Article:
Ibraimovic, T and Hess, S orcid.org/0000-0002-3650-2518 (2018) A latent class model of residential choice behaviour and ethnic segregation preferences. Housing Studies, 33 (4). pp. 544-564. ISSN 0267-3037

https://doi.org/10.1080/02673037.2017.1373749

A Latent Class Model of Residential Choice Behaviour and Ethnic Segregation Preferences

Tatjana Ibraimovic [1,2,*], Stephane Hess [1]

[1] Institute for Transport Studies & Choice Modelling Centre, University of Leeds, Leeds, UK
[2] Institute of Economic Research, University of Lugano, Lugano, Switzerland

E-mail: Ibraimovic.Tatjana@gmail.com (T. Ibraimovic), S.Hess@its.leeds.ac.uk (S. Hess)

*Corresponding author. Current address: Via Emilio Rava 1D, CH-6962 Viganello – Lugano, Switzerland, Tel. +41 91 234 61 17, Email. ibraimovic.tatjana@gmail.com

Abstract

The nature of ethnic residential clustering involves diverse population segments which through their location decisions influence the spatial patterns of ethnic settlements. While residential location is in part determined by outside constraints, choice plays a role too, making the study of preferences an important research topic. Along with differences in socio-economic characteristics, literature often emphasises the role of unobserved (behavioural) elements in the formation of preferences for ethnic neighbourhood composition. This paper tests the potential of latent class choice models to examine both observed and unobserved heterogeneity in residential choices across ethnic groups. The empirical example is estimated on stated preferences data from Lugano, Switzerland. The results indicate different ethnic attributes as key choice drivers for households belonging to three latent classes, where the origin of households is the best predictor of class membership. Swiss citizens are mainly concerned about high shares of foreigners, advantaged foreigners favour their co-nationals, while disadvantaged foreigners hold both of such preferences.

JEL classification: D110, D120, R200, R210, R230

Key Words: Ethnic Residential Segregation - Latent Class Choice Model - Ethnic Preferences Heterogeneity - Residential Location Choice Behaviour – Stated Preferences Choice Experiment

Acknowledgements

The authors are grateful to the three anonymous reviewers, participants at the 53rd ERSA Congress 2013 in Palermo, Italy and Applied Urban Modelling Conference 2013 in Cambridge, UK, for their valuable discussions which helped substantially improve the paper. We are grateful to the Municipality of Lugano and Canton Ticino for the data sources provided. The majority of the work reported in this paper was carried out while the authors were working at the Institute for Transport Studies at the University of Leeds, UK and Institute of Economic Research, University of Lugano, Switzerland. This work was supported by the Swiss National Science Foundation and by the European Research Council through the consolidator grant 615596-DECISIONS.

1. Introduction

Among a variety of factors that can influence ethnic residential segregation in urban areas, the debate focuses on two main segregation drivers: households' preferences for a specific ethnic mix for their neighbourhood and the constraints they face in accessing the most desired locations. As documented by many international studies (Farley *et al.* 1978; Clark 1992, 2009; Charles 2000, 2003; Bolt and Van Kempen 2003; Zorlu and Mulder 2008; Schaake *et al.* 2010; Doff and Kleinhans 2011), a key role in explaining the voluntary groupings is played by preferences of households to reside in proximity to their community of origin. The ethnic composition of the neighbourhood thus represents one of the characteristics people account for when choosing their residential location (Åslund 2005; Ibraimovic and Masiero 2014). On the other hand, involuntary segregation can be caused by accessibility constraints, such as discrimination in the housing market or the weaker socioeconomic position of some ethnic minorities with respect to the mainstream hosting society (Darden 1986; Massey and Denton 1993; Charles 2003; van der Laan Bouma-Doff 2007).

The debate over voluntary and involuntary causes of ethnic segregation in Europe is still open, in academia as well as in the political debate. In fact, these two forces can and often do act together assuming different importance and magnitude in different countries and urban contexts as well as for different ethnic communities. Indeed, in some contexts and for some ethnic communities, segregation can be largely due to preferences (e.g. whites in the USA and natives in Europe as shown by Charles [2000] and van der Laan Bouma-Doff [2007]), while for other ethnicities, especially those who exhibit a weaker socioeconomic position, choices can be significantly constrained. Nevertheless, the literature has shown how puzzling it is to identify and untangle the real causes driving segregation phenomena, in part due to the opposite effects they may have. A greater understanding of preferences can help understand the role of choice in this process, while also shedding some light on the difference between people's aspirations and their actual situation.

In this study we aim to explore the aspect of voluntary self-segregation, guided by preferences for co-ethnics. To allow us to separate out preferences from involuntary segregation, we need to rely on experimental data rather than real world location choices, a point we discuss in detail in Section 3. Through the analysis of the existence of a "pure" preferences effect, and the quantification of the relative importance of different neighbourhood characteristics (including monetary valuations), we can gain unique insights into preferences. For example, if the willingness to pay (WTP) for living next to the co-ethnics is small, it could indicate that observed segregation of an ethnic community might be determined by involuntary causes, while a high WTP would imply that the voluntary concentration of same ethnicity members may be a more probable reason.

A large body of literature on ethnic segregation not only shows the existence of ethnic preferences, but also indicates substantial heterogeneity in such preferences among diverse ethnic groups (see for e.g. Clark 1992, 2009; Charles 2000; Sermons, 2000). The literature findings suggest that different ethnic groups exhibit different segregation behaviour, some being more prone to clustering among their own ethnic community, others being open to ethnically mixed residential environments

(see, e.g., Farley *et al.* [1978]). Along with the differences in socio-economic characteristics of immigrant households, the segregation literature often emphasises the role of underlying behavioural elements, such as attitudes and perceptions of different ethnic groups in the formation of preferences for the ethnic composition of a residential environment. In that context, Andersen (2015) highlights immigrants' success in social integration as a factor that can lead to different attitudes in relation to ethnic preferences. He also mentions other attitudinal elements such as feelings of belonging to the country of origin as a possible preferences driver. Homophily is another behavioural element that is argued to influence the preferences for ethnic residential environment (Lin and Harris 2008). However all these elements can be difficult to observe and account for in models (especially if multiple factors underlie certain behaviour) leading to possible omission of important heterogeneity drivers. This could prevent the full understanding of the differences in the residential behaviour of a heterogeneous population and, in particular, the taste dissimilarities for ethnic composition of neighbourhoods, essential for analysing the dynamics of ethnic concentrations. The questions that arises in this context is how to account for diverse heterogeneity sources (observed and unobserved to the researcher), that have an impact on preferences for ethnic neighbourhood composition.

