

## Heterogeneity in mode choice behavior: A spatial latent class approach based on accessibility measures

**Jaime Orrego-Oñate** (corresponding author)  
Department of Geography  
Universitat Autònoma de Barcelona  
[jaime.orrego@uab.cat](mailto:jaime.orrego@uab.cat)

**Kelly J. Clifton**  
School of Community and Regional  
Planning, University of British Columbia  
[kelly.clifton@ubc.ca](mailto:kelly.clifton@ubc.ca)

**Ricardo Hurtubia**  
School of Architecture & Department of  
Transport Engineering and Logistics,  
Pontificia Universidad Católica de Chile  
[rhurtubia@uc.cl](mailto:rhurtubia@uc.cl)

**Abstract:** We propose a method to estimate mode choice models, where preference parameters are sensitive to the spatial context of the trip origin, challenging traditional assumptions of spatial homogeneity in the relationship between travel modes and the built environment. The framework, called Spatial Latent Classes (SLC), is based on the integrated choice and latent class approach, although instead of defining classes for the decision maker, it estimates the probability of a location belonging to a class, as a function of spatial attributes. For each Spatial Latent Class, a different mode choice model is specified, and the resulting behavioral model for each location is a weighted average of all class-specific models, which is estimated to maximize the likelihood of reproducing observed travel behavior. We test our models with data from Portland, Oregon, specifying spatial class membership models as a function of local and regional accessibility measures. Results show the SLC increases model fit when compared with traditional methods and, more importantly, allows segmenting urban space into meaningful zones, where predominant travel behavior patterns can be easily identified. We believe this is a very intuitive way to spatially analyze travel behavior trends, allowing policymakers to identify target areas of the city and the accessibility levels required to attain desired modal splits.

### Article history:

Received: October 8, 2021  
Received in revised form:  
November 4, 2022  
Accepted: December 26, 2022  
Available online: April 6, 2023

## 1 Introduction

Cities face increasing pressure to reduce car dependency, a trend that generates many externalities for urban dwellers and threatens sustainability for future generations (Stevenson et al., 2016). Urban planners are increasingly attuned to land use policy and its potential effect on travel behavior, explicitly with the aim of reducing car dependency (Giles-Corti et al., 2016). Therefore, understanding the human response to urban environmental attributes and their influence on travel behavior is a relevant and timely

---

Copyright 2023 Jaime Orrego-Oñate, Kelly J. Clifton & Ricardo Hurtubia

<http://dx.doi.org/10.5198/jtl.u.2023.2115>

ISSN: 1938-7849 | Licensed under the [Creative Commons Attribution – Noncommercial License 4.0](https://creativecommons.org/licenses/by-nc/4.0/)

The *Journal of Transport and Land Use* is the official journal of the World Society for Transport and Land Use (WSTLUR) and is published and sponsored by the University of Minnesota Center for Transportation Studies.

research topic. Today's sustainability challenges pose an interesting question: Can the built environment effectively influence travel choices?

The literature frequently reports a positive correlation between attributes of the built environment, such as population density or accessibility, and the propensity to walk or to use public transport (Ewing & Cervero, 2010). However, there is still debate regarding the magnitude of these correlations, which depend on the chosen study method, the specification of the models used, the context and location of the case of study, or the quality of the data. Moreover, the role played by residential self-selection (i.e., households with preconceived travel preferences choosing neighborhoods that best suit their mobility habits) is still largely discussed, although several studies indicate that, beyond self-selection, there is a causal influence of built environment and land use on travel patterns, especially regarding the propensity to walk (Cao, 2010; Lin et al., 2017; McCormack & Shiell, 2011).

A common approach to understanding the associations between travel choices and built environment is the estimation of discrete choice models for mode choice (Ben-Akiva & Lerman, 1985; McFadden, 1974) where attributes of the built environment are included as explanatory variables. Traditional travel models, in practice, usually specify a single utility function for each mode in a whole urban area (or in few predefined zones subdividing it), under the assumption that the response to changes in the attributes of travel modes does not vary across space within the same area or zone (although it does so across different individuals). This approach does not adequately accommodate spatial response heterogeneity in the context-specific built environment, therefore limiting the understanding of how the built environment influences the choice of transportation modes, despite the extensive and deep literature on the topic (Kärmeniemi et al., 2018; Salvo et al., 2018). Considering spatial heterogeneity in the modelling approach should allow to better measure the role played by the built environment. Previous efforts to address spatial heterogeneity include segmenting the analysis in different urban typologies (Choi, 2018; Khattak & Rodriguez, 2005) or defining spatial clusters based on neighborhood characteristics (Salon, 2015). However, these methods are based on *ex-ante* segmentations of space, with the estimation of behavioral (mode choice) models done in a subsequent stage. This approach is less likely to correctly capture the interdependence between spatial attributes and other variables explaining travel behavior, such as trip attributes or socioeconomic characteristics of users.

This research aims to fill this gap by simultaneously analyzing spatial heterogeneity and behavioral (mode-choice) data. The proposed method is based on a discrete choice framework, particularly integrated choice and latent class models (Kamakura & Russel, 1989). The method, applied to the metropolitan region of Portland, Oregon, allows to segment behavioral models in different classes, each having its own probability function. In this application, class segmentation will be based on spatial attributes of the location where the trip is originated, particularly accessibility measures. These segmentations characterize a behavioral representation that depends on the location of a trip, where spatial heterogeneity is explicitly incorporated through the probability of that location belonging to a Spatial Latent Class (SLC). The method innovates by segmenting the space using a suite of (both local and regional) accessibility measures and travel data, allowing for a behavioral-based identification of "neighborhood types" which are described by their attributes. This provides a coherent description of urban space in terms of the travel behavior patterns that it is likely to induce, which could be a useful and intuitive input for policy makers and urban planners.

Beyond interpretability and behavioral consistence of the spatial segmentation that emerges from the proposed method, estimation results confirm better model performance when compared with a classical multinomial logit modelling approach, where spatial attributes are directly included in the utility functions of each mode of transport. The results indicate that certain sociodemographic groups may behave differently (e.g., be represented in models with different coefficients) depending on the attributes of their location, something that cannot be captured by spatially homogeneous models. Finally,

the method allows to spatially delimitate the areas associated with different mobility patterns and understanding the levels of accessibility needed to achieve specific travel outputs. To our knowledge, there has been no previous experience applying a Spatial Latent Class modeling approach to a multimodal travel choice model.

The remainder of the paper is organized as follows; Sections 2 and 3, detail previous literature on spatial heterogeneity and latent class methods applied to spatial segmentation, respectively. Next, Section 4 details the methods used in this article, including the SLC approach. Section 5 describes the case study and the data. Section 6 shows the result of the estimations, with an emphasis on the spatial segmentation generated by the SLC approach. Finally, Section 7 discusses the findings and contribution of this work.

## 2 Spatial heterogeneity in travel behavior

The literature often mentions a positive relationship between attributes of the built environment and urban travel behavior (Ewing & Cervero, 2010; Saelens & Handy, 2008; Salvo et al., 2018). Although a significant correlation has been found in a wide variety of studies, some evidence argues that the magnitude of the effect is small (Duranton & Turner, 2018; Stevens, 2017), suggesting that land use interventions could have a limited impact on urban travel patterns. Therefore, the question of how relevant the built environment is to shaping preferences remains a concern.

The usual approach to understanding this relationship is to estimate probabilistic travel behavior models, most commonly mode choice, with socioeconomic characteristics, travel attributes, and built environment measures as independent variables. Modeling approaches tend to assume spatial homogeneity, which refers to an estimation with singular coefficients for each measure of the built environment throughout a region. A common study approach is the “D system” (Cervero & Kockelman, 1997; Ewing et al., 2009), referring to concepts associated with specific attributes of the built environment, such as density of population or employment, diversity of land uses, urban design, accessibility to destination, and distance to transit. For example, Stevens (2017), using this approach in a meta-regression study, argues that the effect of built environment measures is significantly associated with private Vehicle Miles Traveled (VMT) but with a small magnitude effect.

Several studies have questioned the assumption that relationships between the built environment and mode choice follow a unique behavioral response regardless of the context. For example, Lewis and Grande del Valle (2019) found that VMT decreases with density, with evidence pointing to a nonlinear relationship in the shape of a negative exponential, using the San Francisco Bay Area as a case study. In another example, Choi (2018), in a study on Canada, found a limit in the association of density with VMT reductions, suggesting that increasing density in central areas, that have reached a certain threshold of population, has no effect on VMT. This is an indicator of the effect of the built environment on the elasticities of key variables that affect travel behavior.

