25th Australasian Transport Research Forum
Incorporating the BTRE Transport Policy Colloquium
Canberra 2-4 October 2002

Simulating household travel survey data in Australia: Adelaide case study

Peter Stopher, Philip Bullock and John Rose
The Institute of Transport Studies

Abstract

A method has been developed to synthesize household travel survey data from a combination of Census and national transport survey data sources. The procedure, described in other papers, involves creating distributions of pertinent variables (numbers of trips by purpose, mode of travel, time of day of travel, and trip length) that can be used to estimate travel-demand models. A sample of local residents is then drawn from disaggregate census data, providing detailed information on the socioeconomic characteristics of the sample. Using these socioeconomic characteristics, travel data are simulated from the transport data distributions using Monte Carlo simulation. This procedure was developed in the United States in the past four years.

The paper describes the application of this procedure to Adelaide, South Australia, for which an actual household travel survey exists from 1999. The paper describes results obtained from applying the generic data as the basis of the simulation. Results are compared between the synthetic and real data to determine the closeness of the match between the data sets. The procedure uses data derived from a nationwide travel survey in the U.S., but uses census data for Adelaide from the 1996 ABS Census, using the one percent sample. The purpose of this research was to determine the extent to which the trip characteristics distributions from the U.S. could be used in Australia. It is concluded that the procedure performs about as well as the process was shown to perform in Dallas, Salt Lake City, and Baton Rouge in the U.S.

This process holds out considerable promise as a means to increase available samples for local and corridor planning, as well as to provide data for regions that have typically not been able to undertake household travel surveys on the scale of those being conducted in the Melbourne and Sydney regions.

Contact author
Peter Stopher, Professor of Transport Planning,
The Institute of Transport Studies, The University of Sydney, NSW 2006.
Phone: (02) 9351 0010   Fax: (02) 9351 0088
e-mail: peters@its.usyd.edu.au

GO BACK TO START SCREEN   GO BACK TO INDEX

Introduction

Household travel surveys (HTSs) are expensive to undertake, costing up to $350 per completed household in Australia, depending on the data collection technique. Given the large sample sizes (a minimum of 2,500 to 3,000 households) required to capture representativeness of existing household types within both large and small metropolitan areas, the costs of HTSs soon become prohibitive to all but large well-funded government organisations. Despite the absolute and relative costs, HTSs are critical to travel-demand forecasts, used to estimate the impact of proposed transport policies, and to defining regional travel patterns. HTSs involve collecting socio-demographic data along with associated travel or activity data for each of the surveyed households that are then used to derive travel patterns for the survey region.

As with any survey method, public resistance to undertaking the survey task can result in high non-response rates and associated non-response errors. This may be compounded if the survey is undertaken using a data collection method that has been shown to impact response rates. Also, increasing personal mobility makes it more difficult to find individuals at their place of residence, where they can undertake the survey at a convenient time.

The research reported here represents a continuation of work undertaken by Greaves, Stopher, and Bullock (Greaves, 1998; Greaves, 2000, Greaves and Stopher, 2000, Stopher *et al.*, 2001) on the use of Monte Carlo simulation to create simulated HTS data in the U.S.A. This research applies the methodology used by Greaves *et al.* to a metropolitan area in Australia. The method uses a sample of real households for which there is known socio-demographic data, and then simulates key characteristics of travel patterns from distributions of these characteristics obtained from a national HTS.

A sample of households, with socio-demographic descriptors, was generated from the 1996 Census Household Sample File (HSF), derived from the 1996 Australian Bureau of Statistics (ABS) Census Data for the Adelaide region. Distributions of key variables influencing travel behaviour were derived from US data mainly because no sufficiently recent national HTS exists in Australia[1]. The simulated data were then compared to the 1999 Metropolitan Adelaide Household Travel Survey (AHTS) to determine the ability of simulated data to simulate the actual travel patterns observed in the AHTS survey.

Data Preparation

The Metropolitan Adelaide Household Travel Survey (AHTS) conducted in 1999 by Transport South Australia of the Department for Transport, Urban Planning and the Arts sampled 5,886 households with 14,004 associated individuals out

[1] Apparently, there was a national survey in 1985-6, but, even if the data from that survey were available and included the needed variables, the data are too old to be useful for this process.

of a then total population of 1,045,854 individuals. The survey collected demographic and socio-economic data for each household as well as information on travel activities undertaken over a two day period.

These data were collected using several surveys of each household which required cleaning and combining for use. Several categorical variables were created for the analysis including a lifestyle variable consisting of the number of workers and children within a household. After cleaning the data and removing households that had not completed all of the surveys, 5,615 households remained. After removing weekend data, because the other data sets to which comparisons are made are weekday only data, 4,774 households remained.

The 1996 Census Household Sample File (HSF) represents a 1% sample from the 1996 census household data collected by the Australian Bureau of Statistics (ABS). Data covering the same region of Adelaide were selected for this project. The data were in three files covering private dwellings, their associated household and individual records as well as a 1% sample of persons from non-private dwellings. These were combined so that the final data file format was similar to the AHTS data file. Significant cleaning of the data was required.

