TRANSPORTATION DATA COLLECTION FOR THE FUTURE: A CASE STUDY IN THE GREATER GOLDEN HORSESHOE, CANADA

Matthew J. Roorda, Dept. of Civil Engineering, University of Toronto, 35 St. George Street, Toronto, Ontario, Canada, M5S 1A4, Tel: 416.978.5976, roordam@ecf.utoronto.ca

Amer Shalaby, Dept. of Civil Engineering, University of Toronto, 35 St. George Street, Toronto, Ontario, M5S 1A4, Tel: 416.978.5907, amer@ecf.utoronto.ca

Sheyda Saneinejad, Dept. of Civil Engineering, University of Toronto, 35 St. George Street, Toronto, Ontario, Canada, M5S 1A4, Tel: 416.978.3141, s.saneinejad@utoronto.ca

ABSTRACT

This paper addresses two questions regarding transportation data collection in the Greater Golden Horseshoe, based on analysis of data needs and a review of international data collection practices. First, will the current data collection efforts provide sufficient information to answer key transportation questions that will arise over the next decade? Second, how should the household travel survey be effectively and responsibly conducted in the future, recognizing significant technological and societal changes that are occurring?

Major established data collection efforts be maintained, prioritizes gaps in transportation should be addressed, and an on-going research program dedicated to survey methods research should be developed. The paper outlines a strategy for upgrading the household travel survey. An internet version of the survey, a dual-frame sampling approach, a supplementary GPS-assisted survey, and consistency of the survey with previous implementations are recommended.

The transportation data challenges faced in the Greater Golden Horseshoe are not unique to the region, and the recommendations made here are relevant to data collection programs in other peer jurisdictions world-wide.

Keywords: Travel demand, Surveys, Data collection, Global Positioning Systems, Planning

1. INTRODUCTION

A well-conceived and well-executed transportation data collection program is central to good public-sector decision-making towards a prosperous, liveable and sustainable city. Looking ahead at the future of data collection in many developed regions of the world, a few major challenges are apparent in terms of the content of the collected data and methodology of data collection.

Over the next decade, it is expected that vigorous public debate and discussion worldwide will surround major transportation-related issues such as greenhouse gas emissions and air quality, transit investment, transportation pricing, cost of fuel, alternate fuels and electric autos, changing demographics, costly impacts of the prevalent transportation patterns on road safety and population health, changes in nature of the regional economy, and the growing significance of personal travel to destinations other than work or school. Although no single travel survey can provide enough information to answer these concerns there are significant opportunities to develop integrated systems of data collection that minimize important data gaps, maximize compatibility of data sources, and allow for comprehensive modelling and analysis that provides better decision-support to tackle these major transportation policy issues, both on an individual basis as well as comprehensively.

Another critical challenge to the existing data collection programs is posed by the significant technical and societal changes occurring in most developed nations. There is, on the one hand, a need to maintain continuity in survey instruments over time to allow for trend analysis, to prevent changes in instrument bias, to support legacy modelling techniques, and to build upon the knowledge gained from previously collected data. On the other hand, the social and technological context for data collection is clearly changing. Methods that have worked in the past are no longer working as well. For example, telephone interview response rates have declined substantially over time and directories of land telephone lines that have traditionally been used do not capture an increasing population of individuals that only use mobile telephones or voice over internet protocol (VoIP). At the same time, data collection programs can also take advantage of technological developments. Several new instruments for data collection are available and are being increasingly accepted internationally as robust and reliable state-of-practice tools, such as Global Positioning System (GPS) assisted surveys, internet-based data collection and a variety of ITS-based passive data collection techniques.

Similar to peer jurisdictions across the world, the transportation planning community in Southern Ontario is faced with the above challenges to its existing transportation data collection program for the Greater Golden Horseshoe (GGH). In the GGH, a heavily populated area (approximately 8 million residents) which includes the Greater Toronto Area and the surrounding regions (see Figure 1), there exists a long history of high quality transportation data collection to support many forms of transportation planning and research performed by a wide range of organizations. The centrepiece of this history is the Transportation Tomorrow Survey (TTS), a 24-hour retrospective telephone interview of the personal travel conducted by all members (aged 11 or older) in the household. This survey has presumably the largest sample of its kind in the world, with a sample of 5.2% in 2006,

resulting in interviews of about 150,000 households, involving 400,000 persons and 865,000 trips (DMG, 2007a).

Figure 1 - Greater Golden Horseshoe in Southern Ontario

The following two questions need to be addressed by the transportation planning community in Southern Ontario:

1. Will the TTS, in combination with other data sources, provide sufficient information to answer key transportation questions that will arise over the next 10 years?
2. How should the TTS be effectively and responsibly conducted in the future, recognizing the significant technological and societal changes that are occurring?

The objective of this study is to provide insights towards answering these two questions based on an analysis of data needs in the GGH and a review of international data collection practices

2. TRANSPORTATION DATA COLLECTION FRAMEWORK

In order to make good decisions about future implementations of the TTS and other data collection efforts to be undertaken in the GGH, it is critical to a) understand what the "universe" of travel is, b) understand which components of the "universe" are being observed in established data collection programs, c) prioritize those components of travel that require better understanding in order to inform good public policy decisions, and d) identify improvements to the current transportation data collection strategy that would allow us to better achieve that understanding.

Defining the Universe of Travel

Figure 2 depicts the universe of travel in the GGH, categorizing travel by traveller (who), the purpose of travel (why), the travel location (where), the mode of transportation (how), and various time periods (when). Different travel behaviour occurs at different times of day, on different days of the week (with the primary difference between weekdays and weekends), and over different seasons.

