SAMPLE SIZE REQUIREMENTS FOR REGIONAL HOUSEHOLD TRAVEL SURVEYS

2015-09-30

Transportation Tomorrow Survey 2.0

Khandker Nurul Habib & Wafic El-Assi

UNIVERSITY OF TORONTO
FACULTY OF APPLIED SCIENCE & ENGINEERING
Transportation Research Institute

SAMPLE SIZE REQUIREMENTS FOR REGIONAL HOUSEHOLD TRAVEL SURVEYS

TRANSPORTATION TOMORROW SURVEY 2.0

Table of Contents

1 INTRODUCTION

Since the 1980s, trip diaries of household members have been collected as part of household travel surveys (Harvey 2003). Practitioners have always had issues with these surveys, especially with regards to data quality and low response rates. There have been numerous efforts to improve household surveys, most of which have been targeted at reducing missing/omitted trip information and response burden. However, the reduction of non-response rates in household travel surveys may not be correlated with such endeavors. Non-response rate is a consequence of the evolving lifestyle, technology reliance and increasing time pressure of modern urban life.

Considering the aforementioned factors, Stopher and Greaves (2007a) suspected that future household travel surveys will not be restricted to their current form of diaries, but offered no alternative. They predicted that household travel surveys, especially in the form of travel diaries, will continue to be the only reliable passenger travel data for urban transportation planning in the foreseeable future. The use of GPS, Smart phones, panel surveys, continuous surveys and other innovative approaches are recommended with caution. It is, however, often expected that such advanced approaches would complement the core cross-sectional household travel surveys rather than replace them.

The household travel survey provides basic information on household and individual level characteristics, and activity-travel information of household members, all fundamental to the development of any regional travel demand model (Goulias 2013). However, sample size determination for household travel surveys has proven to be a controversial element in the urban planning process, as statistical considerations are often dominated by cost and political considerations (NCHRP 2008). As a result, there is no consensus on sample sizes for household travel surveys in practice.

Additionaly, despite many valid reasons for switching to continuous surveys (Ampt and Ortuzar 2004; Ortuzar et al 2011), large cross-sectional surveys remain the dominant method across the major metropolitan areas in North America and abroad. Nonetheless, examples of empirical investigation on the adequacy of different sample sizes of cross-sectional household travel surveys are few. Moreover, continuous surveys have been pooled to represent a pseudo cross-sectional survey of sufficient sample size to represent travel demand of the population of a study area. This adds to the complexity of implementing continuous surveys in many North American cities.

Almost all travel survey researchers recommend a combination of data sources to replace large cross-sectional travel surveys. These data sources include small sample panel surveys that utilize GPS or smartphones and continuous cross-sectional surveys. Data fusion is considered to be the statistical tool to combine all such datasets to produce a core database, comparable to the large scale household travel survey. One of the key arguments for replacing large sample cross-sectional surveys by continuous, panel or repeated cross-sectional surveys is the lower sample size requirement. If the rolling average of aggregate travel information is being considered (e.g. trip rates, modal share, etc.), a smaller repeated cross-sectional or continuous travel survey can provide data of similar statistical strengths to that of a once-in-a-while large cross-sectional survey. Some regions (e.g. Montreal) in Canada have tested the feasibility of replacing large sample household travel surveys by continuous repeated cross-sectional surveys. Although countries like Australia (e.g. the Greater Sydney continuous survey) have adopted the continuous survey method, there is no firm evidence that this is a practical alternative to serve the passenger travel data needs of all large urban areas. In either case (continuous or cross-sectional), the sample size of a household travel survey remains a critical element that has yet to reach a consensus.

This report investigates the issue of sample size requirements for household travel surveys from the perspective of adequate data availability to support data-driven evidence-based planning processes in large metropolitan areas. The report is inspired by the prospect of re-designing one of the oldest and most regular (every 5 years since 1986) household travel surveys in North America, the Transportation Tomorrow Survey (TTS) of the Greater Toronto and Hamilton Area (GTHA) (DMG 2015). The sample size of the TTS has traditionally covered approximately 5% of the GTHA household population. The TTS started in 1986, and the latest cycle (5th) was in 2011-2012. The TTS sampling frame leverages a land-line telephone directory to conduct telephone interviews with prospective respondents. In its latest cycle (2011-2012), a web version of the telephone interview was introduced as an alternative option for respondents.

The TTS, however, is now facing issues concerning the under coverage of certain population cohorts from the use of the land-line telephone directory as a sample fame. This has resulted in the under-representation of key population segments, despite the survey's large sample size. This has also prompted investigation into the issue of sample size adequacy and the need of an alternative approaches. Therefore, this report focuses on sample size requirements of large household travel surveys.

The report is organized as follows. Section 2 presents a literature review on household travel survey's sample size requirements. Section 3 discusses the differences in practice of household travel survey sample size determination in Canada and abroad. Section 4 presents an empirical investigation on the representativeness of a large scale household travel survey in Canada - the Transportation Tomorrow Survey. Section 5 presents a recommendation for the determination of adequate sample sizes for household travel surveys. The report concludes with a summary of key findings and recommendations for further research.

2 EXISTING LITERATURE ON SAMPLE SIZE

Statistical procedures for estimating the required sample size of different variables of interest are well-established. Kish (1965), Richardson et al (1995) and the NCHRP report (2008) are a few examples of many other similar sources that explain systematic approaches for estimating appropriate sample sizes while considering specific objective variables of measurements. They illustrate sample size determination processes for random and stratified random sampling, along with various other combinations of methods for both continuous and discrete objective variables. Hence, the calculation procedures (for a single objective variable) for sample size determination are not under debate. The question that is still un-answered is: what is the most appropriate sample size for a multi-objective household travel survey for large urban areas? This is with respect to household travel surveys that are conducted by one or multiple planning agencies of an urban area to collect data necessary to drive various evidence-based planning processes.

Sample size requirements of household travel surveys have been a concern for transportation planners since the 1970s (Stopher and Meyburg 1979). It has been established that the sample size of a household travel survey depends on the purpose of the survey, population representation, variability of key measured variables (e.g. trip generation rate, trip length distribution, trip distribution patterns, modal shares etc.), allowable tolerance of errors in measurement, and the desired confidence limit on the estimates from the sample. After more than a decade-long pause, the next phase of research on this topic showed up in the mid-1990s. Interestingly, during this time period (in the 1980s and early 1990s), the concept of travel demand underwent a paradigm shift, transitioning from an aggregate trip-based approach to the disaggregate tour or activity-based approach.

