11th International Conference on Transport Survey Methods

Impact of Multiple Sample Frames on Data Quality of Household Travel Surveys: The Case of the 2016 Transportation Tomorrow Survey

Albert Lo, Siva Srikukenthiran*, Meishi Chen, Khandker N. Habib, Eric J. Miller

University of Toronto, Civil Engineering Department, 35 St. George St., Toronto M5S 1A4, Canada

Abstract

The Transportation Tomorrow Survey (TTS) is a household travel survey conducted in the Greater Golden Horseshoe Area (GGH) every 5 years. To attempt to alleviate underrepresentation due to decline in landline use, the 2016 iteration of the TTS used multiple sample frames consisting of address/phone, address-only, and phone-only. The sample frames were compared to the census on variables such as age and sex. The phone-only sample frame did little to increase representativeness of the dataset due to a small sample size. Respondents of the address-only sample frame were more likely to respond by web than the other frames.

Keywords: Multiple sample frame; Household travel survey; Transportation tomorrow survey; Data quality

1. Introduction

Since 1986, the Transportation Tomorrow Survey (TTS) has surveyed five percent of households in the Greater Golden Horseshoe area (GGH) every five years (DMG, 2014). It is a household travel survey that collects information such as the mode of travel and trip distance, along with the demographics of household members. Past iterations of the TTS have used a landline sample frame where notification letters were sent to selected households. Representatives of these households could call or, for the 2011 iteration, respond online to the survey. While this methodology has been sound in the past, the representativeness of recent iterations has declined with the reduction of household landline use and rise of cellphone use. Improving demographic representation of the travel data collected in the GGH is critical to properly design and implement infrastructure plans and projects to reflect the travel behaviour of all segments of society.

In response to these issues in the landline frame, the most recent TTS conducted from September to December 2016 sampled households with a multiple sample frame approach (Malatest, 2017). These sample frames include the address/phone, address-only, and phone-only. Address/phone samples corresponded to households with both a known address and phone number, whereas address-only and phone-only corresponded to households where only an address and a phone number were known, respectively. The phone-only sample frame consisted of a

* Corresponding author. Tel.: +1-416-946-8299
E-mail address: siva.srikukenthiran@mail.utoronto.ca

white page listing, random digit dialling (RDD), and a verified cellphone sample.

The objective of this paper is to present the respondent profiles and analysis of the TTS 2016, with a focus on the multi-sample frame approach. The survey data is compared to the 2016 census to examine differences in proportions of shared variables. A similar comparison is conducted for each of the three sample frames. This analysis is done from both the perspective of the entire GGH, as well as separately for sub-regions. The paper acts to inform about recent demographic trends in the region and improvements made in representativeness of the sample as part of an overall set of recommendations to improve future TTS iterations. A general insight into the impact of using multiple frames is provided as well.

2. Current state of surveys using multiple survey frames

Several large-scale surveys have been conducted in the past with multiple sample frames, specifically in the public health and household travel survey fields. Use of this approach has the potential of increasing survey data representativeness by recruiting survey respondents through different channels and possibly reducing coverage bias. This section investigates the effectiveness of these surveys' methodologies in terms of their response rates and representativeness to the population.

The dual cellphone and landline sample frame utilizes landline numbers with a supplemental cellphone phone list. Phone numbers for both sample frames can be obtained from commercial vendors and/or RDD. Depending on information associated with the phone numbers, one can sample by stratifying the phone numbers by geography and/or certain demographics. With decline in landline use, a cellphone sample frame is included in the sampling design to attempt to reduce underrepresentation. This multi-sample frame has been used in several surveys in multiple fields around the world.

In Australia, surveys investigating drug use, gambling habits, and personal health were provided phone numbers by third parties (Livingston, et. al., 2013; Barr, et. al., 2012; Jackson, et. al., 2013). Landline numbers were sampled by geographic stratification, specifically size quotas by cities. Cellphone numbers were selected randomly as there were no geographic identifiers. Using the American Association for Public Opinion Research (AAPOR) response rate formula, response rates ranged from 16% to 33%. AAPOR's response rate formula not only takes into account completed responses and total samples sent, but also includes partial completions, refusals, non-contact among other measures. The response rate for the cellphone frame was about 10% lower relative to the landline frame. These surveys found that a combined dual frame sample was more representative of census data than either sample frame alone. With respect to demographics, cellphone-only samples were more likely to be young, male, and living in metropolitan areas.

In the United States, medical accessibility and health insurance surveys have in the past typically only used a RDD landline sample frame approach. Surveys like the Behavioural Risk Factor Surveillance System (BRFSS) and the Minnesota Health Access Survey (MHAS) now currently include a cellphone sample frame in addition to the landline to reduce coverage bias (Hu, et. al., 2010; SHADAC, 2013; Lu, et. al., 2013; Lee, et. al., 2010; Abt Associates Inc., 2012). This notion of coverage bias arises from prior studies that have shown that US cellphone-only households are more likely to have household members being young adults of black and/or Hispanic ethnicity and have lower income compared to landline households.

Selecting cellphone samples in these US health-related surveys were either stratified geographically or randomly sampled. Most surveys implemented a screening regimen to filter for cellphone-only households or filter out landline households with only seniors for a certain timeframe (MHAS). Monetary incentives were provided for completion of surveys. Specifically for the BRFSS, cellphone users were more likely to be male, 18-34 years old, not married, working or not retired, non-Hispanic white, and to have an annual household income less than $35,000 compared to the landline group. Response rates ranged from 15% to 40% based on AAPOR's response rate formula.

