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

UTTRI

IMPACT OF MULTIPLE SURVEY FRAMES ON DATA QUALITY OF HOUSEHOLD TRAVEL SURVEYS

2017-12-01

Transportation Tomorrow Survey 2.0

Albert Lo, Siva Srikukenthiran, Meishi Chen
Khandker Nurul Habib & Eric J. Miller

EXECUTIVE SUMMARY

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 uses multiple sample frames consisting of address/phone, address-only, and phone-only. The purpose of this report is to understand the impacts of this approach on the demographics of the TTS dataset.

Data used for this analysis includes the 2011 and 2016 TTS and the census. It is important to note that the analysis used a preliminary version of the 2016 TTS, but no major deviations are expected. In terms of collecting the 2016 data, the address/phone and address-only sample frames were sent a notification letter(s), while the phone-only sample frames were contacted by phone. Phone-only samples were discontinued two weeks after introduction due to a low response rate and low interviewer morale. This results in a low proportion of phone-only sample completes (<1%).

Methodology to analyze impact includes the chi-squared test, RMSE, percent error maps, response rates, and proportion comparison. These methods are chosen for their simplicity and ease of use. These tools are used to analyze the demographic trends of the entire GGH, to compare the 2011 and 2016 TTS, compare representativeness of the 2016 TTS and census, and a comparison between the different sample frames.

Analysis of GGH trends between 2011 and 2016 show relatively stable proportions of gender, age, household size, and type of dwelling.

Analysis between the TTS 2011 and 2016 dataset shows a higher proportion of apartments, 1 to 2 person households, and 20-39 and 60+ year olds in 2016 compared to 2011. RMSE values show that the 2016 TTS has better representation than 2011 in the inner GGH regions (Toronto, Greater Golden Horseshoe Area) than in the outer regions (905 area, extended region).

The comparison between the TTS and the 2016 census reveals households in the TTS are more likely to be 2 person households than 5+ person households. This may be due to respondent fatigue. There is a smaller proportion of the TTS 2016 population under the age of 40 compared to the census.

The sample frame comparison yields several findings. In general, the address-only sample frame proportions are close to the 2016 census which translates to a lower RMSE value compared to the other two frames. The address-only sample frame has a 14.1% RMSE value for the GGH compared to the 27.3% and 22.9% for address/phone and phone-only respectively. Respondents of the address-only sample frame are more likely to respond by web than the other frames. The address-only sample frame appears to be well representative of the 2016 population in terms of demographics. However, it is unknown whether this is because of non-response or the demographics the sample frame covers. Moreover, the phone-only sample frame does little to increase the representativeness of the dataset due to the small sample size.

Key recommendations include investigating the trip and transit dataset of the 2016 TTS to determine the dataset's usability. It is also recommended to investigate the demographic profile of all samples used to determine whether the address-only sample frame's low RMSE value compared to the other two is due to non-response or the sample frame itself.

Impact of Multiple Survey Frames on Data Quality of Household Travel Surveys

TRANSPORTATION TOMORROW SURVEY 2.0

Table of Contents

List of Tables

List of Figures

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 the 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 sample correspond to households with a known address and phone number whereas the address-only and phone-only correspond to a known address and a known phone number respectively. The phone-only sample frame consisted of a white page listing, random digit dialing (RDD), and a verified cellphone sample.

The objective of this report is to present the respondent profiles and analysis of the TTS 2016. The survey data is compared to the 2016 census to examine differences in proportions of shared variables. A similar comparison is conducted for the 2011 iteration of the survey and 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 report acts to inform about recent demographic trends in the region and improvements made in the 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.

The structure of this report is as follows. A literature review of prior large-scale surveys is done to study methodology where multiple sample frames, especially phone-only, were used. Response rates and demographic profiles are examined. The following chapters detail the analysis of the GGH population trends from 2011 to 2016, comparing the TTS 2016 to the 2016 census, the 2011 TTS, and itself (the different sample frames). Results are shown and discussed in each chapter. The report concludes with respondent profiles and the effects of the sample frames on the TTS.

2 CURRENT STATE OF SURVEYS USING MULTIPLE SAMPLE 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 utilize landline numbers with a supplemental cellphone phone list to recruit participants. 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 the 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 an 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 (ABS) 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 for 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 undersampled 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 has 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 household are significantly different from the landline population and thus should be included in 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 report adds to this literature, showing results of using a multiple sample frame to collect household travel information in the GGH.

This report builds upon the work done by Chen et. al. for the TTS 2.0 project (Chen et. al.). That document reviews sample frame options including address-based, web-based and mixed. The authors conclude that the cellphone and landline sample frame would increase representation, especially among males between ages of 18 to 34. Chen's report along with this document is a part of the TTS 2.0 project, an initiative at the University of Toronto to develop next-generation survey methodology and tools for the upcoming TTS to be conducted in 2021.

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.

