
10	5.52	4.62	69,581	58,336	-19%
11	5.97	4.99	43,553	36,433	-20%
12	5.94	5.81	48,930	47,869	-2%
13	6.99	6.18	38,270	33,840	-13%
Total			516,563	508,190	-2%

Chi-Square = 76.9, df = 12, p = 0.00; RMSE = 0.72

Transferability of the Approach

Introduction

The purpose of this part of the research was to determine whether the approach works as well for other locations as for Baton Rouge, particularly where there may also not be a similar survey to the NPTS against which to make comparison. As described earlier, two metropolitan areas were selected for the tests of simulation household travel survey data: the North Central Texas region and the Wasatch Front and Mountainland region in Utah.

Having prepared the PUMS data for these two regions and using the same distributions as were developed for the Baton Rouge simulation, samples were drawn from each of the two regions that corresponded in sampling characteristics to the samples drawn in the most recent household travel surveys in the two regions. The simulation procedure followed exactly along the lines described for Baton Rouge in the preceding sections of this report.

Sample Comparison

Table 21 shows some basic statistics for comparing the HTS and simulation samples for each of the two regions. The distributions of households by household size, vehicles, and geographic area for each of Dallas and Salt Lake City were identical and are not shown in the table. The simulation samples match very closely on the tabulated statistics to the original surveys. However, there is one important issue on which the simulation and original samples do not match well. In the Dallas data, there are 278 non-mobile households, while there are only 68 such households in the simulation data. This arises because of differences in the survey techniques that resulted in lower non-mobility rates in the NPTS than in the Dallas data. This affects comparisons of trip rates but does not affect the other characteristics materially. For Salt Lake City, the match on non-mobile households is closer than for Dallas with 102 households reporting no trips on the diary day while the simulation produced 63 such cases.

Table 21: Comparison of Statistics of the HTS and Synthetic Samples

Statistic	Dallas		Salt Lake	
	HTS	Simulation	HTS	Simulation
Sample Size (Households)	3,996	3,988	3,082	3,082
Average Household Size	2.47	2.44	3.14	3.34
Average Vehicle Ownership per Household	1.84	1.76	1.97	1.97
Average Workers per Household	1.38	1.40	1.31	1.42
Average Household Income	$49,255	$44,276	$20,001-$30,000	$20,001-$30,000

Trip Rate Simulations

Table 22 compares the person trip rates by purpose between the simulated and the HTS data for each region. All trip rates except home-school and home-college are overestimated by the simulation for Dallas and are significantly different at the 99% confidence level. This also results in the total number of trips being overestimated. The overestimation is expected, given the difference in non-mobility rates. The home-work trip rate for mobile households for Dallas is 1.86 (compared to 1.73 averaged across mobile and non-mobile households), while the mobile households rate for the simulation is 1.89. Similarly, the home-other trip rate would increase from 2.91 to 3.16 for the Dallas data, while the simulation rate increased only from 3.19 to 3.25. These results significantly reduce the apparent differences in the simulation and actual data.

The trip rates for Salt Lake City are all significantly different statistically at beyond 99 percent except for home-school and home-college (95%) and other-work trips, which are not significantly different. The home-work trip rate would be expected to be lower, and the home-other and other-work rates would be expected to be higher. This is due to the larger families in Utah than the national average. These effects are insufficiently captured by the socio-demographic characteristics used in the simulation.

Table 22: Comparisons of HTS and Simulated Person Trip Rates per Household

Trip Purpose	Dallas		Salt Lake		Baton Rouge		
	HTS Mean	Simulated Mean	HTS Mean	Simulated Mean	Simulated Mean	Diff. From Dallas	Diff. From Salt Lake
Home-Based Work	1.73	1.86**	1.66	1.83**	1.83		
Home-Based School	0.57	0.60	1.20	1.07*	0.74	**	**
Home-Based College	0.17	0.16	0.28	0.23*	0.17		*
Home-Based Shop	0.63	1.14**	1.25	1.38**	1.32	**	
Home-Based Other	2.91	3.19**	4.93	4.17**	3.69	**	**
Other-Work	1.17	1.35**	1.29	1.33	1.34		
Other-Other	1.29	1.86**	2.67	2.26**	2.02	*	*
TOTAL TRIPS	8.47	10.17**	13.28	12.28**	11.11	**	**