Prior household travel surveys have used multiple sample frames in their sampling design. Both the Chicago Regional Household Travel Inventory and the California Household Travel Survey used address-based sampling frames (address/phone and address-only), with the former also using a RDD landline sample frame (NuStats, 2007; NuStats, 2013). The idea of using address-based sampling was to capture 'hard-to-reach' populations, as traditionally RDD landline samples have a disproportionally high number of upper income homeowners. Conversely, 'hard-to-reach' populations (in an American context) are Hispanic / Black, young population and low income. In both cases, oversampling and monetary incentives were used to increase completion rates with these 'hard-to-reach' households. Advanced notification mailing was used for the address-based sampling frame and respondents were able to answer the survey via online, calling, and an optional GPS tracker. Both surveys achieved seemingly low response rates of 11% and 4.9% for Chicago and California respectively. The Council of American Survey Research Organizations or CASRO response rate formula was used which includes the number of samples with unknown eligibility along with the number of completed surveys. For reference, the 2000-2001 iteration of the California Household Travel Survey was 20.0% (NuStats, 2002).

Others forms of multi-sample frames use an email sample frame in conjunction with the typical landline sample frame. A study done by Verreault and Morency used an email sample frame to exclusively target post-secondary students in four post-secondary institutions within Sherbrooke, Québec (Verreault & Morency, 2016). The purpose was to augment the surveyed '20 to 29' year old population of the landline travel survey done in parallel as this age category was under sampled in the landline sample frame. The students, responding to a web survey, achieved a 10% response rate and were concentrated among 20-24 year old females. Workers in the same age group were underestimated in the combined landline and email sample frame dataset.

The success of surveys using multiple sample frame to correct deficiencies in the individual frames have been variable. The purpose of including a cellphone, address-based, email or other type of sample frame is generally due to a potential bias of non-coverage if only a landline sample frame is used. It is suggested that, particularly in the United States, that cellphone-only households are significantly different from the landline population and thus should be included to the landline sample frame to have a more representative sample. From a Canadian perspective, 21% of households were cellphone-only in 2013 (Statistics Canada, 2014). Canadian cellphone-only households tended to have all household members being less than 35 years old. Other efforts to increase representativeness include incentives for 'hard-to-reach' populations and screening out samples based on set criteria.

Specific to the TTS, the survey relied on a landline sample frame up to the 2016 iteration. Underrepresented populations have included males between the ages of 18 and 32, as well as post-secondary students. This paper adds to this literature, showing results of using a multiple sample frame to collect household travel information in the GGH.

3. Data

3.1 Greater Golden Horseshoe context

The GGH is located in southern Ontario around the western portion of Lake Ontario. The area includes the major cities Toronto, Hamilton and Mississauga within the Greater Toronto-Hamilton Area (GTHA), as well as the surrounding counties and regions, including the Region of Waterloo, as shown in Figure 1. The GTHA is often separated into the City of Toronto and the rest of the regions, known as the 905 area. Transportation infrastructure in the GGH includes the 400 series highways and several regional and municipal transit systems such as the Toronto Transit Commission (TTC). The area contains over 9 million people with 3.3 million households and has seen a growth of 8% in population since 2011. The largest growth percentage has been in the extended regions, especially in Dufferin (73%) and Simcoe (37%) counties. The GTHA has experienced a growth of 400,000 people since 2011 with a quarter of this growth being in the City of Toronto.

3.2 Dataset description

The data consists of frequency tables from 2011 and 2016 TTS and the Canadian census. The TTS data is unexpanded in this analysis and the census is used as the reference population. Variables include dwelling type, respondent status, age, sex, household size amongst others. For the first time in the TTS, household