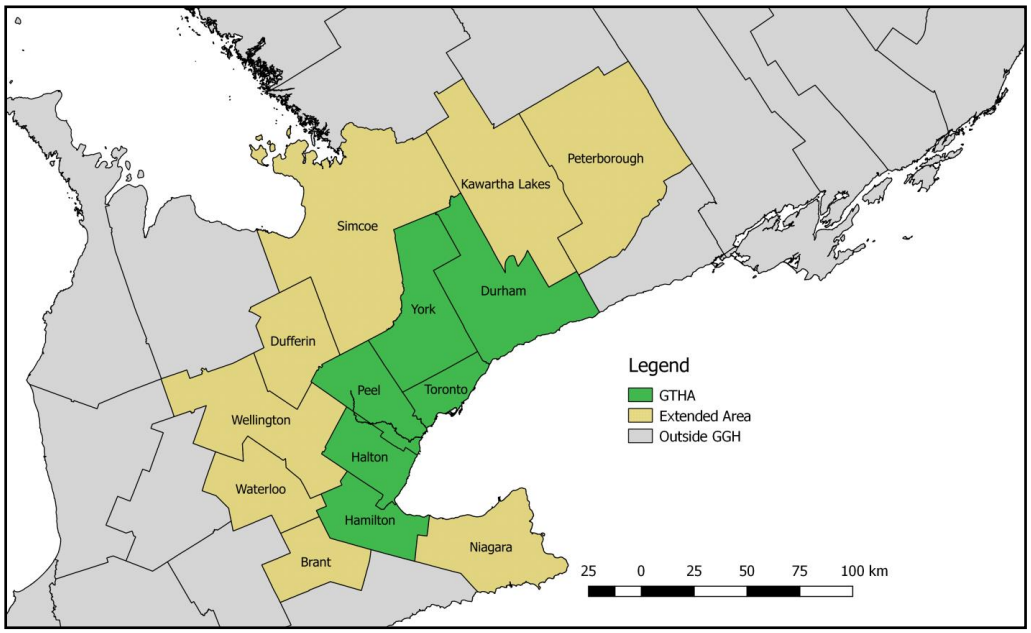

FIGURE 1 – THE GGH AND GTHA

3.2 Dataset Description

The data consists of frequency tables from 2011 and 2016 TTS and the Canadian census. A list of variables in the TTS dataset (specifically the house and person tables) and their possible values can be found in Table 1. 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 income was included in the 2016 survey. In terms of data integration and cleaning, ARCGIS is 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 are outside of the GGH and/or are labelled as Indian Reserve are removed.

It is important to note that at the time of writing this report, the 2016 dataset analyzed is 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. A separate database was provided by MALATEST to display response rates for the TTS 2016 by the municipality (CSD).

TABLE 1 – VARIABLES AND THEIR POSSIBLE VALUES IN THE TTS DATASET

Category	Values
Sample Frame	Address/Phone, Address-only, Phone-only, Volunteer
Response Mode (From beginning to end of survey)	Phone to phone, Web to web, Phone to web, Web to phone
Dwelling Type	House, Apartment, Townhouse, Unknown
Number of trips per person	Integer values from 0 to 33
Number of person per household	Integer values from 1 to 12
Age of person	Integer values from 0 to 99
Sex of person	Male, Female, Unknown
Year TTS conducted	2011, 2016
Income level (only available for the TTS 2016)	$0 to $14,999, $15,000 to $39,999, $40,000 to $59,999, $60,000 to $99,999, $100,000 to $124,999, $125,000+, Declined/I don't know

3.3 Method of Data Collection

Due to the difference in sample frames, the 2016 TTS method of collection is 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) are sent a notification letter. These households can respond to the survey via phone, complete it on the web, or wait for a call (restricted to the address/phone sample frame). Notification letters are 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, samples are 'cold-called' and can 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 the notification letter.

The proportions of the three sample frames are shown in Table 2. The low phone-only proportion is 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 is 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 2 – SAMPLE FRAMES 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 involves a comparison of the 2011 and 2016 census, the 2016 TTS and census, the 2011 and 2016 TTS, and between the 2016 TTS sample frames. The analysis is based on the chi-squared test, data observations, and non-statistical experiments. The comparison is 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 are excluded where appropriate. For example, the volunteer sample frame is 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 is no relationship between the samples and variables.

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

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

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

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 critical value of the chi-squared distribution with same degrees of freedom (calculated as a product of the 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 2016 TTS and census, the 2011 and 2016 TTS, and the three sample frames were the independent samples compared to the demographic variables.

4.2 RMSE

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 the variable i in category j

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

Table 3 shows the NCHRP's suggested variables and categories 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 report's analysis, household size, age, and gender are used to calculate RMSE as shown in Table 4 (Habib & El-Assi, 2015). Vehicle availability and the race could not be used. The latter was not asked in the TTS and the census values of the former 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. This is shown in Table 5.

This paper does not include household income in the RMSE. The most geographic disaggregated level by income bracket offered by the Canadian census is the census metropolitan area (CMA). The CMA level does not cover the entire GGH which means it is impossible to do a detailed analysis of household income with this data.

TABLE 3 – 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 4 – 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

The RMSE is calculated to compare representativeness of the TTS dataset and its three sample frames against the census. To determine the representativeness of 'hard-to-reach' populations, namely males between 20 and 29 years of age, an average is taken of RMSE values that are calculated as follows. For each combination of individual dwelling type and household size (15 combinations in total), an RMSE value is 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 is 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).

$$\%error = \frac{TTS\_proportion - census\_proportion}{census\_proportion} \quad (5)$$

For each planning district of the GGH, a percent error value is 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 are displayed on the maps from -100% to 100% in 5 bins using the graduated colour symbol.

