# Ethnic segregation and residential location choice in Cape Town

Tatjana Ibraimovic[1], Stephane Hess[1], Hazvinei T.T. Moyo[2], Mark Zuidgeest[2], Romain Crastes dit Sourd[1]

[1] *Choice Modelling Centre and Institute for Transport Studies, University of Leeds, UK*

[2] *Centre for Transport Studies, University of Cape Town, South Africa*

**Abstract**

The Group Areas Act of 1950 left post-apartheid South African cities with residential spatial patterns where diverse ethnic groups largely stayed in separated neighbourhoods. In the recent past, public housing policies have been implemented in order to achieve a more integrated society, encouraging social and ethnic mixing. However, when this is perceived as an imposed desegregation, such policies are ineffective as ethnic polarisation often persists even within mixed neighbourhoods. In this paper, we aim to explore whether the preferences for ethnic neighbourhood composition of households from different ethnic backgrounds indicate a voluntary desire for residential mixing. Using a stated preference survey of neighbourhood choice, we analyse the location attributes driving the residential location choices in Cape Town. We find that ethnic neighbourhood composition plays a big role in influencing the choice behaviour in this particular urban context. In particular, we find that each ethnic group holds preferences for co-ethnic neighbours, while exhibiting different sensitivities to the presence of other ethnic groups, where these sensitivities far outweigh the sensitivities to other factors such as travel time, cost, safety and school quality.

**Keywords**: ethnic residential segregation; Cape Town; residential location choice models; ethnic preferences; stated choice; housing policies; post-apartheid

## 1. Introduction

The residential choice behaviour of different ethnic communities can be guided at least in part by preferences for ethnic neighbourhood composition, and this can be the cause of the onset of urban ethnic segregation (Ibraimovic and Hess, 2017). However, in some countries segregation has developed historically through an imposed spatial separation process. This is the case of South Africa, where the demarcation of neighbourhoods into Groups in 1950's (Group Areas Act) and the forced displacement of different ethnic communities produced ethnically homogeneous areas. In Cape Town, this resulted in a residential composition where the Central Business District (CBD) was surrounded by areas with a high percentage of white residents, while exurban high density suburbs were dominated by black Africans, coloured, Indian and Asian ethnicities who were disconnected from the centres of economic activity.

In the post-apartheid period, Cape Town inherited city structures that were mired in inefficiencies, ranging from an unequal access to economic and social opportunities, peripheral location of low income households and segregated access to transport infrastructure (Boraine et al., 2006). Following the end of Apartheid, enforcement by law of segregation has been replaced by economic forces which continued to sustain ethnic separation. As Turok (2001) states: "Institutional practices and market forces are tending to reinforce spatial

divisions, with costly consequences for the poor majority of the population and for the wider urban economy and society." Today, the severe income inequalities, with South Africa topping the international inequity list with a GINI coefficient of 0.69 (Statistics South Africa, 2014), constrain the neighbourhood choice for a vast segment of underprivileged inhabitants, with the consequent sorting in space mainly through the housing market.

Relatedly, several policy initiatives were developed by the South African government trying to address some of these inefficiencies. The national housing plan *Breaking New Ground* (2004) was proposed with the aim of solving housing issues for previously marginalized groups. *Inclusionary Housing* projects under this initiative were implemented in order to overcome accessibility constraints by dedicating 20% of all new housing developments to diverse or disadvantaged ethnic groups. In Cape Town this promoted neighbourhoods such as Delft and Westlake, partially reaching ethnic integration. Even though it has to some extent encouraged mixed race neighbourhoods, the outreach was more limited than expected (Klug et al., 2013). The success of such policies requires people's willingness to move to mixed developments, both by disadvantaged and wealthier ethnic groups. Indeed, if perceived as imposed desegregation, these measures might be less effective.

The above discussion highlights the importance of understanding the willingness of different ethnic groups to relocate and live in ethnically diverse areas and neighbourhoods. Hence, the aim of this study is to explore the role of ethnic preferences in residential location choice decisions for households from different ethnic backgrounds. Following Ibraimovic and Hess (2017), the main goal here is to reveal the "pure preferences" effect, free from choice-constraints (accessibility, affordability, discrimination, etc.) so as to understand whether such preferences indicate a desire for residential mixing. Given that the severe income inequalities within ethnic groups impose substantial constraints on neighbourhood choices, especially for underprivileged ethnic groups, data on real world choices (i.e. revealed preferences data) is not suitable for analysis in this context. We therefore rely on a Stated Choice (SC) approach where respondents are asked to make trade-offs between ethic neighbourhood composition and other location characteristics. This permits us not only to reveal the ethnic preferences but also explore their importance among other location choice drivers. Our analysis draws on a similar study in Switzerland using a SC survey of neighbourhood choice (Ibraimovic, 2013; Ibraimovic and Masiero, 2014), which was adapted and conducted in Cape Town in February 2017.

The remainder of this paper is organised as follows. We first discuss the spatial context before turning our attention to survey design (Section 3) and model specification (Section 4). After discussing the results of our analysis in Section 5, we present some brief conclusions and an outlook for further work in Section 6.

## 2. Spatial context

### 2.1. History of ethnic segregation and todays spatial inequalities

This study focuses on neighbourhoods of the Greater Cape Town area, the second most populated city in South Africa and one of the most ethnically diverse cities in the world. According to the 2011 South African population Census, the major ethnic groups in Cape Town were the coloured ethnic group with 42.4%, followed by black African ethnic group with 38.6%, white ethnic group with 15.7% and other ethnic groups (Asian, etc.) with 3.3%. Today, the city's population still experiences high levels of ethnic segregation

inherited from the apartheid era. Even if the spatial separation on an ethnic basis started in the early 1900s, many historically mixed neighbourhoods continued to exist until the 1950s. In these years, a new National Party government implemented the Group Areas Act which produced the most notable example of racial spatial separation. It took place in three phases, where the initial stage entailed the demarcation of neighbourhoods into Groups. In the Cape Town context, this was followed by the forced removal of established communities, with the black African and coloured ethnicities being dislocated from neighbourhoods earmarked for white inhabitants. A wave of housing construction in the peripheries for displaced populations followed, producing a clear spatial separation of the main ethnic groups (Wilkingson, 2000). Examples of neighbourhoods built for black Africans include Nyanga and Gugulethu, while Blue Downs, Delft and Mitchell's Plain were subsequently developed for the coloured population. Moreover, several informal settlements such as Crossroads emerged, populated by Africans with no housing permission. The inhabitants of these so called "squatter camps" were later on relocated into new townships such as Khayelitsha (Turok, 2001).

Today, over 20 years after the end of apartheid, Cape Town remains a severely polarised city. The inherited ethnic segregation continues to exist driven by severe income imbalance across ethnic groups and the economic forces which took over the legislative ones in sustaining the separation. Figure 1 highlights the important inequalities across South African ethnic communities.

