CIRRELT

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

Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation

Modeling the Effect of Land Use on Activity Spaces

Christopher Harding
Zachary Patterson
Luis F. Miranda-Moreno
Seyed Amir H. Zahabi

January 2013

CIRRELT-2013-02

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

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

www.cirrelt.ca

Modeling the Effect of Land Use on Activity Spaces

Christopher Harding[1], Zachary Patterson[1,2,*], Luis F. Miranda-Moreno[2,3], Seyed Amir H. Zahabi[3]

[1] Department of Geography, Planning and Environment, Concordia University, 1455 de Maisonneuve W., H 1255-26 (Hall Building), Montreal, Canada H3G 1M8

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

[3] Department of Civil Engineering and Applied Mechanics, McGill University, 817 Sherbrooke Street West, Room 268, Montréal, Canada H3A 2K6

Abstract. Historically, when analyzing the effect of land-use on travel demand, research has concentrated on a few key indicators, notably mode choice, VMT and number of trips. At the same time, this literature has primarily focused on the effects of individual land-use variables: e.g. what is the effect of land-use mix or population density on mode choice. It is becoming increasingly clear however that the isolated impact of particular measures of land-use on individual and household transportation behavior is small, but that when dealt with using a clustered approach, their combined influence becomes both less ambiguous in direction and greater in magnitude. This paper contributes to the transportation and land-use literature by examining the effect of clusters of land-use indicators on activity spaces, an emerging but traditionally ignored, transportation behavior indicator. Regression analysis results point to a significant relationship between large and dispersed activity spaces, low levels of population and employment density, and low levels of public transit accessibility and land use mix.

Keywords. Activity space, clusters, land-use, travel demand, dispersion, convex hull, minimum convex polygon.

Acknowledgements. We acknowledge the financial support received from the Fonds de recherche du Québec - Nature et technologies (FRQNT), under the program “Recherche en partenariat contribuant à la réduction et la séquestration des gaz à effet de serre”.

---

* Corresponding author: Zachary.Patterson@cirrelt.ca

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2013

INTRODUCTION

Research on the effects of land-use on transportation has historically concentrated on a few key indicators, notably mode choice, VMT and number of trips. The focus of such research has also overwhelming been concerned with the effects of individual land-use variables: e.g. what is the effect of public transit accessibility or residential density on distances travelled. Recent literature has however brought to light that when modeled using a clustered approach, which typifies areas based on combinations of land-use variables, as opposed to dealing with them individually, their combined influence on individual and household transportation behavior is less ambiguous in direction and greater in magnitude.

In line with such findings and using the Metropolitan region of Montreal as an application environment, this paper examines the effect of clusters of land-use indicators on activity spaces, an emerging but traditionally ignored, transportation behavior indicator.

The paper begins with a review of the literature on land use variables and travel behavior, followed by a summary of the work on clustering, and finally that which pertains to activity spaces. The data used for this paper is described, as well as the ways in which it was employed to quantify the impact of land use variables on activity spaces. Regression model results and data analysis follow, and the paper concludes with a summary of key findings and suggestions for future research.

LITERATURE REVIEW

The following literature review outlines the different approaches taken to measuring the effect of land use variables on transportation behavior, both individually and as clusters, and ends with the material related to activity spaces.

Traditional Land-use and Travel Behavior Literature

The traditional approach to linking land use variables to transportation behavior looks to the levels of either mix or density and links these to common measures of travel activity such as vehicle kilometers travelled (VKT), vehicle hours travelled (VHT), number of trips and mode choice. Ewing and Cervero's (*1*) (*2*) seminal works looked at this body of literature in both 2001 and 2010, highlighting the links found between different indicators and travel behavior. They point out those with the strongest correlation, but also highlight areas where links have proven either difficult to quantify or demonstrate as significant.

Travel behavior variables are usually broken down into categories for individual, household and built form characteristics. Commonly used individual variables include gender, age, income (*3*) and education (*4*), whereas household variables, or indicators, commonly used are number of persons or children per household (the latter acting as a proxy for stage in the life-cycle) (*5*), income and number of vehicles owned (*6*).

Built form characteristics can also be divided into a few categories. There is widespread agreement within the literature that the three Ds proposed by Cervero and Kockelman act as the basic categories of urban form indicators, notably density, diversity and design (*7*) (*6*). One can find residential and employment densities quantified as simple measures of individuals per unit area (*8*) or retail employment per area (*4*), but more elaborate methods are also employed. Many papers outline different ways to address public transit accessibility, dealing with it as proximity to stations or bus stops (*6*), rail and bus line coverage (*3*), headway (*9*), etc.

Literature on clustering of urban form and public transit variables

More recent literature in the field deals with the effect of multiple land-use variables on transportation behavior through clusters, or neighborhood typologies.

In the literature which links specific urban form characteristics to travel behavior, three distinct problems are encountered, namely that of biased elasticities (*3*) (*4*), results which are not statistically significant (*10*) (*6*) and issues of causation or self-selection (*11*) (*9*). Neighborhood typologies, combined with household-level control variables, enable researchers to deal with urban form attributes while circumventing issues of biased coefficients, statistical significance and causation (*9*) (*5*) (*6*) (*8*) (*12*).

Measuring levels of the three Ds is a common approach to linking travel behavior to land use, however, authors such as Krizek have argued that interpreting such measures individually disregards the inherent relationships which exist between them (*7*). By combining indicators, one can better describe activity density (*13*) and more clearly understand the effect that changing levels of urban form and public transit can have (*3*). Techniques such as k-means clustering (*14*) can be employed to define these typologies and, when combined with control variables such as income or life-cycle characteristics, aid in building more accurate models for predicting travel demand.

