The Importance of the Social Environment on Leisure Destination Choice

A Mixed Multinomial Analysis of Homophilic Preferences

Author(s):
Gramsch-Calvo, Benjamin; Axhausen, Kay W.

Publication date:
2024-07

Permanent link:
https://doi.org/10.3929/ethz-b-000682385

Originally published in:
Arbeitsberichte Verkehrs- und Raumplanung 1853

The Importance of the Social Environment on Leisure Destination Choice

A Mixed Multinomial Analysis of Homophilic Preferences

Benjamin Gramsch-Calvo

Kay W. Axhausen

Preprint **July 2024**

IVT Institut für Verkehrsplanung und Transportsysteme
Institute for Transport Planning and Systems

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

The Importance of the Social Environment on Leisure Destination Choice

Benjamin Gramsch-Calvo
IVT
ETH Zürich
CH-8093 Zurich
benjamin.gramsch@ivt.baug.ethz.ch

Kay W. Axhausen
IVT
ETH Zurich
CH-8093 Zurich

July 2024

Abstract

Individuals are fond of belonging to a social environment with a similar social background, which can impact the individual's decision to visit specific venues for leisure activities. Using data from Zurich, we have measured the preference for a social environment in four categories of leisure venues: restaurants, cafes, bars, and nightclubs; the estimation was performed using a mixed multinomial logit model to see how homophily for socioeconomic characteristics can impact the decisions of choosing a leisure venue. The models included three homophilic preferences: age, income, and cultural origin as variables of interest. The results show a positive impact of the three variables studied: age is the most relevant in all venue categories, income shows a higher relevance when individuals choose restaurants or cafes, and cultural background is more important in bars and nightclubs. These results show that the socioeconomic characteristics of the social environment are relevant for the choice of leisure destination. These findings can contribute to the formulation of policies to create more diverse leisure environments and socially cohesive communities.

Keywords

Leisure destination choice, Social Environment, Leisure Segregation, Self-selection

Suggested Citation

B. Gramsch-Calvo and K.W. Axhausen (2024) The Importance of the Social Environment on Leisure Destination Choice: A Mixed Multinomial Analysis of Homophilic Preferences, ETH Research Collection, ETH Zürich.

1 Introduction

Leisure travel is an essential part of urban mobility, accounting for 43% of trips in Switzerland (Bundesamt für Statistik, 2023), with increasing importance over the last decades. One difference between leisure travel and commuting is its heterogeneity in motivations, schedules, and routines; leisure travel has less rigid spatial and temporal constraints, giving more freedom to the individual to perform it at different times and places (Schlich *et al.*, 2004). Regardless of this heterogeneity of motivations, one common characteristic is its social motivation, since most of these activities are carried out in pairs or groups to maintain social connections or expand one's social network (Axhausen, 2005, 2008). The two main types of social influences studied to understand leisure travel have been the ego-alter and the ego-network. The first is mainly used to understand how specific individuals belonging to the ego's social network impact travel behavior, while the second is used to understand macro-level influences of the ego's network. In this paper, we propose a third type of social influence for social-leisure travel: the social environment, which is the sociodemographic characteristics of a group of unknown individuals that share a common space, in this case an urban leisure venue. We base our hypothesis on the interest of individuals to belong to a social context founded on shared experiences, beliefs, or personal characteristics (Mahar *et al.*, 2013), which can be a primary motivation for leisure. However, preferences for venues with specific social environments can lead to self-selection in those venues, contributing to an overall increase in urban segregation.

Since the mid 1970s, segregation has been on the rise in many countries, and understanding its causes and effects has become a significant topic of interest in urban studies (Musterd *et al.*, 2017). Segregation is a complex spatial phenomenon, as it responds to multiple factors, including social, economic, and political, and cannot be reduced to purely residential clusters, as residential segregation does not necessarily imply social segregation (Vaughan and Arbaci, 2011). The use of public and semi-public spaces for leisure can have a significant impact on interactions between different social groups in daily life, as leisure spaces have replaced firms and workplaces as organizing units in society (Florida, 2003), becoming an essential contribution to the social sustainability of cities and a source of face-to-face interaction between unknown people, generating a sense of community (Jacobs, 1961; Nasar and Julian, 1995) and improving the general quality of life of the population (Bramley and Power, 2009). Therefore, understanding the importance of self-selection for the social environment in leisure activities can help create more socially diverse venues to help heterogeneous social interactions thrive.

To test the importance of the social environment in the leisure destination choice process,

we have conducted a survey in Zurich, Switzerland, asking about regularly visited leisure venues. Later, a Mixed Multinomial Logit was estimated; this model measures the impact that homophily, as our variable of interest, has on the decision to go to a venue. Homophilic preferences are multifactorial, with age, income, education, race, and religion being the most important sociodemographic characteristics of tie formations (McPherson *et al.*, 2001). For this manuscript, we focus on three characteristics of the social environment: age, income, and cultural origin; we do not include education, as it is closely related to income, while we have selected cultural origin as a proxy of race and religion, as both characteristics could be hard to disentangle in unknown individuals. These preferences are studied in four types of leisure venues: restaurants, cafes, bars, and nightclubs. The models show that the three homophily variables studied positively and significantly impact the decision to choose these venues for leisure activities and can be relevant to understanding how segregation dynamics are generated in daily activities.

2 Literature review

Urban leisure segregation has been a topic of growing interest in recent years, thanks to the increased availability of mobility data. Most of the research has focused on leisure segregation through access inequalities, which has shown that there are divisions in accessibility to leisure. For example, in Brazil, high-income white individuals have the highest access to leisure activities due to the concentration of venues around the neighborhoods where this group lives and the higher access to motorized private transport. In comparison, low-income black populations live further away from leisure venues and have less access to mobility tools (Bogado Tomasiello and Giannotti, 2023). Wu *et al.* (2023) has studied the gathering capacity of different areas of Shenzhen, showing that high-tech areas of the city have a high gathering capacity for heterogeneous populations, while in terms of specific venues, food and beverage services vary in their capacity to gather individuals from different socioeconomic backgrounds, as consumption habits vary significantly between individuals. The results align with studies on the ethnolinguistic co-presence of individuals, showing that the city center can be an essential space to interact with culturally different individuals (Toomet *et al.*, 2015). These studies have focused on how the built environment and accessibility can impact the segregation of leisure activities.

To study leisure behavior, the literature has focused on the social context and the need

for interaction with other individuals. In order to understand travel decisions for social activities, the main focus has been on ego-alter and ego-network characteristics (Kim *et al.*, 2018). In terms of ego-alter characteristics, the models of social activity participation have estimated different parameters that impact the generation of face-to-face interaction, such as the distance between individuals (Van Den Berg *et al.*, 2010) and the type of relationship (Carrasco and Miller, 2009). Han *et al.* (2023) has focused on the impact of cliques on restaurant choice, showing that cliques tend to choose venues based on the mean distance to the venue of individuals who belong to that clique. The social influence of travel affects the overall individual's mobility patterns; individuals with a more dispersed social network also generate a more dispersed leisure activity space (Gramsch-Calvo and Axhausen, 2024). Furthermore, individuals with more extensive social networks tend to have greater heterogeneity in the type of sites visited and to perform more socially motivated travel (Baburajan, 2019).

