DISAGGREGATE COMPARISION OF TRAVEL REPORTED BY TWO AMERICAN SURVEYS, THE NHTS AND THE ATUS

Ramakrishna Yennamani, Graduate Student, University of Florida, USA

Sivaramakrishnan Srinivasan, Assistant Professor, University of Florida, USA[1]

Draft paper[2] submitted for presentation at the

8th International Conference on Survey Methods in Transport

Annecy, France, May 25-31, 2008.

BACKGROUND AND OBJECTIVES

Data from household activity-travel surveys constitute a fundamental input to modeling travel-demand for the purposes of transportation planning and policy analysis. In practice, such surveys typically differ from one another in several ways. One such difference is the type of survey instrument used. Specifically, some surveys use the "trip-based" instrument (respondents are queried on each trip undertaken during the day) whereas others use the "activity-based" or "time-use oriented" instrument (respondents are queried on different types of activities pursued in-home and out-of-home during the day). It has been argued that activity-based surveys are more intuitive and therefore help better recall of short and infrequent trips (Pendyala, 2003) whereas the trip-based surveys require the respondents to report their travel "out-of-context of the activities undertaken" (Harvey 2003). Consequently, the expectation is that activity-based

[1] Corresponding author (siva@ce.ufl.edu)

[2] Please do not cite without the consent of the authors

surveys would result in greater number of trips and total travel duration per respondent compared to trip-based surveys. Alternately, it could also be stated that trip-based instruments may result in under-reporting of trips. Such differences could be critical from the standpoint of travel modeling. For example, Bricka and Bhat (2006) point out that although under-reporting of travel at the individual level might appear to be insignificant (1-2 trips), it would magnify to several hundreds of trips when accumulated to the universe of the survey. Further, analysis undertaken by Wolf *et al.* (2003) indicates that under-reporting of trips could result in as much as 40% underestimation of Vehicles Miles of Travel (VMT). These findings underscore the need to quantify and correct the under-reporting of travel. The focus of this study is on the former aspect.

Harvey (2003) presents some empirical evidence to support the claim that trip-based surveys under-estimate travel relative to activity-based surveys based on past research by Jones *et al.* (1983) and Vilhelmson (1997). He reports Jones *et al.* (1983) found that activity-based surveys identified, on an average, 0.51 trips per person more than trip-based surveys whereas Vilhelmson (1997) found this difference to be 0.3 trips per person. In recognition of this deficiency, there have been improvements in the administration of the trip-based surveys. For instance, Sharp and Murakami (2004) report that additional "probe" questions were added to the 2001 National Household Travel Survey (a trip-based survey) conducted in the United States to better capture trips that were thought to be traditionally under-reported in the trip-based surveys. Subsequent research undertaken by Bose and Sharp (2005) indicates that aggregate travel estimates (trip rates and travel durations) obtained from the NHTS were very similar to those from the 2003 American Time Use Survey (an activity-based survey). However, it should be noted that these researchers did not control for other differences (such as rules relating to proxy reporting and the provision of a diary to use as a memory jogger) between the two surveys. Thus, while the research concludes that the two national surveys produce similar aggregate estimates of travel, the impact of the survey instrument itself is not readily apparent.

It is also interesting to note that, much (if not all) of past research has involved aggregate comparisons across surveys. With the field of travel-demand modeling moving towards disaggregate, behaviorally-oriented methods, it is also important for us to understand disaggregate (or individual-level) differences in activity-travel patterns obtained from different

types of surveys. Further, disaggregate analysis allows better control for the differences in socio-economic and other characteristics of the respondents across the surveys being compared. Consequently, one can be more confident that the observed differences in the magnitudes of travel are because of survey differences.

In the light of the above discussions, the objective of this research is to undertake disaggregate (individual-level) comparison of travel reported from two recently conducted national-level surveys in the Unites States: the NHTS and the ATUS. We measure the magnitude of travel in terms of the total number of daily trips as well as the total duration invested in traveling during the day. Specifically, our research questions can be formulated as follows:

1. What are the differences in the magnitude of travel reported by an "identical" respondent across the two surveys?
2. To what extent does the survey instrument (trip-based versus activity-based) contribute to these differences?
3. How do these differences vary based on the socio-economic characteristics of the respondent?

The rest of this paper is organized as follows. The next section outlines the analysis methodology and describes the econometric structure of the models developed. This is followed by a detailed discussion of the data used in the analysis. The next section presents the empirical results. The paper ends by identifying the major conclusions from this research and avenues for further study.

METHODOLOGY

The total number of daily trips and the total daily travel duration are each modeled using the hetroskedastic linear-regression structure. In the rest of this section, this approach is explained in the context of modeling number of trips.

$$T_i = \beta_0 + \sum_{k=1}^{K} \beta_k X_{ki} + \gamma_0 S_i + \sum_{k=1}^{K} \gamma_k \left(X_{ki} S_i\right) + \varepsilon_i \quad (1)$$

In Equation (1), T_i is the total number of trips for respondent i. X_{ki} is the value of the k^{th} explanatory variable corresponding to respondent i. S_i equals 1 if respondent i is from the ATUS and 0 if (s)he is from the NHTS. β and γ are the coefficients on the explanatory variables including constant terms. The β parameters are called the "base effects" as they capture the effect of explanatory factors (for example, socio-economic characteristics) on travel behavior whereas the γ terms are called "interaction effects" and they capture how the effects of the explanatory factors on travel behavior are different across the two surveys. Note that, for an individual i, characterized by $\{X_{ki} \ \forall k = 1,2,..K\}$, the expected difference in the number of trips between ATUS and NHTS is given by: $\gamma_0 + \sum_{k=1}^{K} \gamma_k X_{ki}$.