Earlier studies note that household travel surveys used to be 5% to 10% of population size (Smith 1979). However, Smith (1979) argued that if the main purpose of household travel surveys is to develop travel demand models, the repetition of such large surveys at regular time intervals is redundant. Instead, if stable estimates of key variables from previously conducted large surveys are available, a small sample size may be sufficient for updating the different components of a travel demand modelling system. Smith (1976) proposed a method of estimating such sample size in situations of purely random sampling. Stopher (1982) extended the proposed procedure of Smith for stratified random sampling, where an assumption could be made of the availability of accurate estimates of mean and variance of key variables. Nonetheless, the availability of such input estimates of mean and variance of key variables are often difficult to obtain. For example, Kollo and Purvis (1984) collected household travel survey data over a 20-year period and found that trip rates only remained stable over time when they were aggregated. Specially, disaggregation of trip rates by purpose caused instability of estimates over time.

In terms of critically defining the sample size requirement, Smith identified that trip distribution is the most critical element that may nullify the sample size estimates based on trip generation. He found that at least a 4% sample is necessary to achieve a 90% confidence interval, with a 25% error for trip interchanges between Origin-Destination (O-D) pairs with less than 1100 trip population in between two zones. So, he suggested the use of secondary data sources (e.g. cordon counts, etc.) to create and update O-D matrices in place of collecting large scale (e.g. 4% or more) household travel surveys samples.

The next significant document that has dealt with household travel survey sample size determination is TMIP (1996). The report states that sample rates for household travel surveys receive little, if any, analytical consideration, as sample size is primarily defined by budget considerations. The cost of a household travel survey is considered more important than the accuracy in representation of collected data. It also reports that the exhaustive objectives of household travel surveys inhibit the optimization of sample size estimation. Further, the document recommends that one out of every hundred households (1% of the population) for large urban areas, and one of every ten households (10%) for small suburban areas should be the minimum sample size for household travel surveys. The report capitalizes on the fact that the drop of household travel surveys' sample sizes from over 4% to less than 1% of households happened during the late 1980s without necessarily affecting the accuracy of demand modelling. On the other hand, it also recognizes the importance of large sample sizes for increasing the reliability of sample statistics. It provides a step-by-step procedure for sample size estimation of various types of target variables, and for different sampling procedures. However, it provides no definite guideline for sample size determination for a generalized multi-objective household travel survey that could be used by different planning agencies for various purposes.

Greaves and Stopher (2000) highlighted the importance of large sample household travel surveys while recognizing the increasing cost of collecting larger sample sizes. Large sample sizes are being increasingly demanded for developing advanced disaggregate travel demand models. The authors proposed a simulation technique to generate synthetic household travel survey data in the absence of large sample household travel surveys. Simulation takes, as input, the conditional distributions from the National Personal Travel Survey (NPTS) and Public Use Microdata Sample (PUMS) to generate artificial sample. The PUMS is of a 5% sampling rate and so is considered a reliable data source. Pointer et al (2004) also used the same procedure to generate synthetic household travel survey data for Sydney. They used the Sydney household travel survey, a relatively small continuous survey of 3000 households per year. They noted that, although estimating a travel demand model for a region may not need a large household travel survey, portraying an accurate picture of the spatial distribution of travel demand within the region requires a large sample size.

Ampt and Ortuzar (2004) presented a comprehensive discussion on the sample size requirements of household travel surveys. The authors investigated the sample size of O-D trips from a group of only 34 zones in Santiago by using data from the 1991 Santiago O-D survey. They re-confirmed that they would have needed at least a 4% sample to have achieved a 90% confidence and 25% standard error if they were to conform to Smith's (1979) proposition. A 4% sample size was identified as too large considering trip distribution as a meagre objective of the overall household travel survey in Santiago. Instead, they proposed an alternative heuristic algorithm based on stratified random sampling of selected socio-economic variables. However, they also recognized the fact that actual sample size requirements may be very large if geographic distributions of key variables (e.g. zonal or sub-regional estimates of household car ownership) are of concern. The authors also proposed that large metropolitan areas should implement small sample continuous household travel surveys with once in a while large sample cross-sectional surveys. Stopher and Greaves (2007b) further proved that if a continuous panel survey is to be the method of choice, sample size requirements reduce drastically. The combination of one of the aforementioned approaches with the use of GPS devices, and weeklong surveys instead of a one-day survey is capable of further reducing sample size requirements for household travel surveys (2007a).

In addition, Stopher et al (2008) proved that even with increasing response burden and the possibility of attrition, a week-long household travel survey could be more efficient than a 24-hour travel survey. A week-long survey necessitates a smaller sample size requirement. It also provides a rich dataset that can reflect the dynamics of travel behaviour. As an empirical anecdote, the authors proved that a 7 day GPS assisted household travel survey would require a sample size that is 35% less than that of a typical 1-day household travel survey. Similarly, Bolbol et al (2012) suggested a procedure for estimating sample size requirement for GPS assisted household travel surveys. They suggested that the temporal variability of travel mode choices has to be carefully considered for sample size determination. Furthermore, Goulias et al (2013) considered a week-long GPS assisted household travel survey as the core for their core-satellite approach of urban travel data collection. They recommended small yet detailed household travel surveys as the core, which should follow the form of week-long travel diaries of household members. However, the small sample would have to be complemented by a series of carefully designed satellite (synonymous to an augment survey) surveys targeting specific variables that were under or unrepresented in the core. Nevertheless, their proposal provides no guidelines on sample size requirements.

The NCHRP report (2008) stated that even strictly designed (statistically efficient) sample sizes may not be sufficient for serving many of the critical objectives. The 1990 Southern California household travel survey was presented as a case study. A statistically adequate sample size was estimated at 3,500 to 5,000 households. However, the actual sample size was selected to be 15,000 households, partly due to political reasons. Interestingly, even with such a large sample size, data were not adequate. Low transit modal shares proved to be a major problem, resulting in a small number of observed transit trips. The number of trips was not large enough to estimate a reasonable mode choice model. Therefore, NCHRP (2008) suggested that the sample sizes of household travel surveys should be based on proper stratification of the key variables of concern (socio-economic variables, modal shares, etc.). The report also suggested that, as an alternative to larger sample sizes, a designed sample size should be complemented by augment samples collected for certain zones/sub-regions with a small number of observations.