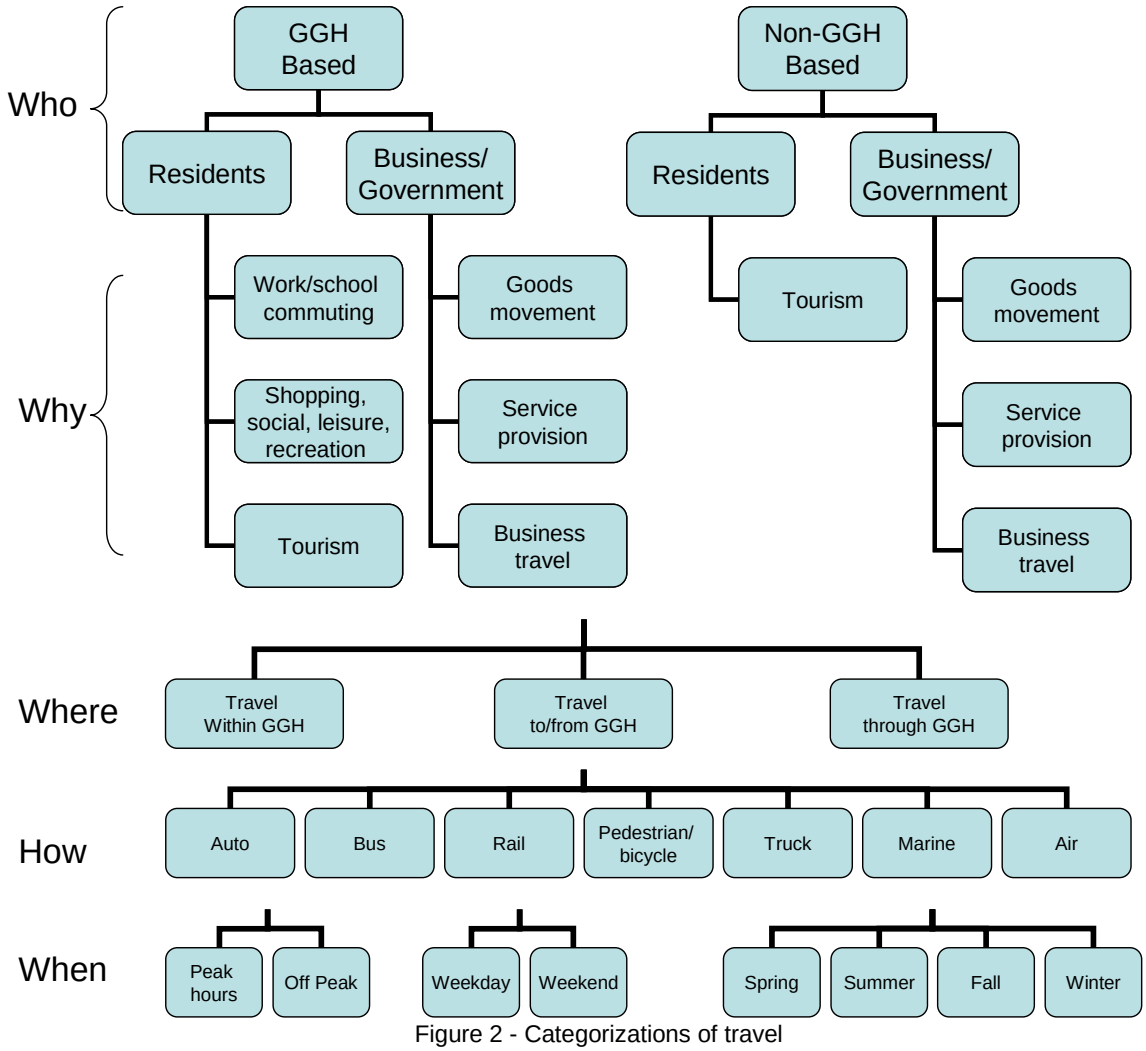

Figure 2 - Categorizations of travel

Transportation Data Collection in the GGH

Significant information is already collected about travel in the GGH. Several high-quality established data collection programs focus on important parts of the "universe" of travel in the GGH. Rich databases are also collected in the private sector. However, the transportation data collection cannot be considered comprehensive because significant gaps remain. This section briefly describes the scope of major data sources available in the GGH for transportation planning purposes, their key limitations and the remaining gaps.

TTS

The TTS is the centrepiece for collection of data about the personal travel behaviour of GGH residents (in Canada, most transportation related data are collected at the regional or municipal level, rather than in the Canadian census). The TTS collects information about trips made on a single fall weekday by all household members 11 years of age or older, by all modes of transportation. The TTS has been conducted on approximately a 5% sample of the population every 5 years from 1986 to 2006. A complete description of the TTS data can be found in (DMG, 2007a). Despite the rich travel data it collects, the TTS does not attempt to capture:

- Non-motorized trips (i.e. walk and cycle) for non-work/non-school purposes;
- Weekend travel;
- Seasonal variations in travel (the TTS collects information for fall, and in rare cases spring, weekdays only);
- Travel information for children under the age of 11;
- Economic information including income and costs of travel and parking;
- Information about the types and ages of vehicles owned by the household; and
- Detailed information on activities and travel routes (e.g. activity start time, travel route for auto and other non-transit trips).

Traffic Count Programs

Extensive cordon count programs are undertaken every 2-3 years by various regional governments within the GGH in addition to intersection and roadway traffic counts (DMG, 2007b). Traffic counts separately classify automobiles, truck configurations and buses and while they provide information about the number of vehicles travelling at specific times, they do not provide any information about the origin and destination of travel, who the traveller is or the purpose of travel. Thus, count information can only provide limited insight for policy assessment in most cases, and is most useful for model validation purposes.

Transit Surveys

The major transit agencies in the GGH carry out field surveys such as speed-and-delay surveys, transit ridership surveys and attitudinal surveys, some regularly and others on as-needed basis, to gather data on transit operations and usage for service planning and

scheduling purposes. Transit ridership surveys are performed regularly (a few times a year by some agencies) using manual methods.

Ontario Ministry of Transportation (MTO) Commercial Vehicle Survey (CVS)

The most extensive survey of travel initiated by businesses is the MTO CVS, which is supplemented by the National Roadside Survey (NRS). The MTO CVS is a province-wide road-side vehicle survey, conducted in 2001 at over 150 road-side directional sites in the province of Ontario, in which drivers are asked to report on truck activity characteristics related to the trip, driver, carrier, commodity and vehicle. This survey does not collect travel information on any business-related travel except that which moves by trucks over 4500kg GVW, and that passes one of the CVS data collection sites. Thus, little information is collected about urban pick-up and delivery trips, and no information is collected about commercial travel by automobile or other modes.

Private sector databases

Private sector databases are collected in the GGH for railway companies (CN/CP), the Greater Toronto Airport Authority (GTAA), port operators, and truck fleet management firms. The challenge with all private sector databases is that they include proprietary data, such that the databases are not necessarily available for all types of analysis. Furthermore, private data sources do not necessarily maintain the same standards of sample control such that it can be difficult to make inferences about the full population.

