EJTIR

Issue 18(1), 2018
pp. 36-59
ISSN: 1567-7141
http://tlo.tbm.tudelft.nl/ejtir

Effect of land use and survey design on trip underreporting in Montreal and Toronto's regional surveys

Chris Harding[1]
University of Toronto Transportation Research Institute (UTTRI), Canada.

Monika Nasterska[2]
University of Toronto Transportation Research Institute (UTTRI), Canada.

Leila Dianat[3]
University of Toronto Transportation Research Institute (UTTRI), Canada.

Eric J Miller[4]
University of Toronto Transportation Research Institute (UTTRI), Canada.

This paper contributes to the literature on travel survey methods by quantifying the relationship between land use, data collection protocol and trip under-reporting in regional travel surveys. While under-reporting more broadly is a recognized problem, the significant increase in under-reporting in denser, more urban-type environments identified here has never before been demonstrated or measured. Consequences of this land use-related bias for transportation planning and modelling are explored.

The work is carried out by comparing the results of two very similar household travel surveys conducted in 2011 and 2013, in Toronto and Montreal respectively. Using data on over 350,000 persons, a binary logit model for discretionary trip making is estimated and the effects of land use and data collection protocol on under-reporting are isolated. This is done by controlling for mobility tool access, household type and other key determinants of travel demand. Counter-intuitive effects for urban type environments found indicate the under-reporting effect is equivalent to a reduction in the pre-existing odds of reporting discretionary trip making in more urban environments of 19 to 29%. When combined with Toronto's data collection protocol effect, the range increases to 39 to 55%.

Results should be of use to transportation planning authorities wishing to make better use of the data collected in large surveys. Recognizing some of the flaws and biases in what is reported, these authorities can complement existing sources of data or modify their approaches to demand-based infrastructure provision to better account for the large number of, largely pedestrian, unreported trips.

Keywords: *data collection, reporting bias, land use, active transportation, discretionary trips.*

[1] 35 St George, Toronto, Canada, M5S 1A4. T: 1-647-963-6950 E: chris.harding@mail.utoronto.ca
[2] 35 St George, Toronto, Canada, M5S 1A4. T: 1-647-886-4044. E: m.nasterska@mail.utoronto.ca
[3] 35 St George, Toronto, Canada, M5S 1A4. T: 1-647-786-0866. E: dianat.leila@gmail.com
[4] 35 St George - Room 305A, Toronto, Canada, M5S 1A4. T: 1-416-978-4076. F: Fax: 1-416-978-6813 E: miller@ecf.utoronto.ca

1. Introduction

Travel survey methods are of considerable interest at the moment. Departments of transportation and transit planning agencies across the world are dealing with increasing cost and difficulty in collecting high-quality, representative travel data. At the same time, active transportation (AT) is also gaining visibility in the planning field as a result of concerns over equity (Currie, et al., 2010), emissions and air pollutants (Zahabi, et al., 2013) and health (Carmichael, et al., 2012). Governments around the world are actively trying to reduce their greenhouse gas emissions, promote healthful walking and cycling, and fight traffic congestion. One of the ways by which these goals can be achieved is by getting people out of their cars and onto transit, bicycles, or into their walking shoes.

Land use and transportation linkages have been investigated at the macro (Newman & Kenworthy, 1989) as well as neighbourhood scale (Shay & Khattak, 2007) for quite some time, and there is compelling evidence that living or working in dense, mixed-use and high connectivity environments can lead to reductions in kilometres travelled, as well as increase the share of trips made by active modes (Zhang, 2004). Both Ontario and Quebec (the Canadian provinces within which Toronto and Montreal are located) have ambitious emission-reduction targets (Ministry of the Environment and Climate Change, 2014; Ministère du Développement durable, Environnement et Lutte contre les changements climatiques, 2016), but in order to assess the potential of land use interventions in helping meet these emissions-reductions targets through changing travel patterns, it is important to have high-quality data on current behaviour and response to the built environment.

This knowledge is typically derived from the analysis of travel surveys. Both Toronto and Montreal have among the world's largest cross-sectional surveys. The Transportation Tomorrow Survey (TTS) has been conducted every 5 years since 1986 in the Toronto region, and the Enquête Origine Destination (EOD) has been conducted every 4 or 5 years in the Montreal region since 1970. In both these surveys, respondents provide demographic information and a complete single-day travel diary for every member of their household. One key difference with the TTS, however, is that simple home-based tours made for discretionary purposes on foot are not recorded (Data Management Group, 2013); this particularity will be described in greater detail in subsequent sections, it making comparison of TTS and EOD data particularly fruitful.

Having both these surveys, conducted in nearly identical ways, at our disposal allowed us to investigate the magnitude of under-reporting specific to each city, data collection effort, as well as more broadly with household telephone surveys. To design more effective land use policies and better plan investments in infrastructure, a better understanding of the factors that influence mode and destination choice is required. Most of the existing travel demand models are inelastic to built environment variables such as density, land use diversity and urban design. As a result, these models are likely to over-predict future auto trips and vehicle kilometres travelled (VKT) by distributing trips to destinations further afield. Therefore, effective multi-modal trip generation models are required to enable evaluation of various land use policies aimed at affecting travel patterns. This is a major challenge due to lack of data on active transportation demand. While it is known that under-reporting is a problem in travel surveys overall, and that under-reporting disproportionately occurs with short, walked and discretionary trips (Jin, et al., 2014), the effect of land use on under-reporting has not yet been estimated.

To offer a concrete example, Institute of Transportation Engineers (ITE) trip generation rates, which are widely used, are based on data collected in low-density areas with low shares of active mode use and a lack of walk/bike friendly infrastructure (Millard-Ball, 2015). One consequence of using these rates when planning for new developments that aren't in low-density areas is then the generation of 'phantom trips' and over-investment in road and parking infrastructure capacity for the automobile (Millard-Ball, 2015). This is both environmentally and economically wasteful.

There are other tangible policy implications to the lack of reliable AT data. The City of Toronto is in the process of developing a 10-year cycling plan, for example, in which one of the criteria for future infrastructure investment is current cycling demand (City of Toronto, 2015). If this demand is inaccurately assessed due to under-reporting of discretionary trips, infrastructure provision could be ineffectively allocated. Although AT investments tend to be more cost-effective than infrastructure projects for cars or transit, it can still be challenging for municipalities to secure the funding they require. The regional transportation plan for the Greater Toronto and Hamilton Area lists AT as a priority for instance, however, the regional planning agency has committed less than 5% of its funding to AT. AT projects in the region have been hindered by the lack of funding and support from decision-makers, and an overall prioritizing of other modes (Hess, et al., 2014). Without reliable or significant numbers of AT trips reported, it can be difficult to justify such investments. As a result, municipalities turn to smaller-scale studies or pilot projects, which require additional funding to be carried out (City of Toronto, 2015) and slow down the overall process of change.

The comparison of Toronto and Montreal offers an interesting case for analysis, as the outright non-recording of most discretionary walk trips in Toronto allows us to look for gaps in the data and relate these to the built environment. More specifically, by looking at discretionary trip-making at the individual level and relating this to land use, the method employed seeks to identify unexplained decreases in trip making and investigate whether these may be the under-reported walk trips referred to above. Discretionary trips are looked at because they are the type of trip most often made on foot (Driscoll, et al., 2013), as well as the type of trip whose frequency is most affected by accessibility (Ewing, et al., 1996). As a result, an anomaly related to this type of trip is an indication we may have isolated our problem.

Given the particularities related to the TTS, it is possible to contrast this extreme case with that of the EOD, where all walk trips are supposed to be reported and recorded, and look for land use effects on under-reporting that may be present at lower or more difficult to discern levels in other datasets.

The objectives of the paper are thus twofold. First, to assess the direction and magnitude of the relationship between land use and trip under-reporting using discretionary trip-making as a dependent variable. Second, to compare effect sizes for land use effects in two cities with differences in recording methods, but otherwise similar data, to isolate the effects of data collection protocol from that of land use.

Comparing summary statistics and logit model estimation results, controlling for demographics, household type, mobility tool ownership and distance from the core, we find evidence of a statistically significant relationship between land use and trip under-reporting for discretionary purposes in Toronto, with a smaller magnitude effect in Montreal. This is consistent with our hypothesis. After model estimates are presented, the paper concludes with a discussion of potential ramifications on the planning and provision of infrastructure.

2. Literature review

2.1 Trip Under-reporting

Trip under-reporting in travel surveys is well-documented (Wolf, et al., 2013). Previous authors have quantified under-reporting using passive GPS or other location traces and found significant differences when compared to the responses provided in household and personal travel surveys (Shen & Stopher, 2014; Wolf, 2000). GPS traces have been used in assessments of under-reporting, as the "information derived through the GPS-assisted prompted recall method is considered the closest reflection of the actual travel activity behaviour, or the 'ground truth'" (Jin, et al., 2014, p. 3403). While GPS recorded trips are subject to their own issues, whether signal loss, spatial accuracy or processing algorithm failures (Wolf, 2000), they remain sources of data not affected by respondent memory, fatigue, or personal perspective or understanding on what is to be reported.

More specifically, authors comparing GPS to survey data have found correlations between under-reporting and short (Jin, et al., 2014) or AT trips (Chen, et al., 2014), especially when these trips were undertaken for discretionary purposes. This might reflect biases in respondents' minds or an explicit encouragement on the part of interviewers to pay closer attention to those trips which are made using motorized modes. While there has been consistent under-reporting across all household types, the prevalence of discretionary trip-making, and, by association, of trip under-reporting, has been found to be much higher among retired individuals, as well as with part-time workers and non-working adults (Jin, et al., 2014). We could not find, however, any work that directly related residential location to under-reporting.

There is also existing evidence of trip under-reporting specific to the City of Toronto. TTS data indicate that only 8.2% of all trips within the city are made on foot or by bicycle (Data Management Group, 2013), while in a Harris/Decima survey, 31% of Torontonians reported that walking was their main mode of transportation to work, school, shopping or leisure activities (2008).