The error term assumed to be independently Normal-distributed across the respondents with a mean of zero: $\varepsilon_i \sim N(0, \sigma_i^2)$. The variance of the error term (σ_i^2) is assumed to depend on the survey type and parameterized as follows: $\sigma_i^2 = \sigma^2 \exp(\theta S_i)$. Thus, the variance of the error term is σ^2 for NHTS and $\sigma^2 \exp(\theta)$ for ATUS (θ is called the scale parameter). Note that the specification guarantees positive variance values irrespective of the signs of the estimated coefficients $(\sigma$ and $\theta)$. Further, a positive θ implies higher variance for the ATUS surveys and $\theta = 0$ implies that the variances are the same across the two surveys.

The parameters of the model $(\beta_k \ \forall k = 0,1,2..,K; \ \gamma_k \ \forall k = 0,1,2..,K; \ \sigma$ and $\theta)$ can be estimated by maximizing the following log-likelihood function:

$$\sum_i Ln\left\{ \frac{\exp\left[\frac{-1}{\left(2\sigma^2 \exp(\theta S_i)\right)} \left(T_i - \beta_0 - \sum_{k=1}^{K} \beta_k X_{ki} - \gamma_0 S_i - \sum_{k=1}^{K} \gamma_k \left(X_{ki} S_i\right) \right)^2 \right]}{\sqrt{2\pi\left(\sigma^2 \exp(\theta S_i)\right)}} \right\} \quad (2)$$

In addition to the differences in the survey instrument, there are other key survey-administration differences between the NHTS and ATUS (discussed in further detail in the section on Data) which can influence trip reporting. Specifically, the NHTS respondents may choose to self-report their travel or report by proxy. They may or may not use the diary provided to them for recording their travel. Finally, the number of days from the "travel day" to the day of the interview could depend on the characteristics of the respondent. This is in contrast to the ATUS which does not allow proxies, does not provide a diary, and always collect data for "yesterday" (*i.e.*, the number of days between the "travel day" and interview day is always 1). Consequently, the methodology as described above captures the disaggregate differences across surveys that differ not only in the survey-instrument but also other survey-administration protocols. This approach can be extended to capture the marginal influence of the survey instrument by controlling for the effects of proxy-reporting, use of diary, and the number of days between "travel day" and interview day using instrumental variables (as the choices of proxy, diary usage, etc. themselves could be endogenous – *i.e.,* unobserved factors impacting the number of trips could also impact of choice of use of diary, reporting by proxy, etc.).

$$T_i = \beta_0 + \sum_{k=1}^{K} \beta_k X_{ki} + \gamma_0 S_i + \sum_{k=1}^{K} \gamma_k \left(X_{ki} S_i\right) + \delta_1 PR_i + \delta_2 DU_i + \delta_3 MD_i + \varepsilon_i \qquad (3)$$

In Equation (3), PR_i, DU_i, and MD_i are respectively the predicted probabilities of (1) proxy reporting, (2) use of diary, and (3) having multiple days between travel-day and interview for any NHTS respondent i. δ_1, δ_2, and δ_3 are the coefficients on these probability variables to be estimated. The model parameters $\left(\beta_k \;\; \forall k = 0,1,2..,K;\;\; \gamma_k \;\; \forall k = 0,1,2..,K;\;\; \delta_1, \delta_2, \delta_3, \sigma \text{ and } \theta\right)$ can be estimated using maximum-likelihood methods.

Note that, based on the revised specification, $\gamma_0 + \sum_{k=1}^{K} \gamma_k X_{ki}$ captures the marginal effect of the survey instrument (activity-based versus trip-based) on the reported differences in the number of trips for any individual i, characterized by $\left\{X_{ki} \; \forall k = 1,2,..K\right\}$. The same value

$\left(i.e., \gamma_0 + \sum_{k=1}^{K} \gamma_k X_{ki} \right)$ from model in equation 1 captures the "net" difference between the ATUS and NHTS.

The probabilities of proxy reporting, use of diary, and having multiple days between travel-day and interview are determined from binary-logit models estimated separately using data from the NHTS.

For example, $PR_i = \dfrac{\exp\left(\sum_{t=0}^{T} \lambda_t Z_{ti} \right)}{1 + \exp\left(\sum_{t=0}^{T} \lambda_t Z_{ti} \right)}$ (4)

In Equation (4), Z_{ti} is the value of the t^{th} explanatory variable corresponding to respondent i and λ_t is the corresponding model coefficient to be estimated. A positive coefficient implies that the corresponding explanatory variable increases the probability of proxy reporting. Once again, the model parameters can be estimated using the maximum-likelihood techniques.