These clusters and neighborhood typologies can be built in different ways, with or without the use of weights, and can include any indicator one finds pertinent, be it population or employment density, street grid connectivity, sidewalk provision, transit availability, etc. See Gershoff, Pederson, and Aber (2009), Lin and Long (2008), Manaugh, Miranda-Moreno, and El-Geneidy (2010) or Shay and Khattak (2007) for an overview of different techniques and indicators used.

Activity Space Literature

Activity spaces can be used to represent the areas individuals or households interact with as they travel (*15*). They can be used to measure either access to certain resources or the spread of activities throughout space, bringing a new dimension to travel demand modeling. Created using standard deviational ellipses (SDE), minimum bounding geometry or other means (*16*) (*17*), these spaces have been employed in fields as varied as criminology (*18*), transit planning (*13*), nutrition exposure (*15*) and healthcare (*19*). Different types of data have also been used to generate them, some accounting only for routine activities based on interviews (*19*), others using travel diaries (*13*).

Fan and Khattak for example used the indicators of building density, retail accessibility and street grid connectivity to quantify the impact of land use variables on individual spatial footprints and found that downtown residents generated smaller spaces than their suburban counterparts (*20*) (*21*). Smaller activity spaces are commonly viewed as beneficial from an energy and environmental perspective (*22*) (*23*); this is also true from a health (*24*) (*25*) and economic perspective (*26*). Activity spaces can therefore aid in developing policy to guide cities towards more sustainable mobility futures.

The idea of moving from traditional transportation demand measures to activity spaces is supported by a growing recognition of the importance of non-commuting trips to the total travel of households (*16*). The link between land use characteristics and distances travelled has already been investigated by many scholars, but a strong body of literature on the relationship between urban form, transit accessibility and activity spaces is not yet available. This paper will begin to fill that void by demonstrating the effect clustered indicators can have on activity spaces.

STUDY AREA AND DATA USED

The methods proposed in this paper are applied to the greater Montreal region of Quebec, Canada. Montreal is the second largest Census Metropolitan Area in Canada with a population of more than 3.6 million inhabitants in the latest (2006) census. It is an old city by North-American standards, characterized by an urban form built up over many phases. It also has a varied housing stock and a heterogeneous mix of transportation options, offering both heavy and commuter rail, and extensive bus service in addition to a well-developed highway network (see Figure 1). This heterogeneity in urban form and transit accessibility creates the landscape which makes Montreal a perfect case study for the effects of urban form on travel demand.

Seven different sources of data were required for this analysis, which builds upon the methodological approach of Miranda-Moreno et al. (*9*) : census tract (CT) population and employment counts, demographic characteristics, land use data, public transit data, personal and household mobility data, and finally CT shapefiles.

Land use shapefiles were obtained from Desktop Mapping Technologies Inc. (DMTI), a recognized GIS content provider. DMTI categorizes land use into seven categories, including water, open areas, residential, commercial, governmental and institutional, industrial and parks and recreation. Census tract shapefiles were in turn obtained from Statistics Canada's Census Tract Digital Boundary Files (*27*). These boundaries, as well as those for land use data, were used to delimit the study area.

With respect to public transit, geocoded transit lines and stops tagged with unique identifiers linking them to weekday AM-peak headways were used. The transit network used as the source for this information is a hybrid network. Its base comes from an existing TransCAD transit network of the Island of Montreal created in 2003 by Dr. Murtaza Haider of Ryerson University. The development of this digital network was supported by a grant from the National Sciences and Engineering Research Council (NSERC) as well as infrastructure provided by the Canada Foundation for Innovation (CFI). Transit lines off the island were added to the existing network in the summer of 2011. Both parts of the network were geocoded by hand since network information (property of five main transit operators) is not generally available outside of those institutions. That said, while the networks have changed over time, their main characteristics have remained similar. The principal difference between the 2003 and updated network was the addition of three metro stations on the island of Laval.

For household mobility data, the Montreal 2003 and 2008 Origin Destination surveys (OD), which are comprehensive travel demand surveys carried out every 5 years by the Agence Métropolitaine de Transport (AMT) – Montreal's public transport planning agency- were used. Montreal's OD surveys contain data on approximately 5% of the households in the study area, collecting time, mode and motive specific travel descriptions, as well as origin and destination XY coordinates for all trips carried out by persons aged over 4 years. They also collect household and personal characteristics for the individuals in each surveyed household.

It should be noted that in 2003, household domicile coordinates were coded as the XY coordinates of the actual home, whereas in 2008, the domicile coordinates were instead entered as the XY coordinates of the dissemination area (census subdivision smaller than a CT) within which the household was found. To ensure compatibility between datasets, we recoded the 2003 home-based trips to indicate the dissemination area centroid as opposed to the domicile. Data concerning 56,965 households in 2003 and 66,124 in 2008 were used in this analysis (*28*).

A grid consisting of cells 500 meters wide, as well as a nine-cell grid encompassing the host and references to the eight surrounding cells was also used; the latter to average indicator values over a larger area, avoiding peaks.

Census level data was acquired via StatsCan's E-Stat website, which provides information at the CT level regarding both the socio-demographics of populations– including average income and education attainment, employment sector activity, etc.-, as well as built form – including building type, age, condition- and other variables (*29*) (*30*). Employment data was obtained through the 2001 and 2006 "Enquête sur le travail et le milieu de travail et les employés," produced by Statistics Canada (*31*) (*32*). This was provided by Statistics Canada as a 'special order' from a consortium of provincial government ministries and agencies. Statistics Canada uses census information to infer employment information (number of jobs by NAICS sector by CT).

METHODOLOGY

This section provides a description of the generation of clusters from the four selected indicators, the calculation of activity spaces and statistical methods employed to estimate the effect of land-use on activity space.

Whereas much of the research previously published has made use of aggregated data for their analyses, either at the transportation analysis zone or CT level, this paper uses highly disaggregate data for cluster analysis. For example, data such as population and employment may be obtained at the CT level, but by isolating land uses which could contain them and calculating their density after an adjustment to area, much more accurate information on the locations and densities of indicators is obtained.