Analysis of recent data collected using passive smartphone applications actually found that over half of all trips made by Toronto residents in the sample were made on foot (Harding, et al., 2015); while the population that made up this sample were primarily downtown residents recruited in coffee shops, the difference remains stark. Finally, in a web survey of Toronto-area post-secondary students with 15,226 respondents, 17.2% of these respondents indicated having omitted information on walk or bike trips immediately after completing a single-day travel diary, in total accounting for 0.41 omitted trips per respondent (StudentMove, 2015).

2.2 *Trip generation rates, mode choice and land use*

There would appear to be a general consensus in the literature that increased accessibility (as seen in more central, denser or more mixed environments) either leads to increased overall trip generation rates (Ewing, et al., 1996) or has negligible effects after controlling for demographics and vehicle ownership (Williams, 1989). It does not, however, lead to decreases in trip rates.

This is not merely a theoretical claim, but has been measured using travel survey data. Referring to work by colleagues in the United Kingdom, Germany, Austria and Sweden, Williams stated that "after controlling for household size and automobile ownership, no systematic relationship between household location and trip-making frequency was confirmed" (Williams, 1989, p. 374). Looking to more recent data, summary statistics extracted from the 2010 Swiss microcensus on mobility and transportation (nearly 60,000 households interviewed) indicate trip rates are slightly higher in the core of cities (3.5 trips per person, per day) than in outlying areas (3.4 trips per person, per day) (Office Fédéral de la Statistique, 2012, p. 41). In analysis of the same data, trip rates, as well as number of unique locations visited, were both found to be equal or greater in almost every inner urban area, when compared to the same region's outer urban area (Harding, et al., 2013).

Closer to our study area, comparing trip rates extracted from GPS-validated time use diaries in Halifax, Millard and Spinney (2011) concluded that "inner-city residents make most trips, but have trips of shortest duration, and spend least time in travel. […] There are significantly fewer trips in the [outer commuting belt], which [the authors] attribute to lack of need, lack of opportunities, and adjustments in discretionary behaviour" (2011, p. 51). In other words, inner-city residents respond rationally to increased opportunities, and decreased trip distances and durations, by making more discretionary trips. What's more, unlike self-reported OD datasets such as the TTS or EOD, the data analysed by Millard and Spinney was comprised of time-use diaries validated with GPS, and as such did not suffer from under-reporting in the same way.

At its core, the relationship between land use and trip generation is the following: ceteris paribus, if accessibility increases, the cost of travel decreases, which in turn means the number of trips made should increase. Depending on the elasticity of the demand, this effect may be small or large. Crucial to the analysis carried out in this paper, "if accessibility has any effect on trip rates, [Ewing et al. state,] it will be most apparent for "discretionary" trips, that is, trips that can be postponed or not made at all" (1996, p. 3).

Boarnet et al. discussed the effect of land use on mode choice and found that residents living near a vibrant centre took more than double the amount of walking trips as compared with similar households along a nearby corridor (Boarnet, et al., 2011). The same regressions that show residents of walk-amenable centres making more trips on foot did not, however, show these residents making fewer car trips (Boarnet, et al., 2011). This would indicate both that walk trips are not always substitutes for car trips, and that increased accessibility, ceteris paribus, leads to an increase in trip-making. This is logical given that the cost of making a trip decreases with an increase in absolute accessibility. With respect to retail densities specifically, Boarnet et al. found that "persons living in neighbourhoods with more business establishments per acre take more daily walking trips, have higher trip capture rates and are more likely to travel to their centre or corridor by walking and less likely to travel to their centre/corridor by car" (Boarnet, et al., 2011, p. 141).

Ewing and Cervero, in their 2001 meta-analysis, found that availability of amenities or attractive destinations within walk/bike distance had been found to be the most important predictor of reductions in vehicle trips and VKT. "Despite higher rates of midday trip making, [for example,] downtown environments generated fewer vehicle trips and less VMT per employee than suburban office parks because of the preponderance of walking trips" (2001, p. 104). Ewing and Cervero also found that while non-work trip generation was most influenced by socio-demographic characteristics and VKT by land use attributes, mode choice was a function of both (2001). Boarnet et al. also showed that land use attributes were as important as socio-economic characteristics and affect both trip frequency and VKT (Boarnet, et al., 2004).

While the exact effect sizes and operationalization of built environment measures or unit of analysis can be debated, one thing agreed upon in the literature is that at some level, there exist relationships between land use, trip generation and mode choice.

Touching briefly on residential location choice and its relation to mode choice, work by Badoe and Miller, as well as Currie et al. has shown that certain households purposefully choose to locate near transit to avoid having to own a vehicle, thereby increasing the share of walk, bike and transit trips that can be taken to satisfy their needs (Badoe & Miller, 2000; Currie, et al., 2010). Land use has also been found to affect choice set formation (Zhang, 2004), removing certain modes from contention in the traveler's mind if the relative utility is below a threshold of consideration. This blurs the direction of causality within transportation and land use research, but should have no effect on the actual reporting of trips.

2.3 Modelling implications

In transport models, trip generation rates are often taken from present behaviour, and classified by demographic or geographic characteristics. They do not account for behavioural changes due to, for example, a change in level of service (Martens & Hurvitz, 2009), or, as is relevant to this paper, changing land use patterns. Subsequent model iterations thus result in increased investment for infrastructure reported to be highly utilized. Through this investment, the favoured mode may become even more attractive, creating a self-fulfilling prophecy. As a perverse result of this, Martens and Hurvitz wrote, there is a real potential for "the consecutive application of the four-step model over a number of years, and successive investments in transport infrastructure consistent with the model results, [to] widen existing gaps between high-mobile and low-mobile groups" (2009, p. 1). The impacts of successive waves of demand-based modelling are poorly understood and there may very well be a self-fulling prophesy aspect to them if a systematic bias exists in what data are properly recorded. At its most obvious, road infrastructure may be built to meet vehicle demand that may not otherwise have materialized had investments not been made (Millard-Ball, 2015), while proper pedestrian and bicycle infrastructure that would allow for the sustainable movement of persons may not be built, limiting the potential for adoption of AT.

"Despite recent increased interest in planning for walking, current forecasting tools —namely regional travel demand models— incompletely represent pedestrian behaviours" (Clifton, et al., 2015, p. 1). Without assigning a specific 'portion' of influence for preferences or land use, one thing that has been shown is that trip distances, and walking and cycling convenience and opportunities are two crucial elements forming peoples' destination choice in dense areas (Clifton, 2013). Changes in urban design shift peoples' destination choices most for discretionary trips (ibid), so having good data on the distances actually travelled in conjunction with good land use is critical.

In order to address the problem of under-reporting, past authors have relied on application of trip correction factors (Wolf, 2000; Wolf, et al., 2013), "however, this is a simplified method that does not reflect the multifaceted dimensions in misreporting (…). In addition, this approach does not really provide a solution to address trip misreporting directly in travel demand models" (Jin, et al., 2014).

3. Data and methods of analysis

To investigate our hypothesis of a relationship between land use and trip under-reporting, we make use of two travel surveys: Montreal's regional origin-destination survey (EOD) (Agence Métropolitaine de Transport, 2013) and Toronto's Transportation Tomorrow Survey (TTS) (Data Management Group, 2013).

There are many similarities between the two surveys' data collection protocols, which allows for ease of comparison between the two datasets and eliminates many sources of bias when comparing results. Both include a (roughly) 5% sample of households in their study area. The surveys are conducted over the phone, primarily through calling land lines, although Montreal did recently begin piloting cell phone-only random digit dialing. Both surveys attempt to replicate a 'typical fall day' and collect information on respondent demographics, as well as mobility tool ownership and full-day travel diaries.

In both the EOD and TTS, respondents are asked to provide full travel diaries for the previous day for all members of the household. The person who begins the interview typically answers for all household members, which may lead to some proxy bias if this individual is not aware of, or does not report the travel of other household members accurately – this aspect is controlled for later on in the analysis stage. There are no monetary incentives for participation and the surveys are voluntary. In Toronto, only trips for persons over the age of 11 are recorded, whereas in Montreal the age cut-off is four. To ensure our model estimates are comparable, Montreal respondents between the ages of 4 and 10 are removed.

Another difference is that the TTS has fewer trip purpose categories than the EOD: work, school, facilitate passenger, marketing (shopping) and other. 'Marketing' and 'other' trips are generally considered to be discretionary, although in reality, 'other' trips could involve less strictly-discretionary purposes, such as visits to the doctor. Montreal's survey includes more categories, such as health, visiting friends/family, leisure, etc. To ensure modelling results are comparable, trip purpose categories in Montreal have been aggregated to reproduce TTS categories.

The most important difference between the two survey methodologies, for our purposes, is the approach to the recording of walk trips. In the TTS, walk trips are only recorded if they are for work or school purposes, unless they are part of a complex chain in which other modes are used, or work or school appear later in the home-based trip chain. This is a crucial difference with regard to the analysis carried out here, as the hypothesis we put forth is that this non-recording of walk trips not only affects the overall quality of the data, but also specifically leads to an under-reporting of travel in denser, more 'urban' environments. This 'urban' under-reporting results from walk trips essentially replacing motorized trips for discretionary purposes in these environments.

While one could argue that this difference in discretionary walk trip reporting would preclude use of TTS data for the purposes of our analysis, this is not the case. Because discretionary trips by all other modes are recorded in the TTS, comparison of recorded travel behaviour in dense environments in EOD and TTS, we posit, serves to quantify the trips not recorded that happen to be on foot.

For the purposes of this paper, the most recent surveys were selected for each city: 2013 for Montreal, and 2011 for Toronto (which was actually conducted in both 2011 and 2012). Respondents living within the Montreal and Toronto census metropolitan areas (CMA) were chosen for model estimation to ensure there was as little effect as possible related to commuting and distribution of activities within the region (CMAs being defined by commuting flows). We also removed households living in Toronto's planning districts (PD) and Montreal's municipal sectors (SM), where land use information was missing for more than 10% of all person-days of travel.