In summary, it is evident that there is a lack of consensus on the appropriate guidelines for establishing sampling rates for household travel surveys. Although, theoretically, the sample size could be set quite low, the actual sample sizes of urban household travel surveys vary widely. Different trends are observed in different parts of the world. The following section presents a discussion on this.

3 COMPARISON OF RECENT HOUSEHOLD SAMPLING RATES

Table 1 presents a list of recent household travel surveys from the US, Canada, Australia, Europe and South America. The selection of this list is based on web-accessible information. Although it does not provide an exhaustive list of all household travel surveys around the world, it portrays the distinctive approaches in major cities/urban regions.

TABLE 1: SAMPLE SIZES OF RECENT HOUSEHOLD TRAVEL SURVEYS AROUND THE WORLD

City/Region	Survey	Year	Sampling Rate
Canada			
Calgary	Calgary Travel and Activity Survey[1]	2012	3.4% of households
Edmonton	Edmonton Household Travel Survey[2]	2005	4.5% of households
Greater Montreal Region	Greater Montreal Area Origin-Destination Survey[3]	2013	4.6% of households
Greater Toronto and Hamilton Area: GTHA	Transportation Tomorrow Survey: TTS[4]	2011-2012	5.0% of households
National Capital Region: NCR	NCR Origin-Destination Survey[5]	2011	5.0% of households
Saskatchewan	Saskatoon Household Travel Survey[6]	2013	3.0% of households
Vancouver	Metro Vancouver Regional Trip Diary Survey[7]	2011	2.2% of households
Winnipeg	Winnipeg Area Travel Survey[8]	2007	3.3% of households
United States			
Atlanta Region	Regional Travel Survey[9]	2011	0.5% of households
Chicago Metropolitan Area	Regional Household Travel Inventory[10]	2007-2008	0.44% of households
Dallas Metropolitan Area	Household Travel Survey[12]	2008	0.24% of households
New York and New Jersey Metropolitan Area	Regional Household Travel Survey[11]	2010-2011	0.24% of households
Southeast Florida	Household Travel Survey[12]	2007-2008	0.11% of households

1 http://wwwsptest.calgary.ca/Transportation/TP/Pages/Planning/Forecasting/Forecasting-surveys.aspx
2 http://www.edmonton.ca/transportation/RoadsTraffic/2005_HTS_Region_Report_FINAL_Oct24_06.pdf
3 https://www.amt.qc.ca/fr/a-propos/portrait-mobilite/enquetes-en-cours
4 http://www.dmg.utoronto.ca/transportationtomorrowsurvey/
5 http://www.ncr-trans-rcn.ca/surveys/o-d-survey/o-d-survey-2011/
6 https://www.saskatoon.ca/sites/default/files/documents/transportation-utilities/transportation/planning/Attachment3%20Technical%20Report%20HTS_FollowUp_report.pdf
7 http://www.translink.ca/en/Plans-and-Projects/Transportation-Surveys.aspx
8 http://transportation.speakupwinnipeg.com/WATS-Final-Report-July2007.pdf
9 file:///C:/Users/khandker-admin/Downloads/tp_2011regionaltravelsurvey_030712.pdf
10 http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/34910
11 http://www.nymtc.org/project/surveys/survey2010_2011RTHS.html
12 http://www.fsutmsonline.net/images/uploads/mtf-files/Southeast_Florida_Household_Travel_Survey_0205_2014.pdf

State of California	California Household Travel Survey[13]	2010-2012	0.34% of households
Utah State	Household Travel Survey[14]	2012	1.0% of households
Australia			
Adelaide	Travel Survey[17]	1999	5,886 households
Brisbane	Travel Survey[15]	2009	10,000 households
Canberra	Travel Survey[17]	1997	3,054 households
Central Melbourne	Travel Survey[16]	2012	0.7% of total weekday population
Hobart	Travel Survey[17]	2008-2009	2,400 households
Sydney Greater Metropolitan Area	Continuous Household Travel Survey[17]	2015	5,000 households per year
Europe			
France	National Transport and Travel Survey[18]	2007-2008	20,220 households: 10,700 of common national sample, the rest from regional augment samples
Germany	Mobilitat in Deutschland (MiD)[19]	2008	25,922 households
The Netherlands	Onderzoek Verplaatsingen in Nederland (OViN)[20]	2011	43,400 individuals
Spain	Movilia[21]	2007	49,027 households
Switzerland	Microcensus on Travel Behavior[22]	2010	62,868 individuals
South America			
City of Rosario, Argentina	Household Travel Survey[23]	2002	3% of households
Greater Santiago Area	Origin-Destination Survey[24]	2012-2013	1% of households

The first observation worth noting is that different regions/countries have developed their own patterns of household travel survey sample sizes. It is evident that almost all regions which had been collecting regular household travel survey data are gradually moving towards (or at least experimenting with) continuous

[13] http://www.dot.ca.gov/hq/tsip/FinalReport.pdf
[14] http://www.wfrc.org/new_wfrc/publications/Utah_FinalReport_130228.pdf
[15] Stopher et al (2011)
[16] https://www.melbourne.vic.gov.au/AboutMelbourne/Statistics/Documents/Central_Melbourne_Travel_Survey_2012.pdf
[17] http://www.bts.nsw.gov.au/Statistics/Household-Travel-Survey/default.aspx#top
[18] http://www.insee.fr/en/methodes/default.asp?page=sources/ope-enq-transports-deplac-2007.htm
[19] http://mobilitaet-in-deutschland.de/02_MiD2008/index.htm
[20] http://www.cbs.nl/nlnl/menu/informatie/deelnemersenquetes/personen-huishoudens/ ovin/doel/default.htm
[21] http://www.fomento.gob.es/mfom /lang_castellano/estadisticas_y_p ublicaciones/informacion_estadis tica/movilidad
[22] Ohnmacht et al 2012
[23] Ortuzar (2004)
[24] http://datos.gob.cl/datasets/ver/31616

surveys. Large scale continuous travel surveys are normally small-scale repeated cross-sectional surveys collected in an ongoing fashion, rather than once every 5 or 10 years. In either case, household travel survey sample size determination is an important concern. Even for continuous surveys, it is recommended to pool the ongoing surveys in large intervals (3 or 5 years) to form a large pseudo cross-sectional survey (Ampt and Ortuzar 2004).

