CIRRELT

Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation

Comparison Between Trip and Trip Chain Models: Evidence from Montreal Commuter Train Corridor

Gabriel Sicotte
Catherine Morency
Bilal Farooq

June 2017

CIRRELT-2017-35

Bureaux de Montréal :
Université de Montréal
Pavillon André-Aisenstadt
C.P. 6128, succursale Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :
Université Laval
Pavillon Palasis-Prince
2325, de la Terrasse, bureau 2642
Québec (Québec)
Canada G1V 0A6
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

Comparison Between Trip and Trip Chain Models: Evidence from Montreal Commuter Train Corridor

Gabriel Sicotte[1,*], Catherine Morency[1,2], Bilal Farooq[1,2]

[1] Department of Civil, Geological and Mining Engineering, Polytechnique Montréal, P.O. Box 6079, Station Centre-ville, Montréal, Canada H3C 3A7

[2] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT)

Abstract. Typical mode choice models consider trips independently without the complete structure of the trip chain. This paper aims to further demonstrate the importance of considering the spatiotemporal structure of trip chains by comparing two multinomial discrete choice models. First, transportation modes are classified in two categories, "With anchor" modes and "No anchor" modes, to allow the enumeration of the possible mode sequences for each trip chain that becomes the set of alternatives for a traveler for a given trip chain. Two mode choice models are then estimated for chains of two trips: one processing trips independently and one based on entire trip chains. The case study is a commuter train corridor of the Greater Montreal Area. With a success rate increasing from 12.74% (trip based) to 72.15% (trip chain based), the results clearly confirm that modeling the trip chain at the mode choice step is much more coherent. It considers the alternatives for the return trip while examining the mode choice for the morning commute.

Keywords: Modal choice, chain based model, trip chain, tour.

Acknowledgements. The authors wish to acknowledge the contribution and support of the Chaire Mobilité partners (Ville de Montréal, Ministère des transports du Québec, Agence métropolitaine de transport and Société de transport de Montréal). They also wish to thank Jean-Simon Bourdeau and Pierre-Léo Bourbonnais for their contributions.

* Corresponding author: gabriel.sicotte1@gmail.com

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2017

1. Introduction

Traditional mode choice models don't take into account the spatiotemporal structure of trip chains and instead process trips independently based on the assumptions of the Four-Step models (McNally 2007). The work of Asensio and Javier (2002) is a good example of trip based model used to analyze the determinants of mode choice. Miller et al. (2005) and Dong et al. (2006) note that this can significantly affect the ability of the model to correctly reproduce observed behaviors, especially in situations where mode choice is highly influenced by the preceding or following trip of a chain. For instance, an individual that goes to work using his own car in the morning will certainly be using the same mode for his trip back home if he wants to use it again the next day. This paper aims to compare the results of these two models. The objective is to assess the hypothesis that taking into account the trip chain structure in mode choice modeling helps to better reproduce observed behaviors.

Valiquette (2010) proposed a trip chain typology as well as a method to classify observed trip chains from large-scale Origin-Destination (OD) surveys. Our research builds upon this previous work and proposes a conceptual framework to first exhaustively enumerate possible mode sequences for various types of trip chains. We then use this typology to develop a mode choice model. For this purpose, the research also proposes a travel mode classification, centered on the anchor point concept, allowing to create possible mode sequences for trip chains.

This paper proposes a comparison between two mode choice models, one based on trips and the other on trip chains. Two models are developed for a suburban commuter train corridor located in the Greater Montreal Area (GMA). Since its service is mainly directional and concentrated during the peak periods, it is a perfect and unique case study to assess the contribution of modeling the entire chain instead of a single trip.

The paper is organized as follows. First, the background elements are proposed that are required for the conceptual framework. Then, the general methodology is presented which includes the main concepts of the conceptual framework as well as the underlying hypothesis. The case study is then presented along with the choice set preparation and the models formulation. Modeling results are discussed. Paper concludes with a discussion on the research work, limitations and research perspectives.

2. Background

2.1 Mode choice modeling

Random utility based discrete choice models are widely used to model mode choice behaviors. They allow estimating the probability function that can predict if an individual will choose a specific mode for a particular trip. Discrete choice models rely on the formulation of utility functions adapted to each transportation mode and individual. They furthermore rely on the hypothesis that the traveler chooses the mode that maximizes its utility (Train 2009; Ortúzar and Willumsen 2011). Depending on its formulation, the utility function will include attributes describing the individual, the household, the neighborhood and the available alternatives (Martel-Poliquin 2012).

The relationship between trip chains and mode choice has been studied extensively in recent literature. Currie and Delbosc (2011) studied the behavior of public transport users in Melbourne and found that the complexity of chain is larger for train and tram trips than for car trips. Hensher and Reyes (2000) produced models of mode choice (car and public transit) by complexity of the chain (simple and complex) and found that, as the complexity of the trip chain increases, the gain in utility from using the public transit decreases.

Ye et al. (2007) found that the chain complexity drives mode choice, both for work and non-work chains. This conclusion was refined by Islam and Habib (2012): for weekdays work related trips, trip chaining and mode choice are simultaneous, while for non-work trip chains during weekdays, mode choice is made before trip chaining. For weekend non-work chains, trip chaining decisions drive mode choice. The travel behavior of household member seems to be different during weekends compared to weekdays, with the intra-household ride sharing interactions being higher during weekdays, due to purpose and mobility constraints (Ho and Mulley 2013).

Xianyu (2013) explores the interdependency between trip chain complexity and mode choice. He found that commuters using cycling, transit or car to work are less likely to make non-work stops on a work tour than those who are walking. In a previous study, Strathman and Dueker (1995) found that when transit is combined with car the rate of complex work tours increases compared to simple work tours. The rate of complex work tours increases even more when transit is combined with non-auto modes, like walking. However, Harding et al. (2015) argue that it’s not the complexity of the trip chain that determines the mode choice, but distance between destinations.

Walle and Steenberghen (2006) studied trip chains in the city of Brussels and found that a group of trip chains were made using public transit and car at the end of the chain due to missing links of public transport.

Extensive literature can be found on estimation and simulation of mode choice models, but very few studies have taken into account the complete spatiotemporal structure of the trip chain. Yun et al. (2011) present a relatively simple framework based on the concept of “main mode” and propose a comparison between a model based on chains and a model based on trips to confirm the advantage of the first one. The study by Miller et al. (2005) proposes a more complex framework that integrates the Travel/Activity Scheduler for Household Agents (TASHA) model developed for the city of Toronto. Using the same framework, Roorda et al. (2009) propose a model that includes minor modes of transportation such as taxi, bicycle, commuter rail and subway.

Our research innovates by proposing a classification of transportation modes that supports the construction of a more exhaustive list of possible mode sequences given a particular type of chain.

2.2 Trip chains

The trips composing a chain have different activity locations and can also involve different transportation modes. Switching from one mode to another during a trip does not correspond to a chain as trips can be multimodal (Primerano et al. 2008; Steenberghen, Toint, and Zuallaert 2005). Defining the concept of chain is not as trivial as it seems and multiple definitions have been proposed. Still, there is a general agreement on the concept of having a chain defined using the home location as starting and end points. This is the definition we have adopted in our work. If another location is visited more than once during the chain, then it creates a loop and this point is defined as an anchor. Several reasons explain why this definition is more appropriate. First, according to the Metropolitan Adelaide Household Survey 1999, most first trips of the day begin at the home location and most of last trips return home. Also, leaving the home location for a set of trips implies that multiple decisions are taken a priori and that these choices influence the sequence of trips and activities occurring before returning home. The set of choices to be made after leaving home is smaller. For instance, the first mode selected when leaving home impacts the possible modes for the other trips of the chain. Hence, this definition puts more emphasis on the activities than on the trips, which is more in phase with up-to-date activity models (Primerano et al. 2008).

