Spatio-Temporal Analysis of Car Distance and Greenhouse Gases and the Effect of Built Environment: A latent class regression analysis

Seyed Amir H. Zahabi (corresponding author)
Department of Civil Engineering Applied Mechanics, McGill University
Room 398, Macdonald Engineering Building, 817 Sherbrooke Street West
Montreal, Quebec H3A 0C3
Tel: (514) 649-4579
seyed.zahabi@mail.mcgill.ca

Luis F. Miranda-Moreno
Department of Civil Engineering and Applied Mechanics, McGill University
Room 268, Macdonald Engineering Building, 817 Sherbrooke Street West
Montreal, Quebec H3A 0C3
Tel: (514) 398-6589
Fax: (514) 398-7361
luis.miranda-moreno@mcgill.ca

Zachary Patterson
Department of Geography, Planning and the Environment, Concordia University
1255-15, Hall Building, 1455 De Maisonneuve W.,
Montreal, Quebec, Canada
Tel: (514) 848-2342 ext 3492
Fax: (514) 848 2032
zachary.patterson@concordia.ca

Philippe Barla
Centre for Data and Analysis in Transportation (CDAT)
Université Laval
Département d’économique
1025 av. Des Sciences-Humaines, Québec, QC G1V 0A6 Canada
Phone: 418-656-7707
philippe.barla@ecn.ulaval.ca

Submission date: 31 July, 2011
5703 Words + (7 *250) Figures and Tables = 7453 Words

Submitted for presentation/publication to the 92st Annual Meeting of the TRB
ABSTRACT

This paper has two main objectives: i) to estimate GHG emission inventories at the household level using link-level average speeds for three origin-destination surveys in Montreal, Canada for the years 1998, 2003 and 2008; and ii) to estimate the temporal and spatial variation of built environment characteristics and socio-demographics on automobile distance traveled and GHG emissions at the household level. To estimate emissions, different sources of data are combined including road network, link-level average speeds during different hours, trip-level information and vehicle fleet characteristics. Urban form indicators over time such as population density, land use mix and transit accessibility are generated for each household in each of the three waves. A Latent Class (LC) regression modeling framework is then implemented to investigate the link between the built environment, socio-demographics, GHGs and automobile distance traveled. The latent class modeling approach results in three population subclasses and as a result, three separate models of household-level automobile distance traveled and transport-related GHG emissions. This demonstrates the utility of using this approach when the nature of the dataset has spatial and temporal variation. Our findings on the effect of built-environment (BE) and transit accessibility (TA) on GHG and car distance travel are consistent with the literature. Also overall, we observe a declining trend in travel-related GHG emissions over time. By setting all other variables to a base case, we observe that a given household produced 15% and 10% more GHG in 1998 and 2003 compared to 2008. This is due to improved fuel economy of auto-vehicles over time and an increase in transit trips. Employment status also significantly affects household GHGs (with elasticities as high as 51% for each additional full-time worker). As expected, and consistent with the literature, low and medium income households pollute less than high-income households (42% less GHG for low income class compared to high income).
INTRODUCTION

Transportation makes up 27% of Canada’s total greenhouse gas (GHG) emissions (1). In addition, transport-related GHGs have risen consistently, resulting in an increase of nearly 33% between 1990 and 2005 (2). In order to limit climate change, local and worldwide policy makers are looking for strategies to reduce vehicular emissions. In Quebec, the provincial government is aiming to reduce GHG emissions by 20% with respect to 1990 levels by the year 2020.

An extensive literature has developed on different strategies and policy options related to how the built environment (often represented by population density, land use diversity, and transit accessibility) might be used to reduce automobile transportation and transport-related GHG emissions. In this literature, it is contended that planning with the “D’s” in mind will make it possible to reduce automobile dependence by developing dense, diverse, and well-designed neighborhoods with efficient public transportation options. Much of the empirical literature on this topic has found that density, land use and transit accessibility are important factors in determining household travel outcomes such as travel distance and GHGs. The empirical research is typically approached using cross-sectional data. Research using disaggregate data (individual or household level data) is mostly based on cross-sectional analyses comparing mobility patterns across cities or neighborhoods at a single point in time. Moreover, despite having been recognized as an important consideration in properly understanding transportation behavior, past studies have not considered the differences that might exist in the built environment – travel outcome across household subpopulations or classes.

As such, the purpose of this research is to estimate the effect of the built-environment on household automobile distance traveled and transport-related GHG emissions, over time and across population subgroups. In order to do this, a unique 3-wave transport related household GHG inventory is built and used with the Montreal Origin-Destination surveys of 1998, 2003 and 2008. A latent class modeling approach is adopted to estimate the differences in built-environment effects across population subgroups.

This paper is structured as follows: in the following section background literature is discussed, then a description of the study area. This is followed by a description of the methods used to estimate distance traveled and GHG emissions at the trip level, and the determination of BE attributes. Next comes a section on summary statistics and figures regarding the input data. Afterwards another section presents the empirical results of the statistical models. The final chapter concludes with policy implications.

BACKGROUND

Due to the size of the literature as well as paper length restrictions, it is not possible cover the entire literature on the link between travel behavior and the built environment. Instead, this section summarizes the primary literature on the topic, and in particular on the literature considering the effect of built-environment characteristics on vehicle miles traveled (VMT) and GHG elasticities.