All estimations were performed using the GAUSS matrix-programming language and MAXLIK library of routines for maximum-likelihood estimations. Specifically, the log likelihood functions and the gradient functions were coded in GAUSS and maximized using the MAXLIK functions.

It is useful to acknowledge here that we use rather simple econometric methods for the analysis so as to facilitate ease of interpretation of results. There are at least two major avenues for methodologically enhancing this analysis. First, discrete-choice methods (such as ordered probit) can be used in place of linear regression to model number of trips. Such approaches can also be applied to model the trip durations to account for respondents' inherent rounding-off of time to the nearest 10 or 15 minute. Second, choices of proxy-reporting, diary use, and number of days before interview can be estimated simultaneously with the model number of trips/duration. Such a procedure might be a more efficient control for potential endogeneity than the two-stage instrumental-variable-type method adopted in this paper.

DATA

This section describes the data used in this analysis. First, we present a detailed description of the two surveys. Second, the sample formation procedure is explained. Finally, the sample characteristics are presented and discussed.

SURVEY CHARACTERISTICS

Data used in this analysis come from two sources: (1) The National Household Travel Survey (NHTS) conducted in 2001-2002 and (2) The American Time Use Survey (ATUS) conducted in 2003-2006. Both surveys collected data from a national sample of civilian, non-institutionalized population (*i.e.,* people not living in college dormitories, nursing homes, other medical institutions, prisons, and military bases) of the United States.

The NHTS is a U.S. Department of Transportation (DOT) effort sponsored by the Bureau of Transportation Statistics (BTS) and the Federal Highway Administration (FHWA) to collect data on both long-distance and local travel by the American public. Most recently, the NHTS data were collected in 2001-2002. Additional details on the NHTS and the resulting data can be obtained from: http://www.bts.gov/programs/national_household_travel_survey/

The ATUS is conducted by the Census Bureau under contract with the Bureau of Labor Statistics and collects detailed individual-level daily time use information. The sample is drawn from a subset of households responding to the Current Population Survey (CPS) interviews. Data collection began in January 2003 and has continued yearly since. Additional details on the ATUS and the resulting data can be obtained from: http://www.bls.gov/tus/home.htm.

In addition to the fundamental difference in the objectives for collecting the data, the two surveys have the following other differences:

1. NHTS obtained data for all household members but only one individual aged 15 years or older is randomly selected from each sampled household for the ATUS.

2. The ATUS includes households without a telephone whereas the NHTS includes only households with telephones.
3. The NHTS surveys are distributed uniformly across the seven days of the week whereas the ATUS over samples the weekend days (50% records are for the two weekend days and 50% are for the five week days).
4. The NHTS used a trip-based survey instrument. Data on the start time, end time, distance, mode, and purpose are collected for each trip. The ATUS used a time-use oriented survey instrument. Data on the start time, end time, location, and type are collected for each activity episode (for travel episodes, the location refers to the mode of travel).
5. NHTS provided respondents with a diary to record their travel whereas the ATUS did not. The actual use of the diary by the respondent in the case of NHTS was, however, optional.
6. NHTS allowed proxy reporting while ATUS did not.
7. In the case of ATUS, the activity-travel information was necessarily retrieved on the day after a random "travel day". However, in the case of NHTS, the data could have been retrieved several days after the pre-assigned "travel day".

SAMPLE FORMATION

In the light of the differences between the two surveys discussed in the previous section, the procedure for generating "comparable" samples from the two surveys for model estimation is outlined here. First, the following restrictions are imposed:

- Only individuals aged 21 or over from both surveys were included. (Although the ATUS provides data for individuals aged 15 and older, we find high proxy-reporting for persons of age 15-20 in the NHTS.) Further, one individual was randomly selected from each household in the NHTS for inclusion in the analysis.
- Only households with telephones from the ATUS were included.
- Only weekday records from both surveys were included.

The next and arguably the major step in the sample formation procedure was the identification of trips from the ATUS that are consistent with the NHTS definition of trips. This procedure involved the following two steps:

- Contiguous travel episodes were "chained" to create a single trip. Most of such continuous travel episodes involved a mode change (for example, the contiguous episodes may be walk to bus stop, wait for bus, travel in bus, and walk to destination, all of which would be reported as a single trip in the NHTS). A second, but much less common, situation involving multiple, successive travel episodes was when the respondent was pursuing some other activity during a part of a trip (for example, talking on the phone for 5 minutes in the middle of a 25-minute trip might get reported as three episodes in the ATUS whereas activities pursued during travel were not explicitly recorded in the NHTS).
- Round-trips with no specific destination (such as biking/walking/jogging around the block or walking the dog) get reported as two travel episodes in the NHTS (with each trip having approximately one-half of the total travel duration). However, such episodes were not even classified as travel in the ATUS. Biking, jogging, and walking activities were classified under "Sports, Exercise, and Recreation" episodes and walking the dog was classified under "Household Activities / Pet Care". Hence, walking, biking, jogging, and pet-care episodes taking place at an "out-doors" location that were immediately preceded and succeeded by episodes at home were classified as two trips each.

Once the trips were identified, they were appropriately aggregated to determine the total number of trips and the total daily travel duration for each respondent. The explanatory variables were created ensuring consistent definitions across the two surveys. Finally, the data were subjected to additional cleaning to ensure internal consistency among the explanatory variables and to remove records with missing data.

