Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 32 (2018) 482–494

Transportation Research Procedia

www.elsevier.com/locate/procedia

International Steering Committee for Transport Survey Conferences

An online survey to enhance the understanding of car drivers route choices

Hamzeh Alizadeh[a]*, Pierre-Léo Bourbonnais[a], Catherine Morency[a], Bilal Farooq[b], Nicolas Saunier[a]

[a]*Department of Civil, Geological and Mining Engineering, Polytechnique Montréal, C.P. 6079, succ. Centre-Ville Montréal, H3C 3A7, Canada*
[b]*Department of Civil Engineering, Ryerson University, 350 Victoria Street, Toronto, M5B 2K3, Canada*

Abstract

The increasing application of advanced behavioral choice models, reflecting the stochasticity of individuals' preferences and the complex nature of human decision-making process, requires enhanced data collection methods to obtain detailed data without significantly increasing the respondent burden. In this study, we present the development and deployment of a general data collection framework adapted for behavioral route choice studies. The main objectives of the proposed framework are to observe drivers' route choices, and to identify important factors, including observable attributes and latent behavioral traits, affecting those decisions. More specifically, the survey has been designed to reveal drivers' consideration set of route alternatives from which they pick their final choices. The detailed analysis of survey's response behavior will help improve the framework to gather travel data even more efficiently.

Keywords: Revealed preference survey, Route choice modeling, Consideration sets, Survey response behavior

1. Main text

Demand in a road network, i.e. the traffic flow, is the direct result of individuals' travel and route choice decisions. These decisions are captured through route choice models, which are considered as an important part of traffic

* Corresponding author. Tel.: +1-(514)-632-2381.
E-mail address: hamzeh.alizadeh-aliabadi@polymtl.ca

10.1016/j.trpro.2018.10.042

assignment procedures, and play a key role in transportation planning to forecast the traffic flow, design new transportation infrastructures, and investigate new policies. It is therefore of utmost importance to understand drivers' route choice behaviors and factors affecting them. Route choice is probably one of the most complex and challenging problems in traffic assignment. The complexity of this process is partly due to factors such as the sophisticated nature of human behavior, the ambiguity of the decision-making process, and the stochasticity of individuals' preferences.

Previous route choice studies have mostly focused on the effects of observable factors on drivers' route choice decisions (Alizadeh, Farooq, Morency, & Saunier, 2017; Dalumpines & Scott, 2017; Jan, Horowitz, & Peng, 2000). Instances of these factors include route features (i.e. travel distance, number of turns, etc.) and drivers' characteristics (such as age, gender, etc.), which are tangible and can be directly observed. However, numerous studies have shown that latent factors such as attitudinal traits, perceptions and lifestyle preferences play a major role in the decision-making process. For instance, in addition to the observable factors, drivers' decisions might also be influenced by factors such as safety concerns, driving habits and spatial abilities (Alizadeh, Farooq, Morency, & Saunier, 2018; Prato, Bekhor, & Pronello, 2012; Sarkar & Mallikarjuna, 2017; Walker, 2001).

The complexity of route choice modeling is also attributed to the high density of the road network, the large number of possible alternatives between OD pairs, and the correlations among these alternatives. Since it is not computationally feasible to enumerate all the possible routes connecting a given OD pair (i.e. the universal choice set) in a real world road network, nor behaviorally accurate to assume that drivers are aware of all of them, a two stage choice modeling process is usually adopted (Manski, 1977). In the first stage of this process, a subset of the universal choice set is selected to form the collection of feasible travel alternatives considered by the driver (i.e. the considered choice set). Then, in the second stage, drivers make their final route choices from the considered set of route alternatives. However, defining a proper consideration set is a serious challenge in route choice modeling. The consideration set of route alternatives is usually latent to the analyst and alternatives are therefore usually generated using variations of the shortest path algorithm (Hoogendoorn-Lanser, van Ness, & Bovy, 2005; Prato et al., 2012). The generated set should include alternatives that are attractive to the driver in a real world choice situation and the misspecification of the size and composition of the considered choice set greatly affect model's estimates and may lead to fallacious predicted demand levels (Bliemer & Bovy, 2008; Geda, 2014; Nadine Schuessler & Axhausen, 2009; Peters, Adamowicz, & Boxall, 1995; Prato & Bekhor, 2006, 2007; Swait & Ben-Akiva, 1987).

Considering the abovementioned challenges of route choice modeling, and in order to improve the estimation and prediction of drivers' route choice decisions, it is imperative to first, observe drivers' revealed preferences in real route choice situations, second, to identify behavioral and attitudinal factors as additional sources of heterogeneity affecting their decisions, and finally, to get a better grasp of the formation process, size and composition of drivers' considered sets of route alternatives. Accordingly, this study proposes a framework of data collection for route choice studies, with the objective of satisfying the aforementioned requirements.

The remainder of the paper is structured as follows. First, we briefly review some of the previous route choice studies and their data collection methods, and in that context further clarify the contributions of the presented data collection framework. Next, we present the proposed survey framework and its implementation. Survey participants, their response behaviors, completion rates and dropouts are discussed in the next section. In the end, we highlight the possible implications and applications of this survey framework, underscore its limitations, and suggest further research directions.

Nomenclature

Behvr.	Behavioral traits
CB	Computer-Based
CCS	Considered Choice Set
Fact.	Factors
GMA	Greater Montreal Area
HTS	Household Travel Survey
Obs.	Observed choices
OD	Origin - Destination
PB	Paper-Based
Percp.	Perception
RP	Revealed Preference

SP	Stated Preference
WB	Web-Based

2. Previous Studies

Table 1 summarizes the data collection methods adopted in some of the previous route choice studies and enumerates the attributes that have been found to significantly affect drivers' decisions. Also, it puts into perspective the specifications and characteristics of the proposed framework. Although the table does not encompass all the previous route choice studies, this list has been selected to provide a wide spectrum of research on that matter.

Studies have been compared using four criteria, namely Medium, Method, Collected data, and Significant attributes. Medium refers to the type of interface that has been used to collect the data. Different types of media have been used during the past few decades. Telephone, mail, face-to-face, and web surveys are among the typical methods that have been extensively used to collect choice data. In order to reduce the respondent burden, computer-assisted self / telephone interviewing have been adopted (Dhakar, 2012; Papinski, 2011; Srinivasan & Dhakar, 2013). In Table 1, four types of interfaces have been identified for route choice data collection. Household Travel Survey (HTS) data, as a traditional source of data, has been used in few studies. Such diaries are not very effective in collecting detailed, long-term or large scale route choice data, due to the excessive respondent burden of declaring the exact routes (Chen, 2013). Therefore, the detailed trajectory is usually not available in HTS data, and a shortest path algorithm (based on some generalized cost function) is used to simulate the chosen route. Paper-based (PB), computer-based (CB), and web-based (WB) route choice surveys are among other types of data collection media adopted in route choice studies.