* Statistically significant difference in trip rates at the 95 percent confidence level
**Statistically significant difference in trip rates at the 99 percent confidence level

Comparing Dallas and Salt Lake City to Baton Rouge indicates whether the simulation process is responsive to the characteristics of the local area. Table 22 shows the Baton Rouge simulated data and the significance of differences between each of the three locations. Work related trips are the least likely to be different and, for both Dallas and Salt Lake City, home-work and other-work trips are not significantly different at the 95% confidence level. For the other trip purposes, home-college trips in Dallas and home-shop trips in Salt Lake City are also not significantly different, while all other purposes are significantly different, showing that the simulation procedure has responded to differences in household characteristics in the regions.

Table 23 shows differences in the data by household size. Household size is a particularly important variable to consider because it is both a classification variable for the simulation for most trip purposes and a prime variable in trip-production modeling. Compared to Baton Rouge, there are many more significant differences in trip rates between survey and simulation data for both Dallas and Salt Lake City. There are significant differences in Dallas trip rates for almost all household sizes relating to home-shop, home-other, and non-home-based travel. Overall, trip rates are overestimated by the simulation, again, as a result of the nonmobility differences between the actual and simulation data. Numerical differences are generally small, though this is less so for home-shop and other-other trips. The pattern of overestimation appears to be consistent across household sizes.

Table 23: Comparisons of Person Trip Rates per Household by Household Size

Trip Purpose	Data Source	Persons per Household				
		1	2	3	4	5+
Home-Work	Dallas HTS	0.95	1.78	2.22	2.23	2.34
	Simulated Data	0.96	1.97**	2.35**	2.27**	2.77**
	Salt Lake HTS	0.75	1.41	1.95	2.16	2.23
	Simulated Data	0.75	1.57**	2.01	2.39**	2.61**
Home-School	Dallas HTS	0.00	0.09	0.61	1.57	2.86
	Simulated Data	0.00	0.09	0.60	1.66*	3.06**
	Salt Lake HTS	0.00	0.04	0.55	1.35	4.15
	Simulated Data	0.00	0.07**	0.52	1.33	3.29**
Home-College	Dallas HTS	0.06	0.09	0.31	0.28	0.38
	Simulated Data	0.12**	0.13**	0.21**	0.20**	0.32**
	Salt Lake HTS	0.06	0.29	0.29	0.38	0.40
	Simulated Data	0.08	0.18**	0.30	0.28**	0.34*
Home-Shop	Dallas HTS	0.35	0.73	0.73	0.69	0.85
	Simulated Data	0.59**	1.11**	1.28**	1.69**	1.88**

	Salt Lake HTS	0.51	1.04	1.31	1.46	2.00
	Simulated Data	0.60**	1.21**	1.38	1.60**	2.08
Home-Other	Dallas HTS	1.08	2.40	3.38	5.07	6.70
	Simulated Data	1.26**	2.51	3.78**	5.48**	7.69**
	Salt Lake HTS	1.64	3.26	4.39	5.74	9.82
	Simulated Data	1.33**	2.72**	3.70**	5.40**	7.77**
Other-Work	Dallas HTS	0.72	1.07	1.54	1.68	1.46
	Simulated Data	0.73	1.31**	1.85**	1.71	2.09**
	Salt Lake HTS	0.63	1.09	1.54	1.82	1.59
	Simulated Data	0.60	1.07	1.43*	1.76	1.93**
Other-Other	Dallas HTS	0.62	1.18	1.46	1.95	2.64
	Simulated Data	0.81**	1.64**	2.15**	2.89**	4.15**
	Salt Lake HTS	1.00	2.06	2.61	3.10	4.69
	Simulated Data	0.77**	1.66**	1.99**	3.03	3.92**
All Purposes	Dallas HTS	3.78	7.34	10.26	13.46	17.24
	Simulated Data	4.47**	8.75**	12.23**	15.89**	21.97**
	Salt Lake HTS	4.60	9.19	12.63	16.00	24.88
	Simulated Data	4.13**	8.47**	11.3388	15.79	21.94**