SAMPLE CHARACTERISTICS

The final estimation sample comprises 17,429 individuals from the NHTS and 20,311 individuals from the ATUS. Descriptive statistics on the number of trips and total travel duration are presented in Table 1. Overall, the distribution appears to follow similar trends across the two surveys. The average number of trips per person and the average daily travel duration are both slightly higher in the NHTS compared to the ATUS. This is despite the observation that fewer persons report undertaking no trips in the ATUS compared to the NHTS (10.39% versus 12.36%). Table 2 presents the sample distributions for the different explanatory variables (individual and household characteristics, day of the week, and location variables) to be used in the analysis. Once again, the samples appear to be reasonably similar. It is however interesting to note the following differences: (1) ATUS has a larger share of non-Whites [15.7% versus 11.4%] compared to the NHTS, (2) ATUS has considerably more single parent households than the NHTS [6.6% versus 3.5%], (3) The NHTS has more households without children compared to the ATUS [64.1% versus 53.9%], and (4) The ATUS records are evenly distributed across the weekdays whereas the NHTS has more records for Wednesdays than any other weekday. Overall, the two survey samples appear reasonably similar in terms of their socio-economic, spatial, and temporal distributions.

EMPIRICAL RESULTS

This section of the paper discusses the empirical model results. As already indicated, we measure travel in terms of (1) the total number of daily trips and (2) the total daily travel duration. For each measure of travel, two types of models are developed. The first type does not control for survey-administration differences across the two surveys and hence captures the "net" difference in travel. These are presented in Table 3 (number of trips) and Table 4 (travel duration). The second type controls for proxy reporting, use of diary, and the number of days between the survey day and the interview day and hence captures the marginal influence of the survey instrument on the difference in reported travel. These are presented in Table 5 (number of trips) and Table 6 (travel duration). In addition, binary-logit models were estimated to determine probabilities of proxy-reporting, diary usage, and having multiple days between travel-day and

interview (these are required as inputs to the second type of models, see discussion in the methodology section). These models are presented in Table 7 but are not discussed further in the text. The two types of models are however discussed further in separate subsections.

NET DIFFERENCES ACROSS SURVEYS

Table 3 presents the model for number of trips whereas Table 4 presents the model for total travel duration. Only statistically significant effects are presented. In both tables, the parameters under the column "Base Effects" capture the impacts of socio-economic, day-of-the-week, and spatial-location characteristics on trip-making. Overall these effects appear reasonable and consistent with previous findings in travel-demand modeling literature. The parameters under the column "Interaction Effects" are of direct relevance to this research and hence discussed further.

In Table 3 (model for number of trips), all the statistically significant interaction effects with the exception of the coefficient on "Citizen" are negative and almost all of them are greater in magnitude than the coefficient on "Citizen". Thus, it appears that ATUS surveys will generally result in *fewer* trips per person than the NHTS for several demographic segments (such as males, single-persons, and couples) and *higher* trips per person in other demographic segments (such as mothers who are US Citizens).

In Table 4 (model for travel duration), all the statistically significant interaction effects are negative with the exception of the coefficient on "Constant" term. It may be inferred that unemployed /low-income persons might report *lesser* travel duration in the ATUS compared to the NHTS. In contrast employed females with children will report *higher* duration in the ATUS compared to the NHTS.

Overall, the models in Tables 3 and 4 indicate systematic heterogeneity in the "net" impact of the survey on the magnitude of the travel reported. Further, it is also interesting to note that the ATUS can also result in fewer trips and lesser travel durations for certain demographic segments.

However, on observing the coefficients on the interaction variables (note that all the explanatory variables in the specification are binary), it appears as though the implied magnitude of differences in the number of trips and the travel duration is not large.

Finally, the scale parameters (θ) are estimated to be negative (albeit only marginally significant). This indicates that the variance in the magnitude of travel among an observationally-identical population is greater when measured using the NHTS. This result appears reasonable considering that the NHTS allowed respondents to report by proxy, did not enforce the use of the diary, and varied in the number of days between the assigned diary day and the survey day.

EFFECT OF SURVEY INSTRUMENT

Table 5 presents the model for number of trips whereas Table 6 presents the model for total travel duration. Only statistically significant effects are presented. In both tables, the parameters under the column "Base Effects" capture the impacts of socio-economic, day-of-the-week, and spatial-location characteristics on trip-making. This column also includes the NHTS-specific parameters (δ). The corresponding probability values were determined using the binary-logit models presented in Table 7. The signs of the estimated coefficient on these parameters indicate that (1) a person with a higher probability of having his/her travel proxy-reported will have fewer trips reported, (2) a person who is more likely to use the diary is also likely to report more trips, and (3) a person who is more likely to be contacted 2 days or more after the diary day also reports more trips (perhaps the busier people are the ones who are more likely to be contacted later – a hypothesis also supported by the models in Table 7).