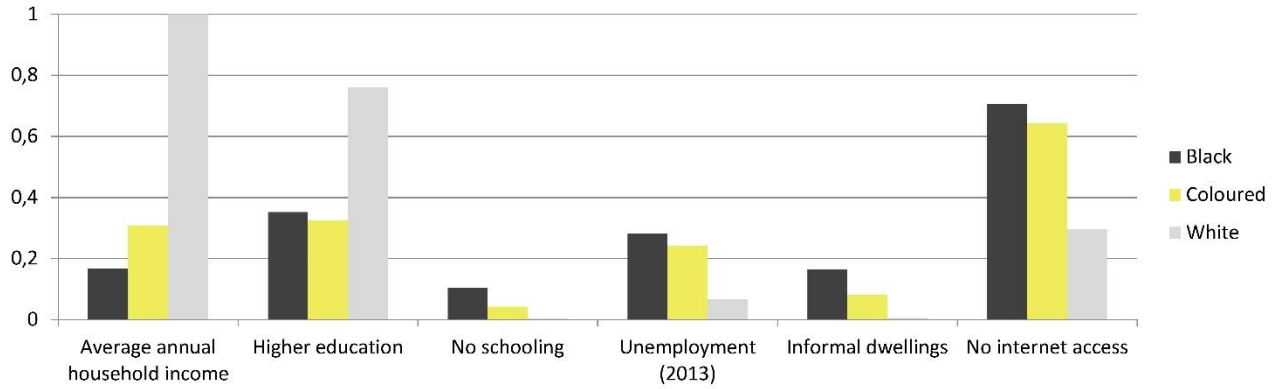


*Figure 1: Socio-economic differences across ethnic groups in South Africa:*

Source: Community Profiles (Statistics South Africa, Census 2011); Quarterly Labour Force Survey (Statistics South Africa, Q3 2013).

The average annual income for the black African population is less than 20% of that of whites, whereas coloured reach just above 30%. Differences in education and unemployment are also notable, with 16% of black African and 8% of coloured residents still living in informal settlements. Figure 2 shows how ethnic distribution highly correlates with the income distribution in Cape Town's neighbourhoods.

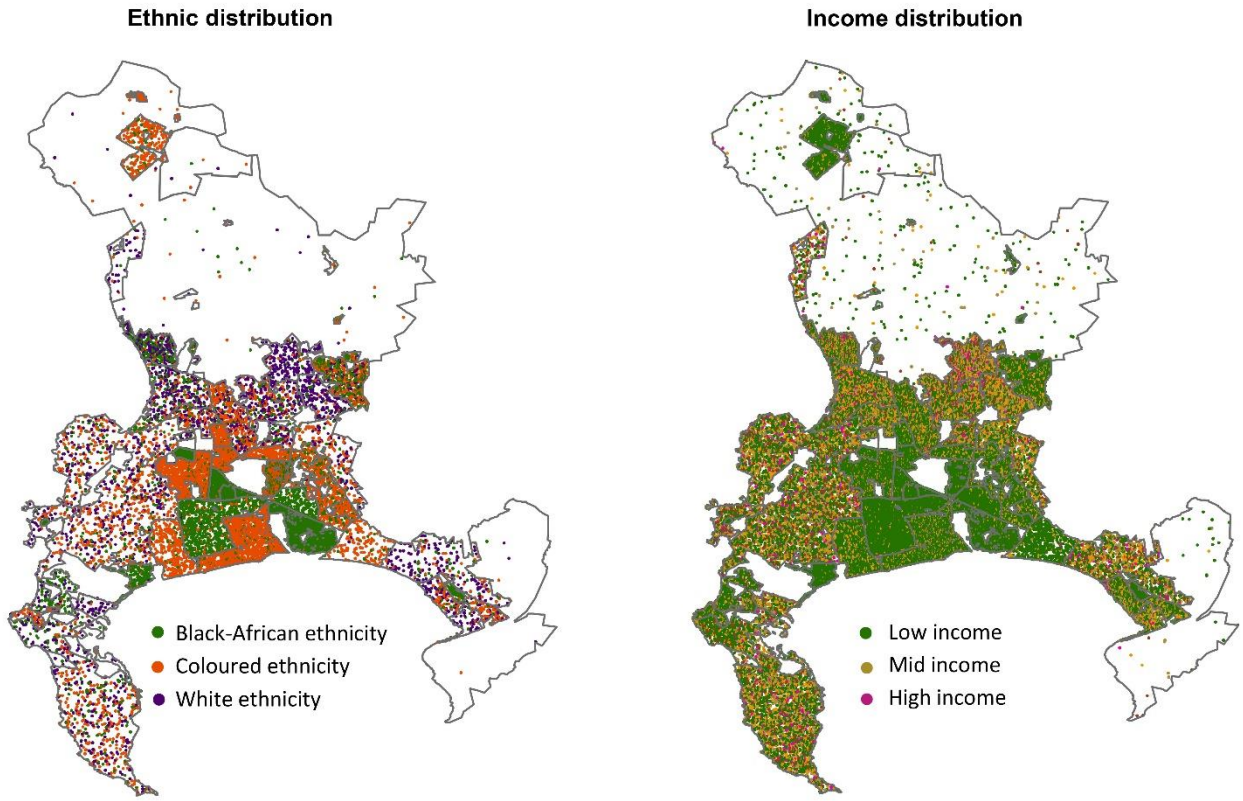


*Figure 2: Residential ethnic and income segregation in Cape Town (population Census 2011)*

Source: Statistics South Africa, Census 2011 (georeferenced data). Elaboration: Centre for Transport Studies, University of Cape Town; South Africa.

Such spatial separation continues to produce repercussions at the economic, social and urban level, especially for the underprivileged ethnic groups. Various studies report effects including the limited use and privatization of public space (Landman, 2004), displacement of crime (Atkinson and Blandy, 2005), unequal provision of services and amenities (Low, 2008), reinforcement of socio-spatial inequalities (Newton and Schuermans, 2013) and low interaction and contact between ethnic groups (Atkinson and Flint, 2004). Moreover, the segregation and unequal distribution of economic activities forces a large number of periphery inhabitants into long-distance commuting, with high costs in terms of congestion, pollution and travel time (Turok, 2001). Indeed, the concentration of employment in the middle and high income areas, examples being Claremont, Century city and Tyger Waterfront, has precipitated the exclusion of low income individuals to access employment. This has resulted in the low income (public transport) cohort spending about 43%, against the international norm of 10%, of their income on transport (TDA, 2017; Department of Transport, 1996, 2015), thus living in transport poverty.

### 2.2. Housing market and desegregation policies

Today's housing market in Cape Town is mainly influenced by liberal economic policies and international investments. Gentrification in some areas such as Woodstock has resulted in the urban poor being pushed further to the peripheries of the city where they are socially and economically excluded (Booyens, 2012). This sustained the persistent polarisation in the location of centres of economic activities and provision of public services. Such forces tend to maintain or even reinforce segregation through socio-economic

inequalities, imposing important accessibility constraints to higher quality neighbourhoods for underprivileged population segments.