In social networks, a recurring topic is homophily; individuals tend to interact and create connections with similar individuals, creating personal networks that are homogeneous in terms of sociodemographic, behavioral, and intrapersonal characteristics, generating clusters in social space (McPherson *et al.*, 2001). Homophily has been studied in different aspects of social life, such as housing markets (Galster *et al.*, 2021; Pinchak *et al.*, 2021), friendship (Currarini *et al.*, 2009), and migrant communities (Vacca *et al.*, 2022), and can have positive and negative effects on social systems; on one hand, it promotes cooperation between individuals and the diffusion of knowledge (Melamed *et al.*, 2020; Korkmaz *et al.*, 2020) but on the other hand, it can generate social segregation (Schelling, 1971).

The decision to conduct leisure in a specific venue can be influenced by a wide variety of temporal, spatial, personal, and social factors, making it almost impossible to estimate choice models that can be generalizable to all choice situations, generating a variety of approaches, methodologies, and theoretical structures adjusted for the context of the models (Barnard, 1987). One of the most common methodologies is the multinomial logit model (McFadden, 1973), which can individually disaggregate the alternatives in the set of choices. Despite this benefit, there are still challenges associated with destination choice models that are not directly related to Random Utility Theory, but to the data used to estimate them. Most research using this technique still uses assumptions from gravity theory and zonal-level characteristics, such as the number of venues, population, and OD impedance variables of interest (Wang and Miller, 2014), which can be misleading for low-density areas (Molloy and Moeckel, 2017). To improve the information about the choice process, researchers have started using social networks to collect data on variables of interest, allowing them to increase the available information (Rashidi *et al.*, 2017);

however, these models still rely primarily on zonal characteristics. In this paper, we combine data collected using a survey with venue data collected from a Social Network Service (Google Places) to understand the choice process at the venue level, improving the level of detail of the models by using information that was previously expensive to collect, such as prices and ratings of the venues studied.

Urban leisure segregation has been studied mainly as the co-presence of dissimilar individuals in the same areas of a city. However, this is not sufficient to ensure the mixing of social groups, as leisure activities can be performed in adjacent venues with different social environments. For this reason, it is essential to integrate leisure behavior to understand how social travel can be influenced by the preference to be part of a specific homophilic social environment at the venue level, separating the segregation generated by individual behavior from segregation generated by the built environment.

3 Data description

The data used for this research have been collected from a random sample of individuals living in the Zurich Metropolitan Area. It consisted of a survey to measure participation in leisure activities. For this purpose, a methodology was developed that included a *place generator*, consisting of seven open-ended questions that asked the respondent to name the venues they regularly visit by category. Later, a *place interpreter*, with questions about those venues, including information on activity patterns, such as the time and day of the visit, the reasons to go to those venues, and the mode used to go (Gramsch-Calvo and Axhausen, 2022). One of the questions of interest was to describe the sociodemographic characteristics that the respondent believes they have in common with the other visitors to the venue.

The survey has 975 respondents, with 9,721 venues mentioned, in various categories, including parks, museums, and gyms. The focus of this manuscript is on restaurants, cafes, bars, and nightclubs; we have chosen restaurants and cafes because there is a large offer of venues and it is a widely popular leisure activity, while bars and nightclubs were chosen because of the social nature of the activity, with the possibility to interact with unknown people being a component of the experience. After filtering these categories, there are 3,340 observed choices from 813 individuals, totaling 1,990 different venues. As cafes and nightclubs do not have enough observations to be estimated separately, we have

Table 1: Modelling categories used for the estimation.

Original category	Number of unique venues	Number of choice situations*	Grouped category
Restaurant	1291	1893	Food venues
Café	240	388	
Bar	384	864	Nightlife
Nightclub	75	195	

* The categories were combined for the estimation, but the choice sets were generated only with the same original category (i.e. nightclubs are considered in the nightlife but their choice set is composed only other nightclubs)

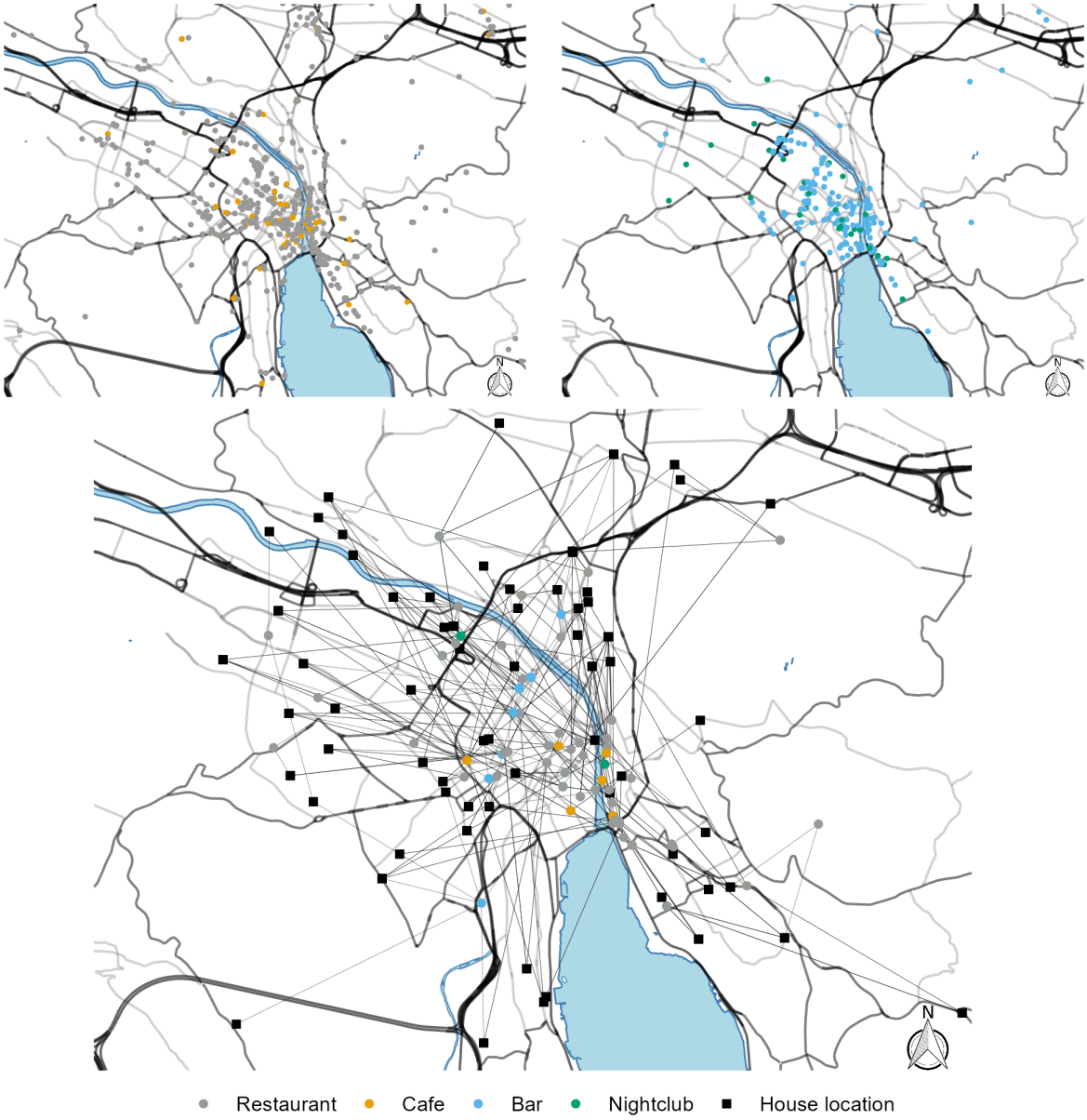
created two joint categories: *food venues* and *nightlife venues*. The grouping and details of each category are presented in table 1.