The parameters under the column "Interaction Effects" are of direct relevance to this research and hence discussed further. As already discussed, these parameters capture the effect of the survey instrument as the other differences among the two surveys are controlled for in the model. In both the models (Tables 5 and 6), the "Constant" terms in the "Interaction Effects" are large, positive, and statistically significant. Further, the magnitudes of these parameters are greater all other statistically-significant interaction effects in the corresponding models. This result

indicates that the activity-based survey instruments result in *higher* estimates of travel (both in terms of number of trips and travel duration) for all demographic segments. Further, the difference is the largest for younger persons and smallest for males, low income persons, and low-education persons. Since younger persons, females, high-income persons, and highly-educated persons may be expected to make more "incidental" and/or "non-mandatory" travel, the result appears reasonable and consistent with the popularly-held notions about the impact of activity-based surveys on travel reporting. Further (and in contrast to the results from the models in Tables 3 and 4), the magnitudes of the interaction effects are larger indicating differences in the number of trips and travel duration that may be practically significant.

Finally, the scale parameters (θ) are estimated to be negative, lesser in magnitude compared to the corresponding coefficients in Tables 3 and 4, and statistically significant at a lesser confidence level compared to the corresponding coefficients in Tables 3 and 4.

SUMMARY AND CONCLUSIONS

The accuracy of travel-demand models and the forecasts made using them are critically dependent on the quality of the data obtained from household-travel surveys. Further, it has been well recognized that the type of survey instrument used has a significant impact on the magnitude of travel that gets reported. In this context, this study contributes by examining disaggregate (individual-level) differences in travel reported from two recently conducted national-level surveys in the Unites States: the NHTS and the ATUS. Substantial processing was undertaken to produce comparable samples from the two surveys. Hetroskedastic linear regression models were estimated to compare the differences in terms of the number of trips and total travel duration.

On examining the net differences between the two surveys, we find that ATUS surveys can result in *fewer* trips (*lower* travel durations) per person than the NHTS for certain demographic segments and *more* trips (*higher* travel duration) per person in other demographic segments.

However, it also appears as though the implied magnitude of differences across surveys (in both the number of trips and the travel duration) is not large. Further, the variance in the magnitude of travel among an observationally-identical population (*i.e.,* unobserved heterogeneity) is greater when measured using the NHTS. This result appears reasonable considering that the NHTS allowed respondents to report by proxy, did not enforce the use of the diary, and varied in the number of days between the assigned diary day and the survey day.

Empirical results from a second set of models that control for differences in survey administration indicate that the activity-based survey instrument results in *higher* estimates of travel (both in terms of number of trips and travel duration) for all demographic segments. Further, the difference is the largest for younger persons and smallest for males, low income persons, and low-education persons. Since younger persons, females, high-income persons, and high-education persons may be expected to make more "incidental" and/or "non-mandatory" travel, the result appears reasonable and consistent with the popularly-held notions about the impact of an activity-based survey instrument on travel reporting.

Overall, we find that the two national-level surveys (the NHTS and the ATUS) result in reasonably similar estimates of the magnitude of travel despite differences in the survey instrument. This is possibly a manifestation of the provision of travel diaries in the NHTS "offsetting" the effects of using a trip-based instrument.

There are several interesting avenues for further research. First, it would be useful to examine whether the results from this study are robust to alternate methods used to create "NHTS-equivalent" trips from the ATUS data. Second, in this study we only examine the total daily travel. Extending this research to examine specific trip types and non-motorized travel is necessary. Third, the analysis methodology can be enhanced. For example, (1) ordered-response discrete-choice methods can be used in place of linear regression and (2) choices of proxy-reporting, diary use, and numbers of days before interview can be estimated simultaneously with the model number of trips/duration.

REFERENCES

Bricka, S. and Bhat, C.R. (2006) *A Comparative Analysis of GPS-Based and Travel Survey-based Data*, Transportation Research Record, Vol. 1972, pp. 9-20, 2006.

Bose, J. and Sharp, J (2005) *Measurement of Travel Behavior in A Trip-Based Survey versus a Time-Use Survey: A Comparative Analysis of Travel Estimates Using the 2001 National Household Travel Survey and the 2003 American Time Use Survey*, Presented at the American Time Use Survey: Early Results Conference, Washington D.C.

Harvey, A. S. (2003). Time-Space Diaries: Merging Tradition, In Transport Survey Quality and Innovation, (Pergamon 2003) Elsevier Science Ltd, Oxford, UK.

Jones, P. M., Dix, M.C., Clarke, M. I., and Heggie, I.G. (1983). *Understanding travel behaviour*. Great Britain: Antony Rowe Limited.

Pendyala, R. (2003) Quality and Innovation in Time Use and Activity Surveys. In P.R. Stopher and P. Jones (eds) *Transport Survey Quality and Innovation*. Elsevier Science Ltd, Oxford, UK, pp. 181-190.

Vilhelmson, B. (1997). *On travel in time use surveys: A comparison with other approaches*. Draft paper tabled at the XIX Conference of the International Association for Time-Use Research (IATUR). Stockholm Sweden, October 8-10, 1997.

Sharp, J. and Murakami, E (2004) *Travel Survey Methods and Technologies: Resource Paper*, presented at the National Household Travel Survey Conference: Understanding our Nation's Travel, Washington D.C.