TABLE 5 – DWELLING TYPE DEFINITION ASSOCIATION

Census Type of Dwelling Definitions	TTS Type of Dwelling Definitions
Single-Detached House	House
Apartment in a building that has 5 or more storeys	Apartment
Semi-Detached house	House
Row House	Townhouse
Apartment or Flat in a Duplex	Apartment
Apartment in a building that has fewer than 5 storeys	Apartment
Other single-attached house	House
Movable Dwelling	removed

4.4 Response Rate

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

5 GGH POPULATION TRENDS

To investigate the impact of multiple sample frames on the 2016 TTS, it is useful to consider population trends of the GGH between 2011 and 2016. This is because the 2011 TTS serves as the baseline for comparison to the 2016 dataset to determine if there is an improvement in representation (this analysis is in chapter 6). To consider these trends, the 2011 census is compared to the 2016 census. The results of the GGH population trend analysis are as follows.

TABLE 6 – DEMOGRAPHICS BY CENSUS 2011 AND 2016, GGH

	Census 2011	Census 2016
Gender		
Male	48.6% (4,140,210)	48.6% (4,379,630)
Female	51.4% (4,375,540)	51.4% (4,622,965)
Type of Dwelling		
Apartment	34.6% (1,078,760)	35.9% (1,194,550)
House	56.6% (1,764,950)	54.8% (1,827,120)
Townhouse	8.9% (276,820)	9.3% (309,605)
Household Size		
1	24.0% (749,430)	24.7% (822,055)
2	30.1% (939,620)	30.4% (1,013,465)
3	17.2% (538,225)	17.0% (567,380)
4	17.4% (543,115)	16.8% (560,685)
5+	11.3% (353,880)	11.1% (370,865)
Age Categories		
0-19	24.0% (2,047,900)	22.7% (2,046,745)
20-29	13.4% (1,137,165)	13.6% (1,225,305)
30-39	13.5% (1,145,650)	13.3% (1,197,870)
40-59	29.9% (2,547,370)	28.9% (2,599,310)
60+	19.2% (1,637,570)	21.5% (1,933,365)

TABLE 7 – CHI-SQUARE TEST RESULTS BY CENSUS 2011 AND 2016, GGH

	Chi-Squared Test Statistic	Cramér's V
Gender	1.603101	0.00
Type of dwelling	1924.791	0.02
Household size	756.013	0.01
Age categories	15955.63	0.03

Table 6 summarizes the demographics of the 2011 and 2016 census. Gender and household size proportions are shown to remain relatively stable. The apartment proportion is slightly increased between the two years. With respect to age categories, there is a small decrease in the proportion of 0-19 year olds and a small increase of the 60+ year olds. This most likely signifies an aging population for the GGH.

Chi-squared tests are conducted for each of the four demographic variables and the results are displayed in Table 7. All but the gender variable exhibit significance and all variables have a weak association with the 2011 and 2016 census. This is likely due to the large sample sizes of the dataset, which enables the test to detect small changes which are deemed significant. Similar work is done for the disaggregated GGH regions (i.e. GTHA, extended region, Toronto, 905) and results were similar.

Comparing trends between the 2011 and 2016 census of the GGH yields little change due to the low association between each demographic variable and the 2011 and 2016 census, and the relatively stable proportions between the variables' categories.

6 2011 AND 2016 TTS COMPARISON

The purpose of comparing the TTS 2016 and 2011 datasets is to determine if there has been an improvement in representation between the two surveys. Proportions, chi-squared test results, response rates, and RMSE values are investigated. As a side note, when analyzing at the disaggregated GGH region levels, CSDs for Barrie and Innisfil are aggregated due to the difference in the boundary definition between 2011 and 2016.

TABLE 8 – DEMOGRAPHICS BY TTS 2011 AND 2016, GGH

	TTS 2011	TTS 2016
Gender		
Male	48.8% (200,380)	48.0% (190,256)
Female	51.2% (209,885)	52.0% (205,805)
Type of Dwelling		
Apartment	26.0% (41,397)	29.2% (47,545)
House	66.5% (105,647)	61.1% (99,474)
Townhouse	7.5% (11,891)	9.6% (15,667)
Household Size		
1	21.6% (34,424)	25.3% (41,229)
2	36.5% (58,040)	37.7% (61,418)
3	17.2% (27,310)	15.8% (25,712)
4	16.0% (25,533)	13.9% (22,664)
5+	8.7% (13,850)	7.2% (11,768)
Age Categories		
0-19	22.1% (90,698)	19.1% (75,842)
20-29	8.6% (35,128)	9.6% (37,900)
30-39	10.9% (44,565)	11.9% (47,291)
40-59	31.0% (126,992)	29.5% (116,854)
60+	27.5% (112,529)	29.9% (118,415)

Table 8 shows percentage proportions of the demographic variables. There are notable differences between the two survey datasets. A higher proportion of apartments is exhibited in 2016. The 2016 TTS also exhibits a higher percentage of 1 and 2 person households, a lower proportion of 3+ person households, and a large proportion of 20-29, 30-39, and 60+ year olds compared to 2011.