This literature has concentrated on how the built environment, commonly represented by population density, land use diversity, and transit accessibility, might be used to reduce GHG emissions. In particular, the literature concentrates on whether planning with the “D’s” in mind would make it possible to reduce automobile dependence by developing dense, diverse, and well-designed neighborhoods with efficient public transportation options. Previous literature on travel behavior suggests that density, land use and transit accessibility are contributing factors in determining household travel outcomes such as travel distance and GHGs. For the most part, the literature also agrees that these effects are predictable: higher density, land-use mix and transit accessibility all tend to reduce travel distances and GHG emissions. This literature is so large that there now exist a number of literature reviews on these topics, such as Badoe and...
Miller (3), Handy et al. (4), Ewing and Cervero (5) and TRB report 298 (6). Most of the empirical evidence in this literature is based on cross-sectional studies that have analyzed travel behavior, while controlling for measures of the built environment (BE) and socio-economic control variables. Different statistical approaches and degrees of data aggregation have been used for studying the impacts of built-environment and transit accessibility on travel distance and GHGs. In some studies, the problem of residential self-selection has also been accounted for by using simultaneous equation modeling approaches Brownstone and Golob (7), Eluru et al. (8), Miranda-Moreno et al. (10). Also, the relationship between the built environment and travel outcomes has typically been studied at the regional scale using cross-sectional data (Donoso et al., (11); Ewing et al., (12); Hunter, (13)). Some studies have considered data from several cities (Bento et al. (14)).

Research using disaggregate data (individual or household level data) is mostly based on cross-sectional analyses comparing mobility patterns across cities or neighborhoods at a single point in time. It has been recognized in the literature that if temporal trends are not taken into consideration, built environment – travel outcome relationships may be spurious. Moreover, the presence of subpopulations has been demonstrated in other travel outcomes (e.g. Greene and Hensher (15)), but has yet to be explored in the built-environment travel behavior literature. For example, it may happen that car distance and GHGs evolve in different ways across different subgroups, or classes, over time. If subpopulations exist, the effects of built environment factors (density, diversity and transit accessibility) can have different effects on the different groups. Few studies have looked at the temporal and spatial variations of car distance and GHGs across subpopulations in the same region using disaggregate data (household level). The research presented here aims to fill this gap in the literature by estimating the effect of built-environment characteristics on household automobile distance traveled and transport-related GHG emissions over time as well as across subpopulations.

METHODOLOGY

The methodology proposed for this research builds on previous research dealing with the disaggregate analysis of urban travel GHG emissions (9, 10, 19). For this work, several sources of data are needed, including: link-level speeds and trip-level data from a household survey, motor-vehicle fleet characteristics, land-use data, etc. In this empirical analysis, the main sources of trip data are three O-D surveys for the years 1999, 2003 and 2008 from the region of Montreal, Canada. These surveys provide urban travel information for a very large sample of households in the region – about 5% in each survey (~65,000 households). The OD survey collects a great deal of information from households. The main part of the data collected is related to trips undertaken by all family members (over the age of 4) in the day previous to the survey. The information collected for each trip includes: origin and destination x-y coordinates, transportation mode(s), purpose, transit lines used, time of departure, car occupancy, etc. Socio-demographic information at the individual and household level includes gender, age, work status, family structure, number of vehicles at home and household income. Since O-D survey data do not include information on the make, model, or year of vehicles owned by each household, this was estimated using motor-vehicle fleet inventory data from the Quebec automobile insurance corporation - SAAQ. Also, to estimate average link-based speeds for the different traffic conditions during the day, a road network and traffic assignment model built in EMME and calibrated and maintained by the provincial ministry of transportation was used.

This research consists of 4 four main tasks:

a) The calculation of trip GHGs using a link-level average speed approach: Emissions for each O-D trip were estimated using link-level attributes such as link speed, link-distance as well as trip-level
attributes such as vehicle occupancy and fuel consumption rate. Trip emissions were then aggregated to the household level.

b) The definition of the BE indicators: The three main factors studied are residential density, land use mix and transit accessibility. These factors were estimated using different sources of data as close as possible to the years of the 3 years O-D surveys. Other factors such as employment density were tested but they are highly correlated with these other three.

c) The estimation of the impact of the BE on both car distance and GHGs across years and subpopulations: For this a latent-class regression approach with fixed-temporal effects is adopted and compared to basic models.

d) The comparison of elasticities across years and subpopulations: This is to determine the impact of the BE on the spatial-temporal variability of GHGs.

Trip-Level GHGs:

For each trip in the three O-D surveys (1998, 2003 and 2008), two GHG emitting mode categories are distinguished, private motor vehicles and public transit including transit buses and commuter trains. Some trips can involve more than one mode. The procedure for GHG emissions estimation is described as follows:

i. From a traffic assignment model developed and calibrated by the Quebec provincial ministry of transportation (MTQ) (18), the congested times for each link of the network were obtained along with their distances. Link travel times were obtained hourly for all periods of the day.

ii. Associate each trip according to the departure time to a particular (time-of-day) network described in the previous step. Then calculate shortest paths (based on congested times) at the link level to obtain routes and link distances and speeds.

iii. Obtain ridership, fuel consumption rates and emission factors for each trip.

iv. Compute GHGs emission according to equations 1 and 2.

For trips involving motor vehicle as a unique or combined mode, the emissions are estimated using distance and average speed at the link level, vehicle fuel consumption rate (FCR) and GHGs emission factors. This procedure is fully defined in Barla, et al (19) and Barla, et al (9). Then, emissions for a given trip j departing in a particular hour t is estimated as:

\[ GHG_{Ajt} = \sum_{i=1}^{N} [SP_{ij} \times D_{ij}] \times \frac{FC_{Aj} \times EF_{A}}{R_{ij}} \]  (1)

Where:

- A - automobile
- i - Link (i=1,…,N links used by trip j)
- j - Trip
- t - Departure time

\( GHG_{Ajt} \) = GHGs for automobile trip j (kg of CO\(_2\)) departing at time t.

\( D_{ij} \) = Travel distance on segment (link in network) i in 100km.