In 2011, the TTS collected information on 858,848 trips made by 410,404 persons living in 159,157 households. In the 2013 EOD, information was collected on 410,741 trips made by 188,746 persons living in 78,731 households. After filtering for persons living within their respective CMAs, and then again for persons living in Planning Districts (Toronto) and Municipal Sectors (Montreal) where fewer than 10% of households made trips for which land use information was not available, there remain 217,815 (Toronto) and 138,131 (Montreal) persons in our modelling sample. The percent of respondents over the age of 11 not reporting any trips on the day of the survey in Toronto ranges from 19% in more suburban areas to 31% in more urban environments. In Montreal, a much smaller spread is found, with 18% reporting no travel in more suburban areas and 19% at the core – see Table 1.

3.1 Land use

The data used to characterize land use are derived from two main sources: Statistics Canada (Statistics Canada, 2011) and transit agency-provided General Transit Feed Specification (GTFS). These sources provide us with population and employment densities, as well as retail density and, finally, transit service density. For population, dissemination area population figures derived from the 2011 census are used. For retail and overall employment NAICS-classified job counts at the census tract level (larger than dissemination areas) are used. This allows us to calculate general employment density, as well as retail density.

Retail density is used in our model because shopping trips are classified as discretionary. Other data sources were considered, but an inspection of both DMTI enhanced points of interest (EPOI) and land use classified vectors revealed considerable gaps and problematic coding. EPOI data contained far too many points classified as 'unknown'. While this may not be a problem if they are evenly distributed, there were unexplainable geographic concentrations that precluded our using that data. As for land-use classified vector data, 'mixed-use' was not part of the coding scheme and our research team noticed important commercial arteries labeled as residential. Because of these concerns, we decided to use retail employment density, which other authors have successfully done when relating the built environment to travel behaviour (Clifton, et al., 2015).

For transit, GTFS data are used to derive the sum of scheduled stops per hectare on a weekday afternoon between peaks – 9AM to 3PM. This time period was chosen as it is the time of the day during which the highest concentration of discretionary trips occurs.

3.2 Indicator generation

Indicators were developed at traffic analysis zone (TAZ) and census tract (CT) levels to relate land use to travel behaviour. In the case of Toronto, TTS trip destinations, as well as home locations, were only available at the TAZ level. Thus, to relate land use to behaviour, we developed indicators at this scale. In Montreal, EOD contained exact coordinates for both home and destination locations. Given that the land use data are defined at the census tract level, however, we chose to use these as analogs to Toronto's TAZs. This ensures greater compatibility in results.

To reduce the effect of the modifiable area unit problem (MAUP), which occurs when there is spatial unit variation in size, shape and orientation (Openshaw, 1984), we employed a three-step process. First, land use indicators were generated at a grid cell level, after which these grid cells were intersected with TAZ or CT shapefiles. Finally, weighted averages of cell values were used to calculate TAZ and CT values.

As opposed to using square cells, as have been employed by co-authors Harding and Miller in the past, hexagonal cells with 500-meter sides were employed. This approach was used since hexagons "have the lowest edge-effects of any shape that can completely cover a region" (Wilkinson Davis, et al., 2015, p. 4). 500 meters in turn was chosen given that we are looking at discretionary activities with an implied focus on walking trips: the resulting cells are approximately 15 minutes on foot from centre to edge, a reasonable walking distance for shopping and other purposes. There is no one correct size in our view, however, and each type of analysis or regional context may have its own 'ideal'. The 90th percentile of walking distance or some other sample-based value could have been employed, but in the context of a study area so large and diverse, this may have been a false-'objective' solution. What constitutes a 'neighbourhood' and what distance one is willing to walk is determined by land use at a given location, the predominant modes used by residents and the barriers that exist (whether natural, physical or social).

The hexagonal cells were clipped to account for study area boundaries and a weighted average for cell values was calculated using the net area of each cell and surrounding cells. Calculating an average simplifies the process of determining an exact cell size, while also reducing potential edge effects and avoiding any peaks or sudden spikes in cell values. All TAZ and CT-level land use values, later associated with individuals and households, are calculated based on hexagonal cell values. Distance from the core is a final spatial variable calculated using the centroid of home TAZs in Toronto and actual home location in Montreal. Union station in Toronto and Station Centrale in Montreal were chosen to represent the CBD for each.

Land use values for each TAZ and CT were normalized and run through a k-means cluster analysis in STATA to find neighbourhood types, or clusters of similar combinations of land use. Different numbers of clusters were generated (from 2 to 10); 4 and 6 were selected for additional testing in Toronto, while 5 was chosen for Montreal. In all cases, these were found to be local maxima when using the Calinski-Harabasz test. This test is a measure of the efficiency of the clusters (neighbourhood types in this context) at maximizing inter-group variation while minimizing intra-group variation (Dimitriadou, et al., 2002). The 6-neighbourhood typology performed better when modelling our dependent variable in Toronto, so all information presented below will relate to this classification, with 5 being the number in Montreal. High value clusters refer to higher densities and will be referred to using higher numbers (5 and 6) or the terms 'urban' and 'core' (see Figure 1).

Frank and Pivo indirectly demonstrated the benefit of using clustered land use indicators in a linear model (1994). Their work clearly showed that the effect of land use on mode choice is non-linearly related to density. While one could look for significant levels for each explanatory land use variable in the dataset, clustering carries an additional benefit over such a variable-by-variable categorization: interactions and combined effects are better accounted for using clusters or neighbourhood types. A particularly high population density might be insufficient to cause a shift away from single occupancy vehicle travel (SOV), for example, but as researchers have shown, a simultaneous or combined improvement in density, diversity and accessibility could lead to noticeable reductions in the number of vehicle kilometres travelled (Miranda-Moreno, et al., 2011).

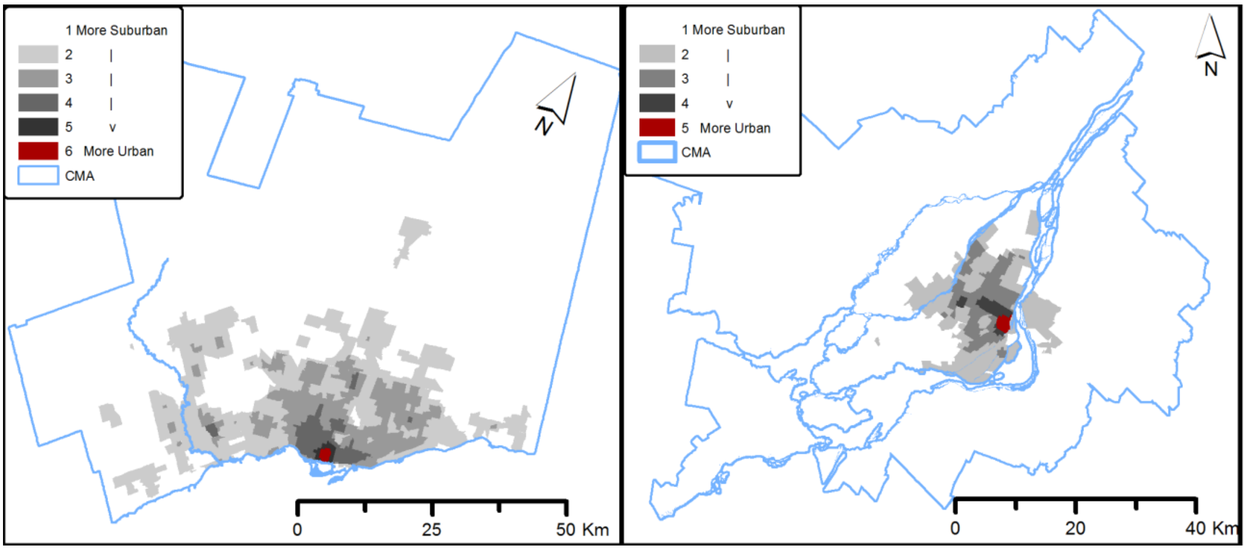

Figure 1. Neighbourhood types in Toronto and Montreal.

3.3 Methods

As we posited trips are under-reported more frequently in denser, more mixed environments, these being the locations where walk trips and discretionary trips are most common, we looked for signs of this first by comparing discretionary trip rates per person across the region by household and neighbourhood type, then by modelling discretionary trip-making as a binary dependent variable at the person-level. Controlling for household type, demographics, mobility tool ownership and distance from the core, logit model coefficient estimates are used to investigate whether there is a relationship between land use characteristics and trip under-reporting. More specifically, by identifying counter-intuitive coefficient estimates for land use, we set out to investigate whether these might be indicative of either response-behaviour motivated or methodological under-reporting. While in TTS data there is an expectation of decreased trip rates as a result of the data collection protocol, *differences* in decreased trip rates in urban environments between the TTS and EOD can be looked at to disentangle the effect of land use and behavioural non-reporting from that of data collection protocol exclusion for discretionary trips.

Our hypothesis was that home location land use variables would be estimated with counter-intuitive signs. While one would expect denser, more mixed environments to be related with more discretionary trip making, ceteris paribus, we posited that walk trip under-reporting would lead to negative signs for variables related to density and accessibility at the home location.

The assumption underlying this approach to quantification of under-reporting, detailed in Section 2.2 (Trip generation rates, mode choice and land use), is that there is no rational explanation why increases in density or land use mix should lead to fewer discretionary trips being made, controlling for household type and vehicle access. More specifically, as the literature would suggest that a decrease in the cost of travel accompanied by an increase in the wealth of opportunities should lead to an increase in travel, finding lower discretionary trip generation rates or likelihood of discretionary trip-making for households living in denser environments we take as a sign of trip under-reporting related to land use.

In Montreal (EOD), the link is that denser, more urban land use leads to increased discretionary walk trips that are merely more often forgotten or omitted, leading to a decreased discretionary trip rate, all other variables controlled for. In Toronto (TTS), urban land use would also lead to an increase in walk trips, but given that they remain systematically unrecorded, whether or not remembered, the decrease in the overall trip rate will be more pronounced.

Discretionary trip making was chosen as our dependent variable as one's choice in where to buy a litre of milk, for example, is relatively unconstrained (a discretionary shopping trip). Where one works or goes to university, however, is restricted to places where one's skills are needed or where particular opportunities exist. This means discretionary trips are much more likely to be affected by land use than anchored mandatory trips.