Method indicates the methodology to observe and quantify respondents' preferred choice. Route choice surveys are either Revealed Preference (RP), in which respondents reveal their actual choices in real route choice situations, or Stated Preference (SP), in which respondents are asked to choose between several hypothetical route alternatives based on some provided details on each choice. In recent years, the prevalent use of GPS technology has provided researchers with an abundance of high-resolution geospatial data. An important advantage of GPS data collection over other methods is that it can record travel information for several days without any additional respondent burden. However, working with GPS data brings several complexities including the large size of the dataset, the absence of data due to signal loss, the challenges in constructing a representative sample, and the technological issues such as battery life and record accuracy amongst others. Furthermore, studies based on GPS data often lack personal information on the decision maker, his attitudes, experiences, and preferences. Even though GPS traces can be considered as RP data, we considered them as a separate method of data collection because of their different nature and data processing requirements, and their prevalence of use in route choice studies.

Regardless of the method used, the main objective of all the above-mentioned data collection methods is to record the observed choices (Obs.). The other types of collected data depends to a large extent on the objectives of the survey. For instance, to make an in-depth analysis of the effect of behavioral traits (Behvr.), attitudinal questions and psychometric indicators should be the focus of the survey, while to analyze respondents' perception bias towards a particular choice, questions regarding the perceived values of different attributes (Percp.) is of prior importance. It is also a common procedure to ask respondents to reveal the most important factors affecting their choices (Fact.).

The observation of the considered choice set (CCS) of route alternatives, from which drivers make their final choices, is not as straightforward, and hence, is not very common in practice. In SP surveys, participants make their choices from a series of hypothetical alternatives provided by the analyst, while in GPS surveys, the consideration set of route alternatives mostly remains unidentified. Moreover, the specification of the considered choice set is usually ignored in RP surveys to reduce the response time as well as the respondents' burden.

Finally, factors that have been found to significantly affect drivers' route choice process are compared in the last column. The variety of factors that have been found to significantly affect drivers' route choice behavior further illustrates the importance that survey design should be in line with the objectives of the survey and the expected results.

3. Survey Framework

In this section, we present the development and implementation of the proposed revealed preference web-based survey, designed to observe drivers' revealed route choices towards their most frequently visited destinations, and

identify behavioral and attitudinal factors affecting them. We also intend to observe drivers' consideration sets of route alternatives and to characterize them based on drivers' perceptions.

Drivers residing and driving in the Greater Montreal Area (GMA) have been targeted for this study. This area covers approximately 9840 square kilometers and contains a population of roughly 4 million inhabitants (Transport, 2013). Since it is a bilingual region with both French and English speaking populations, the survey was prepared in both languages. In order to decrease the respondent burden, mitigate the implementation cost, and enhance the data quality, a high performance front-end user interface with an elaborated graphic design has been adopted. For more details on the design of the interface, the reader is referred to (Bourbonnais & Morency, 2013).

To minimize the complexity of questions, where respondents had to specify the origin and destination points of their trips and trace the considered routes, geographical map interfaces were adopted. Moreover, an internal validation mechanism was designed to maximize the quality and completeness of the recorded data, and to minimize the data cleaning effort, by reducing participant errors. In this process, several validation criteria were defined for each question, and responses were required to comply with all the criteria in order to be approved and stored in the database. It should be noted that, to advance to a next section, all the responses in that section should be accepted by the internal validation process. In other words, respondents are required to satisfy all the validation criteria of a particular section to be able to advance to the next section. A red exclamation mark appears beside questions that do not meet the required validation criteria, along with a message box explaining the reasons for which the given answers are not acceptable.

Seven types of questions are used in the design of the survey:

- Dichotomous: provides two options for a statement to select from.
- Text box: requires respondents to enter a number, a text, or a combination of both.
- Select: provides a list of choices, from which respondents can select only one of the several options.
- Multi-select: allows respondents to choose more than one option from a list.
- Slider: Likert scale questions are used to measure attitudes, opinions, perceptions, and levels of agreement with a statement. We adopt a Slider with a continuous scale to obtain more precise recordings.
- Map-point: location (such as origin or destination points) needs to be specified on the map.
- Map-route: respondents are required to specify their routes by dragging and moving an automatically generated route between the predefined OD pairs.

The whole survey is divided into six separate sections, namely Profile, Home, Trips, Routes, Preferences, and End. In the first section, Profile, we collect typical sociodemographic data, such as age, gender, educational attainment, type of work, and salary. Collecting these data provides the possibility of comparing the sample population with the reference population (Ory & Mokhtarian, 2005), and to segment the population, based on factors affecting individuals' route choices. Respondents are also asked to specify the duration that they have been living in the GMA and to indicate their general familiarity with its road network.

In the following section, Home, participants are required to provide their home address. A geographical map is also provided, which geolocates the specified address. The provided address should be precise enough to be automatically pinpointed on the map. Participants can further adjust the pinpointed location on the map by moving the marker to the exact location. This section also includes questions regarding the household size, the number of cars in the household, the duration of living at the same address, and the familiarity with the road network around the specified address. It is worth mentioning that the exact home address is required to explore factors such as the accessibility to the road network, availability of transit services, and land use specifications.

The third section, Trips, also consists of a geographical map, on which respondents specify the destination point to which they drive most frequently, such as work places, shopping malls, etc. They are then required to indicate their familiarity with the road network around the specified destination point, and the purpose for making the declared trip. Respondents are also asked to specify why they have used their cars to make the declared trip, as well as to select the five most important factors affecting their route choices for that particular trip. Moreover, they are asked to select from a list, the type of information that they consult prior to their trip and on the way, if any. This section ends by asking respondents to provide the number of route alternatives that they consider for the declared trip, up to a maximum of five routes. This number was based on the results of the pilot study that took place before the main data collection. Respondents are then asked to specify these alternatives on a map and provide more details regarding their preferences towards them in the next section of the survey. Table 2, provides more details regarding questions in the first three sections.

Table 1: Comparison of selected route choice studies and their data collection methods.