Among all regions, Canadian cities are pioneers in large household travel surveys. Toronto and Montreal have regular (5-year interval) cross-sectional household travel surveys with a sampling rate of over 4.5%. Montreal has piloted a continuous household travel survey from 2009 to 2012 with an annual sample size of 15,000 households. The continuous surveys were introduced between two large cross-sectional surveys conducted in 2008 and 2013. Other Canadian cities also regularly conduct household travel surveys with sampling frames of 2% to 5% households. The City of Calgary is currently piloting a continuous household travel survey of 2000 households in 2 years. Almost all Canadian household travel surveys are predominantly telephone-based with some introducing a web-version of the telephone survey and small scale GPS applications. Vancouver had the smallest sampling rate of all Canadian cities (2.2%). The metro region has stated in the past that the objective of the survey is mainly for model calibration purposes. The 2004 Metro Vancouver report mentioned that for obtaining detailed travel statistics, such as trip rates and mode shares, a larger sampling rate will be required. Nonetheless, the magnitude of such a survey may be too large, adding costs and complexity to the data collection process (TransLink 2010).

On the other hand, cities and regions in the US have moved to small scale household travel surveys since the 1990s, potentially influenced by Smith (1979) and Stopher (1982). Almost all household travel surveys in the US have a sampling rate of less than 1%. However, US surveys are more dynamic in adopting advanced technology, e.g. GPS. The 2010-2011 New York and New Jersey regional household travel survey used a 10% sub-sample of households to collect wearable GPS-based travel diary data. Even though the sample size remained small, the subsample proved to bring socio-economic groups that otherwise would not have participated in the survey. Furthermore, the GPS subsample allowed the New York Metropolitan Transport Council and the North Jersey Transportation Planning Authority to calculate statistically reliable trip rates that would have otherwise been more difficult to determine using a relatively small sample size. Still, the survey report recognized the fact that this sample size may have been too thin for various travel segments. The 2010-2012 California household travel survey employed a 12% sub-sample for a wearable GPS based travel survey. The biggest travel survey in the US is the National Household Travel Survey (NHTS) with a sampling rate of below 1%. However, in many cases, such data alone are not considered sufficient for demand modelling and evidence-based transportation planning exercises. The California household travel survey for example conveyed difficulty in determining detailed observed travel patterns at the county and/or sub-county levels due to the small sample size. Other difficulties reported include the underrepresentation of certain socio-demographic groups.

Australian cities have been implementing both cross-sectional and continuous travel survey approaches (Stopher 2015). Due to the lack of proper statistics, it is difficult to approximate the sample sizes of Australian surveys. However, it is clear that Australian surveys favor small sample sizes (Stopher et al 2011). Nevertheless, Stopher et al (2011) have highlighted the lack of consistency across these surveys, thus limiting the potential of fusing the numerous datasets into one large survey, which the authors listed as an objective of various Australian planning agencies.

One of the oldest household travel surveys in Australia is The Greater Sydney Metropolitan Area survey. Prior to 1997, the Greater Sydney Area used to conduct large scale cross-sectional surveys every 10 years. Since then, the area has been running a continuous survey. The data is pooled every 3 years, where the total sample size equals that of the pre-1997 cross-sectional survey. Other areas, such as the Central Melbourne area, use a cross-sectional household travel survey. The region uses both a land line based interview (55% of total sample) and a road-side intercepts approach (45% of total sample) for data collection.

The European continent has carried out the most consistent national household travel surveys. Bonnel and Armoogum (2005) stated that national household travel surveys in Europe vary widely in terms of sample sizes. In addition, the authors reported that the sample size determination was not correlated with the size or the characteristics of the countries respected population. One of the critical sampling issues of national surveys is that the sampling process follows a variant of the cluster sampling approach. Cluster sampling may leave out several sub-regions from the data collection process. Thus, spatial distribution of travel behaviour at the smaller metropolitan level may become difficult.

Chile, specifically the city of Santiago, has been a global leader in travel surveys. The latest Santiago household travel survey was of around 1% of households in the region. Chile also has been experimenting with various approaches (e.g. continuous surveys, use of GPS technology and panel surveys) (Ortuzar 2004).

Overall, it is clear that there is no consensus on the selection of sample sizes for household travel surveys. There are, however, recommendations on moving to continuous surveys instead of one-off surveys, but the issue of sample size is never tackled on purpose. Lack of proper data due to the small sample sizes of household travel surveys in the US has presented an issue for many researchers due to their inability to investigate detailed disaggregate (at a zonal or sub-regional level) travel behaviour. Some regions in the US have put forward the claim that small sample sizes prevent the observation of detailed travel patterns at the county or sub-county levels, and under represent certain segments of the population (SEFTC 2014).

This report is mainly concerned with Canadian practice, specifically that of the GTHA. The next section considers a large-sample travel survey for empirical investigation on the representativeness of such surveys.

4 EMPIRICAL INVESTIGATION ON REPRESENTATIONS OF A LARGE SAMPLE HOUSEHOLD TRAVEL SURVEY: THE CASE OF TTS

The Transportation Tomorrow Survey (TTS) in the GTHA is one of the largest (5%) and most regularly conducted (every 5 years since 1986) household travel survey in the world. The TTS study area is composed of 30 municipalities in addition to the City of Toronto's 16 planning districts. The City of Toronto is the largest municipality in the GTHA. The TTS has also been extended to include several smaller municipalities outside the borders of the GTHA. The 2011-2012 TTS survey data was used to investigate its representativeness of the various socio-economic characteristics of its population. Figure 1 presents the aggregate region-to-region peak-period trip matrix of the study area of TTS (DMG 2015).

It is clear that over 70% of the origin-destination pairs have less than 1100 trips. According to Smith (1979), origin-destination pairs with lower than 1100 trips require a sample size of 4% or more to achieve a 90% confidence with a 25% error margin. Considering that around 20% of total trips are made by transit, the region-to-region transit trip matrix will be even less reliable, as a thinner distribution of trip interchanges is expected.

Within the GTHA, the City of Toronto is the largest urban area, with an established Central Business District (CBD). Its neighboring regions of Halton, York, Peel and Durham feature independent municipalities. These regions function more or less as suburbs of Toronto. However, the Hamilton region is farther away from the City of Toronto and is also an established urban area. Almost all Origin-Destination pairs of the City of Toronto, Peel Region and Halton Region have more than 1100 peak period trips between them. Hence, a 4% sample for these areas should be sufficient to adequately model trip behavior. However, in the case of the City of Hamilton and the Region of Durham and York, the majority of origin-destination pairs have less than 1100 trips in the peak period. If only peak period transit trips are considered, then the numbers are even worse.