TABLE 9 – CHI-SQUARED TEST RESULTS BY TTS 2011 AND 2016, GGH

	Chi-Squared Test Statistic	Cramér's V
Gender	52.19654	0.01
Type of dwelling	1853.61	0.05
Household size	1054.898	0.06
Age categories	1084.534	0.06

Similar to the results detailed in the previous chapter, chi-square tests show significance but have a low associated Cramér's V value. Again, this may be due to the large sample size which makes what seems insignificant, significant.

Table 10 provides a summary of response rates and RMSE values for the 2011 and 2016 TTS dataset. This iteration of the TTS achieved a response rate of 16.3% (998,515 samples in the field). Generally, both the GTHA and Toronto exhibit a lower response rate compared to the extended region and 905 area, respectively. Disaggregated GGH region response rates were not available. The reason for the sharp decrease in the response rate at GGH level is due to a large number of households being part of the address-only cohort (as shown in Table 2). The response rate was low for this sample frame.

TABLE 10 – RESPONSE RATE BY TTS 2011 AND 2016

	TTS 2011	TTS 2016
Response Rate		
GGH	49.0%	16.3%
GTHA	-	16.0%
Extended Region	-	17.2%
Toronto	-	14.0%
905 area	-	17.9%

Table 11 displays the RMSE values. The TTS 2011 and 2016 datasets are compared to their respective census datasets. Overall, the 2016 TTS has a lower RMSE value which means a better representation than its 2011 counterpart. However, examining the disaggregated GGH regions' RMSE, it seems that the RMSE values are higher in 'outer' regions (i.e. extended region, 905 area) versus the 'inner' regions (GTHA, Toronto).

TABLE 11 – RMSE VALUES BY TTS 2011 AND 2016

	2011	2016
GGH	18.2%	17.0%
GTHA	18.0%	16.2%
Extended Region	18.7%	20.1%
Toronto	24.2%	19.7%
905 area	15.1%	16.3%

7 REPRESENTATIVENESS COMPARISON

The TTS 2016 dataset is compared to the 2016 census to determine whether this TTS iteration is representative of the population. The proportions are presented, followed by the chi-squared test results and several maps.

TABLE 12 – DEMOGRAPHICS BY TTS 2016 AND CENSUS 2016, GGH

	TTS 2016	Census 2016
Gender		
Male	48.0% (190,256)	48.6% (4,379,630)
Female	52.0% (205,805)	51.4% (4,622,965)
Median Age	46	40
Median Respondent age	55	-
Type of Dwelling		
Apartment	29.2% (47,545)	35.9% (1,194,585)
House	61.1% (99,474)	54.8% (1,827,055)
Townhouse	9.6% (15,667)	9.3% (309,595)
Household Size		
1	25.3% (41,229)	24.7% (822,055)
2	37.7% (61,418)	30.4% (1,013,465)
3	15.8% (25,712)	17.0% (567,380)
4	13.9% (22,664)	16.8% (560,685)
5+	7.2% (11,768)	11.1% (370,865)
Household Response Mode		
Online only	64.1% (103,335)	-
Phone only	35.9% (57,890)	-
Household Income		
$0 to $14,999	4.0% (6443)	-*
$15,000 to $39,999	14.6% (23,815)	-*
$40,000 to $59,999	14.2% (23,102)	-*
$60,000 to $99,999	20.8% (33,818)	-*
$100,000 to $124,999	9.8% (15,943)	-*
$125,000 +	17.4% (28,341)	-*
Decline / Don't know	19.2% (31,330)	-*
Age Categories		
0-19	19.1% (75,842)	22.7% (2,046,910)
20-29	9.6% (37,900)	13.6% (1,225,290)
30-39	11.9% (47,291)	13.3% (1,197,905)
40-59	29.5% (116,854)	28.9% (2,599,220)
60+	29.9% (118,415)	21.5% (1,931,315)

* income not included due to incomplete data in the census

Table 12 summarizes the demographics by TTS 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.

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 by the older median age and a 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.

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.

Table 13 shows the chi-squared test statistic values and their associated Cramer's V value. All chi-squared tests are significant. All relationships have a weak association.

TABLE 13 – CHI-SQUARED TEST RESULTS BY 2016 TTS AND CENSUS, GGH

	Chi-Squared Test Statistic	Cramér's V
Gender	61.55	0.00
Type of dwelling	3055.69	0.03
Household size	5810.56	0.04
Age categories	19750.31	0.05

Figure 2 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 3 to 7 show the percent error of each age group by planning district. Of particular note are the '20-29' and '60+' age groups where there is high under and overrepresentation, respectively. The '20-29' age group ranges from 66.6% to 24.3%. The '60+' percent errors are exclusively positive, ranging from 5% to 117%.