Previous work on clusters has looked at population density, land use entropy, public transit accessibility (*9*), urban design (*7*) (*5*) and other variables, and research by Leck (2006), Bento (2005) and Ewing and Cervero (2010) has demonstrated that employment density is an important predictor of travel demand. As such, clusters were designed to incorporate the following four indicators: population and employments densities, land use mix and public transit accessibility.

Densities

For all calculations involving residential density, only residential land use area was used, likewise for employment density, only commercial, government and institutional, and resource and industrial land use areas were used; to obtain the most accurate information 'net' and not 'gross' density was employed. These net-density employment and residential-only census tract polygons were then intersected with the grid previously described.

Land use mix

A similar process was used in the calculation for land use mix, also at the cell level, where an entropy index was devised based on that of Miranda-Moreno et al. (see equation 1). The more land uses there are in a cell and the more evenly their areas are distributed, the higher the value; its range is 0 (no mix) to 1 (perfect heterogeneity).

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

Where:

A_{ij} : area of land use *i* in cell *j*

D_j : area of cell (excluding water and open area)

n : total number of different land uses

Public transit accessibility

The grid approach was used to calculate the accessibility of cells 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 a smaller headway (calculated using AM peak) would mean a larger contribution to transit accessibility (see equation 2).

$$PTaccess_j = \sum_{i=1}^{n} \frac{1}{(d_{ij} * h_i)} \qquad \text{Eqn. (2)}$$

Where:

$PTaccess_j$: accessibility to public transit at cell *j*

d_{ij} : distance, in km, from cell centroid *j* to nearest bus stop of line *i* (minimum value of 0.1 km)

h_i : average headway, in hours, of line *i* in AM peak (maximum value of 1 hour)

All four indicator values were averaged with those contained in the eight surrounding cells. There are particular ways in which incomplete cells near bodies of water or the boundaries of the study were dealt with, in addition to the weighing of cells that intersected partial land use tracts, but it is beyond the scope of this paper to describe these.

Neighborhood typology, or clustering

After compiling indicator values, k-means cluster analysis was employed to create the typology. Similar to the procedure outlined in Lin and Long, clusters were generated attempting to find a balance not only between predictive power and number of cases (households), but also using visual representations as a 'sanity' check (*5*). Such a verification of face validity was also used later on in the regression stage, combined with a review of the correlation matrix, to aid in determining which independent variables to include in the model.

The four cell values, for population and employment densities, public transit accessibility and land use mix were input in STATA, and to increase the relevance of clusters, only cells which contained OD survey households and at least one non-null value were kept. Excluding cells which contained only null values removed 6,125 of 17,601 cells, or 35%, from the exercise, but only 1% of the valid OD households. Of the remaining cells, 3,007 contained the dissemination area centroids of valid OD households for 2003 and 3,168 for 2008. Since the goal was to predict activity spaces, assigning clusters to areas which were uninhabited was deemed unnecessary.

Although the densest cluster contains very few cells (and only 1.2% of the total number of households in a 7 cluster approach), the large size of the dataset ensures this remains a significant number of cases. See Table 1 for the mean of urban form and public transit characteristics, as well as counts and percentages, for each cluster.

Clusters were reclassified to represent increasing levels of transit accessibility, land use mix and density. From cluster 1 to 2 and so on, the densities (measured in persons or jobs per hectare) increase rather significantly; cluster 7 has four times the mean population density and over 50 times the mean employment density as its transit-less rural counterpart, cluster 1 (see Table 1). Land use mix also increases significantly when one passes from the low value clusters (20% entropy value) to higher ones (60%), and transit increases almost exponentially, from 7 to 550 units. The transit indicator's values are unbounded, but in this case range from a low of 0, which indicates that no public transit stops are within a host cell's search radius, to a high of 775.

To be useful to planners, clusters must not only be significant in modeling travel demand, but must also provide clear and legible descriptions of the neighborhoods they represent. Based on the literature, limiting the generation to less than 10 clusters, was expected to produce a legible typology. The results and discussion sections describe two variations attempted and the problems encountered.

With mean population densities ranging from 19 to 29 persons per hectare, clusters 1 and 2 (see Table 1) could be considered, as Newman and Kenworthy would call them, automobile-oriented outer suburbs (*33*). Clusters 3 through 5, at 45 to 86 persons per hectare would be transit-oriented inner and middle suburbs, and clusters 6 and 7, at 86 to 96 persons per hectare, and much higher employment densities, would be pedestrian-oriented core suburbs (*33*) (see Figure 2 for a visual representation of their distribution). Land use mix, transit supply and employment density also reflect the typical definitions of such neighborhoods.

Activity spaces

With respect to activity spaces, there exist many different tools that one can use to describe the travel behavior of households (*19*). Given the type of data available (daily travel surveys), the convex hull minimum bounded geometry (CVH) was however the best fit; regressions were also run on models using the standard deviational ellipse (SDE), but R^2 values were found to be higher using the CVH, which also ensured all trip locations were accounted for. Because of the joint constraints of the OD survey being a one-day travel diary, and that of activity spaces requiring 3 unique points, household activity spaces were chosen over individual activity spaces. Previous research supports such an approach, household characteristics having been demonstrated to effect travel behavior in previous models (*16*).

The CVH polygons were generated using ArcMap 10. The first step was to isolate individuals whose trips were all performed within the study area, then to map their origins and destinations. Using the Minimum Bounding Geometry tool, convex hull polygons were generated around each household's origin and destination coordinates. Households whose trips only included one valid origin and destination pair were excluded from subsequent statistical analysis, having formed lines with no area as opposed to polygons. These were isolated by removing the CVH polygons with zero width (16,727 of 52,386 valid households in 2008 and 13,400 of 47,053 in 2003).