\( SP_{ij} \) = Speed correction factor for segment i of trip j departing at t. Since fuel consumption also depends upon speed, speed correction factors developed by the MTQ were also used. These factors were produced after a local calibration of MOBILE6 (for further details, see (20)). Link speed was matched with its corresponding speed correction factor.

\( FC_{Aj} \) = Average fuel consumption rate (FCR) in liters of gasoline/100km for the vehicle used in trip j. This was generated using the motor-vehicle fleet inventory of the automobile insurance corporation of
Quebec (SAAQ). For further details see (21). This inventory contains the make, year and model of each vehicle in the province as well as the fuel consumption rate per km. However, the address of the vehicle is provided at the 3-digit postal code. Therefore, FCR at three-digit postal code (FSA-Forward Sortation Area) were generated. An FCR is then associated to each vehicle belonging to the same spatial area.

\[ EF_A = \text{Emission factor for gasoline (2.289 kg of CO}_2/\text{liter of gasoline)} \]. This is obtained from the national inventory report by Environment Canada. Although this number is fixed for all gasoline vehicles for CO\textsubscript{2}, for other GHG emissions such as CH\textsubscript{4} and N\textsubscript{2}O, the emission factor depends on the type of vehicle (ex. Light duty, heavy duty, Oxidation Catalyst, non-catalytic controlled and etc.). Since we didn’t have knowledge of the type of the fleet owned by the household, we were unable to estimate the emission for other GHG emissions.

\[ R_{Bj} = \text{Number of passengers in trip j including the driver. This is determined from the O-D survey data. Car trips in the same household, departing at the same hour and with the same origin-destination are associated to the same motor-vehicle trip.} \]

For uni-modal or multimodal trips involving public bus transit and/or commuter trains, GHGs are estimated in a similar fashion. In this case, however, average speeds at the trip-level are used since link-level speeds were not available, but this speed considers congestion. For the bus portion, GHGs are calculated using the following equation:

\[ GHG_{Bj} = \frac{FC(S)_{Bj} \times D_{Bj} \times EF_B}{R_{Bj}} \]  \hspace{1cm} (2)

\[ GHG_{Bj} = \text{GHGs for bus portion of transit trip j (kg of CO}_2) \]

\[ FC(S)_{Bj} = \text{Average fuel consumption as a function of operating speeds (S) in liters of diesel/100km).} \]

Fuel consumption rates for the typical fuel bus technology operating in real conditions were obtained from a recent field study done by the local transit agencies, the Société de transport de Montréal (STM). The fuel consumption curve according to this study is given by:

\[ FC(S) = 255.33 \times (\text{Bus speed})^{-0.4753} \] \hspace{1cm} (3)

\[ D_{Bj} = \text{Distance traveled by bus in transit trip j (km). For each trip involving transit (bus, metro and commuter trains) in the Montréal region, distances are obtained using the public transit software, MADIGAS (22). Trips were simulated by the Agence Métropolitaine de Transport (AMT).} \]

\[ EF_B = \text{Emission factor for diesel. Here, an emission factor of 2.663 kg CO}_2/\text{liter of gasoline is considered based on the recommendation of Environment Canada for Canadian city conditions}^2. \]

\[ R_{Bj} = \text{Ridership for bus on trip j. In this case we use a mean value for each line used in the trip. This is obtained from the bus provider agencies for each bus line.} \]

For commuter train lines using diesel or diesel-electric locomotives, average fuel consumption for passenger-km (FC/PK) were directly estimated by the local commuter train agency (Agence métropolitaine de transport - AMT). This was done by dividing the annual fuel consumption (liters of diesel) by their respective annual passenger kilometers traveled. Travel distance by rail (DR) is then estimated for each trip (km). By multiplying (DR) by the fuel consumption rate per passenger km (FC/PK),

\[ \text{Page dimensions: 612.0x792.0} \]

---


liters of fuel consumed for the train segment are estimated. To get the kg of CO\textsubscript{2} for each trip, the resulting liters of fuel for each trip is multiplied by the emission factor for CO\textsubscript{2} obtained from Environment Canada. This is equal to 2.663 kg of CO\textsubscript{2} for each liter of diesel fuel combusted in trains. It’s worth mentioning that the GHG emissions from the metro (subway system) are assumed to be zero since it runs on hydro-electric.

To obtain the household inventory, GHGs are estimated for each uni-modal and multimodal trip in the O-D surveys. Trip level emissions are then aggregated at the individual and household level.

**BE Attributes and Neighborhood Typologies**

To generate the BE indicators in the vicinity of each household involved in this analysis, a nine-cell grid approach was undertaken \((10)\). This is done in order to keep the benefits of a region-wide grid but to partly overcome the inaccuracies (the instability of the results) in a normal grid method. The approach involves creating a 500m x 500m grid for the Montreal census metropolitan area (CMA). Each central cell (center of the 9 gridcells), for which the attributes of the eight surrounding cells are also considered equally and applied to this central cell. In defining a grid cell at 500 meters, the nine-cell grid method creates an area that approximates a buffer with an approximately 900 m radius (the minimum “radius” is 750 m, and the maximum is 1061 m).

According to this grid approach, the following indicators are built for each year involved in the analysis (1988, 2003, and 2008):

- **Land use mix**: Using the nine-cell grid approach, land use mix was calculated using the entropy index (Theil et al. \((23)\); Frank et al. \((24)\)). The land uses considered are those defined by Desktop Mapping Technologies Inc. (DMTI) and include residential, commercial, institutional and governmental, resource and industrial, and park and recreation, with water and open area not being considered. The computation of this index is according to the common equation that can be seen in Frank et al. \((24)\) and Miranda-Moreno et al. \((10)\).

  \[
  E_j = -\sum_{i=1}^{n} \frac{\left(\frac{A_{ij}}{D_j}\right)\ln\left(\frac{A_{ij}}{D_j}\right)}{\ln(n)}
  \]

  In this equation \(A\) is the area of land use \(i\) in the nine-cell grid \(j\), \(D_j\) represents the total area of nine-cell grid \(j\), without taking into account water and open area and \(n\) is the total number of different land uses (5 in this paper).