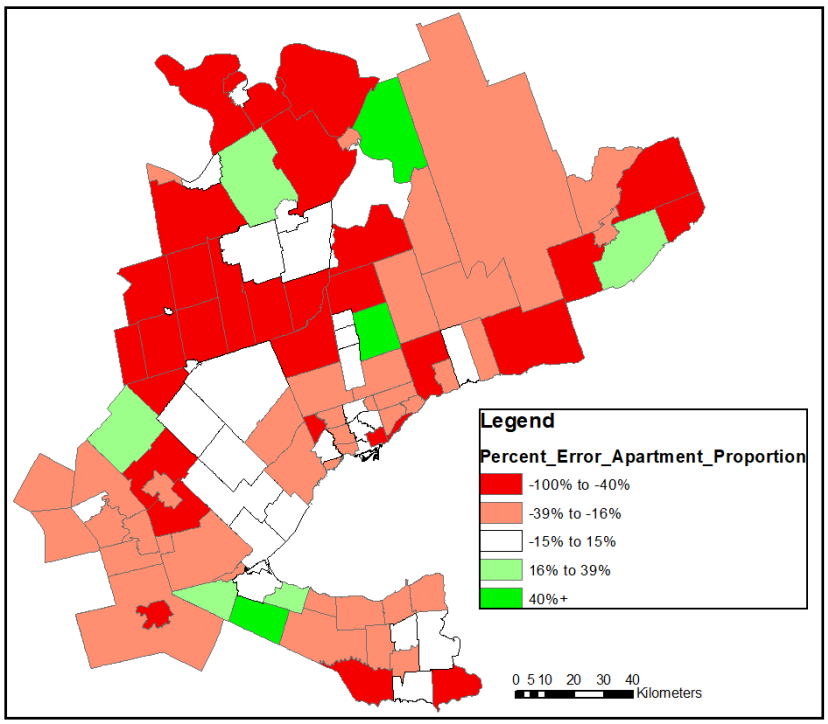

FIGURE 2 – PERCENT ERROR OF APARTMENT PROPORTIONS BY PLANNING DISTRICT, (GGH)

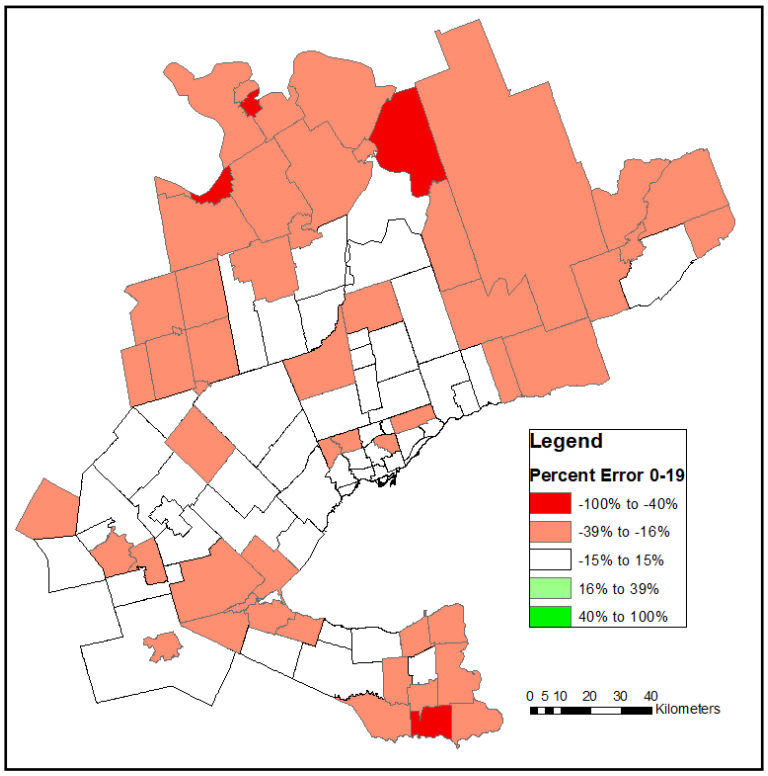

FIGURE 3 – PERCENT ERROR OF AGE CATEGORY 0-19 BY PLANNING DISTRICT, (GGH)

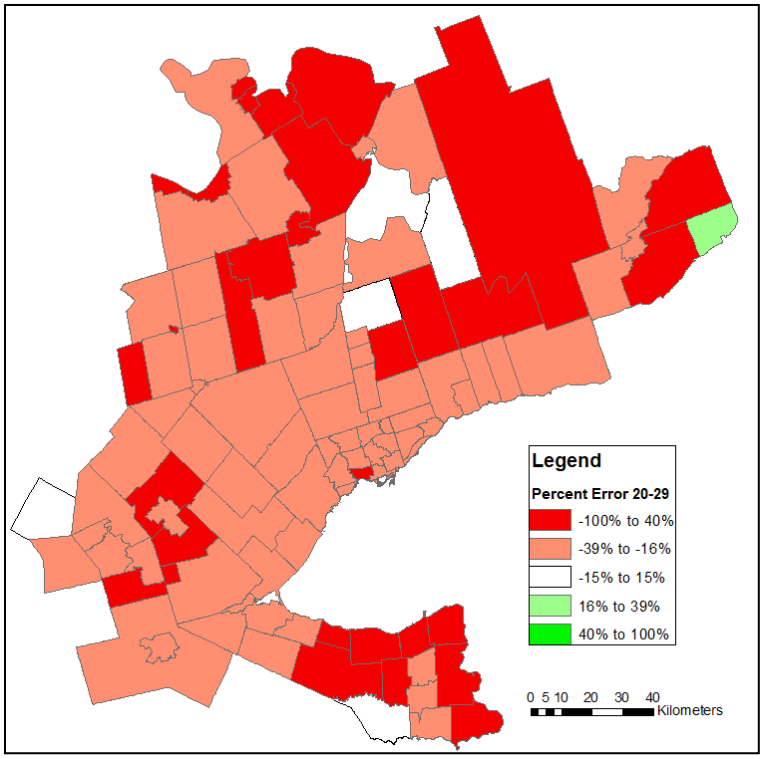

FIGURE 4 – PERCENT ERROR OF AGE CATEGORY 20-29 BY PLANNING DISTRICT, (GGH)