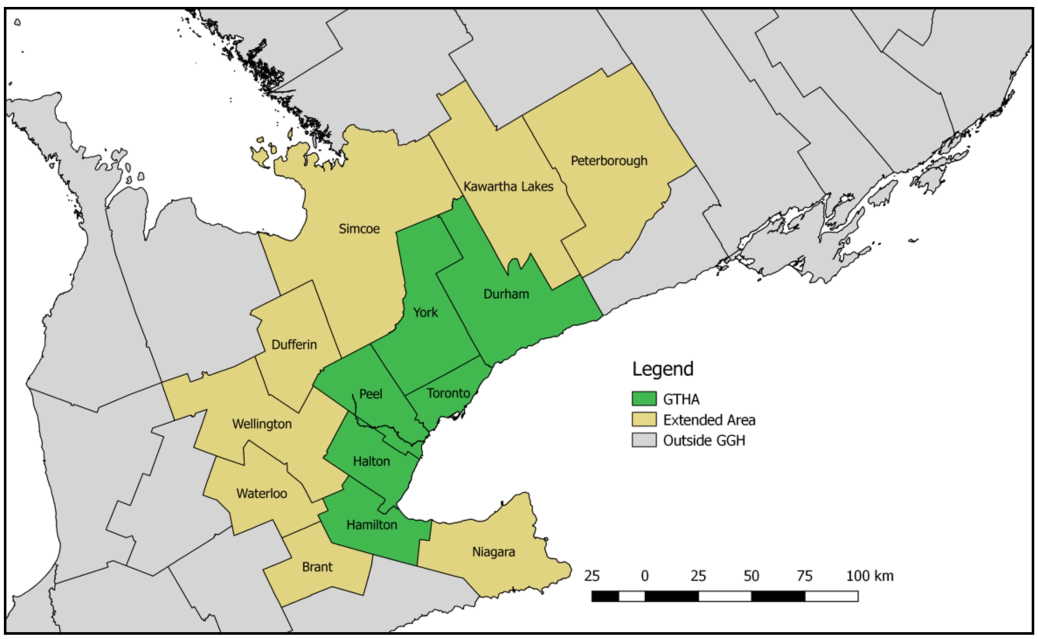

Fig. 1. The GTHA and GGH

income was included in the 2016 survey. In terms of data integration and cleaning, ARCGIS was used to spatially join latitude and longitude points of each household to the census subdivision (CSD) shape file with the NAD 1983 UTM coordinate reference system. Data that were outside of the GGH and/or were labelled as Indian Reserve were removed.

It is important to note that at the time of writing this paper, the 2016 TTS dataset analyzed was a preliminary version. However, significant deviations between the final and preliminary dataset are not expected. Foreseen changes to the preliminary dataset include implementing transit and school codes to origins and destinations and data weighting. Moreover, census income and transportation related variables will not be available until September 13, 2017 and November 2017 respectively. Lastly, a separate database was provided by MALATEST to display response rates for the TTS 2016 by municipality (CSD).

3.3 Method of data collection

Due to the difference in sample frames, the 2016 TTS method of collection was slightly different from past iterations (Malatest, 2017). Similar to the landline sample frame, selected households in address/phone and address-only sample frames (pulled from a Canada Post mailing address database) were sent a notification letter. These households could respond to the survey via phone, completing it on the web, or waiting for a call (restricted to the address/phone sample frame). Notification letters were sent in batches twice per week to targeted areas. Samples sent to areas where there were many respondents were paused, whereas additional notification letters were sent to areas with low response rates. For the phone-only sample frame, instead of a notification letter, the samples were 'cold-called' and could respond by phone or online. It is noted that there was a potential for overlap between the address-based (i.e. address/phone and address-only) and phone sample frames. The phone-only sample frame was an experiment with the intention of reaching out to households in the address-only frame whom would be reluctant to answer to the notification letter.

The proportions between the three sample frames are shown in Table 1. The low phone-only proportion was a result of a low response rate and high refusal rate from phone-only samples, which led to a decay of interviewer morale (Malatest, 2017). Consequently, phone-only samples were discontinued a week after their introduction. The address-only proportion was high due to the sole method of contact being the notification letter. The inability to follow up with these samples may lead to lower response rate and thus more address-only samples were sent in the field.

Table 1. Sample Frame Proportions

Sample Frame	Total Sampled Households	Proportion
Address/Phone	198,027	20%
Address-only	790,367	79%
Phone-only	10,121	1%

4. Methodology

The analysis conducted in this study involved a comparison of the 2016 TTS and census, as well as a comparison between the 2016 TTS sample frames. The analysis was based on the chi-squared test, data observations, and non-statistical experiments. The comparison was done at the levels of the GGH, GTHA versus the extended area, and Toronto versus the 905 area. This section describes the methods that were used.

It should be noted that data were excluded where appropriate. For example, the volunteer sample frame was taken out when comparing between sample frames (making up less than 1% of households). Another example is the unknown housing type and unknown gender removal when comparing housing type and gender respectively (both less than 1%).

4.1 Chi-squared test

The chi-squared test compares two or more independent samples' distributions of a variable's categories (McHugh, 2013). The test statistic is the summation of squared error differences between each variable and samples as shown in equations (1) and (2). This squared error difference is between the observed value and the expected value, the latter being the value if there was no relationship between the samples and variables.

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \tag{1}$$

where
n is the number of cells in the table*
O_i is the observed value for cell i
E_i is the expected value for cell i

$$E_i = M_{r_i} \times \frac{M_{c_i}}{\sum_{i=1}^{n} O_i} \tag{2}$$

where
M_{r_i} is the row marginal sum of the counts for cell i
M_{c_i} is the column marginal sum of the counts for cell i
n is the number of cells in the table*
O_i is the observed value for cell i

*The table refers to nominal frequency data arranged by the variable's mutually independent categories of the independent sample groups

The test-statistic is compared to a chi-squared distribution with the same degree of freedom (calculated as the product of number of rows subtract 1 and number of columns subtract 1) to attain a p-value in order to determine significance. Assumptions of the chi-squared tests include independent sample groups and nominal level of measurements (i.e. categories) for the frequency data. The paper suggests pairing the chi-squared test with a measure of association to determine the strength of the association between the independent samples and categories.

There are several association measures including phi, contingency coefficient and Cramér's V (Gingrich, 2004). The first two measures adjust the test statistic with only sample size, whereas Cramér's V adjusts with both sample size and the number of rows and columns. The most common association measure used is Cramér's V, which is shown in equation (3). The measure ranges from 0 to 1 with a value close to 0 meaning little association and a value close to 1 meaning high association.

$$V = \sqrt{\frac{\chi^2}{n \bullet \min(r-1, c-1)}} \tag{3}$$

where
χ^2 is the chi-squared test statistic
n is sample size
r is the number of rows in the table
c is the number of columns in the table

For the implementation of the chi-squared test, the three sample frames and 2016 TTS and census were the independent samples which were compared to the demographic variables. Cramér's V was used to measure association.