Wolf, J., Oliviera, M. and Thompson. M. (2003) *Impact of Underreporting on Mileage and Travel Time Estimates: Results from Global Positioning System-Enhanced Household Travel Survey*, Transportation Research Record, Vol. 1854, pp. 189-198.

Table 1. Descriptive Statistics on Number of Trips and Travel Duration

		NHTS	ATUS
Total Daily Number of Trips	0	12.36	10.29
	1	1.30	1.05
	2	18.04	21.03
	3	9.26	10.68
	4	15.89	16.60
	5	10.45	10.75
	6	10.56	9.72
	7	6.97	6.34
	8	5.35	4.69
	9	3.43	2.94
	10	2.38	2.06
	11	1.48	1.25
	12	0.93	0.84
	13	0.61	0.64
	14	0.36	0.44
	15	0.25	0.23
	16+	0.37	0.46
Sample Size		17429	20311
Median # Trips		4	4
Mean # Trips		4.41	4.33
Std. Dev. # Trips		3.10	3.04

		NHTS	ATUS
Total Daily Travel Duration (mins.)	0-15	15.60	14.29
	16-30	6.90	7.25
	31-45	10.69	12.88
	46-60	10.03	9.56
	61-75	11.27	12.16
	76-90	8.86	8.40
	91-105	7.78	8.49
	106-120	5.80	5.38
	121-135	5.12	5.58
	136-150	3.55	3.55
	151-165	2.98	2.78
	166-180	2.18	1.99
	181-195	1.71	1.65
	196-210	1.39	1.18
	211-225	1.07	0.96
	226-240	0.87	0.53
	241+	4.19	3.37
Sample Size		17429	20311
Median Duration		66	65
Mean Duration		83.99	80.89
Std. Dev. Duration		83.04	81.16

Table 2. Descriptive Statistics on Number of Trips and Travel Duration

Explanatory Factors		Sample Shares (%)	
		NHTS	ATUS
Age	21-24	4.00	4.30
	25-29	7.30	7.70
	30-34	8.90	10.60
	35-39	10.00	12.30
	40-44	11.40	12.80
	45-49	11.10	11.00
	50-54	10.40	9.70
	55-59	8.30	8.40
	60-64	7.00	6.60
	65-69	6.00	5.10
	70-74	5.70	4.00
	75-79	4.70	3.70
	80+	5.10	3.80
Gender	Female	55.20	56.00
	Male	44.80	44.00
Employment Status	Not Employed	18.40	18.10
	Part-time	10.70	11.80
	Full-time	52.30	56.20
	Retired	18.50	14.00
Highest Education	High School or Lower	41.00	39.40
	High School - Bachelors	45.60	48.90
	Masters or Higher	13.40	11.70
Born in the US	No	9.10	11.40
	Yes	90.90	88.60
Race	White	88.60	84.30
	African American	6.60	11.00
	Other	4.80	4.70

Explanatory Factors		Sample Shares (%)	
		NHTS	ATUS
Household Structure	Single Person	21.20	21.40
	Single Parent - One Child	1.70	3.40
	Single Parent - >=2 Children	1.80	3.20
	Couple	30.80	22.10
	Nuclear - One Child	8.60	10.20
	Nuclear - >=2 Children	16.10	20.60
	Other - No Children	12.20	10.30
	Other - >=1 Children	7.80	8.70
Number of Children (age <= 17)	0	64.10	53.90
	1	14.80	18.50
	2	14.10	18.10
	3	5.00	6.90
	4	1.50	1.90
	5+	0.50	0.70
HH Income	Less than 25K	23.00	24.50
	25-50 K	32.60	28.70
	50-75 K	19.50	20.40
	More than 75 K	24.90	26.40
Day of the week	Monday	20.20	20.50
	Tuesday	19.60	19.90
	Wednesday	26.80	19.90
	Thursday	16.60	19.80
	Friday	16.80	19.90
Household Census Region	Northeast	18.70	18.20
	Midwest	26.10	25.20
	South	33.60	34.70
	West	21.60	21.90

Table 3. Model for Number of Trips (No Control for Survey Administration Differences)

	Base Effects (β)		Interaction Effects (γ)	
	Param.	t stat.	Param.	t stat.
Constant	3.309	26.900	-0.122	-1.305
Age				
21-29	1.339	16.480		
30-39	1.234	15.576		
40-49	1.378	17.911		
50-59	1.146	15.068		
60-69	1.202	15.162		
70-79	0.995	10.182	-0.190	-1.794
Gender				
Male			-0.384	-9.416
Employment Status				
Part-Time	0.796	15.058		
Education				
High School or Lower	-0.722	-13.386		
High School - Bachelors	-0.155	-3.091		
Citizen				
Yes	0.267	3.498	0.269	2.834
Household Income				
Less than 25K	-0.785	-14.675		
25-50 K	-0.250	-5.579		
50-75 K	-0.150	-3.231		
Race				
White	0.313	7.042		
Household Structure				
Single Person	0.508	7.641	-0.322	-4.102
Single Parent	1.293	14.773		
Couple	0.181	3.082	-0.274	-3.781
Nuclear	0.341	6.418		
Multiple Children				
Yes	0.549	10.981		
Day of the Week				
Monday	-0.672	-13.587		
Tuesday	-0.563	-11.283		
Wednesday	-0.451	-9.313		
Thursday	-0.479	-9.379		
Census Region				
Midwest	0.091	2.604		
Std. Deviation (σ)	2.986 (63.769)			
Scale Parameter (θ)	-0.073 (-1.479)			
Log-Likelihood at Convergence	-128749.633			
Log-Likelihood Constant-only	-130474.728			
Number of Cases	37740			