A suitable method for studying preferences for co-ethnic neighbours or for a certain ethnic neighbourhood composition is the analysis of residential location choice behaviour (McFadden 1974, 1977). This theory assumes that households select the location providing them with the highest utility among the available alternatives. Indeed, according to the Random Utility Modelling (RUM) framework, the utility assigned by each household to each of the available alternatives will depend on the characteristics of those alternatives and the preferences (sensitivities) of households to these characteristics. Various advances in such methods have been developed with the aim of representing heterogeneity in choice behaviour [1]. Among the most popular methodological extensions, the Latent Class Choice Model (LCCM) not only analyses heterogeneity due to observed socio-economic characteristics, but is also able to account for unobserved factors which could affect the preferences of diverse population segments. Incorporating such random heterogeneity is especially important if we assume that some behavioural latent factors underlie a specific choice behaviour (Walker and Li 2007; Hoshino 2011). We assume that such unobserved factors exist in the particular context of ethnic preferences, where heterogeneity is not only due to differences in individual-specific variables, but also due to unobserved factors, such as attitudes towards co-ethnics and other ethnic groups.

Relying on the methods developed for modelling residential location choice behaviour, this paper aims to explore ethnic preferences and their impact on residential location choices. In particular, it focuses on accounting for the observed and unobserved sources of heterogeneity that might affect the residential choice behaviour of different ethnic groups. For this purpose, we test the potential of the LCCM method applied on a dataset stemming from a Stated Preferences (SP) experiment of

[1] For an extensive overwiew of the methods see Train (2009).

neighbourhood choice. The choice survey is conducted in the Swiss city of Lugano in 2010 (Ibraimovic, 2013; Ibraimovic and Masiero 2014). In this particular paper, we aim to explore:

i) If latent segments holding diverse preferences for neighbourhood mix exist?

ii) Which household characteristics explain variations in such preferences?

iii) How do these effects translate into monetary valuations (*willingness-to-pay*) for different ethnic and non-ethnic neighbourhood characteristics?

Various socio-economic covariates affecting residential location choice decisions and, in particular, those potentially related to the tastes over ethnic neighbourhood characteristics, are tested as predictors of the class-membership model.

The remainder of the paper is structured as follows. A literature overview is presented in Section 2, followed by a description of the data and the context of the study in Section 3. Section 4 defines the methodological approach and model specification, while the model results are discussed in Section 5. Finally, Section 6 draws conclusions and proposes recommendations for further research.

2. Literature overview

Residential location choice (RLC) models have been extensively used to analyse the determinants of housing choice behaviour, i.e. the preferences over a wide set of location characteristics that drive these choices. Although many studies find origin and ethnicity to be among the significant location choice drivers (Schirmer *et al.* 2104), RLC models have only been used in a limited number of studies to directly examine segregation issues.

In fact, preferences for ethnic neighbourhood mix have been commonly addressed through different stated preferences methods, especially in the US context with studies such as Multi-City Study of Urban Inequality in the '90s. Such methods analyse ethnic preferences from stated choices among hypothetical neighbourhoods with different ethnic composition. The most prominent examples of applications include those by Farley *et al*. (1978, 1997, 2000) and Krysan (2002a, 2002b) in several US cities. The studies find the existence of ethnic preferences, where these differ across diverse ethnic groups. Many other studies follow, developing such methods in various directions. For example, in Charles (2006) and Krysan (2016) the respondents are asked to describe their ideal neighbourhood in terms of ethnic mix. A very interesting attempt to capture and isolate the social class effects that might be embedded in ethnic preferences, i.e. the way people perceive ethnicity as a proxy for social class, was proposed by Krysan *et al.* (2009). They use a video experiment in which ethnic composition and social class characteristics were manipulated to infer the ethnic preferences free from the social class effects. The findings from Detroit and Chicago show that ethnic composition had a significant impact on neighbourhood desirability, net of social class effects.

However, all these survey methods imply ethnic description of the neighbourhood, in terms of shares of different ethnic groups, as the only driver of residential choices. Whereas in the real world

settings, a compensatory behaviour is more likely to be employed by households to trade-off among different location characteristics, ranging from the house prices, school quality to commuting time and local amenities. It is possible that in some contexts of severe segregation, ethnic factors can overtake other choice drivers, while in other contexts they might be much less important as characteristics of consideration when choosing a particular residential location. In either case, it is essential to quantify the relative importance of ethnic versus other neighbourhood characteristics, so to understand whether the segregation patterns are driven by ethnic preferences or not. These are some of the key advantages that RLC and choice experimental methods offer to analyse ethnic preferences in relation to location choices on one hand and the resulting segregation patterns on the other.

Coming back to the RLC applications, Sermons (2000) was among the first to look specifically at the ethnic neighbourhood composition testing various questions over the ethnic preferences and segregation outcomes. Aslund (2005) estimated a choice model on revealed preferences (RP) data from Sweden, finding that both the presence of co-ethnics and the presence of other immigrants act as attractors for new and old immigrants. Ibraimovic and Masiero (2014) analyse ethnic preferences in the Swiss context, and measure the impact of these preferences on residential location choice behaviour among other location choice drivers. In the same context, Ibraimovic and Hess (2016) test and confirm Schelling's (1971) hypothesis relating to asymmetries in ethnic preferences, i.e. the *weak ethnic preferences* assumption. They find that "individuals tend to be averse to decreases in the share of their co-nationals, while being indifferent to any increases". However, such results vary across ethnic groups indicating important heterogeneity. Moreover, Zhang and Zheng (2015) estimate a RLC model on US data, with the aim of investigating the preferences and willingness to pay for increase/reduction of segregation at a city level. They find that both white and black households dislike segregation, whereas the willingness-to-pay for less segregation is bigger for blacks than for whites. All these studies confirm the existence of heterogeneity influencing ethnic preferences for different ethnic groups, which can often result independently from the observed socio-economic characteristics (Ibraimovic and Masiero 2014). The urge is thus on applying the methods able to represent observed and unobserved heterogeneity, measure its impact on residential location choice behaviour and provide valuable insights over the underlying factors affecting such behaviour.

Since the importance of recognizing taste differences in models of choice behaviour has been emphasised (McFadden and Train 2000), several methods for incorporating observed and unobserved heterogeneity have been proposed. Accommodating observed heterogeneity stemming from the differences in individual-specific characteristics through interaction terms or deterministic segmentation (Train 2009) are the first step undertaken in most applications. Beyond the observed sources of heterogeneity, the literature on choice behaviour indicates that an important part of individual-level taste variations results from unobserved factors (Bhat 2000). Similar conclusions have been drawn in the residential location choice context (Walker and Li 2007; Hoshino 2011).

According to such recommendations further model extensions allowing for random heterogeneity were developed, the Random Coefficients Logit Model (RCL) and the Latent Class Choice Model (LCCM). [2]

In a LCCM, the sample population is divided into a number of classes, each class having its own parameters vector, where each individual belongs to a class up to a probability, and where that probability is a function of the individual's characteristics. This model thus combines deterministic and random heterogeneity in preferences. One of the first papers to address the latent heterogeneity in discrete choice models through the segmentation of population was by Swait (1994), followed by Gopinath (1995), Ben-Akiva and Boccara (1995) and several other authors. In the context of residential location choice decisions, Walker and Li (2007) tested the impact of lifestyle preferences on residential choice behaviour. Considering the concept of *lifestyle* as a latent driver of residential location decisions, they investigated its impact on the tastes for different location characteristics. In another application, Ettema (2010) studied the heterogeneity in preferences for commute distance across commuters and telecommuters. The findings suggested that although telecommuting does not significantly impact the relocation choice in the simple MNL model, the distinction between different segments in LCCM can add to the explanation of the impact of telecommuting on residential location choice patterns.