To obtain the geographic locations, categories, prices, and ratings of the venues, we geocoded the responses with the Google Places API (Cambon *et al.*, 2021). The first information collected by geocoding is the spatial location of the venues. Figure 1 shows the distribution of the venues. Plot A is the spatial distribution of food venues in Zurich; plot B is the nightlife venues. Restaurants and cafes are more spatially dispersed than bars and nightclubs, since food venues have a median distance to their nearest neighbor of 906.6 meters, compared to 779.6 meters for nightlife venues. Plot C is a sample of the network formed by leisure activities; representing how unknown individuals (squares) are connected through leisure venues (circles). By participating in leisure activities, individuals generate the social environment of the venue visited, generating a complex network of interactions and communities.

The second set of information obtained from the geocoding process that we use in our model is the rating and price of the venue. This information, combined with the information provided by the respondents, creates a dataset of venues to analyze. Table 2 explains in more detail the variables used in the model. We have excluded temporal and modal restrictions and variables because we are interested in understanding the reasons for choosing one place as part of the individuals' leisure routine. At the same time, these visits are not necessarily constant in schedule and duration and are not performed with the same transport mode.

To create the attributes of the social environment of each venue, we have used the responses mentioned by other respondents in the survey, assuming that the social environment described by the respondents in the survey is accurate. The attributes were created as

Figure 1: Sample of distribution of venues



follows:

- Age: If the respondent answers that the venue social environment's age is similar to theirs, we assume the social environment of the place is in a range of ±5 years to the respondent's age. For example, if an individual aged 34 mentions that the social environment of restaurant X is similar to theirs, the age of the social environment is between 29 - 39. If two or more individuals mention age similarity about the same place, the range goes from the youngest respondent - 5 to the oldest respondent + 5.
- Income: To create the income social environment of the venues, we used a subjective income variable; the question asks about perceived household income level compared

Table 2: Venue-level variables used in the choice model.

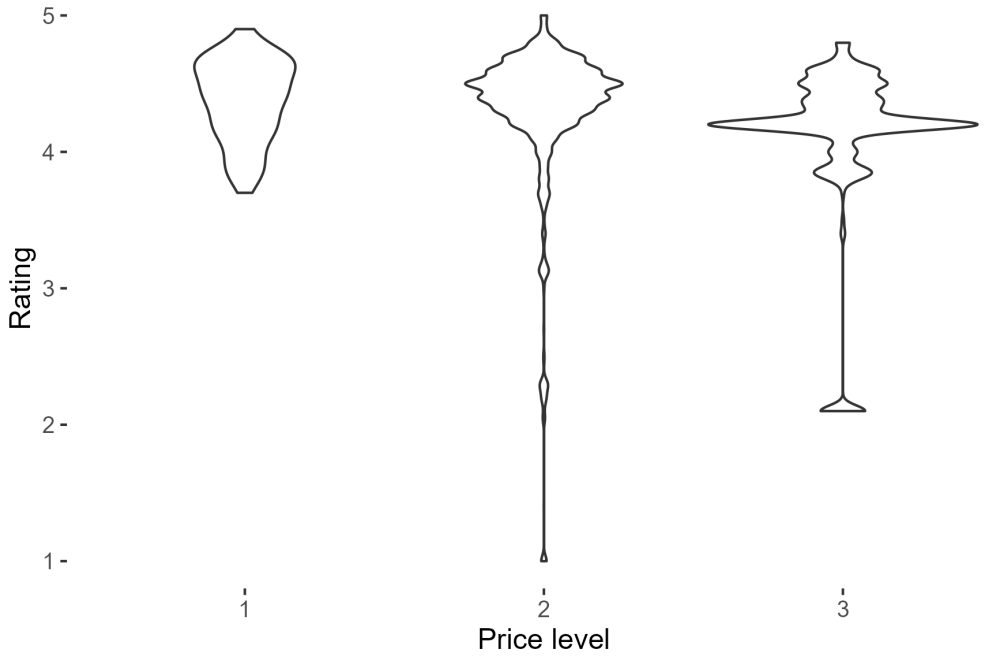
Variable name	Description
Price	Price level of the venues on a scale from 1 to 3.
Rating	Average user rating of the place on a scale from 1 to 5.
Attractiveness	Number of other venues within a 4 km radius of the venue of interest.
Distance	Logarithm of the Euclidean distance from the individual's home location to the venue.
Homophily age	Dummy variable if the individual has an age in the range of the venues' social environment.
Homophily income	Dummy variable if the individual belongs to the socioeconomic group of the social environment of the venues.
Homophily culture	Dummy variable if the individual shares a similar cultural origin with the social environment of the venue.

to the country's average, with five Likert options from *far below average* to *far above average*. We consider this variable instead of the objective income because the perceived socioeconomic position results from a process of acquiring self-perception, and it is not only related to income but to education and individual and family experiences (Ferreira *et al.*, 2018). If an individual mentions that a venue has a similar socioeconomic social environment, we used their answer from the subjective income question as the income level of the venue's social environment. If more than one individual mentions a similar income for the place, we use all answers given.

- Cultural origin: This variable was created using the parents' origin as the variable of interest. The first question asked the respondents was if the parents were foreign-born; if the answer was "no," the person is considered Swiss. If the answer was "yes," the follow-up question asked what macro-region the parents were originally from. Then, if the person answers that the social environment has a similar cultural origin, the venue has a social environment of the culture specified before. Even though the definition of culture has changed with globalization, and geographic borders do not necessarily explain cultural similarities anymore (Sycara *et al.*, 2013), the country's origin of first and second-generation immigrants is still a relevant factor in social connections (Galster *et al.*, 2021). When more than one individual mentions the cultural origin of a venue, we use all answers given.