4.2 Root mean squared error

According to the 2016 NCHRP (National Cooperative Highway Research Program) report on standardized procedures for personal travel surveys, data analysis of household travel surveys should include determining sample bias between the collected survey data and the census (NCHRP, 2016). To this end, the report suggests using the following root mean squared error (RMSE) formula shown in equation (4).

$$RMSE = \sqrt{\frac{1}{n_i}\sum_{1}^{n_i}\frac{1}{n_{ji}}\sum_{1}^{n_{ji}}\left(\frac{r_{ij}-s_{ij}}{r_{ij}}\right)^2} \times 100 \tag{4}$$

where
n_i is the number of variables i
n_{ji} is the number of categories 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

Table 2 shows the NCHRP's suggested variables and categories to be used in determining sample bias. The RMSE is used to determine how well the survey data represents the population, along with being a metric to compare against prior iterations' RMSE values. It should be noted that the RMSE values are subject to non-response bias and availability of the data.

For this paper's analysis, only household size, age, and gender were used to calculate RMSE as shown in Table 3 (Habib & El-Assi, 2015). Vehicle availability, household income, and race could not be used. The latter was not asked in the TTS and the census values of the rest were not available at the time of writing this paper. To overcome the discrepancy in dwelling type definitions between the TTS (4 categories) and census (8 categories), the census dwelling type categories were condensed based on their name into the TTS definitions. The exception to this rule was the category 'row house' was associated with 'townhouse' instead of 'house'. Census counts labeled as 'movable dwelling' were removed.

The RMSE was calculated to compare representativeness of the TTS dataset and its three sample frames against the census. To determine representativeness of 'hard-to-reach' populations, namely males between 20 and 29 years of age, an average was taken of RMSE values that were calculated as follows. For each combination of individual dwelling type and household size (15 combinations in total), a RMSE value was calculated with the male gender category and '20-29' age category held constant. In addition, RMSE values were calculated by CSD and presented on a map.

4.3 Percent error maps

Another method of analysis was creating maps in ARCGIS to visually display percent error in representation between the TTS and census. The general equation for percent error is presented in equation (5).

$$\%error = \frac{TTS\_proportion - census\_proportion}{census\_proportion} \tag{5}$$

For each planning district of the GGH, a percent error value was calculated based on a category of a demographic variable. Chosen categories include the proportion of apartments and all proportion of age groups as shown in Table 3. Percent errors were displayed on the maps from -100% to 100% in 5 bins using graduated colour symbol.

Table 2. NCHRP definitions of variables and categories for RMSE formula

Variable	Categories
Household Size	Mean
Vehicle Availability	0, 1, 2, 3+
Household income	Intervals of $10,000 (e.g. Under 10,000, 10,000 – 19,999, to 140,000- 149,999, 150,000 and over)
Race	White, Black, Native, Asian, Hawaiian, other, two or more
Age	0–5, 6–10, 11–14, 15–17, 18–64, 65–74, 75 and over
Gender	Male, Female

Table 3. Current TTS definitions of variables and categories for RMSE formula

Variable	Categories
Household Size	1, 2, 3, 4, 5+
Dwelling Type	House, Apartment, Townhouse
Age	0–19, 20-29, 30-39, 40-59, 60+
Gender	Male, Female

4.4 Response rate

Instead of using CASRO or AAPOR's response rate formulation, response rates were calculated as the number of completed surveys divided by the number of samples in the field. These values were disaggregated by sample frame in order to calculate response rates by sample frame.

5. Results

Table 4 shows the chi-squared test statistic values and their associated Cramér's V. All chi-squared tests are significant for both comparisons of demographic variables to the sample frames and 2016 TTS and census. However, virtually all of the relationships have a weak association with the exception of the relationship between response mode and sample frame.

Table 4. Chi-squared test results by sample frame and by 2016 TTS and census

	Sample Frame		2016 TTS and Census	
	Chi-Squared Test Statistic	Cramér's V	Chi-Squared Test Statistic	Cramér's V
Gender	76.61	0.01	61.55	0.00
Type of dwelling	3240.58	0.10	3055.69	0.03
Household size	248.30	0.03	5810.56	0.04
Response mode	45292.75	0.53	-	-
Income	2014.53	0.08	-	-
Age categories	17401.77	0.15	19750.31	0.05

Table 5 provides a summary of response rates and RMSE values by sample frame and the 2016 TTS dataset. This iteration of the TTS achieved a response rate of 16.3% (998,515 samples in field). Generally, both the GTHA and Toronto exhibit a lower response rate compared to the extended region and 905 area respectively. The RMSE between the 2016 TTS and census is 17.0%. For comparison purposes, the 2011 RMSE is 18.5% (not shown). The TTS 2016 dataset has an average RMSE value of 19.0% for males aged between 20 to 39 years of age.

At the GGH level by sample frame, the response rate for address/phone is high, with a value of 39.6%, while address-only and phone-only are much lower at 10.5% and 8.6%, respectively. The sample frame response rates by disaggregated GGH regions are generally similar to the GGH level. The address/phone, address-only and phone-only frames have RMSE values of 27.3%, 14.1%, and 22.9%, respectively. Across the GGH, its individual geographic disaggregates and the male 20-29 years old category, the address-only sample frame has a lower RMSE value than the other two frames.