Even some government efforts to provide housing for the poorest, such as the 1990's Urban Development Framework (National Department of Housing, 1997), have been criticised for reinforcing ethnic separation. Such measures indeed resulted in extra-urban development of low income housing, given the lower construction and land costs associated with these locations. The succeeding plan, Breaking New Ground in 2004, thus aimed to change the segregation patterns by building spatially, economically and socially integrated human settlements. The change in the paradigm was the promotion of a people centered approach vs a technocratic approach as well as involvement of beneficiaries through consultation, communication and education. Various housing policies have been adapted to integrate land and transport planning such that there is access to economic and social activities with the aim of reducing marginalisation (South African Government, 2012).

Drawing from experiences in other country contexts, Inclusive Housing (IH) schemes were proposed to overcome accessibility constraints, by compelling the developers to dedicate 20% of new housing projects to diverse and disadvantaged population groups. As in Europe and the USA, there are limits to the extent to which such housing policies can contribute to social mixing. But the dynamics are very different in a context where there are far more severe inequalities, and spatial marginalisation is not a minority concern (Klug et al, 2013). Indeed, the success of such housing policies lies in people's willingness to relocate and live in ethnically mixed residential spaces. This highlights the importance to gain an in-depth understanding of the preferences for ethnic neighbourhood composition which are at the core of this analysis.

## 3. Data

In order to address the objectives of this study, we use data on residential location choice decisions of a sample of Cape Town inhabitants, collected through a face-to-face paper-based survey across 24 neighbourhoods in January and February 2017. Different neighbourhood types were selected for the survey, according to the level of ethnic segregation as well as the average income level. In particular, the city neighbourhoods were classified into segregated and mixed neighbourhoods based on the 2011 Census data (Statistics South Africa, 2011). Then the segregated neighbourhoods were divided according to the dominant ethnic group as black-African, coloured and white neighbourhoods, while the mixed neighbourhoods were classified according to their known average income levels into low, mid and high income neighbourhoods. The survey was conducted in 5 segregated neighbourhoods of each dominant ethnic group and in 3 mixed neighbourhoods of each income type, as presented in Table 1 and shown in Figure 3.

Eligible respondents were all residents in the selected neighbourhoods who were over 18 years old, stratified by their ethnic background. This sampling strategy permitted us to collect a database which includes respondents of different ethnic groups living in different neighbourhood types. The final sample size consisted of 162 respondents, all of which completed a pre-survey questionnaire and the Stated Choice (SC) component of neighbourhood choice.

*Table 1: Sampling strategy in selected Cape Town neighbourhoods*

| 3 segregated neighbourhood types | Cape Town neighbourhoods |
|---|---|
| Black Africans (>75%) | Langa, Nyanga, Khayelitsha, Philippi, Cross Roads |
| Coloured (>75%) | Athlone, Mitchells Plain, Mamre, Belhar, Grassy Park |
| Whites (>75%) | Durbanville, Fishhoek, Pinelands, Brackenfell, Melkbosstrand |
| **3 mixed neighbourhood types** | **Cape Town neighbourhoods** |
| Low income | Klipheuwel, Kraaifontein, Fistantekraal |
| Mid income | Philadelphia, Blackheath, Blue Downs |
| High income | Scarborough, Hout Bay, Somerset West |

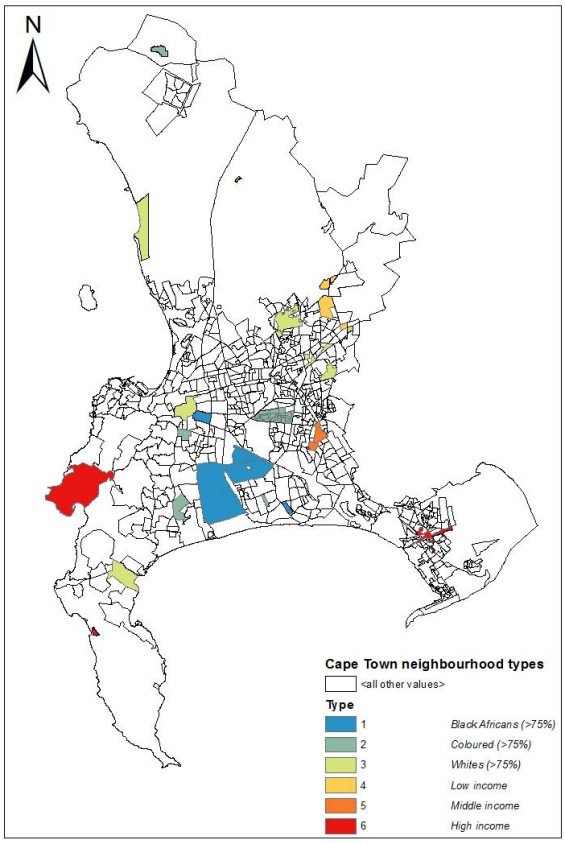


*Figure 3: Survey locations*

## 3.1. Stated choice survey of neighbourhood choice

To reveal the preferences for ethnic neighbourhood composition we use a SC survey of neighbourhood choice, building on the work of Ibraimovic (2013), Ibraimovic and Masiero (2014) and Ibraimovic and Hess (2016, 2017), who employed the approach for studying ethnic preferences in the Swiss urban context. Following these studies we design a similar experiment adapting it to the urban and ethnic context of Cape Town. The experiment involves 12 choice tasks for each respondent, each time with a choice among three alternative neighbourhoods: the respondent's current neighbourhood of residence (which represents the reference alternative) and two unlabelled hypothetical alternative neighbourhoods. Each neighbourhood

alternative is described by attributes representing the main location choice drivers in Cape Town. In order to focus on neighbourhood characteristics only and reduce the influence of other factors and the resulting endogeneity bias, respondents were asked to assume that their present dwelling (in terms of size and quality) and all other neighbourhood characteristics not considered in the experiment, remain the same. Moreover, no relocation costs are assumed.

In each of the 12 scenarios, a respondent was faced with a choice between their current neighbourhood and two hypothetical alternative neighbourhoods. The three alternatives were described to the respondent in terms of travel time to the city centre, monthly expenditure, school quality, property crime and violent crime, and finally ethnic composition. The attributes for the first (reference) alternative were kept invariant across choices for the same respondent, while a D-efficient design was used to determine the attribute combinations for the remaining alternatives (see Rose & Bliemer, 2014). Our design produced 7 blocks of 12 choice tasks each, where each respondent was allocated to one of these 7 blocks.

The approach used in the design process varied across attributes.

- *Travel time to city centre*: we presented the current travel time in minutes for the reference alternative, while the values for the remaining two alternatives were pivoted around that reference value, with five possible changes, namely -50%, -33%, 0%, +33% and +50% (with rounding to the nearest five minutes). The attribute represents the average time that the respondents employ to reach the city centre with the usual transport mode. We use Cape Town city centre as a proxy for travel time to work as most of the jobs are located in the central parts of the city. It is also a convenient variable for the unemployed population, with the expectation of positive utility associated with proximity to the city centre as suggested by the bid rent theory (Alonso, 1964).