The first variable analysis compares the rating with the price level of the venues. Figure 2 shows the relationship between these two variables. The rating variable ranges from 1 to 4.9, with most venues in the 4 to 5 range. Regarding the price levels, the variable ranges from 1 to 4. The price level is also a proxy for the type of service offered; restaurants price level 1 are primarily fast food and budget-focus restaurants, while restaurants price level

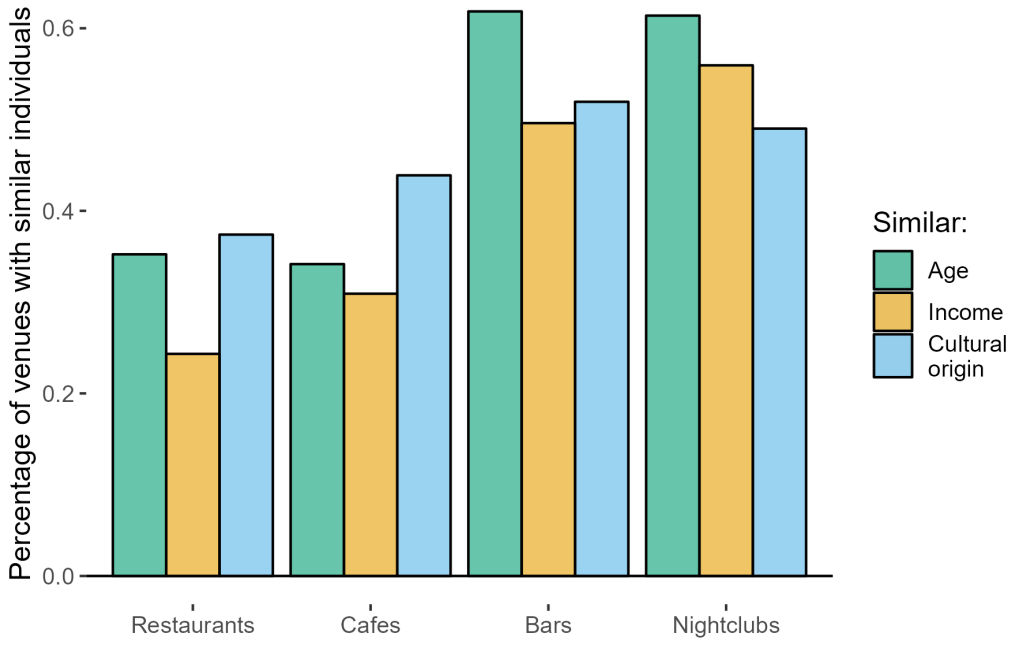
Figure 2: Relation between price and rating of the venues visited



4 are fine dining; but due to the lack of available locations with price-level 4 (10 venues), we have excluded them from the analysis. The correlation between rating and price is negative, but low (-0.08). As the rating is user-based, one could expect that individuals rate venues depending on the price level and service expected. Distance was calculated as the Euclidean distance from the home location of the respondent to the venue; individuals live on a median of 2.9 km from their regular leisure venues.

The homophily variables are presented in figure 3. The figure shows the percentage of places where the individual has stated that there is homophily in the venue's social environment for each characteristic. Nightclubs and bars are the category that has the highest homophily, especially in terms of age, restaurants on the other side have the lowest similarity, with income being the lowest variable. Cafés show a similar distribution of the social environment as restaurants.

Figure 3: Homophily described by individuals



4 Methodology

To estimate the leisure destination choice model, we base our model on the methodology first introduced by Mansky (1977), with a choice-set probabilistically integrated into the choice model. The original model proposed by Mansky is as follows:

$$P_{ni} = \sum_{C} P_n(i|C)\pi_n(C|i,z) \tag{1}$$

In which P_{ni} is the probability that individual n chooses option i, being C_n the evoke set; $P_n(i|C_n)$ is the probability that individual n chooses option i given C_n; and $\pi_n(C_n|i,z)$ is the probability that the choice set of individual n is C_n given the observed choice i and a vector of variables z.

4.1 Choice set selection

As the choice set is unknown, we have to differentiate between the universal set (all available alternatives) and the evoke set (the alternatives that meet certain criteria and are considered by the individual) (Pagliara and Timmermans, 2009). To estimate $\pi_n(C_n|i,z)$ we have generated a choice set with thresholds on the location characteristics (Cantillo and Ortúzar, 2006). The evoke set C_n is :

$$C_n = \{j|X_{nj} \leq H_n\} \tag{2}$$

Where X_{nj} is an attribute of the alternative j for the individual n, H_n is the highest threshold value that the individual n considers for the attribute X. The defined threshold is 1.5 times the distance from the site that the individual visits for leisure. The value was chosen to reduce the possibility that an unchosen alternative is too far from the individual's activity space, generating a model that estimates more discerning individuals regarding this specific attribute, avoiding implausible behaviors (Kimya, 2018), such as always choosing the closest venue. The average choice set size is 398.2 with a maximum of 1182. We tested other choice set generation strategies such as importance and strategic samplings, with this spatial restriction strategy being the one with the most reliable results in terms of rho^2, AIC, and BIC.

4.2 Model description

To estimate the probability of choosing an alternative i, following the formulation of McFadden (1973):

$$P_{ni} = \frac{e^{V_{ni}+ln(\pi_n(C_n|i,z))}}{\sum_{j\in C_n} e^{V_{nj}+ln(\pi_n(C_n|j,z)}} \tag{3}$$

In which V_{ni} is the systematic component of the utility of person n given by the alternative i. As $\pi_n(C_n|i,z)$ includes all alternatives below the threshold H_n, and contains i, it satisfies

the positive conditioning property (if $i, j \in C_n$ and $P(i \in C_n) > 0$, then $P(j \in C_n) > 0$) and the uniform conditioning property ($\pi_i(C_n|i,z) = \pi_i(C_n|j,z)$), then the probability of the choice set is cancelled out. Then, the log-likelihood function is:

$$LL(\Omega,\theta) = \sum_{n=1}^{N} ln\left(\int_{\beta} \left\{ \frac{e^{V_{ni}}}{\sum_{j\in C} e^{V_{nj}}} \right\} f(\beta|\Omega)\, d\beta \right) \tag{4}$$

$$SLL(\Omega,\theta) = \frac{1}{K} \sum_{k=1}^{K} \sum_{n=1}^{N} ln\left(\frac{1}{R} \sum_{r=1}^{R} \left\{ \frac{e^{V_{ni}}}{\sum_{j\in C} e^{V_{nj}}} \right\} f(\beta|\Omega)\, d\beta \right) \tag{5}$$

with

$$V_{n,i} = f_{i,c}^{Price}(x_n^{Age}, x_n^{Income}) \cdot x_{i,c}^{Price} + \beta^{Rating} \cdot x_i^{Rating} + \beta^{Att} \cdot ln(x_i^{Att}) + \beta^{Dist} \cdot ln(x_{i,n}^{Dist}) + \quad \beta_{n,h}^{Homophily} \cdot x_{i,n,h}^{Homophily} \tag{6}$$