- **Population density**: Population was obtained at the census tract level from Statistics Canada \((25)\) for the Montreal CMA. Land use data from DMTI Spatial was then used to more accurately allocate population within each census tract, which then allowed for the calculation of approximate population per grid cell.

- **Transit accessibility**: The grid approach was also used to calculate accessibility to transit by finding the nearest bus, metro and rail line stops to each cell and summing each line’s closest stop’s contribution to a transit accessibility index; a stop closer to a cell centroid or with a smaller headway (calculated using AM peak) would mean a larger contribution to transit accessibility. This is calculated as: \(PT_m = \sum_{k=1}^{n} [d_{km} \times h_k]^{-1}\), where \(PT_m\) denotes accessibility to public transit at cell \(m\), \(d_{km}\) stands for distance, in km, from cell centroid \(m\) to nearest bus stop of line \(k\) (minimum
value of 0.1 km) and $h$ stands for average headway, in hours, of line $k$ in AM peak (maximum value of 1 hour).

Note that these three indicators have been used in our previous research. Other BE indicators can be included (e.g. employment density); however, high correlation among them becomes an issue. Moreover, it is important to mention that these indicators evolve over time across the three time periods.

**Model Setting**

To estimate the effect of BE and socio-demographics on GHGs, a Latent Class (LC) regression technique was used. For model definition, consider $J$ homogenous classes of households, with $j=(1,…J)$ and $J$ is to be defined as part of the calibration process. Then, within each class, we represent the GHGs outcome using a log-linear model with fixed temporal effects for the years, with $i$ representing a household ($i=1,…n$). Then equation 5 represents the log-linear GHGs produced by household $i$ belonging to the class $j$ ($j=1,…J$):

$$\ln(GHG_{ij}) = \beta_j x_i + \gamma_j z_i + \theta_j t_i + \eta_{ij}$$

Where:
- $\ln(GHG_{ij}) = $ natural logarithm of GHGs for household $i$ which belongs to class $j$
- $\beta_j, \gamma_j, \theta_j =$ model parameters (vectors) specific to regression model of class $j$
- $x_i =$ socio-demographics of household $i$ affecting household $i$ GHGs
- $z_i =$ BE attributes (population density, entropy and PT accessibility) in the vicinity of $i$
- $t_i =$ year fixed effect (the year of the OD the household data belongs to), with $t$ being a dummy variables
- $\eta_{ij} =$ random independent error term, normally (Gaussian) distributed for household $i$ and class $j$

Now to assign households to classes, the random utility based multinomial logit structure is used for the household segmentation model. The utility function for assigning a household $i$ to class $j$ is defined as:

$$U_{ij} = \alpha_j w_i + \varepsilon_{ij} \quad j = 1, ..., J$$

where $w_i =$ vector of household attributes that influences the propensity of belonging to class $j$. Also, $\alpha_j$ is the corresponding vector of coefficients and $\varepsilon_{ij}$ is the random error term assumed to be identically and independently distributed (with a Type 1 Extreme Valued error). Then, the probability that household $i$ belongs to class $j$ is determined by (27):

$$P_{ij} = \frac{\exp(\alpha_j w_i)}{\sum_{j=1}^{J} \exp(\alpha_j w_i)}$$

Then, the unconditional probability of a GHGs production at household $i$ is given as $f(GHG_i | $all parameters) = $\sum_{j=1}^{J} P_{ij} f(GHG_i | \beta_j, \gamma_j, \theta_j)$. In order to define the optimal number of classes, different numbers of classes are tested for the same model. The number of classes for the model with the lowest BIC (Bayesian Information Criterion) is selected. In LC regression the special case of 1-class corresponds to the homogeneous population assumption made in traditional regression – which means that population is not segmented. Moreover, the regression parameters are estimated simultaneously for classes and GHGs equations. The estimation is performed in LatentGold software (version 4.5).

**DATA**

This section describes the input data used in the latent class modeling. The main data components are O-D surveys, BE data and fuel consumption rates for the Montreal motor vehicle fleet.
O-D Surveys

A summary of the main household-level socio-demographic characteristics obtained from each O-D surveys is provided in Table 1. The average values are the raw estimates (without taking into account the expansion factors provided to reproduce the entire population of Montreal). Note that numbers are more or less consistent over time. However some variations are observed such as the decrease in the number of children and students.

Table 1. Summary statistics of socio-demographics, BE characteristics at the household