The ability to link heterogeneity in preferences to observable individual-specific characteristics is one of the characteristics of LCCM that could greatly benefit the analysis of ethnic segregation patterns. Several socio-economic and demographic variables have been indicated as potential determinants of the differences in tastes for ethnic neighbourhood mix. However, to the best of our knowledge, there are no existing studies employing this framework in the ethnic segregation domain. Our goal is therefore to exploit the potential of the LCCM for exploring the sources of heterogeneity affecting the residential location choice behaviour of different population subgroups.

3. Data

The main dataset used for the analysis is based on the Stated Preferences experiment of neighbourhood choice conducted in the Swiss city of Lugano in 2010 (Ibraimovic 2013). The survey involved a stratified sample of the city population, which in 2008 comprised 78,025 inhabitants. In particular, all residents (over 18 years of age) were stratified according to their origins[3] (in 10 different nationality groups shown in Table 1) and their neighbourhood of residence. They were then randomly sampled and the female or male household head was interviewed. In order to allow households from diverse national backgrounds to take part in the experiment, some less represented nationalities were oversampled.[4] A final sample of 133 households of different

[2] For advantages of LCCM over RCL see Hensher and Greene (2003) and Hess et al. (2009).
[3] Origin is defined on the basis of citizenship. We define the foreign population as all the residents with a foreign citizenship (i.e. non Swiss citizens).
[4] The sampling procedure did not affect the model results since the sampling criteria was based on exogeneous individual-specific variables and not the choice variable (Manski and Lerman, 1977; Manski and McFadden, 1981).

origins and socio-economic status participated in a SP choice experiment and a household survey which collected information about their present residential location and their socio-economic characteristics.

3.1. Ethnic communities and residential concentration patterns in Lugano

Lugano has a large foreign population (nearly 40%) from over 120 countries world-wide. Even though the immigrant population is very diverse, just a few countries of origin make up a large share of the total (Table 1). Due to its geographical position, the Italian community dominates with 53% of the total foreign population (corresponding to 21.22% of the overall city population). The second most represented nationality group are citizens from former Yugoslavia (nearly 17%), followed by Portuguese (6%), Germans and Turks (each make up less than 3%). For the purpose of this study, the remaining nationalities were clustered in four groups according to geographical and linguistic similarities, namely *rest of EU, USA and Australia* (c.a. 7 % of foreigners), *East Europe and Asia* (c.a. 6% of foreigners), *South America* (3,5% of foreigners) and *Africa and Middle East* (c.a. 3% of foreigners).

Similarly to other European countries, the foreign population in Switzerland is more represented in urban centres. Foreigners are also often clustered in certain neighbourhoods where they sometimes outnumber the native population. In Lugano, the share of foreigners in different neighbourhoods varies from 12% to 30% in suburban neighbourhoods, and reaches up to 57% in central neighbourhoods.

Even though there are no cases of segregation of single nationality groups, there is a tendency of different nationalities to group in specific neighbourhoods. Among the most clustered nationalities are the citizens of South American countries, Turkey and former Yugoslavia, but also some western European and American communities. From the patterns of clustering of single nationality groups, a certain socio-spatial hierarchy emerges. On the one hand, the citizens from wealthier EU and other Western countries predominantly inhabit more attractive city neighbourhoods together with the natives; on the other hand, immigrants from other countries are mostly located in large residential districts with higher than average foreign population. Such distinct patterns of residential location among “advantaged” and “disadvantaged” foreign communities are also present in other major Swiss cities (Arend 1991). Following this categorization of the foreign population in the Swiss context, we aim to explore the heterogeneity in preferences for these two groups of foregners, along with the native population. Accordingly, we define as the “advantaged foreigners” category citizens from the Western EU countries, US, Canada and Australia, whereas the “disadvantaged foreigners” category comprehends nationals of Former Yugoslavia, Turkey, Portugal, Eastern Europe and Asia, South America, Africa and Middle East.

[Table 1 near here]

3.2. The Stated Preference experiment of neighbourhood choice

As already mentioned in the introduction, the use of data on real world residential location choices would not allow us to reliably tease out preferences for specific neighbourhood characteristics, as

real world decisions are affected by constraints other than pure preferences. Moreover, as explained in Sermons (2000) other biases from the use of RP cross-sectional data might arise. Firstly, biases could arise due to the assumption that all households evaluate the same set of location characteristic; secondly due to the potential substantial change in location characteristics including ethnic neighbourhood mix from the time when each of the households has made the present location decision; and thirdly due to the relocation costs that might prevent a household from moving even if the present location is not the one that maximizes its utility. For these reason, unlike the other RLC applications using RP data to study ethnic preferences (e.g. Sermons, 2000; Aslund, 2005), we relied on a stated preference (SP) survey, and in particular a stated choice (SC) setup, which obviates such biases.

SP surveys are a widely used method for identifying preferences in cases where revealed preferences (RP) data are unavailable or inadequate to indentify preferences (Louviere *et al.* 2000). They are used as the key data source for numerous studies that guide policy making across fields as diverse as transport, environmental protection, health and resource and energy economics. In the case of residential markets, we cannot use revealed preferences due to the choice-constraint issue (see. Van der Laan Bouma Doff, 2007), which is related to the socioeconomic position of ethnic minorities and thus price, discrimination and other causes of accessibility constraints which do not allow a free choice among neighbourhoods especially for disadvantaged ethnic communities. These correlations would prevent us from producing meaningful estimates of the relative importance of different neighbourhood characteristics.

SP methods have developed extensively since the early days when studies relied primarily on contingent valuation and other more direct preference elicitation methods (e.g. transfer price questions). When applied carefully, SP methods can produce results that are in line with real world preferences. A key component in this context is to make the surveys realistic for the respondent, allowing them to relate to the choice they are faced with. In our specific context, we do this by pivoting (framing) the SP experiment around respondents' real neighbourhood of residence, permitting us to adapt the hypothetical choice settings to respondents' housing situation as well as the urban context, increasing realism and reducing the risk of biased responses.

In order to reveal preferences for ethnic characteristics of the neighbourhood, the SP experiment presents respondents with multiple choice scenarios, each time involving a hypothetical situation where they are asked to choose one among three alternative neighbourhoods. The first neighbourhood option being the respondent's actual area of residence, where the attribute values correspond to real observed values. Given the pivoted experimental settings, the present neighbourhood of residence constituted the reference alternative in the choice experiment design. The second and the third neighbourhoods were represented by two hypothetical alternatives with attribute levels pivoted around the values of the reference neighbourhood. This setting permitted respondents to recognize a familiar choice situation, thus making the choice experiment more realistic and reliable.