And,

$$f_{i,c}^{Price} = (\beta_{i,c}^{Price} + \beta_{i,c}^{Price,Age} * x_n^{Age} + \beta_{i,c}^{Price,Income} * x_n^{Income}) \tag{7}$$

$$\beta_{n,h}^{Homophily} = exp\big(\mu_{ln(\beta_h^{Homophily})} + \sigma_{ln(\beta_{n,h}^{Homophily})} \cdot r_N\big) \tag{8}$$

Where β is the taste coefficient of each variable, x is the vector of attributes for individual n at venue i, when applicable. Price has a subindex c as it has three levels from 1 to 3, while homophily sub-index h represents the three homophilic preferences studied. The scale parameter has a mean value $\mu_{ln(\beta_h^{Homophily})}$, a standard deviation $\sigma_{ln(\beta_{n,h}^{Homophily})}$ and a log-normal distributed individual-specific random component r_N. We tested different distributions of taste heterogeneity, with the log-normal fitting best to the data. The model was estimated using the R package Apollo (Hess and Palma, 2019), the BGW algorithm (Bunch *et al.*, 1993), with 1,000 Halton draws.

5 Model results

5.1 Multinomial logit

Before the mixed model, we have estimated a model with $\beta_{n,h}^{Homophily} = \mu_{\beta_h^{Homophily}}$. The results of the food venues and nightlife models are presented in table 3. For the food services models, the control variables show expected signs; the estimates for rating and attractiveness are positive, while the estimates for price and distance are negative. The interaction terms between price and the socioeconomic variables are positive (except price = 2 and its interactions that are non-statistically significant), showing that older and wealthier individuals prefer more expensive venues than younger and lower income individuals. The nightlife venues show similar results, but income being non-statistically significant in any of the price levels. Compared with food services, we see that age has a higher impact on the preference for more expensive nightlife venues, while income is more relevant for the food services' price preference. Both models show negative estimates regarding distance, with nightlife venues having a lower preference for closer venues. In terms of attractiveness, the estimate for nightlife venues is more than double the estimate for food venues; this could be related to the interest of individuals to visit nightlife venues that are in more vibrant parts of the city.

In terms of the preference for homophily, the nightlife venues model, on average, has higher estimates. In both models, age homophily is the most important variable, with estimates of 1.15 and 1.33, respectively. The second most important homophily variable for food venues is income, with an estimate of 0.69, followed by cultural origin with an estimate of 0.60. For nightlife venues, homophily in cultural origin is more important than income homophily with estimates of 1.13 and 0.36 respectively.

5.2 Mixed Multinomial Logit

After estimating the multinomial logit model, we have included the variables to measure the taste heterogeneity. The results are present in table 4. The goodness of fit of the mixed model presents an improvement in ρ^2 and slight improvements in AIC and BIC. Regarding the results, the mixed model of the food venues shows estimates similar to the

Table 3: MNL Model of homophily in destination choice

	Food services			Nightlife		
	Estimate	s.e.	p-value	Estimate	s.e.	p-value
Price = 2	-0.53	0.70	0.22	-0.63	1.15	0.29
Price = 3	-1.29	0.90	0.08	-2.26	1.49	0.06
Distance	-1.00	0.03	0.00	-0.95	0.07	0.00
Rating	0.26	0.07	0.00	0.42	0.10	0.00
Attractiveness	0.09	0.01	0.00	0.21	0.03	0.00
Age * Price = 2	0.01	0.01	0.14	0.03	0.03	0.09
Age * Price = 3	0.02	0.02	0.09	0.06	0.03	0.02
Income * Price = 2	0.12	0.28	0.33	-0.24	0.68	0.36
Income * Price = 3	0.54	0.34	0.06	0.03	0.74	0.48
Age homophily	1.15	0.08	0.00	1.33	0.09	0.00
Income homophily	0.69	0.08	0.00	0.36	0.10	0.00
Culture homophily	0.60	0.08	0.00	1.13	0.10	0.00
rho-squared	0.11			0.14		
AIC	17738.7			7210.73		
BIC	17805.6			7268.84		

MNL model, except for the estimate of price = 3, which is slightly more negative. In the nightlife model, the value of the estimate of price = 2 is reduced from -0.63 to -1.37, the estimate of income * price = 2 goes from -0.24 to -0.04, and the estimate of income * price = 3 changes from -0.03 to 0.26, but this estimate is not statistically significant.

Due to the interactions that the price attribute has in the model, table 5 presents a comprehensive examination of the utilities of each demographic group in terms of preferences for price levels. As explained above, older and wealthier individuals tend to have preferences for more expensive venues; more specifically, low-income individuals have an average negative preference for venues with price = 2 (compared to price = 1) for all age groups except for 40 and older, who have a positive preference for venues with price = 2. Conversely, high-income individuals over 30 prefer food venues with a price level = 2. Low-income individuals under 60 years of age have a negative preference for food venues with a price level = 3. In contrast, high-income individuals older than 40 prefer these venues over cheaper venues. While for nightlife venues, even if the estimate for price = 2 is negative, only 20-year-olds have a negative preference for this type of venue, as older individuals have a positive preference for venues with price level = 2. For nightlife venues with price = 3, individuals have a positive preference when they are 40 years or older. Finally, we have excluded the income variable from nightlife venues because, as shown in the table 4, the income of the individual does not play a role in the preference for nightlife

Table 4: Mixed MNL Model with heterogeneity in preferences for homophily in destination choice

	Food venues			Nightlife venues		
	Estimate	s.e.	p-value	Estimate	s.e.	p-value
Price = 2	-0.57	0.72	0.22	-1.37	1.17	0.12
Price = 3	-1.47	0.96	0.06	-3.25	1.61	0.02
Distance	-1.04	0.03	0.00	-0.94	0.08	0.00
Rating	0.25	0.07	0.00	0.43	0.10	0.00
Attractiveness	0.09	0.01	0.00	0.20	0.02	0.00
Age * Price = 2	0.02	0.01	0.12	0.05	0.03	0.04
Age * Price = 3	0.03	0.02	0.07	0.08	0.03	0.01
Income * Price = 2	0.12	0.29	0.33	-0.04	0.75	0.48
Income * Price = 3	0.54	0.36	0.07	0.26	0.83	0.38
μ Age	-0.35	0.14	0.00	-0.03	0.13	0.40
μ Income	-1.69	0.42	0.00	-6.40	3.04	0.02
μ Culture	-1.38	0.33	0.00	-0.24	0.19	0.11
σ Age	-1.00	0.11	0.00	-0.93	0.15	0.00
σ Income	-1.91	0.36	0.00	-4.65	1.67	0.00
σ Culture	-1.49	0.24	0.00	-1.30	0.26	0.00
Rho-squared	0.13			0.15		
AIC	17,427			6,816		
BIC	17,511			6,888		