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Socio-demo</td>
<td>Number of cars</td>
<td>1.36</td>
<td>0.86</td>
<td>1.26</td>
<td>0.91</td>
<td>1.27</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>Number of persons</td>
<td>2.65</td>
<td>1.29</td>
<td>2.42</td>
<td>1.26</td>
<td>2.36</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>Number of children (younger than 15)</td>
<td>0.64</td>
<td>0.98</td>
<td>0.48</td>
<td>0.87</td>
<td>0.48</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>Number of fulltime workers</td>
<td>1.16</td>
<td>0.82</td>
<td>1.06</td>
<td>0.84</td>
<td>0.97</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>Number of part time workers</td>
<td>0.12</td>
<td>0.36</td>
<td>0.11</td>
<td>0.33</td>
<td>0.10</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>Number of students</td>
<td>0.67</td>
<td>0.96</td>
<td>0.58</td>
<td>0.92</td>
<td>0.52</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>Number of retirees</td>
<td>0.26</td>
<td>0.58</td>
<td>0.31</td>
<td>0.62</td>
<td>0.46</td>
<td>0.70</td>
</tr>
<tr>
<td>BE</td>
<td>Population density * (people per hectare)</td>
<td>42.77</td>
<td>31.91</td>
<td>47.04</td>
<td>34.36</td>
<td>47.57</td>
<td>34.30</td>
</tr>
<tr>
<td></td>
<td>Transit accessibility *</td>
<td>108.48</td>
<td>119.00</td>
<td>120.87</td>
<td>126.44</td>
<td>121.43</td>
<td>123.85</td>
</tr>
<tr>
<td></td>
<td>Land use mix (entropy)*</td>
<td>0.32</td>
<td>0.16</td>
<td>0.343</td>
<td>0.1740</td>
<td>0.337</td>
<td>0.17</td>
</tr>
</tbody>
</table>

*In the vicinity of each household (a buffer of nine-cell grids, 500m by 500m each)

It is also important to mention that in order to make the 3 waves comparable in terms of area used; the O-D survey region of the survey 1998 was used as the reference region. Then, households of the O-Ds 2003 and 2008 located outside the 1998 O-D survey were excluded. Moreover, trips with destinations outside of region were considered, however, only the emissions associated to the part of the trip done in the region were included in the inventory. The inventory is then a weekday inventory of travel in the region where the day of the trips is also recorded.

Vehicle Fleet Inventory

The original data from the Quebec provincial vehicle inventory comes from SAAQ. This data was then treated by CDAT (Centre for Data and Analysis in Transportation-Laval University) in order to obtain fuel consumption rates according to the make, year and model. The distribution of the FCR across the 3 waves is shown in Figure 1. It is worth noting that the fuel economy of the fleet has been increasing over the 10 year period. The average FCR for the year 2008 is 9.19 lit/100km, 9.36 lit/100km for 2003 and 9.57 lit/100km for the year 1998.
BE Indicators

Data used for the BE attributes include: Land use shape files that were obtained from Desktop Mapping Technologies Inc. (DMTI) and have been described above. The population density data come from the Statistics Canada Censuses for the year 1996, 2001 and 2006. With respect to public transit accessibility, geo-coded transit lines and stops tagged with unique identifiers linking them to weekday AM-peak headways were used.

Table 1 shows the summary statistics of the BE indicators in the vicinity of the households across each O-D survey. Note that transit accessibility and population density increase over time. Unsurprisingly, land-use mix stays very stable. The maps in Figure 2 show population density, land use mix and transit accessibility, in Montreal for the year 2003. For the other years of OD surveys (1998 and 2008) similar maps have been generated, but not reported here due to lack of space. Each household is assigned a value for each of these indicators based on the grid cell in which their dwelling is located and the year of OD survey in which they were observed.

RESULTS

This section starts by introducing the emission inventory results. Then, the outcomes of the LC model are provided with a discussion

Household travel distance and emission inventory

Summary statistics of travel distances and emissions, averaged at a 500-meter grid level, is presented in Table 2.

From table 2 one can see that both the distance and GHG emissions are decreasing over the 10 year period. Also as mentioned in figure 1, as a result of the introduction of increasingly fuel efficient cars over the years, the average FCR has also declined. As such, a decrease in GHGs is observed, not only as a result of decreasing travel distances, but also greener cars (lower FCR).
Table 2. Summary of expanded GHG, distance traveled at the household level*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Average total (car+transit) GHG (kg/day)</td>
<td>10.70</td>
<td>10.32</td>
<td>10.02</td>
</tr>
<tr>
<td>Average car GHG (kg/day)</td>
<td>10.25</td>
<td>10.00</td>
<td>9.61</td>
</tr>
<tr>
<td>Average total distance at household (km/day)</td>
<td>53.65</td>
<td>51.08</td>
<td>47.82</td>
</tr>
<tr>
<td>Average car distance at household (km/day)</td>
<td>45.23</td>
<td>43.68</td>
<td>39.66</td>
</tr>
<tr>
<td>Number of households in OD survey</td>
<td>53,810</td>
<td>56,959</td>
<td>66,124</td>
</tr>
<tr>
<td>Average expansion factor</td>
<td>21.16</td>
<td>26.15</td>
<td>24.98</td>
</tr>
</tbody>
</table>

These values are expanded using the expansion factors for each household, to represent a picture of the total population of the area.

The spatial distribution of GHGs at the household level is represented in Figure 3. The map represents average emissions for total household travel GHGs, for all households falling inside each grid cell. From this figure it can be seen that the central neighborhoods emit the least, and as one goes towards the suburbs...
the GHG footprint of the households increases. This can be explained by the relative increase in distance traveled by these households (suburban) and their predominant use of the car.

FIGURE 3: Household GHG emission inventory for different OD years (average GHG, kg of CO2, at 500 by 500 meter grid)

OLS results

Log-linear (OLS) regression models are presented first. In this first attempt, BE attributes (residential density, PT accessibility and land use mix represented by the entropy index) were directly entered in the GHG and distance model as explanatory variables. Ln(GHG) and ln(distance) are the dependent variables.
Table 3 presents the results for the two models. The regression models were estimated take into account the expansion weights corresponding to each household.