It should be noted that the prevalence of households with zero-width polygons was slightly higher in the dense urban clusters, where they account for 35% of cases, against 28% in the more sprawling suburban and rural clusters (2008 numbers). These polygons, however, also occur most often in smaller households (46% in households of 1 person and 41% in households with 2 persons, 2008 also), and these small households are more prevalent in dense urban clusters, where the mean household size is 2.27 as opposed to 3.22 in more rural clusters. Since the major influence is household, and not cluster,-based, it was determined results would be more accurate if the model were built without taking zero-width polygons into account.

Out of an awareness of the importance of household and life-cycle characteristics, over 25 different variables were run alongside clusters in the regression model; only the final set will be reported here. Since census tracts define areas which are "designed to be homogeneous with respect to population characteristics, economic status, and living conditions" (5 p. 741) (*8*), CT-level information was included by matching households to the census tracts in which they reside. This made up for the absence of socio-demographic information, such as income and employment, in the OD survey.

Distance to central business district (CBD) was considered, but as had been demonstrated in Shearmur (*34*), although the downtown core attracts high numbers of commuters, the concentrations of employment present in other centers, combined with the changing demographics of society (the increasing number of dual-income households for instance), would have required a much more elaborate model be devised.

RESULTS

Looking at Table 2, one can see that all signs for the reported coefficients carry face validity and only variables with significance levels above 95% were kept. The model's dependent variable is the logarithm of the area occupied by the CVH polygons of households which had more than one unique OD XY coordinate pair and performed both mandatory as well as non-mandatory trips - such an approach was also taken in Manaugh and El-Geneidy (*22*). This left us with 20,703 valid household polygons for analysis in 2008 and 20,413 in 2003.

As the model demonstrates, more urban clusters lead to consistently smaller activity spaces; the absolute value of the coefficients for cluster dummy variables become larger as cluster values increase (see Table 2). In addition, not only are the cluster binary variables statistically significant, but the confidence intervals for these variable coefficients exhibit no overlap. This is a clear indication that, not only are these clusters significant in improving the predictive power of the model, but they are also statistically significantly different one from another.

For example, were we to build a sample household using the mean observed values for each individual model variable and then move this *typical* household across clusters, the predicted activity spaces produced would vary from a high of 67.11 km^2 in the base case (cluster 1, or low density transit-less rural), to 35.42 km^2 in cluster 4 (outer suburb), all the way to a low of 14.72 km^2 in cluster 7 (dense, downtown core). These differences in predicted activity space are the result not of changing household demographics or tract level properties, but merely moving a hypothetical *typical* household from one cluster to another. The values mentioned above are bias-corrected for logarithmic back transformation using the technique described by Newman (*35*).

Model results indicate a significant link between clusters and activity space, and Figure 3 shows a 3D representation of this. In it, darker colors represent low cluster values and heights represent the average actual activity space at a given cell. The cells which appear flat on the map represent values for activity space below a certain threshold (for display purposes the heights are multiples of the square root of activity space). In contrast to the large dark peaks, many flat cells appear in the central portion of the island of Montreal and many low-height peaks are in white and light-grey (high-value clusters, or dense, highly mixed and well-served by transit areas).

The inclusion of F and G categories of employment (percentage of persons per CT working in occupations in art, culture, recreation and sport, and sales and service occupations) as a CT-level variable was based upon trial and error, but also previous work which found that these sectors were consistently overrepresented outside of employment centers (*34*); i.e. more dispersed, leading to smaller distances traveled on average to access work locations. The lower level of specialization within these sectors, means local workforces are more likely to fill these positions.

A variable whose predictive power and significance proved very high was "homemakers". This CT-level variable indicated the percentage of women aged 15 and over in a CT spending more than 15 hours a week performing unpaid child care. When tracts with high homemaker values were displayed in ArcMap, a pattern emerged where most were rural CTs and the remainder high average-income CTs. Rural populations would intuitively have to travel long distances to reach activities, while high incomes would justify one partner's ability to stay home tending to children, while the other partner (most likely working in a specialized field or occupying a managerial position, would need to travel long distances to commute to his or her high income position).

Number of trips was included in the model despite higher trip generation in rural and suburban clusters because their numbers were found to be more closely tied to household size than urban form and transit indicators. Conversely, household sizes in the more rural and suburban clusters are on average larger, and as such it would be expected that their activity spaces be larger across the board, but these households also contain more children, who, as Shay and Khattak describe, lead to increases in household size without adding drivers (*6*). As such they are unlikely to travel large distances for work or school, and by their influence on the time budget of adults, actually decrease average activity space (*14*).

"From an economic perspective, distance to work is conceptualized as a cost, and greater travel distances are associated with higher earnings (and/or lower residential costs)" (*34* p. 332), as such it was odd to find that the average income variable attempted in the model resulted in a very small coefficient. The aggregated nature of data may be one explanation, it having come from the CT as opposed to the household, but the fact that many high income tracts are found near the CBD is more likely the determinant factor.

High percentage of detached housing led to larger activity spaces and high rental-housing proportions led to smaller activity spaces, but these and many other CT-level variables were excluded from the model because they were not found to be statistically significant, possibly due to high collinearity. These housing indicators merely stand as poor proxies of urban form and transit characteristics, without taking into account the subtle variations that make the clusters more accurate.

DISCUSSION

Data analysis had the objectives of quantifying the relationship between clusters and travel behavior, and in particular activity spaces.

The number of clusters to include in the final model was not only based on face validity when looking at the maps produced by assigning clusters to cells, nor was it determined purely on the basis of regression results. It is important in any study of the effect of urban form on travel behavior to bear in mind that the goal is to provide planners with easy to interpret and apply templates for neighborhoods, not merely to increase statistical significance.