Table 5: Utilities of the socioeconomic groups for the price level of the venues

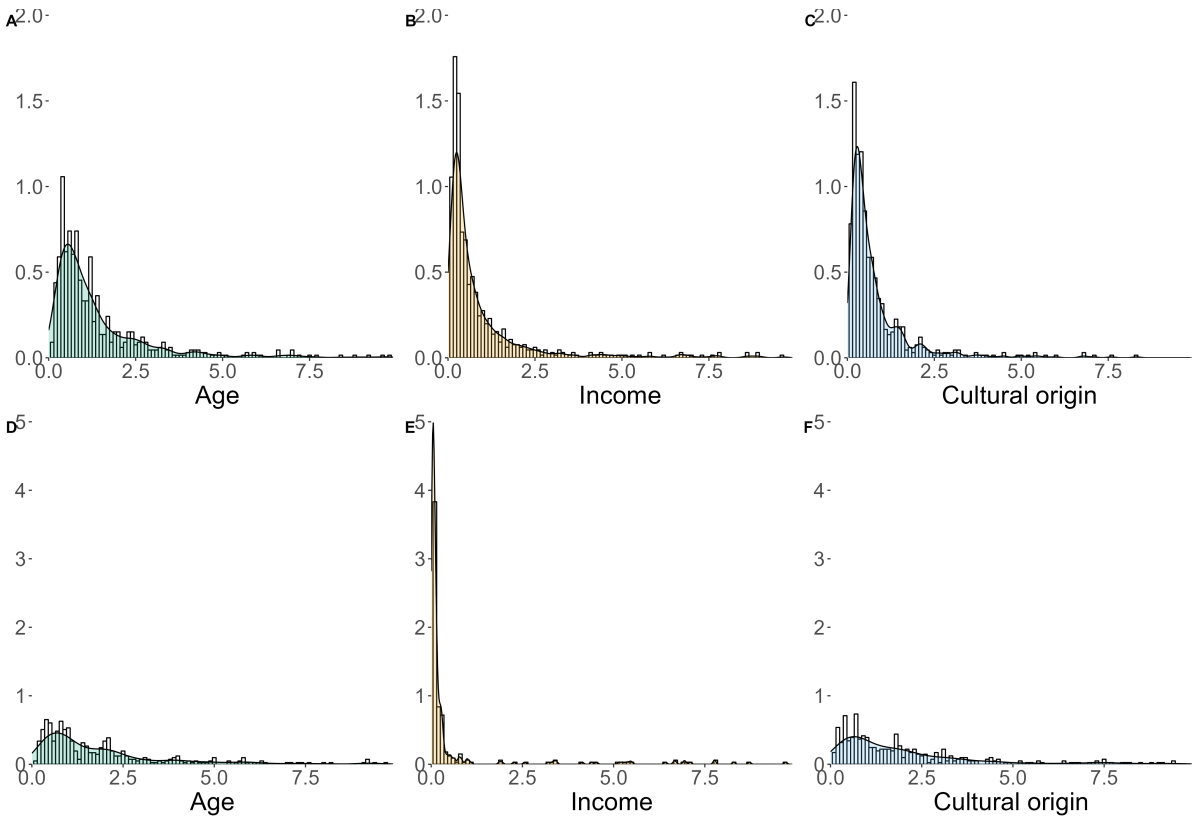
	Food venues					
	Age:	20	30	40	50	60
Price = 2	Low income	-0.26	-0.11	0.04	0.19	0.34
	High income	-0.14	0.01	0.16	0.31	0.47
Price = 3	Low income	-0.96	-0.70	-0.45	-0.20	0.06
	High income	-0.42	-0.17	0.09	0.34	0.59
	Nightlife venues*					
	Age:	20	30	40	50	60
Price = 2		-0.36	0.15	0.65	1.16	1.66
Price = 3		-1.62	-0.80	0.01	0.83	1.65

* Income has been excluded of the Nightlife table as it shows no statistical significance in any of the interactions with price

venues. Changes in preferences, from negative to positive, are due to the price level being a category related to the service offered, as well as the price.

Concerning the variables of interest, the value $exp(\mu)$ provides the mean preference for each homophily estimate, while $exp(\sigma)$ measures the heterogeneity of the preferences. The

Figure 4: Conditional and unconditional distribution of taste heterogeneity



combination of both estimates provides the range of preferences for each of the variables of interest, which is assumed to be log-normal. Regarding food venues, interest in homophily of income shows the lowest heterogeneity of preferences, followed by cultural origin homophily and age homophily. In nightlife venues, the highest preference heterogeneity is age homophily, while income homophily has the lowest value. A visual representation of the heterogeneity estimates is shown in figure 4, where the conditional (bars) and unconditional (density) distributions of the three preferences for homophily are presented by model. Regarding food services, cultural origin and income have high peaks, showing greater homogeneity in preferences compared to age homophily. In the case of the nightlife venues model, age and cultural origin homophily show a similar distribution of the conditionals, with both distributions having a high heterogeneity.

5.3 Willingness to travel

Table 6 presents the value of the willingness-to-travel (WTT) indicators calculated from the mean preference attribute for each homophily preference; this value measures how

Table 6: Homophily-distance willingness-to-travel

	Food venues		Nightlife venues	
	MNL $\frac{\beta_h}{\beta_d}$	Mixed MNL $\frac{\exp(\overline{\mu}_h)}{\beta_d}$	MNL $\frac{\beta_h}{\beta_d}$	Mixed MNL $\frac{\exp(\overline{\mu}_h)}{\beta_d}$
Age homophily	1.15	0.68	1.41	1.02
Income homophily	0.69	0.18	0.38	0.00
Culture homophily	0.60	0.24	1.19	0.84

* As the WTT depends on distance, the willingness to travel is estimated by multiplying these values by $x_{i,n}^{Dist}$

much extra the average individual is willing to travel to be with similar others. These indicators allow to directly compare the preference coefficients and measuring the trade-off between distance and social environment, models, and venues. Comparing the MNL and Mixed MNL model, in the latter model, all willingness-to-travel are relatively smaller than in the former model, due to the distribution of the taste heterogeneity, with cultural origin in food services and income in nightlife venues the most notable changes. When comparing venue categories, nightlife has a higher average willingness to travel, in terms of preferences for age homophily, the WTT of nightlife venues is 1.02 versus 0.68. For food services, the average individual is willing to travel an additional 18% of the distance to be in a social environment with similar income. At the same time, this attribute is negligible for nightlife venues (0.00%). Finally, cultural origin has a lower willingness to travel for food services compared to nightlife venues (24% versus 84% extra distance).