Valiquette (2010) proposed the categorization of trip chains according to their main features: the number of trips (activity) and anchor points (number of loops). Using this definition, trips from the OD surveys of the GMA are categorized in four trip chain categories: simple, complex single-loop, complex multi-loop and open. The simple chain has only one anchor point (home location) and one destination (activity location). Thus, this chain consists only of two trips. The complex single-loop chain includes more than one destination, but still a single anchor point (home). The complex multi-loop chain has also more than one destination, but at least one of these destinations is an anchor point (a sub-loop of the chain starts and ends at this destination). It should be noted that the home location is necessarily an anchor point. The open chain is an incomplete one for which there is either no return home or leave home trip.

3. Theoretical framework

3.1 Assumptions

The most important assumption of this work is that the space-time structure of the chain is constant. Hence, in our modeling framework, we assume that the simulated chain has the same number of trips, same departure times and same destinations as the observed chain. The proposed framework can be expanded to relax this hypothesis for further research by integrating, for instance, a chain choice model to a nested logit mode choice model.

3.2 Mode classification

In order to integrate trip chains within mode choice models, it is important to generate the alternatives i.e. the possible mode sequences for each type of chains. Beforehand, modes have to be classified depending on their linkage to an anchor point or not. It is important to take into account the fact that some modes have a "retrieve at anchor point” constraint so the traveler cannot switch mode without making sure that he will come back to the location where this mode was left to retrieve it. Thus from this characteristic, two main categories of transportation modes are defined: the "with anchor" (WA) modes and the "no anchor" (NA) modes.

The WA modes are those that absolutely must be retrieved at an anchor point and brought back to the end of the loop. Otherwise they will not be available for use in the next chain or loop. This is logical in a single day modeling framework. It is to be noted that in the context of a trip chain, an anchor point is defined as the place where a loop starts and ends while in the context of a mode, the anchor point relates to the place where the mode (car for instance) is parked when not in used. This feature acts as a constraint on car (driver), station-based car-sharing and cycling. However, cycling is a special case as it has characteristics of both categories since it can either be left at an anchor point or brought inside other modes (transit, car). For practical purposes (since cycles are typically not allowed in transit during peak periods) it is considered a WA mode in this paper.

The NA modes are independent. If a traveler goes to work by transit, he can select within a wider range of modes to return home. For the GMA, the category of NA modes include walking, bike-sharing, free-floating car-sharing, car passenger, taxi and transit.

A similar classification is performed for bimodal sequences, only here we integrate the concept of intermediate anchor point (IAP), which consists in a place where a loop starts and ends, without being an activity location. A park and ride lot at a commuter train station is a good example. The category of bimodal sequences is determined by the presence of a WA mode within the sequence. Thus, WA sequence includes Park & Ride and Bike & Ride (coherent with our previous hypothesis), while NA sequence includes Kiss & Ride.

This categorization of transportation modes and bimodal sequences is particularly important for our modeling process as the probability that a certain mode is selected for a particular trip must take into account the mode category of the previous trips of the chain. Indeed, if the mode used for the first trip is WA, it is unlikely that the following trips will be done using another mode, unless it's a multi-loop chain with more than one anchor point.

3.3 Data extraction

Every five years since 1970, a large-scale OD survey is conducted in the GMA. The survey provides detailed information regarding trips of people in the households of GMA, aged five years or older, in the form of a one-day trip diary for a single weekday of the fall period. The observations are collected from about 5% of the households of the GMA, surveyed via land-line phone, in which one respondent describes the trips of each person in the household. These observations are then weighted at a smaller scale (i.e. one of the 108 municipal areas of the GMA) to match the census population that has the same characteristics (household size, gender and age). More details from the latest OD survey can be found in (AMT 2010). A trip chain database is generated using the method developed in Valiquette (2010) by aggregating all the trips made by an individual, between the moment he left home and came back to it, into a trip chain. The different chains generated are then enriched with additional information: unique identifier, chain type, chain sequence (within a daily individual activity program), main purpose, number of activities, number of loops, open-chain (true or false), number of unique activity locations, mode sequence, purpose sequence and trip sequence. Data from the 2008 OD survey are used in this study.

3.4 Set of mode sequence alternatives

Once the modes are categorized, it is possible to develop a chain typology to enumerate the different alternatives that can be integrated within the choice model, while taking into account the constraints related to the "with anchor" mode. This exercise is done for each of the chain types observed in the OD. Four types of chain match 97.85% of all 3,599,132 chains derived from trips in the 2008 OD survey. Out of these chains 86.24% are simple chains. To constitute the set of mode choices, the alternatives enumerated must correspond to the observations.

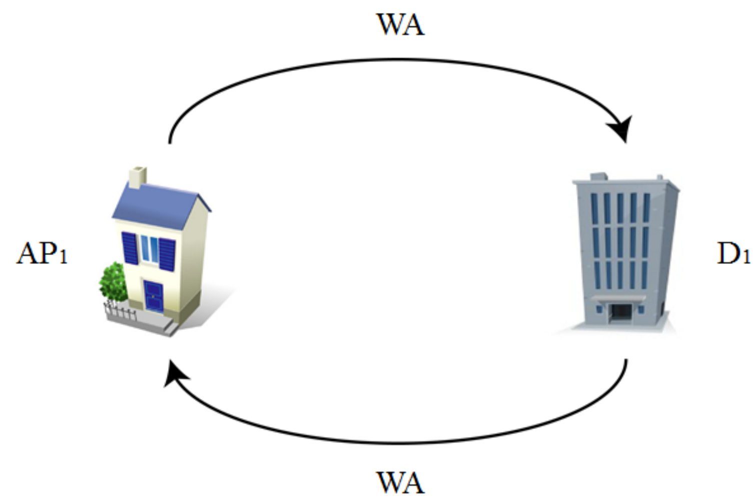
The first step is to determine possible sequences using the two previous classes. Each trip (or segment in the case of bimodal sequences) is given a mode category. It is important to specify the order in which the secondary activities are carried with respect to the anchors of the chain (before or after) as it has a significant impact on the possible alternatives.

For a simple chain, there are only three possible WA mode sequences, excluding bimodal options. With the obligation to return to the starting point of the loop with the same mode, we find the same number of sequences as there are modes in this category. As illustrated in Figure 1a, the most common example is: Car (driver) - Car (driver).

To properly represent the bimodal Park & Ride and Bike & Ride sequences within the available alternatives, an IAP is created in the chain. Thus, for a simple chain as illustrated in Figure 1b where a bimodal sequence WA is used for both trips, there are only two alternatives of different possible bimodal sequences.

Form of the chain: Anchor point 1 (AP_1) – Destination 1 (D_1) – AP_1
Mode sequence: With Anchor (WA) – WA
Number of possible alternatives: 3

a)



Form of the chain: AP_1 – Intermediate anchor point 1 (IAP_1) – D_1 – IAP_1 – AP_1
Mode sequence: (WA-No Anchor (NA)) – (NA-WA)
Number of possible alternatives: 2

b)

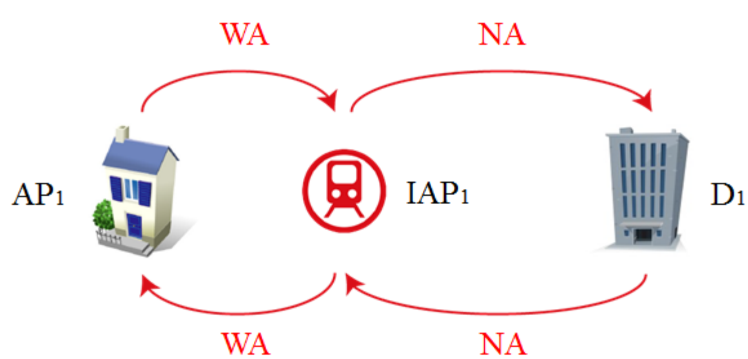


Figure 1 a) Structure of a simple chain for a sequence of "with anchor" modes, b) structure of a simple chain for a bimodal sequence "with anchor"

4. Case study

Data from the OD survey offers the possibility of a case study at the scale of the GMA, but due to the high time investment required in the preparation of the dataset, this case study presents a simplified application of the approach on a suburban commuter train corridor. More importantly, since commuter train service is mainly directional and concentrated during the peak periods, it is a perfect case study to test a model based on single chains, made of two trips.