Table 1 shows comparative statistics for the AHTS and the HSF for the city of Adelaide. For completeness, the table also includes the comparative statistics for the data collected for an earlier study using the 1995 U.S. Nationwide Personal Transportation Survey (NPTS) as well as HTSs undertaken in Baton Rouge (Greaves, 2000) Dallas, and Salt Lake (Stopher *et al.*, 2001). In addition, because of differences between the AHTS and HSF, comparative statistics are included from the ABS for Adelaide, so as to determine if the Adelaide HSF represents the entire population reasonably well.

Table 1 Summary Statistics from the Household Travel Surveys

Statistic	**AHTS**	**HSF**	**ABS Adel.**	**NPTS**	**Baton Rouge**	**Salt Lake**	**Dallas**
Average Age	38	35	35	35	33	34	35
Average Household Size	2.5	2.44	2.5	2.63	2.71	3.14	2.47
Percent in Single Family Dwellings	N/A	N/A	69%	74%	75%	73%	78%
Percent from Non-Car-Owning Households	7%	13.2%	12.5%	8%	8%	4%	5%
Average Vehicles per Household	1.61	1.43	1.44	1.73	1.78	1.97	1.84
Percent Females in Sample	51%	50.9%	51.4%	51%	52%	53%	52%
Percent Home Owners	76.5%	59.5%	68.6%	64%	66%	76%	68%
Average Workers per Household	1.08	1.03	1.08	1.33	1.34	1.31	1.40
Average Vehicles per Worker	1.49	1.39	1.33	1.30	1.33	1.50	1.31

In comparing the Australian data sources, it appears that possible sampling problems exist between the AHTS and the census. In comparison to the HSF and the full census data, the AHTS appears to have under sampled the number of non-car-owning households, whilst over sampling the percent of home owners. For all other variables, the statistics shown appear to be similar across data sets. As a result of the under-sampling of non-car owning households, average cars per household appears to be too high. There is, of course, a three-year difference between the census results and the AHTS (1996 versus 1999), but it seems unlikely that non-car-owning households would have decreased by 5% in that time. However, the over-sampling of home owners would be consistent with a higher level of car ownership, which would confirm a sampling bias in the AHTS.

It can be seen that the number of household workers in both Australian data sets is well below the average represented across the American data. Further, it is evident that the proportion of non-car-owning households in the Australian data is greater than the U.S. cities, whilst the average number of household vehicles is below the U.S. cities. This suggests deviations away from the trip distributions derived in the U.S. towards a greater use of public transport and a lower number of home-work trips. It is worth noting that the car ownership per worker, however, is very similar between the U.S. and Adelaide, with the exception of Salt Lake, which exhibits a significantly higher rate (comparable to the possibly incorrect figure from the AHTS).

Simulating Travel Survey Data

A household sample was drawn from the HSF data, using a procedure similar to that described in Stopher *et al.* (2001), replicating as closely as possible the AHTS sample, controlled on household size and household vehicle availability. This sample was then compared to the actual survey sample from the AHTS, as shown in Table 2. In terms of general household characteristics, the HSF sample compared favourably to the AHTS sample, although the HSF exhibits about a 6% higher average number of workers per household than the AHTS.

Table 2 Comparison of Statistics of the HTS and HSF Samples

Statistic	AHTS	HSF
Sample Size (Households)	4774	4774
Average Vehicles per Household	1.63	1.61
Average Household Size	2.5	2.52
Average Workers per Household	1.08	1.14

The goal of this research is to determine whether or not acceptable data can be produced through a simulation process, particularly one that is based on distributions of travel characteristics obtained from a U.S. nationwide survey. The primary question to be addressed is that of defining what constitutes acceptability. In the previous work of Greaves and Stopher, the primary measure of acceptability used was statistical significance of differences between the simulated data and a real household travel survey. However, statistically significant differences can be found when numerical differences are very small. The purpose of the simulation is to produce numbers that are reasonably close to those that would have been obtained from an actual survey.

More importantly, the data need to be usable to develop travel-demand models, and to estimate population statistics that would have been obtained from an actual survey. Thus. these are more at issue than simply statistical significance in the difference of statistics from the simulated data and an actual survey. Statistical significance of differences is not a necessary or sufficient measure of acceptability. Furthermore, a household travel survey is an imperfect measure of population behaviour, so that it also has to be assumed that differences between simulation and an actual small sample survey could be a result of errors in the survey. The research reported here has not yet included using the simulated data for modelling and comparing this to identically-specified models from the actual survey data. Rather, we report first on the similarities and dissimilarities in the performance of the simulation data compared to real data, and how this performance compares to that found in three U.S. applications. These comparisons indicate whether or not it is feasible to use the U.S. distributions for simulating Australian travel characteristics. Second, we report on some regional statistics derived from each of the simulated and real data in Adelaide, and comment on these figures.