* Statistically significant difference in trip rates at the 95 percent confidence level

**Statistically significant difference in trip rates at the 99 percent confidence level

There are fewer significant differences for Salt Lake City, however. From Table 23, one-person households show significant differences for home-shop, home-other, and other-other trips and for total trips. Two-person households show all rates to be significantly different except for other-work trips. For three-person households, there are significant differences for home-other, other-work, other-other and total trips. Four-person households show significant differences for all home-based trips except home-school, but do not show any differences for the non-home-based categories. The really substantial differences all occur in households with 5 or more people. Because large households are prevalent in Utah and the average household size for Salt Lake City is higher than the national average, this is not an unexpected result. Simulation produces a trip rate of almost 3 fewer trips per day than the survey data, and all purposes except home-shop trips are significantly different. Furthermore, the home-work and other-work trips are too high in the simulation for this group of households, while all other purposes (except home-shop, which is not statistically significantly different) are too low in the simulation. These are the expected results given the difference in family sizes. They also confirm the conclusion drawn on the overall trip rates.

Disaggregate analysis was conducted of household lifecycle and trip rates by numbers of workers, vehicles, and school age children. In general the results were similar to those shown in Table 24, which compares the person trip rates per household across the lifecycle groups. Once again there is a consistent pattern of overestimation by the simulation for Dallas. Although almost all rates are significantly different at the 99% confidence level, numerical differences are generally small. This is the case for home-work trips, where the simulated trip rates are generally only a fraction higher than actual rates. Again, the differences are attributable almost entirely to differences in the mobility rates of the two data sets.

There are a substantial number of significant differences between the simulation and survey data for Salt Lake City also. However, in this case, there are some large numeric differences that help to pinpoint where some of the major differences arise. It appears that the cross-tabulation with life cycle identifies real differences in the data that are not accounted for in the simulation.

Table 24: Comparison of Trip Rates per Household by Household Lifecycle

Trip Purpose	Data Source	Household Lifecycle						
		1	2	3	4	5	6	7
Home-Work	Dallas HTS	1.31	1.09	2.45	2.01	2.25	0.13	0.24
	Simulated Data	1.31	1.29**	2.60**	2.12**	2.38**	0.00**	0.00**
	Salt Lake HTS	1.27	1.02	2.21	1.89	2.26	0.12	0.04
	Simulated Data	1.34**	1.44**	2.64**	2.01**	2.48**	0.00**	0.03**
Home-School	Dallas HTS	0.00	1.62	0.00	0.00	2.02	0.00	2.10
	Simulated Data	0.00	1.74**	0.00	0.00	2.38**	0.00	1.54**
	Salt Lake HTS	0.00	1.88	0.07	0.01	3.68	0.00	2.02
	Simulated Data	0.00	2.10**	0.00**	0.00**	3.01**	0.00	1.81**
Home-College	Dallas HTS	0.06	0.07	0.23	0.18	0.25	0.06	0.37
	Simulated Data	0.13**	0.10**	0.22	0.18	0.17**	0.08*	0.12**
	Salt Lake HTS	0.05	0.12	0.51	0.28	0.30	0.08	0.42
	Simulated Data	0.07*	0.09**	0.45	0.14**	0.22**	0.10	0.08**
Home-Shop	Dallas HTS	0.29	0.51	0.68	0.58	0.74	0.86	0.76
	Simulated Data	0.55**	1.22**	1.13**	1.07**	1.74**	1.09**	1.10**
	Salt Lake HTS	0.41	0.80	1.15	1.06	1.90	0.99	1.46
	Simulated Data	0.50**	1.33**	1.20	1.15*	1.95	1.20**	1.91**
Home-Other	Dallas HTS	0.92	2.86	2.42	2.50	5.59	2.48	3.15
	Simulated Data	1.16**	4.00**	2.60**	3.25**	6.31**	2.31**	3.61**
	Salt Lake HTS	1.47	3.67	3.61	3.80	8.95	3.04	5.34
	Simulated Data	1.12**	4.59**	2.92**	2.89**	7.27**	2.57**	5.46
Work-Other	Dallas HTS	0.99	1.10	1.40	1.48	1.68	0.10	0.05
	Simulated Data	0.98	1.00**	1.69**	1.86	1.91**	0.00	0.00
	Salt Lake HTS	1.13	1.38	1.61	1.38	1.80	0.07	0.04
	Simulated Data	1.08	1.18**	1.68	1.69**	1.87**	0.00	0.00