4.1 Data preparation and extraction

The first step is the extraction, from the 2008 OD file, of the trips, which include at least one segment on the Vaudreuil-Hudson commuter train line. Municipal area of origin and destination of these trips are compiled in pairs to create an OD matrix that represents the attraction zone of the corridor.

The second step is to go back to the 2008 OD survey to extract all the trips available between the identified pairs of zones. For these trips, train is considered to be a possible alternative.

The third step relates to the extraction of all the trips of the chains to which the previous trips belong. In this study, only chains starting and ending at the home location are considered. At this stage, a chain dataset is created and from that a second dataset is built, consisting of trips of the chain. However, it is individually coded with the aim of being used for the trip based model. To simplify the method, the mode choice model based on the chains only estimates simple chains.

4.2 Developing the choice set

4.2.1 Trip based model

For each traveler, eight different modes are used to generate the alternatives:

1. Car (driver): **CD** (4448 observations)
2. Car (passenger): **CP** (1153 obs.)
3. Transit: **PT** (1680 obs.)
4. Walking: **W** (280 obs.)
5. Cycling: **C** (24 obs.)
6. Train: **TR** (684 obs.)
7. Park & Ride (CD and TR): **PR** (338 obs.)
8. Kiss & Ride (CP and TR): **KR** (181 obs.)

4.2.2 Trip chain based model

A list of observed mode sequences for simple chains is constructed using the sample, along with the number of observations. Sequences that are considered impossible are removed. A minimum threshold of 8 observations is used to determine the possible mode sequences from the sample, which amounts to 15. These mode sequences are listed below.

1. Car (driver) – Car (driver): **CD-CD** (2220 obs.)
2. Public transit – Public transit: **PT-PT** (800 obs.)
3. Car (passenger) – Car (passenger): **CP-CP** (543 obs.)
4. Train – Train: **TR-TR** (317 obs.)
5. Park & Ride (Train) – Park & Ride (Train): **PR(TR)-PR(TR)** (169 obs.)
6. Walking – Walking: **W-W** (136 obs.)
7. Kiss & Ride (Train) – Kiss & Ride (Train): **KR(TR)-KR(TR)** (81 obs.)

8. Cycling – Cycling: **C-C** (62 obs.)
9. Car (passenger) – Public transit: **CP-PT** (32 obs.)
10. Public transit – Car (passenger): **PT-CP** (26 obs.)
11. Kiss & Ride (Train) – Train: **KR(TR)-TR** (19 obs.)
12. Public transit – Train: **PT-TR** (11 obs.)
13. Train – Public transit: **TR-PT** (11 obs.)
14. Train – Car (passenger): **TR-CP** (9 obs.)
15. Car (passenger) – Walking: **CP-W** (8 obs.)

4.2.3 Availability of alternatives

The availability of alternatives is calculated at the trip level. It is only after this step that the trips are linked into chains and that the availability of mode sequences is calculated.

Car (driver): Two conditions are used to ensure the availability of the car (driver) option. The household must have a minimum of one car and the individual must own a driving license. Since the driving license is considered in the availability of the mode, it is not used as a parameter of the utility function for the car (driver).

Car (passenger): The availability of the car (passenger) is simpler. Only the presence of at least one car in the household is required. According to the 2008 OD survey, 76% of the rides are provided by drivers of the same household. Thereby this is an acceptable constraint.

Public transit: A feasible route needs to be available for transit to be among the choices of an individual. An access time threshold is also used at the origin and destination points. Value of the threshold is determined based on the concept of threshold distance used by Godefroy (2011), but applied using time. Thus, a threshold time of 17 minutes is calculated from the observations and represents the access and egress duration of walking for 85% of the current transit users.

Walking and cycling: The possibility to walk or cycle to do the trip is determined also using a threshold-based approach. In this case, it is a distance threshold of 1.4 km for walking and 5.7 km for cycling. Note that the cycling option is considered available for all trips within the specified threshold regardless of the actual availability of such vehicles in the household since this information is not available in the survey. This hypothesis is acceptable as Vélo Québec (2011) estimates that the average ownership is 1.7 bicycle per household in Montreal.

Train: Since in the case of train only sequence, the access and egress to the train stations are made by walking, the concept of distance threshold is also used here. Thus, this option is available only if a Vaudreuil-Hudson line station is within 1.4 km of both the origin and the destination of the trip. Also, the train schedule is taken into consideration when assessing if the trip can be done by train.

Park & Ride, Kiss & Ride: The first criteria an individual has to meet to use the Park & Ride and Kiss & Ride sequences respectively correspond to those of the car (driver) and car (passenger). A threshold distance for the segment made by car is set to 7 km. Here too, the train schedule is considered depending on the time of departure and the direction.

4.3 Data analysis

4.3.1 Studied variables

Given the experimental nature of the chain based model, variables with direct and high impact are chosen for the estimation. The variables included in both models relate to the socio-demographic profile of the person and his household, as well as to the trip or chain features. For the chain model, a dummy variable indicates if the first trip is made with a WA mode or not. Table 1a and Table 1b present the variables used for the two different models.

Table 1a Description of studied variables for the trip based model

Variable type	Name of the variable	Description	*Min*	*Max*	$\bar{X}$	σ
	Choice	Observed chosen alternative				
sociodemographic variables	h_nb_car	Number of cars at home	0.00	14.00	1.73	1.11
	h_nb_adult	Number of adults at home	1.00	8.00	2.40	0.95
	h_nb_kid	Number of kids at home	0.00	5.00	0.74	1.00
	i_age_0_15	Individual aged between 0 and 15 years old (0 or 1)	0.00	1.00	0.09	0.28
	i_age_16_24	Individual aged between 16 and 24 years old (0 or 1)	0.00	1.00	0.15	0.36
	i_age_25_64	Individual aged between 25 and 64 years old (0 or 1)	0.00	1.00	0.69	0.46
	i_age_65_over	Individual aged of 65 years old and over (0 or 1)	0.00	1.00	0.07	0.26
	i_gender_m	Individual of masculine gender (0 or 1)	0.00	1.00	0.51	0.50
	i_gender_f	Individual of feminine gender (0 or 1)	0.00	1.00	0.49	0.50
	i_status_work_study	Individual with worker of student status (0 or 1)	0.00	1.00	0.80	0.40
	i_pt_fare_yes	Individual with a public transit monthly fare (0 or 1)	0.00	1.00	0.28	0.45
	i_tr_fare_yes	Individual with a train monthly fare (0 or 1)	0.00	1.00	0.12	0.32
	i_license_yes	Individual with a driving license (0 or 1)	0.00	1.00	0.81	0.39
Characteristics of the trip	t_motive_work_study	Trip with a motive of study or working (0 or 1)	0.00	1.00	0.36	0.48
	t_motive_other	Trip with other motive than study or working (0 or 1)	0.00	1.00	0.14	0.34
	t_total_cost_1	Total cost of the trip with CD ($)	0.01	8.74	2.01	1.44
	t_total_cost_3	Total cost of the trip with PT ($)	0.00	5.50	0.66	1.25
	t_total_cost_6	Total cost of the trip with TR ($)	0.00	9.75	3.30	1.88
	t_total_cost_7	Total cost of the trip with PR ($)	0.05	16.00	3.96	2.53
	t_total_cost_8	Total cost of the trip with KR ($)	0.00	8.50	3.17	1.83
	t_total_dist_4	Total trip distance with W (metres)	16.00	44018.00	1840.51	4216.89
	t_total_dist_5	Total trip distance with C (metres)	144.00	30003.00	7201.31	5929.25
	t_total_time_1	Total trip time with CD (min.)	0.00	60.00	15.47	9.35
	t_total_time_2	Total trip time with CP (min.)	0.00	57.00	12.63	9.37
	t_in_vehicle_time_3	Total in vehicle time with PT (min.)	0.00	126.00	29.06	18.37
	t_in_vehicle_time_6	Total in vehicle time with TR (min.)	4.00	112.00	31.87	14.47
	t_in_vehicle_time_7	Total in vehicle time with PR (min.)	12.00	124.00	43.49	21.80
	t_in_vehicle_time_8	Total in vehicle time with KR (min.)	6.00	93.00	34.64	15.48
	t_access_egress_time_3	Access and egress time with PT (min.)	0.00	57.00	18.26	8.57
	t_access_egress_time_6	Access and egress time with TR (min.)	5.00	275.00	66.96	43.94
	t_access_egress_time_7	Access and egress time with PR (min.)	5.00	419.00	52.10	55.19
	t_access_egress_time_8	Access and egress time with KR (min.)	7.00	350.00	53.91	53.95
	t_waiting_time_3	Waiting time with PT (min.)	0.00	33.00	1.30	2.62
	t_nb_transfers_3	Number of transfers with PT	0.00	4.00	0.60	0.72
	t_activity_duration_1	Duration of activity when completed with CD (min.)	-30.00	1354.00	375.21	242.83
	t_activity_duration_2	Duration of activity when completed with CP (min.)	-5.00	1165.00	363.05	214.74
	t_activity_duration_3	Duration of activity when completed with PT (min.)	-90.00	920.00	398.92	185.65
	t_activity_duration_4	Duration of activity when completed with W (min.)	-64.00	745.00	259.40	191.66
	t_activity_duration_5	Duration of activity when completed with C (min.)	-47.00	859.00	357.91	237.48
	t_activity_duration_6	Duration of activity when completed with TR (min.)	-109.00	902.00	419.41	159.40
	t_activity_duration_7	Duration of activity when completed with PR (min.)	-260.00	840.00	462.69	171.75
	t_activity_duration_8	Duration of activity when completed with KR (min.)	-180.00	732.00	448.22	156.13