Private vehicle choice and VMT have been found to have a similar associations with the built environment, suggesting that those non-linearities would exist when either VMT or car alternative are the modeled variable (Salon et al., 2012). In the case of transit, Ding et al. (2021) found that assuming spatial homogeneity can lead to inconsistent estimates, and a certain threshold of transit supply should be exceeded to be effective. Additionally, they found that densification facilitates transit, but with diminishing returns. In a similar vein, but for walking, Tanishita and van Wee (2017) found an upper threshold for increasing walking with a population density of around 11,000 people/km<sup>2</sup>. Also for the case of walking, Guimpert and Hurtubia (2018) found that the shape and size of the areas that people consider to be part of their usual walking neighborhood are heterogeneous, depending on the attributes

of the built environment and individual characteristics.

Several authors have attempted to incorporate spatial heterogeneity into travel behavior analysis. Some approaches have grouped areas that may show similar features of the built environment and performed a spatially segmented analysis. Using normative urban typologies, Choi (2018) found in Calgary, Canada, that intensifying activities in the center of the city would not have any effect on decreasing VMT, while increasing densification and transit supply may decrease VMT in other consolidated areas outside of downtown. Salon (2015), using *ex-ante* clustering of space, also found evidence on spatial heterogeneity, with total VMT depending on the neighborhood type and marginal effects of built environment attributes changing in magnitude depending on the neighborhood. Furthermore, Feuillet et al. (2018), using a Geographical Weighted Regression, found that walking is also spatially heterogeneous in a study over France. Their findings point out that increasing density in already dense cities would not increase walking, but that the same change could increase walking in smaller and less dense cities. Although these findings support the need to include spatial heterogeneity in the analysis, these methods have not considered clustering or spatial segmentation methods based on behavioral outcomes.

### 3 Latent class models for spatial heterogeneity

One of the most standard modeling techniques applied to mode choice is the multinomial logit model (MNL). Although this model specification can capture heterogeneity by introducing systematic taste variation, the resulting coefficients are not necessarily spatially sensitive. Therefore, the estimation could be biased if there is spatial heterogeneity in the choice of modes, misrepresenting the association between travel choices and environmental measures. Furthermore, the model specification could be estimated by interacting accessibility/built environment measures with socioeconomic characteristics. However, in this approach, the interpretation is more challenging as there is no explicit spatial segmentation.

Latent class modeling (LCM) is a framework that allows for the introduction of heterogeneity into discrete choice models without the need of arbitrary segmentations of individuals (Kamakura & Russell, 1989). LCMs are less restrictive than MNLs and, instead of systematic taste variation, they assume a latent heterogeneity in preferences, which translates into types of users that cannot be directly observed by the analyst, but whom can be probabilistically identified and correlated with observable variables. LCMs implicitly estimate a different behavioral model for each latent class (with different coefficients). Additionally, another model defines the probability of belonging to each class (Greene & Hensher, 2003).

There are several examples of LCM in the travel behavior literature where, in most cases, the partition into classes is based on individual characteristics. For example, Hurtubia et al. (2014) estimated significantly different mode choice parameters for classes defined by household composition variables, as well as income and car ownership levels. Wen et al. (2012) used the method to identify different classes of travelers for high-speed rail. Etzioni et al. (2021) analyzed mode choice between three emerging automated vehicles: ride sharing, car sharing, and automated transit. The study estimates latent classes to capture taste heterogeneity using sociodemographic characteristic, travel habit, and latent variables. Kim and Rasouli (2022) studied the willingness to adopt new and innovative mobility solutions based on metrics of the individuals' lifestyle to differentiate in classes.

The literature proposing latent classes based on spatial attributes and/or environmental factors is sparse. Cox and Hurtubia (2021, 2022), manage latent segmentations of residential locations based on place-specific attributes and observed location choices in Santiago, Chile. Oliva et al. (2018) estimate a SLC model for the frequency of bicycle commuting, where class membership depends on built environment attributes of the neighborhood of residence of each traveler. Beyond providing a richer understanding of the relationship between the built environment and cycling, the results identify

neighborhoods that facilitate or discourage cycling, helping to define and geographically target policies to incentivize bicycle use. Furthermore, Sarrias (2019) found that using discrete classes to model spatial heterogeneity can be more advantageous than using a continuous distribution.

While this approach should allow to better measure spatial heterogeneity in travel behavior, we have not identified previous literature applying SLC to accessibility (or built environment) measures and multimodal choice relationships. This article seeks to better understand how different urban contexts play a role in defining mode choice behavior while acknowledging that the response of travelers to the built environment is not constant in space.

## 4 Methods

The hypothesis underlying this work is that the competition between different travel modes varies across locations, and therefore, the factors that influence behavior differ depending on the spatial context. Areas with low population density and lower intensity of use of space typically have more available road and parking infrastructure, making driving a more accessible option compared to walking or using public transit. Conversely, areas with high population density and more intense use of the built environment generally have more infrastructure and support for walking or using public transit, but less availability of parking and roads for driving. Therefore, the specific travel behavior of an individual may vary depending on where they start and end their trip, as well as the transportation options available to them in that location.

The density of people using a space will depend on how many people can access that place, which depends not only on local accessibility levels but also on the regional access to places. Then both measures (local and regional accessibility) are of interest in characterizing the mode choice and considered in our study. First, local accessibility corresponds to small-scale measures of people and employment within a place, reachable through its pedestrian transportation networks. Regional accessibility measures jobs that can be reached in motorized vehicles (either transit or cars) in a larger urban context.

The combination of regional and local accessibility measures will be used to segment the space into classes. Each class will be associated with different utility functions for mode choice, based on trip distance and individual characteristics. Consequently, the results will help describe the critical components of the built environment in each spatial context, improving the performance assessment of different land use policies and transit infrastructure.

We propose to compare two modeling approaches, Spatial Latent Classes (SLC) and a more traditional MNL. Both models will include the same independent variables that account for local and regional accessibility measures. To the best of our knowledge, this is the first time a Spatial Latent Class approach is applied to a mode choice model.

### 4.1 Multinomial logit (MNL)

For the case of the MNL, the utility associated with transport mode  $m$  for individual  $n$  with a trip starting in zone  $i$  with destination in zone  $j$  is defined as:

$$U_{ijm}^n = V_{ijm}^n + \varepsilon_{ijm} \quad (1)$$

where  $\varepsilon_{ijm}$  is a random term accounting for unobserved factors influencing the choice and  $V_{ijm}^n$  is the systematic part of the utility, which can be specified as a function of individual characteristics, zonal

attributes, and travel distance.

If the error term follows a Gumbel independent and identical distributed (IID), the functional form of the probability of choosing a particular transport mode  $m$  in a set  $M$  of available mode, for an individual  $n$  with trip-origin  $i$  and destination  $j$  is:

$$P_{ijm}^n = \frac{\exp(V_{ijm}^n)}{\sum_{l \in M} \exp(V_{ijl}^n)} \quad (2)$$

#### 4.2 Spatial latent class (SLC) and mode choice model

Latent Class Models allow for the estimation of different choice models conditional on the membership to unobserved latent classes, which are traditionally modelled as a function of individual characteristics. We adapt this approach by making the class membership model a function of spatial attributes, therefore, segmenting space in terms of the likelihood of a belonging to a spatial latent class (SLC) for each location. Because the mode-choice models conditional to each latent class are estimated simultaneously with the class membership model, the spatial segmentation can be interpreted as zones of the city associated with different patterns of mode-choice behavior.

We define the membership function of location origin  $i$  to class  $k$  ( $W_{ki}$ ) similarly to the utility function of Equation 1 but with its systematic part being a function of a vector of local and regional accessibility measures associated with zone  $i$ . Making the same assumptions regarding the error term, the probability of zone  $i$  belonging to class  $k$  is:

$$P_{ki} = \frac{\exp(W_{ki})}{\sum_{k'} \exp(W_{k'i})} \quad (3)$$

We define a mode choice model conditional to each class, where the probability of an individual  $n$  choosing mode  $m$  for a trip between zones  $i$  and  $j$ , conditional on the origin of the trip belonging to spatial class  $k$  ( $P_{ijm/k}^n$ ) as in Equation 2, but with class-specific preference parameters. Then, the unconditional choice probability of mode  $m$  for individual  $n$  travelling between zones  $i$  and  $j$  is

$$P_{ijm}^n = \sum_k P_{ki} P_{ijm/k}^n \quad (4)$$

Both the MNL and the SLC models can be estimated through maximum likelihood methods.

## 5 Case study data

The location and time frame for this case study is the metropolitan region of Portland, Oregon, in 2011. The Urban Growth Boundary (UGB), a normative instrument of the state of Oregon defined by the local Metropolitan Planning Organization (MPO), will define the area for this application. The UGB limits urban sprawl by confining development to a specific area. For this study, the area size is approximately 1,000 km<sup>2</sup> and includes approximately 1.5 million people in 2011. There are several cities within the metro area: Portland is the central city of the region, with a population of approximately 600,000 in 2011. The transportation system comprises an extensive freeway network that connects the central city with the surrounding suburbs. Furthermore, the regional transit system includes five light rail lines

with an length of around 100 km, more than 80 bus routes with a fleet of about 700 buses, and a daily passenger volume (Pre-COVID-19) of more than 300,000.