Other-Other	Dallas HTS	0.55	1.69	1.16	1.00	2.15	1.13	2.02
	Simulated Data	0.74**	2.34**	1.58**	2.07**	3.33**	1.58**	2.46**
	Salt Lake HTS	0.95	2.26	2.27	2.23	4.37	1.75	3.32
	Simulated Data	0.67**	1.93**	1.79**	1.77**	3.75**	1.47**	2.46**
All Purposes	Dallas HTS	4.13	8.93	8.33	7.75	14.69	4.74	8.68
	Simulated Data	4.87**	11.70**	9.81**	10.55**	18.22**	5.06**	8.83
	Salt Lake HTS	5.29	11.12	11.42	10.64	23.26	6.06	12.64
	Simulated Data	4.79**	12.66**	10.67**	9.65**	20.56**	5.34**	11.73**

** Statistically significant difference in trip rates at the 95 percent confidence level*

**Statistically significant difference in trip rates at the 99 percent confidence level

Overall, the results of the comparisons of trip rates by the different demographic groupings suggest that simulation for Salt Lake City is not as effective as was simulation for Baton Rouge for the trip rates. In general, simulation has worked fairly well for school and college trips and for overall trip rates. However, there is a consistent overestimation of work and shop trips and underestimation of home-other and other-other trips. For Dallas, the major problem lies in the difference in non-mobility rates. There would appear to be grounds to either investigate potential data problems in the Salt Lake City data or a necessity for local updating of the distributions to take into account specific differences in Salt Lake City and Dallas compared to national data. There is also a need to develop a method to deal with differing levels of non-mobility.

Mode-Share Comparisons

Table 25 compares the mode shares by purpose and mode between the HTS and simulation data for Dallas and Salt Lake City. It also shows the simulation results for Baton Rouge and the significance of differences between the HTS and simulation and between the Baton Rouge and the two new simulations. For this and subsequent tables, the difference in mobility rates between the surveys no longer has any effect because the comparisons deal only with those who make trips.

Table 25: Comparisons of Simulated Data by Mode and Purpose

Trip Purpose	Mode	Dallas		Salt Lake		Baton Rouge		
		HTS	Simulation	HTS	Simulation	Simulation	Vs. Dallas	Vs. Salt Lake
Home-Work	Auto Driver	88.8%	88.8%	82.9%	89.9%**	89.0%		
	Auto Pass.	5.7%	7.4%**	9.8%	7.4%**	7.5%		
	Transit	3.9%	2.2%**	2.3%	1.2%**	1.8%		