Table 1b Description of studied variables for the chain based model

Variable type	Name of the variable	Description	*Min*	*Max*	$\bar{X}$	σ
	Choice	Observed chosen alternative				
Sociodemographic variables	h_nb_car	Number of cars at home	0.00	14.00	1.73	1.11
	h_nb_adulte	Number of adults at home	1.00	8.00	2.40	0.95
	h_nb_kid	Number of kids at home	0.00	5.00	0.74	1.00
	i_age_0_15	Individual aged between 0 and 15 years old (0 or 1)	0.00	1.00	0.09	0.28
	i_age_16_24	Individual aged between 16 and 24 years old (0 or 1)	0.00	1.00	0.15	0.36
	i_age_25_64	Individual aged between 25 and 64 years old (0 or 1)	0.00	1.00	0.69	0.46
	i_age_65_over	Individual aged of 65 years old and over (0 or 1)	0.00	1.00	0.07	0.26
	i_gender_m	Individual of masculine gender (0 or 1)	0.00	1.00	0.51	0.50
	i_gender_f	Individual of feminine gender (0 or 1)	0.00	1.00	0.49	0.50
	i_status_work_study	Individual with worker of student status (0 or 1)	0.00	1.00	0.80	0.40
	i_pt_tr_fare_yes	Individual with a PT or TR monthly fare (0 or 1)	0.00	1.00	0.09	0.28
	i_pt_tr_fare_no	Individual without a PT or TR monthly fare (0 or 1)	0.00	1.00	0.69	0.46
	i_license_yes	Individual with a driving license (0 or 1)	0.00	1.00	0.81	0.39
Characteristics of the chain	ch_motive_work	Chain with a work motive (0 or 1)	0.00	1.00	0.53	0.50
	ch_motive_non_work	Chain with a non-work motive (0 or 1)	0.00	1.00	0.27	0.44
	ch_total_cost_1	Total monetary cost of alternative 1 ($)	0.02	17.11	4.01	2.87
	ch_total_cost_2	Total monetary cost of alternative 2 ($)	0.00	8.50	1.27	2.46
	ch_total_cost_4	Total monetary cost of alternative 4 ($)	0.00	13.50	6.71	3.54
	ch_total_cost_5	Total monetary cost of alternative 5 ($)	28.23	216.52	83.58	34.05
	ch_total_cost_7	Total monetary cost of alternative 7 ($)	0.00	12.25	6.56	3.28
	ch_total_cost_9	Total monetary cost of alternative 9 ($)	0.00	3.00	1.41	1.50
	ch_total_cost_10	Total monetary cost of alternative 10 ($)	0.00	3.50	1.40	1.52
	ch_total_cost_11	Total monetary cost of alternative 11 ($)	0.00	9.75	4.04	3.29
	ch_total_cost_12	Total monetary cost of alternative 12 ($)	3.00	6.75	3.89	0.95
	ch_total_cost_13	Total monetary cost of alternative 13 ($)	2.25	4.50	3.45	0.79
	ch_total_cost_14	Total monetary cost of alternative 14 ($)	0.00	6.00	3.58	1.76
	ch_total_dist_6	Total distance of alternative 6 (metres)	32.00	88036.00	3682.06	8545.03
	ch_total_dist_8	Total distance of alternative 8 (metres)	288.00	59960.00	14402.61	11854.34
	ch_total_time_1	Total trip time of alternative 1 (min.)	0.00	120.00	30.93	18.59
	ch_total_time_3	Total trip time of alternative 3 (min.)	0.00	114.00	25.58	19.19
	ch_total_time_9	Total trip time of alternative 9 (min.)	16.00	116.00	60.59	28.75
	ch_total_time_10	Total trip time of alternative 10 (min.)	38.00	128.00	71.92	25.19
	ch_total_time_12	Total trip time of alternative 12 (min.)	95.00	315.00	173.55	61.80
	ch_total_time_13	Total trip time of alternative 13 (min.)	60.00	252.00	164.00	57.48
	ch_total_time_14	Total trip time of alternative 14 (min.)	54.00	155.00	98.22	31.51
	ch_total_time_15	Total trip time of alternative 15 (min.)	10.00	60.00	26.00	18.08
	ch_in_vehicle_time_2	Total in-vehicle time of alternative 2 (min.)	0.00	222.00	57.29	35.25
	ch_in_vehicle_time_4_5_7_11	Total in-vehicle time of alternatives 4,5,7 et 11 (min.)	12.00	248.00	71.63	35.45
	ch_access_egress_time_2	Access and egress time of alternative 2 (min.)	0.00	107.00	36.18	13.93
	ch_access_egress_time_4	Access and egress time of alternative 4 (min.)	10.00	550.00	135.21	89.31
	ch_access_egress_time_5_7	Access and egress time of alternative 5 et 7 (min.)	10.00	517.00	71.39	66.19
	ch_access_egress_time_11	Access and egress time of alternative 11 (min.)	23.00	193.00	86.37	40.85
	ch_nb_transfers_2	Total number of transfers with alternative 2 (min.)	0.00	6.00	1.17	1.13
	ch_duration_activity_1	Duration of activity when completed with alternative 1 (min.)	-65.00	1327.00	378.57	235.83
	ch_duration_activity_2	Duration of activity when completed with alternative 2 (min.)	-115.00	869.00	367.57	172.13
	ch_duration_activity_3	Duration of activity when completed with alternative 3 (min.)	-36.00	1129.00	354.77	212.78
	ch_duration_activity_4	Duration of activity when completed with alternative 4 (min.)	-397.00	781.00	344.95	142.26
	ch_duration_activity_5	Duration of activity when completed with alternative 5 (min.)	27.00	778.00	434.41	126.19
	ch_duration_activity_6	Duration of activity when completed with alternative 6 (min.)	-455.00	740.00	233.63	204.78
	ch_duration_activity_7	Duration of activity when completed with alternative 7 (min.)	-482.00	681.00	395.26	176.05
	ch_duration_activity_8	Duration of activity when completed with alternative 8 (min.)	-154.00	817.00	376.15	214.90
	ch_with_anchor_mode_1_5_8	First trip made with a "with anchor" mode (0 or 1)	1.00	1.00	1.00	0.00

4.3.2 Calculation of travel time

Car, cycling and walking: Car, cycling and walking travel times are determined using a shortest path algorithm over the OSM (Open Street Map) networks. Then, by assigning an average speed to each mode, we obtain travel time. The tool used is developed by our research group based on the code of Luxen and Vetter (2011).