FROM: \ TO:	CITY OF TORONTO	REGION OF DURHAM	REGION OF YORK	REGION OF PEEL	REGION OF HALTON	CITY OF HAMILTON	REGION OF NIAGARA	REGION OF WATERLOO	CITY OF GUELPH	COUNTY OF WELLINGTON	TOWN OF ORANGEVILLE	CITY OF BARRIE	COUNTY OF SIMCOE	CITY OF KAWARTHA LAKES	CITY OF PETERBOROUGH	COUNTY OF PETERBOROUGH	CITY OF ORILLIA	COUNTY OF DUFFERIN	CITY OF BRANTFORD	COUNTY OF BRANT	TOTAL
CITY OF TORONTO	510,000	7,300	62,600	47,900	5,000	900	400	600	200	*	*	300	400	*	100	*	*	*	100	*	636,000
REGION OF DURHAM	51,800	72,800	15,900	3,200	300	100	*	200	*	*	*	*	100	400	800	100	*	*	*	*	145,900
REGION OF YORK	123,600	4,300	127,800	20,500	1,500	300	100	400	100	*	*	700	1,800	100	100	*	100	*	*	*	281,500
REGION OF PEEL	92,600	700	18,200	188,600	15,900	1,400	500	1,200	600	200	600	100	700	*	*	*	100	200	100	*	321,800
REGION OF HALTON	28,000	200	2,800	37,700	52,300	7,900	800	2,000	700	400	*	100	100	*	*	*	*	*	200	100	133,300
CITY OF HAMILTON	4,600	100	600	6,100	21,000	72,700	2,600	2,000	500	*	*	*	*	*	*	*	*	*	1,600	400	112,300
REGION OF NIAGARA	1,300	100	100	1,000	3,800	6,300	75,900	100	100	*	*	*	*	*	*	*	*	*	100	*	88,700
REGION OF WATERLOO	1,300	100	200	3,100	1,800	1,200	100	104,400	7,600	1,300	*	*	100	*	*	*	*	*	800	400	122,400
CITY OF GUELPH	800	*	200	1,400	1,500	400	100	3,400	21,000	1,700	100	*	*	*	*	*	*	*	*	*	30,800
COUNTY OF WELLINGTON	400	*	200	1,900	1,100	100	*	2,100	4,200	3,600	300	*	*	*	*	*	*	*	*	*	14,000
TOWN OF ORANGEVILLE	400	*	400	3,400	200	*	*	*	100	100	2,100	*	100	*	*	*	*	600	*	*	7,400
CITY OF BARRIE	1,600	100	2,800	1,000	100	*	*	*	*	100	*	17,500	5,400	*	*	*	600	*	*	*	29,300
COUNTY OF SIMCOE	4,000	400	8,300	3,000	200	100	*	100	*	100	300	10,500	29,300	*	*	*	3,600	300	*	*	60,300
CITY OF KAWARTHA LAKES	300	2,600	500	100	*	*	*	*	*	*	*	100	200	8,600	1,500	300	100	*	*	*	14,500
CITY OF PETERBOROUGH	100	700	*	*	*	*	*	*	*	*	*	*	*	500	12,400	2,100	*	*	*	*	15,800
COUNTY OF PETERBOROUGH	100	500	*	*	*	*	*	*	*	*	*	*	*	400	5,700	2,000	*	*	*	*	8,800
CITY OF ORILLIA	*	*	100	*	*	*	*	*	*	*	*	500	1,200	*	*	*	3,400	*	*	*	5,300
COUNTY OF DUFFERIN	300	*	200	2,000	200	*	*	100	100	200	1,400	*	400	*	*	*	*	1,700	*	*	6,600
CITY OF BRANTFORD	200	*	100	300	600	2,100	100	1,200	200	*	*	*	*	*	*	*	*	*	12,000	2,400	19,200
COUNTY OF BRANT	100	*	*	100	300	900	*	1,700	100	*	*	*	*	*	*	*	*	*	2,800	2,400	8,500
REGION TOTALS	821,600	90,000	240,900	321,400	105,900	94,400	80,700	119,400	35,600	7,700	5,000	29,900	39,900	10,200	20,700	4,600	7,900	3,000	17,800	5,700	2,062,400

NOTE: All numbers have been rounded to the nearest one hundred. As a result, there may be some discrepancy in the column and row totals.

FIGURE 1: PEAK PERIOD TRIP MATRIX OF 2011-2012 TTS

In order to further investigate how well the 5% TTS sample represented the whole population, the Root Mean Square Error (RMSE %) was used, as proposed by Kish (1965) and NCHRP (2008):

$$Percent\ RMSE\ = \sqrt{\frac{1}{n_i}\sum_{1}^{n_i}\frac{1}{n_{ji}}\sum_{1}^{n_{ji}}\left(\frac{r_{ij}-s_{ij}}{r_{ij}}\right)^2}\ x\ 100, \quad (1)$$

where

n_i is the number of variables

n_{ji} is the number of category j in variable i

r_{ij} is the reference value of variable i in category j

s_{ij} is the sample value of variable i in category j

As it is clear in this equation, the higher the error for a particular variable, the higher is its representation bias of the whole population. The socio-economic and household specific variables were selected that were common between 2011 TTS and the 2011 census, considering census data as a reference. The following variables were used to estimate the RMSE of 2011 TTS data:

- Number of males
- Number of females
- Number of employed people
- Number of unemployed people
- User of modes:
  - Private car driver; Private car passenger; Transit users; Pedestrians; Bicycle users and Other mode users
- Age groups:
  - Under 14 years; 14+, up to 24 years; 24+, up to 44 years; 44+, up to 64 years and 64+ years
- Household sizes:
  - 1 person; 2 persons; 3 persons; 4 to 5 persons and 6 or more persons

Figure 2 presents the results of 6 cities in the GTHA. The cities are Toronto, Hamilton, Mississauga (Peel Region), Brampton (Peel Region), Oshawa (Durham Region), and Markham (York Region). Please note that the RMSE estimation of the cities and variables was dependent on variable availability, and the commonality of spatial boundaries in both TTS data and corresponding census data (Stats Can 2011). It is also important to note that the 2011 TTS featured a consistent sampling rate of approximately 5% across all regions. On the other hand, the 1991 TTS adopted a differential sampling rate distinguishing between "high growth" and "low growth" areas, where the former was sampled at a 4.5% rate and the latter at 0.5%. The mean sampling rate of the 1991 TTS was 1.4% (DMG 2012).