**TABLE 3: OLS model for Ln (household trip GHG & car distance traveled)**

<table>
<thead>
<tr>
<th></th>
<th>Model 1 –GHG</th>
<th></th>
<th>Model 2 – Car distance</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>P&gt;</td>
<td>t</td>
<td></td>
</tr>
<tr>
<td>Residential density *</td>
<td>-0.004</td>
<td>0.00</td>
<td>-2.08</td>
<td>-0.004</td>
</tr>
<tr>
<td>PT accessibility *</td>
<td>-0.001</td>
<td>0.00</td>
<td>-1.56</td>
<td>-0.001</td>
</tr>
<tr>
<td>Entropy *</td>
<td>-0.882</td>
<td>0.00</td>
<td>-2.97</td>
<td>-1.083</td>
</tr>
<tr>
<td>Year 1998</td>
<td>0.140</td>
<td>0.00</td>
<td>15.03</td>
<td>0.146</td>
</tr>
<tr>
<td>Year 2003</td>
<td>0.096</td>
<td>0.00</td>
<td>10.18</td>
<td>0.220</td>
</tr>
<tr>
<td>Year 2008</td>
<td>Base case</td>
<td></td>
<td></td>
<td>Base case</td>
</tr>
<tr>
<td>Number of retirees **</td>
<td>-0.025</td>
<td>0.00</td>
<td>-2.52</td>
<td>-0.016</td>
</tr>
<tr>
<td>Number of students **</td>
<td>0.078</td>
<td>0.00</td>
<td>7.88</td>
<td>0.027</td>
</tr>
<tr>
<td>Number of part time workers **</td>
<td>0.311</td>
<td>0.00</td>
<td>31.13</td>
<td>0.385</td>
</tr>
<tr>
<td>Number of fulltime workers **</td>
<td>0.510</td>
<td>0.00</td>
<td>51.07</td>
<td>0.621</td>
</tr>
<tr>
<td>Number of children **</td>
<td>0.012</td>
<td>0.00</td>
<td>1.27</td>
<td>0.075</td>
</tr>
<tr>
<td>Single adult family</td>
<td>-0.385</td>
<td>0.00</td>
<td>-31.98</td>
<td>-0.604</td>
</tr>
<tr>
<td>Low income (less than 40k)</td>
<td>-0.557</td>
<td>0.00</td>
<td>-42.72</td>
<td>-0.629</td>
</tr>
<tr>
<td>Medium income (40k to 80k)</td>
<td>-0.219</td>
<td>0.00</td>
<td>-19.70</td>
<td>-0.205</td>
</tr>
<tr>
<td>High income (more than 80k)</td>
<td>Base case</td>
<td></td>
<td></td>
<td>Base case</td>
</tr>
<tr>
<td>Constant</td>
<td>2.008</td>
<td>0.00</td>
<td>-</td>
<td>3.041</td>
</tr>
</tbody>
</table>

*(10% increase for elasticity)*
***(1 unit increase for elasticity)**

From table 3 we can see that BE variables (population density, PT accessibility and LU mix) are statistically significant and negatively associated to both household travel GHGs and distance traveled. From the elasticities we can observe that increasing population density and PT accessibility by 10% (one at a time) would cause 2.08% and 1.56% reduction in the households' GHG and 2.20% and 2.12% in the distance traveled respectively. For land use mix (entropy index), elasticities are slightly higher, -2.97% and -3.64 for GHGs and distance respectively. Although distance elasticities are slightly higher than GHGs elasticities, both outcomes are consistent. These results are in accordance with the literature, in terms of sign and significance; however the magnitude of these parameters seems to be slightly greater than most of the past North American studies. This may be linked to the fact that the Montreal region has a higher population density and better transit supply than many US cities on which studies reported in the literature are based.

Another interesting observation is captured by looking at the fixed effect for the year of the OD survey. The positive and decreasing elasticities for these variables suggest that both GHG and distance traveled have a declining trend over the 10 year period. More precisely, keeping everything to the base values, the GHG for year 1998 and 2003 are by 15% and 10% higher than 2008, respectively. Distance is 15.8% and 24.7% higher than 2008 for 2003 and 1998, respectively. Since GHGs have been decreasing at a faster rate than distances, decreases in GHG appear partly explained by better fuel economy of the automobile fleet over the 10 year period (Fig 1).
With respect to employment, different employment status variables are statistically significant. Increasing the number of full time or part time workers by one unit causes about 51% (62%) and 31% (38%) increase in the total household trip GHG (or distance). This shows the important link between labor force participation and transportation-related GHGs and distance traveled at the household level. The single adult family variable (household with only one member which is adult) is also found to be statistically significant. This type of household has a much smaller (32% less in model 1 and 45% in model 2) carbon footprint and distance traveled comparing to households with more than one member.

**Latent class results**

For the definition of the best number of classes (subgroups) in the LC regression analysis, different numbers of classes were attempted. The BIC values were then used as a goodness of fit (selection) criterion. The aim was to identify the model with the lowest BIC and with a reasonable number of observations in each class. Also, for the regression models (for both GHG and classes equations) different combinations of variables were tested. Particular attention was paid to avoid high correlation among explanatory variables. The regression models were estimated take into account the expansion weights corresponding to each household.

After trying different numbers of sub-groups (classes) and variable combinations, a 3-class model was selected and is reported in Table 4. In the GHG model results, in addition to the built environment variables (population density, land use mix and transit accessibility), several socio-demographics and year fixed effects are statistically significant. In the class model, 3 variables were retained as covariates to assign households to classes. These are the number of cars, number of children, and household size (number of persons in the household) – these are also reported in the second part of Table 4.

The following observations can be made about the three classes:

- Class 1 has the lowest car ownership, number of children and people among the 3 classes. It mainly consists of households on the island of Montreal (84%) and towards central areas.
- Class 2 has intermediate values for these attributes (higher than class 1 but lower than class 3) and is mostly located between the other two mentioned classes geographically.
- Class 3 has the highest value for these attributes (the highest number of cars, children and people) and is mostly located outside the island of Montreal in suburbs.