	Bike/Walk	1.6%	1.6%	4.2%	1.4%**	1.7%		
Home-School	Auto Driver	5.0%	3.8%*	7.0%	5.4%**	4.3%		
	Auto Pass.	49.2%	39.0%**	33.1%	36.4%**	37.6%		
	Transit	25.9%	49.2%**	25.1%	49.7%**	49.1%		
	Bike/Walk	20.0%	8.1%**	34.2%	8.5%**	9.0%		
Home-	Auto Driver	69.3%	71.5%	65.8%	68.5%	73.0%		
College	Auto Pass.	16.2%	11.0%**	13.2%	8.2%**	9.2%		
	Transit	7.2%	7.8%	4.9%	7.8%*	6.3%		
	Bike/Walk	7.3%	9.6%	14.3%	14.9%	11.2%		
Home-Shop	Auto Driver	77.7%	73.0%**	68.5%	72.3%**	71.6%		
	Auto Pass.	17.5%	21.7%**	26.0%	22.6%**	22.6%		
	Transit	0.7%	1.4%**	0.7%	1.2%*	1.6%		
	Bike/Walk	4.1%	3.9%	4.3%	3.8%	4.2%		
Home-Other	Auto Driver	69.5%	62.4%**	63.9%	59.4%**	60.6%		
	Auto Pass.	25.8%	30.7%**	26.1%	33.6%**	31.8%		*
	Transit	0.5%	1.9%**	0.8%	1.4%**	2.0%		*
	Bike/Walk	4.1%	5.1%**	8.4%	5.5%**	5.5%		
Other-Work	Auto Driver	84.0%	85.1%	82.8%	85.1%**	84.4%		
	Auto Pass.	9.7%	8.9%	9.5%	9.7%	10.1%		
	Transit	0.7%	1.5%**	0.9%	1.1%	1.3%		
	Bike/Walk	5.5%	4.5%*	5.5%	4.0%**	4.2%		
Other-Other	Auto Driver	68.3%	63.2%**	64.0%	60.2%**	63.0%		*
	Auto Pass.	26.0%	29.8%**	28.5%	33.0%**	30.1%		*
	Transit	1.8%	1.7%	1.8%	1.5%	1.8%		
	Bike/Walk	3.9%	5.3%**	5.0%	5.3%	5.1%		
All Purposes	Auto Driver	71.5%	68.2%**	63.4%	63.8%	66.2%	**	**
	Auto Pass.	20.3%	22.5%**	23.3%	25.5%**	23.8%	**	**
	Transit	3.3%	4.7%**	3.5%	5.7%**	5.1%		*
	Bike/Walk	4.9%	4.5%**	9.0%	5.0%**	4.8%		

* Statistically significant difference in trip rates at the 95 percent confidence level
**Statistically significant difference in trip rates at the 99 percent confidence level

Overall, the mode shares are simulated quite well for Dallas with the differences between actual shares and the simulation only being on the order of one or two percent of trips. The simulation shows more auto passenger and transit and less bike/walk and auto driver than the actual data. However, almost every mode share is significantly different between the simulated data and the survey data for Salt Lake City. It can also be seen in Table 25 that the Baton Rouge simulation is not statistically different from the Salt Lake City simulation except for home-other auto passenger and transit, other-other auto driver and auto passenger, and auto driver and passenger, and transit for all purposes combined. Numerically, however, many of the Baton Rouge values are different than the Salt Lake City simulation.

Disaggregate analysis of the mode shares by different segments of the population helps demonstrate how well or how poorly the simulation has worked for the each region. As an example, analysis of the data by life-cycle groups (Table 26) is also useful to examine because life cycle is the main categorization variable. For Dallas, in life-cycle groups 1 and 2, POV use is over-predicted by the simulation data, and transit and bike/walk are under-predicted. In contrast, in life-cycle groups 4, 5, and 6, POV shares are under-estimated while transit is over-predicted. Otherwise, the mode shares are fairly well simulated.