The majority of the RMSE was below 20% for both 2011-2012 and 1991 TTS. In other words, TTS data represented its target population with an 80% accuracy margin. Part of the 20% error margin was germane to its sampling frame (land line phone directory), which could not be eliminated by simply increasing

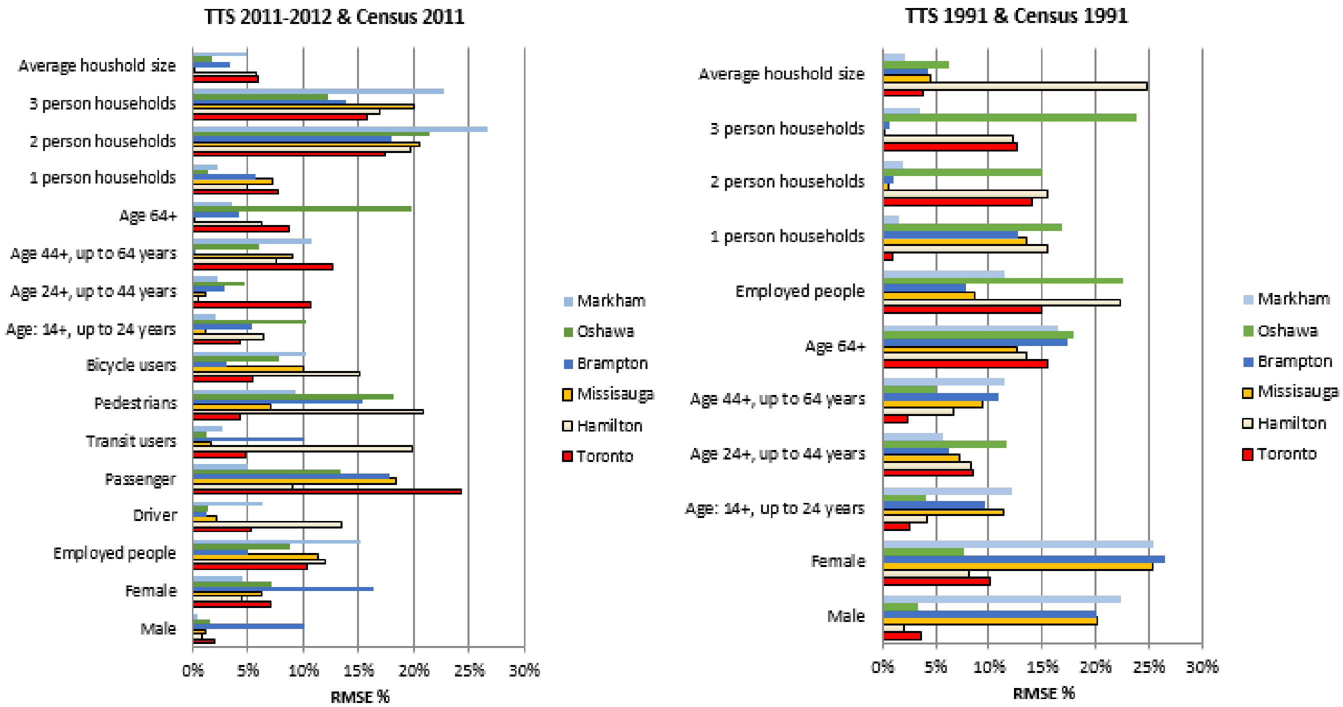

FIGURE 2: RMSE OF SELECTED AREAS IN THE GTHA

sample size. Results showed that non-motorized modes and transit modal shares had a higher error percentage than private automobile use. Error dispersion was higher in 1991 for cities other than Toronto. This may have been due the adoption of a differential sampling strategy in 1991. Since the 1991 census did not capture modal share, it was not possible to assess the accuracy of the 1991 TTS data. It was also difficult to assess the adequacy of the TTS by comparing it to the census due to the incompatibility of the spatial units of analysis in use. In other words, the traffic analysis zones (TAZ – special units of analysis) used in the TTS, did not map on a 1 to 1 scale with the census tracts of the Census, thus adding a second element of bias. Nevertheless, it seems that a 5% sample can produce data representing the target population with more than a 80% level of accuracy. However, it appears that sample sizes of TTS can be improved for better representation. The following section presents a recommendation on this.

5 RECOMMENDATION FOR THE TTS

The Canadian regions have been uniquely and consistently conducting large scale household travel surveys compared to other locations around the world. The GTHA has been consistently conducting 5 percent household sample survey over last three decades in regular 5-year time intervals. However, with increasing under-coverage concerns with the land telephone-based sampling frame and the increasing cost per completed survey responses, it has become imperative for the GTHA to identify a more efficient sample size than current practice of 5 percent sampling.

Leveraging the lessons learned from the review conducted in this report on sampling sizes, and an empirical investigation on previous TTS datasets, a compound sampling approach is proposed that allows the regions/cities in the GTHA the flexibility to acquire sufficient data needed to maintain a good representation of the target populations. The GTHA has been able to bear the cost of a 5% TTS survey over the last three decades. There is an intention to accommodate a similar percentage in the future, notwithstanding the increasing costs. However, it is necessary that an expensive survey should provide a representative sample, while meeting the data needs of individual regions/cities within the GTHA. As a result, it is proposed that a core-augment approach should be followed for survey sample size determination. The core refers to a common base sampling rate of 4% to develop a level ground for trips among the sub-regions of the entire study area. However, such a sampling rate will not be sufficient for the sub-regions with small population sizes, as certain travel modes or segments of the community will be severely underrepresented. Augment samples should be used to complement the data needed for travel demand modelling, specifically that of trip destination choice and travel mode choices for disaggregate activity-based models.

Having a common core sample is also important to acquire a general understanding of the market shares of different competing modes in the region, especially in the absence of a mandatory census long form (that collects household travel information from a representative sample of the whole population) in Canada. The recommendation of reducing the core sample from 5% to 4% is based on the understanding that an additional core of 1% does not necessarily provide a better representation than custom made region-specific augment samples.