As Krizek described, "higher [neighbourhood accessibility] households travel shorter distances for maintenance-type errands (personal, appointment, and shopping) (…), [however a large portion of households'] "maintenance travel is still pursued outside the neighbourhood; a mere 20 percent of their simple maintenance tours are within 3.2 km (2.0 mi) of their home (in contrast to a mere 1.7 percent for households in the lower half of neighbourhood accessibility)." (Krizek, 2003, p. 406). We did not expect our results to show all discretionary trips by urban residents were missing because they were all made on foot, but even a fraction of discretionary travel unreported could be enough to help us hone in on statistically significant land use effects.

4. Results and discussion

4.1 Summary statistics[5]

Table 1. Trip rates and land use, by neighbourhood type (filtered datasets for both cities, as described in the Data section).

	Neighbour-hood Type		Trips/ Person	% 0 Trips Recorded	Discr Trips /Person	Walk Trips /Person	Transit Density /ha	Persons /ha	Jobs/ ha	Retail Jobs/ ha	Dist from Core (km)	Count
Toronto	More	1	2.51	19%	0.60	0.07	0	13	4	0	38	33,530
	Suburban	2	2.40	21%	0.54	0.10	2	33	9	1	24	89,550
	↓	3	2.15	26%	0.53	0.11	5	50	16	2	15	60,745
		4	2.12	26%	0.53	0.17	10	81	38	4	7	26,913
	More	5	1.78	32%	0.41	0.38	14	129	214	13	2	4,170
	Urban	6	1.80	31%	0.42	0.49	19	123	480	24	1	2,907
Montreal	More	1	2.30	18%	0.49	0.10	1	18	6	1	23	70,310
	Suburban	2	2.26	21%	0.54	0.22	6	49	16	2	10	33,498
	↓	3	2.36	19%	0.56	0.42	9	82	30	3	7	27,485
	More	4	2.53	17%	0.66	0.81	11	100	95	8	3	5,723
	Urban	5	2.43	19%	0.69	1.00	14	70	351	16	1	1,115

* For transit density, as explained in section 4.1, the values shown are the sum of scheduled stops per hectare, on a weekday afternoon between peaks

As one can see in both Table 1 and Figure 1, there is a correlation between distance from the core and our land use variables of interest. This is inevitable in regions such as Toronto and Montreal,

[5] Note: For methodological reasons, expansion factors were considered to be uniform across all respondents. If there further exists a correlation between sampling bias and trip under-reporting within a given household type, this aspect was not corrected for.

as they may have secondary employment centres, but still have strong cores. It is because of these strong cores, and the history of analysis of monocentric cities that we in fact chose to focus on TAZs and CTs within each region's census metropolitan area (CMA). CMAs are analogous to metropolitan statistical areas (MSAs) in the US, and their boundaries are defined by commuting flow. These CMAs, then, enable us to observe the behaviour of households whose travel patterns revolve around a common core. There is significant suburb-to-suburb commuting in Toronto, but given our focus on discretionary trips and the additional filtering employed, we believe our suite of spatial units is appropriately defined.

Looking more closely at Table 1, we see that an unexpected pattern of decreases in trip rates for all trips, as well as a more pronounced decrease in discretionary trip rates, can be seen in Toronto when moving from more suburban clusters to more urban ones. From looking only at the Toronto data, it is not clear that the TTS survey methodology results in trip under-reporting; however, when the data are compared with Montreal, as well as contrasted to the Halifax (Millward & Spinney, 2011) and Swiss (Office Fédéral de la Statistique, 2012) examples described in Section 2.2, it is clear that there is a significant overall under-reporting. Seeing polar opposite patterns in overall trip rates, and the significant increase in walk trip rates in Montreal as one gets closer to the core, it is clear that the absence of non-work, non-school walk trips in Toronto has a significant impact on overall reported trip rates. In Montreal, discretionary trip rates increase for residents living in more 'urban-type' environments, while the percent of individuals reporting no-travel days stays constant. In Toronto, the percent of non-travelling individuals increases by over 50%.

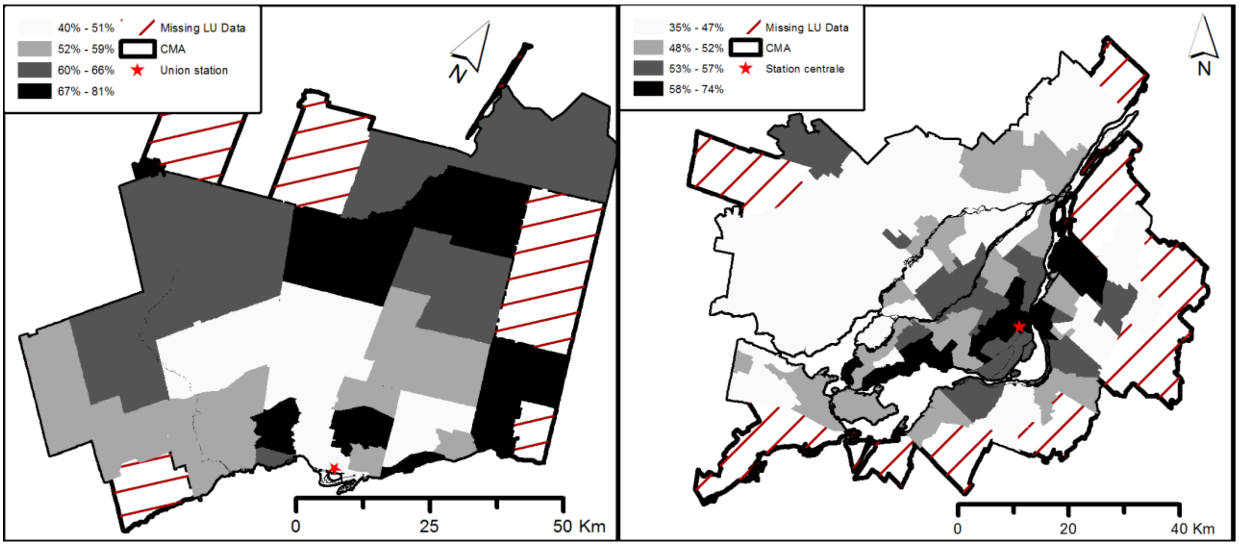

Figure 2. Study area discretionary trip rates.

Looking at the travel patterns of residents across the region, we see certain spatial trends. To begin, in Figure 2, the number of discretionary trips per person per PD and SM decrease near the core in Toronto and increase in Montreal. Either Toronto's downtown has a tendency to attract more homebodies than Montreal's, or there is a difference in reporting patterns. Given the similar dynamics observed in both cities, with high rents paid to access the amenities at the core, it seems much more likely that trips in Toronto are simply not being recorded.

Table 2. Neighbourhood type composition

	Neighbour-hood Type		2 Adults, < 2 incomes	Dual Income, No Kids	2+ Adults w Kid(s)	Single Parent	Retiree	Retirees	Single Adult	Student(s) Only	Other
Toronto	More Suburban	1	7.1%	5.6%	44.8%	2.0%	2.4%	6.5%	2.3%	0.2%	29.1%
		2	5.5%	4.5%	42.7%	2.3%	3.1%	6.7%	2.6%	0.4%	32.2%
	↓	3	5.9%	4.1%	34.5%	2.8%	5.7%	9.2%	4.4%	0.9%	32.7%
		4	9.0%	6.8%	29.4%	2.6%	5.9%	7.0%	8.6%	1.0%	29.6%
	More Urban	5	12.3%	9.7%	18.8%	2.5%	8.4%	5.6%	17.0%	2.7%	23.2%
		6	14.5%	12.2%	13.2%	2.1%	7.5%	8.5%	19.0%	2.4%	20.4%
Montreal	More Suburban	1	8.4%	8.2%	37.1%	2.4%	3.6%	8.7%	4.0%	0.2%	27.5%
		2	8.2%	7.5%	29.7%	2.8%	7.0%	9.9%	7.2%	0.8%	26.9%
	↓	3	8.2%	7.8%	30.8%	3.4%	7.6%	7.2%	10.0%	1.9%	23.2%
	More Urban	4	10.2%	13.7%	25.0%	3.0%	7.4%	4.1%	16.0%	2.9%	17.7%
		5	10.2%	12.0%	15.2%	1.0%	10.4%	4.8%	14.9%	7.3%	24.1%

Aggregate statistics at the PD or SM scale, while informative, can be misleading, however. This is because different areas are occupied by different types of households, as shown in Table 2. As one can see from this table, households with children are more prevalent in lower density, more suburban areas, while the proportions of single-person and other household types increase in denser, more urban-type environments. As there is a correlation between household type and travel patterns, it is important to look more closely at the trip making of these different household-type residents in conjunction with neighbourhood type.

Figure 3 shows how stark these spatial differences in household type representation can be using the example of single-person households.

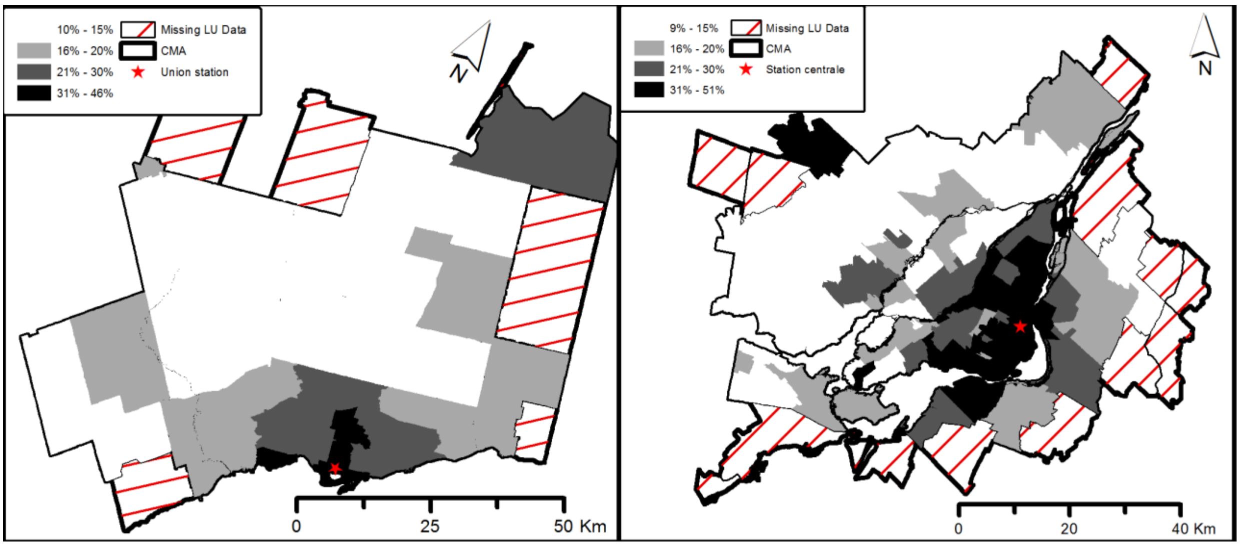

Figure 3. Proportion Single Person Households

As one would expect, there is a significant concentration of singles in the core. Because of this uneven distribution, we cannot draw conclusions on under-reporting of walk or discretionary trips based purely on distance to the core, land use or household type, but must employ more sophisticated methods in order to isolate the effect of each variable.