Study	Medium				Method				Collected data					Significant Attributes *
	HTS[a]	PB[b]	CB[c]	WB[d]	RP[e]	SP[f]	GPS	Obs[g]	CCS[h]	Demo[i]	Fact[j]	Behvr[k]	Percp[l]	
Iida, Uno, and Yamada (1994)	☐	☐	☒	☐	☐	☒	☐	☒	☐	☐	☐	☐	☐	28, 30
Abdel-Aty and Jovanis (1997)	☐	☒	☒	☐	☒	☒	☐	☒	☐	☐	☒	☐	☒	1, 2, 3, 4, 5, 6
Peeta, Ramos, and Pasupathy (2000)	☐	☒	☐	☐	☐	☒	☐	☒	☐	☒	☐	☒	☐	9, 12, 28
Cascetta, Russo, Viola, and Vitetta (2002)	☐	☒	☐	☐	☒	☐	☐	☒	☒	☒	☐	☐	☐	1, 12, 17
Parkany,Du,Aultman-Hall,andGallagher (2006)	☐	☒	☐	☐	☐	☐	☒	☒	☐	☒	☒	☒	☐	3, 9, 10
Cools, Moons, and Wets (2009)	☒	☐	☐	☐	☒	☐	☐	☒	☐	☒	☐	☐	☐	1, 9, 10, 23
Papinski, Scott, and Doherty (2009)	☐	☒	☒	☒	☐	☐	☒	☒	☐	☒	☒	☐	☐	1, 3, 4, 6, 21, 22
Ben-Elia and Shiftan (2010)	☐	☐	☒	☐	☐	☒	☐	☒	☐	☒	☐	☐	☐	14, 15, 25, 29
Tawfik, Rakha, and Miller (2010)	☐	☒	☐	☐	☒	☐	☐	☒	☐	☒	☐	☒	☒	1, 6, 9, 15, 24, 25
Schlaich (2010)	☐	☐	☐	☐	☐	☐	☒	☒	☐	☐	☐	☐	☐	3, 28
Prato et al. (2012)	☐	☐	☐	☒	☒	☐	☐	☒	☒	☒	☐	☒	☐	1, 5, 6, 7, 9, 13, 27
Kaplan and Prato (2012)	☐	☐	☐	☒	☒	☐	☐	☒	☐	☒	☐	☒	☐	1, 5, 6, 7, 9, 13, 27
Gan and Chen (2013)	☐	☐	☐	☐	☐	☒	☐	☒	☐	☒	☐	☐	☐	14, 15, 16
Jou and Yeh (2013)	☐	☐	☐	☐	☐	☒	☐	☒	☐	☒	☐	☐	☐	1,5,6,8,10,12,18,19
Tawfik and Rakha (2013)	☐	☒	☐	☐	☒	☐	☐	☒	☐	☐	☐	☒	☒	1, 6, 9, 15, 24, 26, 27
Ramaekers, Reumers, Wets, and Cools (2013)	☐	☒	☐	☐	☐	☐	☒	☒	☐	☒	☒	☐	☐	8, 9, 10, 11
Koller-Matschke, Belzner, and Glas (2013)	☐	☐	☐	☒	☒	☒	☒	☒	☐	☐	☐	☐	☐	1, 18, 28, 29
Habib, Morency, Trépanier, and Salem (2013)	☒	☐	☐	☐	☒	☐	☐	☒	☐	☒	☐	☐	☐	1, 8, 10, 20
Vacca and Meloni (2014)	☐	☐	☐	☐	☐	☐	☒	☒	☐	☒	☐	☐	☐	1, 5, 9, 12, 15, 22, 24
Hess, Quddus, Rieser-Schüssler, and Daly (2015)	☐	☐	☐	☐	☐	☐	☒	☒	☐	☐	☐	☐	☐	1, 6, 11, 19
Manley, Addison, and Cheng (2015)	☐	☐	☐	☐	☐	☐	☒	☒	☐	☐	☐	☐	☐	20
Lai and Bierlaire (2015)	☐	☐	☐	☐	☐	☐	☒	☒	☐	☐	☐	☐	☐	5, 6, 22
Dalumpines and Scott (2017)	☐	☐	☐	☐	☐	☐	☒	☒	☐	☒	☐	☐	☐	1, 5, 6, 7, 11
Alizadeh et al. (2017)	☐	☐	☐	☐	☐	☐	☒	☒	☐	☐	☐	☐	☐	4, 5, 6, 20
Proposed Framework	☐	☐	☐	☒	☒	☐	☐	☒	☒	☒	☒	☒	☒	-

***Factors affecting route choice**

1- Travel time	7- Number of turns	13- Delay	19- Toll rate / Cost	25- Learning process
2- Travel time reliability	8- Time of day	14- Network familiarity	20- Anchor points	26- Drivers' categories
3- Traffic conditions (level of service)	9- Socio-demographic	15- Driving experience	21- Stop signs	27- Personality traits
4- Number of segments	10- Trip purpose	16- Education	22- Traffic lights	28- Availability of information
5- Percentage of highway	11- Road type	17- Topology	23- Holidays	29- Type of information
6- Travel distance	12- Socio-economic	18- Choice inertia	24- Travel speed	30- Quality of information

[a] Household Travel Survey [b] Paper Based [c] Computer Based [d] Web based [e] Revealed Preference [f] Stated Preference [g] Observed Choices
[h] Considered Choice Set [i] Demographics [j] Revealed Factors [k] Behavioral traits [l] Perception

Table 2. List of questions in sections Profile, Home, and Trips.

ID	Question Type	Description
First Section: Profile		
101	Dichotomous	Gender
102	Text box / Select	Age / Age group
103	Select	Educational attainment
104	Select	Main occupation
105	Text box	Age of first driving licence
106	Dichotomous	If question ID 104 equals "Worker" è Whether work on the road regularly
107	Dichotomous	If question ID 104 equals "Worker" è Whether work at home regularly
108	Dichotomous	If question ID 104 equals "Worker" or "Student" è Whether flexible arrival time
109	Select	Living time in Montreal
110	Slider	Familiarity with Montreal road network
Second Section: Home		
201	Text box	Postal code
202	Text box	Apartment number (optional)
203	Text box	Street Address
204	Text box	City
205	Map-point	Home location
206	Text box	Household size
207	Text box	Household vehicle number
208	Select	Living time at the specified address
209	Slider	Familiarity with the road network of the neighborhood they live in
Third Section: Trips		
301	Map-point	Specify destination point
302	Slider	Familiarity with the road network around the destination
303	Dichotomous	Is origin home location?
304	Map-point	If question ID 303 equals "No" è Specify origin point (if not home location)
305	Slider	Familiarity with the road network around the specified origin
306	Select	Purpose of the trip
307	Text box	Frequency of driving to the specified destination by car (per week)
308	Multi-select	Why choose car for this trip?
309	Multi-select	Factors affecting route choice to this destination

Route alternatives are specified in the fourth section, Routes. First, an automatically generated route connecting the predefined origin and destination points appears on a geographical map. Then, respondents are required to drag the generated route and adjust it to match their actual considered route. Every time respondents drag the route to a new place on the map, a way point is created on that new location. A minimum of three way points are required for the route to be validated by the internal validation process of the questionnaire. Specified routes are followed by several questions, to gather more details on their main features (see Table 3). At first, respondents are required to indicate how frequent they use the declared alternative, on a five level Likert scale ranging from rarely to frequently. Then they provide information regarding the day (i.e. weekdays / week-ends) and the specific time period, during which they start the trip. They also indicate the importance of habit and the effect of weather conditions on their use of the declared route. Moreover, respondents are asked to pinpoint their regular stop (if any) on a map, and specify the amount of toll paid for that particular trip. Finally, drivers' perception regarding the characteristics of the declared alternatives is evaluated based on several factors such as travel time and its reliability, safety, traffic conditions, scenery, pavement quality, and the number of traffic lights.