Finally, we recommend that the core 4% sample should be based on a stratified random sampling approach to maintain a good representation of population cohorts (age, gender, income, household car ownership, employment, occupation, etc.). Augment samples can be choice-based, if specific mode users are under-represented, or attribute-based augmentation, if general socio-economic cohort representation is of concern. Different regions/cities within GTHA can define their augment samples based on their transportation planning priorities (e.g. additional sample of transit riders, non-motorize mode users, low-income people, no-car households, students, empty nesters, etc.). This will provide regions/cities within the GTHA the flexibility in meeting their individual data needs, in addition to having a common database across the GTHA. The approach of Smith (1979) and Ampt and Ortuzar (2007) are recommended for the augment sample. An empirical exercise is presented in the following section.

6 EMPIRICAL INVESTIGATION ON THE CALCULATION OF SAMPLE SIZE FOR AN ATTRIBUTE-BASED AUGMENT SAMPLE

Smith (1979) proposed a heuristic to calculate sample size for household travel surveys. Smith's approach was regenerated by Stopher (1982) for different area zones in Michigan for a population of 1.6 million households, only to conclude that the approach is adequate to estimate sample size for trip generation purposes. Ortuzar (2004) took sample size estimation one step further by developing an optimization strategy to calculate sample sizes required for the trip generation estimation of household surveys. His strategy, which is a multi-stage stratified random sampling technique, calculates an even smaller sample size by ordering socio-demographic variables by zone, and subsequently selecting a random sample from these zones. Smith's approach was chosen over Ortuzar's advancement, however, due to its simplicity, permitting reproducibility among the different planning agencies. Furthermore, Ortuzar's optimization of size and number of zones sampled may not be of great benefit to organizers who have the objective of designing an attribute (versus spatial) -based augment sample.

Consequently, leveraging the simplicity and adequate accuracy of the Smith's method, the sample size for two major cities (City of Toronto and Region of Hamilton) was calculated to provide an example for the proposed augmented sample recommended in this report. The variables were chosen arbitrarily for the sake of demonstrating the calculation method and the expected results of an augment sample. The augment sample is intended to play one of two roles: to provide statistically reliable data on variables that were not considered in the core household survey; and, to compensate the underrepresentation of certain socio-demographic or trip-characteristic variables that are already present in the core survey. Examples of the latter objective for the TTS include the underrepresentation of young adults, and/or the thin public transit counts present in a number of smaller sized planning districts in the GTHA.

The 2011 TTS data was used for the estimation process. Households in each city were stratified in accordance with three major criteria: household size, income status and number of vehicles. This is an addition to the original Smith approach, which only focused on two variables. The planning organization of interest can further expand the number and vary the selection of variables as per their pre-determined objectives. A 95% confidence interval was adopted along with a 5% acceptable error margin. Following the stratification, the sample size for each class was calculated. The sample sizes were then summed up to provide the total number of household random sample required to estimate statistically reliable trip generation rates.

The following steps outline Smith's method for estimating the augment sample:

1- Stratify survey or trip data in accordance to individual or household socio-demographic or trip attributes.

2- Select variables of interest and designate every category representing a possible combination of the selected variables a number.

3- Calculate average trip rate per category.

4- Calculate the standard deviation of the average trip rate for each category.

5- Determine the coefficient of variation (CV) per category by dividing the standard of deviation by the overall average trip rate.

6- Estimate the frequency of each category within the sample.

7- Normalize the CV by multiplying it with the category frequency.

8- Sum up step 7 to determine C^*

9- Choose a desired level of accuracy (e.g. 5%) and a corresponding confidence interval (e.g. 95%) in reference to 90% confidence and 25% error of the core 4 percent sample base

10- Calculate the initial sample size using the following equation:

$$F = C^{*2} \frac{Z^2}{E^2}, \qquad (2)$$

11- Divide the categorical factors calculated in step 7 by their sum C^* to determine the weight of every category.

12- Multiply the weight of every category by the initial sample size calculated using formula number (2) to determine the optimal allocation of sample size by category.

13- Multiply each category frequency by the initial sample size to determine the expected frequency of the sample to be surveyed.

14- Identify the critical category in which the sample has the highest shortfall in estimating sample size. This may be determined by comparing the CVs and identifying the one with the largest value.

15- Divide the optimally allocated sample (step 12) of the critical category by its expected frequency sample distribution (step 13). This identifies the shortfall percentage.

16- Multiply the expected frequency sample distribution by the percentage increase identified in step 15 to calculate the required representative full random sample.

The resulting augment sample size for the city of Toronto is 5344 households, which is around 0.5% of Toronto's household population. This number is more than fourfold the recommended sample size by Smith. This is mainly due to the stricter confidence interval and associated error margin. Nonetheless, Toronto is a vibrant mobility hub with population of 2.6 million residents of various socio-demographics and trip behaviours, and is a central city to its neighboring region suburbs. Thus, a larger sample size may be required to capture the variation between the pre-specified classes. On the other hand, the total sample size required for Hamilton is 1255 households, which is around 0.6% of Hamilton's population. This is equivalent to the upper end of sample sizes suggested by Smith.

7 CONCLUSION AND SUGGESTION FOR FUTURE RESEARCH

The report examined the issue of sample size determination for household travel surveys. An extensive review of existing literature revealed varying global practices. Only Canadian regions have been able to maintain large sample household travel surveys, while most other countries have faltered, instead choosing a small sampling rate, or switching to new approaches like continuous surveys. The Toronto and Montreal areas are prime examples of cities implementing large cross-sectional surveys. While the move towards small sample household travel surveys is mainly driven by budget limitations, theoretical justification have not been necessarily neglected. However, although small sample sizes are theoretically acceptable, the approach often fails to provide sufficient data for long-term trend analysis and disaggregate travel demand modelling. The Canadian examples have proven that, even with the increasing cost of implementing surveys, it is possible to maintain large sample household travel surveys. However, the review and empirical exercises revealed that the practice of a common percentage of population as the basis for sample size determination may not be very efficient. It is clear that even with a 5% random sample, many regions/cities within the study may have lower than the necessary data points for statistically justified analyses. Empirical investigation revealed that even a 5% sample could have a bias of over 15% in representing basic population cohorts and attributes.