Table 26: Mode-Share Comparisons by Household Lifecycle

	Life-Cycle Grouping	Data Source	Mode Shares (Percent)			
			Driver	Passenger	Transit	Bike/Walk
1	1 person, 1 worker	Dallas HTS	92.5	3.0	4.4	0.2
		Simulated Data	93.7*	2.0*	4.2	0.0
		Salt Lake HTS	89.7	2.8	6.9	0.6
		Simulated Data	94.1**	1.9	3.9**	0.0**
2	Single working parent	Dallas HTS	79.3	11.1	9.5	0.1
		Simulated Data	84.4**	8.9*	6.8**	0.0
		Salt Lake HTS	82.0	4.3	12.7	1.0
		Simulated Data	81.0	11.7**	7.2**	0.0**
3	Multiple adults, 1+workers, 0 children	Dallas HTS	96.0	1.1	2.7	0.2
		Simulated Data	96.0	1.1	2.8	0.1**
		Salt Lake HTS	94.3	0.9	4.2	0.6
		Simulated Data	94.9	1.2*	3.8	0.1**
4	Multiple adults, 1+ workers, 1+children 0-4	Dallas HTS	94.8	1.2	3.9	0.1
		Simulated Data	92.2**	3.3**	4.4	0.1
		Salt Lake HTS	97.6	2.4	0.0	0.0
		Simulated Data	94.8	1.6	3.6**	0.1
5	Multiple adults, 1+ workers, 1 +School age child	Dallas HTS	89.1	4.7	6.1	0.1
		Simulated Data	86.1**	8.3**	5.6	0.0**
		Salt Lake HTS	81.2	5.2	12.7	1.0
		Simulated Data	86.3**	8.3**	5.4**	0.0**
6	1+ adults, no workers, no children	Dallas HTS	94.8	1.1	3.9	0.3
		Simulated Data	92.7**	2.0**	5.2*	0.0*
		Salt Lake HTS	92.3	2.5	5.0	0.1
		Simulated Data	93.8*	1.6**	4.6	0.0
7	1 adult, no workers,	Dallas HTS	69.7	10.4	19.9	0.0

1+ children	Simulated Data	75.4	15.2	9.4**	0.0
	Salt Lake HTS	76.0	10.6	11.8	1.6
	Simulated Data	78.7	11.7	9.6	0.0**

For Salt Lake City, the fewest significant differences occur for lifecycle groups 4 and 7 and then 3 and 6. The other groups all have three significant differences out of the four possible ones, or all four in the case of lifecycle group 5. For the most part, when the differences are significant, the simulation overestimates transit and underestimates bike/walk shares while the problem with the "other" mode continues as before.

Departure-Time Comparisons

Table 27 shows the departure time comparisons for the Dallas and Salt Lake City HTS data and the simulations. The Kolmogorov-Smirnov statistic provides a test to determine if the distribution of trips over the day is significantly different between the HTS and simulated data.

For Dallas, only in the cases of the home-other and all purpose trips is there a significant difference in the diurnal distributions of the trips. Numerical differences are generally small even for the home-other trips where the primary difference appears to occur in the 6 - 9 a.m. and 9a.m.- 4 p.m. categories, which appear to be the source of the differences. This is also reflected in the all purposes category.

For Salt Lake City, home-school and home-college show the largest numerical differences in fractions by time period. This probably reflects local conditions with respect to the time at which classes begin, which may be different in the Salt Lake City region from the national average. From the percentages shown, it appears that both schools and colleges tend to start later in Salt Lake City then on average across the nation. There also appears to be a shift towards more work trips and more shopping and other-other trips in Salt Lake City being made later in the day. Home-based other trips tend to be made earlier in the day in Salt Lake City. There are almost no significant differences in the other trip purposes.