The Oregon Household and Activity Survey of 2011 (Oregon Modeling Steering Committee, 2011) is used for the travel data estimation in the models. The one-day survey includes 6,108 households, adding up to 56,534 trips. These data reflect the travel patterns of what would be a normal week-day in the metropolitan area of Portland, Oregon.

To define our basic spatial analysis unit, the region was subdivided into 80 by 80 meters square cells, resulting in approximately 160,000 zones. The population and employment levels were calculated for each cell. The cell-based employment was derived from Metro, the Regional Metropolitan Planning Organization. For population, we used 2010 block-level population census data, scaled according to the percentage of a block within each cell (areal interpolation).

**5.1 Travel data**

The 56,534 trips included in the survey correspond to the Metropolitan Statistical Area of Portland, which includes parts of the southern state of Washington. The data set is filtered to include only trips that start and end in Oregon, with walking, transit, or automobile as travel modes (the sample for bicycle trips was too small to merit inclusion), and travelers older than 16 years of age. Observations without income data, trip distance, and a specified age were excluded. Data processing resulted in a final data set of 27,252 trips/observations. Each trip origin is matched with a unique cell. The travel patterns of the sample are shown in Table 1.

**Table 1.** Travel patterns of the sample

	Car	Transit	Walk
Sample	22,779	1,760	2,713
Trips (%)	3.01M (82%)	0.28M (8%)	0.39M (10%)
Mean distance [km]	7.5	11.5	0.8
Median distance [km]	5	9.9	0.5

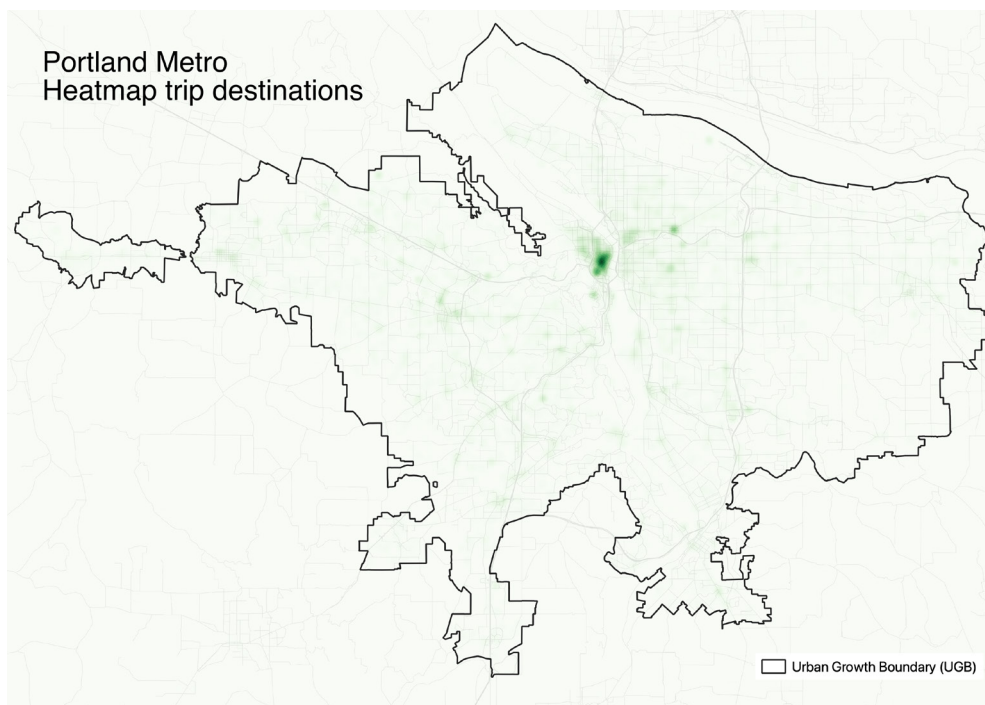
Each observation is a trip, described by the travel distance, the origin coordinates, and the chosen primary mode. The individual traveler characteristics included are the income level of the home in three categories (\$0 to \$35,000; \$35,000 to \$75,000; and \$75,000 or more), gender as a binary variable (0 represents female; 1 means male), and age in three levels (16-24, 25-65 and 65 and older). The socio-economic distribution of the sample is shown in Table 2.

**Table 2.** Socioeconomic distribution of the sample

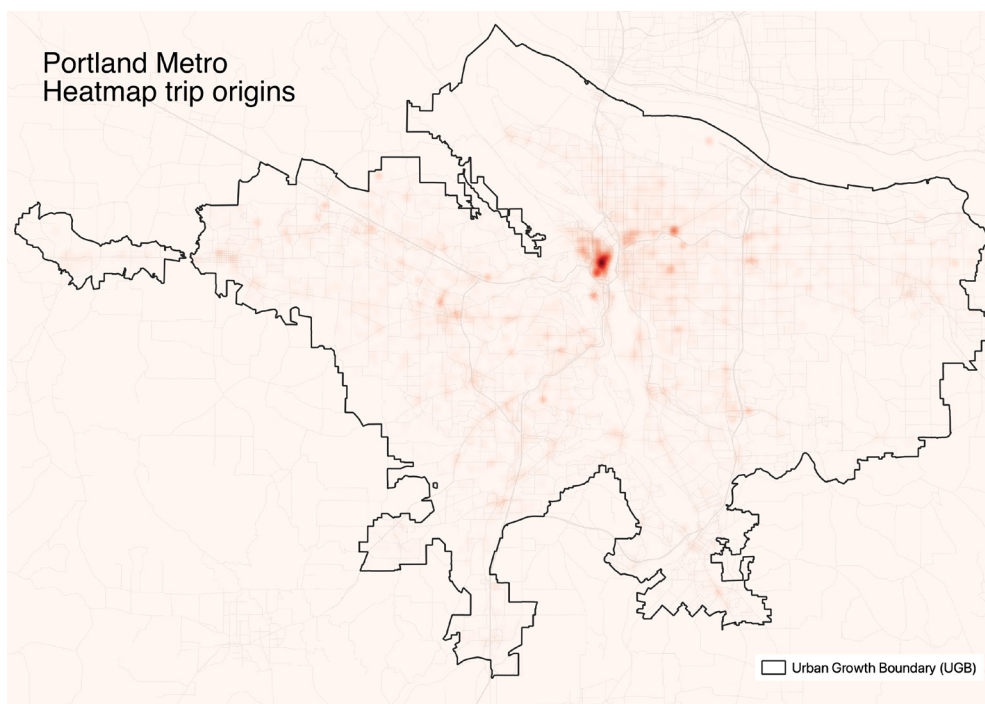
<b>Socioeconomic distribution</b>	People	Trips per person
Total sample	6,564	4.2
Income (annual household income - US dollars)		
Category 1: <\$35,000	1,016 (15%)	4.0
Category 2: \$35,000 - \$75,000	2,136 (33%)	4.1
Category 3: >\$75,000	3,412 (52%)	4.3
Age		
Category 1: 16-24	426 (7%)	3.5
Category 2: 25-64	5,003 (76%)	4.2
Category 3: 65+	1,135 (17%)	4.1
Gender		
Woman	3,553 (54%)	4.3
Man	3,011 (46%)	4.0

Finally, as a reference, we include a heat map of origin and destinations on Figure 1 and Figure 2. The maps show that the origin and destinations are similarly concentrated in the downtown and in several regional subcenters across the metropolitan area.





**Figure 1.** Heatmap of origins



**Figure 2.** Heatmap of destinations

## 5.2 Local accessibility

For local accessibility, all measures are defined in a 750-meter isodistance buffer from the centroid of each grid cell, following the walkable road network. This 750-meter isochrone accounts for approximately 10-minutes walking, usually considered as the threshold of proximity travel (Marquet & Miralles-Guasch, 2015). The buffer area acts as the pedestrian catchment area (PCA) or a network measure and correlates with mode choice (Adams et al., 2015; Koohsari et al., 2015; Stevens, 2017; Stockton et al., 2016). The total population and employment within each PCA are computed and used as local accessibility measures. Employment is an indicator of land use intensity and has been associated with mode choice (Brown et al., 2016; Clifton et al., 2016; Huang et al., 2019; Lefebvre-Ropars et al., 2017). Furthermore, population density has been identified as a relevant covariate of modal choice (Eom & Cho, 2015; Merlin, 2018; Tanishita & van Wee, 2017). Combining all these measures will represent a proxy of activity capacity in an area. The final suite of variables included are:

- Pedestrian Catchment Area Surface (PCA)
- Employment in PCA
- Population in PCA

## 5.3 Regional accessibility

Variables for regional accessibility are constructed for the number of jobs accessible by motorized vehicles in a time frame of 30 minutes for each centroid cell grid. Previous studies have found that these measures are correlated with the use of automobiles and transit (Bento et al., 2005; Owen & Levinson, 2015). The accessibility measure for cars is constructed using the Portland MPO estimates from the regional transportation model at the TAZ level. The travel time from the centroid of TAZ to every other TAZ centroid is scaled down to each cell. The assumption is that the travel time between two TAZ centroids is the average between the origin and destination grid cells within a TAZ. Then, with the same threshold of 30 minutes, all reachable TAZ employment is added. Around 90% of car trips in the sample are under 30 min. Therefore, the threshold is chosen as a round number with high representation.