Simulation of Trip Rates

Table 3 compares the trip rates for several trip purposes across the AHTS and synthetic data using the Z test for equal population means. As can be seen from the table, home-based work, home-based school, home-based college, home-based other and other-work trips were significantly different at the 99% level between the two data sets. The total number of trips was also significantly different at the 99% level. With the exception of home-based other trips, the simulation overestimated the trip rates. However, comparing the AHTS trip rates with the Sydney HTS trip rates, shown in Table 3, it appears that the AHTS shows very significantly lower trip rates than the Sydney HTS. In general, the Sydney trip rates are similar to or higher than the simulation figures.

Table 3 Comparisons of Person Trip Rates per Household

Purpose	AHTS		Simulation		Sydney HTS
	Mean	Std. Dev.	Mean	Std. Dev.	
Home-Based Work	1.08	1.56	1.52**	1.79	1.62
Home-Based School	0.34	1.04	0.59**	1.35	0.82
Home-Based College	0.05	0.34	0.18**	0.68	
Home-Based Shop	0.95	1.52	1.30**	1.74	1.82
Home-Based Other	3.14	3.93	3.41**	3.62	N/A
Other-Work	1.16	2.55	1.10	1.94	0.98
Other-Other	1.58	2.91	2.02**	3.00	N/A
TOTAL TRIPS	8.32	8.16	10.13**	7.49	10.35

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

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

One possible explanation for the over-estimation of trip rates by the simulation is the under-estimation of the number of households not making any trips on the survey day. For the AHTS survey, 492 households or 8.8% of the total sample, made no trips on the day of the survey whilst the simulation included a relatively smaller 180 households or 3.2% of the total sample. Correcting this by computing trip rates only for mobile households (i.e., households making trips on the survey day) produces the results shown in Table 4. Most trip rates have moved closer to each other. Home-based other trips have shifted from being overestimated by the simulation to underestimated, while other-work are more substantially underestimated than before. The overall trip rate is no longer significantly different and is much closer than before. This might suggest that the Adelaide survey classified more trips as being home-based other, possibly because of problems in determining the real purpose, with a resulting underestimate of the work, school, college, and shopping trips. This result bears further investigation.

Table 4 Comparisons of Person Trip Rates per Mobile Household

Purpose	AHTS		Simulation	
	Mean	Std. Dev.	Mean	Std. Dev.
Home-Based Work	1.34	1.63	1.56**	1.79
Home-Based School	0.42	1.15	0.61**	1.37
Home-Based College	0.07	0.38	0.19**	0.69
Home-Based Shop	1.18	1.61	1.33**	1.75
Home-Based Other	3.89	4.03	3.49**	3.63
Other-Work	1.44	2.77	1.13**	1.96
Other-Other	1.96	3.13	2.07	3.02
TOTAL TRIPS	10.30	7.87	10.37**	7.41

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

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

Table 5 compares the above findings to that of Stopher, Greaves and Bullock (2001). Given that the distributions of trips used in each synthetic study were derived from the NPTS, differences between the synthetic data generated for different localities demonstrates that the simulation is responsive to differences in the local population characteristic. Table 5 shows that significant differences exist between Adelaide and Salt Lake City for all trip purposes at the 99% level. Significant differences between Adelaide and both Baton Rouge and Dallas do not exist for home-based college and other-other trips, while significant differences do exist for at least four of the trip purposes.

Table 5 also shows the significant differences between simulation and the actual HTS data for each of Baton Rouge, Dallas, and Salt Lake. Baton Rouge had only one significant difference, which was at 95% not 99%. Most purposes for Dallas and Salt Lake show significant differences. One of the primary differences is that the Baton Rouge survey was conducted using identical instruments and protocols to the NPTS survey from which the distributions have been derived, while Salt Lake and Dallas used markedly different survey procedures. This may suggest that there is an issue here in the methods used for the surveys that are being used to benchmark the simulations. Also, the number and magnitude of significant differences for Dallas, Salt Lake, and Adelaide between the simulated and actual HTS data for each region are about the same. On these results, which are uncorrected for differences in mobility rates, Dallas has two trip rates that are not significantly different, while Adelaide and Salt Lake each have one, although the latter also has two trip rates that are significantly different at 95% but not at 99%. All three regions show significantly different overall trip rates than their respective HTSs.