As such, 7 clusters were generated, but a look at the Percent on-island Montreal (dummy variable) column of Table 1 reveals an important point related to scales of analysis; over 98% of the households represented by clusters 3 through 7 are on the island of Montreal (as can also be seen in Figure 2), this despite on-island observations representing only 62.5% of the valid household population. When 6 clusters were generated, this geographic difference was even more pronounced, with 92% of households represented by clusters 2 through 6 being Montreal households. In essence, the difference that exists between the landscapes of Montreal and its surrounding areas is so large that bringing their urban form and transit characteristics together to generate clusters leads to an almost complete disappearance of the subtleties present off-island. Clustering still leads to intuitively consistent predictions, but it does not leave much room for off-island tracts to learn from on-island ones; the differences in urban form and transit accessibility being so stark between the two that off-island municipalities aiming to emulate characteristics of denser Montreal clusters to reduce excessive travel demand would face landscape redesign challenges worthy of Haussman's transformation of Paris. With respect to Montreal on the other hand, this reaffirms the vast differences which exist between geographically proximate, but dissimilar, neighborhood types.

An interesting notion to keep in mind when interpreting results is the concept of diminishing returns. As Krizek stated, once a certain level of service provision or density is exceeded, an increase in the number of businesses or transit stops may have negligible impact on travel behavior (*7*). This is reflected in Table 1, wherein the population density actually decreases between clusters 6 and 7; a non-linear relationship case in point for using clustered land use indicators as opposed to individual ones.

Another important point is that the measure for transit could still be refined, as infrequent regional bus stops, as well as commuter rail stops, exhibit high spatial correlation with large activity spaces. Future research should thus try to separate local transit from regional and express transit, which by their very nature carry people over large distances, but all the while create an upward bias to cluster cells near them.

To reiterate, looking to the regression results in Table 2, one can see that the influence of clusters was high and all the included coefficients were intuitive and right-sided: household licences, high homemaker CTs, more trips, full time students and workers, and coming from Laval (an island just North of Montreal island, separated from it by bridges) increasing activity space, while high cluster values (which are associated with dense, mixed use and transit rich environments), service sector employment, and children and seniors decreased activity spaces. Finally, the OD 08 dummy variable produced a statistically significant, albeit slight, negative coefficient, which would indicate that activity spaces decreased somewhat from 2003 to 2008. Further analysis would be necessary however to confirm this as a trend.

CONCLUSION

In summary, this paper has demonstrated the pertinence of using a clustered approach to relate urban form and public transit to activity spaces in a context-sensitive way. Results point to a significant link between land use clusters and activity spaces, and imply that efforts to increase density, mix and transit accessibility are valid investments for cities seeking to reduce travel demand they deem excessive, environmentally detrimental or unproductive. Since household and CT characteristics were used as control variables, the regression results make a strong case for promoting densification, increased land use mix and better transit provision.

An approach which could bear fruit to improve model accuracy in the future may be to use latent-class linear regression, which would combine the land use clustering approach with a form of household clustering. Instead of using continuous household or even CT variables like income, number of cars, persons and children to predict activity spaces, these could instead be treated as subpopulations. Another improvement could be to endogenize household location choice to account for residential self-selection.

Future research aimed at developing land use and transportation policy could definitely make use of the clustered approach combined with activity spaces, but what this case study has demonstrated is that the scale and heterogeneity of the region studied must be carefully considered before undertaking such an endeavor. Smaller scales or an altered methodology would improve the likelihood of clear policy being written from these analyses. Cluster analysis provides an effective means by which the potential impacts of urban form and transit interventions can be assessed, and thus their costs and benefits properly evaluated.

REFERENCES

1. Ewing, R., and R. Cervero. Travel and the Built Environment. *Transportation Research Record*: *Journal of the Transportation Research Board*, No. 1780, Transportation Research Board of the National Academies, Washington, D.C., 2001, pp. 87-114.
2. Ewing, R., and R. Cervero. Travel and the Built Environment. *Journal of the American Planning Association*, Vol. 76, No. 3, 2010, pp. 265-285.
3. Bento, A. M., M. L. Cropper, A. M. Mobarak, and K. Vinha. The Effects of Urban Spatial Structure on Travel Demand in the United States. *The Review of Economics and Statistics,* Vol. 87, No. 3, 2005, pp. 466-478.
4. Boarnet, M. G., and S. Sarmiento. Can Land-use Policy Really Affect Travel Behaviour? A Study of the Link between Non-work Travel and Land-use Characteristics. *Urban Studies*, Vol. 35, No. 7, 1998, pp. 1155-1169.
5. Lin, J., and L. Long. What neighborhood are you in? Empirical findings of relationships between household travel and neighborhood characteristics. *Transportation*, Vol. 35, 2008, pp. 739-758.
6. Shay, E., and A. Khattak. Automobiles, Trips, and Neighborhood Type: Comparing Environmental Measures. *Transportation Research Record*: *Journal of the Transportation Research Board*, No. 2007, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 73-82.
7. Krizek, K. (2003). Operationalizing Neighborhood Accessibility for Land Use-Travel Behavior Research and Regional Modeling. *Journal of Planning Education and Research*, Vol. 22, 2003, pp. 270-287.
8. Riva, M., P. Apparicio, L. Gauvin, and J.-M. Brodeur. Establishing the soundness of administrative spatial units for operationalizing the active living potential of residential environments: an exemplar for designing optimal zones. *International Journal of Health Geographics*, Vol. 7, No. 43, 2008, pp. 1-13.
9. Miranda-Moreno, L. F., L. Bettex, S. A. Zahabi, and T. Kreider. A Simultaneous Approach to Evaluate the Endogenous Influence of Urban Form and Public Transit Accessibility on Distance Travelled. *Transportation Research Record: Journal of the Transportation Research Board*, (in press), Transportation Research Board of the National Academies, Washington, D.C.
10. Ross, N. A., S. Tremblay, S. Khan, D. Crouse, M. Tremblay, and J.-M Berthelot. Body Mass Index in Urban Canada: Neighborhood and Metropolitan Area Effects. *American Journal of Public Health*, Vol. 97, No. 3, 2007, pp. 500-508.
11. Leck, E. The Impact of Urban Form on Travel Behaviour: A Meta-Analysis. *Berkeley Planning Journal*, Vol. 19, 2006, pp. 37-57.
12. Gershoff, E. T., S. Pederson, and J. L. Aber. Creating Neighborhood Typologies of GIS-Based Data in the Absence of Neighborhood-Based Sampling: A Factor and Cluster Analytic Strategy. *Journal of Prevention and Intervention in the Community*, Vol. 37, 2009, pp. 35-47.
13. Kamruzzaman, M., J. Hine, B. Gunay, and N. Blair. Using GIS to visualize and evaluate student travel behaviour. *Journal of Transport Behaviour*, Vol. 19, 2009, pp. 13-32.
14. Manaugh, K., L. F. Miranda-Moreno, and A. El-Geneidy. Neighborhood travel behaviour: A factor and cluster analysis to understand travel behavior in the Montréal metropolitan region. *Transportation*, Vol. 37, No. 4, 2010, pp. 627-646.