The transit accessibility is constructed from the GTFS data standard and the R package tidytransit (Poletti et al., 2020) to identify all the transit stops that are within a 30-minute travel time of each transit stop. The 30-minute threshold is chosen as it captures most transit trips (~60%) and the same timeframe used for car accessibility measure. Additionally, a series of 250-meter network buffers is created at each stop, for which total employment is calculated. The final variable is generated by adding all the jobs for each stop to all other stops within 30 minutes of transit travel (understanding the jobs of each stop as the buffer of the employment network of each stop).

Therefore, the two regional accessibility metrics included are:

- Number of jobs accessible by public transport in 30 min
- Number of jobs accessible by car in 30 minutes

Table 3 shows the basic statistics of accessibility measures.

**Table 3.** Statistics of accessibility measures

<b>Number of cells</b>	160,148			
<b>Local accessibility</b>	<b>Mean</b>	<b>Standard Deviation</b>	<b>Median</b>	<b>Max</b>
Pedestrian Catchment Area Isodistance 750m (PCA) [ha]	38	25	34	122
Employment in PCA [# employees]	444	1787	377	55,445
Population in PCA [people]	802	958	506	9,569
<b>Regional accessibility</b>	<b>Mean</b>	<b>Standard Deviation</b>	<b>Median</b>	<b>Max</b>
Total employment accessible by transit in 30 minutes [# employees]	14,172	40,288	0	227,459
Total employment accessible by car in 30 minutes [# employees]	427,334	144,586	313,257	718,650

All variables are scaled by their standard deviation without centering to make the estimation process more efficient.

## 6 Results

As explained in the methods sections, the SLC has a class membership model and a class-specific mode choice model, making it conditional to the class membership. The MNL will include the same independent variables as the SLC models, but with environmental variables added directly to the utility function. The class membership model is set up for three spatial latent classes associated with each zone (grid cell) according to their local and regional accessibility measures. We explored specifications with more than three classes, but then the interpretation of each class became challenging, without much gain in model fit. The models were estimated using Pandas Biogeme (Bierlaire, 2018). Table 4 shows the results for SLC and MNL.

**Table 4.** Coefficients for SLC and MNL

Coefficients	Spatial Latent Class (SLC)						Multinomial Logit (MNL)	
	Class 1		Class 2		Class 3		Value	p-value
<b>CLASS MEMBERSHIP MODEL</b>	Value	p-value	Value	p-value	Value	p-value	Value	p-value
Intercept	-2.25	0.00	Ref.		0.948	0.00	-	-
<b>Local Accessibility</b>								
Pedestrian Catchment Area - 750m Isodistance (PCA) [ha]	-0.43	0.00	Ref.		0.41	0.00	-	-
Employment in PCA [#1000s jobs]	10.00	0.00	Ref.		-4.00	0.00	-	-
Population in PCA [people]	0.05	0.00	Ref.		-1.10	0.00	-	-
<b>Regional accessibility</b>								
Total employment accessible by transit in 30 min (# jobs)	0.08	0.00	Ref.		-0.06	0.00	-	-
Total employment accessible by car in 30 min (# jobs)	-0.87	0.00	Ref.		-0.40	0.00	-	-
<b>CAR</b>	Value	p-value	Value	p-value	Value	p-value	Value	p-value
Intercept	2.92	0.00	14.74	0.07	25.99	0.00	5.66	0.00
<b>Socioeconomics</b>								
Age (25-65)	0.13	0.00	-12.00	0.00	3.86	0.00	-1.34	0.00
Age (16-25)	-0.28	0.00	0.57	0.00	-11.02	0.00	-0.76	0.00
Gender (male=1)	-0.79	0.00	0.00	0.00	-0.51	0.00	-0.22	0.00
Low household annual income (\$0-\$35K)	-9.49	0.00	-1.68	0.00	-11.72	0.00	-1.49	0.00
Mid household annual income (\$35K-\$75K)	-1.58	0.00	-0.96	0.00	-9.76	0.06	-0.82	0.00
<b>Trip characteristics</b>								
Distance to the destination (km)	-2.35	0.00	-0.31	0.00	-0.92	0.00	-0.59	0.00
<b>Local Accessibility</b>								
Pedestrian Catchment Area - 750m Isodistance (PCA) [ha]	-	-	-	-	-	-	0.28	0.00
Employment in PCA [# jobs]	-	-	-	-	-	-	-0.08	0.00
Population in PCA [people]	-	-	-	-	-	-	-0.18	0.00
<b>Regional accessibility</b>								
Total employment accessible by transit in 30 min (# jobs)	-	-	-	-	-	-	-0.07	0.00
Total employment accessible by car in 30 min (# jobs)	-	-	-	-	-	-	-0.80	0.00
<b>WALKING</b>	Value	p-value	Value	p-value	Value	p-value	Value	p-value
Intercept	5.64	0.00	16.34	0.00	19.40	0.00	3.87	0.00
<b>Socioeconomics</b>								
Age (25-65)	4.15	0.00	-9.34	0.00	-6.79	0.00	-0.27	0.00
Age (16-25)	0.00	0.00	1.12	0.00	-10.01	0.00	-0.41	0.00
Gender (male=1)	-0.91	0.00	0.76	0.00	0.01	0.00	-0.01	0.23

Coefficients	Spatial Latent Class (SLC)						Multinomial Logit (MNL)	
	Class 1		Class 2		Class 3			
Low household annual income dummy (\$0-\$35K)	-1.22	0.00	-0.87	0.00	-9.12	0.00	-0.77	0.00
Mid household annual income dummy (\$35K-\$75K)	-1.17	0.00	-0.39	0.00	-9.44	0.00	-0.42	0.00
<b>Trip characteristics</b>								
Distance to the destination (km)	-18.27	0.00	-40.73	0.00	-1.65	0.00	-13.88	0.00
<b>Local Accessibility</b>								
Pedestrian Catchment Area - 750m Isodistance (PCA) [ha]	-	-	-	-	-	-	0.36	0.00
Employment in PCA [# jobs]	-	-	-	-	-	-	-0.03	0.00
Population in PCA [people]	-	-	-	-	-	-	0.06	0.00
<b>Regional accessibility</b>								
Total employment accessible by transit in 30 min (# jobs)	-	-	-	-	-	-	-0.04	0.00
Total employment accessible by car in 30 min (# jobs)	-	-	-	-	-	-	-0.09	0.00
Pseudo R <sup>2</sup>	0.659						0.643	
Log-likelihood	-1,376,624						-1,442,723	

In the class membership model, the first class has positive coefficients for the local population and employment with a negative coefficient for the PCA. In contrast, class three has opposite coefficients signs for the same attributes. Employment and population within a PCA have a very skewed distribution with the mean much higher than the median while, in comparison, the PCA appear more centered in the mean. Therefore, Class 1 is more likely when employment and population are high and especially when density is high due to smaller PCA. Additionally, Class 1 membership increases with transit employment accessibility, suggesting that more central areas can be associated with this class. In contrast, class 3 would be less likely in the case of high employment and population in local accessibility, and more likely when the density is lower.

The characterization of each class can be described as follows:

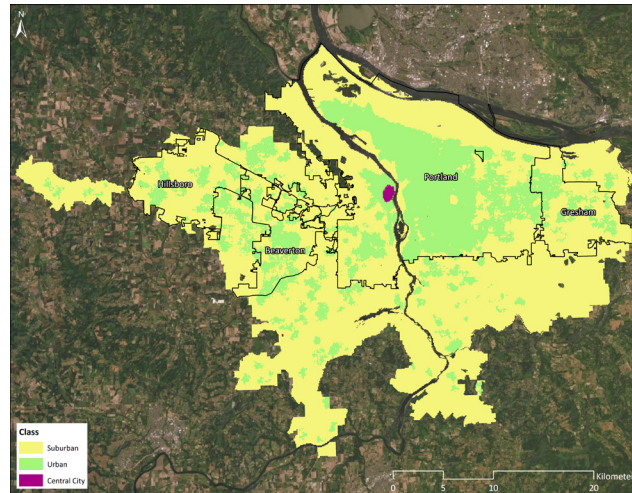
- Class 1: High population and employment density; high regional transit accessibility; lower levels of accessibility by car.
- Class 2: Moderate levels of population and employment density; moderate levels of transit and car accessibility.
- Class 3: Low density of population and employment; low level of transit accessibility; moderate levels of car accessibility.

To tag the classes according to our interpretation, we use the first three cluster types from Salon (2015): City Center, Urban and Suburban. Class 1 corresponds to the City Center; Class 2 corresponds to Urban areas and Class 3 to Suburbs.