Table 5 Comparisons of Simulated Trip Rates per Household across Data Sets

Purpose	Adelaide	Baton Rouge	Dallas	Salt Lake	Adelaide		
	Simulated Mean	Simulated Mean	Simulated Mean	Simulated Mean	Diff. from Baton Rouge	Diff. from Dallas	Diff. from Salt Lake
Home-Based Work	1.52**	1.83	1.86**	1.83**	**	**	**
Home-Based School	0.59**	0.74	0.60	1.07*	**		**
Home-Based College	0.18**	0.17	0.16	0.23*			**
Home-Based Shop	1.30**	1.32*	1.14**	1.38**		**	*
Home-Based Other	3.41**	3.69	3.19**	4.17**	*	**	**
Other-Work	1.10	1.34	1.35**	1.33	**	**	**
Other-Other	2.02**	2.02	1.86**	2.26**		*	**
TOTAL TRIPS	10.13**	11.11	10.17**	12.28**	**		**

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

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

The other important result in Table 5 is that there are significant differences between the simulations for all four locations: Adelaide, Baton Rouge, Dallas, and Salt Lake, indicating that the simulation procedure takes account of some of the specific differences between different cities and does not produce statistically similar trip rates for each purpose, irrespective of the locality.

Because of the difference in household size between Adelaide and the U.S. cities, Table 6 shows a comparison of the per person trip rates, similar to the household trip rates in Table 5. Table 6 shows that the Adelaide trip rates per person are more comparable to the U.S. cities. All Adelaide trip rates fall within the range represented by the U.S. cities, although most are still significantly different from the other cities.

Table 6 Comparisons of Simulated Trip Rates per Person across Data Sets

Purpose	Adelaide	Baton Rouge	Dallas	Salt Lake
	Simulated Mean	Simulated Mean	Simulated Mean	Simulated Mean
Home-Based Work	0.60**	0.68	0.59**	0.74**
Home-Based School	0.23**	0.27	0.19	0.43*
Home-Based College	0.07**	0.06	0.05	0.09*
Home-Based Shop	0.52**	0.49*	0.36**	0.56**
Home-Based Other	1.35**	1.36	1.02**	1.69**
Other-Work	0.44	0.49	0.43**	0.54
Other-Other	0.80**	0.75	0.59**	0.91**
TOTAL TRIPS	4.02**	4.10	3.24**	4.97**

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

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

Simulation of Mode Shares

Table 7 shows a comparison of the mode shares for the Adelaide HTS and simulation data as well as comparing the Adelaide simulation mode shares with those of the simulation data from Dallas and Salt Lake City. Significant differences are also shown. Baton Rouge was not compared due to the non-availability of data. As can be seen for the Adelaide study, significant differences exist for nearly every mode share, the exceptions being for home-school drivers, home-college transit and bike/walk, and home other car passenger (significant at 95%, not 99%).

Examination of the differences between both the US simulation data and the Adelaide simulation data show few significant differences, until all purposes are combined. Salt Lake City shows a few more differences than Dallas.
The simulation consistently over-estimates car driver shares for all trip purposes and significantly under-estimates the volume of bike/walk trips. Transit and other trips are also relatively poorly simulated. This is possibly the result of US trip distributions not being representative of travel patterns in Australia. There is a lower level of reliance on car driver travel than is probably the case in the US.

Table 7 Comparisons of Simulated Data by Mode and Purpose

Trip Purpose	Mode	Adelaide HTS	Adelaide Simulation	Dallas Simulation	Salt Lake Simulation	Adelaide Simulation Diff From Dallas	Adelaide Simulation Diff From Salt Lake
Home-Work	Driver	85.3%	88.9%**	89.3%	89.92%	*	
	Passenger	7.3%	7.8%	6.9%	7.42%		
	Transit	3.8%	1.6%	1.7%	1.22%		
	Bike/Walk	3.6%	1.5%**	2.0%	1.43%		
Home-School	Driver	1.8%	3.1%**	4.0%	5.43%		**
	Passenger	66.9%	36.7%**	38.9%	36.40%	*	
	Transit	1.4%	51.3%**	48.3%	49.68%		
	Bike/Walk	29.9%	8.7%**	8.8%	8.48%		
Home-College	Driver	62.4%	73.3%**	74.9%	68.90%		
	Passenger	18.2%	10.9%**	11.5%	8.27%		
	Transit	5.8%	8.3%	6.8%	7.85%		
	Bike/Walk	13.6%	7.3%**	6.8%	14.98%	**	**
Home-Shop	Driver	65.6%	71.6%**	74.2%	72.33%	**	
	Passenger	16.1%	22.7%**	20.4%	22.65%		
	Transit	1.8%	1.1%**	1.2%	1.22%		
	Bike/Walk	16.4%	4.4%**	4.2%	3.80%		
Home-Other	Driver	50.9%	62.8%**	63.2%	59.40%		**
	Passenger	32.2%	30.4%**	30.1%	33.62%		**
	Transit	3.8%	1.5%**	1.7%	1.45%		
	Bike/Walk	13.2%	5.2%**	5.0%	5.54%		
Other-Work	Driver	72.5%	84.1%**	83.6%	85.12%		
	Passenger	6.4%	10.2%**	9.8%	9.72%		
	Transit	12.0%	1.4%**	1.6%	1.13%		
	Bike/Walk	9.0%	4.1%**	4.9%	4.04%		
Other-Other	Driver	48.4%	63.8%**	63.6%	60.26%		**
	Passenger	32.2%	29.6%**	29.5%	32.99%		**
	Transit	3.3%	1.6%**	1.6%	1.46%		
	Bike/Walk	16.0%	4.8%**	5.3%	5.28%	**	
All Purposes	Driver	57.7%	67.1%**	68.9%	63.80%	**	**
	Passenger	24.8%	23.7%**	22.1%	25.55%		**
	Transit	4.5%	4.5%	4.3%	5.70%		**
	Bike/Walk	12.9%	4.6%**	4.6%	4.95%	*	*

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

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

Departure time Comparisons

Table 8 shows the departure time comparisons for all trip purposes. The Kolmogorov-Smirnov D-value provides a test to determine if the distribution of trips over the day is significantly different between the HTS and simulated data.