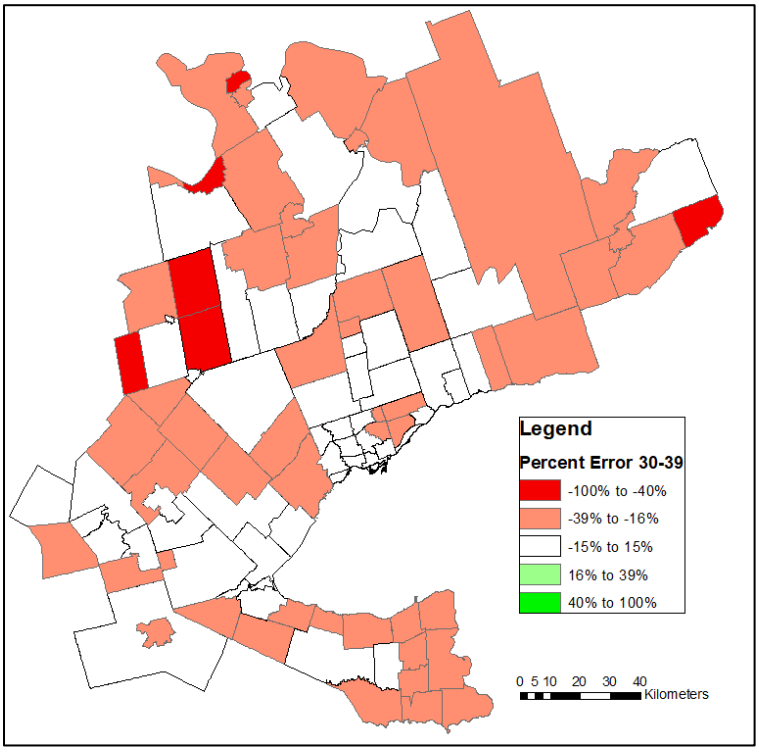

FIGURE 5 – PERCENT ERROR OF AGE CATEGORY 30-39 BY PLANNING DISTRICT, (GGH)

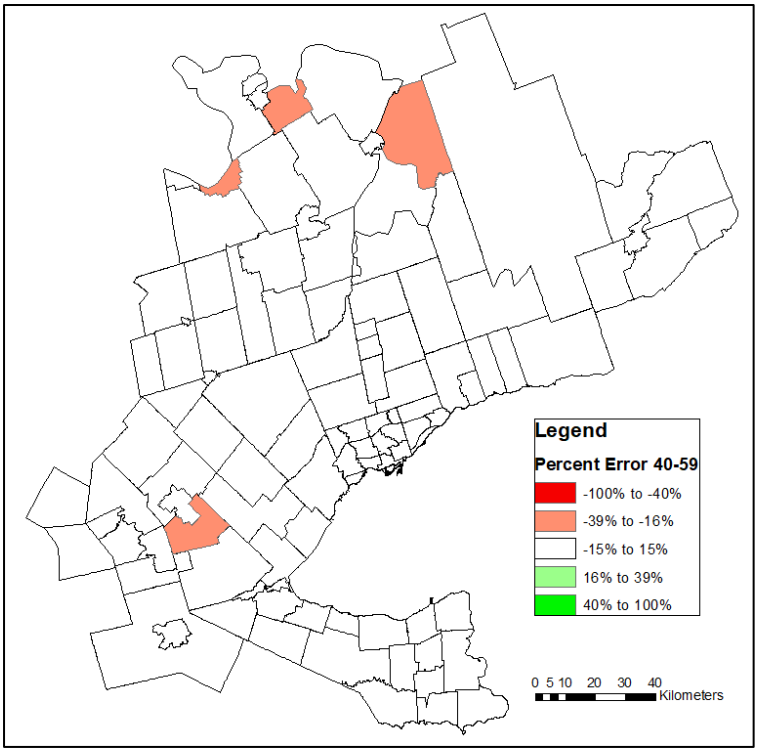

FIGURE 6 – PERCENT ERROR OF AGE CATEGORY 30-39 BY PLANNING DISTRICT, (GGH)