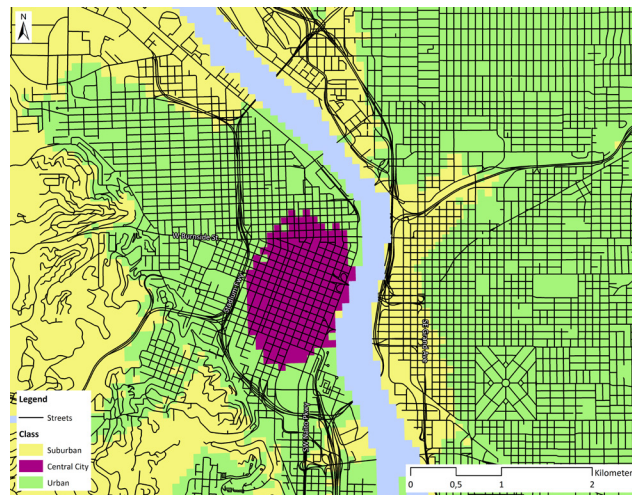
To further characterize the classes, Figure 3 shows the map of the most probable class membership for each cell. The Central City class (purple in the figure) coincides with the center of the metropolitan area, including downtown, the central financial district and the older area of the city (shown with zoom in Figure 4). As the city center, it is the hub of the transit network and therefore has the region's highest

level of transit service. This area is delineated by the Willamette River to the east, a trench freeway that cuts the surface connectivity to the west.

Surrounding the Central City class is the Class 2 - Urban (green in Figure 3 and Figure 4), which consists of neighborhoods characterized by a moderate population density and some corridors that provide retail and services. Finally, the outer part of the urban area is Class 3 - suburban (yellow in Figure 3 and Figure 4), which corresponds to zones with low activity intensity and dwellers. Some smaller urban areas (green spots in Figure 3) surrounded by Suburban class areas correspond to a location with some level of concentration of activities of people (regional subcenters).



**Figure 3.** Portland most probable class membership for each cell



**Figure 4.** Zoom in to Portland Central City for the most probable class for each cell from previous figure

Figures 5, 6, and 7 show the spatial distribution of the membership probability for each of the three classes, respectively. Central city is the most probable class in under 1% (0.16%) of the total land, while the Urban class is the most probable in 26% of the cells. The suburban class dominates the region, being the most likely in 73% of the land area. The Central City class (Figure 5) is not only the most probable in the central area, but it is also only likely in the central city and adjacent areas. In the rest of the metropolitan area, a Central City class membership is quite improbable (<15%). The urban class (Figure 6) has little competition from other classes in the areas where it is the most likely. In the areas where the suburban class is the most probable, there appears to be more competition with the urban class, except for the limits of the metropolitan area into rurality where the suburban class is highly likely.

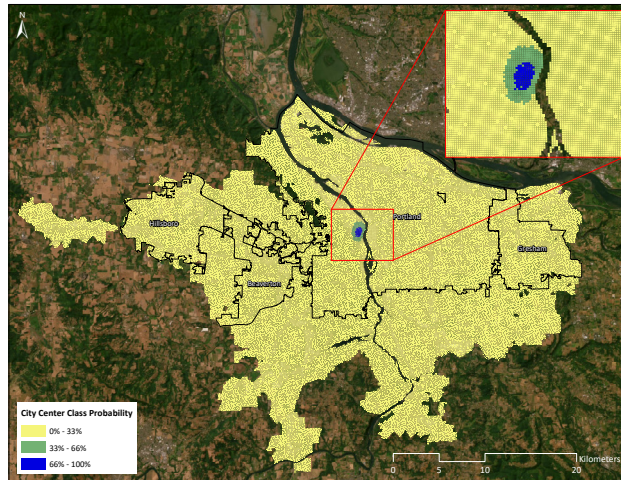


Figure 5. Portland Central City class membership probability

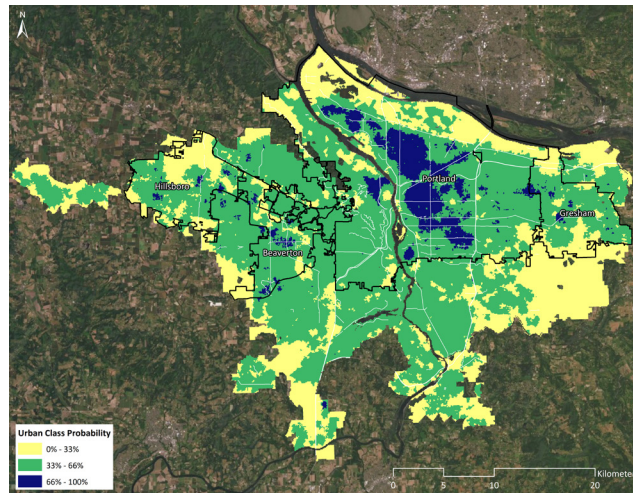


Figure 6. Portland urban class membership probability

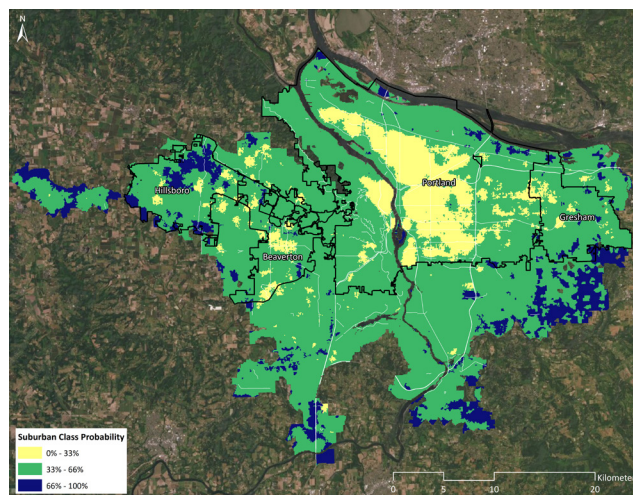


Figure 7. Portland suburban class membership probability

Figure 8 shows the distribution of accessibility measures by highest probability class for each cell. Figures A, B, and C show local measures, while D, and E display regional measures. The Central City class has the highest accessibility in every measure, the urban class tends to be in the middle, and the suburban class displays the lowest values. Employment in the pedestrian catchment area (A) has lower magnitudes in the Suburban and the Urban classes compared to the Central City class, meaning that a large concentration of these measures is in the central city. The pedestrian catchment area (B), which works as a connectivity measure, shows a clear difference between the three classes. If the area were only an open space, it would be a circle of 750 meters radii, approximately 176 hectares. Because the urban form of the city has blocks, the maximum possible connectivity is less than 125 hectares. The median value of the buffer area of the network is around 25 hectares for the Suburban class, 65 hectares for the urban class and approximately 100 hectares in the Central City class. These differences indicate how relevant local connectivity is to class membership.

The Urban class also shares a range of values with those of the Central City class in other measures, such as the population in the pedestrian catchment area (C) and the regional accessibility of automobile (D) and transit (E). Then, while controlling for all other variables, higher local employment levels define the Central City class. The suburban class also shares a range of measures with the values of the Urban class, except for regional transit accessibility, which is particularly low, with most of the values close to zero, indicating that the transit service is nonexistent there.