Table 8 Comparisons of HTS and Simulated Departure Times by Trip Purpose

Purpose	Time Period	AHTS	Simulation	D Stat
Home-Based Work	6.01 a.m. - 9.00 a.m.	40.1%	36.1%**	*
	9.01 a.m. - 4.00 p.m.	23.8%	24.8%	*
	4.01 p.m. - 7.00 p.m.	24.8%	24.8%	*
	7.01 p.m. - 6.00 a.m.	11.3%	14.2%**	*
Home-Based	6.01 a.m. - 9.00 a.m.	53.7%	49.0%**	
School	9.01 a.m. - 4.00 p.m.	41.9%	44.7%	
	4.01 p.m. - 7.00 p.m.	4.2%	5.5%*	
	7.01 p.m. - 6.00 a.m.	0.2%	0.8%**	
Home-Based	6.01 a.m. - 9.00 a.m.	29.5%	33.0%	*
College	9.01 a.m. - 4.00 p.m.	39.5%	42.4%	*
	4.01 p.m. - 7.00 p.m.	19.8%	15.2%	*
	7.01 p.m. - 6.00 a.m.	11.2%	9.4%	*
Home-Based Shop	6.01 a.m. - 9.00 a.m.	8.4%	6.2%**	**
	9.01 a.m. - 4.00 p.m.	62.4%	53.5%**	**
	4.01 p.m. - 7.00 p.m.	22.8%	24.2%	**
	7.01 p.m. - 6.00 a.m.	6.4%	16.1%**	**
Home-Based Other	6.01 a.m. - 9.00 a.m.	15.8%	12.1%**	**
	9.01 a.m. - 4.00 p.m.	44.6%	41.1%**	**
	4.01 p.m. - 7.00 p.m.	24.4%	26.7%**	**
	7.01 p.m. - 6.00 a.m.	15.2%	20.1%**	**
Other-Work	6.01 a.m. - 9.00 a.m.	10.4%	13.3%**	
	9.01 a.m. - 4.00 p.m.	64.8%	65.6%	
	4.01 p.m. - 7.00 p.m.	20.2%	17.0%**	
	7.01 p.m. - 6.00 a.m.	4.7%	4.2%	
Other-Other	6.01 a.m. - 9.00 a.m.	8.1%	6.8%**	**
	9.01 a.m. - 4.00 p.m.	69.1%	59.8%**	**
	4.01 p.m. - 7.00 p.m.	15.7%	19.0%**	**
	7.01 p.m. - 6.00 a.m.	7.1%	14.5%**	**
TOTAL TRIPS	6.01 a.m. - 9.00 a.m.	17.5%	16.6%**	**
	9.01 a.m. - 4.00 p.m.	51.3%	46.9%**	**
	4.01 p.m. - 7.00 p.m.	21.2%	22.1%	**
	7.01 p.m. - 6.00 a.m.	10.05%	14.51%**	**

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

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

Table 8 shows that for all but home-based school trips and other-work trips significant differences are observed between the two over the day. However, the Z test between the individual fractions in each time period demonstrates slightly better results with 14 out of the 32 fractions showing statistically different results. Interestingly, the differences largely occur in the home-based shopping, home-based other and other-other trips suggesting that for all other categories Australians and Americans share similar travel time patterns. Clearly examination of the percentages suggest that residents of Adelaide prefer to, or are forced by necessity to shop between business hours whereas in America, a greater percentage of shopping activity probably occurs outside this period.

Trip-Length Comparisons

Table 9 shows the comparison for the Adelaide HTS and simulation trip lengths. Statistical differences are observed for all trips purposes with the exception of other-other trips. Home-based work, college, and other, and both categories of non-home-based trips are underestimated by the simulation, while home-based school and shop are overestimated.