- *Monthly expenditure*: this attribute represents the living costs based on the average expenses (e.g. communal taxes, transportation costs, etc.) of living in a particular Cape Town neighbourhood. As we did not have reliable data on current expenses, we presented an “as now” level for the reference alternative, with positive or negative shifts presented for the two hypothetical alternatives. Here, we assume that such costs amount to 10% of the average income as per national benchmark (Department of Transport, 1996) in each neighbourhood, and we pivot the cost values on the living costs in the respondent’s current neighbourhood, as depicted in the 2011 Census, with changes of -25%, -12.5%, 0%, +12.5% and +50%, where the changes were shown in rounded values in South African Rand (ZAR)[1].

- *School quality*: this attribute represents the average 2016 matriculate pass rate (News24, 2016) among the schools in each of the neighbourhoods. We define five levels of school quality, ranging from the lowest (bachelor pass rate below 20%) to the highest (bachelor pass rate above 80%). Access to quality education is deemed to be an important factor for the residential location choice. This is due to the unequal distribution of education quality and type (private vs public) which correlates with the neighbourhood income and thus the ethnic composition (Motala, 2009; Sayed

[1] 1 ZAR ≈ 0.06 EUR (December 2017)

and Ahmed, 2011; Frempong et al, 2011). As such, this variable is also suitable as a proxy for the overall neighbourhood quality. The school quality attribute was kept fixed for the reference alternative, but was chosen freely by the design optimisation process for the remaining two alternatives, choosing amongst the five possible levels (i.e. no pivoting).

- *Crime occurrence*: this attribute is defined by two distinct variables, one for *violent crime* and the other for *property crime*. The variables are computed from 2016 and 2017 statistical data (Cape Town database) on the rates and types of crime in Cape Town neighbourhoods (CrimeStats, 2017). Each type of crime has three levels representing low, average and high occurrence. As shown by many studies in the South African context (see e.g. Lemanski, 2004, 2006a) the fear of crime constitutes one of the main elements influencing the residential choices. As for school quality, the crime attributes were kept fixed for the reference alternative, but were chosen freely by the design optimisation process for the remaining two alternatives, choosing amongst the three possible levels (i.e. no pivoting).

Small priors were used for these neighbourhood characteristics, with negative priors for time, cost, property and violent crime, and a positive prior for school quality, where the difference in utility between the best and worst level was kept constant across attributes.

A different approach was used for the sixth ethnic composition attribute. *The ethnic neighbourhood mix* represents the percentage of each of the main three ethnic groups living in Cape Town, i.e. black African, coloured and white ethnicity, as derived from the 2011 Census. The survey visually presents the ethnic neighbourhood composition for each of the alternatives, with the aim to improve the comprehension of the concept of ethnic shares for the respondents and reduce the cognitive burden over different choice tasks.[2] For the reference alternative we use the actual ethnic shares in the respondent's current neighbourhood, whereas the values for the alternative neighbourhoods are selected by the design process from 21 combinations of different ethnic mix, going from the most segregated to the fully balanced neighbourhoods, as presented in Table 2. This ensures that even for respondents who currently live in a very segregated neighbourhood, a mixed neighbourhood can occur in the choice tasks, with the reverse for respondents who currently live in mixed neighbourhoods. The main aim is to identify preferences for ethnic neighbourhood mix for different levels of ethnic segregation and especially the preferences for self-clustering within the own ethnic group as well as the willingness to live with other ethnic groups constituting the Cape Town urban population. No priors were used for this attribute for this specific reason.

Figure 4 shows an example of a choice task presented to respondents. Respondents were asked to indicate their preferred option amongst the three alternatives. If they indicated their current neighbourhood as their preferred option, then they were additionally asked to make a "forced" choice between the two remaining options.

[2] We acknowledge the possibility that visual representation of a specific variable can draw the attention to such variable, potentially influencing its importance among different attributes. However, we deem such representation necessary for the comprehension of the choice tasks by the respondents in this specific contexts.

*Table 2: Combinations of different ethnic neighbourhood mix*

| Type | black % | coloured % | white % | |
|---|---|---|---|---|
| 1 | 100 | 0 | 0 | fully segregated |
| 2 | 0 | 100 | 0 | |
| 3 | 0 | 0 | 100 | |
| 4 | 80 | 8 | 12 | nearly segregated |
| 5 | 12 | 80 | 8 | |
| 6 | 8 | 12 | 80 | |
| 7 | 60 | 20 | 20 | quite segregated |
| 8 | 20 | 60 | 20 | |
| 9 | 20 | 20 | 60 | |
| 10 | 44 | 28 | 28 | nearly balanced |
| 11 | 28 | 44 | 28 | |
| 12 | 28 | 28 | 44 | |
| 13 | 36 | 32 | 32 | fully balanced |
| 14 | 32 | 36 | 32 | |
| 15 | 32 | 32 | 36 | |
| 16 | 40 | 40 | 20 | balanced for two, nearly segregated for other |
| 17 | 40 | 20 | 40 | |
| 18 | 20 | 40 | 40 | |
| 19 | 48 | 48 | 4 | balanced for two, segregated for other |
| 20 | 48 | 4 | 48 | |
| 21 | 4 | 48 | 48 | |

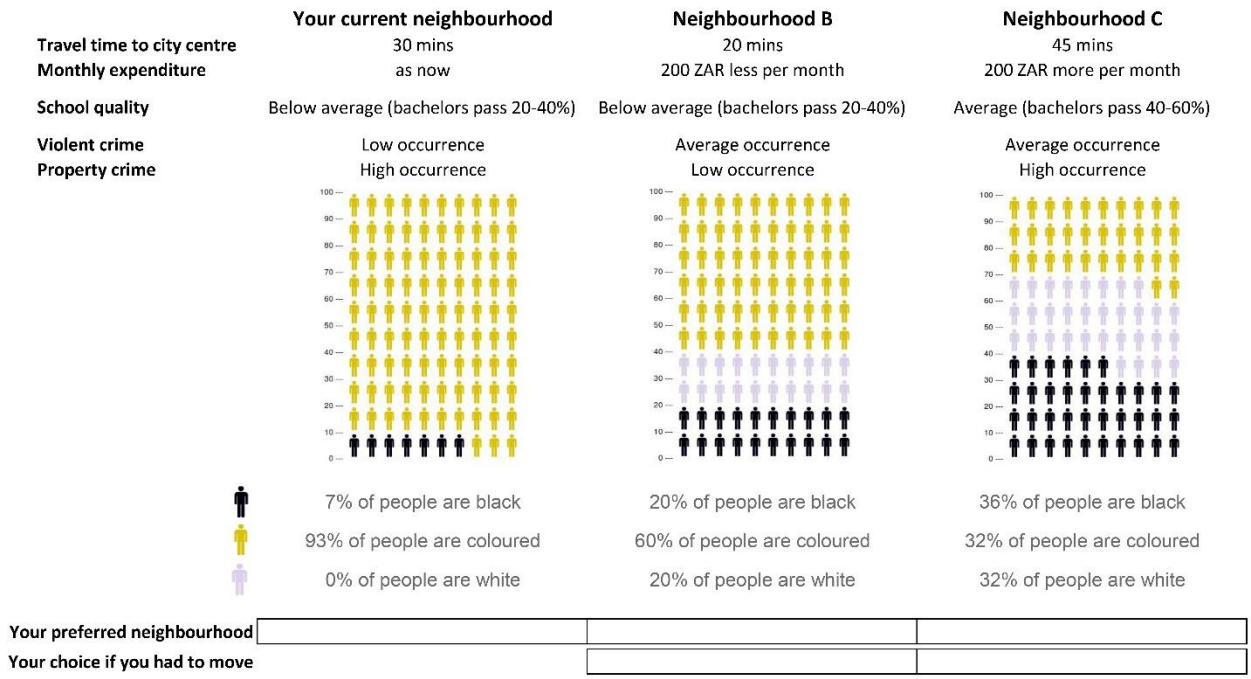


*Figure 4: Example of a choice task from the SC survey*

In addition to the choice data, the survey also collected information on respondents' socio-economic and demographic characteristics.