Figures 4-7 combine these two ideas of representation and discretionary trip making by showing the number of discretionary trips per person for different household types and in different PDs and SMs. Toronto is shown on the left, Montreal on the right. The scale is different for Toronto and Montreal, but the legend remains the same.

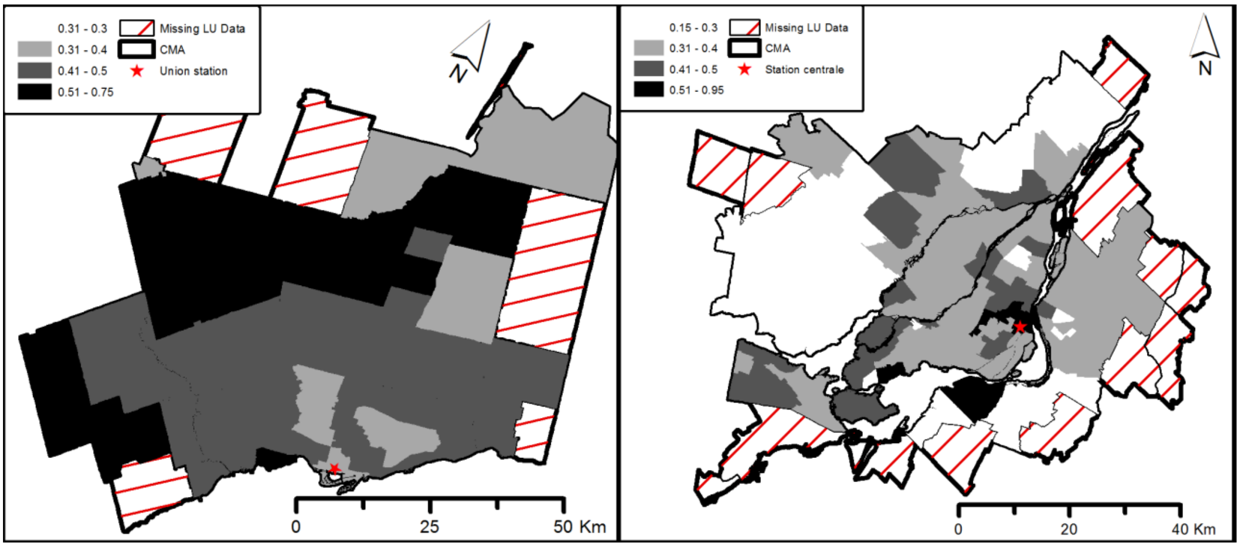

Figure 4. Discretionary trips per person for 'Dual Income No Kids' households

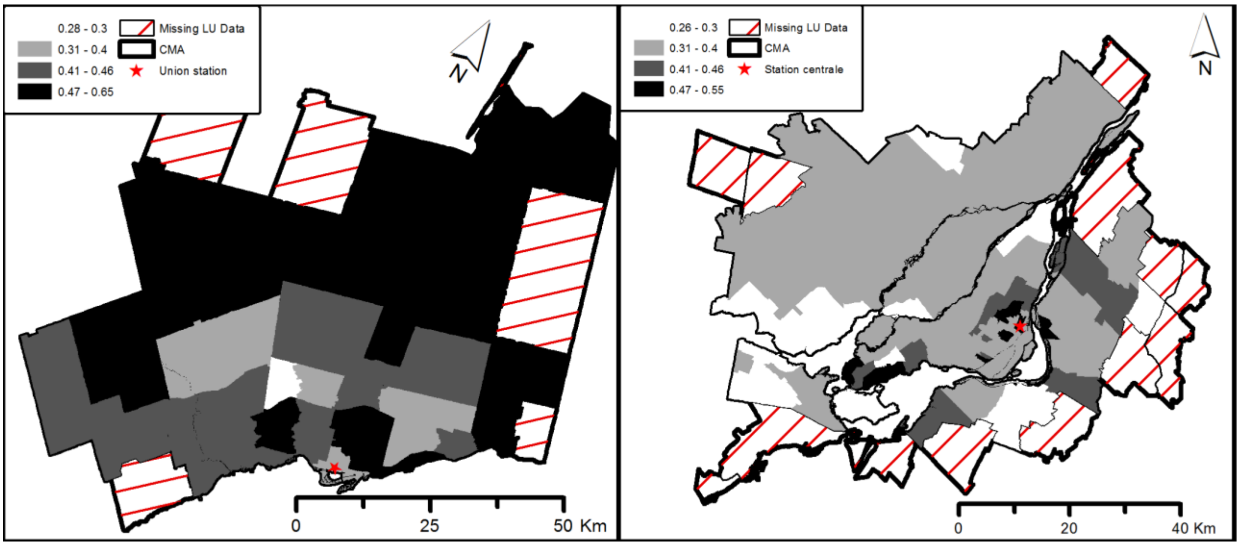

Figure 5. Discretionary trips per person for '2+ Adults with Kid(s)' households

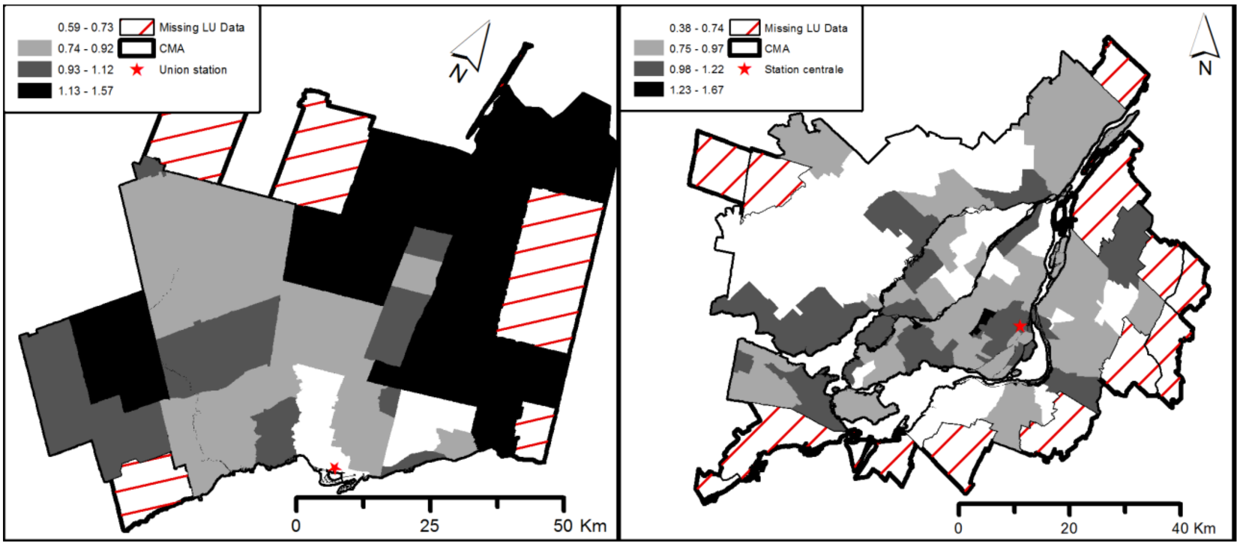

Figure 6. Discretionary trips per person for 'Retiree'' households

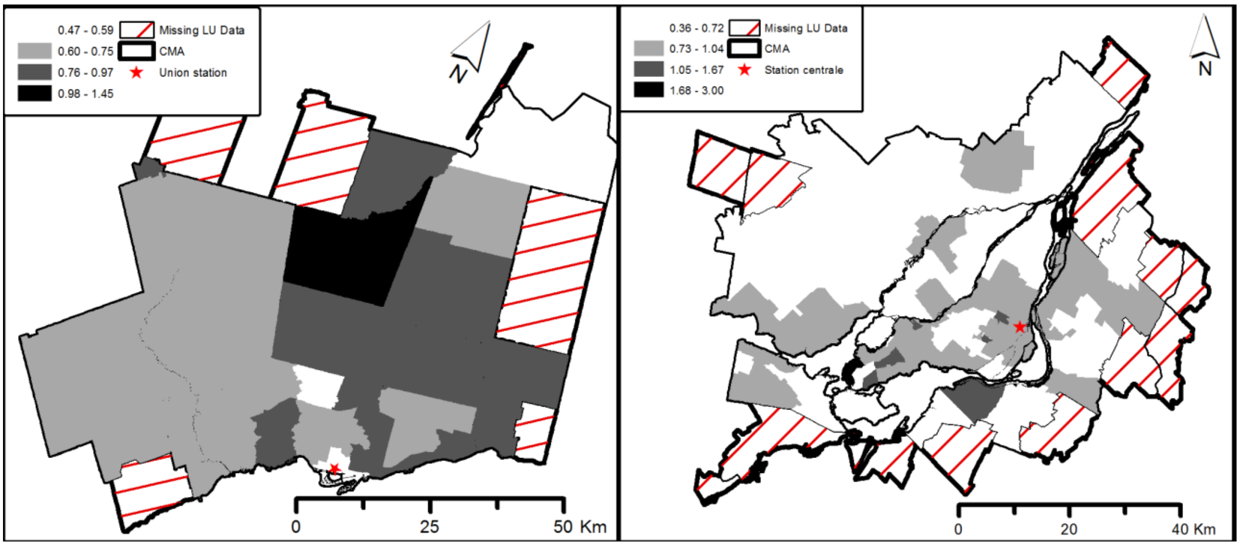

Figure 7. Discretionary trips per person for 'Single Person'' households

As can be seen by the many pale grey and white cells near Toronto's core (near Union Station, indicated with a red star), there would appear to be a significant under-reporting there that becomes more obvious when looking only at the travel of persons one household type at a time. Certain types of households (notably retired persons or part-time workers) are known to make considerably more discretionary trips than others (Jin, et al., 2014), but are not evenly distributed across the region.