Table 4. Model for Travel Duration (No Control for Survey Administration Differences)

	Base Effects (β)		Interaction Effects (γ)	
	Param.	t stat.	Param.	t stat.
Constant	1.100	26.339	0.133	3.684
Age				
21-29	0.441	9.340		
30-39	0.400	8.716		
40-49	0.461	9.979		
50-59	0.367	8.067		
60-69	0.406	10.356		
70-79	0.346	10.007		
Gender				
Male	0.199	9.213	-0.107	-3.642
Employment Status				
Unemployed	0.074	1.679	-0.140	-3.218
Part-Time	0.181	4.829	-0.098	-2.897
Retired	0.199	4.893		
Education				
High School or Lower	-0.140	-9.020		
Household Income				
Less than 25K	-0.219	-6.610	-0.155	-4.077
25-50 K	-0.149	-5.642	-0.109	-3.397
50-75 K	-0.111	-5.220		
Race				
African-American	0.063	2.544		
Household Structure				
Single Person	0.076	3.371		
Single Parent	0.221	7.151		
Couple	0.072	2.782	-0.087	-2.657
Multiple Children				
Yes	0.067	3.380		
Day of the Week				
Monday	-0.216	-9.320		
Tuesday	-0.218	-9.516		
Wednesday	-0.160	-7.111		
Thursday	-0.191	-8.146		
Census Region				
Midwest	-0.101	-5.810		
South	-0.038	-2.320		
Std. Deviation (σ)	1.357 (67.815)			
Scale Parameter (θ)	-0.0458 (-1.065)			
Log-Likelihood at Convergence	-99260.729			
Log-Likelihood Constant-only	-100023.454			
Number of Cases	37740			

Table 5. Model for Number of Trips (Controlled for Survey Administration Differences)

	Base Effects (β)		Interaction Effects (γ)	
	Param.	t stat.	Param.	t stat.
Constant	2.170	7.311	1.000	3.425
Age				
21-29	0.600	4.534	0.741	5.118
30-39	0.483	3.912	0.712	5.430
40-49	0.727	6.247	0.541	4.313
50-59	0.633	5.910	0.375	3.169
60-69	0.723	6.816	0.339	2.787
70-79	0.690	8.396		
Gender				
Male	0.192	3.611	-0.576	-8.558
Employment Status				
Part-Time	0.617	7.946	0.248	2.336
Education				
High School or Lower	-0.275	-5.086	-0.292	-4.146
Citizen (ref. = No)				
Yes	0.139	1.761	0.396	4.024
Household Income				
Less than 25K	-0.515	-6.171	-0.338	-3.357
25-50 K	-0.107	-1.718	-0.249	-3.211
50-75 K	-0.177	-3.842		
Race				
White	0.275	6.008		
Household Structure				
Single Person	-0.267	-2.840	0.503	4.709
Single Parent	0.780	4.727	0.457	2.410
Couple				
Nuclear	0.331	6.683		
Multiple Children				
Yes	0.549	11.070		
Day of the Week				
Monday	-0.621	-12.155		
Tuesday	-0.477	-9.052		
Wednesday	-0.412	-8.264		
Thursday	-0.444	-8.584		
NHTS-Specific Variables (δ)				
Probability of Proxy	-3.037	-10.821	NA	NA
Probability of Diary	2.305	11.699	NA	NA
Probability of Multiple Days	1.450	4.007	NA	NA
Std. Deviation (σ)	2.966 (65.227)			
Scale Parameter (θ)	-0.060 (-1.241)			
Log-Likelihood at Convergence	-128631.884			
Log-Likelihood Constant-only	-130474.728			
Number of Cases	37740			

Table 6. Model for Travel Duration (Controlled for Survey Administration Differences)

	Base Effects (β)		Interaction Effects (γ)	
	Param.	t stat.	Param.	t stat.
Constant	0.607	5.377	0.682	5.813
Age				
21-29	0.296	5.707	0.136	2.600
30-39	0.253	5.420	0.146	3.487
40-49	0.340	7.259	0.108	2.638
50-59	0.331	8.881		
60-69	0.369	10.777		
70-79	0.309	9.060		
Gender				
Male	0.239	10.347	-0.154	-5.119
Employment Status				
Unemployed			-0.089	-2.583
Part-Time	0.131	4.499		
Full-Time	0.086	3.166		
Education				
High School or Lower	-0.072	-2.956	-0.076	-2.360
Household Income				
Less than 25K	-0.162	-4.040	-0.175	-3.751
25-50 K	-0.121	-4.313	-0.121	-3.549
50-75 K	-0.107	-5.030		
Race				
African-American	0.059	2.355		
Household Structure				
Single Parent	0.171	5.617		
Multiple Children				
Yes	0.051	2.761		
Day of the Week				
Monday	-0.190	-8.052		
Tuesday	-0.179	-7.502		
Wednesday	-0.138	-6.040		
Thursday	-0.173	-7.331		
Census Region				
Midwest	-0.124	-5.834		
South	-0.064	-3.089		
West	-0.053	-2.268		
NHTS-Specific Variables (δ)				
Probability of Proxy	-0.476	-5.377		
Probability of Diary	0.552	5.949		
Probability of Multiple Days	0.788	4.997		
Std. Deviation (σ)	1.355 (67.907)			
Scale Parameter (θ)	-0.043 (-0.994)			
Log-Likelihood at Convergence	-99237.33			
Log-Likelihood Constant-only	-100023.454			
Number of Cases	37740			