Therefore, a core-augment approach of sample size determination has been proposed in this report for household travel surveys. The proposed approach suggests a core 4% common to all regions/cities of the study area, defined as a percentage of the population. In addition, member regions/cities can implement custom-made augment surveys specific to their data needs. It is recommended that the core common sample follows a stratified random approach to have a good representation of the whole population of the study area. The size of the core sample should be statistically able to provide adequate (e.g. defined acceptable errors and confidence limits) population representation while factoring budget limitations. For the case of the Greater Toronto and Hamilton Area Transportation Tomorrow Survey, it is suggested that the current practice of common 5% sample can be replaced by a 4% core plus individual regions/cities specific augment samples.

This report investigated solely the issue of sample size requirements of household travel surveys without considering the issues of survey cost, sampling frame and continuous versus one-off cross-sectional survey choice. While budget limitations are an unavoidable reality, it is important to investigate the direct and indirect benefits of large scale household travel surveys, including potential future money savings from limiting the implementation inefficient infrastructure investments. Such savings can offset and justify the high cost of large scale surveys. Moreover, identifying an appropriate sampling frame is another critical factor that can inhibit the representation of large scale household travel surveys. It is necessary to investigate whether any innovative or hybrid sample frames and survey modes (e.g. smart phone, GPS, etc.) can further reduce the base cost of household travel surveys.

8 BIBLIOGRAPHY

Ampt, E.S., Ortuzar, J.d.D. 2004. On best practice in continuous large-scale mobility surveys. Transport Reviews 24(3): 337-363

Badoe, D.A., Steuart, G.N. 2002. Impact if interviewing by proxy in travel surveys conducted by telephone. Journal of Advanced Transportation 36(1): 43-62

Bolbol, A., Cheng, T., Tsapakis, I., Chow, A. 2012. Sample size calculation for studying transportation modes from GPS data. Procedia-Social and Behavioural Science 48: 3040-3050

Bonnel, P. and J. Armoogum (2005) National Transport Surveys – What can we learn from international comparisons? Paper presented at the European Transport Conference, Strasbourg, October.

Data Management Group: DMG. 2015. http://www.dmg.utoronto.ca/transportationtomorrowsurvey/

Goulias, C., Pendyala, R., Bhat, C.R. 2013.Total design data needs for the new generation large scale activity microsimulation model. In *Transport Survey Methods: Best Practice for Decision Making*, J. Zamud, M. Lee-Gosselin, J.A. Carrasco and M.A. Munizaga (eds). Emareld Group Publishing, UK

Greaves, S. P., Stopher, P.R. 2000. Creating a synthetic household travel and activity survey: Rationale and feasibility analysis. Transportation Research Record 1706: 82–91.

Harvey, A.S. 2003. Time-space diaries: merging traditions. In: Stopher, P., Jones, P. (Eds.), Transport Survey Quality and Innovation. Pergamon Press, pp. 151–180.

http://dc.chass.utoronto.ca/cgi-bin/census/2011nhs/displayCensus.cgi?year=2011&geo=da

JRC. 2013. Analysis of National Travel Statistics in Europe. JRC Technical Report. https://ec.europa.eu/jrc/sites/default/files/tch-d2.1_final.pdf (accessed in July 2015)

Kish, L. 1965. Survey Sampling. John Wiley & Sons, New York.

Kollo, H.P.H., Purvis, C.L. 1984. Changes in regional travel characteristics in the San Francisco Bay Area: 1960-1981. Paper presented at 1984 Transportation Research Board Annual Meeting, Washington DC, January 1984

National Cooperative Highway Research Program: NCHRP. 2008. Standardized procedure for personal travel surveys. Transportation Research Board. Washington DC

Ohnmacht, R., Rebmann, K., Brugger, A. 2012. Swiss microcensus on mobility and transport. Paper presented at the 12th Swiss Transport Research Conference, Monte verita, May 02-04, 2012

Ortuzar, J.d.D. 2004. Travel survey methods in Latin America. Paper presented at the 7th International Conference on Survey Methods in Transport. Costa Rica 2004. http://www.isctsc.cl/archivos/2004/Keynote%20paper%20Ortuzar%20rev.pdf (Accessed in July 2015)

Ortuzar, J.d.D., Armoogum, J., Madre, J.L., Potier, F. 2011. Continuous mobility surveys: The state of practice. Transport Reviews 31(3): 293-312

Pointer, G., Stopher, P.R., Bullock, P. 2004. Monte Carlo simulation of household travel survey data for Sydney, Australia: Bayesian updating using different local sample sizes. Transportation Research Record 1870:102-108

Richardson A. J., Ampt, E. S. and Meyburg, A. 1995. Survey Methods for Transport Planning (Melbourne: Eucalyptus).

Smith, M.E. 1979. Design of small sample home interview travel surveys. Transportation Research Record 701: 29-35

Southeast Florida Transportation Council: SEFTC. Southeast Florida Household Travel Survey. 2015. Modeling and Regional Transportation Technical Advisory Committee.

Statistics Canada: StatsCan. Census of Canada. 2011. From Canadian Census Analyser.

Stopher, P. 1982. Small sample home-interview travel surveys: Application and suggested modification. Transportation Research Record 886: 41-47

Stopher, P. Zhang, Y., Armoogum, J., Madre, J-L. 2011. National household travel surveys: The case for Australia. Proceedings of 2011 Australian Transport Research Forum Conference, 28-30 September 2011, Adelaide, Australia.

Stopher, P., Greaves, S.P., 2007a. Household travel surveys: Where are we going? Transportation Research Part A 41: 367-381

Stopher, P., Greaves, S.P., 2007b. Guidelines for samples: measuring a change in behaviour from before and after survey. Transportation 34:1-16

Stopher, P., Kockelman, K., Greaves, S.P., Clifford, E. 2008. Reducing burden and sample sizes in multi-day household travel surveys. Paper presented at the 87th Annual Meeting of the Transportation Research Baord. Washington DC, January 2008

Stopher, P., Meyburg, A.H. 1979. Survey sampling and multivariate analysis for social scientists and engineers. Heath, Lexinton, MA, USA

TransLink. 2010. 2008 Regional Trip Diary Survey Final Report, http://www.translink.ca/-/media/Documents/plans_and_projects/trip_diary/2008%20TransLink%20Trip%20Diary%20Survey%20Report.pdf (Accessed in Sep 2015)

Travel Model Improvement Program: TMIP. 1996. Travel Survey Manual. Prepared by Cambridge Systematic Inc. for US DOT, FTA, FHA, OoC and US EPA. www.fsutmsonline.net/images/uploads/mtf-files/Southeast_Florida_Household_Travel_Survey_0205_2014.pdf