Public transit: A tool developed from the code of Dibbelt et al. (2013) in our research group estimates a time for transit trips using General Transit Feed Specification (GTFS) data. Different components of the total travel time are estimated: walking time, waiting time and in-vehicle time. The number of transfers is also provided.

Train: To calculate the train travel time, we consider that access and egress are made by walking. The choice of departure and arrival train stations is determined by proximity to origin and destination points. Then a travel time matrix between stations is derived from the train schedule.

4.3.3 Calculation of monetary travel cost

Car (driver): The average price of gasoline in the Montreal Area is 137.7 Canadian cents per liter in January 2014 (Québec 2014). The average consumption of vehicles is 8 L/100km (EPA 2014; Berg 2014). Statistics on the 2014 consumption are considered despite the fact that the data are from 2008 to ensure consistency between prices of alternatives as transit fares of 2014 are used.

Train and public transit: 2014 transit fairs are used: when an individual owns a monthly fare for a given mode, the cost is null as the subscription fees are considered long term investments.

5. Models and results

5.1 Data sample

After the various filtering steps, the sample of data corresponds to 8,888 trips (4,444 simple chains). The estimation of the models is made using 80% of the observed data, determined randomly, which corresponds to 7,100 trips and 3,550 chains. The simulation is made from the 20% remaining data, but here they are weighted, corresponding to 39,969 trips and 19,938 chains.

5.2 Definition of models

Both estimated models are multinomial (MNL) discrete choice models (Train 2009; Ortúzar and Willumsen 2011). These models are based on the fact that an individual q wishes to maximize its utility U. An individual will choose alternative j if $U_{jq} \geq U_{iq} \; \forall \; j \neq i$. Thus, the probability that an individual chooses alternative j from the set of alternatives **A** is expressed as follows:

$$P_{jq} = \frac{e^{\beta \, U_{jq}}}{\sum_{A, i \, \epsilon \, A} e^{\beta \, U_{iq}}}$$

Since the utility U is unknown, the concept of representative utility is used where $U_{jq} = V_{qj} + \varepsilon_{qj}$. The different variables described above are used to calculate V_{qj}. The models are estimated using the open source software BIOGEME (2003).

5.3 Estimations

5.3.1 Trip based model

Several tests were conducted with various sets of variables and the results of the selected model are presented in Table 2. Some of the parameters may not be significant, but are kept in the model, as they are considered important in the understanding of the phenomena under study.

The results of the trip based model are generally consistent with the literature except for few elements, namely access and egress times of bimodal sequences (Park & Ride, Kiss & Ride) which have small positive coefficients with lower significance. One possible reason for that is the important variability of the values, as can be seen in Table 1a (average access time varying between 18 and 67 minutes and variation coefficient above 100% for some combinations).

Table 2 Results of the estimation for the trip based model (part 1)

Name	*Value*	*Robust Std err*	*Robust t-test*	*p-value*
ASC_CD	0.0000			
ASC_CP	-1.3200	0.2570	-5.1200	0.0000
ASC_KRTR	-1.2200	0.4780	-2.5400	0.0100
ASC_W	1.1900	0.5570	2.1300	0.0300
ASC_PRTR	-1.3000	0.5580	-2.3300	0.0200
ASC_PT	3.6600	0.4040	9.0600	0.0000
ASC_TR	1.2000	0.7080	1.6900	0.0900
ASC_C	-3.8600	0.5980	-6.4600	0.0000
AGE_0_15 and AGE_25_64 (alt. 1)	0.0000			
AGE_16_24_1	-0.5660	0.3510	-1.6100	0.1100
AGE_65_1	0.3550	0.2740	1.3000	0.1900
AGE_0_15 and AGE_65 (alt. 2)	0.0000			
AGE_16_24_2	-0.7910	0.2790	-2.8300	0.0000
AGE_25_64_2	-1.3000	0.2620	-4.9500	0.0000
AGE_16_24 and AGE_65 (alt. 3)	0.0000			
AGE_0_15_3	-1.4800	0.2940	-5.0200	0.0000
AGE_25_64_3	-0.7750	0.3310	-2.3400	0.0200
AGE_0_15, AGE_25_64 and AGE_65 (alt. 4)	0.0000			
AGE_16_24_4	-0.7970	0.5930	-1.3400	0.1800
AGE_0_15, AGE_25_64 and AGE_65 (alt. 5)	0.0000			
AGE_16_24_5	1.1200	0.4610	2.4200	0.0200
AGE_0_15 and AGE_65 (alt. 6)	0.0000			
AGE_16_24_6	0.4850	0.5640	0.8600	0.3900
AGE_25_64_6	1.4200	0.4560	3.1100	0.0000
AGE_0_15, AGE_16_24 and AGE_65 (alt. 7)	0.0000			
AGE_25_64_7	0.9820	0.4650	2.1100	0.0300
AGE_0_15, AGE_16_24 and AGE_65 (alt. 8)	0.0000			
AGE_25_64_8	-0.6850	0.4140	-1.6500	0.1000
TOTAL_COST (alt. 4,5)	0.0000			
TOTAL_COST_1_2_3_6_7_8	-0.9180	0.0363	-25.3200	0.0000
TOTAL_DIST (alt. 1,2,3,6,7,8)	0.0000			
TOTAL_DIST_4	-0.0025	0.0003	-7.5100	0.0000
TOTAL_DIST_5	-0.0003	0.0001	-2.5900	0.0100
DURATION_ACTIVITY_OTHER (alt. 4,5,8)	0.0000			
DURATION_ACTIVITY_OTHER_1	0.0035	0.0027	1.3000	0.1900
DURATION_ACTIVITY_OTHER_2	0.0061	0.0027	2.2500	0.0200
DURATION_ACTIVITY_OTHER_3	0.0067	0.0033	2.0600	0.0400
DURATION_ACTIVITY_OTHER_6	-0.0057	0.0078	-0.7200	0.4700
DURATION_ACTIVITY_OTHER_7	-0.0162	0.0089	-1.8200	0.0700

Table 2 Results of the estimation for the trip based model (part 2)

Name	*Value*	*Robust Std err*	*Robust t-test*	*p-value*
PURPOSE_OTHER (alt. 1,2,6,7,8)	0.0000			
PURPOSE_WORK_STUDY_1	-0.4460	0.1180	-3.7700	0.0000
PURPOSE_WORK_STUDY_2	-0.5640	0.1340	-4.2100	0.0000
PURPOSE_WORK_STUDY_6	0.5930	0.2890	2.0500	0.0400
PURPOSE_WORK_STUDY_7	1.2900	0.3190	4.0400	0.0000
PURPOSE_WORK_STUDY_8	0.9980	0.3630	2.7500	0.0100
H_NB_CAR (alt. 3,4,5,6,7,8)	0.0000			
H_NB_CAR_1_2	0.4830	0.0730	6.6200	0.0000
H_NB_ADULT (alt. 1,3,7,8)	0.0000			
H_NB_ADULT_2	0.1710	0.0507	3.3800	0.0000
H_NB_ADULT_4	0.5630	0.1940	2.9000	0.0000
H_NB_ADULT_5	0.5740	0.1910	3.0000	0.0000
H_NB_ADULT_6	0.3440	0.1400	2.4600	0.0100
H_NB_KID_OTHER (alt. 3,4,5,6,7,8)	0.0000			
H_NB_KID_1	0.4230	0.1600	2.6500	0.0100
H_NB_KID_2	0.2750	0.1530	1.8000	0.0700
NB_TRANSFERS (alt. 1,2,4,5,6,7,8)	0.0000			
NB_TRANSFERS_3	-0.1280	0.0970	-1.3200	0.1900
LICENSE_NO (alt. 2)	0.0000			
LICENSE_YES_2	-0.5600	0.1690	-3.3100	0.0000
GENDER_M (alt. 2)	0.0000			
GENDER_F_2	1.0700	0.1010	10.6000	0.0000
GENDER_F (alt. 3,4,6,7,8)	0.0000			
GENDER_M_3	-0.4270	0.1180	-3.6100	0.0000
GENDER_M_4	-1.0800	0.2680	-4.0400	0.0000
GENDER_M_6	-1.0700	0.2750	-3.9100	0.0000
GENDER_M_7	-0.8810	0.2610	-3.3800	0.0000
GENDER_M_8	-0.8360	0.3280	-2.5500	0.0100
ACCESS_EGRESS_TIME (alt. 1,2,4,5)	0.0000			
ACCESS_EGRESS_TIME_3	-0.0681	0.0104	-6.5300	0.0000
ACCESS_EGRESS_TIME_6	-0.0669	0.0200	-3.3400	0.0000
ACCESS_EGRESS_TIME_7_8	0.0103	0.0083	1.2400	0.2100
WAITING_TIME (alt. 1,2,4,5,6,7,8)	0.0000			
WAITING_TIME_3	-0.0413	0.0174	-2.3700	0.0200