Gaps in Behavioural Transportation Data Collection in the GGH

Figure 3 summarizes those components of travel that are currently captured in a systematic large-scale behavioural data collection effort and those that are not. “Behavioural” travel data refers to data that not only counts vehicles/passengers, but also captures sufficient background information about the individuals and the trips to be able to model or otherwise analyze the underlying behaviour that might be influenced by policy intervention. Gaps in the behavioural transportation data currently collected in the GGH are summarized as follows:

- Local (within GGH) goods movement and service provision by trucks;
- Goods movement and service provision by automobiles;
- Non-motorized personal trips to/from non-work/non-school trip purposes;
- Tourist and business travel, aside from such travellers captured through private surveys at the airports, on transit systems, or intercity tourism studies;
- Travel information on children under the age of 11;
- Weekend travel and seasonal variation in travel;
- Economic elements of personal travel, including income and costs of travel and parking;
- Information about the types and ages of vehicles owned by the household;

- Detailed information on activities (e.g. activity start time, activity scheduling), to enable modelling and assessment of policies and technologies that target activity participation (e.g. telecommuting); and
- Detailed information on trip routes, specifically auto trip routes, to enable modelling and assessment of congestion pricing schemes.

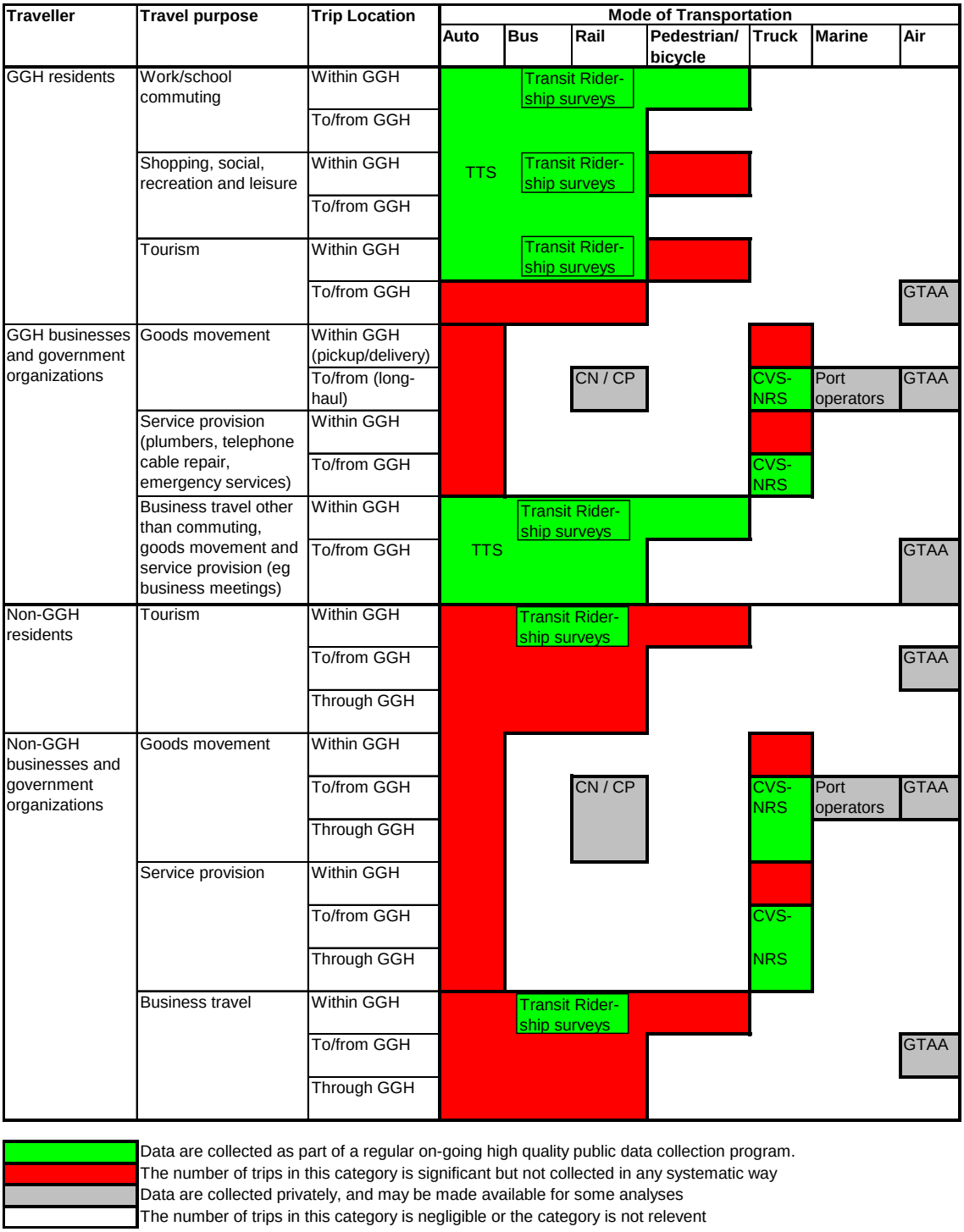

Figure 3 - Summary of behavioural travel data collection in the GGH

Which Data Are Important for Public Policy Decision-Making?

A gap in information does not necessarily warrant a new data collection effort. A rational transportation data collection program captures information that is most useful for providing decision-makers with the means to make good public policy decisions on issues that have high impact on the residents and businesses of the GGH. Thus, the following criteria can provide guidance about the sub-markets of travel that should be targeted for data collection.

- The amount of travel occurring within the transportation sub-market;
- The positive and negative impacts of that sub-market;
- The influence that public policy interventions can have on that sub-market; and
- The extent to which quality information can be cost-effectively gathered to link policy interventions to outcomes.

Amount of Travel by Sub-Market

An in-depth quantitative analysis of travel in each transportation sub-market is beyond the scope of this study. Furthermore, it is not possible to precisely assess the amount of travel in each transportation sub-market without holistic data collection. The authors make the following general observations/ judgements:

- Personal travel by GGH residents (commuting and travel to shopping, leisure, and social and personal business activities) is clearly the largest component of travel-kilometres in the GGH. In the AM Peak period, home-based work and school trips dominate, while in the PM peak, mid-day, and evening periods, a greater variety of travel purposes is found.
- Travel initiated by GGH businesses and government organizations is most significant during business hours with some avoidance of peak periods, especially in the AM. It is expected that a very significant proportion of light goods and service movements is by automobile, although little local information is available to assess this.
- Tourist travel by non-GGH residents is greatest on weekends, holidays, and the summer, and is more concentrated at tourist locations. There exists little information to assess the amount and type of tourist travel in the GGH aside from the Travel Survey of Residents of Canada (Statistics Canada, n.d.), which does not acquire detailed geographic information.
- Non-GGH businesses initiate travel into, out of, and through the GGH. Through travel is small compared to the amount of travel with trip origins or destinations in the GGH, since the GGH is both a major supplier and consumer of transported goods and services.