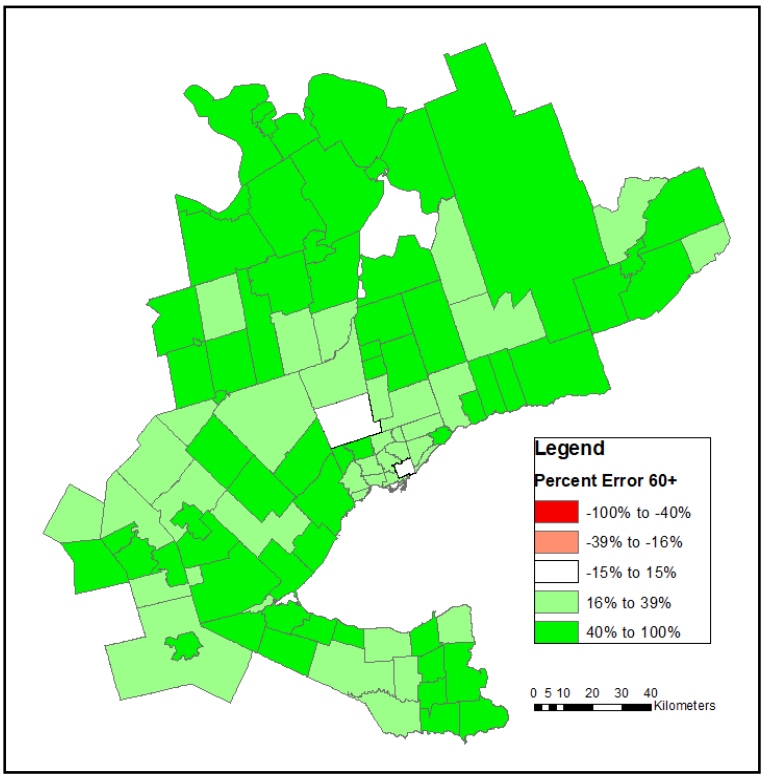

FIGURE 7 – PERCENT ERROR OF AGE CATEGORY 30-39 BY PLANNING DISTRICT, (GGH)

Figure 8 shows RMSE values comparing TTS and census 2016 by CSD. Values range from 11% to 84% with the median RMSE value being 23.4%. The figure 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 townhouses. The 60+ age category and the household size of 5+ have high error as well.

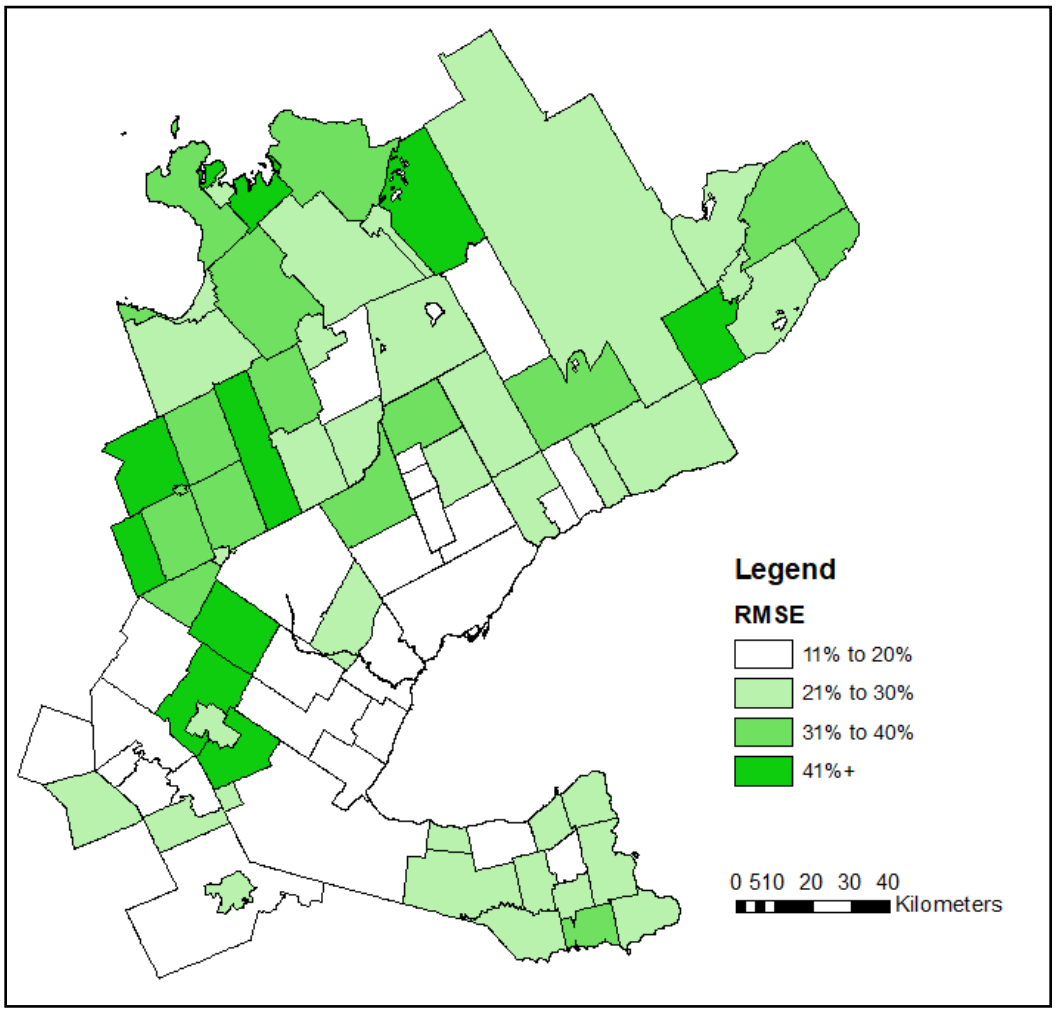

FIGURE 8 – RMSE BY CSD, (GGH)

8 SAMPLE FRAME COMPARISON

For the first time in the TTS, the 2016 iteration introduces ABS into its sampling frame. To determine the impact of this approach, proportion percentage of demographic variables, response rates, chi-squared test results, and RMSE values are compared by sample frame.