## 4. Modelling approach

We use Mixed Multinomial Logit models to model the responses to the stated choice scenarios, with a particular emphasis on attempts to capture what we expect to be high levels of heterogeneity in sensitivities across individual respondents. Let the utility that individual *n* obtains from alternative *j* in neighbourhood choice situation *t* be denoted as:

$$U_{jnt} = \delta_{jn} + \beta_n' x_{jnt} + \varepsilon_{jnt} \tag{1}$$

Where $\delta_{jn}$ is a constant for alternative *j* (set to zero for *j=3*), $x_{jnt}$ is a vector of observed attributes and $\beta_n$ is a corresponding vector of utility coefficients. Let $\theta_n = (\delta_{nj}, \beta_n)$ be a vector grouping together $\delta_{nj}$ and $\beta_n$, where all elements in $\theta_n$ are allowed to vary across individuals, a point we return to below. Finally, $\varepsilon_{jnt}$ is a random error term that corresponds to the unobserved component of utility. We assume that the unobserved term $\varepsilon_{jnt}$ is *iid* extreme value, so that the probability that individual *n* chooses neighbourhood alternative *i* in choice situation *t* conditional on $\theta_n$ corresponds to the logit formula:

$$P_{int}(\theta_n) = \frac{e^{\delta_{in} + \beta_n' x_{int}}}{\sum_{j=1}^{3} e^{\delta_{jn} + \beta_n' x_{jnt}}} \tag{2}$$

As we discussed in the data section, if a respondent chose their reference neighbourhood in a given choice situation, then they were asked to additionally make a "forced" choice between the remaining two alternatives. In that case, Equation (2) would be rewritten as:

$$P_{int_2}(\theta_n) = \frac{e^{\mu_2\left(\delta_{in} + \beta_n' x_{int}\right)}}{\sum_{j=2}^{3} e^{\mu_2\left(\delta_{jn} + \beta_n' x_{jnt}\right)}} \tag{3}$$

where $\mu_2$ allows for differences in the model error between the main choice and forced choice.

Let $i_{nt}^*$ be the alternative chosen by respondent *n* in choice task *t* and let $i_{nt_2}^*$ be the alternative chosen in the "forced" choice, which applies only if $i_{nt}^* = 1$. We then have that the probability of the outcome for choice task *t* for person *n* is given by:

$$P_{nt}(\theta_n) = P_{i_{nt}^*}(\theta_n) \cdot \left[(i_{nt}^* = 1) \cdot P_{int_2}(\theta_n) + (i_{nt}^* \neq 1)\right] \tag{4}$$

where $P_{int_2}(\theta_n)$ only contributes to $P_{nt}(\theta_n)$ if $i_{nt}^* = 1$.

As mentioned above, the elements in $\theta_n$ are allowed to vary randomly across respondents, using a joint distribution $f(\theta_n \mid \Omega)$, where $\Omega$ is a vector of parameters to be estimated, relating to the means and covariance structure of the elements in $\theta_n$. As the actual value of $\theta_n$ for a given respondent is not observed by the analyst, the choice probabilities are given by a multi-dimensional integral of the MNL probabilities in

Equation (4) over the distribution of $\theta_n$. We assume that the sensitivities are stable across the 12 choices for each given individual, and the probability of the sequence of choices observed for person *n* is then given by:

$$L_{int} = \int_{\theta_n} \prod_{t=1}^{12} P_{nt}(\theta_n)\, f(\theta_n|\Omega) d\theta \qquad (5)$$

Our utility specification includes alternative specific constants for the first two alternatives, a continuous time and cost coefficient, as well as a continuous coefficient for the ethnic composition attribute. As the share of the three ethnic groups needs to sum to 100%, we use the coefficient for the black group as the base by fixing it to zero. For the school quality, violent crime and property crime attribute, we used dummy coded coefficients, setting the coefficient for the lowest category to zero. In our specification of the heterogeneity, we imposed a monotonic effect across levels, a point we return to below. For all coefficients except the time and cost coefficients, we used separate coefficients for each of the three main ethnic groups.

We now look at our distributional assumptions for the individual components of the utility function:

- for the alternative specific constants, we constrain the constant for the third alternative to zero, and use a Normal distribution for the remaining two constants;
- for travel time and cost, we use negative lognormal distributions, ensuring a negative contribution to utility for increased travel time or cost;
- for school quality, we fix the utility of the lowest quality level to zero, and estimate four additional coefficients that follow positive lognormal distributions, where these are specified in an additive manner to ensure monotonically increasing utilities. The first coefficient is used for all school quality levels that are equal or better than “below average”, the second is used for all levels that are equal or better than “average”, etc. These can be seen as “shift” coefficients, where e.g. the contribution to utility for the “above average” level would be given by $\beta_{school\ below\ average} + \beta_{school\ average} + \beta_{school\ above\ average}$;
- for property and violent crime, we again fix the utility of the lowest level to zero, and then estimate two additional coefficients (for each) that negative lognormal distributions, where these are again specified in an additive manner, ensuring that the highest (worst) level had a more negative value than the middle or the lowest level.

Given our split of all parameters except time and cost by ethnicity, the above specification leads to the use of 38 randomly distributed coefficients, along with the non-random scale parameter for the forced choice $\mu_2$. The flexibility of the model is further enhanced by estimating a full covariance matrix between these 38 random coefficients. This flexibility comes at the cost of increased estimation complexity, and we thus use Bayesian estimation of our model, using RSGHB (Dumont et al., 2015). For an in-depth discussion of MMNL models and HB estimation, see Train (2009).

## 5. Model results

The final model estimation was carried out on the choices of 162 respondents, making 12 main choices each, leading to an initial sample of 1,944 observations. Amongst those, 1,296 were for the reference alternative, showing a strong status quo effect, and justifying the use of the “forced” second choice. The final

log-likelihood of the model was -2,452.68, with a log-likelihood at zero of -3,034.021, giving a $\rho^2$ measure of 0.19.

Bayesian estimation produces posterior distributions for the model parameters, i.e. the means of the individual coefficients as well as the covariance matrix. For each one of these components, we can calculate a posterior mean (which has similar properties to a maximum likelihood estimate) and a posterior standard deviation (which has similar properties to a maximum likelihood standard error). We start by looking at the means of the underlying Normal distributions for the individual randomly distributed coefficients (i.e. also for the Lognormally distributed coefficients, which are the exponentials of Normals) and the fixed scaled parameter.