Impact of Travel by Sub-Market

Impacts of travel are diverse, including environmental, social, economic, health and safety, operational, etc. Differences in impacts between modes of transportation are well known and are not discussed in detail here. However, to provide insights for data collection, the authors recognize the following salient facts:

- Auto, transit and other motorized personal travel by GGH residents, as collected by the TTS, cause the lion's share of most major categories of impact.
- While passenger and bicycle travel involve less impact, they are also very important to understand because they are healthy and "sustainable" alternatives to higher impact (motorized) modes for a large number of short trips.
- Commercial vehicles, and in particular loaded trucks, have a disproportionately large effect on infrastructure deterioration, traffic safety, congestion, air quality and greenhouse gas emissions, noise, and vibration. Many of these impacts are experienced most acutely when trucks enter into urban areas. Efficiency of commercial vehicle travel also has a direct bearing on the region's economy.
- The composition and age distribution of the vehicle fleet in the GGH has a very significant relationship with GHG emissions and air quality. Especially, with the coming to market of alternative fuel technologies and electric vehicles, there is potential for mitigating the impacts of automobile and truck transportation.

Data Needs to Support Policy Interventions

Agencies funding data collection should focus on the data most relevant to high impact decisions that are within their jurisdiction/mandates. Over its history, the TTS has been used to support a wide variety of modelling and analysis efforts to provide decision support for land use, road infrastructure and transit improvements.

While it is essential that data collected in the GGH continue to support current modelling and analysis efforts, we expect that over the next decade vigorous public debate and discussion will surround the following major transportation-related policies and issues:

- Major investment in transit infrastructure;
- Policies and strategies towards improvement of air quality and reduction of greenhouse gas emissions;
- Transportation pricing mechanisms (congestion pricing / tolls / carbon taxes);
- Promotion and regulation of new transportation technologies (e.g. ITS, electric vehicles);
- Meeting transportation needs of a retired baby boom generation;
- Movement from a manufacturing-based to a service/knowledge-based economy; and
- Large increases in fuel prices for both personal travel and goods movement.

Arguably, the TTS and other major established data collection programs provide substantial information to support decision-making about transit investment, however, they do not provide enough information to inform decisions on other policies and issues on this list. For example:

- The TTS does not collect information about household income or the cost of travel, both of which are primary determinants of an individual household's response to transportation pricing, fuel price increases, and the purchase of new transportation technology such as hybrid-electric vehicles.
- Transportation air quality and greenhouse gas emissions are heavily influenced by vehicle type and age, neither of which is linked to household or travel information in any of the established data collection systems. Particulate matter and NOx

emissions largely originate from trucks, which are not well observed in the urban areas of the GGH.

- No comprehensive information is collected about urban goods pick-up and delivery or service truck movements, which could be influenced by pricing policy, transportation infrastructure and regulation of various sorts.
- The transportation needs of an aging baby boom generation can be reasonably assessed through analysis of historical data from the TTS. However, non-work non-school trips by non-motorized modes, which are not collected by the TTS, can be expected to represent a greater proportion of this population segment's travel.

The judgement of the authors of this study is to examine further the collection of additional data as shown in Table 1. These additional data are ranked in order of priority of importance by the authors for the purposes listed above. The best measure of the feasibility of new or expanded data collection is to observe the success of other metropolitan areas, or individual regional municipalities within the GGH, that have engaged in different data collection practices. Thus, Table 1 also briefly describes some of the most relevant precedents in Canada and the US where the additional data elements are collected.

Table 1 – Priorities and Precedents for New Data Collection

Prior-ity	**Data Collection Need**	**Discussion of precedents** (this is not a comprehensive review, but rather a selection of the most relevant studies)
1	Goods movement and service provision by automobile and truck, within the GGH.	Urban goods and service movements have been collected through in-person interviews with business establishments in the Cities of Calgary and Edmonton (City of Edmonton, 2003), mail-out surveys of business establishments in Peel Region (Roorda et al., 2007) and Durham Region (in progress), and a mail-out O-D driver survey in Vancouver. Trade-offs exist between the higher quality of data obtained from in-person interviews and the lower cost of mail-out surveys.
2	Economic elements of personal travel, including income and costs of travel and parking	Household income questions are commonplace in household travel surveys. Income questions are considered to be sensitive by some, resulting in lower response rates to this type of question. Nevertheless, asking the question at the end of the survey and in the form of a categorical question are methods that improve the question response rate and prevent bias to the rest of the survey. The Montreal household travel survey (AMT, n.d.), and the US National Household Travel Survey (NHTS) (US NHTS, 2004), are two examples of major surveys that do include an income question. Costs of travel are imputed, with greater precision in surveys where the type of vehicle is collected.
3	Information about the types and ages of vehicles owned by the household	The US NHTS (US NHTS, 2001) asks questions about the make, model and vintage of vehicles owned by the household (in addition to determining which household vehicle was used for each trip), allowing for much more refined analysis of fuel economy, fuel costs, and emissions.
4	Non-motorized personal trips to/from non-work/non-school trip purposes	The Montreal household travel survey and the US NHTS do not restrict the collection of non-motorized trips to work and school destinations. As of 2001, the US NHTS includes a specific reminder to include walk and bike trips for trips that start and end in the same place.
5	Travel of children under the age of 11	The Montreal household travel survey collects travel information for children 5 years old and older. The US NHTS collects travel information (by proxy) for all children, as of 2001. In the TTS, travel of children under 11 years of age is only collected indirectly (and incompletely) through the reporting of "daycare" or "facilitate passenger" trips of parents.
6	Weekend travel, and seasonal variation in	Weekend travel surveys have been conducted in Calgary (IBI Group, 2002), as well as other cities in the US (see Hunt et al, 2005). Little information

	travel	appears to be collected showing seasonal variations in travel, aside from traffic counts and intercity travel survey data from the Travel Survey of Residents of Canada (see below).
7	Tourism and intercity business travel	The Travel Survey of Residents of Canada (Statistics Canada, n.d) is a quarterly survey that obtains information about intercity travel behaviour of Canadian residents over a one month period, however, geographic details of tourist related trips are not collected. Similarly, the US NHTS (US NHTS, 2004) includes a module in which trips over 50 miles are collected. Tourist travel information at a refined level of geographic detail within the GGH appears to be very difficult to collect.

3. ADDRESSING CHALLENGES FACING PERSONAL TRAVEL SURVEYING

The previous section offers a broad overview of transportation data collection priorities and needs in the GGH. This section focuses on the main challenges facing future implementations of personal travel surveys in the GGH and recommends strategies to address such challenges.