The fifth section of the survey, entitled Preferences, focuses on behavioral and attitudinal variables affecting drivers' route choice behaviors. A list of different statements is provided to respondents, who were asked to specify their level of agreement with each statement on a five-point Likert scale ranging from total agreement to total disagreement (see Table 3). These statements are based on psychometric indicators and on some behavioral assumptions on drivers' attitudes, to reveal the most important latent variables affecting drivers' route choice behaviors (Atasoy, Glerum, & Bierlaire, 2013; Ory & Mokhtarian, 2005).

The survey ends with few optional questions in the final section, End (see Table 3). First, respondents are asked to provide their household's gross income level. Then, they are asked to provide their e-mail address if they desire to

participate in other transportation surveys. Finally, respondents can provide their comments and opinions regarding the survey in a blank box.

Table 3. List of questions in sections Routes, Preferences, and End.

ID	Question Type	Description
Fourth Section: Routes		
401	Map-route	Specify considered route
402	Slider	Frequency of using the specified route
403	Select	Weekdays / Weekend
404	Select	Departure time
405	Text box	Perceived travel time
406	Multi-select	Effect of weather conditions
407	Dichotomous	Use tolled route?
408	Text box	If question ID 407 equals "Yes" è Amount of toll
409	Slider	Perception of safety
410	Slider	Perception of scenery
411	Slider	Perception of travel time reliability
412	Slider	Perception of pavement quality
413	Slider	Perception of traffic conditions
414	Select	Perception of the number of traffic lights
415	Slider	Effect of habit in choosing the specified route
416	Dichotomous	Have regular stop?
417	Map-point	If question ID 416 equals "Yes" è Specify the location of the regular stop
Fifth Section: Preferences		
501	Slider	Driving to my destination, I prefer to take freeways, whenever I have access to them.
502	Slider	Driving to my destination, I prefer to take local routes, even when freeways are available.
503	Slider	The pavement quality is an important factor in my route choice.
504	Slider	I always look for shortcuts to minimize the travelled distance.
505	Slider	I do my best to avoid traffic lights.
506	Slider	Minimizing the travel time is my principal goal while choosing my route.
507	Slider	I prefer taking a longer route with a fluid traffic flow rather than being stuck in traffic in a shorter one.
508	Slider	I have the tendency to follow the same route over and over.
509	Slider	I have the tendency to try new routes.
510	Slider	I tend to avoid routes with narrow lanes.
511	Slider	I prefer to take routes with higher speed limits.
512	Slider	I am not comfortable driving next to trucks and I try to avoid them.
513	Slider	I prefer to choose a more beautiful and scenic route, even if it takes longer to get to work.
514	Slider	I prefer to take tolled routes because they are less congested and much faster.
515	Slider	I inform myself about road construction sites to avoid them.
516	Slider	I have the tendency to avoid turns and take the most direct route to get to work.
517	Slider	I have a good sense of direction and I can easily find my way in a road network.
518	Slider	When I'm informed by radio or variable message signs, of an accident causing traffic jam on my route, I change my itinerary and choose an alternative route to avoid the congestion.
519	Slider	I can easily remember a route which I took once.
520	Slider	I use landmarks to remember a route that I took once.
521	Slider	I prefer to choose a route which has a more reliable travel time even if it takes me more time.
522	Slider	I take the route suggested by Google Maps (or other route planners).
Sixth Section: End		
601	Select	Household gross income level
602	Dichotomous	Would like to participate in other mobility surveys
603	Text box	If question ID 602 equals "Yes" è Put e-mail address
604	Text box	General comments on the interview

To evaluate the simplicity and clarity of the questions, to assess the accuracy of the provided directions on how to complete the survey, and to detect the weaknesses of the designed interface, graduate students of the Transportation Research Group of Polytechnique Montreal took part in a pilot test in February 2017. The revised version was launched in March 2017, and data was collected over a period of three months.

4. Survey Response Behavior

By the end of the three-month data collection period, 843 individuals started the survey from which 539 (64 %) completed it, while the remaining 304 (36 %) dropped out at various points of the survey. In this section, we present the survey recruitment methods and the obtained response rates. We also investigate participants' characteristics and their response behaviors. Finally, we explore participants who dropped out of the survey before finishing it.

4.1. Recruitment and response rates

To be eligible to take part in the survey, participants were required to reside in the GMA and drive regularly to at least one specific destination in this area. These criteria were clearly outlined in several occasions, including the informed consent form, which was mandatory to read and accept before starting the survey. To disseminate the survey, three target groups were identified: i) graduate students, postdocs, faculty members, and staff of Polytechnique Montreal, ii) users of social media, such as Facebook, LinkedIn, etc., and iii) volunteer participants who previously agreed to participate in surveys conducted by the Mobility Chair of Polytechnique Montreal and provided their e-mail addresses. A total of 4000 volunteers were contacted on different occasions, and a recall e-mail has been sent to those who have received the first invitation letter, few days later. Out of the 95 % of recipients who received the first invitation letter, 45.1 % opened the e-mail and 12.7 % clicked on the survey link. However, for the recall e-mail, these statistics were 98.6 %, 45.3 %, and 10.1 %, respectively. Invitation e-mails were sent between 9:00 AM and 11:00 AM, and consequently higher response rates between these hours were observed. The completion rate decreases throughout the afternoon, increases slightly around 9:00 PM to 10:00 PM, and reaches its minimum overnight.

4.2. Participants Characteristics

Table 4, illustrates sociodemographic and socioeconomic characteristics of the 539 respondents who have completed the survey. It should be noted that the sample includes mainly young and middle aged full time workers with a university level of education. This may partly be because the survey was also disseminated among scholars, faculty members, and staff of Polytechnique Montreal. Moreover, the prevalence of young participants explains to some degree the higher frequency of smaller households.

Table 4. Characteristics of the Survey Participants.