These differences in reported discretionary trip rates are most quickly grasped in map form, but to get at a further explanatory factor with particular relevance in Toronto, Table 3 summarizes these trip rates by neighbourhood and household type, as well as by automobile access. Certain sections highlighted.

Table 3. Discretionary trips/person, by neighbourhood type and auto access

	2 Adults, < 2 incomes		Dual Income, No Kids		2+ Adults w Kid(s)		Single Parent		Retiree		Retirees		Single Adult		Student(s) Only		Other		All	
	TO	Mtl	TO	Mtl	TO	Mtl	TO	Mtl	TO	Mtl	TO	Mtl	TO	Mtl	TO	Mtl	TO	Mtl	TO	Mtl
No Car	0.42	0.55	0.26	0.47	0.23	0.42	0.36	0.45	0.43	0.66	0.34	0.53	0.48	0.75	0.35	0.49	0.30	0.48	0.35	0.57
ExUrban	0.31	0.38	0.67	0.26	0.32	0.32	0.09	0.31	0.37	0.44	0.32	0.48	0.30	0.58	0.00	0.16	0.32	0.35	0.33	0.43
Sprawl	0.39	0.53	0.19	0.34	0.21	0.33	0.30	0.44	0.41	0.61	0.35	0.47	0.38	0.63	0.28	0.44	0.28	0.48	0.32	0.52
Suburban	0.39	0.56	0.20	0.41	0.21	0.43	0.40	0.50	0.42	0.76	0.32	0.57	0.51	0.77	0.25	0.50	0.29	0.49	0.34	0.60
Semi-urban	0.53	0.65	0.37	0.66	0.28	0.59	0.42	0.38	0.46	0.85	0.37	0.81	0.54	0.95	0.58	0.59	0.34	0.60	0.41	0.73
Urban	0.30	0.50	0.17	0.75	0.23	0.30	0.39	0.71	0.45	0.98	0.32	0.63	0.43	0.90	0.36	0.50	0.27	0.54	0.34	0.68
Core	0.29		0.16		0.22		0.20		0.43		0.46		0.40		0.24		0.32		0.34	
Car Access	0.70	0.70	0.45	0.37	0.44	0.35	0.64	0.53	1.24	1.09	1.04	0.89	0.82	0.77	0.66	0.50	0.53	0.49	0.56	0.52
ExUrban	0.74	0.70	0.47	0.34	0.50	0.34	0.67	0.49	1.35	1.05	1.11	0.89	0.80	0.73	0.53	0.47	0.59	0.47	0.60	0.49
Sprawl	0.73	0.68	0.47	0.38	0.42	0.35	0.66	0.53	1.32	1.14	1.12	0.91	0.87	0.77	0.66	0.49	0.52	0.51	0.56	0.54
Suburban	0.69	0.75	0.41	0.43	0.41	0.40	0.57	0.59	1.16	1.08	0.99	0.87	0.84	0.81	0.62	0.55	0.52	0.51	0.56	0.55
Semi-urban	0.64	0.67	0.43	0.48	0.45	0.43	0.75	0.73	1.16	1.27	0.87	0.96	0.81	0.90	0.73	0.27	0.50	0.65	0.56	0.61
Urban	0.48	0.90	0.39	0.50	0.30	0.48	0.37	0.75	0.99	1.22	0.60	1.03	0.59	0.85	0.71	0.67	0.49	0.69	0.47	0.70
Core	0.37		0.38		0.22		0.62		0.75		0.82		0.59		0.83		0.56		0.48	
All	0.67	0.69	0.43	0.38	0.43	0.36	0.58	0.51	0.84	0.88	0.95	0.86	0.70	0.76	0.53	0.49	0.51	0.49	0.54	0.52

As can be seen in Table 3, individuals living in 'No Car' households in more urban environments in Toronto have dramatically lower reported discretionary trip rates than their car-owning counterparts. This is in stark contrast to Montreal, where reported discretionary trip rates are much more uniform and where much less pronounced differences exist by neighbourhood type. In the cells highlighted in red, we can see that discretionary trip rates for carless households living in more urban-type environments are more than twice as high in Montreal, but then looking to the cells highlighted in yellow (lower portion of the table), we see that discretionary trip rates for households with access to a car in more suburban type environments are actually higher in Toronto in most cases. These particular cells, while not the only ones of interest, are highlighted to clearly show that the difference in reported rates of discretionary trip making is not due to a higher overall discretionary trip rate in one city or the other, but is much more likely due to the particular way in which walk trips are treated in the TTS.

Table 4. Discretionary Trip Rates by Neighbourhood type and Mode of Transportation

	Neighbourhood Type		Walk	Bike	Transit	Car
Toronto	More Suburban	*1*	0.00	0.00	0.01	0.58
		2	0.00	0.00	0.03	0.51
	↓	*3*	0.01	0.00	0.07	0.45
		4	0.02	0.02	0.11	0.37
		5	0.03	0.02	0.14	0.22
	More Urban	*6*	0.03	0.02	0.13	0.23
	Average		0.01	0.01	0.05	0.48
Montreal	More Suburban	*1*	0.02	0.00	0.02	0.44
		2	0.06	0.01	0.07	0.39
	↓	*3*	0.13	0.02	0.12	0.30
		4	0.28	0.05	0.12	0.20
	More Urban	*5*	0.31	0.01	0.15	0.22
	Average		0.07	0.01	0.06	0.39

Table 4 presents discretionary trip rates per person, per mode of transportation used to make said discretionary trips and clearly demonstrates a definite relationship between land use, preferred mode of transportation and reported discretionary trip rates. Comparing the number of discretionary walk trips per person in more urban environments in Toronto (0.03 trips per person) to Montreal (0.31 trips per person), it is clear that the systematic non-reporting of walk trips is non-trivial. If we instead compare car trips per person in the suburbs, we see that higher discretionary trip rates are found, and reported, in Toronto. The difference is 0.58 to 0.44 discretionary trips per person per day.

Finally, Figure 8 shows the percent of individuals within each PD or SM making trips for which land use information is not available. Most of the trips for which we were unable to associate land use values are trips outside the study area. This explains their prevalence near the study area boundary, where people are much more likely to travel outside the CMA even for 'short' trips. If we were to include all households where land use information for trips was complete and exclude those living in the same tracts with missing information, this might cause a bias in our assessment of the causes of under-reporting. As such, for our logit modelling, PDs and SMs where more than 10% of individuals have incomplete land use information were removed.

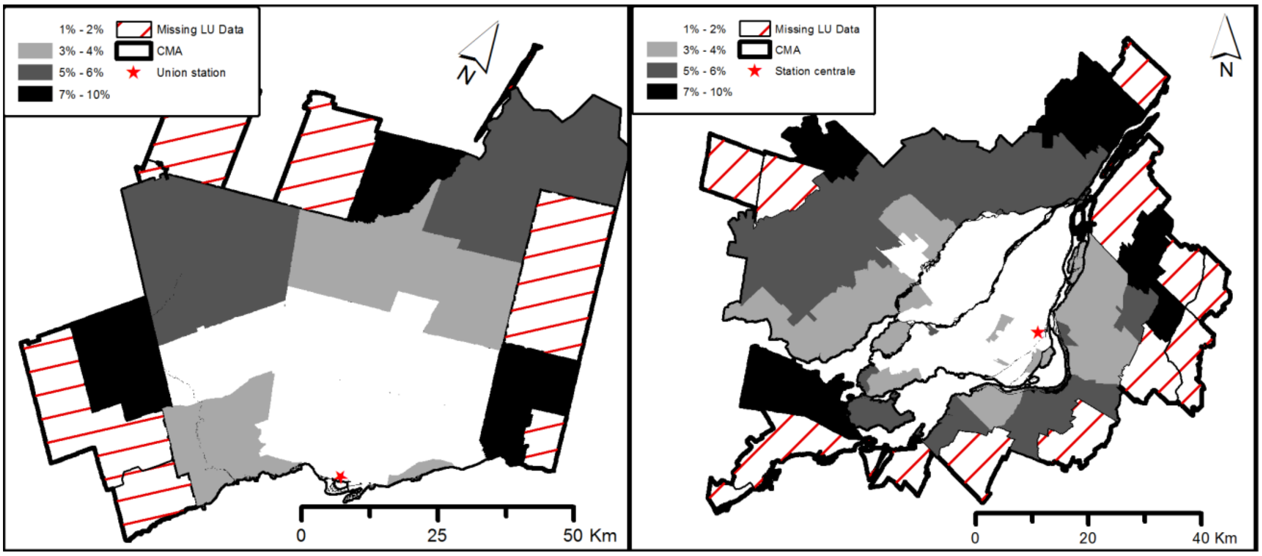

Figure 8. Percent person-days with missing land use data by PD and SM

Using simple summary statistics and looking at subsets of the population already provides a strong case for there being a land use/neighbourhood-type effect. There is, again, no obvious explanation why persons living in more densely populated, and retail and transit accessible areas would be less likely to report discretionary trips, as is found in the Toronto dataset. Whether they have access to cars or not, it appears central city residents are, according to Toronto's TTS, not making use of the amenities at their disposal – recognizing that a large share of these trips are not simply unreported (behavioural effect), but rather are not recorded (method). This also goes against the findings of Boarnet et al. who find that increased accessibility increases overall travel (Boarnet, et al., 2011).