What are the Challenges?

Like any personal travel survey, the TTS involves selecting a representative sample of the population and subsequently contacting the sample subjects to gather the desired data using an appropriate survey instrument. The challenges facing the future conduct of the TTS, as well as similar travel surveys, are related to sample selection issues, contact and recruitment methods, and survey instruments used for gathering the data.

Sample Selection Challenges

Sample selection is typically done by drawing a random sample from a comprehensive "sampling frame" which should ideally include a list of all population units being surveyed. In the case of the TTS, which is a household travel survey, the sampling frame used thus far has been the directory of residential telephone land-lines in the survey area. Although this sampling frame provided in the past an adequate base to draw a representative sample, it is increasingly becoming an incomprehensive list of all households in the survey area, seriously affecting the representativeness of the sample. Several emerging developments have contributed to this problem.

A major contributing factor has been the growing number of households with no land-lines, where members rely solely on their cellphones. Cellphone-only households are not listed in residential phone registries, and are therefore underrepresented in TTS-type samples. The growing numbers of cellphone-only households and the resulting sampling issues are noted in the literature on survey design of many developed regions (e.g. Bricka et al., 2007, Abi-Habib et al., 2003). Research has shown that the socioeconomic characteristics and trip patterns of individuals in cellphone-only households are different from those with land-lines, which makes this a significant sampling issue due to the potential sample bias that might be introduced if not treated carefully. Russell *et al*. (2004) found that

households with no land-lines are more likely to have less than a high school education, have children in the household, be younger than 35 years old, be males, have no household vehicles, and have a lower than average household size.

Another issue with phone land-line sampling is the growing subscription to phone services through Voice over Internet Protocol (VoIP). VoIP is a technology that allows users to make voice calls using a broadband internet connection instead of a regular (or analog) phone line service. VoIP providers allow users to keep their phone numbers, including the original area code, when moving to a different city or country. This causes issues for surveys using a phone land-line directory sampling frame, as households outside the survey area may be contacted while households with external phone numbers but residing in the survey area may not be.

It is difficult to predict accurately the use of the above technologies over the next decade. Currently, cellphone-only households are probably the largest group of households without listed land-line phones. Since such households are more likely to live in apartments and to consist of young residents, survey samples drawn exclusively from land telephone lists will under-represent these household types and their members, introducing some bias into the socioeconomic and travel characteristics of the sample. This warrants an explicit treatment of sampling cellphone-only households to avoid sample bias. At this point, there are no research results to show whether VoIP only households have distinct characteristics that might introduce sample bias without explicit treatment.

The authors expect the challenges associated with the selection of a representative sample from land telephone lists to grow in significance over the years as the above technologies and services take a bigger hold of the market.

Respondent Contact and Recruitment Challenges

Even if a representative sample is selected, there remain several issues related to how survey subjects are contacted/recruited and how the data are gathered. The TTS is conducted through a telephone interview with a resident of each household in the sample. In the early versions of the TTS, households were contacted directly by phone without any prior notification. The proliferation of telemarketers and growing use of call screening services have posed a major challenge to keep response rates of travel surveys at reasonable levels. As such, the TTS has started since 1996 a practice of sending an invitation letter by regular mail to each household in the sample prior to the interview in order to explain the objective and significance of the survey and to specify the targeted day for the interview. The invitation letter has proven to be effective in improving the response rate, specifically of households living in single family housing units. Nevertheless, one lingering challenge is to contact the apartments in the sample by regular mail prior to the interview. This is a problem because of the lack of information on apartment numbers in the used sampling frame (i.e. telephone land-line lists), so letters sent to apartments are not forwarded to the intended dwelling units but instead are usually kept in a common area in the apartment building. Some letters may be noticed and picked up by dwellers of the targeted apartments while others are not.

Another challenge has been making successful phone contacts for interview and reminder purposes with apartment households, since young highly-mobile apartment dwellers are hard to contact at home.

The above discussion points to the fact that representative survey samples are becoming increasingly harder to select, and sample subjects are proving more challenging to successfully recruit and interview using telephone as the main survey instrument. This is a particularly acute problem for households living in apartments and where mainly young people reside.

Survey Instrument Challenges

Another limitation of the TTS telephone interview method is its collection of retrospective data of the interviewee's travel on the previous day. In addition, the interviewee is asked to report (by proxy) on the previous-day travel of each other household member. The retrospective and proxy reporting employed by the TTS has long been known to produce inaccurate estimates of travel to destinations other than work and school, namely home-based discretionary trips and non home-based trips. These types of trips, as well as short trips, are known to be under-represented by the TTS due to retrospective and proxy reporting. This travel market has grown over the years in significance (measured by size and impact), and is expected to continue to grow in the future. In 1986, home-based work trips were the single largest sector of personal weekday travel (about 38% of all personal trips), and by 2006 this percentage had declined to 32%. In 2006, home-based discretionary trips became the largest personal travel sector (about 37% of all personal daily trips). This percentage is definitely larger than 37% because of the current under-representation of this travel market in the TTS database. The rise in home-based discretionary and non home-based trips reflects the increasingly complex trip chains and travel activity patterns that people engage in.

Strategies to Address Personal Travel Survey Challenges

As mentioned in earlier parts of this paper, it is very important to maintain, at least in the short term, the general sampling approach and instrument of the TTS in order to allow for trend analysis, to minimize changes in instrument bias, and to support legacy modelling systems. Nevertheless, it is crucial to address the identified challenges so as to maintain and enhance the quality of the collected travel data in the future. This should be done through a gradual approach, phasing in new survey methods and technologies.

The challenges identified above are not unique to the TTS, as demonstrated by a detailed review of the recent travel survey literature (see Roorda & Shalaby, 2008). The review also sheds light on recent efforts and strategies in other parts of the world to address the emerging personal travel survey challenges. This review has shown that several jurisdictions around the world are experimenting with dual frame sampling techniques and with various survey instruments to overcome the limitations of traditional telephone interviews. Below, a summary of those techniques and instruments is provided, followed by a proposed approach for future implementations of the TTS.