Table 7. Binary Logit Models for Proxy Reporting, Diary Use, and Multiple Days Between Diary-day and Interview

	Proxy Reporting		Diary Use		Multiple Days	
	Param.	t stat	Param.	t stat	Param.	t stat
Constant	-0.528	-3.321	-1.530	-11.188	0.809	6.485
Age						
21-24	-0.274	-2.615			0.806	8.363
25-29	-0.520	-5.230			0.673	7.414
30-34	-0.585	-5.987			0.560	6.259
35-39	-0.616	-6.349	0.179	3.675	0.580	6.513
40-44	-0.521	-5.548	0.229	4.852	0.478	5.480
45-49	-0.651	-6.998	0.189	4.007	0.408	4.723
50-54	-0.605	-6.532	0.114	2.376	0.312	3.643
55-59	-0.541	-5.698	0.141	2.675	0.224	2.567
60-64	-0.475	-4.715	0.283	4.725	0.167	1.849
65-69	-0.769	-7.612	0.219	3.598		
70-74	-0.628	-6.169	0.131	2.098		
75-79	-0.482	-4.523				
Gender						
Male	0.674	21.303	-0.075	-2.808	0.092	3.707
Citizen						
Yes	-0.363	-6.805			-0.213	-4.947
Drivers License						
Yes	-1.137	-18.375	0.766	14.903		
Employment Status						
Not Employed	-0.293	-6.550	-0.077	-2.133	-0.265	-7.737
Part Time Employed	-0.304	-5.632			-0.155	-3.688
Retired					-0.249	-2.942
Highest Education						
High School or Lower	0.526	10.051	-0.386	-8.348		
High School - Bachelors	0.109	2.181	-0.131	-2.963		
Race						
White			0.332	5.718	-0.147	-2.607
African American	-0.437	-6.005	-0.210	-2.755	0.202	2.717
Household Structure						
Single Person			0.125	2.796	-0.477	-8.977
Single Parent - One Child					-0.292	-2.329
Couple	-0.313	-7.984	0.103	3.119	-0.323	-7.305
Nuclear - One Child	-0.156	-2.820	0.096	1.987	-0.309	-5.759
Nuclear - >=2 Children	-0.125	-2.615			-0.255	-5.352
Other - No Children					-0.168	-3.554

Table 7 (Continued). Binary Logit Models for Proxy Reporting, Diary Use, and Multiple Days Between Diary-day and Interview

	Proxy Reporting		Diary Use		Multiple Days	
	Param.	t stat	Param.	t stat	Param.	t stat
Household Income						
25-50 K	-0.278	-4.887	-0.586	-12.766	-0.088	-2.292
50-75 K	-0.157	-3.751	-0.222	-5.970	-0.103	-3.606
More than 75 K	-0.116	-2.728	-0.094	-2.345		
Residential Unit Tenure						
Own	0.208	4.179	0.106	2.967	-0.087	-2.527
Residential Unit Type						
Detached Unit	0.126	2.575				
Multiple Land Line Phones						
Yes	0.088	2.486			0.073	2.521
Day of the Week						
Monday					-0.351	-8.874
Tuesday			-0.130	-3.958	-0.436	-10.876
Wednesday					-0.328	-8.683
Thursday					-0.228	-5.512
Month of the Year						
January			0.558	9.200	-0.694	-11.416
February			1.075	18.439	-0.575	-10.182
March			1.231	21.431	-0.515	-9.335
April			1.110	20.609	-0.368	-6.975
May			1.269	18.644	-0.531	-8.318
June			1.445	19.399	-0.600	-8.893
July			1.261	16.185	-0.342	-4.785
August			1.433	21.308	-0.506	-8.171
September			1.284	17.748	-0.411	-6.114
October			1.204	16.857	-0.352	-5.271
November			-0.194	-2.980	-0.252	-3.887
Year						
2001	-0.200	-6.667				
Housheold Census Region						
Midwest			0.073	1.821		
South			-0.113	-2.952	0.154	5.556
West			-0.112	-2.708	0.304	9.661
Language of Survey						
English	0.907	7.200	0.630	6.635		
Log-Likelihood at Converegence	-13510.324		-17246.139		-19742.658	
Number of Cases	25547		29834		29834	

These models use all respondents aged >= 21 who reported travel for a weekday (not just one random person per houehold)

Single person and single parent households are excluded from the model for Proxy

Sample shares: 24.8% proxy reported, 67.9% diary used, and 45.7% muliple days between travel day and CATI