TABLE 14 – DEMOGRAPHICS BY SAMPLE FRAME, GGH

	Address/Phone	Address-only	Phone-only*
# of households	82,383	79,256	1,097
# of people	203,004	190,296	2,872
Gender			
Male	47.4% (96,044)	48.8% (92,768)	48.5% (1390)
Female	52.6% (106,746)	51.2% (97,506)	51.5% (1477)
Median Age	52	40	45
Median Respondent age	60	49	53
Type of Dwelling			
House	67.7% (55,732)	54.1% (42,902)	74.1% (813)
Apartment	23.8% (19,611)	35.0% (27,714)	17.8% (195)
Townhouse	8.5% (7001)	10.8% (8575)	8.1% (89)
Household Size			
1	24.2% (19,897)	26.6% (21,091)	20.1% (221)
2	38.5% (31,753)	36.9% (29,230)	38.2% (419)
3	15.3% (12,611)	16.3% (12,924)	15.3% (168)
4	14.3% (11,802)	13.5% (10,674)	17.0% (186)
5+	7.7% (6320)	6.7% (5338)	9.4% (103)
Household Response Mode			
Online only	39.8% (32,484)	90.1% (70,698)	14.0% (153)
Phone only	60.2% (49,163)	9.9% (7731)	86.0% (941)
Household Income			
$0 to $14,999	3.8% (3127)	4.1% (3261)	4.6% (51)
$15,000 to $39,999	15.8% (13,025)	13.4% (10,623)	14.2% (156)
$40,000 to $59,999	14.5% (11,944)	13.9% (10,986)	14.9% (163)
$60,000 to $99,999	18.5% (15,257)	23.2% (18,348)	19.0% (208)
$100,000 to $124,999	8.5% (7040)	11.1% (8820)	7.4% (81)
$125,000 +	16.2% (13,372)	18.6% (14,761)	18.3% (201)
Decline / Don't know	22.6% (18,618)	15.7% (12,458)	21.6% (237)
Age Categories			
0-19	18.2% (37,039)	20.0% (38,126)	22.6% (648)
20-29	7.5% (15,184)	11.8% (22,447)	9.1% (262)
30-39	7.3% (14,860)	16.9% (32,146)	9.7% (279)
40-59	29.6% (60,138)	29.3% (55,777)	31.4% (902)
60+	37.3% (75,783)	22.0% (41,800)	27.2% (781)

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

Table 14 shows the demographics by sample frame. 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 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 the 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.

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.

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.

TABLE 15 – RESPONSE RATES BY SAMPLE FRAME

	Address/Phone	Address-only	Phone-only
Response Rate			
GGH	39.6%	10.5%	8.6%
GTHA	38.7%	10.4%	8.2%
Extended Region	42.4%	10.8%	9.6%
Toronto	39.1%	9.8%	8.0%
905 area	38.4%	11.1%	8.3%

At the GGH level by sample frame, table 15 provides 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.

Table 16 shows the chi-squared test statistic values and their associated Cramér's V by sample frame. 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 16 – CHI-SQUARED TEST RESULTS BY SAMPLE FRAME, GGH

	Chi-Squared Test Statistic	Cramér's V
Gender	76.61	0.01
Type of dwelling	3240.58	0.10
Household size	248.30	0.03
Response mode	45292.75	0.53
Income	2014.53	0.08
Age categories	17401.77	0.15

Table 17 summarizes RMSE values by sample frame. 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.

TABLE 17 – RMSE VALUES BY SAMPLE FRAME

	Address/Phone	Address-only	Phone-only
RMSE			
GGH	27.3%	14.1%	22.9%
GTHA	27.3%	14.5%	22.1%
Extended region	29.1%	15.4%	28.6%
Toronto	36.2%	17.4%	31.0%
905 area	23.7%	13.9%	20.1%

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 a 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.

The address-only sample frame is seemingly more representative than the address/phone and phone-only sample frames. 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 were chosen of the sample frame or non-response bias. Hypothetically speaking, if samples chosen in the address/phone sample frame are skewed to an older population, the sample frame would, therefore, be representative of its population. However, comparing this sample frame to the census would result in the sample frame being deemed as weak in representation. Nonetheless, each sample frame can serve a different purpose depending on the need.

9 CONCLUSIONS

Overall, in terms of only demographic variables, the 2016 dataset is somewhat representative with a 17.0% RMSE, meaning that the TTS 2016 represents the census with an 83% accuracy; 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 subsections 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.

Examining RMSE values by disaggregated GGH region and CSD reveals that the inner regions of the GGH are better represented than the outer regions.

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.

To augment this analysis on the demographic variables, it is recommended to examine the travel behavior variables of the 2016 dataset (i.e. the trip and transit datasets). Initial assessments show that trip rates for the 2016 dataset is lower than the 2011 dataset and may be a cause for concern. It is also recommended to investigate whether participants have different travel behavior (e.g. primary mode of travel, number of trips, length of trips) and by which sample frame they were recruited. Not only will this determine the usability of the 2016 dataset but inform future TTS iterations' methods of how to recruit participants.

In regard to the impact of the multiple sample frame, the demographic profile of each sample frame should be analyzed based on the samples sent in the field. This information is needed in order to understand if the address-only sample frame is more representative than the other sample frames due to non-response or the demographic that the address-only sample frame recruits.

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 Dlgit 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.