Variable	Categories	N	%	Variable	Categories	N	%	Variable	Categories	N	%
Gender				*Household size*				*Education*			
	Male	306	56.7		1	100	18.5		None	0	0.0
	Female	234	43.3		2	218	40.4		Less than university	62	11.5
Age (years old)					3	101	18.7		University	472	87.4
	Young (15 to 39)	306	56.7		4	88	16.3		Other	6	1.1
	Middle age (40 to 59)	195	36.1		+5	33	6.1				
	Old (more than 60)	39	7.2								
Occupation				*Income (Thousand CAD per capita)*				*Household car number*			
	Full time worker	393	72.8		< 30	152	28.2		0	122	22.6
	Partial time worker	38	7.0		> 30 and < 60	202	37.4		1	290	53.7
	Student	69	12.8		> 60 and < 90	75	13.9		2	106	19.6
	Retired	25	4.6		> 90 and < 120	12	2.2		+3	22	4.1
	At home	6	1.1		Not declared	99	18.3				
	Other	9	1.7								

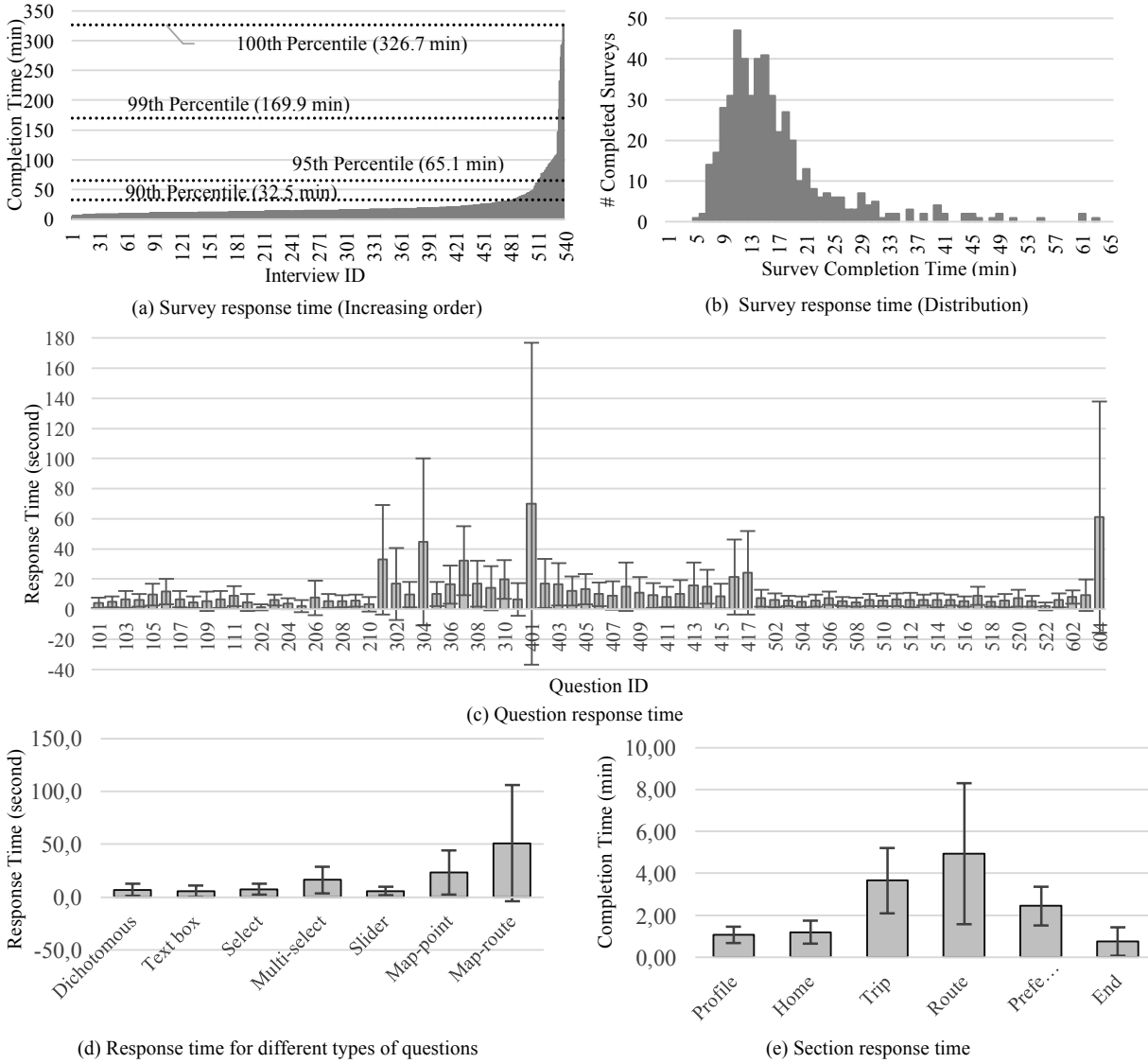
4.3. Response Behavior

Usage information showed that 61% of participants used Windows devices to complete the survey, while Macs (20 %), IOS (12%), Android (5%), Linux (1%), and Chrome OS (1%) accounted for the remaining 39%. Moreover, information on the variety of web browser illustrates that Chrome (59%), Firefox (18%), and Safari (15%) account for around 92% of the completed surveys, while the remaining 8% have been completed on Internet Explorer, Microsoft

Edge, and Opera. These statistics emphasize the importance of making the interface friendly and easy to use for a wide range of devices and browsers to increase the response rate of a survey.

Fig. 1(a), illustrates survey completion times (in minutes) in an increasing order. The completion time of the survey is expressed as the summation of the completion time of all the sections for each interview. It can also be thought of as the difference between the starting time and ending time of the survey excluding the time that respondents had left the survey platform. Considering a 95th percentile threshold, the average and maximum completion time of the survey are found to be 16.1 and 65.1 minutes, respectively and its distribution is illustrated in Fig. 1(b).

Considering a 95th percentile threshold, the average response time for different questions, question types, and sections are illustrated in Fig. 1(c), 1(d), and 1(e), respectively. It can be noted that question ID 401 has the highest response time and variation. Given the complexity of the question, which involves a geographical map and requires respondents to drag and adjust a suggested route, the high response time of this question is not surprising. It is noticed that questions involving a geographical map (i.e. Map-route and Map-point types of questions) have longer response times, while Dichotomous, Text-box, Select, and Slider questions require shorter response times (Fig. 1(d)). Consequently, sections including more geographical maps (Trips, and Routes) have longer response times (Fig. 1(e)).



(a) Survey response time (Increasing order)

(b) Survey response time (Distribution)

(c) Question response time

(d) Response time for different types of questions

(e) Section response time

Fig. 1. Response times (Vertical bars represent the standard deviation).