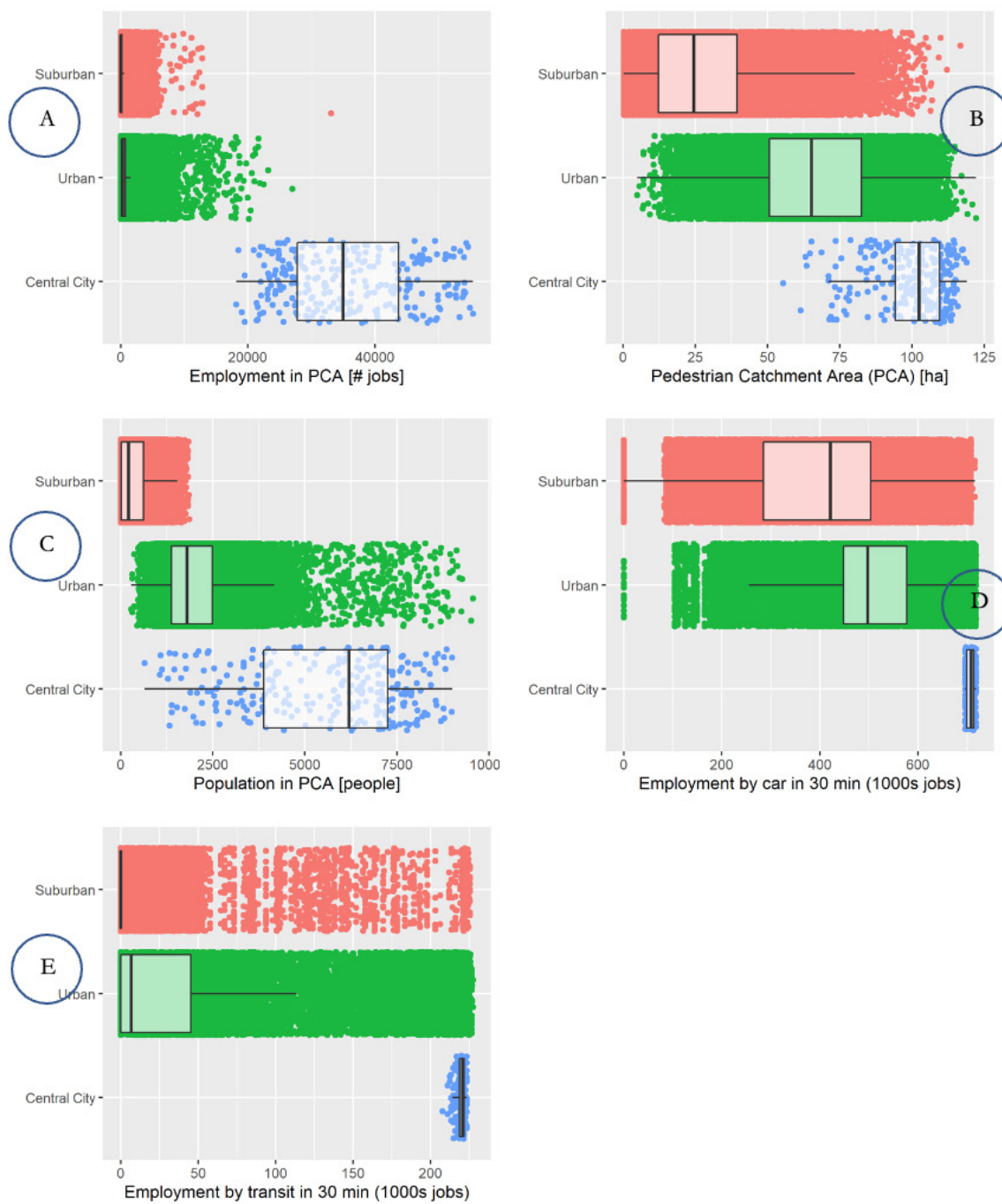


Figure 8. Distribution of local and regional accessibility measures in each cell of the most probable class

## 6.1 Mode choice

Table 4 also shows the results for the class-specific mode choice model of the SLC approach and the MNL. In both approaches, transit is set up as the reference alternative. The MNL includes the local and regional accessibility variables directly in the walking and car utility functions, unlike the SLC, where they are included in the class membership model.

The effect of spatial heterogeneity is most notable when the direction of the effect (i.e., the sign of the parameter) of socioeconomic attributes is different across spatial classes. For example, while being in the age group 25 to 65 increases the probability of walking in the Central City (when compared with the reference, 65+ group), it has the opposite effect in Urban and Suburban class areas. Similarly, belonging to the 25-65 age group diminishes the probability of using the car in Central and Suburban areas, while the opposite happens in Urban class areas. Another interesting example of this is the fact that men are less likely to walk in Central City than women while, in Urban and Suburban areas the effect is the opposite.

Less notable, but still relevant, is the difference in magnitude of the preference parameters across classes. For example, households of low income are less likely to use the car everywhere, but this effect is much stronger in Central and Suburban areas.

The MNL model, while able to capture the effect of socioeconomic characteristics and accessibility levels in mode choice, is unable to measure spatial heterogeneity unless an interaction between all possible combinations is explicitly included in the model specification. While this is possible to do, results become difficult to interpret and spatial segmentation is not possible.

For validation, we re-estimated the model with a randomly chosen subset of 80% of the sample and left the remaining 20% for testing purposes. The pseudo R2 of the models when used in prediction mode on the observations of the left-out-sample is 0.653 for the SLC model and 0.637 for the MNL, indicating that the SLC approach not only performs better in terms of fit, but also provides more accurate forecasting.

## 6.2 Scenario analysis

We develop a scenario analysis by plotting all possible combinations of socioeconomic attributes versus median levels in accessibility measures. The models are applied to a range of distances for 18 scenarios by combining all the levels of age (3 categories), income (3 levels), and gender (2 categories) totaling 18 scenarios. Figure 9 shows the results for the SLC and Figure 10 shows the results for MNL. Each line in every subfigure corresponds to a unique case of age, income, and gender totaling 18 lines. Each column in each of the plots corresponds to a travel mode (walk, transit, or car), and each row corresponds to a class (Central City, Urban, and Suburban).

For each class, we propose a scenario using the median values of accessibility measures in each area where each class is the most probable. These median values correspond to the middle line of the boxplots in Figure 8. These values are applied to both the SLC and MNL models.

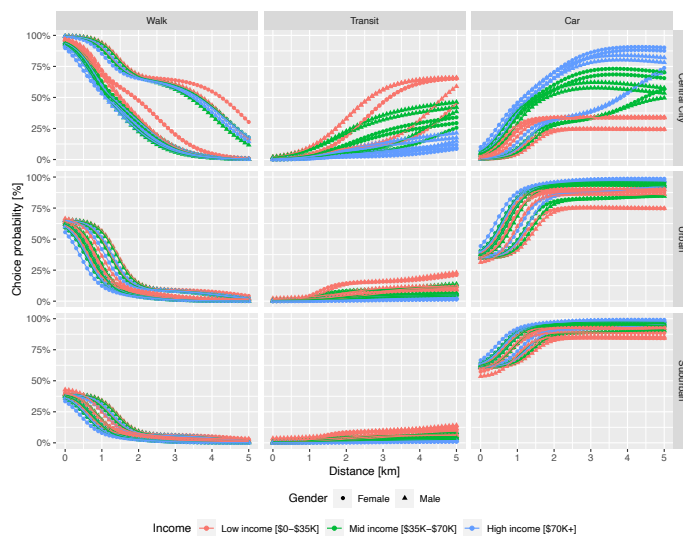


Figure 9. Probabilities of SLC mode choice versus travel distance – median accessibility values

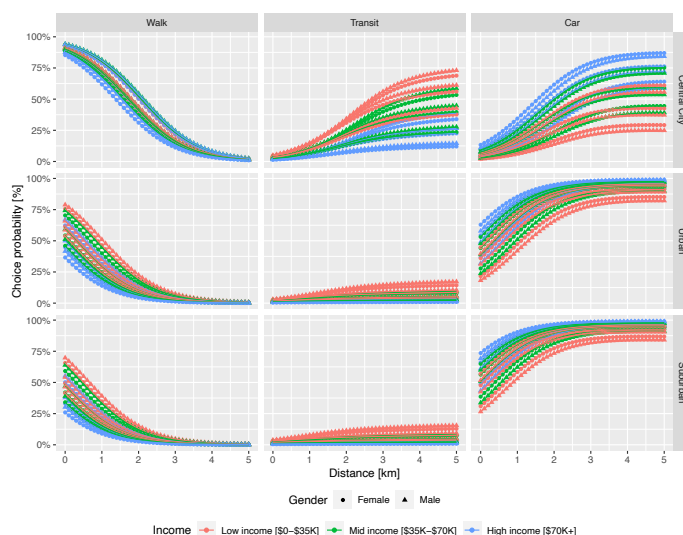


Figure 10. Probabilities of MNL mode choice versus travel distance – median accessibility values

In Central City, with high local and regional accessibility, it is very likely that walking is chosen for short trips in all socioeconomic groups in SLC and MNL. However, in the SLC, there are socioeconomic groups in which walking is much more likely than in the MNL. Therefore, the interaction of accessibility measures and socioeconomic characteristics of this class is less prominent in the MNL than in the SLC.

Furthermore, the difference between SLC and MNL is that the former has a lower variability across socioeconomic attributes than the latter for very short trips. While SLC has a walking probability for trips under 250 meters between 55-65% in urban and between 30%-40% in the suburban case for all socioeconomic groups, the difference in the MNL is between 30-80% in urban and 25-75% in suburban.

The same differences occur in the case of the car. In SLC, the likelihood of choosing a car is between

30-45% for short trips in urban areas and between 55-70% in suburban areas. On the contrary, the MNL case can be between 20-60% for the Urban and 25-75% for Suburban. Additionally, in the case of car and walking, the probability tends to be more stable in the very short trips section and steeper in the case of the SLC, while the change in MNL is more gradual. Although MNL is sensitive to the environment, it is less precise, especially on short trips that are more likely to be walked.

The exercise in this research illustrates how the spatial heterogeneity affects the characterization between mode choices and accessibility measures. Higher accessibility is associated with high use of transit and walking in both model approaches. However, in the case of the MNL, these differences are not as precise as in the SLC model.