Figure 2 shows RMSE values comparing TTS and census 2016 by CSD. Values range from 11% to 84% with the median RMSE value being 23.4%.

Table 5. Response rates and RMSE values by sample frame and TTS 2016

	Address/Phone	Address-only	Phone-only	TTS 2016
Response Rate				
GGH	39.6%	10.5%	8.6%	16.3%
GTHA	38.7%	10.4%	8.2%	16.0%
Extended Region	42.4%	10.8%	9.6%	17.2%
Toronto	39.1%	9.8%	8.0%	14.0%
905 area	38.4%	11.1%	8.3%	17.9%
RMSE				
GGH	27.3%	14.1%	22.9%	17.0%
GTHA	27.3%	14.5%	22.1%	16.3%
Extended region	29.1%	15.4%	28.6%	19.9%
Toronto	36.2%	17.4%	31.0%	19.7%
905 area	23.7%	13.9%	20.1%	16.5%
Average RMSE				
Male 20-29 year olds	31.7%	13.5%	26.0%	21.7%

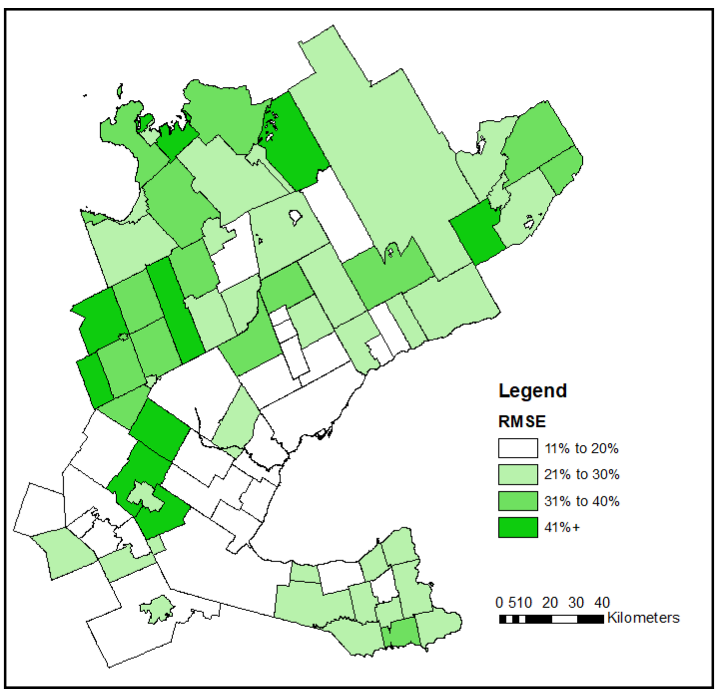

Fig. 2. RMSE by CSD, (GGH)

Table 6 summarises the demographics by sample frame, TTS 2016, and census 2016. Comparing the TTS dataset with the census, discrepancies are observed. Households are more likely to be 2 person households (37.7% TTS to 30.4% census) and less likely to have 5 or more people (7.2% to 11.1%). In addition, there is a smaller proportion of the age category '0 to 39' years old (40.6% to 49.7%), with a corresponding higher proportion of older people 40 and above.

In terms of the respondents, they are likely to be an older member of each household, illustrated by a gap between the median age and median respondent age. 64.1% of respondents answered the survey online from beginning to end. 80.8% of respondents answered the household income question.

Figure 3 shows the percent error of apartment proportions between the 2016 TTS and census by planning district. Percent errors range from -100% (no apartments captured in the TTS for a particular planning district) to 238%. Discarding outliers, the average percent error is -23.1%. Figures 4 and 5 show the percent error of age groups '20-29' and '60+' by planning district. These two particular age groups are presented due to particularly high under and overrepresentation, respectively. The '20-29' age group ranged from -66.6% to 24.3%. The '60+' percent errors are exclusively positive, ranging from 5% to 117%.

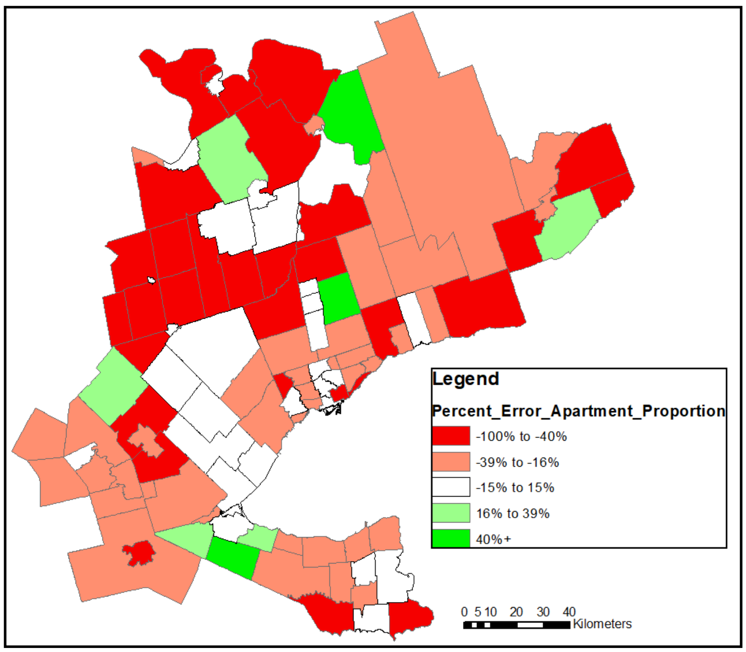

Fig. 3. Percent error of apartment proportions by planning district, (GGH)

The proportion of households by sample frame is as follows: 50.6% for address/phone, 48.7% for address-only, and 0.7% for phone-only. This translated to roughly 80,000 households for both address/phone and address-only, and 1,000 households for phone-only.