Table 27: Comparisons of HTS and Simulated Departure Times by Trip Purpose

Trip Purpose	Departure Time Period	Dallas		Salt Lake	
		HTS (% of Trips)	Simulated (% of Trips)	HTS (% of Trips)	Simulated (% of Trips)
Home-Work	6 a.m. – 9 a.m.	36.7%	35.2%	34.7%	36.3%
	9 a.m. – 4 p.m.	22.2%	25.4%	25.7%	24.2%
	4 p.m. – 7 p.m.	27.2%	24.5%	25.3%	24.5%
	7 p.m. – 6 a.m.	14.0%	14.8%	14.3%	14.9%
Home-School	6 a.m. – 9 a.m.	51.6%	50.2%	45.7%	51.5%**
	9 a.m. – 4 p.m.	42.7%	44.1%	50.1%	42.2%**
	4 p.m. – 7 p.m.	4.8%	5.1%	3.0%	5.6%**
	7 p.m. – 6 a.m.	0.9%	0.7%	1.1%	0.7%**
Home-College	6 a.m. – 9 a.m.	36.0%	30.6%	27.0%	33.3%**
	9 a.m. – 4 p.m.	41.2%	46.3%	41.7%	44.8%**
	4 p.m. – 7 p.m.	11.8%	14.2%	16.8%	13.8%**
	7 p.m. – 6 a.m.	10.9%	8.9%	14.5%	8.1%**
Home-Shop	6 a.m. – 9 a.m.	4.2%	5.7%	2.9%	6.1%**
	9 a.m. – 4 p.m.	44.1%	45.6%	42.4%	48.0%**
	4 p.m. – 7 p.m.	32.3%	27.5%	33.3%	26.1%**
	7 p.m. – 6 a.m.	19.5%	21.1%	21.4%	19.9%**
Home-Other	6 a.m. – 9 a.m.	17.2%	12.0%**	12.3%	11.7%**
	9 a.m. – 4 p.m.	31.0%	35.9%**	32.9%	37.9%**
	4 p.m. – 7 p.m.	28.9%	29.3%**	29.5%	28.2%**
	7 p.m. – 6 a.m.	22.8%	22.8%**	25.3%	22.1%**
Other-Work	6 a.m. – 9 a.m.	14.4%	13.9%	13.9%	13.4%
	9 a.m. – 4 p.m.	64.5%	63.9%	63.1%	64.8%
	4 p.m. – 7 p.m.	18.2%	17.9%	19.7%	17.9%
	7 p.m. – 6 a.m.	2.9%	4.3%	3.3%	3.9%
Other-Other	6 a.m. – 9 a.m.	9.4%	7.0%	6.0%	7.1%
	9 a.m. – 4 p.m.	53.5%	56.6%	57.8%	56.0%
	4 p.m. – 7 p.m.	22.9%	21.2%	22.4%	20.9%
	7 p.m. – 6 a.m.	14.2%	15.3%	13.8%	16.0%
All Purposes	6 a.m. – 9 a.m.	21.4%	17.4%**	16.4%	17.9%**
	9 a.m. – 4 p.m.	39.3%	43.2%**	42.6%	43.8%**
	4 p.m. – 7 p.m.	24.4%	23.5%**	24.3%	22.7%**
	7 p.m. – 6 a.m.	14.9%	15.8%**	16.7%	15.6%**

Trip-Length Comparisons

Table 28 shows the trip-length statistics from each of the Dallas and Salt Lake City data and the simulation of each region's households. The simulation consistently underestimates trip lengths for all trip purposes in Dallas. Similarly, for Salt Lake City, all purposes except other-work are significantly different between the simulation and the actual data. It appears that the simulation does not pick up city size factors, as would be expected, given the attributes used to categorize households for this simulation.

Table 28: Comparison of HTS and Simulated Vehicle Trip Lengths (minutes)

Trip Purpose	Dallas		Salt Lake	
	Mean	Simulated	Mean	Simulated
Home-Work	29.29	19.59**	18.60	19.90**
Home-School	20.24	16.56**	13.15	16.52**
Home-College	24.18	18.15**	16.31	19.07**
Home-Shop	14.52	11.63**	10.77	11.64**
Home-Other	17.15	13.92**	12.44	13.63**
Other-Work	19.80	15.12**	14.09	15.03
Other-Other	16.21	14.25**	12.20	13.87**
All Purposes	20.05	15.14**	13.31	14.89**

CONCLUSIONS AND RECOMMENDATIONS

This research proposes a new approach designed to assist practitioners in preparing travel-forecasts for their region. Rather than focusing on aggregate relationships and models, the idea is to synthetically derive the data that drives these models by combining local sociodemographic information with simulated travel data. The appeal of the approach lies in its low-cost, relative ease of use, the fact that the data required are freely available (NPTS and PUMS90), and the potential advantages it offers over the use of borrowed models that incorporate no local demographic element directly in their estimation process. It also provides regions with the capability of specifying and estimating their own models rather than being tied to structures employed in transferred relationships. The approach offers a flexibility that will enable regions to update their travel data base and modeling efforts regularly as new demographic and travel data sources become available such as the 2000 waves of the PUMS and the NPTS. Finally, the approach also offers the potential to increase sample sizes in sub-regions and corridors, where standard HTS data may be too sparse to allow detailed study.