These three neighbourhoods were described by a number of different characteristics as follows.

- The experiment was primarily aimed at analyzing the preferences for ethnic neighbourhood characteristics, particularly exploring the self-segregation propensity among the own community of origin, represented by the *concentration of co-nationals* as the first attribute describing the three alternative neighbourhoods. This was defined as the number of co-national inhabitants in the neighbourhood over the total number of co-nationals in the city.
- Secondly, the sensitivity to different levels of foreign population in the neighbourhood was represented by the *share of foreigners* attribute. This was defined as the number of non-Swiss residents over the total number of residents in the neighbourhood
- Finally, the experiment included variables indicating the *travel time to work* and the *monthly dwelling rent*, which are, according to the literature, among the main drivers of residential location choice decisions. The actual value of the monthly rent and travel time to work by habitually used transport mode were asked to each respondent in a set of pre-survey questions, thus they were input into the choice experiment, creating the values for alternative neighbourhoods based on the attribute levels and the experimental design.

Table 2 presents some summary details of the experiment. Each neighbourhood attribute contained the reference value level (the observed value in the residential location of respondents) and four additional levels expressed as positive and negative percentage deviations from the reference value. The range of percentage deviations was established according to the urban context under exam.

[Table 2 near here]

Based on fractional factorial orthogonal design, the experiment contained 25 choice tasks divided in two blocks of 12 or 13 choice tasks, as presented in the Figure 1. The values of attributes describing the alternative neighbourhoods varied across every choice task. Each respondent was assigned one of the two blocks, thus responding to 12 or 13 choice tasks. The SP choice experiment[5] was conducted through face-to-face computer assisted interviews where the respondents were asked to select their favourite neighbourhood of residence among three alternative options across the different choice tasks. The question asked to respondent in the choice experiment was the following: "We present you the characteristics of your present neighbourhood and those of two other neighbourhoods in the city of Lugano, in 10 years time. Imagine that you can choose to live in the dwelling same as your, situated in one of these neighbourhoods. In which of these neighbourhoods would you want to live?" They were instructed to make their choices only on the basis of the attributes used to describe the neighbourhoods and asked to imagine that other attributes, such as for example the characteristics of their dwelling, would remain unchanged in the other neighbourhoods. Moreover, the respondents were told to assume no relocation costs. A total of 1,626 valid observations, collected from 133 households, were used for the empirical analysis.

[Figure 1 near here]

[5] For an extensive review on SP choice experiments and experimental design see Louviere et al. (2000) and Hensher et al. (2005).

To recap, the advantage of this hypothetical setting is the ability to produce "net" relative valuations of the different characteristics of neighbourhoods, free of the influence of involuntary causes. The neighbourhoods are described by different ethnic and non-ethnic characteristics, such that the resulting choice relies on trade-offs among these characteristics, revealing respondents' preferences for the factors driving their choices.

3.3. Additional data

For analysing the heterogeneity in preferences, a set of socio-economic and demographic variables were collected in a previously conducted household survey on the same set of respondents. The data included information on the origin, income, education, religion, national language proficiency, language used in free time, years lived in Switzerland, occupational status, age, and other characteristics. The descriptive statistics of socio-economic variables of the SP survey sample are presented in Table 3.

[Table 3 near here]

As discussed in the Section 3.1, we base our analysis on three distinct segments of population, namely advantaged foreigners, disadvantaged foreigners and native population. Advantaged foreigners tend to exhibit similar characteristics and status as the native population. They have lived in the country for 31 years on average and are slightly younger that the Swiss. Disadvantaged foreigners are on the other hand the most recent immigrants in the country with 18 years of average residence, and are considerably younger than the other two groups (nearly 38 years on average). This is mainly due to their immigration motivation, namely economic migration or asylum seekers and refugees. The level of income between the three groups on average shows slightly lower values for disadvantaged ethnic minorities, despite their education level being higher than that of the natives. Advantaged foreigners have the highest education and income level, while their Italian language proficiency is higher than that of the disadvantaged foreign communities. This could be a result of a considerably longer residency in Switzerland compared to disadvantaged foreigners. On average, the Swiss are the oldest group which also shows the lowest residential mobility. Together with advantaged foreigners, the Swiss tend to live longer in the same neighbourhood and dwelling (more than 14 and 13 years respectively). Disadvantaged foreigners show higher mobility levels, residing on average just over 10 years in the same neighbourhood, and 8 years in the same dwelling.

4. Modelling framework

4.1. Choice modelling framework

As already mentioned in the introduction, we conduct our analysis using mathematical structures belonging to the family of random utility models, and specifically discrete choice models. For a detailed overview of these models, see Train (2009). Discrete choice models are used to analyse individual choices made by decision makers, in this case the choice between different neighbourhoods. They explain that choice by the notion that each of the neighbourhoods has an associated utility for the decision maker, and that a rational decision maker chooses the

neighbourhood which provides him/her with the greatest utility. The utility is a function of the characteristics of the neighbourhood and the sensitivities of the decision maker, which vary in the population. The outputs from the modelling work are coefficients (say β_k for characteristic k) which represent the marginal utilities, showing the change in utility resulting from a change in the associated attribute (x_k). The utility that a given decision maker n obtains from choosing neighbourhood i is then given by $U_{i,n} = \sum_k \beta_k x_k + \varepsilon_{i,n}$ where $\varepsilon_{i,n}$ is an error term that accounts for the limited knowledge of the analyst, and which is assumed to follow an extreme value distribution in logit models. In a simple MNL model, the probability of decision make n choosing a given neighbourhood i (out of $j=1,...,I$) is given by $P_{i,n} = \frac{e^{V_{i,n}}}{\sum_{j=1}^{I} e^{V_{j,n}}}$, where $V_{i,n} = \sum_k \beta_k x_k$, i.e. the probabilities are independent of the error term.

A key characteristic of the models is that behaviour is compensatory, in that poor 'performance' on one characteristic can be compensated by good performance on another (so for example lower rent can compensate for a longer commute). By comparing the individual coefficients with each other, an analyst can obtain insights into the relative importance of different neighbourhood characteristics in driving preferences. By comparing the sensitivities of non-cost characteristics of the neighbourhood to the sensitivity to cost, an analyst can compute willingness to pay (WTP) measures which show how much a decision maker is willing to pay extra in rent (in our case) to improve a given characteristic of his/her neighbourhood. For example, a WTP for increasing the share of co-nationals would be given by $\frac{\beta_{co-nationals}}{\beta_{rent}}$, showing how much extra respondents would be willing to pay in rent per month to increase the share of co-nationals by one unit (i.e. one percentage point).