Correlation is only indicative however. To gain a better understanding of whether the relationship between land use and travel holds up when control variables are employed, we present results of our binary logit model below. The dependent variable is whether a person reported one or more discretionary trips on the survey day. The logic in employing such a model is to then look for any counter-intuitive land-use and discretionary trip making relationship. Such a relationship would indicate having isolated the effect of land use on under-reporting discretionary trips, whether by outright omission of walk trips (in Toronto) or as a result of non-reporting of short discretionary trips as is common in the literature (in both Toronto and Montreal). The counter-intuitive coefficient expected in Montreal, we interpret as a measure of the effect of under-reporting in a behavioural sense, as relates to more urban locations where walk trips are more prevalent, with the difference between effects in Toronto and Montreal being the additional effect of systematic under-reporting resultant from non-recording of discretionary walk trips.

4.2 Logit model

As described in the Methods section (3.3), the way we sought to quantify the effect of land use on under-reporting was by estimating a binary logit model. Within the model, the dependent variable indicates whether a respondent reports or does not report making one or more discretionary trips on the day of the survey. To create the dependent variable, all the trips reported in the TTS and EOD were first coded as being made for discretionary (1) or non-discretionary (0) purposes. The maximum value was then taken for each respondent. Records represent respondents in the estimation.

In addition to this straightforward data preparation, land use at the home and all destination locations needed to be represented. This was done first by associating every home and trip destination with a census tract (Montreal) or traffic analysis zone (Toronto). Land use and neighbourhood type data associated to the tract or zone in question were then pulled in to the respondent dataset. For destination neighbourhood types included in the model, the highest cluster value was kept, labeled as 'most urban dest'(ination). For individuals who made no trips on the survey day, the highest destination cluster is represented as the home cluster. 0-trip individuals were included in the model estimation, as the likelihood of not recording any travel was associated with the non-recording of discretionary trips. This relationship was clearly demonstrated in the TTS' summary statistics, where, to reiterate, the percent of individuals reporting no travel increased by over 50% at the core when compared with lower density, less mixed and transit accessible environments (see Table 1). We wanted to make sure our model captured this effect.

The estimated logit model odds-ratios are shown in Table 5. Odds ratios are shown as they can easily be interpreted, as odds multipliers. To give a simple example, an estimated coefficient value of 0.5 for a given dummy variable (coded 0 or 1), means that if an observation in the data has a '1' for this independent variable, the pre-existing odds of an outcome 'y' (the dependent variable) are halved. If, on the other hand, the estimated coefficient value is 2, an observation in the data with a '1' for this variable would indicate a doubling of the pre-existing odds for outcome 'y'. An estimated coefficient of 1 would mean that the variable has no effect. In our case, the dependent variable is making at least 1 discretionary trip on the reported survey day.

Table 5. Binary logit model results, with most relevant sections highlighted. Dependent variable is making at least 1 discretionary trip

		Toronto			Montreal		
		Odds Ratio	z	P>z	Odds Ratio	z	P>z
Home Variables	Distance* from Union	0.993	-4.53	0	0.993	-2.37	0.018
	Distance^2 from Union	1.000	8.66	0	1.000	2.46	0.014
	Home = Sprawl	0.767	-13.67	0	0.810	-7.51	0
	Home = Suburban	0.613	-17.02	0	0.719	-7.22	0
	Home = Semi-urban	0.625	-9.93	0	0.706	-4.82	0
	Home = Urban	0.554	-6.16	0	0.711	-2.12	0.034
	Home = Core	0.445	-5.09	0			
	Transit service density	0.864	-6.59	0	1.010	1.92	0.055
	Population density	0.997	-5.58	0	1.000	-0.06	0.952
	Job density	1.000	1.01	0.314	1.001	1.49	0.136
	Retail Density	1.004	0.96	0.338	0.992	-1.25	0.213
Destination Variables	Most urban dest = Sprawl	1.982	39.73	0	1.278	12.81	0
	Most urban dest = Suburban	3.099	60.76	0	1.747	26.63	0
	Most urban dest = Semi-urban	4.020	62.91	0	2.730	37.01	0
	Most urban dest = Urban	3.599	43.09	0	1.899	27.31	0
	Most urban dest = Core	3.921	58.73	0			
Mobility Tools	Auto available	1.659	19.9	0	1.082	2.58	0.01
	License	1.681	15.83	0	1.588	12.86	0
	License+Auto	1.253	6.26	0	1.164	3.65	0
Person Characteristics	PT Worker	0.694	-19.99	0	0.536	-22.58	0
	FT Worker	0.441	-63.56	0	0.286	-69.38	0
	Youth + Student	0.964	-1.62	0.104	0.460	-23.27	0
	Adult + Student	0.417	-33.54	0	0.343	-38.43	0
	Male	0.897	-11.04	0	1.005	0.38	0.701
Household Characteristics	2 Adults, < 2 incomes	1.025	1.2	0.228	1.132	5.1	0
	Dual Income, No Kids	0.927	-3.03	0.002	0.954	-1.72	0.086
	2+ Adults w Kid(s)	1.014	1.01	0.314	0.971	-1.47	0.143
	Single Parent	1.195	5.51	0	1.277	6.17	0
	Retiree	1.065	2.23	0.026	0.807	-6.42	0
	Retirees	1.712	26.03	0	1.191	6.81	0
	Single Adult	1.098	3.48	0.001	1.225	6.63	0
	Student(s) Only	1.089	1.38	0.169	1.229	2.96	0.003
	Nb persons	0.888	-22.23	0	0.933	-8.36	0
Proxy	Is Survey Respondent	1.788	52.97	0	1.699	38.6	0
	Constant	0.234	-31.61	0	0.523	-11.56	0
* All distances expressed in km			Obs.	217,815		Obs.	138,131
* All densities in units/hectare			Pseudo R2	0.097		Pseudo R2	0.092

As was our hypothesis, the estimated model presented in Table 5 shows intuitive signs for almost all variables except land use at the home location. Specifically, we find that after controlling for demographics and vehicle access (which here is also somewhat correlated with making discretionary trips via motorized modes and thus being recorded), denser, more urban environments lead to a decrease in discretionary trip making. This is found for both cities. In contrast, travel to a denser urban environment increases the pre-existing odds of making a discretionary trip ('destination' land use neighbourhood-type binary variables).

An important distinction between Toronto and Montreal estimated coefficients, however, is the magnitude of the neighbourhood type effect for the home location. Controlling for demographics, mobility tools and distance from the core, living in a denser, more mixed environment reduces pre-existing odds of reporting discretionary trip making by 19 to 29% in Montreal (suburban to semi-urban, urban and core) and 39 to 55% in Toronto (suburban to core). This despite the fact travel to the core increases the pre-existing odds of making a discretionary trip by 170% for Montreal and 300% for Toronto.

To answer our main research question, then, our results indicate that the effect of land use on under-reporting can be as high as a 29% decrease in the pre-existing odds of discretionary trip making when walk trips are not explicitly ignored, with effects estimated as high as 55% when they are. These are non-insignificant effects.

The way we interpret these findings is that persons living in dense, urban core-type environments are likely to make short, simple-tour walking trips for discretionary purposes (which would have no possibility of being recorded in Toronto, but are potentially recorded in Montreal). Persons living elsewhere than these dense environments, in turn, are more likely to make complex tours (as was found in Harding, et al. 2015) and make the most of the amenities present by chaining trips together. Again, when chained within complex tours, walk trips do get reported in both datasets.

In addition to the significant neighbourhood type coefficients, we find statistically significant, but counter-intuitive signs for population density and transit access in Toronto. By this we mean that a given increase in density or transit access would be expected to also lead to an increase in the likelihood of discretionary trip making, but after controlling for household type, the reverse is found. Furthermore, the fact that retail density does not have a statistically significant effect in either city – or a reasonable magnitude of effect- is also a clear indication there may be a significant problem resultant from under-reporting of walk or discretionary trips overall. If this under-reporting is indeed the cause for the non-significant coefficient estimate, then the implications from a modelling standpoint are considerable. This would mean that trip generation and mode choice models based on this data will produce biased estimates of number of trips and the relative shares for each mode, while trip distribution models in turn will also produce erroneous predictions because attraction for non-mandatory activities is not properly measured.

Regarding the significant and positive coefficient estimates for auto available, license and license plus auto, results are consistent with findings by Driscoll et al., who find "an attachment to the car as a mode of transport for non-commuting journeys when available" (Driscoll, et al., 2013). This in turn means we should be picking up all the more discretionary trips in the suburbs and by car-owning households – which is exactly what was shown in Table 5.

A final variable worth mentioning is "Is survey respondent", with its large and positive coefficient that can be read as a significant proxy bias effect. Not having the individual on the phone to report their own travel (Is survey respondent = 0) greatly decreases the likelihood of recording a discretionary trip, where the pre-existing odds of making a discretionary trip are 79% greater in Toronto and 70% greater in Montreal if the travel data are self-reported. This variable is not always provided with regional OD datasets, but reminds us of a significant limitation with the EOD, TTS, and telephone household surveys more broadly. We report this effect here as it is significant, but as proxy bias is already well documented, unlike land use effects, we do not focus on this aspect.

5. Conclusion and policy recommendations

As per our results, it would appear there is indeed a relationship between land use and trip under-reporting for discretionary purposes. Knowing that walk trips specifically are lost as a result of the way in which TTS data are collected and recorded, it then follows that many of these trips not reported may be trips made on foot. A recent paper by co-authors Harding and Miller showed the prevalence of walk trips in the core using a smartphone-based, passive location trace dataset. In that particular case, while the sample was small and not necessarily representative of the broader City population, over 50% of all detected trips were made on foot (Harding, et al., 2015). Many of these were short, but if that is what travel looks like in the dense core of our cities, not making every effort to collect that data may lead to significant problems when one wishes to draft evidence-based policy to effect change in travel behaviour. These are not merely trips made that do not require vehicle lanes or additional transit capacity, these are trips potentially replacing motorized trips.

While our results are fairly stark and raise concerns with regard to the accuracy of reported discretionary travel in the Toronto core specifically, and in the region more broadly to a lesser extent, the case we have looked at is an extreme; follow-up studies should be conducted to obtain additional estimates on the effect of land use in this regard. That is not to say, however, that trip under-reporting does not have profound effects or is not associated with land use in other contexts, irrespective of whether the true effect is a 10, 20 or 30% reduction in the pre-existing odds of reporting discretionary trips in more urban environments. The trends found here of a concentration in problematic reporting are likely also found in other cities' surveys where short, walked, discretionary trips are under-reported. To the extent that there is any correlation between land use and the propensity to make these kinds of trips, then the bias in model estimates will be greater in denser, mixed environments and reduce the potential for regional survey data to guide policy aimed at increasing the share of AT.