For each class, the GHG regression model parameters are presented in Table 4. Also, the models for each of the classes are presented at the end of this table. Among other results, it is observed that:

- The three built environment factors are found to be significant at 0.05 significant levels. As was the case in the previous OLS models, BE attributes have a negative effect on GHG and car distance traveled.
- More important, the size of the parameters is statistically different across household types (subpopulations). This suggests that the effectiveness of built environment policies may vary importantly across household classes.
### TABLE 4: LC regression model for Ln (household trip GHG & car distance traveled)

<table>
<thead>
<tr>
<th></th>
<th>Model 3 – GHG</th>
<th>Class 1</th>
<th>Elast%</th>
<th>Class 2</th>
<th>Elast%</th>
<th>Class 3</th>
<th>Elast%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential density (in 1000s)*</td>
<td>-2.1</td>
<td>-0.95</td>
<td>-2.4</td>
<td>-1.09</td>
<td>-2.9</td>
<td>-1.32</td>
<td></td>
</tr>
<tr>
<td>PT accessibility (in 1000s)*</td>
<td>-0.1</td>
<td>-0.12</td>
<td>-0.9</td>
<td>-1.05</td>
<td>-0.8</td>
<td>-0.94</td>
<td></td>
</tr>
<tr>
<td>Entropy *</td>
<td>-0.381</td>
<td>-1.29</td>
<td>-0.540</td>
<td>-1.83</td>
<td>-0.561</td>
<td>-1.9</td>
<td></td>
</tr>
<tr>
<td>Year 1998</td>
<td>-0.256</td>
<td>-22.65</td>
<td>0.339</td>
<td>40.4</td>
<td>0.178</td>
<td>19.5</td>
<td></td>
</tr>
<tr>
<td>Year 2003</td>
<td>-0.182</td>
<td>-16.7</td>
<td>0.224</td>
<td>25.2</td>
<td>0.145</td>
<td>15.7</td>
<td></td>
</tr>
<tr>
<td>Year 2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of retirees **</td>
<td>0.079</td>
<td>7.93</td>
<td>-0.229</td>
<td>-22.9</td>
<td>-0.010</td>
<td>-1.02</td>
<td></td>
</tr>
<tr>
<td>Number of students **</td>
<td>0.052</td>
<td>5.27</td>
<td>0.03</td>
<td>3</td>
<td>0.077</td>
<td>7.8</td>
<td></td>
</tr>
<tr>
<td>Number of part time workers **</td>
<td>0.062</td>
<td>6.23</td>
<td>0.1663</td>
<td>16.6</td>
<td>0.116</td>
<td>11.6</td>
<td></td>
</tr>
<tr>
<td>Number of fulltime workers **</td>
<td>0.076</td>
<td>7.62</td>
<td>0.3626</td>
<td>36.2</td>
<td>0.175</td>
<td>17.55</td>
<td></td>
</tr>
<tr>
<td>Number of children **</td>
<td>-0.024</td>
<td>-2.41</td>
<td>-0.107</td>
<td>-10.7</td>
<td>-0.067</td>
<td>-6.76</td>
<td></td>
</tr>
<tr>
<td>Single adult family</td>
<td>-0.095</td>
<td>-9.11</td>
<td>-0.0944</td>
<td>-9.01</td>
<td>-0.137</td>
<td>-12.8</td>
<td></td>
</tr>
<tr>
<td>Low income (less than 40k)</td>
<td>-0.062</td>
<td>-6.04</td>
<td>-0.384</td>
<td>-31.9</td>
<td>-0.202</td>
<td>-18.3</td>
<td></td>
</tr>
<tr>
<td>Medium income (40k to 80k)</td>
<td>-0.025</td>
<td>-2.56</td>
<td>-0.1454</td>
<td>-13.5</td>
<td>-0.094</td>
<td>-9.03</td>
<td></td>
</tr>
<tr>
<td>High income (more than 80k)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.265</td>
<td>-1.618</td>
<td>-2.867</td>
<td>-0.478</td>
<td>-2.721</td>
<td>-3.673</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Model 4 – Car distance</th>
<th>Class 1</th>
<th>Elast%</th>
<th>Class 2</th>
<th>Elast%</th>
<th>Class 3</th>
<th>Elast%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential density (in 1000s)*</td>
<td>-0.6</td>
<td>-0.27</td>
<td>-1.9</td>
<td>-0.86</td>
<td>-3.7</td>
<td>-1.7</td>
<td></td>
</tr>
<tr>
<td>PT accessibility (in 1000s)*</td>
<td>-0.3</td>
<td>-0.35</td>
<td>-1.1</td>
<td>-1.3</td>
<td>-1.1</td>
<td>-1.3</td>
<td></td>
</tr>
<tr>
<td>Entropy *</td>
<td>-0.146</td>
<td>-0.5</td>
<td>-0.639</td>
<td>-2.16</td>
<td>-0.668</td>
<td>-2.3</td>
<td></td>
</tr>
<tr>
<td>Year 1998</td>
<td>0.146</td>
<td>15.7</td>
<td>0.415</td>
<td>51.6</td>
<td>0.104</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Year 2003</td>
<td>0.077</td>
<td>8.1</td>
<td>0.364</td>
<td>44</td>
<td>0.220</td>
<td>24.6</td>
<td></td>
</tr>
<tr>
<td>Year 2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of retirees **</td>
<td>-0.023</td>
<td>-2.3</td>
<td>-0.161</td>
<td>-16.1</td>
<td>-0.105</td>
<td>-10.6</td>
<td></td>
</tr>
<tr>
<td>Number of students **</td>
<td>-0.003</td>
<td>-0.38</td>
<td>-0.037</td>
<td>-3.7</td>
<td>0.071</td>
<td>7.1</td>
<td></td>
</tr>
<tr>
<td>Number of part time workers **</td>
<td>0.038</td>
<td>3.82</td>
<td>0.164</td>
<td>16.5</td>
<td>0.213</td>
<td>21.3</td>
<td></td>
</tr>
<tr>
<td>Number of fulltime workers **</td>
<td>0.078</td>
<td>7.9</td>
<td>0.322</td>
<td>32.3</td>
<td>0.342</td>
<td>34.2</td>
<td></td>
</tr>
<tr>
<td>Number of children **</td>
<td>-0.061</td>
<td>-6</td>
<td>-0.203</td>
<td>-18.4</td>
<td>-0.891</td>
<td>-59</td>
<td></td>
</tr>
<tr>
<td>Single adult family</td>
<td>-0.045</td>
<td>-4.47</td>
<td>-0.194</td>
<td>-17.6</td>
<td>-0.333</td>
<td>-28.4</td>
<td></td>
</tr>
<tr>
<td>Low income (less than 40k)</td>
<td>-0.008</td>
<td>-0.81</td>
<td>-0.010</td>
<td>-1.04</td>
<td>-0.119</td>
<td>-11.2</td>
<td></td>
</tr>
<tr>
<td>Medium income (40k to 80k)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High income (more than 80k)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.478</td>
<td>-2.721</td>
<td>-3.673</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