The results presented here suggest that this concept of creating simulated HTS data is worth pursuing further. The procedure was able to provide trip rates that were generally comparable to actual data. The other salient trip characteristics (mode, departure time, and reported trip length) were less effectively replicated in the simulation procedure. These discrepancies were attributed to contextual differences between regions (e.g., population size, transit service), which intuitively affect these attributes to a greater extent than trip rates. Clearly, further research is needed to determine how local characteristics can be incorporated in the simulation process to capture these unexplained differences.

Several issues affected the simulation. First, the approach is based on the premise that persons of similar demographic characteristics living in similar environments display similar travel behavior. If this is so, then the problem should simply be to define these characteristics correctly. However, the goodness-of-fit indicators for the segmentation schemes suggest a limited ability to explain the variation in certain types of trips (particularly shop and other) with the characteristics available. These differences must be attributed to other factors that are not captured in conventional demographic measures. A further complication is added by how the dependent measures are defined.

A second issue that affected the simulation was the use of the NPTS as the source for the simulated travel data. While it proved a particularly “clean” data set, the collection methodology was somewhat different from the typical travel survey providing data for a national cross-section for every day of the year. More significantly, however, use of this data source assumes that nationally-derived relationships were maintained at a local level. This appears to be only partially true and suggests that some local information must be employed to ensure the simulated data displays sensitivity to a particular locale. Another possibility might be to apply the same logic as model transfer and use a travel survey from a region of similar characteristics (if available) or an amalgam of several travel surveys as the source of the simulated travel data.

Another issue that affected the simulation was the decision to work at the household level. For trip rates, it is arguably more appropriate to work at a household level because of the effect of interactions among household members. However, for mode, departure time, and trip length, it could be argued one should work at an individual level, although some household-level variables are still needed. Comparisons between the simulation results using households and individuals is recommended as a potential future research topic. Finally, the issue of non-mobility rates needs to be addressed for future applications and also needs to be considered in relation to any form of updating from local data.

While the concept of simulated HTS data is a seemingly valid one, further research is needed. The following recommendations are made for future research activities:

- Validation of the results should include comparisons of the traditional outputs of the travel-forecasting process such as link volumes.
- Updating the simulated data is clearly critical particularly for the simulation of mode, departure time, and trip length. Future work should assess the various options available for updating including use of aggregate (e.g., mode shares) as well as disaggregate (individual/household-level) data.
- The procedures could conceivably be used to generate future HTS data sets. This would require the prediction of future household demographics, possibly using a microsimulation approach to forecast the characteristics of each person in the sample in the future year (e.g., Chung and Goulias, 1997). There is no apparent reason why the procedure could not be applied to simulate travel data for an entire regional population.
- In this application, travel data only are simulated. One could conceivably simulate the locations of households and trip-ends as is currently being attempted in the TRANSIMS project (TMIP, 1995).
- The simulation procedure produces a set of trips and their associated attributes that could have been collected in a household travel survey. However, no continuity is maintained between individual trip records. To achieve this continuity, one would need to incorporate the constraints imposed by the previous trip into the simulation. For instance, if the individual took five minutes to get to work, they should take approximately five minutes to return home. This could be achieved by tour-based simulation rather than trip-based.
- The approaches presented here were intended for application using the PUMS and the NPTS. However, in the future, another possibility for updating the demographic database will come from the American Community Survey (ACS). The ACS will provide PUMS-like data for three million households per year sampled from across the nation. This could provide a means to update the database annually rather than every ten years.