The first output in Table 3 relates to the fixed scale parameter $\mu_2$ for the "forced" choice, where the value larger than 1 indicates that the amount of noise in the forced second choice is less than in the first choice, i.e. a greater influence for the modelled part of utility. This could suggest that when making a choice between the reference neighbourhood and the competing neighbourhoods, a larger number of factors outside the model play a role than for the forced choice.

The remaining values in the tables are the posterior means and standard deviations for the means of the underlying Normal distributions. For the alternative specific constants (ASC), we see a strong preference for the reference neighbourhood, with no major additional utility for the second alternative compared to the third (which has a constant of zero). For the school and crime parameters, the values shown relate to the means of the Normal distributions used inside the exponential to produce Lognormals. The negative sign simply implies that the median of the resulting Lognormals is less than 1. The implications of the shift parameters will be discussed in more detail below when we turn to the transformed parameters. Finally, for ethnic composition parameters, these again follow Normal distributions, so that the means can be interpreted directly. We see that each ethnic group has the highest preference for its own co-ethnics at the mean. For whites, there is little difference between coloured and black, nor is there between white and coloured for black respondents. However, coloured respondents strongly prefer white over black. The estimates shown here relate solely to the mean parameters, and the full covariance matrix of 38 by 38 terms is available on request from the authors.

We next used the Bayesian estimates of the means and the full covariance matrix between the 38 randomly distributed coefficients to produce means and standard deviations for the actual marginal utility components. This involves transforming from Normals to Lognormals and also adding the shift parameters to obtain level specific coefficient values for the school and crime coefficients. These results are shown in Table 4. We see the already discussed result that the ASC for the present neighbourhood of residence is positive across all three ethnic groups, but also note each time a high level of random heterogeneity as shown by the standard deviation.

The coefficient estimates for travel time to city centre (in minutes) and for neighbourhood living costs (in ZAR) are negative as a result of using a negative Lognormal distribution. They are very small in magnitude, for both the mean and the standard deviation, highlighting the dominance of ethnic variables and other location characteristics as the main drivers of neighbourhood choice. Looking at school quality, we see the

monotonic increases in utility with increasing school quality that are the result of the use of additive Lognormal distribution, with the opposite applying to the two crime attributes. Substantial levels of random heterogeneity are retrieved across respondents for all these attributes.

*Table 3: Raw estimation results - means and standard deviations (in brackets) for Bayesian posteriors for means of underlying Normal distributions and for fixed scale parameter*

| | ***Generic*** | ***White ethnicity*** | ***Coloured ethnicity*** | ***Black African ethnicity*** |
|---|---|---|---|---|
| **Scale "forced" choice** | 1.37 (0.16) | | | |
| **Travel time** | -3.29 (0.37) | | | |
| **Monthly expenditure** | -10.09 (0.28) | | | |
| **ASC current neighbourhood** | | 2.00 (0.48) | 2.06 (0.31) | 2.38 (0.38) |
| **ASC neighbourhood B** | | 0.09 (0.14) | 0.18 (0.09) | 0.09 (0.09) |
| **School quality (base: lowest = 1)** | | | | |
| *below average (change from 1)* | | -7.6 (0.43) | -4.86 (0.57) | -4.37 (0.48) |
| *average (change from 2)* | | -7.23 (0.35) | -4.65 (0.58) | -3.12 (0.28) |
| *above average (change from 3)* | | -5.39 (0.36) | -6.07 (0.28) | -2.61 (0.22) |
| *high (change from 4)* | | -2.14 (0.35) | -4.02 (0.27) | -4.21 (0.16) |
| **Violent crime (base: low occurence = 1)** | | | | |
| *average occurrence (change from 1)* | | -6.47 (0.4) | -1.97 (0.21) | -3.46 (0.39) |
| *high occurrence (change from 2)* | | -2.3 (0.25) | -6.15 (0.44) | -3.09 (0.43) |
| **Property crime (base: low occurrence = 1)** | | | | |
| *average occurrence (change from 1)* | | -3.86 (0.33) | -5.04 (0.52) | -3.48 (0.3) |
| *high occurrence (change from 2)* | | -4.51 (0.24) | -4.2 (0.32) | -4.8 (0.37) |
| ***Ethnic preferences*** *(base: % black Africans)* | | | | |
| % whites | | 2.41 (0.3) | 1.31 (0.21) | -4.22 (0.44) |
| % coloured | | -0.21 (0.2) | 3.22 (0.28) | -4.09 (0.51) |

We already touched on the findings for the ethnic composition attributes and now investigate those in more detail. As mentioned earlier, the strongest effects are found for the share of co-ethnics, suggesting strong self-segregation preferences for all three ethnic groups. All ethnic groups are more probable (at the mean) to choose neighbourhoods with a higher shares of their co-ethnics, as denoted by the positive coefficients for white respondents for the share of whites and for coloured respondents for the share of coloured. For black respondents, the negative coefficients for the shares of the two other ethnic groups suggest that they too prefer neighbourhoods with a larger share of black-Africans. Moreover, we can note that the preferences for co-ethnics are strongest for the black-African respondents, followed by coloured and white ethnic groups. Nevertheless, there is also a substantial degree of variability, denoted by the large standard deviations, which suggests a spectrum of self-segregation preferences for different ethnic groups, ranging from no preferences for co-ethnics to strong self-clustering.

We also highlight again how the ethnic preferences follow a particular hierarchy of desired ethnic groups. Respondents of white ethnic background are almost indifferent between the share of black-Africans and coloured, those of coloured ethnicity prefer whites to black African neighbours, while black-African

respondents slightly prefer coloured to white ethnicity. These results are also shown in Figure 5 featuring coefficient distributions for the share of white and coloured ethnicity (where black African ethnicity is the base group) in the neighbourhood across the three ethnicities.

*Table 4: Implied means and standard deviations for utility parameters*

| Attributes | description | generic | | White ethnicity | | Coloured ethnicity | | Black African ethnicity | |
|---|---|---|---|---|---|---|---|---|---|
| | | mean | std. dev. | mean | std. dev. | mean | std. dev. | mean | std. dev. |
| Alternative specific constants *(base: neighbourhood C)* | current neighbourhood | | | 1.973 | 3.295 | 2.102 | 2.998 | 2.276 | 3.633 |
| | neighbourhood B | | | 0.088 | 0.574 | 0.167 | 0.438 | 0.099 | 0.373 |
| Travel time | to city centre | -1.01E-04 | 1.36E-04 | | | | | | |
| Monthly expenditure | living costs | -6.67E-07 | 8.02E-07 | | | | | | |
| *School quality (base: lowest)* | below average | | | 0.001 | 0.002 | 0.038 | 0.164 | 0.019 | 0.021 |
| | average | | | 0.140 | 2.950 | 0.060 | 0.167 | 0.200 | 0.657 |
| | above average | | | 0.212 | 2.993 | 0.113 | 0.401 | 0.322 | 0.664 |
| | high | | | 0.384 | 3.001 | 0.244 | 0.944 | 0.351 | 0.674 |
| Violent crime *(base: low occurrence)* | average occurrence | | | -0.006 | 0.015 | -0.199 | 0.199 | -0.106 | 0.415 |
| | high occurrence | | | -0.394 | 1.480 | -0.217 | 0.205 | -0.347 | 1.125 |
| Property crime *(base: low occurrence)* | average occurrence | | | -0.060 | 0.164 | -0.044 | 0.171 | -0.132 | 0.587 |
| | high occurrence | | | -0.074 | 0.164 | -0.101 | 0.240 | -0.327 | 2.027 |
| Ethnic composition *(base: % Black African ethnicity)* | % White ethnicity | | | 2.480 | 2.652 | 1.341 | 0.946 | -4.275 | 4.211 |
| | % Coloured ethnicity | | | -0.207 | 0.840 | 3.154 | 2.395 | -4.119 | 4.783 |
| Scale second choice | | 1.374 | n/a | | | | | | |