Table 9 Comparison of HTS and Simulated Vehicle Trip Lengths (minutes)

Purpose	Adelaide		Simulation	
	Mean	Std. Dev.	Mean	Std. Dev.
Home-Based Work	22.85	15.72	19.69**	16.81
Home-Based School	11.03	8.344	16.50**	13.75
Home-Based College	23.01	22.31	18.10**	15.39
Home-Based Shop	11.01	8.988	12.06**	11.47
Home-Based Other	15.50	17.30	14.05**	15.79
Other-Work	20.41	20.83	14.67**	16.89
Other-Other	15.08	21.60	14.40	16.02
TOTAL TRIPS	16.41	17.97	15.00**	15.68

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

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

Data Expansion

A more useful way to assess the effectiveness of the simulation is to estimate regional statistics from each of the household travel survey and the simulation data. To do this, the HSF data were used as the basis for calculating expansion factors for each of the observations in the household travel survey and the simulation data. Because the HSF is a simple one percent sample of the census data, we expanded the HSF to the full population of Adelaide, and then estimated the occurrence of households in each of the sampling cells for the household travel survey and the simulation data. Dividing these numbers by the sample size in the cells of the sampling design provided the expansion factors for each household. These expansion factors were then applied to the trips measured or simulated for each household.

With this expansion, the number of trips by purpose was expanded first for each of the simulation and HTS data. We also corrected for the non-mobility differences by expanding the Simulation data only to the number of mobile households in the AHTS data, using the mobile household trip rates from the simulation. The results are shown in Table 10. While some individual purposes are mis-estimated by rather substantial margins, the total trips are mis-estimated by slightly less than five percent. School and college trips are substantially over-estimated, while home-based other and other-work trips are under-estimated. This could partly be a result of differences in the way in which trip purposes are classified in each of the U.S. NPTS and the Adelaide HTS data sets. Home-based shopping trips and other-other trips are mis-estimated by small margins (nine and three percent, respectively).

Table 10 Comparison of Total Trips By Purpose

Trip Purpose	AHTS	Simulation	Difference
home-work	375,806	393,610	17,804 (4.73%)
home-school	131,812	189,386	57,573 (43.7%)
home-college	19,174	50,556	31,383 (163.68%)
home-shopping	364,062	396,568	32,506 (8.9%)
home-other	1,193,970	1,031,670	-162,301 (-13.6%)
other-work	415,799	277,606	-138,193 (-33.2%)
other-other	602,560	619,342	16,782 (2.8%)
Total	3,103,184	2,958,739	-144445 (-4.7%)

Some issues arise from this table that bear further investigation. Statistics from the Commonwealth Department of Education indicate that there were 49,037 university students in Adelaide in 1999. The simulation figures in Table 10 are more consistent with these statistics than are the AHTS figures. Each student should make two home-based college trips each day, which may be reduced both by stops on the way to and from the university (which would create home-based other and other-other trips) and by absence from the university on some days, particularly for part-time students. If it were assumed that most students attended university 4 days per week and that 70 percent of these trips were non-stop between home and the university, then one would expect about 55,000 home-based college trips, which is much closer to the simulation of 50,500 than to the AHTS of 19,000. Further investigation of the AHTS data is warranted to determine reasons for the low numbers of school and college trips. It has been ascertained that the survey period from 29 March to 31 July included four weeks of school holidays (out of a total of 18 weeks). The NPTS, from which the simulation distributions are derived, is collected over all 12 months of the year, so also includes school holidays and other seasonal variations.

Table 11 shows the differences in estimated mode shares by purpose and for the total of all purposes. Because of the discrepancies in Table 10, it must be expected that there will be over-estimates for most purposes, and under-estimates for home-based other and other-work trips.

The simulation over-estimates car drivers for all purposes and for the total, where the over-estimation is 145,405, or 8.6 percent. Car passengers are estimated fairly well, with simulation underestimating car passengers by 29,411 or 3.7 percent. Public transport ridership is over-estimated by 8,739 or 5.4 percent, which is more than accounted for by school trips, which are over-estimated by 93,000, while school trips for other modes are generally fairly reasonable. Also, other-work trips are underestimated on public transport by 53,400, which is an unexpected result. Bicycle and walk trips are substantially under-estimated by the simulation, with less than half of the correct number estimated. The biggest component of these occurs in home-other trips, where 103,000 more bicycle and walk trips are reported than simulated, together with another 63,000 for other-other trips.

Table 11 Mode by Purpose for the Two Data Sets

Mode	Source	Home-work	Home-school	Home-college	Home-shopping	Home-other	Other-work	Other-other	Total
Driver	HTS	313,796	1,642	11,176	226,134	576,631	291,560	278,638	1,699,577
	Simulation	336,886	3,155	33,195	266,699	608,737	227,417	368,893	1,844,982
Pass-enger	HTS	29,164	86,514	3,501	60,820	393,435	27,171	199,140	799,745
	Simulation	37,333	69,720	5,568	100,374	332,291	31,158	193,890	770,334
Transit	HTS	16,532	2,004	1,302	8,032	52,440	57,838	22,668	160,816
	Simulation	10,517	95,970	5,343	10,731	25,031	4,478	17,485	169,555
Bike/ walk	HTS	15,883	41,653	3,195	67,877	168,345	38,601	101,116	436,670
	Simulation	8,874	20,540	6,370	18,764	65,191	14,496	38,849	173,084
Total	HTS	375,375	131,813	19,174	362,863	1,190,851	415,170	601,562	3,096,808
	Simulation	393,610	189,385	50,476	396,568	1,031,250	277,549	619,117	2,957,955

Table 12 shows the results of the two data sets with respect to time of day and trip purpose. Again, the misestimates by purpose remain in the table.