Table 2 Results of the estimation for the trip based model (part 3)

Name	***Value***	***Robust Std err***	***Robust t-test***	***p-value***
TOTAL_TIME (alt. 3,4,5,6,7,8)	0.0000			
TOTAL_TIME_1	-0.0356	0.0118	-3.0100	0.0000
TOTAL_TIME_2	-0.1730	0.0127	-13.6300	0.0000
IN_VEHICLE_TIME (alt. 1,2,4,5,6,7,8)	0.0000			
IN_VEHICLE_TIME_3	-0.0631	0.0051	-12.3800	0.0000
STATISTICS OF THE MODEL				
FINAL LOG-LIKELIHOOD	-3051.0950			
LIKELIHOOD RATIO TEST	6895.7980			
AJUSTED R^2	0.5220			
NUMBER OF OBSERVATIONS	7100			

5.3.2 Chain based model

The results of the chain based model are presented in Table 3. The method used is the same as for the trip based model.

We observe that access and egress times of alternative 2 (transit-transit), 4 (Train-Train), 5 (PR-PR), 7 (KR-KR) and 11 (KR-TR) have positive coefficients. For the bimodal sequences, it may be explained by the fact that a traveler may wish to maximize the portion of his trip done by car before switching to transit; wide dispersion of values or small samples (see Table 1b) for these access times can also affect the estimations.

The total time has a negative impact on the probability to use the car, either driver (1) or passenger (3). Alternatives 12 (PT-TR) and 15 (CP-W) have fewer observations so the small positive coefficients are not to be worried about.

In-vehicle time has a negative effect on the use of the alternative 2. This variable appears to have a positive effect for alternative 4, 5, 7 and 11. One possible explanation is the fact that train commuters can do something else while travelling (e.g. checking emails, reading, napping etc.), consequently increasing the utility of this alternative. Such results has already been observed in other studies (Mokhtarian and Salomon 2001). Since those with longer travel times on the train have higher probability of having a seat, it could further explain the positive impact as time increases.

Table 3 Results of the estimation for the chain based model (part 1)

Name	*Value*	*Robust Std err*	*Robust t-test*	*p-value*
ASC (alt. 2,4,5,6, 9 to 15)	0.00000			
ASC_1	0.47500	0.22700	2.09000	0.04000
ASC_2	0.39500	0.41100	0.96000	0.34000
ASC_7	-0.31800	0.44300	-0.72000	0.47000
ASC_8	-0.37200	0.51900	-0.72000	0.47000
NB_KID_PURPOSE_OTHER (alt. 2, 4 to 15)	0.00000			
NB_KID_PURPOSE_OTHER_1	0.28100	0.10300	2.72000	0.01000
NB_KID_PURPOSE_OTHER_3	0.31400	0.11600	2.70000	0.01000
CH_TOTAL_COST (alt. 2,6,8,15)	0.00000			
CH_TOTAL_COST_1_3_4_5_7_9_10_11_12_13_14	-0.07520	0.00708	-10.62000	0.00000
CH_TOTAL_DIST (alt. 1 to 5, 9 to 15)	0.00000			
CH_TOTAL_DIST_6	0.00070	0.00015	4.61000	0.00000
CH_TOTAL_DIST_8	0.00009	0.00003	2.80000	0.01000
CH_DURA_ACTIVITY_OTHER (alt. 4,5,7, 9 to 15)	0.00000			
CH_DURA_ACTIVITY_OTHER_1	0.00441	0.00156	2.83000	0.00000
CH_DURA_ACTIVITY_OTHER_2	0.01010	0.00261	3.87000	0.00000
CH_DURA_ACTIVITY_OTHER_3	0.00694	0.00179	3.88000	0.00000
CH_DURA_ACTIVITY_OTHER_6	0.00693	0.00295	2.35000	0.02000
CH_DURA_ACTIVITY_OTHER_8	-0.00388	0.00421	-0.92000	0.36000
CH_PURPOSE_OTHER (alt. 2)	0.00000			
CH_PURPOSE_WORK_STUDY_2	0.52200	0.27700	1.89000	0.06000
CH_NB_TRANSFERS (alt. 1,3,4,5,6,7,8,14,15)	0.00000			
CH_NB_TRANSFERS_2_9_10_12_13	-0.13800	0.06330	-2.19000	0.03000
CH_ACCESS_EGRESS_TIME (1,3,6,8,9,10, 12 to 15)	0.00000			
CH_ ACCESS_EGRESS_TIME_11	0.01580	0.00764	2.07000	0.04000
CH_ ACCESS_EGRESS_TIME_2	0.03550	0.00640	5.55000	0.00000
CH_ ACCESS_EGRESS_TIME_4	0.04270	0.00376	11.37000	0.00000
CH_ ACCESS_EGRESS_TIME_5_7	0.02810	0.00688	4.08000	0.00000
CH_TOTAL_TIME (alt. 2, 4 to 11)	0.00000			
CH_TOTAL_TIME_1_3	-0.00501	0.00545	-0.92000	0.36000
CH_TOTAL_TIME_12	0.02170	0.00260	8.35000	0.00000
CH_TOTAL_TIME_13	0.02050	0.00278	7.37000	0.00000
CH_TOTAL_TIME_14	0.02240	0.00537	4.17000	0.00000
CH_TOTAL_TIME_15	0.04950	0.02110	2.34000	0.02000
CH_IN_VEHICLE_TIME (alt. 1,3,6,8,9,10, 12 to 15)	0.00000			
CH_IN_VEHICLE_TIME_2	-0.00203	0.00255	-0.80000	0.43000
CH_IN_VEHICLE_TIME_4_5_7_11	0.04340	0.00443	9.81000	0.00000
NB_CAR (alt. 2,4,6,8,12,13)	0.00000			
NB_CAR_1_3_5_7_9_10_11_14_15	0.13900	0.07180	1.94000	0.05000

Table 3 Results of the estimation for the chain based model (part 2)

Name	*Value*	*Robust Std err*	*Robust t-test*	*p-value*
NB_ADULT (alt. 4,9,10, 12 to 15)	0.00000			
NB_ADULT_1	0.10800	0.09190	1.17000	0.24000
NB_ADULT_11	-0.19300	0.14800	-1.30000	0.19000
NB_ADULT_2	0.44400	0.10500	4.22000	0.00000
NB_ADULT_3	0.53000	0.09210	5.76000	0.00000
NB_ADULT_5	-0.41100	0.12500	-3.29000	0.00000
NB_ADULT_6	0.61600	0.17600	3.51000	0.00000
NB_ADULT_7	-0.27100	0.18800	-1.44000	0.15000
NB_ADULT_8	-0.28300	0.22700	-1.24000	0.21000
NO_ANCHOR_MODE (alt. 1,5,8)	0.00000			
WITH_ANCHOR_MODE_1_5_8	4.31000	0.07870	54.75000	0.00000
P_AGE_0_15, P_AGE_16_24, P_AGE_65_OVER (alt. 1)	0.00000			
P_AGE_25_64_1	0.66100	0.11200	5.89000	0.00000
P_GENDER_M (alt. 3)	0.00000			
P_GENDER_F_3	0.31700	0.20600	1.54000	0.12000
P_GENDER_F (alt.1,2)	0.00000			
P_GENDER_M_1	0.31900	0.13900	2.30000	0.02000
P_GENDER_M_2	0.16700	0.17700	0.94000	0.35000
P_GENDER_F_LICENSE_NO (alt. 3)	0.00000			
P_GENDER_M_LICENSE_YES_3	-0.41100	0.23700	-1.73000	0.08000
P_STATUS_OTHER (alt.2,3)	0.00000			
P_STATUS_WORK_STUDY_2	0.28400	0.24600	1.15000	0.25000
P_STATUS_WORK_STUDY_3	0.27600	0.15100	1.83000	0.07000
STATISTICS OF THE MODEL				
FINAL LOG-LIKELIHOOD	-1253.56800			
LIKELIHOOD RATIO TEST	6995.13000			
AJUSTED R^2	0.72700			
NUMBER OF OBSERVATIONS	3550			

5.3.3 Comparison and validation

Once the odds from the estimation of the trip based and chain based models are estimated, the probability of various alternatives is calculated and the choice of transport mode is simulated using the Monte Carlo method on a sub-sample not used in estimation. This method determines the probability that a mode is selected from the group of choices with a random value.