15. Kestens, Y., A. Lebel, M. Daniel, M. Thériault, and R. Pampalon. Using experienced acivity spaces to measure foodscape exposure. *Health and Place*, Vol. 16, 2010, pp. 1094-1103.
16. Buliung, R. N., and P. S. Kanaroglou. A GIS toolkit for exploring geographies of household activity/travel behaviour. *Journal of Transportation Geography*, Vol. 14, 2006, pp. 35-51.
17. Rai, R. K., M. Balmer, M. Rieser, V.S. Vaze, S. Schonfelder, and K. W. Axhausen. Capturing New Activity Spaces, New Geometries. *Transportation Research Record*: *Journal of the Transportation Research Board*, No. 2021, Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 70-80.
18. LeBeau, J. L. The Methods and Measures of Centrography and the Spatial Dynamics of Rape. *Journal of Quantitative Criminology*, Vol. 3, No. 2, 1987, pp. 125-140.
19. Sherman, J. E., J. Spencer, J. S. Preisser, W. M. Gesler, and T. A. Arcury. A suite of methods for representing activity space in a healthcare accessibility study. *International Journal of Health Geographics,* Vol. 4, 2005, pp. 24.
20. Lee-Gosselin, M., L. F. Miranda-Moreno, M. Thériault, and T. Kreider. Evolution of travel and daily activity patterns and ICTs. Presented at the Fourth Specialist Meeting of the Network ICTs: mobilizing persons, places and spaces, Île d'Orléans, Québec., 7-9 October 2009
21. Fan, Y., and A. J. Khattak. Urban Form, Individual Spatial Footprints, and Travel. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2082, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 98-106.
22. Manaugh, K., and A. El-Geneidy. What makes travel local? Defining and understanding local travel behaviour. *Transportation Research Record: Journal of the Transportation Research Board*, Pending Publication, Transportation Research Board of the National Academies, Washington, D.C.
23. Freund, P., and G. Martin. Hyperautomobility, the Social Organization of Space, and Health. *Mobilities*, 2007, pp. 37-49.
24. Frumkin, H. Urban Sprawl and Public Health. *Public Health Reports*, Vol. 117, 2002, pp. 201-217.
25. Marshall, J. D., T. E. McKone, E. Deakin, and W. W. Nazaroff. Inhalation of motor vehicle emissions: effects of urban population and land area. *Atmospheric Environment*, Vol. 39, 2005, pp. 283-295.
26. Calthorpe, P. *The Next American Metropolis.* Princeton Architectural Press, New York, 1993.
27. Statistics Canada. Canada Census Digital Boundary File. Ottawa, 2006.
28. Agence Métropolitaine de Transport. *Enquête origine-destination 2003 et 2008.* Montreal, Quebec, 2003 and 2008.
29. Statistics Canada. (2006). *2006 Cumulative Profile, Montréal (878 Census tracts).* E-STAT (distributor).
30. Statistics Canada. (2001). *2001 Census of Population, Montréal (862 areas).* E-STAT (distributor).
31. 2001 Census - Statistics Canada. *Enquête sur le travail, le milieu de travail et les employés.* CD-ROM
32. 2006 Census - Statistics Canada. *Enquête sur le travail, le milieu de travail et les employés.* CD-ROM

33. Newman, P., and J. Kenworthy. *Sustainability and Cities: Overcoming Automobile Dependence.* Island Press, Washington, D.C., 1999.
34. Shearmur, R. Travel From Home: An Economic Geography of Commuting Distances in Montreal. *Urban Geography*, Vol. 27, No. 4, 2006, pp. 330-359.
35. Newman, M. C. Regression Analysis of Log-Transfomed Data: Statistical bias and its correction. *Environmental Toxicology and Chemistry*, Vol. 12, 1993, pp. 1129-1133.

List of Tables and Figures:

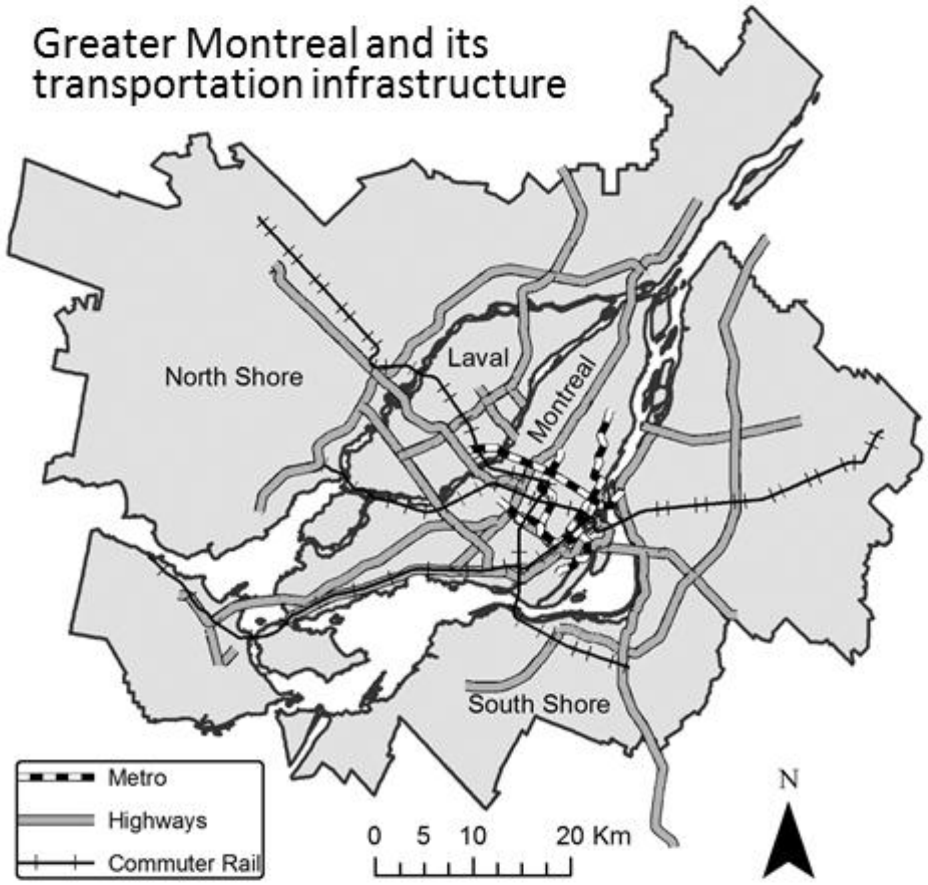

FIGURE 1 Study area.

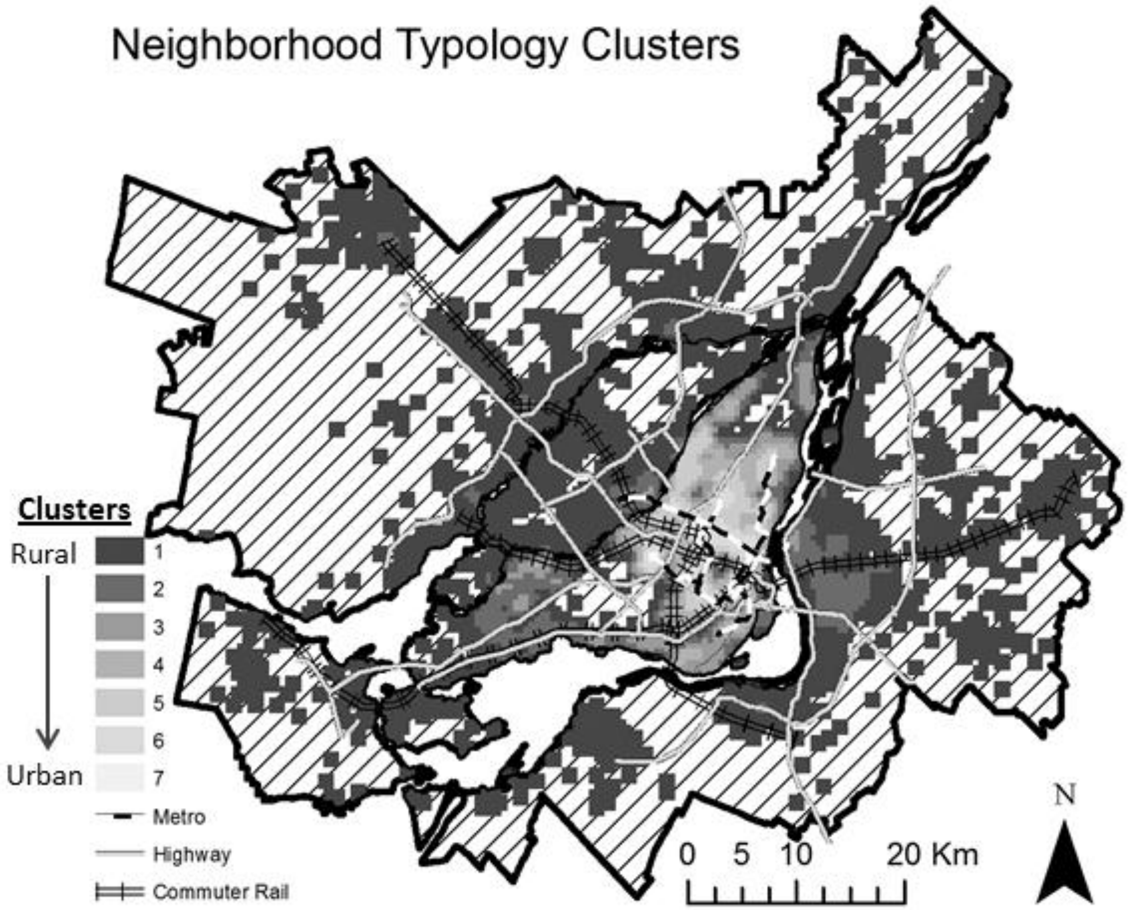

FIGURE 2 Neighborhood typology and the city's main transportation infrastructure.