4.2.Treatment of heterogeneity

In a simple MNL model such as discussed in Section 4.1, the treatment of heterogeneity across individual decision makers is limited to interactions with respondent characteristics, such as income and age, for example. In particular, instead of assuming a constant sensitivity to rent, i.e. $\beta_{rent} x_{i,rent}$ for neighbourhood i, we might use $\beta_{rent} \left(\frac{inc_n}{av.inc}\right)^{\lambda} x_{i,rent}$, where, with inc_n being the income of respondent n and *av.inc* being the average income in the sample, λ is now an estimate income elasticity. If λ is 0, there is no income effect, while, if $\lambda = -1$, we have that the rent sensitivity halves when income doubles.

While socio-demographic interactions such as an income effect can help explain some of the differences in preferences across individuals, a large share of heterogeneity will remain unexplained, due to intrinsic differences in preferences across decision makers as well as a failure in many datasets to capture all relevant respondent characteristics that could explains differences.

In our paper, we make use of a Latent Class Choice Model (LCCM) to capture such differences. The main assumption of LCCM is that the population can be classified in S latent classes, where the decision-makers belonging to the same class exhibit homogeneous preferences, while these tastes are allowed to vary across different classes revealing different sensitivities to alternatives' attributes

for each population segment. Decision-makers are not deterministically segmented into defined classes, but are assumed to belong to each latent class with a certain membership probability. Each respondent has a non-zero probability of falling into every class, thus allowing for random heterogeneity in preferences. The way in which a respondent belongs to a given class is not imposed by the analyst, but is estimated alongside all other model parameters. A membership probability is defined for each person and for each class, and the class membership model probabilistically assigns individuals to different classes according to their individual-specific characteristics.[6]

In a given LCCM model, we will have *S* different latent classes, where the optimal number for *S* is determined by the analyst using an iterative model fitting exercise. We would first try a simple MNL model, where *S=1*, before moving to a LCCM model with *S=2*, and then increasing *S* until no significant further gains in model fit or behavioural insights are obtained. For this purpose, some goodness of fit criteria such as the Bayesian Information Criterion (BIC) or the Akaike Information Criterion (AIC) are widely used. However, as discussed in Scarpa and Thiene [2005], it is important for researchers to also opt for selecting the model that offers the most meaningful and interpretable results, an approach that is used in our analysis.

In a model with *S* classes, we will have *S* different class membership probabilities for each respondent *n*, given by $\pi_{s,n}$ for class *s* and respondent *n*. We would have that $\sum_{s=1}^{S} \pi_{s,n} = 1, \forall n$. Since the individual-specific characteristics of decision-makers are used as determinants of the class membership probabilities, it is possible to describe and characterize different classes adding insight into the interpretation of model results. For the class membership model component we define a MNL model where the probability of individual *n* belonging to class *s*, is given as function of a set of characteristics of person *n*, say z_n, with $\pi_{s,n} = \frac{e^{\sum_l \gamma_{s,l} z_{n,l}}}{\sum_{c=1}^{S} e^{\sum_l \gamma_{c,l} z_{n,l}}}$ where the different $\gamma_{s,l}$ are parameters to be estimated, with an appropriate normalisation, typically setting $\gamma_{s,l}$ to 0 for one class *s*, and for all l. In the earlier example with two classes and an impact of income, we might for example set $\gamma_{1,inc} = 0$, i.e. using the impact of income in class 1 as the base, and would then expect a negative estimate for $\gamma_{2,inc}$ given our expectation that low income people fall into class 2.
Within each class *s*, the choice probabilities would again be given by a MNL model, with $P_{s,i,n} = \frac{e^{V_{s,i,n}}}{\sum_{j=1}^{I} e^{V_{s,j,n}}}$, where $V_{s,i,n} = \sum_k \beta_{s,k} x_k$, i.e. having separate β parameters in each class.

The class allocation and class specific choice models are combined and estimated simultaneously forming a LCCM, where the probability of individual *n* selecting alternative *i* is given by the sum over classes *s* of the product between the probabilities of the class-specific choice model and the

[6] For example, a model might reveal two classes, one of which captures more cost sensitive respondents than the other, where the cost sensitive people maybe care less about travel time. Each respondent would have a non-zero probability of belonging to either class, but higher income people would be expected to have a higher probability of falling into the low cost sensitivity class.

class membership model, i.e. $P_{i,n} = \sum_{s=1}^{S} \pi_{s,n} P_{s,i,n}$, with $\pi_{s,n}$ and $P_{s,i,n}$ defined as above. The model was coded and estimated in OX (Doornik 2001).

4.3. Hypotheses

The hypothetical neighbourhoods are described by four characteristics, and we have a number of a priori expectations about the impact of these on preferences.

- *Co-nationals concentration* is employed to study households' self-segregation preferences, i.e. tendencies of ethnic clustering in specific neighbourhoods. The vast literature on residential segregation shows that the presence of co-ethnic neighbours is one of the major determinants of residential location choices of immigrants as well as of the native population. While we expect this attribute to contribute positively to the probability of selecting a particular neighbourhood of residence, some taste heterogeneity is to be expected across households from different origins and socio-economic status.
- *Foreigners' share* tests the hypothesis that a higher share of foreign population is related to an unfavourable perception of the location environment in terms of quality and safety, thus impacting negatively on choice probabilities. According to this hypothesis, households would be willing to pay a premium in order to live in neighbourhoods with a larger share of natives and a smaller share of foreigners. However, this attribute is likely to exhibit heterogeneity in sensitivities across population clusters, with the main expected determinants being the country of origin, the income and the education level of respondents. In fact, past empirical evidence suggests that natives are the ones to hold the strongest preferences for neighbourhoods with a predominant native population, while weaker preferences are found to exist also for ethnic minorities.
- *Travel time to work* denotes the commuting time by the habitually used mode of transport, expressed in minutes. We expect a negative effect of longer commuting time on the utilities of alternative locations.
- *Monthly dwelling rent* in this study represents the cost variable based on which the WTPs are computed for the residential location attributes. Income is supposed to be the major driver of variation in sensitivity towards the dwelling price.

5. Model results

We report and compare results for two residential location choice models, the base multinomial logit model (MNL, referred as model M1) and the Latent Class Choice Model (LCCM, referred as M2). Both models involve the estimation of a choice between three alternative neighbourhoods (present neighbourhood of residence and two hypothetical neighbourhoods, A and B), each described by four attributes as defined by the choice experiment. With no sample segmentation, model M1 represents homogeneous preferences for residential location alternatives and attributes, albeit that we incorporate an income elasticity on cost, as described in Section 4. The model results are presented in Table 4.