Another problem resulting from current data collection methods is that 'typical' trip distances by mode are often used as a way to estimate how many trips could be diverted (example from car to bike or transit to walk) were some improvement made to the built environment. In Toronto's case to a great extent, and in the case of Montreal to a lesser degree, these 'typical' distance values are based on a biased sample of actual travel.

The reason cycling advocates go out and collect data on bike traffic along a corridor is that this information can be used to make a case for better infrastructure provision. If we have no data on the prevalence of walk trips for simple tours, how do we know where sidewalks are being utilized? How do we know if a pedestrian connection over rail tracks is worth the investment required for its maintenance? These are non-trivial questions.

Providing infrastructure for different types of mobility also comes at varying costs. If a new development is being built, a decision on whether or not to connect it with LRT is not made by flipping a coin. Data are collected and ridership estimated, because that decision will have long term repercussions. Likewise, what type of pedestrian facilities and the amount of road-space allocated to pedestrians and cyclists should be evidence-based decisions.

Planning infrastructure investments with only ITE-style trip generation rates can have serious implications, not only on the potential to cause road conflicts in areas or misunderstand travel patterns and their associated emissions, but also can have the effect of overspending on infrastructure for certain modes. As Millard-Ball has stated, we talk of "conservative" estimates when planning for parking and road space to be allocated to cars, but it is difficult to see how spending too much on infrastructure that will be poorly utilized on the off-chance congestion may occur is a "conservative" way of doing things (Millard-Ball, 2015). We need a shift in the conversation. Better or more attentive data collection methods that reduce the biases inherent in travel surveys and further research on the effects of under-reporting on forecasts are needed.

Acknowledgements

We would like to thank Yunfei Zhang of UTTRI for her help early on. We would also like to acknowledge the financial support provided by both the Social Sciences and Humanities Research Council (SSHRC) and Natural Science and Engineering Research Council (NSERC) of Canada. Finally, we would like to thank our reviewers for their invaluable feedback.

References

Agence Métropolitaine de Transport, 2013. *Enquête Origine Destination 2013 de la Région de Montréal version 13.2b,* s.l.: s.n.

Badoe, D. & Miller, E., 2000. Transportation-land-use interaction: empirical findings in North America, and their implications for modeling. *Transportation Research Part D,* pp. 235-263.

Boarnet, M. G. et al., 2011. Retrofitting the Suburbs to Increase Walking: Evidence from a Land-use Travel Survey. *Urban Studies,* 48(1), pp. 129-159.

Boarnet, M. G., Nesamani, K. & Smith, C. S., 2004. *Comparing the Influence of Land Use on Nonwork Trip Generation and Vehicle Distance Traveled: An Analysis using Travel Diary Data.* Washington, DC, s.n.

Carmichael, L. et al., 2012. Integration of health into urban spatial planning through impact assessment: Identifying governance and policy barriers and facilitators. *Environmental Impact Assessment Review,* Volume 32, pp. 187-194.

Chen, C., Bian, L. & Ma, J., 2014. From traces to trajectories: How well can we guess activity locations from mobile phone traces?. *Transportation Research Part C: Emerging Technologies,* Volume 46, pp. 326-337.

City of Toronto, 2015. *Richmond-Adelaide Cycle Track Planning and Design Study Pilot Project Preliminary Evaluation,* s.l.: s.n.

Clifton, K. J., 2013. *Improving the representation of the pedestrian environment in travel demand models – Phase I,* Oregon, Portland: s.n.

Clifton, K. J., Singleton, P. A., Muhs, C. D. & Schneider, R. J., 2015. *Development of destination choice models for pedestrian travel.* Windsor, UK, s.n., p. 16.

Currie, G. et al., 2010. Investigating links between transport disadvantage, social exclusion and well-being in Melbourne - Updated results. *Research in Transportation Economics,* Volume 29, pp. 287-295.

Data Management Group, 2013. *Transportation Tomorrow Survey (TTS) 2011, Version 1.0 Data Guide,* s.l.: s.n.

Dimitriadou, E., Dolnicar, S. & Weingessel, A., 2002. An examination of indexes for determining the number of clusters in binary data sets. *Psychometrika,* 67(3), pp. 137-160.

Driscoll, A., Lyons, S., Morgenroth, E. & Nolan, A., 2013. *Comparing the Determinants of Mode Choice across Travel Purposes,* Dublin: Munich Personal RePEc Archive.

Ewing, R. & Cervero, R., 2001. Travel and built environment: A synthesis. *Transportation Research Record,* Volume 1780, pp. 87-114.

Ewing, R., DeAnna, M. & Li, S.-C., 1996. Land Use Impacts on Trip Generation Rates. *Transportation Research Record: Journal of the Transportation Research Board,* Volume 1518, pp. 1-6.

Frank, L. D. & Pivo, G., 1994. Impacts of Mixed Use and Density on Utilization of Three Modes of Travel: Single-Occupant Vehicle, Transit, and Walking. *Transportation Research Record,* Volume 1466, pp. 44-52.

Harding, C., Miller, E. J., Patterson, Z. & Axhausen, K. W., 2015. *Multiple Purpose Tours and Efficient Trip Chaining: An Analysis of the Effects of Land use and Transit on Travel Bahavior in Switzerland.* Washington, DC, s.n.

Harding, C., Patterson, Z. & Axhausen, K. W., 2013. *Trip dispersal, efficient chaining and modal propensity: an analysis of the effects of neighborhood type, regional accessibility and activity space geometry on travel behav-iour in Switzerland,* Zurich: Swiss Federal Institute of Technology (ETH).

Harding, C., Zhang, Y. & Miller, E. J., 2015. *Multiple purpose tours and efficient trip chaining in Toronto: an analysis of the effects of land use and transit provision on mode choice and trip chaining using smartphone data.* Windsor, UK, s.n.

Harding, C., Zhang, Y. & Miller, E. J., 2015. *Waterfront Toronto Smartphone Data Collection Project,* Toronto: University of Toronto Transportation Research Institute.

Harris / Decima, 2008. *City of Toronto: Walking Habits and Attitudes Report,* s.l.: s.n.

Hess, P., Lea, N. S., Bidordinova, A. & Klassen, J., 2014. *Identifying and Overcoming Barriers to the Implementation of Active Transportation Policiies: Final Report,* Toronto: s.n.

Jin, X., Asgari, H. & Hossan, M. S., 2014. *Understanding Trip Misreporting in Household Travel Surveys by Comparing GPS-Assisted and Diary-Based Samples.* Changsha, China, s.n., pp. 3401-3412.

Krizek, K., 2003. Neighborhood services, trip purpose, and tour-based travel. *Transportation,* Volume 30, pp. 387-410.

Martens, K. & Hurvitz, E., 2009. Distributive impacts of demand-based modelling. *Transportmetrica,* Volume iFirst, pp. 1-20.

Millard-Ball, A., 2015. Phantom Trips: Overestimating the traffic impacts of new development. *Journal of Transport and Land Use,* 8(1), pp. 31-49.

Millward, H. & Spinney, J., 2011. Time use, travel behavior, and the rural–urban continuum: Results from the Halifax STAR project. *Journal of Transport Geography,* Volume 19, pp. 51-58.

Ministère du Développement durable, Environnement et Lutte contre les changements climatiques, 2016. *Québec's Commitments, Our GHG Emissions Reduction Targets.* [Online] Available at: http://www.mddelcc.gouv.qc.ca/changementsclimatiques/engagement-quebec-en.asp

Ministry of the Environment and Climate Change, 2014. *Ontario's climate change update 2014,* s.l.: Government of Ontario.

Miranda-Moreno, L. F. et al., 2011. Simultaneous Modeling of Endogenous Influence of Urban Form and Public Transit Accessibility on Distance Travelled. *Transportation Research Record,* Volume 2255, pp. 100-109.

Newman, P. & Kenworthy, J., 1989. Gasoline Consumption and Cities: a comparison of U.S. cities with a global survey. *Journal of the American Planning Association,* 55(1), pp. 24-35.

Office Fédéral de la Statistique, 2012. *La Mobilité en Suisse,* s.l.: s.n.

Openshaw, S., 1984. *The modifiable areal unit problem (Concepts and Techniques in Modern Geography).* Norwich: Geo books.

Shay, E. & Khattak, A., 2007. Automobiles, Trips, and Neighborhood Type: Comparing Environmental Measures. *Transportation Research Record,* pp. 73-82.

Shen, L. & Stopher, P., 2014. Review of GPS travel survey and GPS data-processing methods. *Transport Reviews,* 34(3), pp. 316-334.

Statistics Canada, 2011. *National Houshold Survey 2011,* s.l.: s.n.

StudentMove, 2015. *StudentMoveTO (student travel survey),* Toronto: Data Management Group.

Wilkinson Davis, A. et al., 2015. *Perception and Reality: Linking Metrics of Place Perception to Measurable Attributes of Place Using Cross-Classified SEM Analysis.* Windsor, UK, s.n., p. 22.

Williams, P. A., 1989. The influence of residential accessibility on household trip-making. *Socio-Economic Planning Sciences,* 23(6), pp. 373-385.

Wolf, J., 2000. *Using GPS data loggers to replace travel diaries in the collection of travel data,* s.l.: s.n.

Wolf, J., Wilhelm, J., Casas, J. & Sen, S., 2013. A Case Study: Multiple Data Collection Methods and the NY/NJ/CT Regional Travel Survey. In: *Transport Survey Methods: Best Practice for Decision Making.* s.l.:Emerald Group Publishing.

Zahabi, A., Miranda-Moreno, L., Patterson, Z. & Barla, P., 2013. Urban Transportation Greenhouse Gas Emissions and its Link with Urban Form, Transit Accessibility and Emerging Green Technologies. *Journal of Transportation Research Record.*

Zhang, M., 2004. The Role of Land Use in Travel Mode Choice. *Journal of the American Planning Association,* 70(3), pp. 344-360.