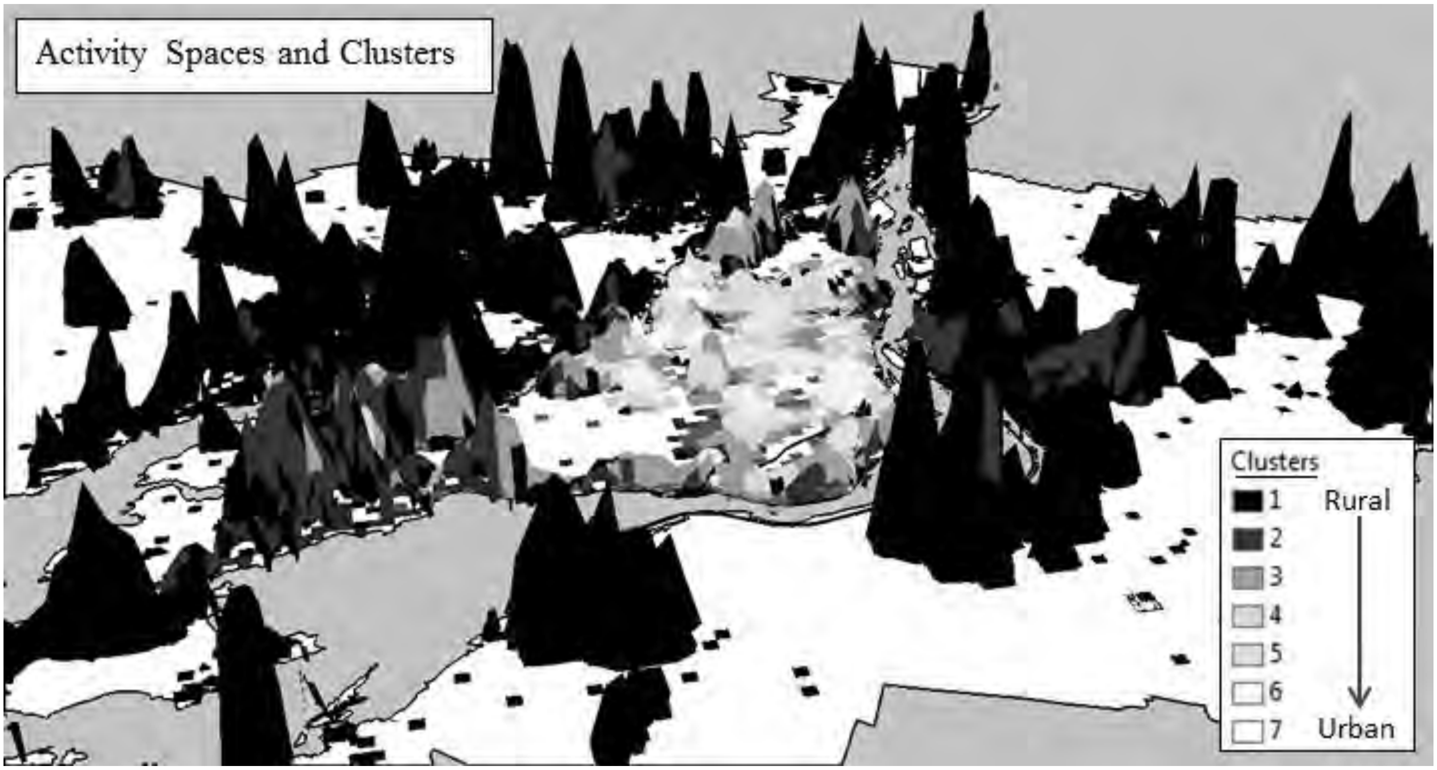

FIGURE 3 Clusters and activity spaces, 2008 data displayed – heights represent activity space.

TABLE 1 Summary statistics for clusters (valid households only)

2003 and 2008 Clusters	Observations	Persons / Hectare	Employment / Hectare	Land Use Mix	PT Accessibility	House-hold Trips	Percent on Island of Montreal	Activity Space (km2)
1, Rural no Transit	44% / 18,210	19.66	4.70	21%	6.98	8.60	24%	68.79
2, Rural	11% / 4,685	29.78	8.90	30%	46.42	8.54	68%	40.57
3, Rural / Suburban	11% / 4,493	44.90	12.18	37%	109.54	8.20	100%	33.31
4, Outer Suburb	12% / 4,771	66.46	21.66	44%	171.75	7.82	99%	22.86
5, Inner Suburb	13% / 5,165	86.72	29.82	50%	247.95	7.53	100%	18.27
6, Urban Core	8% / 3,294	96.22	67.79	56%	362.37	7.16	100%	14.87
7, Downtown Core	1% / 499	86.38	250.55	59%	554.52	6.55	100%	11.41
Mean	41,116 Total	44.37	19.16	33%	107.19	8.18	62.5%	45.01

TABLE 2 Regression results, the logarithm of the CVH area is the dependent variable

	Number of observations		41,116
	R-squared		0.28
Variable	Coefficient	t-stat	95% Confidence Interval
Cluster 2, Rural	-0.25	-8.91	-0.31 to -0.2
Cluster 3, Rural/Suburban	-0.44	-15.17	-0.5 to -0.39
Cluster 4, Outer Suburb	-0.64	-21.54	-0.7 to -0.58
Cluster 5, Inner Suburb	-0.77	-25.55	-0.82 to -0.71
Cluster 6, Urban Core	-0.97	-26.57	-1.04 to -0.9
Cluster 7, Downtown Core	-1.52	-19.37	-1.67 to -1.36
FG workers per CT (%)	-1.07	-11.42	-1.25 to -0.88
# of Children	-0.32	-24.65	-0.35 to -0.3
# of Seniors	-0.27	-5.58	-0.37 to -0.18
# of Full-time Students	0.15	11.67	0.12 to 0.17
# of Full-time Workers	0.40	31.78	0.38 to 0.42
Licences per Household	0.32	25.94	0.3 to 0.34
# of Trips	0.10	34.79	0.09 to 0.1
Resident of Laval	0.15	5.47	0.09 to 0.2
Homemakers per CT (%)	2.95	18.29	2.63 to 3.27
OD 08 Household	-0.05	-3.03	-0.08 to -0.02
Constant	14.79	208.88	14.65 to 14.93

*Reference cluster is cluster 1, the most rural cluster