Two different methods are used to compare simulations and observations. The first one is more general and simply compares resulting mode shares (Table 4). The second is more specific and relies on the success rate of the simulation. It is based on the confusion matrix, which allows comparing observed and simulated mode choice (Table 5).

In order to draw comparison between the two models, the results of the chain based model are analyzed at the trip level. Thus, alternatives of mode sequences are divided and used to study the simulation results through eight different modes rather than 15 different chains.

Comparison of modal shares: From the results of both models, we clearly see that the modal shares of the simulated alternatives from the chain based model are much closer to the observed ones than for the trip based model. The scale of differences, for all modes, is much lower for the chain based model with the most important percentage of errors being observed for the less popular sequences. Hence, the percentage of error drastically decreases for the dominant alternative (car driver): only 1.8% for the chain model against 78.8% for the trip model. Generally, a large improvement in the prediction of modal shares is observed from the chain based model, particularly for the main modes: car (driver), car (passenger), transit, walking and bike. The use of an efficient train corridor with peak-based schedules enhances the differences between the trip-based and chain-based models.

Table 4 Comparison of modal shares for a) the trip based model and b) chain based model

a) Trip based		***Observed***		***Simulated***		***Difference***	
No	**Alternative**	**Nb**	**%**	**Nb**	**%**	**Nb**	**%**
1	Car (driver)	18,350	45.9%	3,889	9.7%	-14,461	-78.8%
2	Car (passenger)	5,223	13.1%	1,246	3.1%	-3,977	-76.1%
3	Public transit	9,381	23.5%	4,528	11.3%	-4,853	-51.7%
4	Walking	957	2.4%	7,231	18.1%	6,274	655.6%
5	Cycling	524	1.3%	5,914	14.8%	5,390	1028.6%
6	Train	3,375	8.4%	6,752	16.9%	3,377	100.1%
7	Park & Ride	1,441	3.6%	5,416	13.6%	3,975	275.9%
8	Kiss & Ride	718	1.8%	4,993	12.5%	4,275	595.4%
	Total	**39,969**	**100.0%**	**39,969**	**100.0%**	**-**	**-**

b) Chain based		***Observed***		***Simulated***		***Difference***	
No	**Alternative**	**Nb**	**%**	**Nb**	**%**	**Nb**	**%**
1	Car (driver)	18,220	45.7%	17,896	44.9%	-324	-1.8%
2	Car (passenger)	5,365	13.5%	4,272	10.7%	-1,093	-20.4%
3	Public transit	9,334	23.4%	7,986	20.0%	-1,348	-14.4%
4	Walking	960	2.4%	1,301	3.3%	341	35.5%
5	Cycling	526	1.3%	702	1.8%	176	33.5%
6	Train	3,358	8.4%	4,636	11.6%	1,278	38.1%
7	Park & Ride	1,442	3.6%	2,048	5.1%	606	42.0%
8	Kiss & Ride	671	1.7%	1,035	2.6%	364	54.2%
	Total	**39,876**	**100.0%**	**39,876**	**100.0%**	**-**	**-**

Confusion matrices: Comparing the results of the two models from their respective confusion matrix (Table 5), we see that 12.74% of trips were correctly simulated from the trip based model compared to 72.15% with the chain based model. The accuracy of simulated modal shares is reflected in the confusion matrix, because here again the results are particularly good in the simulation of the car (driver). In fact, 89.35% of the observed trips whose choice is car (driver) are simulated correctly. Nevertheless, even if the results of the chain based model are far better than those of the trip based model when estimated with about 80% of the same variables, some modes are still not correctly simulated. Indeed, the car (passenger), walking, cycling and the Kiss & Ride have low success rates, not exceeding 46.98%.

While for most of these modes, the erroneous simulations are distributed relatively evenly throughout the other modes; those of cycling appear particularly to be absorbed by the car (driver), where 57.79% of the observed cycling trips have been simulated as car (driver) trips. This phenomenon could be explained by the fact that these two modes have similar characteristics, but the travel time by car (driver) is much shorter and the availability of the bike is limited to trips of less than 5.7 km. In fact, with the exception of car (driver) and train, each mode seems to be more absorbed by one or two specific modes: car (passenger) is absorbed by transit, transit by car (passenger), walking by car (passenger) and transit, Park & Ride by car (driver) and finally, Kiss & Ride by transit.

Table 5 Confusion matrix of the a) trip based model and b) chain based model

a)	**Simulated choice**								
Observed choice	**CD**	**CP**	**PT**	**W**	**C**	**TR**	**PR**	**KR**	**Total**
CD	**10.84%**	1.41%	8.90%	18.83%	16.44%	17.78%	12.75%	13.04%	100%
CP	10.89%	**5.32%**	9.50%	16.64%	11.49%	16.20%	16.08%	13.88%	100%
PT	8.77%	5.82%	**19.09%**	16.10%	10.47%	14.70%	13.43%	11.62%	100%
W	35.01%	12.33%	11.60%	**5.96%**	1.78%	11.70%	12.85%	8.78%	100%
C	10.31%	2.10%	8.78%	20.99%	**14.89%**	18.13%	11.83%	12.98%	100%
TR	2.58%	0.80%	8.62%	21.93%	23.64%	**18.10%**	13.24%	11.08%	100%
PR	1.67%	0.42%	8.33%	22.76%	21.03%	21.93%	**12.91%**	10.96%	100%
KR	1.11%	0.14%	5.43%	22.42%	16.57%	18.25%	22.01%	**14.07%**	100%
Total	9.73%	3.12%	11.33%	18.09%	14.80%	16.89%	13.55%	12.49%	**100%**

b)	**Simulated choice**								
Observed choice	**CD**	**CP**	**PT**	**W**	**C**	**TR**	**PR**	**KR**	**Total**
CD	**89.35%**	2.34%	2.78%	0.44%	1.11%	1.31%	1.87%	0.80%	100%
CP	7.72%	**38.30%**	25.13%	6.86%	2.42%	11.37%	3.69%	4.51%	100%
PT	4.72%	14.83%	**59.94%**	3.92%	1.45%	8.79%	3.40%	2.96%	100%
W	5.63%	20.31%	15.00%	**46.98%**	0.42%	7.19%	2.50%	1.98%	100%
C	57.79%	0.00%	0.76%	1.14%	**37.64%**	0.76%	1.52%	0.38%	100%
TR	1.34%	3.16%	5.99%	0.54%	0.42%	**84.57%**	2.44%	1.55%	100%
PR	22.61%	0.97%	2.43%	0.00%	0.42%	0.35%	**73.09%**	0.14%	100%
KR	4.77%	13.71%	22.65%	1.64%	1.94%	7.30%	3.73%	**44.26%**	100%
Total	44.88%	10.71%	20.03%	3.26%	1.76%	11.63%	5.14%	2.60%	**100%**

6. Conclusion

In this paper, we develop an approach that includes the full spatiotemporal structure of the trip chain in the modeling of mode choice. Our method, based on the classification of modes, contributes in developing the integration of trip chains in mode choice models by using the concept of anchor point. Thus, modes are classified in two categories: "with anchor" and "no anchor". This classification allows an enumeration of the possible sequences of modes an individual can choose for a given chain. The category of the mode chosen for the first trip of a chain has an impact on the mode choice for the next trips. Once this fact is taken into account, a notable improvement of the prediction is observed.