Dual-Frame Sampling

Recently, this sampling technique has been implemented successfully in major jurisdictions (specifically Washington and Chicago) as an effective solution to some of the contemporary issues faced in surveys relying on landline-based samples and contact methods. Paskota (2004) suggests that the best approach to dealing with sample selection problems is to combine more than one sampling frame and target different types of people using various methods. There are generally two approaches to capturing cellphone-only households in dual-frame sampling. The first approach involves matching the names and addresses of all cell-phone users (assuming a comprehensive list is available) in the survey area to those in the telephone land-line directory in the same area so as to identify cellphone-only households. Subsequently, a sample of these households is selected and contacted (*via* cell-phone) to complete the survey using the cell-phone or another survey instrument option (e.g. internet) if made available. This sample augments the main sample of households with land-lines which is selected and surveyed using telephone interview or optionally another method (such as the internet). The challenge with this approach is the requirement of a comprehensive cell-phone list, which may not be possible to compile and obtain because of the potential reluctance of the numerous service providers to disclose the lists of their customers. Additionally, it may not be feasible to distinguish between residence-based and business-based numbers, which would pose a challenge to household survey samples, such as that of the TTS.

The second approach involves identifying households without listed land-line services, which is achieved through address matching of a sample drawn from an address-based sampling frame (e.g. Census list of all residential households, compiled list of residential properties from municipal taxation and assessment agencies, or list of residential addresses from Canada Post) against the land-line phone list in the same survey area. Through the matching procedure, it is possible to identify households in the sample without listed land-line phones, which include not only households relying exclusively on cell-phones but also those having VoIP phone services, those subscribing to the DNC registry, and combinations of the three types. The sampled households without listed land-line phones would then be contacted through regular mail. This is known as “passive recruitment” because it does not follow the recruitment letter with another contact by phone (as no phone contact information is available at that point) but relies on the sampled household to respond to the letter. In order to improve the response rate, sampled households would be sent numerous reminders and possibly offered an incentive to provide their contact information. In addition, such households could be offered alternative methods to complete the survey (e.g. cell-phone, internet), which helps improve the response rate.

The Washington Council of Governments Household Travel Survey in 2007 is one example of a dual-frame sample survey, using the address-based method, and has been shown to provide considerable savings in costs compared to a single frame sample with a similar level of precision. In this survey, an address-based sample was obtained, and addresses and names were matched with the list of all land-lines. Following the address matching, sampled households were assigned to one of two groups, those with and those without land-lines. Households in both sample groups were contacted initially through mail, while the latter group was offered a $50 incentive for agreeing to participate, asking

households to send further contact information (Zmud, 2007). Bricka *et al*. (2007) provides a description and assessment of the Chicago Regional Household Travel Inventory, which also had a similar dual-frame sample.

New Survey Instruments

As noted above, new survey instruments have been used in some surveys to supplement traditional telephone interview methods, mostly when dual frame sampling is employed and in some cases as the sole method. Like the telephone survey instrument, new instruments have their advantages and disadvantages, and none is superior in the absolute sense. As such, targeting different sample sub-groups in a single survey with different instruments and offering each sub-group multiple instrument options to complete the survey should have a positive effect on the survey response rate and should minimize sample bias. Below, new survey instruments are introduced and the advantages and disadvantages of each are discussed briefly.

Internet Lately, there has been a growing interest in web/internet-based surveys for travel data collection. Several researchers have outlined the benefits of such surveys, but also pointed out risks and challenges that need to be carefully addressed. Advantages of using internet-based surveys mentioned in various papers include the following:

- The cost of conducting these surveys (including personnel, communication and data acquisition) is relatively low;
- They enable the incorporation of interactive features, visual aids, animation, automated skip patterns and randomization of questions;
- They have good potential to capture non-respondents to conventional travel surveys (young and busy people);
- They enable automated data entry and checking;
- Interviewer bias is avoided;
- They provide privacy to respondents; and
- They result in quick response times.

Several papers suggest that incorporating internet-based questionnaires to a multi-instrument survey can help attract younger and more affluent respondents, while using them exclusively is advised against (Adler et al., 2002, Alsnih, 2004, Dijst et al., 2006, Potoglou & Kanaroglou, 2008).

Despite the above advantages, internet-based surveys bear some potential risk that need to be addressed and minimized in practical applications in order to reduce socioeconomic and demographic bias. These concerns include:

- Inadequate data quality due to misleading/inaccurate/incomplete information by some respondents who may be frustrated with the survey forms, and by some who may view the internet as means of entertainment (Alsnih, 2004, Dillman & Smyth, 2007).
- Server availability and browser compatibility (Alsnih, 2004, . Li & Shalaby, 2008).
- Access to a computer and the internet, computer literacy, and familiarity with the software (Alsnih, 2004).

Assuming the internet is used to complement other instruments in any given survey, comparability of data can affect the quality of results and the potential for trend analysis. Some research suggests that if responses are similar across instruments, lower response rates in internet-based surveys would not be as critical (Alsnih, 2004*,* Manfreda et al., 2008). Consistency checks, reminders, and graphical shortcuts available in internet questionnaires have resulted in more trips being reported than through either telephone interview or mail-back (Adler et al., 2002). Also, differences between visual and audio stimuli as well as the significant effect of design elements on how respondents interpret and answer questions must be taken into consideration (Dillman & Smyth, 2007). Several papers recommend that further research be done into whether data collected using different survey media result in comparable data.

Cell-Phones Cell-phones have also attracted recent attention as a medium for travel data collection. The greatest potential of this medium is to capture cellphone-only households and young people who are the main residents of such households. Possible recruitment methods include postal mail, cell-phone calls and SMS (Short Messaging Services).
Cell-phones share many features with land-line phones as a tool for travel data collection. However, there are some differences between the two instruments. Keeter and Kennedy (*2006*) have conducted a study on the feasibility of conducting a telephone survey in a cell-phone sampling frame. Results of this research suggest that such surveys are feasible, but they are more difficult and expensive to conduct than land-line surveys. As part of a pilot survey, it was found that while it was easier to contact individuals through cell-phones, the rate of cooperation was about 30%, compared to 50% in land-line phone contacts. The following is a list of potential issues with surveying individuals/households using cell-phones:

- Charges associated with the calls and the need for offering incentives. Research by Yuan *et al*. (*2005*) suggests that higher incentives result in higher response rates.
- Safety of respondents in case they are involved with another activity when contacted, (e.g. driving).
- Privacy of conversations, since the person might be in a public location when interviewed.
- Higher percentage of ineligible individuals (e.g. very young cell-phone owners), compared to land-line samples.

Some research has been conducted on improving cell-phone surveys. Brick *et al*. (2006) found that cellphone-only households are more likely to respond to cell-phone surveys than households that have both types of service. In order to avoid non-response bias, households with both phone types should be identified and contacted through their land-line service only.