Table 12 Time of Day by Purpose for Simulation and HTS

Time Period	Source	Home-work	Home-school	Home-college	Home-shop	Home-other	Other-work	Other-other	Total
6am – 9 am	HTS	150,468	70,289	6,741	27,797	168,089	35,730	44,104	491,029
	Sim.	143,213	93,047	14,839	22,704	121,768	39,236	39,921	474,728
9am – 4 pm	HTS	90,118	56,045	7,373	229,944	547,670	270,209	419,208	1,620,567
	Sim.	95,928	83,820	22,133	219,249	438,886	178,093	393,041	1,431,150
4pm – 7 pm	HTS	92,612	6,020	5,806	106,659	363,349	52,292	123,940	769,208
	Sim.	96,151	10,971	7,902	92,313	268,781	48,844	109,203	634,165
7pm – 6 am	HTS	42,608	230	2,283	21,918	170,590	19,426	40,546	297,601
	Sim.	58,318	1,548	5,682	62,303	202,235	11,433	77,177	418,696
Total	HTS	375,806	131,812	19,174	364,061	1,193,971	415,798	602,560	3,103,182
	Sim.	393,610	189,386	50,556	396,569	1,031,670	277,606	619,342	2,958,739

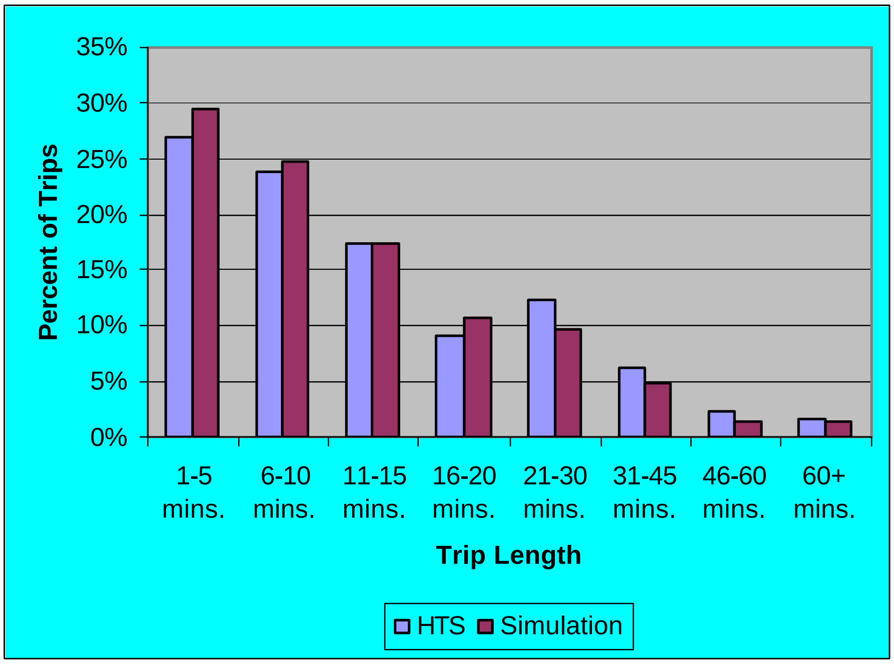
Both data sets agree that the largest proportion of trips occur in the 9 a.m. to 4 p.m. period, although the proportion of trips in this period in the HTS is higher than in the simulation data. Interestingly, the overnight period from 7 pm to 6 am, shows a simulation overestimate of 120,000 trips, or almost 50 percent. The morning and evening peaks show some underestimation by the simulation, more markedly in the evening peak, where about 130,000 more trips are included in the evening peak in the AHTS than the simulation.

Finally, Table 13 and Figure 1 show the trip length distributions by purpose from the two data sets. Allowing for the differences in numbers of trips by purpose, the trip-length distributions by purpose in Table 13 do not appear dissimilar.