The demographics of the individual sample frames show some important differences. First, a median age of 40 in the address-only sample frame matches with the census. The address/phone and phone-only sample

Table 6. Demographics by sample frame, TTS 2016 and census 2016

	Address/Phone	Address-only	Phone-only**	TTS 2016	Census 2016
# of households	82,383	79,256	1,097	162,791	-
# of people	203,004	190,296	2,872	396,302	-
Gender					
Male	47.4% (96,044)	48.8% (92,768)	48.5% (1390)	48.0% (190,256)	48.7% (4,426,030)
Female	52.6% (106,746)	51.2% (97,506)	51.5% (1477)	52.0% (205,805)	51.3% (4,667,240)
Median Age	52	40	45	46	40
Median Respondent age	60	49	53	55	-
Type of Dwelling					
House	67.7% (55,732)	54.1% (42,902)	74.1% (813)	61.1% (99,474)	54.8% (1,827,055)
Apartment	23.8% (19,611)	35.0% (27,714)	17.8% (195)	29.2% (47,545)	35.9% (1,194,585)
Townhouse	8.5% (7001)	10.8% (8575)	8.1% (89)	9.6% (15,667)	9.3% (309,595)
Household Size					
1	24.2% (19,897)	26.6% (21,091)	20.1% (221)	25.3% (41,229)	24.7% (822,055)
2	38.5% (31,753)	36.9% (29,230)	38.2% (419)	37.7% (61,418)	30.4% (1,013,465)
3	15.3% (12,611)	16.3% (12,924)	15.3% (168)	15.8% (25,712)	17.0% (567,380)
4	14.3% (11,802)	13.5% (10,674)	17.0% (186)	13.9% (22,664)	16.8% (560,685)
5+	7.7% (6320)	6.7% (5338)	9.4% (103)	7.2% (11,768)	11.1% (370,865)
Household Response Mode					
Online only	39.8% (32,484)	90.1% (70,698)	14.0% (153)	64.1% (103,335)	-
Phone only	60.2% (49,163)	9.9% (7731)	86.0% (941)	35.9% (57,890)	-
Household Income					
$0 to $14,999	3.8% (3127)	4.1% (3261)	4.6% (51)	4.0% (6443)	-
$15,000 to $39,999	15.8% (13,025)	13.4% (10,623)	14.2% (156)	14.6% (23,815)	-
$40,000 to $59,999	14.5% (11,944)	13.9% (10,986)	14.9% (163)	14.2% (23,102)	-
$60,000 to $99,999	18.5% (15,257)	23.2% (18,348)	19.0% (208)	20.8% (33,818)	-
$100,000 to $124,999	8.5% (7040)	11.1% (8820)	7.4% (81)	9.8% (15,943)	-
$125,000 +	16.2% (13,372)	18.6% (14,761)	18.3% (201)	17.4% (28,341)	-
Decline / Don't know	22.6% (18,618)	15.7% (12,458)	21.6% (237)	19.2% (31,330)	-
Age Categories					
0-19	18.2% (37,039)	20.0% (38,126)	22.6% (648)	19.1% (75,842)	22.7% (2,046,910)
20-29	7.5% (15,184)	11.8% (22,447)	9.1% (262)	9.6% (37,900)	13.6% (1,225,290)
30-39	7.3% (14,860)	16.9% (32,146)	9.7% (279)	11.9% (47,291)	13.3% (1,197,905)
40-59	29.6% (60,138)	29.3% (55,777)	31.4% (902)	29.5% (116,854)	28.9% (2,599,220)
60+	37.3% (75,783)	22.0% (41,800)	27.2% (781)	29.9% (118,415)	21.5% (1,931,315)

**Please note the number of phone-only households is very low compared to the other two sample frames.

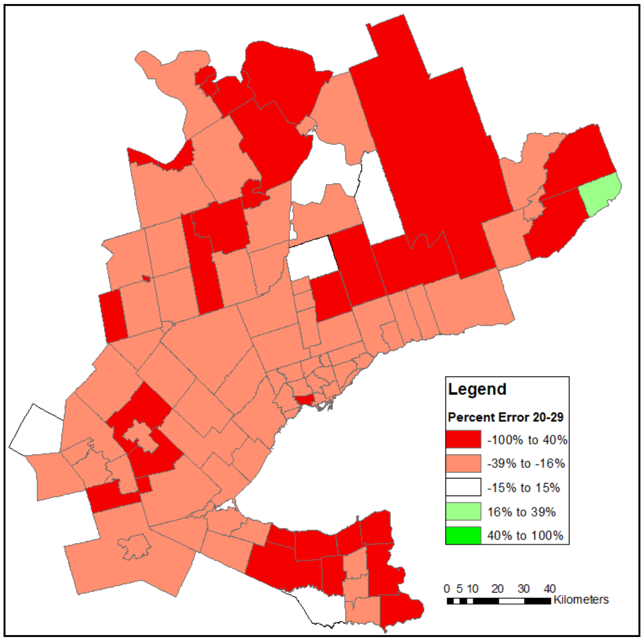

Fig. 4. Percent error of age category 20-29 by planning district, (GGH)