GPS-Based Surveys GPS (Global Positioning System) technologies have recently seen growing levels of interest in their application to travel and activity data collection. Over the past decade, GPS-enabled devices have enjoyed a continuing trend of improved accuracy, lighter weight, better power management and cheaper price, which have given rise to a proliferation of applications in various fields. Over the same period of time, real-world applications of GPS to travel data collection have taken place around the world,

accompanied by a surge of research and pilot studies to further develop and enhance enabling tools.

Several recent studies (Li & Shalaby, 2008, Chung and Shalaby, 2005, Tsui and Shalaby, 2006, Stopher and Greaves, 2007, Stopher, 2008) indicate that the main advantages of GPS-based travel/activity surveys are:

- They provide *accurate information* on activity locations, start and end times of trips, travel route paths, and the breakdown of transit trip components (access walk time/distance, waiting time, in-vehicle time, transfer time and egress walk time/distance).
- They collect *complete information* of all trips, addressing the problem of underreporting of trips (specifically short and discretionary trips) typical of conventional methods.
- They enable *ease of response*, because of the reduced respondent burden involved in data collection compared to the conventional travel diary.
- They are appropriate for *data collection over extended time periods*.

GPS-based travel surveys are typically conducted using a combined GPS receiver and data logger held by the survey participant for a specified period of time (e.g. a day). Upon retrieval of the unit, raw data are downloaded and automatically processed using a system of post-processing algorithms that filters the data and decomposes the GPS data trail into activities and trips, with each trip further decomposed into its components (e.g. access walk, waiting, in-vehicle, etc.). Following data processing, a *prompted recall* interview with the participant is typically conducted in order to confirm trip and activity details and collect further information. Lately, there has been a significant amount of algorithm and system development for data processing and prompted recall surveys (see Li and Shalaby, 2008, Chung and Shalaby, 2005, Tsui and Shalaby, 2006, Stopher and Greaves, 2007). Also, some technical issues that challenged GPS-based travel surveys have largely been addressed of late. For example, commercial GPS sensors are now so sensitive that problems of signal detection and loss are no longer a major issue at locations of tall buildings or inside surface vehicles (there are even ongoing technological advancements of in-door GPS). Another challenge has been limited battery life, however, many new GPS units used for surveys have built-in accelerometers that detect motion which allows putting the unit in sleep mode during inactive periods, enabling collection of data for a few days on a single charge.

The remaining challenges and limitations of GPS-based surveys include the respondent burden of conducting the prompted recall interview, privacy, logistics for delivery and pick up of GPS units, timing of prompted recall interview, and survey cost.

GPS-based surveys have been implemented in several jurisdictions around the world, namely Halifax in 2006-2007 with a sample of 2000 households (GPS as the sole instrument) and France in 2007-2008 where GPS has been used to supplement the main survey of 20,000 households.

4. RECOMMENDATIONS

At the outset of this paper, two questions were framed regarding challenges for data collection in the GGH. Our recommendations are structured to respond to these questions.

> Question 1) Will the TTS, in combination with other data sources, provide sufficient information to answer key transportation questions that will arise over the next 10 years?

Strategic expansion and modification of the current data collection program is needed to answer these transportation questions. The study recommends the following:

Recommendation 1) Continue and improve the major established data collection systems including the TTS, the cordon count programs, the MTO commercial vehicle survey and transit surveys.

Recommendation 2) Undertake new or expanded data collection efforts to address important gaps in transportation data for the GGH. The following additional data should be collected either by adding questions to existing surveys, or developing new surveys. This list is ranked in order of priority:

1) Goods movement and service provision by automobile and truck, within the GGH.
2) Economic elements of personal travel, including income and costs of travel and parking.
3) Information about the types and ages of vehicles owned by the household.
4) Non-motorized personal trips to/from non-work and non-school trip purposes.
5) Travel of children under the age of 11.
6) Weekend travel and seasonal variation in travel.
7) Tourist and intercity business travel.

Recommendation 3) Develop an on-going research program that is dedicated to the identification and testing of practical data collection instruments that focus on unmet travel data needs of the GGH.

> Question 2) How should the TTS be effectively and responsibly conducted in the future, recognizing the significant technological and societal changes that are occurring?

Based on a detailed review of the most promising solutions that others have found world-wide, this study recommends that the 2011 TTS be conducted with the following modifications:

Recommendation 4) Develop and implement an internet version of the 2011 TTS. Respondents should be given the option of responding to the TTS either by computer aided telephone interview or through the internet version of the survey. Extensive testing should be undertaken to fully understand and account for the differences in instrument bias between these two retrieval methods.

Recommendation 5) Enhance the 2011 TTS by using a dual-frame sampling approach, as follows:

- A list of all residential households in the survey area should be obtained.
- A land-line telephone list should also be obtained and compared with this list.
- Those residences with listed land-line numbers should be contacted/recruited as in previous TTS implementations (mail-out followed by telephone calls). Those residences without listed land-line numbers should be contacted/recruited by mail.
- If a list of residential cell-phone numbers can be obtained, it should be compared with the land-line telephone list to identify cellphone-only households as well as households with both cell-phones and unlisted home phones. These households should then be contacted/recruited by cell-phone and/or SMS in addition to regular mail.
- The sub-sample of households with listed land-lines should be given the option to undertake the survey either by land-line phone interview or using the internet version of the survey, while the sub-sample of unlisted land-line numbers should be given the options of a phone interview (by cell-phone or land-line phone) and the internet.

Recommendation 6) Conduct a supplementary GPS-assisted survey in parallel with the main 2011 TTS. The survey results should be used to estimate the magnitude of the under-reporting of non-work/school trips and other instrument biases in the main survey.

Recommendation 7) Ensure that the capability for long term transportation trend analysis is not lost, regardless of what changes are made to the TTS. This means that a major portion of the 2011 TTS should be continued with no more than minor changes.

Most of these recommendations involve investment of limited funds, time and energy. The authors of this investigation strongly feel that such investment is warranted given the importance of data for transportation planning decision-making and ultimately for the long-term prosperity, liveability, and sustainability of the GGH.

5. CONCLUSIONS

The objective of this paper has been to investigate potential solutions to the main challenges facing the travel data needs in the GGH with regards to content of data and methodology of data collection.