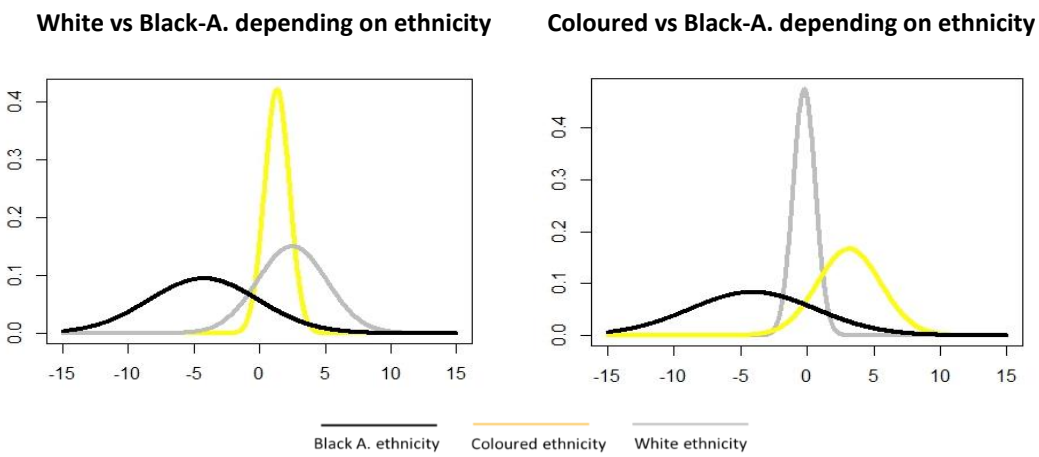


*Figure 5: Distribution of preferences for shares of white and coloured ethnicity for the three ethnic groups*

Even if other location characteristics have much less influence on residential choice decisions, some interesting heterogeneity can be found when differentiating across ethnic groups. The school quality increases the utility and thus the probability of a neighbourhood to be chosen, while crime occurrence has the opposite effect. Figure 6 shows the effects of these attributes on choice decisions for the three ethnic groups. School quality increases the utility for all three groups, however the effect is stronger for the black-

African respondents (especially in the average and above average range) and white respondents (especially in the high school quality range). Crime, on the other hand, negatively impacts the choices, in particular for the black-African ethnicity for both types of crime, while the white ethnic group is mainly concerned about the violent crime.

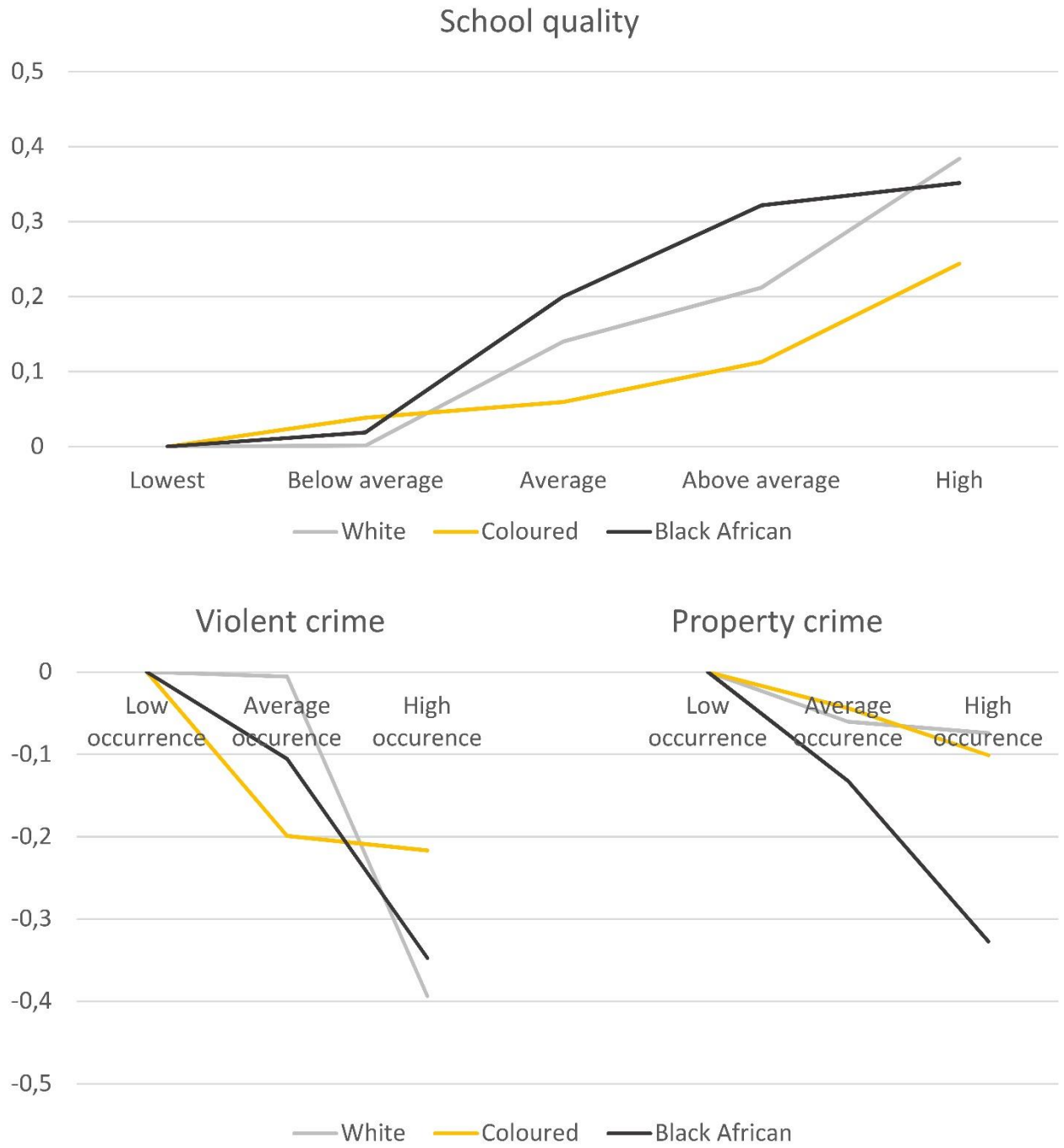


*Figure 6: The impact of school quality and crime occurrence across ethnic groups*

We can gain more insight into the preferences heterogeneity across ethnicities by analysing the trade-offs between different neighbourhood characteristics (Table 5). A positive ratio is obtained for coefficients that are either both for desirable attributes or both for undesirable attributes. A negative ratio is obtained when one attribute is desirable and the other is not. We first examine the impact of an increase in non-ethnic

attributes versus the increase of school quality from the lowest to the highest level, as shown in the first part of the Table 5. As said before, travel time and living cost variables have a very low impact on location choices which might be due to the poor engagement of the respondents with these two variables in comparison with other neighbourhood characteristics, especially the ethnic mix.