## 7 Discussion

The proposed modeling approach helps to better understand the choice of mode depending on income and other socioeconomic characteristics, conditional on the attributes of the origin and the distance of the trip. Previous studies indicate that travel behavior is spatially heterogeneous and that different socioeconomic groups react differently depending on the spatial context and can have different sensitivities to distance, especially in walking choice. The proposed method helps confirming this trend and provides a framework that measures heterogeneity in travel preferences, while simultaneously segmenting space according to the travel patterns that are more likely in each location, as a function of spatial attributes.

In the case study for this research, the Central City class is associated with a greater use of sustainable transportation. However, the areas where this class is likely only represents a small part of the urban area in the study. The total area adds up to approximately 1.5 km<sup>2</sup> (less than 0.2% of the surface of the study area). It is neither necessary nor plausible to think that the whole city should transform to high intensity in order to decrease the dependence on cars. However, it is remarkable how small this area is—where transit and walking are an alternative to driving—and how little is the use of sustainable transportation outside of it. The Urban class is the transition between the Central City and the Suburban class. Both, Urban and Suburban classes, have generally low regional transit accessibility. Therefore, the likelihood of using transit is much lower.

Furthermore, it is interesting that walking shows a lower distance threshold in the Urban class than in the Central City class. In the example from Figure 9 and Figure 10, the chances of walking are almost 0 when the length of the trip reaches 2 km for the urban class. On the contrary, the probability of walking for trips over 2 km is significant in the Central City. The higher probability in the Central City could be due to the more substantial number of walking destinations within a short distance. This process reflects something counterintuitive: people are willing to walk further when they have more options nearby. It is important to notice how the proposed approach (SLC) is able to capture this trend while also, in general, it allows to measure more heterogeneity in behavior across individuals than the MNL approach (compare Walk in the Central City between figures 9 and 10).

The approach proposed here could be used to improve the Portland Metro Regional Demand Model, which does not account for spatial heterogeneity but, instead, considers the same utility function specifications for the whole area of study. The agency's approach is unlikely to capture the interaction between socioeconomic attributes and the location where the trip starts, especially in the central city where short trips (relevant for attaining sustainability goals) can be misestimated.

The SLC approach improves the precision in the characterization of the mode choice when compared with the MNL. The MNL could lead to results that over- or underestimate modal splits. This imprecision in the relationship could support the notion of a weak relationship between urban accessibility and travel choices. In addition, the SLC approach not only allows for a more refined interpretability of

model results, but also allows to delimit spatial areas associated with certain travel patterns and improves the prediction accuracy.

A limitation of this study is the absence of any treatment for residential self-selection. There is evidence that attitudes towards specific travel modes are endogenous to the built environment, making its use in models problematic (Kroesen, 2019). Furthermore, the analysis considers all travel purposes, which means that trip ends are not necessarily anchored to the residential location of the traveler. Therefore, no particular treatment for self-selection was included. Regarding the discussion of self-selection in other studies, it is crucial to consider that highly walkable environments in cities like Portland may be so scarce that identifying different typologies of highly walkable urban environments can be very challenging. Thus, future research considering self-selection should acknowledge that, if the supply of housing in highly walkable areas is less than 1% of the region, which are not necessarily affordable, self-selection may not always be feasible.

It is essential to acknowledge that these results are valid only for the Portland region. It is a mid-sized metropolitan area in the US and one of the few with an urban growth boundary. In the US context, the region may be more compact than others of a similar size. The internal variability of contextual attributes is not the same across regions, especially if they are in different countries or continents. Thus, the results of class membership and metropolitan structure in Portland, with a clear high-activity concentration center and a smooth gradient toward the skirts of the region, could be unique to this case.

Finally, transit could have even more variability across urban areas than population or activities. For example, Portland has a small area with frequent transit service, which tends to be better than other cities in the US, but worse than other metropolitan areas in Latin America, Europe, or Asia. As the class membership model here is transit sensitive, its variability may play an essential role in changing how the classes are defined and the relative importance of other attributes that should be further analyzed in.

Future research should identify and characterize areas with different behavioral responses. This effort must come with a proper conceptualization of the mechanism that affects a specific behavioral response in each context. Characterizing different types of urban environment associated with a specific behavioral response could evolve into urban planning standards that can encourage certain types of travel. Consequently, this could help to (partially) avoid the need for complex (and expensive) mathematical modeling of urban transport systems. Future work should aim to develop these guidelines.

## **Acknowledgments**

We would like to express our sincere gratitude to the Centre for Sustainable Urban Development (CE-DEUS), funded by the Chilean National Agency for Research and Development (ANID) through its FONDAP program (1522A0002), for providing us with the necessary resources and support to carry out this research. We also thank the Becas Chile program, also funded by ANID, for providing financial support to the first author during their doctoral studies.

We would also like to extend our appreciation to the numerous peer reviewers who provided valuable feedback and constructive criticism during the review process. Their input helped us to significantly improve the quality of this manuscript.

## References

- Adams, M. A., Todd, M., Kurka, J., Conway, T. L., Cain, K. L., Frank, L. D., & Sallis, J. F. (2015). Patterns of walkability, transit, and recreation environment for physical activity. *American Journal of Preventive Medicine*, 49(6), 878–887. <https://doi.org/10.1016/j.amepre.2015.05.024>
- Ben-Akiva, M. E., & Lerman, S. R. (1985). *Discrete choice analysis: Theory and application to travel demand*. Cambridge, MA: MIT Press.
- Bento, A. M., Cropper, M. L., Mobarak, A. M., & Vinha, K. (2005). The effects of urban spatial structure on travel demand in the United States. *Review of Economics and Statistics*, 87(3), 466–478. <https://doi.org/10.1162/0034653054638292>
- Bierlaire, M. (2018). *PandasBiogeme: A short introduction (TRANSP-OR 181219; Series on Biogeme*, p. 22). Lausanne, Switzerland: Transport and Mobility Laboratory School of Architecture, Civil and Environmental Engineering Ecole Polytechnique Fédérale de Lausanne.
- Brown, B. B., Werner, C. M., Smith, K. R., Tribby, C. P., Miller, H. J., Jensen, W. A., & Tharp, D. (2016). Environmental, behavioral, and psychological predictors of transit ridership: Evidence from a community intervention. *Journal of Environmental Psychology*, 46, 188–196. <https://doi.org/10.1016/j.jenvp.2016.04.010>
- Cao, X. (2010). Exploring causal effects of neighborhood type on walking behavior using stratification on the propensity score. *Environment and Planning A: Economy and Space*, 42(2), 487–504. <https://doi.org/10.1068/a4269>
- Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, 2(3), 199–219. [https://doi.org/10.1016/S1361-9209\(97\)00009-6](https://doi.org/10.1016/S1361-9209(97)00009-6)
- Choi, K. (2018). The influence of the built environment on household vehicle travel by the urban typology in Calgary, Canada. *Cities*, 75, 101–110. <https://doi.org/10.1016/j.cities.2018.01.006>
- Clifton, K. J., Singleton, P. A., Muhs, C. D., & Schneider, R. J. (2016). Representing pedestrian activity in travel demand models: Framework and application. *Journal of Transport Geography*, 52, 111–122. <https://doi.org/10.1016/j.jtrangeo.2016.03.009>
- Cox, T., & Hurtubia, R. (2021). Latent segmentation of urban space through residential location choice. *Networks and Spatial Economics*, 21, 199–228. <https://doi.org/10.1007/s11067-021-09520-1>
- Cox, T., & Hurtubia, R. (2022). Compact development and preferences for social mixing in location choices: Results from revealed preferences in Santiago, Chile. *Journal of Regional Science*, 62(1), 246–269. <https://doi.org/10.1111/jors.12563>
- Ding, C., Cao, X., Yu, B., & Ju, Y. (2021). Non-linear associations between zonal built environment attributes and transit commuting mode choice accounting for spatial heterogeneity. *Transportation Research Part A: Policy and Practice*, 148, 22–35. <https://doi.org/10.1016/j.tra.2021.03.021>
- Duranton, G., & Turner, M. A. (2018). Urban form and driving: Evidence from US cities. *Journal of Urban Economics*, 108, 170–191. <https://doi.org/10.1016/j.jue.2018.10.003>
- Eom, H.-J., & Cho, G.-H. (2015). Exploring thresholds of built environment characteristics for walkable communities: Empirical evidence from the Seoul metropolitan area. *Transportation Research Part D: Transport and Environment*, 40, 76–86. <https://doi.org/10.1016/j.trd.2015.07.005>
- Etzioni, S., Daziano, R. A., Ben-Elia, E., & Shifan, Y. (2021). Preferences for shared automated vehicles: A hybrid latent class modeling approach. *Transportation Research Part C: Emerging Technologies*, 125, 103013. <https://doi.org/10.1016/j.trc.2021.103013>
- Ewing, R., & Cervero, R. (2010). Travel and the built environment: A meta-analysis. *Journal of the*