All parameters of the base MNL model (M1) have the expected sign and are statistically significantly, i.e. the associated attributes have a demonstrable impact on the choice. The concentration of co-national neighbours positively affects the utility of (and hence probability of choosing) a specific neighbourhood, while the share of foreigners has the opposite effect, exercising a negative impact on utilities. This indicates that households prefer residential environments with a larger presence of their co-nationals, yet a lower presence of (other) foreign communities. Travel time to work and the monthly dwelling rent both exhibit negative coefficient estimates which show disutility associated with such attributes. The positive and significant alternative specific constant for the reference alternative (ASC neighb. ref) indicates the preference of households for staying in the present neighbourhood of residence. Since no relocation costs are assumed in the experiment, such result shows the existence of positive utility effects related to the current residential location (such as social ties, habits or attachment to the territory). The only socio-demographic characteristic which was included in the MNL model following preliminary testing was a continuous income elasticity on rent. This was incorporated as $\beta_{rent}\left(\frac{inc_n}{av.inc}\right)^{\lambda} x_{i,rent}$, and the value of -0.9455 is very close to -1, showing that we see clear evidence of decreasing sensitivity to rent with increasing income, with an almost 1% decrease in sensitivity for a 1% increase in income.

[Table 4 near here]

After estimation of numerous intermediate models, we settled on a LCCM model with three latent classes, i.e. $S = 3$. Each respondent is probabilistically assigned to each latent class, where the probability of a given respondent falling into a given class is a function of his/her characteristics. The allocation is estimated, rather than imposed by the analyst, i.e. it does not relate to some a priori social class structure. Since the LCCM is composed of the choice and class-membership model components, we obtain two sets of estimates related to each of these components. Thus, we define three latent classes based on their specific tastes for neighbourhood attributes (see the first set of seven coefficient estimates relative to the choice model component). Following initial tests, the income elasticity was kept generic across classes allowing for the same sensitivity to rent for all respondents.

Comparing M1 and M2, the significant gains in the model fit obtained for the LCCM model (M2) over the base MNL model (M1) indicate an improvement in explanatory power when accounting for observed and unobserved sources of heterogeneity across population classes. The existence of diverse segments is also supported by dissimilar parameter estimates across classes, suggesting that different individuals exhibit substantial differences in tastes for all attributes studied in our model.

Each respondent in the data has a non-zero probability of belonging to each of the three classes, where the probability varies across respondents as a function of socio-demographic characteristics (see the second set of three coefficient estimates relative to the class-membership model component). In a preliminary analysis, a wide range of individual-specific variables were tested as determinants of the class-membership probabilities. Among these, the model highlights the origins

of the households (distinguishing between natives, advantaged and disadvantaged foreigners) as the strongest predictors of probabilistic class allocation. Thus, these covariates are included in the final model specification as the main drivers of heterogeneity in residential choice behaviour. This means that households of different origins have different tastes for ethnic and other location attributes which then define their choices of diverse neighbourhoods.

We see that being Swiss increases the probability of falling into class 2 and less so class 3 (the values for $\gamma_{2,swiss}$ and $\gamma_{3,swiss}$ are positive), but the estimates for these covariate effects are not statistically significant. Disadvantaged foreigners are less likely to fall into class 2 than class 1, and more likely to fall into class 3, but again, the effects are not very significant. On the other hand, for advantaged foreigners, we see a substantially higher probability of falling into class 3 ahead of class 1 and class 2 (with $\gamma_{1,advantaged}$ normalised to zero, the values for $\gamma_{2,advantaged}$ and $\gamma_{3,advantaged}$ are negative and positive, respectively). The nationality shares in the data along with these γ estimates also allow us to work out sample level class allocation probabilities (using a logit style class allocation model), which are reported in Table 4, showing that class 3 is the largest.

First insights into the heterogeneity recovered by the LCCM model can be observed by comparing the results across classes. We see that:

- Increases in the concentration of co-nationals have a positive and significant impact of choosing a given neighbourhood only in the first and third classes.
- Increases in the concentration of foreigners have a negative and significant impact of choosing a given neighbourhood only in the second class.
- Increases in travel time have a significant negative impact only in the second and third class of the LCCM model, while cost has negative impact in all three LCCM classes.
- The income elasticity recovered by the LCCM model is much weaker (and less significant) than in the MNL model, potentially suggesting confounding with unobserved random heterogeneity in the MNL model.

We next use the results to compute willingness-to-pay (WTP) measures shown in Table 5, which contrast changes in neighbourhood composition and travel time with change in rent, thus giving an indication of the monetary value of such changes. For the MNL model, point values are obtained, showing a positive WTP for increases in the share of co-nationals, while reductions in rent are required to accept increases in the share of foreigners or in travel time (negative WTP). For the LCCM model, there are substantial differences in WTP across the three classes. For example, we note that the WTP for an increase in co-nationals is almost six times as large in the first class as in the third class, where other than cost, it is the only driver of choice in the first class. While the weighted averages in WTP across the three classes are similar for the share of co-nationals and travel time between LCCM and MNL, the weighted WTP for reducing the share of foreigners in the LCCM is more than twice the value from the MNL model. This highlights the importance of allowing for random heterogeneity.

[Table 5 near here]

While the results in Table 5 give some overall indication of behaviour, it is far more insightful to link these with the socio-demographic characteristics. The class allocation parameters explain the role of socio-demographic characteristics in driving the probabilities of belonging to the different classes. Together with the observed choices in the data for different types of neighbourhoods by different individuals, it is then possible to work out posterior class characteristics (cf. Hess 2014). These would give the most likely class for each specific respondent, conditional on their socio-demographics and the choices they have made. We can then use this information to describe the specific composition of each of the three latent classes, as show in Table 6. Disadvantaged foreigners are the largest group in the data, and also the largest group within each of the latent classes. However, their importance is smaller in the third latent class, where they are likely to account for just 54.99%[7] of that class. A difference also arises between Swiss citizens and advantaged foreigners. The former have a much larger presence than the latter in the second latent class, with the opposite applying in the third.

[Table 6 near here]

We next use these new weights to calculate WTP measures corresponding to those from Table 5, but doing this for the three different socio-demographic groups as a function of their specific posterior class allocation probabilities (Table 7).

[Table 7 near here]

Overall, the three population groups value the presence of co-nationals, disliking a higher foreigners' concentration in their neighbourhood of residence. This could indicate a hierarchy in ethnic preferences where the co-nationals are the most desired ethnic group, while other foreigners are the least desired group. Nevertheless, the LCCM highlights some taste differences among the three ethnic groups which are shown by their WTP for different neighbourhood characteristics. Swiss citizens are mainly sensitive to the concentration of foreigners and the travel time to work, exhibiting aversion to increases in these attributes. They indeed require a compensation of CHF 65 (corresponding to ca. 5% lower average monthly rent) for a 10% increase of the share of foreigners in the neighbourhood, and CHF 121 (ca. 9% of average monthly rent) for 10 minutes longer commuting time. To offer some external validation of these results, with an average of 40 commute journeys per month, this would equate to a value of travel time of CHF 18.12 per hour (CHF 12.08 x 60 minutes per hour / 40 journeys per month). This value compares well with the values obtained in official value travel time studies in Switzerland (e.g. around CHF 19 per hour in Axhausen *et al.* [2008]).

[7] 54.99% of 48.86% (sample level class allocation probability for class 3, see Table 5) gives 26.87% (posterior class allocation for disadvantaged foreigners and class 3, see Table 6).