4.4. Dropouts

The number of dropouts per section and question is illustrated in Fig. 2. Most of the dropouts occur in the Trips section (46.1 %). This may mostly be because some respondents started the questionnaire without satisfying the required participation criteria, i.e. residing and driving in the GMA. In this section, respondents are asked specific questions regarding a destination to which they drive frequently (such as questions 301, 302, and 306). It may be at this point of the survey that they realize that they are not fit to continue the survey. We received several e-mails, Facebook messages, and survey comments supporting the claim that some respondents failed to pay sufficient attention to the participation criteria. As mentioned before, these criteria were repeatedly mentioned in the invitation letter, survey starting page, as well as the informed consent form. The same argument stands for the higher rate of dropouts in question 106, in which respondents are required to declare the age at which they got their driving license.

The second section with the highest dropout rate is the Routes section, in which the first question (i.e. specifying the considered route) has the highest dropout rate of the section. Considering that detailed instructions were provided on how to specify routes on a geographical map, both in the introductory page of the Routes section as well as on top of question 401, the high rate of dropouts may be related to the innate complexity of working with geographical maps and the longer response time required for this question.

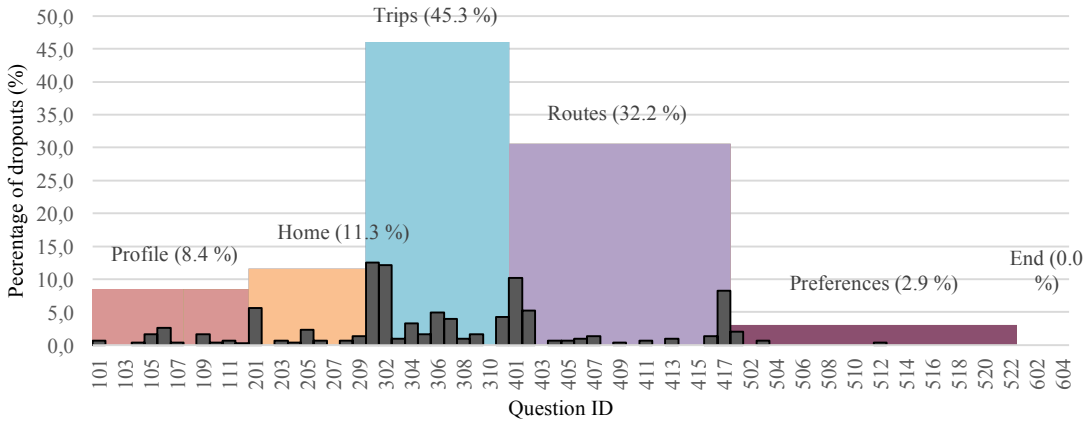


Fig. 2. Percentage of dropouts from the survey per section and question.

Interestingly, we notice that the first question of every section (except for the first and last sections) has the highest rate of dropouts in that section. This indicates that reducing the number of new sections may decrease the total number of dropouts in the survey. To compare the dropout rates of different question types, the total number of dropouts for each question type is divided by the number of recurrence of that particular question type in the whole survey (see Table 5). For instance, 20 dropouts occurred within the 8 recurrences of Dichotomous questions, resulting a dropout rate of 2.5. Results demonstrate that map questions induce higher dropouts compared to other types of questions, with Map-route questions having a higher dropout rate. This is associated to the intrinsic complication of working with geographical maps. Moreover, the effect of Multi-select questions is found to be more pronounced than Select questions. It can also be noted that Slider and Dichotomous questions have the least effects on the number of dropouts.

Table 5. Dropout rates for different types of questions

Question Type	Dropouts	Recurrence	Dropout rate
Dichotomous	20	8	2.5
Text box	47	13	3.6
Select	30	10	3.0
Multi-select	23	5	4.6
Slider	68	33	2.1
Map-point	79	4	19.8
Map-route	31	1	31.0

4.5. Survey Comments

Respondents were asked to provide their general comments, if any, in the last question of the survey (question ID 604). It should be noted that only respondents who advanced to the sixth section of the survey could leave a comment and those who abandoned the survey before that section could not access the comment box to leave their comments. A total number of 149 respondents left comments, most of which were generally positive and encouraging, commending the objectives, the question design and the appearance of the survey. Few respondents, however, criticized the length and complexity of the questionnaire and the hardships of specifying a route trajectory on geographical map. They also reported some technical issues regarding some browsers and occasional difficulties with phone interfaces. We have also received few messages and e-mails from respondents concerned about privacy aspects, most of whom left the survey in the second section. Interestingly, a significant number of respondents who left comments were not happy about the exclusive focus of the survey on drivers route choices, and asked for a more comprehensive travel survey, considering other modes of transportation such as public transit, walk, and bike.

5. Conclusion

The increasing application of advanced choice models, reflecting the stochasticity of individuals preferences and the complex nature of human decision-making behavior, requires enhanced data collection methods collecting detailed data without significantly increasing respondent burden. This paper details the development and deployment of a general survey framework for route choice studies with three main objectives: i) to observe drivers revealed route choices, ii) to identify important factors including behavioral and attitudinal factors affecting them, and iii) to observed and characterize drivers consideration sets of route alternatives.

A web-based survey has been designed to provide researchers with a rich dataset, based on which they can produce reliable behavioral models. A graphical interface is adopted to augment response precision and to reduce the burden of declaring all the considered alternatives. Moreover, the analyst obtains the exact trajectories considered for each trip and will not face the challenges and uncertainties associated with GPS datasets such as trip extraction, map-matching, and path inference. In short, the analyst will be able to investigate more closely some major challenges facing route choice modeling, such as the definition of an alternative route and how it is perceived by drivers, the characteristics of a considered set of route alternatives, and the role of different attributes (observable and latent) in route choice decisions.

Considering the high number of questions included in the survey, the variety of question types, and the spectrum of the collected data, the overall survey completion percentage of 64 % suggest a successful implementation of the survey framework. An internal validation system has been applied to minimize participants errors and maximize the completeness of survey responses. As a result, a small number of interviews were discarded (26 out of 539, 4.8 %), which indicate the high quality of the collected data. The principal reasons for exclusion were twofold: first, living or driving outside the study area, and second, failing to specify a logically sound route between the predefined OD pairs. To increase the quality of the final dataset, unusually short or long response times can be used as proxy indicators to identify measurement errors (Couper & Kreuter, 2013).