6 Discussion

In this study, we have analyzed the importance of the social environment in destination choice, measured as the preference for homophily of three characteristics of the social environment: age, income, and cultural origin. The two models, food venues and nightlife venues, have shown a positive and significant impact of the three socioeconomic homophilic preferences when choosing a venue to perform leisure, age homophily being the most relevant variable in both models. In contrast, income homophily is relevant when choosing a restaurant or cafe, whereas its importance is negligible when choosing nightlife venues, which coincides with a statistical insignificance of the income interaction with price in those venues; homophily of cultural origin shows higher relevance when choosing bars or nightclubs than when choosing food venues. The difference in homophilic preferences shows

that individuals have different preferences for the social environment depending on the context of the social activity. The higher preferences for homophily in nightlife venues can be related to the social expectations of the activity: bars and nightclubs tend to have more interaction between unknown individuals, and they are better venues to potentially meet new people, while restaurants and cafés depend more on clique interaction than interactions 'between tables.' These distinct preferences are consistent with previous research on social networks showing that homophilic preferences have baseline patterns with differences depending on the type of relationship formed between individuals (McPherson *et al.*, 2001). The distribution of the preferences for homophily also depends on the type of venues; preference for income homophily has a higher variance in food venues than in nightlife venues, while the opposite happens with cultural origin, which shows that the distribution of preferences in food services tends to be more homogeneous than in nightlife venues.

The study makes two contributions to the literature. First, the measurement of homophilic preferences in leisure activities; homophily is an essential aspect of social networks and can be measured in many types of social interactions, but as far as the authors are concerned, this is the first time it has been measured as a characteristic that affects the choice of the specific venue to perform leisure activities. The results can help us understand how individuals choose where to share common spaces with unknown individuals, these interactions are an essential aspect of urban social life as they generate a sense of community and belonging, which could result in socialization and the formation of friendship ties. Previous studies have found that some areas of the city tend to gather individuals from different social backgrounds, but visiting the same areas does not necessarily mean that they perform leisure activities in the same venues. Therefore, more research is needed to understand the origins and measure leisure segregation and define whether segregation can still occur in physical proximity; this is a first approach to this task.

The second contribution is to show the potential of using venue-level information to generate models of leisure destination choice. In this paper, we have mixed survey data with social network services (SNS) data to include three types of characteristics: individual level (age and income), venue level (category, price, area attractiveness and rating), and individual venue level (distance and homophily). The use of venue-level characteristics for leisure destination choice is a methodological advancement in comparison to previous models that analyze meso-destination characteristics, such as demographics and agglomeration effects, because using venue characteristics such as price or rating can increase the number of options available for individuals as well as make comparisons between spatially contiguous venues possible, increasing the level of details the model can provide.

The literature on homophily has shown that preferences lead to network segregation, even when preferences are minor (Goodreau *et al.*, 2009; Hatna and Benenson, 2015); for this reason, it is relevant to understand the preferences of individuals for the social environment of their leisure venues. However, a second potential venue characteristic in leisure activities can lead to age and income segregation: the price category. The preference for price category varies depending on age and income, with the sign of the parameters varying from negative to positive as individuals get older and wealthier; this could generate higher levels of segregation in leisure activities than if only homophilic preferences were to be considered.

We want to point out two main limitations of this study. The first is how the social environment is defined, as collecting data on the social environment for more than 1,500 venues is resource intensive; we had to rely on the information provided by the respondents, which can be prone to measurement errors. The second limitation is the number of socioeconomic characteristics used for homophilic preferences; homophily is a complex phenomenon that depends not only on the socioeconomic characteristics and interests of the individual but on shared values, beliefs or a multifactorial combination of all these variables (Block and Grund, 2014). However, due to the complexity of the topic, this study has shown the importance of the three homophilic preferences studied in the individual's process of leisure destination choice, and it is a first step in understanding how segregation can be generated in daily leisure activities.

7 References

Axhausen, K. W. (2005) Social networks and travel: Some hypotheses, 22 p., ETH Zurich.

Axhausen, K. W. (2008) Social Networks, Mobility Biographies, and Travel: Survey Challenges, *Environment and Planning B: Planning and Design*, **35** (6) 981–996, ISSN 0265-8135, 1472-3417.

Baburajan, V. (2019) An investigation into the influence of land-use, social networks and information and communication technologies on destination choice for social activities, *Transactions of the Association of European Schools of Planning*, 14.

Barnard, P. O. (1987) Modelling shopping destination choice behaviour using the basic

multinomial logit model and some of its extensions, *Transport Reviews*, **7** (1) 17–51, ISSN 0144-1647, 1464-5327.

Block, P. and T. Grund (2014) Multidimensional homophily in friendship networks, *Network Science*, **2** (2) 189–212, ISSN 2050-1242, 2050-1250.

Bogado Tomasiello, D. and M. Giannotti (2023) Unfolding time, race and class inequalities to access leisure, *Environment and Planning B: Urban Analytics and City Science*.

Bramley, G. and S. Power (2009) Urban form and social sustainability: the role of density and housing type, *Environment and Planning B: Planning and Design*, **36** (1) 30–48, ISSN 0265-8135, 1472-3417.

Bunch, D. S., D. M. Gay and R. E. Welsch (1993) Algorithm 717: Subroutines for maximum likelihood and quasi-likelihood estimation of parameters in nonlinear regression models, *ACM Transactions on Mathematical Software*, **19** (1) 109–130, ISSN 0098-3500, 1557-7295.

Bundesamt für Statistik (2023) Mikrozensus mobilität und verkehr.

Cambon, J., D. Hernangómez, C. Belanger and D. Possenriede (2021) Tidygeocoder: An r package for geocoding, *Journal of Open Source Software*, **6** (65) 3544. R package version 1.0.5.

Cantillo, V. and J. D. D. Ortúzar (2006) Implications of Thresholds in Discrete Choice Modelling, *Transport Reviews*, **26** (6) 667–691, ISSN 0144-1647, 1464-5327.

Carrasco, J.-A. and E. J. Miller (2009) The social dimension in action: A multilevel, personal networks model of social activity frequency between individuals, *Transportation Research Part A: Policy and Practice*, **43** (1) 90–104, ISSN 09658564.

Currarini, S., M. O. Jackson and P. Pin (2009) An economic model of friendship: Homophily, minorities, and segregation, *Econometrica*, **77** (4) 1003–1045, ISSN 0012-9682.

Ferreira, W. D. A., L. Camelo, M. C. Viana, L. Giatti and S. M. Barreto (2018) Is subjective social status a summary of life-course socioeconomic position?, *Cadernos de Saúde Pública*, **34** (5), ISSN 1678-4464, 0102-311X.

Florida, R. (2003) The rise of the creative class: And how it's transforming work, leisure,

community and everyday life, *Canadian Public Policy / Analyse de Politiques*, **29** (3) 378, ISSN 03170861.