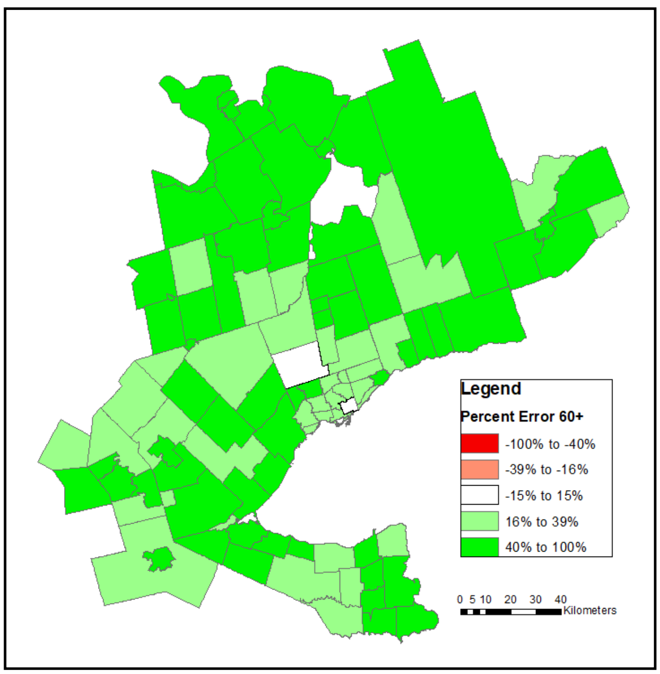

Fig. 5. Percent error of age category 60+ by planning district, (GGH)

frames exhibit median ages' of 52 and 45 respectively, much higher than the census median. In general, median respondent ages are much higher than their population counterparts. The address-only frame has type of dwelling proportions that match fairly well to the census proportions. Calculating RMSE for only the 'type of dwelling' variable yields a value of 9.6% for address-only, 24.1% for address/phone and 36.2% for phone-only. Address-only households are very likely to answer the survey online (90.1%), whereas answering by phone is the most popular method of response for address/phone (60.2%) and phone-only (86.0%). This finding corresponds to the high Cramér's V value found between response mode and sample frame. Address-only households are more likely to contain individuals of age categories 20-39 (28.7%), compared to 14.8% of households in the address/phone and 18.8% of phone-only frames. Lastly, there is a high proportion of 40+ year olds in the address/phone (67.0%) and phone-only (58.6%) frames. Similarities between sample frames are seen in some household characteristics, namely household size and income, where the medium income bracket is $60,000 to $99,999 across all sample frames.

6. Discussion

In terms of demographics, several findings are noteworthy. The smaller proportion of households with 5 or more members can be attributed to respondent fatigue from completing trip information for each household member. It is hypothesized that this translates to the observed low proportion of the younger age groups, illustrated in the older median age and high proportion of 60+ year olds. Conducting a chi-squared test between variables age category and household size yields a chi-squared value of 104,650 and a Cramér's V value of 0.27 which shows a moderate relationship.

Most responses were done online. This is driven largely by households contacted only via mail, namely the address-only frame. One of the interesting results of the address-only frame is it exhibits a higher response rate to the income question than the other sample frames. It is speculated that not being 'put on the spot' during a call to answer household income leads to a higher response rate to the question. On the other hand, households contacted from the address/phone and phone-only sample frames primarily use phone calling as a means to respond, and have a lower response rate to providing their income level. Conducting a chi-squared test between variables response mode and income yields a chi-squared value of 6481.14 and a Cramér's V value of 0.20 which shows a moderate relationship.

The addition of the phone-only sample frame was insignificant to the representativeness of the 2016 TTS dataset. Removing the frame from the data set had negligible impact on RMSE values, lowering it slightly to 16.9% from 17.0% for the full data set with all 3 frames. This insignificance can be attributed to the experimental nature of the phone-only sample frame and its small sample size.

Overall, the dataset is somewhat representative with a 17.0% RMSE, meaning that the TTS 2016 represents the census with an 83% accuracy margin; however, this is not uniform across all variables used in the formula. The gender variable is fairly well represented across all sample frames, GGH and the individual sub sections of the GGH; as a result, it is not a differentiator for comparison between sample frames and years. The driver behind the RMSE value for the collective GGH data is instead a disproportionately high 60+ age category, followed by high and low proportions of household sizes of 2 and 5+, respectively. On the other hand, household size is the main source of error when examining RMSE values of individual frames. This makes sense because each sample frame has a similar proportion of household sizes. Both the address/phone and phone-only sample frames over-represent the 60+ age group and under-represent the 20-29 and 30-39 age groups.

The RMSE map in Figure 2 shows a fair representation of CSDs in the inner core of the GGH and less so for the extended regions, especially CSDs to the north. CSDs with RMSE values in the 31% to 40% and 41%+ categories have large proportion discrepancies for dwelling type, especially apartments and town houses. The 60+ age category and the household size of 5+ has high error as well.

There are limitations to the RMSE formula. Based on the NCHRP guidelines, the RMSE value is calculated in such a way as to treat each category of each variable equally. For example, the squared error of the male gender proportion is valued the same as the squared error of the 20-29 age category. This is fine as a relative measure comparing between years or sample frames. As a measure of error, however, there may be an opportunity to modify the equation or variables to attain a better idea of representation of the reference data.