Despite the successful application of the proposed survey framework, the authors acknowledge its limitations and the uncertainties associated with certain responses. Although web-based interfaces offer more flexibility for designing the questionnaire and can be used as a tool to improve paper-based surveys, it comes with particular limitations. A thorough comparison of web-based, paper-based, and face-to-face route choice survey allows us to explore in detail the quality of the collected data. According to Bayart and Bonnel (2012), in a comparative analysis of different travel survey interfaces, we have to distinguish between three effects: 1) the socio-demographic and socioeconomic difference of the respondents; 2) the difference in travel behaviors of the respondents; and 3) the differences that are merely due to the survey medium and do not necessarily reflect the difference in travel behavior. Although it is practically very difficult to isolate each of these effects (Bayart & Bonnel, 2012), a comparative analysis of the collected data with the data collected using other survey interfaces can shed some light on some of these aspects. For instance, the comparison of the reported number of route alternatives allows us to make some hypothesis about the effect of survey interface on under-reporting route alternatives. Therefore, similar to HTS, in which trip under-reporting is an undeniable issue (Bayart & Bonnel, 2012; Stopher, FitzGerald, & Xu, 2007), the possibility of under-reporting the number of considered route alternatives for the declared trip is recognized in this study.

Four major sources of errors are recognized in web-based surveys, namely the coverage error, the sampling error, the measurement error, and the non-response error (Bayart & Bonnel, 2012; Dillman & Bowker, 2001). It is not straightforward to ensure a representative sample of the population using web-based surveys. For instance, households without a computer or without access to the internet cannot participate in the survey. Recruitment methods in web-based surveys generate bias in the sample population that should be taken into account while interpreting the results. Also, there is a substantial difference in response rates of different socio-demographic groups to web-based surveys which might also affect the representativeness of the results (Christensen, 2013). It has been shown that younger people are over-represented in the web-based survey while elderly are under-represented (Bech & Kristensen, 2009; Bourbonnais & Morency, 2013). Moreover, some individuals may lack the skills to use the internet or the technical knowledge to answer certain types of questions (Bourbonnais & Morency, 2013; Christensen, 2013). For instance, the complexity of questions involving geographical maps might also affect their response quality. It has been clearly stated in the literature that some people are better map navigators than others (Amy, 2007). In this study, the high rate of dropouts associated to these questions underscores their intrinsic complexity. In order to maintain the interest of the respondents and keep them in the survey, appropriate visual aids (such as video and charts) might be helpful. Moreover, sampling strategies such as random postings on social media or web-pages relating to the subject of interest, or emails to employees in relevant organizations will produce samples of people with special interests that are not socio-economically representative of the population.

Further research could include a comparison of the proposed framework with other route choice data collection frameworks, with different types of questions and various lengths. This comparison would lead to a better evaluation of the performance and data quality of different frameworks. Moreover, completion time can be treated as an indicator of respondent burden, and the effect of different completion times can be studied on dropout rates. Another possible extension of this effort can be the integration of the proposed survey framework with smartphones and GPS devices.

Acknowledgement

The authors would like to express their gratitude to the partners of the Mobilité Research Chair of Polytechnique Montréal for their financial support and their help in the dissemination of the survey.

References

Abdel-Aty, M., & Jovanis, P. (1997). *Using GIS Capabilities to Improve The Understanding Of Route Choice Behavior*. Paper presented at the Transportation Research Board, Washington D.C.

Alizadeh, H., Farooq, B., Morency, C., & Saunier, N. (2017). On the role of bridges as anchor points in route choice modeling. *Transportation*. doi:10.1007/s11116-017-9761-7

Alizadeh, H., Farooq, B., Morency, C., & Saunier, N. (2018). *Frequent Vs Occasional Drivers: A Hybrid Route Choice Model*. Retrieved from

Amy, K. L. (2007). Navigational Map Reading: Predicting Performance and Identifying Relative Influence of Map-Related Abilities. *Annals of the Association of American Geographers, 97*(1), 64-85.

Atasoy, B., Glerum, A., & Bierlaire, M. (2013). Attitudes towards mode choice in Switzerland. *disP - The Planning Review, 49*(2), 101-117. doi:10.1080/02513625.2013.827518

Bayart, C., & Bonnel, P. (2012). Combining web and face-to-face in travel surveys: comparability challenges? *Transportation, 39*(6), 1147-1171.

Bech, M., & Kristensen, M. B. (2009). *Differential response rates in postal and Web-based surveys in older respondents.* Paper presented at the Survey Research Methods.

Ben-Elia, E., & Shiftan, Y. (2010). Which road do I take? A learning-based model of route-choice behavior with real-time information. *Transportation Research Part A: Policy and Practice, 44*(4), 249-264. doi:http://dx.doi.org/10.1016/j.tra.2010.01.007

Bliemer, M., & Bovy, P. (2008). Impact of route choice set on route choice probabilities. *Transportation Research Record: Journal of the Transportation Research Board*(2076), 10-19.

Bourbonnais, P.-L., & Morency, C. (2013). Web-Based Travel Survey: A Demo. In M. L.-G. Johanna Zmud, Marcela Munizaga, Juan Antonio Carrasco (Ed.), *Transport Survey Methods: Best Practice for Decision Making* (pp. 207-224).

Cascetta, E., Russo, F., Viola, F. A., & Vitetta, A. (2002). A model of route perception in urban road networks. *Transportation Research Part B: Methodological, 36*(7), 577-592. doi:http://dx.doi.org/10.1016/S0191-2615(00)00041-2

Chen, J. (2013). *Modeling route choice behavior using smartphone data.* (PhD), Polytechnique de Lausanne, Suisse.

Christensen, L. (2013). *The role of web interviews as part of a national travel survey*: Emerald Group Publishing Limited, Bingley, UK.

Cools, M., Moons, E., & Wets, G. (2009). Assessing the Impact of Public Holidays on Travel Time Expenditure Differentiation by Trip Motive.

Couper, M. P., & Kreuter, F. (2013). Using paradata to explore item level response times in surveys. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 176*(1), 271-286.

Dalumpines, R., & Scott, D. M. (2017). Determinants of route choice behavior: A comparison of shop versus work trips using the Potential Path Area - Gateway (PPAG) algorithm and Path-Size Logit. *Journal of Transport Geography, 59*, 59-68. doi:http://dx.doi.org/10.1016/j.jtrangeo.2017.01.003