Galster, G. C., L. M. Turner and A. M. Santiago (2021) Neighbourhood selection by natives and immigrants: Homophily or limited spatial search?, *Housing Studies*, 1–27, ISSN 0267-3037, 1466-1810.

Goodreau, S. M., J. A. Kitts and M. Morris (2009) Birds of a feather, or friend of a friend? using exponential random graph models to investigate adolescent social networks, *Demography*, **46** (1) 103–125, ISSN 0070-3370, 1533-7790.

Gramsch-Calvo, B. and K. W. Axhausen (2022) Place generator & place interpreter: A new survey method to understand regular destination choice, *ETH Research Collection*, 17 p.

Gramsch-Calvo, B. and K. W. Axhausen (2024) Exploring the influence of social relations and knowledge of the urban environment on leisure travel, *Journal of Transport Geography*, **117**, 103860, ISSN 09666923.

Han, C., L. Luo, G. Parady, K. Takami, M. Chikaraishi and N. Harata (2023) Modeling joint eating-out destination choices incorporating group-level impedance: A case study of the Greater Tokyo Area.

Hatna, E. and I. Benenson (2015) Combining segregation and integration: Schelling model dynamics for heterogeneous population, *Journal of Artificial Societies and Social Simulation*, **18** (4) 15, ISSN 1460-7425.

Hess, S. and D. Palma (2019) Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application, *Journal of Choice Modelling*, **32**, 100170, ISSN 17555345.

Jacobs, J. (1961) *The death and life of great American cities*, Vintage Books, New York, NY.

Kim, J., S. Rasouli and H. J. Timmermans (2018) Social networks, social influence and activity-travel behaviour: a review of models and empirical evidence, *Transport Reviews*, **38** (4) 499–523, ISSN 0144-1647, 1464-5327.

Kimya, M. (2018) Choice, consideration sets, and attribute filters, *American Economic Journal: Microeconomics*, **10** (4) 223–247, ISSN 1945-7669, 1945-7685.

Korkmaz, G., C. J. Kuhlman, J. Goldstein and F. Vega-Redondo (2020) A computational study of homophily and diffusion of common knowledge on social networks based on a model of Facebook, *Social Network Analysis and Mining*, **10** (1) 5, ISSN 1869-5450, 1869-5469.

Mahar, A. L., V. Cobigo and H. Stuart (2013) Conceptualizing belonging, *Disability and Rehabilitation*, **35** (12) 1026–1032, ISSN 0963-8288, 1464-5165.

Mansky, C. (1977) The structure of random utility models, *Theory and Decision*, **8**, 229 – 54.

McFadden, D. (1973) Conditional logit analysis of qualitative choice behavior, in *Analysis of qualitative choice behavior*.

McPherson, M., L. Smith-Lovin and J. M. Cook (2001) Birds of a feather: Homophily in social networks, *Annual Review of Sociology*, **27** (1) 415–444, ISSN 0360-0572, 1545-2115.

Melamed, D., B. Simpson, A. Harrell, C. W. Munn, J. Z. Abernathy and M. Sweitzer (2020) Homophily and segregation in cooperative networks, *American Journal of Sociology*, **125** (4) 1084–1127, ISSN 0002-9602, 1537-5390.

Molloy, J. and R. Moeckel (2017) Improving destination choice modeling using location-based big data, *ISPRS International Journal of Geo-Information*, **6** (9) 291, ISSN 2220-9964.

Musterd, S., S. Marcińczak, M. Van Ham and T. Tammaru (2017) Socioeconomic segregation in European capital cities. Increasing separation between poor and rich, *Urban Geography*, **38** (7) 1062–1083, ISSN 0272-3638, 1938-2847.

Nasar, J. L. and D. A. Julian (1995) The psychological sense of community in the neighborhood, *Journal of the American Planning Association*, **61** (2) 178–184, ISSN 0194-4363, 1939-0130.

Pagliara, F. and H. Timmermans (2009) Choice set generation in spatial contexts: a review, *Transportation Letters*, **1** (3) 181–196, ISSN 1942-7867, 1942-7875.

Pinchak, N. P., C. R. Browning, C. A. Calder and B. Boettner (2021) Activity locations, residential segregation and the significance of residential neighbourhood boundary perceptions, *Urban Studies*, **58** (13) 2758–2781, ISSN 0042-0980, 1360-063X.

Rashidi, T. H., A. Abbasi, M. Maghrebi, S. Hasan and T. S. Waller (2017) Exploring the capacity of social media data for modelling travel behaviour: Opportunities and challenges, *Transportation Research Part C: Emerging Technologies*, **75**, 197–211, ISSN 0968090X.

Schelling, T. C. (1971) Dynamic models of segregation, *The Journal of Mathematical Sociology*, **1** (2) 143–186, ISSN 0022-250X, 1545-5874.

Schlich, R., S. Schönfelder, S. Hanson and K. W. Axhausen (2004) Structures of leisure travel: Temporal and spatial variability, *Transport Reviews*, **24** (2) 219–237, ISSN 0144-1647, 1464-5327.

Sycara, K., M. Gelfand and A. Abbe (eds.) (2013) *Models for intercultural collaboration and negotiation*, vol. 6 of *Advances in Group Decision and Negotiation*, Springer Netherlands, Dordrecht, ISBN 978-94-007-5573-4 978-94-007-5574-1.

Toomet, O., S. Silm, E. Saluveer, R. Ahas and T. Tammaru (2015) Where Do Ethno-Linguistic Groups Meet? How Copresence during Free-Time Is Related to Copresence at Home and at Work, *PLOS ONE*, **10** (5) e0126093, ISSN 1932-6203.

Vacca, R., D. Cañarte and T. Vitale (2022) Beyond ethnic solidarity: the diversity and specialisation of social ties in a stigmatised migrant minority, *Journal of Ethnic and Migration Studies*, **48** (13) 3113–3141, ISSN 1369-183X, 1469-9451.

Van Den Berg, P., T. Arentze and H. Timmermans (2010) Location-type choice for face-to-face social activities and its effect on travel behavior, *Environment and Planning B: Planning and Design*, **37** (6) 1057–1075, ISSN 0265-8135, 1472-3417.

Vaughan, L. and S. Arbaci (2011) The challenges of understanding urban segregation, *Built Environment*, **37** (2) 128–138, ISSN 02637960.

Wang, J. and E. J. Miller (2014) A prism-based and gap-based approach to shopping location choice, *Environment and Planning B: Planning and Design*, **41** (6) 977–1005, ISSN 0265-8135, 1472-3417.

Wu, Y., Y. Yao, S. Ren, S. Zhang and Q. Guan (2023) How do urban services facilities affect social segregation among people of different economic levels? A case study of Shenzhen city, *Environment and Planning B: Urban Analytics and City Science*, **50** (6) 1502–1517, ISSN 2399-8083, 2399-8091.