Disadvantaged foreigners associate the highest value among the three groups to presence of co-nationals in the neighbourhood and are willing to pay a premium of CHF 42.5 (ca. 3% of average monthly rent) for a 10% increase in the concentration of their community of origin. They also dislike the presence of other foreigners and are sensitive to commuting time, to a lesser than Swiss citizens but more than advantaged foreigners. Advantaged foreigners also value neighbours of their same nationality (CHF 39 for a 10% increase), however they place a very low value on the share of foreigners, and are also less sensitive to travel time to work (CHF 60 for 10 minutes longer commuting time). Such results indicate differences in residential location choice drivers for different population segments, which account for heterogeneity in tastes for ethnic and non-ethnic neighbourhood characteristics.

6. Conclusions

Ethnic residential segregation can result from the choices of ethnic communities to group in same urban areas, or to the constraints they might face in accessing desired residential locations. In this study we aim to analyse the voluntary self-segregation, guided by preferences for co-ethnic neighbours. In particular, we focus on the heterogeneity in preferences among diverse ethnic groups, either driven by differences in their socio-economic characteristics (observed heterogeneity), or due to various attitudinal and behavioural elements often difficult to observe for the researcher (unobserved heterogeneity). For this purpose we test the potential of Latent Class Choice Model (LCCM) to account for observed and unobserved heterogeneity components in the analysis of ethnic segregation preferences. For the empirical analysis we use a dataset collected from a specifically designed Stated Preference experiment of neighbourhood choice, conducted in the Swiss city of Lugano in 2010 (Ibraimovic, 2013). This experiment permits us to uncover the impact of preferences for ethnic neighbourhood composition, free from the constraints component usually existing in the real housing markets (such as access barriers, shortage of accommodation options or discrimination effects), by implying a hypothetically free choice among alternative neighbourhoods.

The results support the hypothesis of existence of three latent classes which differ in their housing choice behaviour and tastes for the ethnic residential environment. In particular, different ethnic attributes are considered as key choice drivers by households belonging to different latent classes. Among various socio-economic characteristics, the origin of households is the best predictor of the class membership, especially when differentiating between disadvantaged ethnic communities, advantaged foreigners and the native population. We analyse the differences in ethnic preferences and residential choice across different ethnic groups by translating the parameter estimates of the LCCM into the willingness-to-pay (WTP) measures for different residential location attributes.

The results indicate that disadvantaged foreigners value the presence of their co-nationals in the neighbourhood, however disliking other foreign communities. This could indicate a hierarchy in their preferences which favours their community of origin, but also gives the priority to native

population over foreigners from other communities. The presence of underlying factors such as the impact of their social networks could be at the basis of such preference hierarchy. On the one hand, disadvantaged ethnic communities could prefer their own co-nationals due to the existence of ethnic networks; on the other hand they could privilege exposure to natives because of the potential benefits from integration within the mainstream society and higher chances of social mobility. Swiss respondents are mainly concerned with the presence of foreigners and travel time to work, valuing negatively both of these location characteristics. This suggests the propensity of natives to grouping behaviour, due to an aversion to a high presence of foreign inhabitants rather than self-segregation preferences. Advantaged foreigners value their community of origin and dislike longer commuting time. Unlike the other two groups, they show a very small sensitivity to the presence of other foreign groups, which does not play a big role in guiding their residential location choices.

This study represents a step forward in the attempt to account for and better explain different observed and unobserved elements affecting residential choice behaviour of diverse ethnic and socio-economic groups. A greater understanding of such differences represents a valuable indication for policy measures to counter the onset and further development of ethnic segregation. Policy measures tailored to account for preferences of identified population segments are more likely to be effective and cause shifts in residential location choice behaviour towards greater residential integration. This type of analysis could be further developed to include other peculiarities in residential choice behaviour of a heterogeneous population. We here propose two directions for future studies. Firstly, the inclusion of latent variables capturing and explaining attitudes and perceptions could be used in the class membership model for the class definition. Secondly, widening the sample size to represent and study the preferences of single nationality groups would allow us to examine whether they follow the mainstream theories of segregation dynamics or exhibit specificities in their residential behaviour.

Tables and Figures

Caption: The figure illustrates an example of the stated preferences choice task presented to respondents in face-to-face computer aided questionnaire.

References

Andersen, H.S. 2015. “Explanations for Special Neighbourhood Preferences among Ethnic Minorities.” *Housing, Theory and Society* 32 (2): 196–217.

Arend, M. 1991. “Housing segregation in Switzerland.” In: *Urban Housing Segregation of Minorities in Western Europe and the United States*, edited by E. D. Huttman, W. Blauw and J. Saltman, 155–167. Durham, NC: Duke University Press.

Aslund, O. 2005. “Now and forever? Initial and subsequent location choices of immigrants.” *Regional Science and Urban Economics* 35 (2): 141–165.

Axhausen, K.W., Hess, S., König, A., Abay, G., Bates, J.J. and Bierlaire, M. 2008. “Income and distance elasticities of values of travel time savings: New Swiss results.” *Transport Policy* 15(3): 173-185.

Ben-Akiva, M., and Boccara, B. 1995. “Discrete Choice Models with Latent Choice Sets.” *International Journal of Research in Marketing* 12: 9-24.

Ben-Akiva, M., and Lerman, S. 1985. *Discrete Choice Analysis*. MIT Press.

Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., et al. 2002. “Hybrid choice models: progress and challenges.” *Marketing Letters* 13 (3): 163-175.

Bhat, C.R. 2000. “Incorporating observed and unobserved heterogeneity in urban work travel mode choice modeling.” *Transportation Science* 34: 228-239.

Bolt, G., and Van Kempen, R. 2003. “Escaping poverty neighbourhoods in the Netherlands”. *Housing, Theory and Society* 20: 209–222.

Charles, C. Z. 2000. “Neighborhood Racial Composition Preferences: Evidence from a Multiethnic Metropolis”. *Social Problems* 47: 379–407.

Charles, C. Z. 2003. “The dynamics of racial residential segregation”. *Annual Review of Sociology* 29: 167–207.

Charles, C. Z. 2006. *Won’t You Be My Neighbor: Race, Class, and Residence in Los Angeles*. New York : Russell Sage Foundation.

Clark, W. A. V. 1992. “Residential preferences and residential choices in a multiethnic context”. *Demography* 29 (3): 451–466.

Clark, W. A. V. 2009. “Changing residential preferences across income, education, and age: findings from the multi-city study of urban inequality”. *Urban Affairs Review* 44: 334–355.

Darden, J. T. 1986. “The significance of race and class in residential segregation”. *Journal of Urban Affairs* 8(1): 49–56.

Doff, W., and Kleinhans, R. 2011. “Residential outcomes of forced relocation: lifting a corner of the veil on neighbourhood selection”. *Urban Studies* 48 (4): 661–680.