While in this study the respondent profiles of each frame were compared to the census, it is unknown if the RMSE values are due to the samples chosen of the sample frame or non-response bias. Hypothetically speaking, if samples chosen in the address/phone sample frame were skewed to an older population, the sample frame would therefore be representative against its population. However, comparing this sample frame to the census would obviously result in the sample frame being deemed as weak in representation. Nonetheless, each sample frame can serve a different purpose depending on the need of the researchers.

7. Conclusion

GGH respondent profiles for each sample frame are as follows. The address/phone sample frame captured an older population compared to the census. Samples were more likely to respond to the survey by phone. The address-only sample frame captured a high proportion of 20-39 year olds and apartments which closely matched the census better than the other two sample frames. Samples were very likely to answer by web and more willing to answer the income question. The phone-only sample captured an older population although results are not conclusive due to the small sample size. Samples in the phone-only sample frame were not drastically different from other sample frames.

Using a multi-frame approach, the unexpanded TTS 2016 dataset is somewhat representative of the GGH population. The dataset has a higher proportion of people aged 60+ and a lower proportion of households with 5 or more people compared to the census. The addition of the phone sample did little to change the representation due to its small sample size. Strictly based on the respondent households that were selected in the sample frame, the address-only sample frame results best represent the 2016 census.

Based on this analysis, future research questions include examining the RMSE formula and demographics of the field samples. The investigation of the RMSE may involve modifying the formula to become a better indicator of representativeness error of the survey data to the census. This will include adding additional variables to the RMSE value (namely income) once they become available in the census. In addition, work should be undertaken to determine the demographic profile of each sample frame based on the samples sent in the field. This will help determine whether the RMSE values between each sample frame and the census is a function of the sample frame itself or non-response bias.

Acknowledgements

The study was funded by Transportation Information Steering Committee (TISC) represting all regional and local planning agencies that fund and use the Transportation Tomorrow Survey. Authors acknowledge the guidance, suggestions and comment of the TISC steering committee, Technical Advisory Committee (TAC) members, and the TTS 2.0 group.

References

Abt Associates Inc. (2012). Family and Medical Leave in 2012: Methodology Report. Cambridge: Abt Associates Inc.

Barr, M. L., van Ritten, J. J., Steel, D. G., & Thackway, S. V. (2012). Inclusion of mobile phone numbers into an ongoing population health survey in New South Wales, Australia: design, methods, call outcomes, costs and sample representativeness. BMC Medical Research Methodology, 12:177.

DMG. (2014). Transportation Tomorrow Survey 2011: Design and Conduct of Survey. Toronto: Data Management Group.

Gingrich, P. (2004). Association between Variables. Regina: Gingrich, Paul.

Habib, K. N., & El-Assi, W. (2015). Sample Size Requirements for Regional Household Travel Surveys. Toronto: UTTRI.

Hu, S., Balluz, L., Battaglia, M. P., & Frankel, M. R. (2010). Improving Public Health Surveillance Using a Dual-Frame Survey of Landline and Cell Phone Numbers. American Journal of Epidemiology, 703-711.

Jackson, A. C., Pennay, D., Dowling, N. A., Coles-Janess, B., & Christensen, D. R. (2013). Improving Gambling Survey Research Using Dual-Frame Sampling of Landline and Mobile Phone Numbers. Journal of Gambling Study, 291-307.

Lee, S., Brick, M., Brown, R., & Grant, D. (2010). Growing Cell-Phone Population and Noncoverage Bias in Traditional Rnadom DIgit Dial Telephone Health Surveys. Health Research and Educational Trust, 1121-1139.

Livingston, M., Dietze, P., Ferris, J., Pennay, D., Hayes, L., & Lenton, S. (2013). Surveying Alcohol and other drug use through Telephone Sampling: a Comparison of Landline and Mobile Phone Samples. BMC Medical Research Methodology, 13:41.

Lu, B., Sahr, T., Iachan, R., Denker, M., Duffy, T., & Weston II, D. (2013). Design and Analysis of Dual-Frame Telephone Surveys for Health Policy Research. World Medical and Health Policy, 217-232.

Malatest. (2017). Transportation Tomorrow Survey: TTS 2016 Challenges and Lessons Learned (Draft). Toronto: malatest.

McHugh, M. L. (2013). The Chi-square test of independence. Biochemia Medica, 143-149.

National Cooperative Highway Research Program. (2016). Standardized Procedures for Personal Travel Surveys. Washington: NCHRP.

NuStats. (2002). 2000 - 2001 California Statewide Household Travel Survey. Austin: NuStats.

NuStats. (2007). Chicago Regional Household Travel Inventory. Austin: NuStats.

NuStats. (2013). 2010-2012 California Household Travel Survey Final Report Appendix. Austin: NuStats.

State Health Access Data Assistance Center. (2013). Technical Report for the 2011 Minnesota Healtrh Access Survey: Survey Methodology, Weighting and Data Editing. Minnesota: SHADAC.

Statistics Canada. (2014, 6 25). Residential Telephone Service Survey, 2013. Retrieved 8 30, 2017, from Statistics Canada: http://www.statcan.gc.ca/daily-quotidien/140623/dq140623a-eng.htm

Verreault, H., & Morency, C. (2016). Integration of a phone-based household travel survey and a web-based student travel survey. Transportation, New York.