It is evident throughout this study and review of data collection practices worldwide that the challenges faced in the GGH are not unique to this region. Developing a comprehensive and integrated system of travel data collection is of great importance to answering key transportation planning questions that are expected to arise in the next decades in almost any urban region in the world. The TTS, as currently operated, is faced with growing challenges in sample selection, respondent contact/recruitment and survey instrument non-reporting, primarily because of changes in technology and changing attitudes toward telephone surveys. These exact same issues are being experienced in many other jurisdictions world-wide, and are resulting in changes to the state-of-practice in travel survey methodology. The discussion provided in this paper is informed by recent experiences of

conducting the TTS over the past decade and the insights gathered based on extensive review of the recent travel survey literature elsewhere in the world. As a result, the authors expect that the recommendations made here are widely applicable to data collection programs in other regions of the world comparable with the GGH.

ACKNOWLEDGEMENTS

The authors express gratitude first to two outstanding University of Toronto students that assisted in undertaking the extensive literature review. Their names are Josee Dumont and Zaven Mangassarian. The authors also thank Professor Eric J. Miller and Professor Baher Abdulhai, of the Department of Civil Engineering at University of Toronto for their insightful comments. Funding was provided by the Data Management Group, however, the recommendations in this paper are based on the independent research and opinions of the authors.

REFERENCES

Abi-Habib, N., Safir, A., and Triplett, T. (2003) 2002 National Survey of America's Families: Methods and Data Reliability. The Urban Institute, Washington, DC.

Adler, T., Rimmer, L. ,and Carpenter, D. (2002) Use of Internet-Based Household Travel Diary Survey Instrument. Transportation Research Record: Journal of the Transportation Research Board, No. 1804, Transportation Research Board of the National Academies, Washington D.C., pp. 134-143.

Alsnih, R. (2004) Characteristics of Web Based Surveys and Application in Travel Research. Presented at 7th International Conference on Travel Survey Methods, Costa Rica.

AMT (n.d.) Montreal Household Travel Survey. Aaccessed from www.cimtu.qc.ca/EnqOD/Index.asp on July 28, 2009.

Brick, J.M., Dipko, S., Presser, S., Tucker, C. and Yuan, Y. (2006) Nonresponse Bias in a Dual Frame Sample of Cell and Landline Numbers. Public Opinion Quarterly, 70, pp. 780-793.

Bricka, S., Sen, S., and Arce, C. (2007) Chicago Regional Household Travel Inventory – Sampling Plan. Unpublished manuscript. Accessed from www.nustats.com/chicago. on June 2, 2008

Chung, E. and Shalaby, A. (2005) A Trip Reconstruction Tool for GPS-Based Personal Travel Surveys. Transportation Planning and Technology, 28(5), pp. 381-401.

City of Edmonton. (2003) Edmonton Regional Commodity Flow Study. Project Report. Edmonton Alberta.

Dijst, M., Farag, S.,and de Blaeij, A. (2006) Effects of Data Collection Methods in Travel Behaviour Surveys; Comparing an Internet and a Mail Sample. Presented at 85th Annual Meeting of Transportation Research Board, Washington, DC.

Dillman, D.A. and Smyth, J.D. (2007) Design Effects in the Transition to Web-Based Surveys. American Journal of Preventive Medicine, 32(5).

DMG. (2007a) Transportation Tomorrow Survey 2006 Data Guide. Data Management Group, University of Toronto, Report 112.

DMG. (2007b) Greater Toronto Area Cordon Count Summary: Analysis of Traffic Trends: 1985-2004. Data Management Group, University of Toronto, Report 108.

Hunt, J.D., McMillan, P., Stefan,, K. and Atkins, D. (2005) Nature of Weekend Travel by Urban Households. Paper presented at the 2005 Annual Conference of the Transportation Association of Canada. Calgary, Alberta.

IBI Group (2002) City of Calgary – Transportation: 2001 Household Activity Survey: Preliminary Data Expansion and Validation. Report prepared for the Transportation Planning Business Unit, City of Calgary, Canada.

Keeter, S. and Kennedy, C. (2006) The Cell Phone Challenge to Survey Research. Unpublished manuscript. Retrieved from people-press.org/reports/pdf/276.pdf

Li, Z. and Shalaby, A. (2008) Web-based GIS System for Prompted Recall of GPS-assisted Personal Travel Surveys: System Development and Experimental Study, Presented at the 87th Annual Transportation Research Board Meeting, Washington D.C.

Manfreda, K.L., Bosnjak, M., Berzelak, J., Haas, I., and Vehovar , V.(2008) Web Surveys Versus other Survey Modes: A Meta-Analysis Comparing Response Rates. International Journal of Market Research, 50(1), pp. 79-104.

Paskota, M. (2004) Sample Design and Survey Error. Presented at 7th International Conference on Travel Survey Methods, Costa Rica.

Potoglou, D., and Kanaroglou, P. (2008) Comparison of Phone and Web Based Surveys for Collecting Household Background Information. Presented at the 8th International Conference on Travel Survey Methods, France.

Roorda, M.J. and Shalaby, A. (2008). Transportation Data Collection in the Greater Golden Horseshoe: A Framework and Priorities for Improvement. Data Management Group, University of Toronto.

Roorda, M.J., McCabe, S., and Kwan, H. (2007) A Shipper-Based Survey of Goods and Service Movements in the Greater Golden Horseshoe: Report I Survey Design and Implementation (Final Report). MTO and Region of Peel.

Russell, N., Bose, J.,and Giesbrecht, L. (2004) Nonresponse Bias in a Travel Survey of Nontelephone Households. Unpublished manuscript.

Statistics Canada. (n.d.) Travel Survey of Residents of Canada, quarterly.

Stopher, P.R. (2008) Collecting and Processing Data from Mobile Technologies. Presented at the 8th International Conference on Travel Survey Methods, France.

Stopher, P.R.and Greaves, S. (2007) Household Travel Surveys: Where are we Going? Transportation Research Part A: Policy and Practice, 41(5), pp. 367-381.

Tsui, A. and Shalaby, A. (2006) An Enhanced System for Link and Mode Identification for GPS-Based Personal Travel Surveys, Transportation Research Record: Journal of the Transportation Research Board, No. 1972, Transportation Research Board of the National Academies, Washington, D.C., pp. 38-45.

US NHTS. (2004) 2001 National Household Travel Survey User's Guide. Accessed from nhts.ornl.gov/publications.shtml#usersGuide on July 28, 2009

Yuan, A. Y., Allen, B., Brick, J.M., Dipko, S., Presser, S. and Tucker, C. (2005) Surveying households on cell phones — Results and lessons. Presented at the 60th Annual Conference of the American Association for Public Opinion Research, Miami Beach, FL.

Zmud, J. (2007) Washington Full Survey Design Documentation. Unpublished Manuscript, NuStats, Austin, Texas.