*(10% increase for elasticity)*

**(1 unit increase for elasticity)
Note that the effects of population density and entropy (LU mix) are the highest for class 3. For instance, the impact of density on distance travel is -1.32% and -1.7% for GHGs. This could be due to the fact that class 3 is mainly located on the peripheries (with lower density and LU mix) and therefore having more potential for increase in density and entropy.

For PT accessibility elasticities, class 2 and 3 are very similar for GHGs and the same for distance. This suggests that these household classes, mainly located in new and old suburbs, can react differently to transit improvement strategies than central areas.

From the year-fixed effects, one can see different trends for each class type. For households in classes 2 and 3, the emissions are significantly higher in 1998 and 2003 with respect to 2008. A similar pattern is observed for car distance results (a clear decrease of the distance travel).

However, GHGs in class 1 show an inverse effect, i.e., an increase of GHG emissions over time despite a decrease in distance travel. As mentioned before, class 1 is mainly located in central neighborhood with the lowest values for car ownership. To explore why this is happening, the fuel consumption and average travel speed that affect GHGs were analyzed further. For this, we plotted the FCR (fuel consumption rate) histograms of each class for each OD year. Although the overall FCR for all the data tends to decline over time, in this subpopulation of households, one can observe an inverse effect, FCR is increasing over time. This could be associated to different factors, for instance, households belonging to this class have the lowest distance traveled and therefore they may tend to keep their auto-vehicles over longer time periods. This might cause the relative FCR of their cars to increase. This could also have to do with the fact that these households are becoming richer and getting bigger cars – gentrification has been bringing wealthier people downtown and in rich central neighborhoods.

Moreover, employment status of the household plays an important role in their travel related distance and GHG. More precisely, by adding one fulltime and part time worker to the household, there is an increase of 36% and 16.6% in GHG, in class 2. The single adult family variable has a smaller (13% less in model 3 for class 3, and 59% less in model 4 for class 3) carbon footprint and car distance comparing to households with more than one member. Income also plays an important role in the carbon footprint and car distance travel of the household with low and medium class households having less contribution comparing to the high income class households. This is as much as 32% less for GHG and 28% less for car distance in the low income class households.

CONCLUSIONS

Using 3 OD surveys, or waves of travel behavior data, this research has aimed to investigate the potential impact of built environment on GHGs and car distance traveled at the household level in the region of Montreal, Canada. A household inventory was first determined for the 3 O-D surveys covering a period of 10-year period (1998, 2003, and 2008). A temporal and spatial exploratory analysis was first implemented; then, an LC regression modeling approach was adopted that divides the dataset into household subgroups (classes) and was compared to simple OLS. Among other results, it was found that land use mix, population density and public transit accessibility have statistically significant and negative effects on the carbon footprint and car distance travel of households. This is in accordance with the literature; however, these values are slightly greater than those obtained in past studies involving US & Canadian cities. Moreover, the effects of built environment indicators vary importantly among subgroups (classes) and this is very important from a policy making point of view. It has been shown that different household types are likely to respond to policy initiatives differently, suggesting that it is important target the right policy initiatives to the right population subgroups.
From the exploratory analysis and fixed-effect model outcomes, it is observed that the overall trend of GHGs has declined over the 10 year period considered. However, when looking at different population classes, we observe the opposite for the more central, lower car ownership class (class 1). As explained in the results section, this is due to the fact that auto-vehicles in this class are becoming less fuel efficient.

Employment status and income are also significantly related to household trip GHG emissions; having more full-time and part-time workers in the household and a higher income adds to the GHG contribution and car travel distance of that household. This is consistent with the literature and shows that a large share of GHG is due to everyday commutes. Therefore, GHG reduction policies should target commuter trips and higher income households if they wish to maximize the effectiveness of their efforts.

Some limitations of this work are as follows: a) for the FCR we’ve used the rate at the FSA level instead of the real household’s fleet FCR. This was due to the limit of data available; b) including other socio-demographic characteristics (such as age, gender and etc.) was not possible because of aggregating at the household level.

As a future work, the effect of green technologies could be predicted and compared to BE policies. Also other modeling frameworks such as copula and Bayesian structural equations models could be tested.

ACKNOWLEDGMENT

We would like to acknowledge the financial aid provided by FQRNT under the program “Recherche partenariat contribuant réduction et séquestration gaz effet de serre” and thank AMT and MTQ for providing us with the data necessary for this research.

REFERENCES