Doornik, J.A. 2001. *Ox: An Object-Oriented Matrix Language*. Timberlake Consultants Press, London.

Ettema, D. 2010. “The impact of telecommuting on residential relocation and residential preferences.” *Journal of Transport and Land Use* 3 (1): 7-24.

Farley, R., Danziger, S. H., Holzer, H. J. 2000. *Detroit Divided*. New York: Russell Sage.

Farley, R., Fielding, E.L., Krysan, M. 1997. “The residential preferences of blacks and whites: A four-metropolis analysis.” *Housing Policy Debate* 8 (4): 763-800.

Farley, R., Schuman, H., Bianchi, S., Colasanto, D., and Hatchett, S. 1978. “Chocolate city, vanilla suburbs:” Will the trend toward racially separate communities continue?” *Social Science Research* 7 (4): 319-344.

Gopinath, D.A. 1995. “Modeling Heterogeneity in Discrete Choice Processes: Application to Travel Demand.” PhD diss., Massachusetts Institute of Technology, Cambridge, MA.

Hensher, D., and Greene, W. 2003. “A Latent Class Model for Discrete Choice Analysis: Contrasts with Mixed Logit.” *Transportation Research, Part B* 37: 681–98.

Hensher, D.A., Rose, J.M., and Greene, W.H. 2005. *Applied choice analysis: A primer*. Cambridge: Cambridge University Press.

Hess, S. 2014. *Latent class structures: taste heterogeneity and beyond*, in Hess, S. & Daly, A.J. (Eds.), Handbook of Choice Modelling, chapter 14, pp. 311-330, Edward Elgar publishers, Cheltenham.

Hess, S., Ben-Akiva, M., Gopinath, D., and Walker, J. 2009. “Advantages of latent class models over continuous mixed logit.” Paper presented at the 12th International Conference on Travel Behaviour Research, Jaipur, December.

Hoshino, T. 2011. “Estimation and Analysis of Preference Heterogeneity in Residential Choice Behaviour.” *Urban Studies* 48 (2): 363-382.

Ibraimovic, T. 2013. “Investigating the role of ethnic preferences in residential location decisions: Choice analysis on Stated Preferences data.” PhD diss., University of Lugano, Switzerland. doc.rero.ch/record/32276

Ibraimovic, T., and Masiero, L. 2014. “Do Birds of a Feather Flock Together? The Impact of Ethnic Segregation Preferences on Neighbourhood Choice.” *Urban Studies* 51 (4): 693-711. doi: 10.1177/0042098013493026

Ibraimovic, T., and Hess, S. 2016. “Households’ response to changes in ethnic composition of neighbourhoods: Exploring reference-dependence and asymmetric preference structures.” *Papers in Regional Science* (forthcoming).

Krysan, M. 2002a. “Whites Who Say They'd Flee: Who Are They and Why Would They Leave?” *Demography* 39 (4):675–696.

Krysan, M. 2002b. “Community Undesirability in Black and White: Examining Racial Residential Preferences through Community Perceptions.” *Social Problems* 49 (4):521–543.

Krysan, M., Carter, C., Van Londen, M. 2016. “The diversity of integration in a multiethnic metropolis: Exploring What Whites, African Americans, and Latinos Imagine.” *Du Bois Review: Social Science Research on Race*, 1-37.

Krysan, M., Couper, M. P., Farley, R., Forman, T.A. 2009. “Does Race Matter in Neighborhood Preferences? Results from a Video Experiment.” *American Journal of Sociology* 115 (2): 527-559.

Lin, A. C., and Harris, D. R. 2008. *The colors of poverty: Why racial and ethnic disparities persist*. New York: Russell Sage Foundation.

Louviere, J.J., Hensher, D.A., and Swait, J.D. 2000. *Stated Choice Methods: Analysis and Applications*. first ed. Cambridge University Press.

Manski, C.F., and Lerman, S.R. 1977. “The estimation of choice probabilities from choice-based samples.” *Econometrica* 45: 1977-1988.

Manski, C.F., and McFadden, D. 1981. *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge: The MIT Press.

Massey, D. S., and Denton, N. A. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, MA: Harvard University Press.

McFadden, D. 1974. “Conditional logit analysis of qualitative choice behavior.” In *Frontiers in Econometrics, edited by* Zarembka, P., Academic Press, New York.

McFadden, D. 1977. “Modelling the Choice of Residential Location.” *Cowles Foundation Discussion Papers 477*, Cowles Foundation, Yale University.

McFadden, D. 1984. “Econometric analysis of qualitative response models.” Chap. 24 in *Handbook of Econometrics*, *Volume II*, 1395-1457. Elsevier Science Publishers BV.

McFadden, D., and Train, K. 2000. “Mixed MNL models for discrete response.” *Journal of Applied Econometrics* 15: 447-470.

Scarpa, R., and Thiene, M. 2005. “Destination Choice Models for Rock Climbing in the Northeastern Alps: A Latent-Class Approach Based on Intensity of Preferences.” *Land Economics* 81 (3): 426–444.

Schelling,T.C. 1971. “Dynamic Models of Segregation” *Journal of Mathematical Sociology* 1: 143-186.

Schirmer, P. M., van Eggermond M.A.B., and Axhausen K.W. 2014. “The Role of Location in Residential Location Choice Models: A Review of Literature.” *Journal of Transport and Land Use* 7 (2): 3-21.

Schaake, K., Burgers, J., and Mulder, C. 2010. “Ethnicity at the individual and neighborhood level as an explanation for moving out of the neighbourhood.” *Population Research and Policy Review* 29: 593–608.

Sermons, M. W. 2000. “Influence of race on household residential utility.” *Geographical Analysis* 32 (3): 225–246.

Swait, J.R. 1994. “A structural equation model of latent segmentation and product choice for cross-sectional revealed preference choice data.” *Journal of Retailing and Consumer Services* 1: 77-89.

Train, K. 2009. *Discrete Choice Methods with Simulation*. II edition, Cambridge Books: Cambridge University Press.

Van der Laan Bouma-Doff, W. 2007. “Involuntary isolation: Ethnic preferences and residential segregation”. *Journal of Urban Affairs* 29: 289–309.

Walker, J., 2001. “Extended discrete choice models: Integrated framework, flexible error structures and latent variables.” PhD diss., MIT, Cambridge, MA.

Walker, J., and Ben-Akiva, M. 2002. “Generalized Random Utility Model.” *Mathematical Social Sciences* 43 (3): 303-343.

Walker, J., and Li, J. 2007. “Latent Lifestyle Preferences and Household Location Decisions” *Journal of Geographical Systems* 9 (1): 77-101. doi: 10.1007/s10109-006-0030-0

Zhang, J., and Zheng, L. 2015. “Are people willing to pay for less segregation? Evidence from U.S. internal migration.” *Regional Science and Urban Economics* 53: 97-112.

Zorlu, A., and Mulder, C. H. 2008. “Initial and subsequent location choices of immigrants to the Netherlands”. *Regional Studies* 42: 245–264.