The models developed in the case study are experimental, so some flexibility was granted in the estimation because it was important to make sure that they could be compared. Hence, similar sets of variables were selected. Also, value far from the central distribution were found within the distribution of some variables like the total distance of alternative 6 (walk-walk) and 8 (cycle-cycle). This could have implications on some estimates.

Furthermore, it could be useful to try this approach using a more complex model like the Cross Nested Logit (CNL) instead of the MNL as tested in the case study. Indeed, the CNL model could allow the estimation of different chain types at the same time as the choice of the chain structure would be included in the model. In addition, it could help resolve the correlation between alternatives.

This research aimed to highlight the interdependency of mode choice and trip chain by experiencing the estimation of a chain based model and compare it to a trip base model using a train corridor. Comparing these models, estimated with about 80% of the same variables, the benefits of the proposed method are notables. While the simulation of the trip based model has a success rate of 12.74% that of the chain based model is 72.15%. Thus, a significant improvement is observed by incorporating the concept of trip chain, particularly for car (driver) (89.35% against 10.84%), train (84.57% against 18.10%) as well as Park & Ride (73.09% against 12.91%). In short, integration of chains for modeling transport habits is an interesting way to improve planning tools and it is encouraging that their potential continues to be exploited.

Acknowledgements

The authors wish to acknowledge the contribution and support of the Chaire Mobilité partners (Ville de Montréal, Ministère des transports du Québec, Agence métropolitaine de transport and Société de transport de Montréal). They also wish to thank Jean-Simon Bourdeau and Pierre-Léo Bourbonnais for their contributions.

References

AMT. 2010. "Mobilité des personnes dans la Région de Montréal Enquête Origine-Destination 2008." In *Enquête Origine-Destination*, 210. Montréal: Agence Métropolitaine de Transport.

Asensio, Javier. 2002. "Transport mode choice by commuters to Barcelona's CBD." *Urban Studies* 39 (10):1881-95.

Berg, Frédéric. 2014. "Consommation d'essence: enfin des cotes réalistes." *Protégez-Vous*.

Bierlaire, Michel. 2003. BIOGEME: A free package for the estimation of discrete choice models. Paper presented at the 3rd Swiss Transportation Research Conference, Ascona, Switzerland.

Currie, Graham, and Alexa Delbosc. 2011. "Exploring the trip chaining behaviour of public transport users in Melbourne." *Transport Policy* 18 (1):204-10. doi: http://dx.doi.org/10.1016/j.tranpol.2010.08.003.

Dibbelt, Julian, Thomas Pajor, Ben Strasser, and Dorothea Wagner. 2013. "Intriguingly Simple and Fast Transit Routing." In *Experimental Algorithms*, edited by Vincenzo Bonifaci, Camil Demetrescu and Alberto Marchetti-Spaccamela, 43-54. Springer Berlin Heidelberg.

Dong, Xiaojing, Moshe E. Ben-Akiva, John L. Bowman, and Joan L. Walker. 2006. "Moving from trip-based to activity-based measures of accessibility." *Transportation Research Part A: Policy and Practice* 40 (2):163-80.

EPA. 2014. "Fuel economy guide model 2014." In, edited by US Department of Energy, 49 p. United States: Office of Energy Efficiency and Renewable Energy.

Godefroy, François. 2011. "Méthodologie de caractérisation du vélopartage et d'estimation du marché potentiel du vélo à Montréal." Mémoire, École Polytechnique de Montréal.

Harding, Christopher, Eric J Miller, Zachary Patterson, and Kay W Axhausen. 2015. "Multiple Purpose Tours and Efficient Trip Chaining: An Analysis of the Effects of Land Use and Transit on Travel Behavior in Switzerland." In.

Hensher, DavidA, and AprilJ Reyes. 2000. "Trip chaining as a barrier to the propensity to use public transport." *Transportation* 27 (4):341-61. doi: 10.1023/A:1005246916731.

Ho, Chinh, and Corinne Mulley. 2013. "Tour-based mode choice of joint household travel patterns on weekend and weekday." *Transportation* 40 (4):789-811. doi: 10.1007/s11116-013-9479-0.

Islam, Md Tazul, and Khandker M. Nurul Habib. 2012. "Unraveling the relationship between trip chaining and mode choice: evidence from a multi-week travel diary." *Transportation Planning and Technology* 35 (4):pp 409-26.

Luxen, Dennis, and Christian Vetter. 2011. "Real-time routing with OpenStreetMap data." In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 513-6. Chicago, Illinois: ACM.

Martel-Poliquin, Éric. 2012. "Mieux comprendre les déterminants du choix modal." Mémoire, École Polytechnique de Montréal.

McNally, Michael G. 2007. "The Four Step Model." In, 35-41.

Miller, EricJ, MatthewJ Roorda, and JuanAntonio Carrasco. 2005. "A tour-based model of travel mode choice." *Transportation* 32 (4):399-422. doi: 10.1007/s11116-004-7962-3.

Mokhtarian, Patricia, and Ilan Salomon. 2001. "How Derived is the Demand for Travel? Some Conceptual and Measurement Considerations." In.

Ortúzar, Juan de Dios, and Luis G. Willumsen. 2011. "Discrete Choice Models." In *Modelling Transport*, 227-68. John Wiley & Sons, Ltd.

Primerano, Frank, Michael A. P. Taylor, Ladda Pitaksringkarn, and Peter Tisato. 2008. "Defining and Understanding Trip Chaining Behaviour." *Transportation: Planning, Policy, Research, Practice* 35 (1):pp 55-72.

Québec, Régie de l'énergie du. 2014. "Évolution des prix de détail de l'essence ordinaire en 2013 Région de Montréal." In. Québec.

Roorda, Matthew J., Dylan Passmore, and Eric J. Miller. 2009. "Including Minor Modes of Transport in a Tour-Based Mode Choice Model with Household Interactions." *Journal of Transportation Engineering* 135 (12):pp 935-45.

Sondage, Echo. 2011. "État de la pratique du vélo au Québec en 2010."

Steenberghen, T., P. Toint, and J. Zuallaert. 2005. "Déterminants des choix modaux dans les chaînes de déplacements." In, edited by Politique scientifique fédérale, 01-15.

Strathman, James G., and Kenneth J. Dueker. 1995. "Understanding Trip Chaining." In *1990 NPTS Report Series on Trip and Vehicule Attributes*, edited by U.S Department of Transportation Federal Highway Administration, 25 p. Portland: U.S Department of Transportation Federal Highway Administration.

Train, K.E. 2009. *Discrete Choice Methods with Simulation*: Cambridge University Press.

Valiquette, François. 2010. "Typologie des chaînes de déplacements et modélisation descriptive des systèmes d'activités des personnes." Mémoire, École Polytechnique de Montréal.

Vande Walle, S., and T. Steenberghen. 2006. "Space and time related determinants of public transport use in trip chains." *Transportation Research Part A: Policy and Practice* 40 (2):151-62.

Xianyu, Jianchuan. 2013. "An Exploration of the Interdependencies between Trip Chaining Behavior and Travel Mode Choice." *Procedia - Social and Behavioral Sciences* 96 (0):1967-75. doi: http://dx.doi.org/10.1016/j.sbspro.2013.08.222.

Ye, Xin, Ram M. Pendyala, and Giovanni Gottardi. 2007. "An exploration of the relationship between mode choice and complexity of trip chaining patterns." *Transportation Research Part B: Methodological* 41 (1):96-113. doi: 10.1016/j.trb.2006.03.004.

Yun, Meiping, Jiangyong Liu, and Xiaoguang Yang. 2011. Modeling on Mode Choice Behavior Based on Trip Chaining: A Case Study in Zhongshan City. Paper presented at the 11th International Conference of Chinese Transportation Professionals (ICCTP), USA.