Dhakar, N. S. (2012). *Route Choice Modeling Using GPS Data.* University of Florida.
Dillman, D. A., & Bowker, D. K. (2001). The web questionnaire challenge to survey methodologists. *Online social sciences*, 53-71.
Gan, H., & Chen, S. (2013). Why Do Drivers Change Routes? Impact of Graphical Route Information Panels. *ITE Journal*.
Geda, A. (2014). *Effect of Choice Set Composition on Route Choice Models.* (M.Sc.), University of Florida.
Habib, K. N., Morency, C., Trépanier, M., & Salem, S. (2013). Application of an independent availability logit model (IAL) for route choice modelling: Considering bridge choice as a key determinant of selected routes for commuting in Montreal. *Journal of Choice Modelling, 9*, 14-26. doi:http://dx.doi.org/10.1016/j.jocm.2013.12.002
Hess, S., Quddus, M., Rieser-Schüssler, N., & Daly, A. (2015). Developing advanced route choice models for heavy goods vehicles using GPS data. *Transportation Research Part E: Logistics and Transportation Review, 77*, 29-44.
Hoogendoorn-Lanser, S., van Ness, R., & Bovy, P. H. (2005). *Path Size and Overlap in Multi-modal Transport Networks.* Paper presented at the 16th International Symposium on Transportation and Traffic Theory.
Iida, Y., Uno, N., & Yamada, T. (1994, 31 Aug-2 Sep 1994). *Experimental analysis approach to analyze dynamic route choice behavior of driver with travel time information.* Paper presented at the Vehicle Navigation and Information Systems Conference, 1994. Proceedings., 1994.
Jan, O., Horowitz, A., & Peng, Z.-R. (2000). Using global positioning system data to understand variations in path choice. *Transportation Research Record: Journal of the Transportation Research Board*(1725), 37-44.
Jou, R.-C., & Yeh, Y.-C. (2013). Freeway passenger car drivers' travel choice behaviour in a distance-based toll system. *Transport Policy, 27*, 11-19.
Kaplan, S., & Prato, C. G. (2012). Closing the gap between behavior and models in route choice: The role of spatiotemporal constraints and latent traits in choice set formation. *Transportation Research Part F: traffic psychology and behaviour, 15*(1), 9-24.
Koller-Matschke, I., Belzner, H., & Glas, R. (2013). Traffic Information and Individual Driver Behaviour *Proceedings of the FISITA 2012 World Automotive Congress* (Vol. 200, pp. 179-185): Springer Berlin Heidelberg.
Lai, X., & Bierlaire, M. (2015). Specification of the cross-nested logit model with sampling of alternatives for route choice models. *Transportation Research Part B: Methodological, 80*, 220-234. doi:http://dx.doi.org/10.1016/j.trb.2015.07.005
Manley, E. J., Addison, J. D., & Cheng, T. (2015). Shortest path or anchor-based route choice: a large-scale empirical analysis of minicab routing in London. *Journal of Transport Geography, 43*, 123-139. doi:http://dx.doi.org/10.1016/j.jtrangeo.2015.01.006
Manski, C. F. (1977). The structure of random utility models. *Theory and decision, 8*(3), 229-254.
Nadine Schuessler, I., & Axhausen, K. W. (2009). *Accounting for Route Overlap in Urban and Sub-urban Route Choice Decisions Derived from GPS Observations*. Paper presented at the 12th International conference on travel behavior research, Jaipur, India.
Ory, D. T., & Mokhtarian, P. L. (2005). Don't work, work at home, or commute? Discrete choice models of the decision for San Francisco Bay Area residents.
Papinski, D. (2011). *Investigating Route Choice Decisions Using GPS and Prompted-recall Diary Data.* (PhD), McMaster University.
Papinski, D., Scott, D. M., & Doherty, S. T. (2009). Exploring the route choice decision-making process: A comparison of planned and observed routes obtained using person-based GPS. *Transportation Research Part F: traffic psychology and behaviour, 12*(4), 347-358.
Parkany, E., Du, J., Aultman-Hall, L., & Gallagher, R. (2006). Modeling Stated and Revealed Route Choice: Consideration of Consistency, Diversion, and Attitudinal Variables. *Transportation Research Record: Journal of the Transportation Research Board, 1985*(-1), 29-39.
Peeta, S., Ramos, J., & Pasupathy, R. (2000). Content of Variable Message Signs and On-Line Driver Behavior. *Transportation Research Record: Journal of the Transportation Research Board, 1725*(-1), 102-108. doi:10.3141/1725-14
Peters, T., Adamowicz, W. L., & Boxall, P. C. (1995). Influence of choice set considerations in modeling the benefits from improved water quality. *Water Resources Research, 31*(7), 1781-1787.
Prato, C., & Bekhor, S. (2006). Applying branch-and-bound technique to route choice set generation. *Transportation Research Record: Journal of the Transportation Research Board*(1985), 19-28.
Prato, C., & Bekhor, S. (2007). Modeling route choice behavior: how relevant is the composition of choice set? *Transportation Research Record: Journal of the Transportation Research Board*(2003), 64-73.
Prato, C. G., Bekhor, S., & Pronello, C. (2012). Latent variables and route choice behavior. *Transportation, 39*(2), 299-319.
Ramaekers, K., Reumers, S., Wets, G., & Cools, M. (2013). Modelling Route Choice Decisions of Car Travellers Using Combined GPS and Diary Data. *Networks and Spatial Economics, 13*(3), 351-372. doi:10.1007/s11067-013-9184-8
Sarkar, P. P., & Mallikarjuna, C. (2017). Effect of perception and attitudinal variables on mode choice behavior: A case study of Indian city, Agartala. *Travel Behaviour and Society*. doi:https://doi.org/10.1016/j.tbs.2017.04.003
Schlaich, J. (2010). Analyzing route choice behavior with mobile phone trajectories. *Transportation Research Record: Journal of the Transportation Research Board*(2157), 78-85.
Srinivasan, S., & Dhakar, N. S. (2013). *Route-choice Modeling using GPS-based Travel Surveys*. Retrieved from
Stopher, P., FitzGerald, C., & Xu, M. (2007). Assessing the accuracy of the Sydney Household Travel Survey with GPS. *Transportation, 34*(6), 723-741. doi:10.1007/s11116-007-9126-8
Swait, J., & Ben-Akiva, M. (1987). Empirical test of a constrained choice discrete model: mode choice in Sao Paulo, Brazil. *Transportation Research Part B: Methodological, 21*(2), 103-115.
Tawfik, A. M., & Rakha, H. (2013). *A Latent Class Choice Model of Heterogeneous Drivers Route Choice Behavior Based on a Real-World Experiment.* Paper presented at the 20th International Symposium on Transportation and Traffic Theory, Noordwijk, Netherlands.
Tawfik, A. M., Rakha, H. A., & Miller, S. D. (2010, 19-22 Sept. 2010). *An experimental exploration of route choice: Identifying drivers choices and choice patterns, and capturing network evolution.* Paper presented at the 13th International IEEE Conference.
Transport, A. M. d. (2013). *Enquête Origine-Destination 2013: La mobilité des personnes dans la région de Montréal-FAITS SAILLANTS*. Montreal Retrieved from https://rtm.quebec/Media/Default/pdf/section8/enquete-od-2013-faits-saillants.pdf.
Vacca, A., & Meloni, I. (2014). *Understanding route choice behavior: an analysis using GPS based data.* Paper presented at the Transport Research Arena (TRA) 5th Conference: Transport Solutions from Research to Deployment.
Walker, J. L. (2001). *Extended discrete choice models: integrated framework, flexible error structures, and latent variables*. Massachusetts Institute of Technology.