- American Planning Association*, 76(3), 265–294. <https://doi.org/10.1080/01944361003766766>
- Ewing, R., Greenwald, M. J., Zhang, M., Walters, J., Feldman, J., Cervero, R., & Thomas, J. (2009). *Measuring the impact of urban form and transit access on mixed use site trip generation rates—Portland pilot study*. Washington, DC: US Environmental Protection Agency.
- Feuillet, T., Commenges, H., Menai, M., Salze, P., Perchoux, C., Reuillon, R., , & Oppert, J. M. (2018). A massive geographically weighted regression model of walking-environment relationships. *Journal of Transport Geography*, 68, 118–129. <https://doi.org/10.1016/j.jtrangeo.2018.03.002>
- Giles-Corti, B., Vernez-Moudon, A., Reis, R., Turrell, G., Dannenberg, A. L., Badland, H., ... & Owen, N. (2016). City planning and population health: A global challenge. *The Lancet*, 388(10062), 2912–2924. [https://doi.org/10.1016/S0140-6736\(16\)30066-6](https://doi.org/10.1016/S0140-6736(16)30066-6)
- Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: Contrasts with mixed logit. *Transportation Research Part B: Methodological*, 37(8), 681–698. [https://doi.org/10.1016/S0191-2615\(02\)00046-2](https://doi.org/10.1016/S0191-2615(02)00046-2)
- Guimpert, I., & Hurtubia, R. (2018). Measuring, understanding and modelling the walking neighborhood as a function of built environment and socioeconomic variables. *Journal of Transport Geography*, 71, 32–44. <https://doi.org/10.1016/j.jtrangeo.2018.07.001>
- Huang, R., Moudon, A. V., Zhou, C., & Saelens, B. E. (2019). Higher residential and employment densities are associated with more objectively measured walking in the home neighborhood. *Journal of Transport & Health*, 12, 142–151. <https://doi.org/10.1016/j.jth.2018.12.002>
- Hurtubia, R., Nguyen, M. H., Glerum, A., & Bierlaire, M. (2014). Integrating psychometric indicators in latent class choice models. *Transportation Research Part A: Policy and Practice*, 64, 135–146. <https://doi.org/10.1016/j.tra.2014.03.010>
- Kamakura, W. A., & Russell, G. J. (1989). A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, 26(4), 379. <https://doi.org/10.2307/3172759>
- Kärmeniemi, M., Lankila, T., Ikäheimo, T., Koivumaa-Honkanen, H., & Korpelainen, R. (2018). The built environment as a determinant of physical activity: A systematic review of longitudinal studies and natural experiments. *Annals of Behavioral Medicine*, 52(3), 239–251. <https://doi.org/10.1093/abm/kax043>
- Khattak, A. J., & Rodriguez, D. (2005). Travel behavior in neo-traditional neighborhood developments: A case study in USA. *Transportation Research Part A: Policy and Practice*, 39(6), 481–500. <https://doi.org/10.1016/j.tra.2005.02.009>
- Kim, S., & Rasouli, S. (2022). The influence of latent lifestyle on acceptance of mobility-as-a-service (MaaS): A hierarchical latent variable and latent class approach. *Transportation Research Part A: Policy and Practice*, 159, 304–319. <https://doi.org/10.1016/j.tra.2022.03.020>
- Koohsari, M. J., Sugiyama, T., Sahlqvist, S., Mavoa, S., Hadgraft, N., & Owen, N. (2015). Neighborhood environmental attributes and adults' sedentary behaviors: Review and research agenda. *Preventive Medicine*, 77, 141–149. <https://doi.org/10.1016/j.ypmed.2015.05.027>
- Kroesen, M. (2019). Residential self-selection and the reverse causation hypothesis: Assessing the endogeneity of stated reasons for residential choice. *Travel Behavior and Society*, 16, 108–117. <https://doi.org/10.1016/j.tbs.2019.05.002>
- Lefebvre-Ropars, G., Morency, C., Singleton, P. A., & Clifton, K. J. (2017). Spatial transferability assessment of a composite walkability index: The pedestrian index of the environment (PIE). *Transportation Research Part D: Transport and Environment*, 57, 378–391. <https://doi.org/10.1016/j.trd.2017.08.018>
- Lewis, S., & Grande del Valle, E. (2019). San Francisco's neighborhoods and auto dependency. *Cities*, 86, 11–24. <https://doi.org/10.1016/j.cities.2018.12.017>

- Lin, T., Wang, D., & Guan, X. (2017). The built environment, travel attitude, and travel behavior: Residential self-selection or residential determination? *Journal of Transport Geography*, *65*, 111–122. <https://doi.org/10.1016/j.jtrangeo.2017.10.004>
- Marquet, O., & Miralles-Guasch, C. (2015). The walkable city and the importance of the proximity environments for Barcelona's everyday mobility. *Cities*, *42*, 258–266. <https://doi.org/10.1016/j.cities.2014.10.012>
- McCormack, G. R., & Shiell, A. (2011). In search of causality: A systematic review of the relationship between the built environment and physical activity among adults. *International Journal of Behavioral Nutrition and Physical Activity*, *8*(1), 125. <https://doi.org/10.1186/1479-5868-8-125>
- McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 104–142). New York: Academic Press.
- Merlin, L. A. (2018). The influence of infill development on travel behavior. *Research in Transportation Economics*, *67*, 54–67. <https://doi.org/10.1016/j.retrec.2017.06.003>
- Oliva, I., Galilea, P., & Hurtubia, R. (2018). Identifying cycling-inducing neighborhoods: A latent class approach. *International Journal of Sustainable Transportation*, *12*(10), 701–713. <https://doi.org/10.1080/15568318.2018.1431822>
- Oregon Modeling Steering Committee. (2011). *Oregon travel and activity survey, 2009–2011*. Salem, OR: Oregon Modeling Steering Committee.
- Owen, A., & Levinson, D. M. (2015). Modeling the commute mode share of transit using continuous accessibility to jobs. *Transportation Research Part A: Policy and Practice*, *74*, 110–122. <https://doi.org/10.1016/j.tra.2015.02.002>
- Poletti, F., Buckley, T., Noriega-Goodwin, D., & Padgham, M. (2020). *Read, validate, analyze, and map files in the general transit feed specification* (R package version 0.7.2.). Retrieved from <https://rdrr.io/github/r-transit/tidytransit/>
- Saelens, B. E., & Handy, S. L. (2008). Built environment correlates of walking: A review. *Medicine and Science in Sports and Exercise*, *40*(7 Suppl), S550–S566. <https://doi.org/10.1249/MSS.0b013e31817c67a4>
- Salon, D. (2015). Heterogeneity in the relationship between the built environment and driving: Focus on neighborhood type and travel purpose. *Research in Transportation Economics*, *52*, 34–45. <https://doi.org/10.1016/j.retrec.2015.10.008>
- Salon, D., Boarnet, M. G., Handy, S., Spears, S., & Tal, G. (2012). How do local actions affect VMT? A critical review of the empirical evidence. *Transportation Research Part D: Transport and Environment*, *17*(7), 495–508. <https://doi.org/10.1016/j.trd.2012.05.006>
- Salvo, G., Lashewicz, B., Doyle-Baker, P., & McCormack, G. (2018). Neighborhood built environment influences on physical activity among adults: A systematized review of qualitative evidence. *International Journal of Environmental Research and Public Health*, *15*(5), 897. <https://doi.org/10.3390/ijerph15050897>
- Sarrias, M. (2019). Do monetary subjective well-being evaluations vary across space? Comparing continuous and discrete spatial heterogeneity. *Spatial Economic Analysis*, *14*(1), 53–87. <https://doi.org/10.1080/17421772.2018.1485968>
- Stevens, M. R. (2017). Does compact development make people drive less? *Journal of the American Planning Association*, *83*(1), 7–18. <https://doi.org/10.1080/01944363.2016.1240044>
- Stevenson, M., Thompson, J., de Sá, T. H., Ewing, R., Mohan, D., McClure, R., ... & Woodcock, J. (2016). Land use, transport, and population health: Estimating the health benefits of compact cities. *The Lancet*, *388*(10062), 2925–2935. [https://doi.org/10.1016/S0140-6736\(16\)30067-8](https://doi.org/10.1016/S0140-6736(16)30067-8)
- Stockton, J. C., Duke-Williams, O., Stamatakis, E., Mindell, J. S., Brunner, E. J., & Shelton, N. J.



- (2016). Development of a novel walkability index for London, United Kingdom: Cross-sectional application to the Whitehall II Study. *BMC Public Health*, 16(1), 416. <https://doi.org/10.1186/s12889-016-3012-2>
- Tanishita, M., & van Wee, B. (2017). Impact of regional population density on walking behavior. *Transportation Planning and Technology*, 40(6), 661–678. <https://doi.org/10.1080/03081060.2017.1325137>
- Wen, C.-H., Wang, W.-C., & Fu, C. (2012). Latent class nested logit model for analyzing high-speed rail access mode choice. *Transportation Research Part E: Logistics and Transportation Review*, 48(2), 545–554. <https://doi.org/10.1016/j.tre.2011.09.002>