Table 13 Trip Length By Purpose from HTS and Simulation

Trip Length	Source	Home-work	Home-school	Home-college	Home-shop	Home-other	Other-work	Other-other	Total
1-5	HTS	35,117	47,798	2,519	140,203	332,620	82,825	196,381	837,463
mins.	Sim.	62,618	46,643	6,537	131,761	340,373	90,850	194,534	873,316
6-10	HTS	56,982	42,754	2,953	112,689	303,883	76,791	147,014	743,066
mins.	Sim.	78,790	43,587	13,514	109,390	265,932	64,181	158,891	734,285
11-15	HTS	71,282	22,715	4,241	54,848	219,353	68,266	99,299	540,004
mins.	Sim.	74,885	31,361	10,406	66,781	175,979	47,557	107,325	514,294
16-20	HTS	45,188	10,061	2,993	24,335	109,966	44,054	49,651	286,248
mins.	Sim.	60,329	21,714	5,486	40,022	97,749	28,601	62,446	316,347
21-30	HTS	96,898	5,574	3,265	20,614	130,446	66,000	59,734	382,531
mins.	Sim.	64,936	25,160	8,819	28,000	83,360	24,571	53,414	288,260
31-45	HTS	50,445	2,220	1,582	7,549	59,181	45,414	27,849	194,240
mins.	Sim.	34,847	16,169	4,106	12,239	41,774	12,172	22,374	143,681
46-60	HTS	14,900	216	1,142	2,837	20,524	20,592	10,182	70,393
mins.	Sim.	10,974	3,906	788	5,892	9,796	3,932	8,925	44,213
60+	HTS	4,797	475	479	792	17,527	11,707	11,700	47,477
mins.	Sim.	6,232	846	899	2,483	16,707	5,742	11,434	44,343
Total	HTS	375,609	131,813	19,174	363,867	1,193,500	415,649	601,810	3,101,422
	Sim.	393,611	189,386	50,555	396,568	1,031,670	277,606	619,343	2,958,739

Figure 1 Trip Length Frequency Distributions



The overall trip-length frequency distributions are very similar, as shown in Figure 1. There is an unexpected dip in the HTS distribution at 16-20 minutes, which does not occur in the simulated data. The simulated data shows slightly more short trips and fewer long trips than the HTS, but the differences are generally on the order of no more than one to two percent.

Conclusion

While the simulation produced a considerable number of statistical differences to the AHTS that warrant further investigation as to how the simulation procedure can be improved, the overall performance of the simulation was not much different from that found in application to Dallas and Salt Lake in the U.S. As has been stated, there do appear to be significant differences between the US and Australia in terms of reliance upon cars. As a result, the simulated results are somewhat less close than would be desired when applied to the Australian sphere. This suggests that in the future, perhaps a small local survey be conducted and the results used to update the distributions so as to take into account the local trip distributions. Nevertheless, when one considers that the Adelaide data showed about 30 percent fewer workers per household, an average of 50% more non-car-owning households, and about 25% lower average cars per household, it is remarkable that the simulations for Adelaide did as well as they did. This seems to indicate that the simulation procedure used here is capable of adapting to quite a wide range of differences in socio-demographics and transport supply situations.

Recent work by Kothuri (2002) on Bayesian updating of the simulations suggests that this may be a simple method by which the results can be improved. Kothuri's method for Bayesian updating updates the NPTS distributions by using a small sample (about 300-500 households) from the local area, and uses Bayesian Updating with subjective priors This would be the next logical step in this procedure of simulation. At the same time, it is important to note that statistical significance of differences between the simulated and original survey results is neither a sufficient nor a necessary condition for acceptance of the simulation procedure. These results are inputs to descriptions of the region, obtained through data expansion and weighting procedures, and to models that describe and forecast travel behaviour.

The final section of the paper showed comparisons of the aggregate predictions from the simulation data and the AHTS. These comparisons show that the simulation has generally produced very reasonable aggregate trip totals, although there remains some concern about trip purpose definitions. Alternatively, there is the possibility that people in Adelaide chain trips more than in the U.S., leading to changes in the distribution of trips by purpose. However, for the simulation from U.S. distributions to provide an overall prediction of travel that is within five percent of the estimates from the Adelaide HTS is remarkable and suggests that the simulation procedure has performed very well.

References

Australian Bureau of Statistics Census: 1996 Census Household Sample File (HSF).

Cambridge Systematics, Inc., (1996) *Travel Survey Manual*, prepared for U.S. DOT and U.S. EPA, TMIP Program, Washington DC, July 1996.

Greaves, S.P. (1998) *Applications of GIS Technology in Recent Household Travel Survey Methodologies*. Report prepared for the Travel Model Improvement Program, Federal Highway Administration.

Greaves, S.P. (2000), *Simulation Household Travel Survey Data for Metropolitan Areas*, Unpublished Ph.D. Dissertation, Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, Louisiana.

Greaves, S.P. and P.R. Stopher (2000) Creating a Simulated Household Travel/Activity Survey B Rationale and Feasibility Analysis, *Transportation Research Record No. 1706,* pp. 82-91.

Kothuri, S. (2002), *Bayesian Updating Of Simulated Household Travel Survey Data For Small/Medium Metropolitan Areas*, Master's Thesis, Louisiana State University, Baton Rouge, Louisiana, U.S.A.

Transport South Australia of the Department for Transport, Urban Planning and the Arts: Metropolitan Adelaide Household Travel Survey (AHTS) 1999

Stopher, P.R., Greaves, S.P. and Bullock P. (2001) Synthesizing Household Travel Survey Data: Application to Two Urban Areas, Unpublished Paper.

U.S. Census Bureau Census 2000 Initial Response Rates at http://www.rates.census.gov/