*Table 5: trade-offs between different neighbourhood characteristics*

| **Increase in … vs school quality (1 to 5)** | White | Coloured | Black African |
|---|---|---|---|
| travel time (min) | -5.38E-04 | -1.87E-03 | -6.50E-04 |
| living costs (Rand) | -6.37E-06 | -2.26E-05 | -2.44E-06 |
| violent crime (1 to 3) | -2.381 | -6.354 | -1.160 |
| violent crime (2 to 3) | -2.359 | -2.554 | -0.786 |
| violent crime (1 to 2) | -0.022 | -3.800 | -0.374 |
| property crime (1 to 3) | -0.380 | -0.611 | -1.897 |
| property crime (2 to 3) | -0.126 | -0.373 | -1.600 |
| property crime (1 to 2) | -0.254 | -0.238 | -0.297 |
| school quality (4 to 5) | 0.779 | 0.378 | 0.093 |
| school quality (3 to 5) | 0.890 | 0.466 | 0.520 |
| school quality (2 to 5) | 0.991 | 0.712 | 0.882 |
| **Replace 10% … vs school quality (1 to 5)** | | | |
| black by white | 2.729 | | |
| coloured by white | 2.713 | | |
| black by coloured | -0.016 | | |
| black by coloured | | 17.293 | |
| white by coloured | | 15.478 | |
| white by black | | -1.816 | |
| white by black | | | 2.285 |
| coloured by black | | | 2.289 |
| coloured by white | | | 0.004 |
| **Replace 10% … vs violent crime (1 to 3)** | | | |
| black by white | -6.201 | | |
| coloured by white | -7.486 | | |
| black by coloured | -1.285 | | |
| black by coloured | | -2.943 | |
| white by coloured | | -1.787 | |
| white by black | | 1.156 | |
| white by black | | | -4.388 |
| coloured by black | | | -4.520 |
| coloured by white | | | -0.132 |
| **Replace 10% … vs property crime (1 to 3)** | | | |
| black by white | -14.059 | | |
| coloured by white | -14.164 | | |
| black by coloured | -0.105 | | |
| black by coloured | | -89.866 | |
| white by coloured | | -88.985 | |

| | | |
|---|---|---|
| white by black | 0.881 | |
| white by black | | -11.399 |
| coloured by black | | -11.843 |
| coloured by white | | -0.445 |

Across all three ethnic groups, an increase in violent crime from the lowest to the highest level has a stronger impact than an increase in school quality from the lowest to the highest level, while this is even true for an increase from the middle level to the highest level for white and coloured respondents. The relative sensitivity to violent crime vs school quality is highest for the coloured ethnic group, followed by whites. Among the three ethnic groups, only the black Africans are more concerned about property crime than school quality, especially for high crime occurrence. Regarding school quality levels, we see that for the white ethnic group, a major jump occurs from the second highest to the highest level, where this is not at all the case for black African respondents.

Looking at the trade-offs between ethnic shares and school quality, we compute ratios for replacing a 10% of other ethnic groups by co-ethnics, or by the other remaining ethnicity, compared again to an increase in school quality from the lowest to the highest level. Across all three ethnic groups, we see that replacing 10% of another ethnic group by co-ethnics has a much stronger impact on the utility than an increase in school quality from the lowest to the highest level. The highest ratios are observed for the coloured ethnic group, which is either very sensitive to the share of co-ethnics or least sensitive to school quality, when compared with the other two ethnic groups. Black African and white ethnicities, on the other hand, exhibit similar levels of sensitivity to these two variables. In line with the earlier estimation results, we see that replacing 10% of one non-co-ethnic group by the other non-co-ethnic group has an impact only for coloured respondents, who have a dislike for replacing white neighbours by black African. In other words, they show hierarchy in preferences firstly for their co-ethnics, then for white ethnicity and lastly for black-African neighbours.

Similarly, we explore the trade-offs between the ethnic variables and the different types of crime. Even if the shares of co-ethnics still dominate in the choice decisions, significant differences in trade-offs exist across the ethnic groups. The coloured ethnicity exhibits the highest sensitivity to violent crime amongst the three groups, but lowest to property crime, considering the trade-offs with co-ethnics. The black African group shows the highest relative concern for property crime among the three groups, followed by the whites.

## 6. Conclusions

This study has analysed the residential choices of inhabitants of the Greater Cape Town area. Special focus was given to the preferences for ethnic neighbourhood composition and their importance in driving the residential location choice behaviour of the main ethnic groups, namely the black African, coloured and white groups. A Stated Choice survey of neighbourhood choice was designed, so as to reveal the “pure preferences” effect, free from choice constraints imposed by the severe accessibility limitations for the underprivileged population segments.

The results suggest that preferences for ethnic mix of neighbourhoods still dominate the residential location choice decisions in Cape Town. Indeed, we find that when choosing their neighbourhood of residence, people have specific preferences in terms of the mix among different ethnic groups. Each ethnic group has the strongest preferences for own co-ethnics, but is also sensitive to the presence of the other ethnicities. Indeed, ethnic preferences follow a particular hierarchy of desired ethnic groups. Respondents of white ethnic background are almost indifferent between the share of black-Africans and coloured, those of coloured ethnicity value whites more than black African neighbours, while black-African respondents slightly prefer coloured to white ethnics.

The segregation thus might not only be determined by the exclusion of less privileged ethnic groups, but also to their own preferences to live among their ethnic communities. Even if this study does not investigate the motivations behind such preferences, it would be essential to broaden our knowledge of the underlying attitudes and beliefs which might trigger them. As argued in some studies, the preferences of disadvantaged groups for self-segregation might lie in the neighbourhood being perceived as a “safe heaven” due to the absence of discrimination practices (van der Laan Bouma-Doff, 2007) or in the higher levels of perceived neighbourhood social cohesion (Havekes et al., 2014). On the other hand, the advantaged groups might be concerned about the perceived deterioration of level of security or social issues. In future studies, it would be valuable to integrate soft characteristics such as attitudes, beliefs and perceptions about the ethnically mixed neighbourhoods into the residential location choice analysis.

A high level of heterogeneity in ethnic preferences however indicates a certain willingness for residential mixing by respondents in all three ethnic groups. Such preferences for integration should be sustained and translated into opportunities of ethnic mixing. For example, beyond ethnic preferences, different ethnic groups show different sensitivities to other location attributes such as the school quality and safety. While improving the general quality of a neighbourhood may thus counterbalance some of the desire for segregation, these impacts will be weak given the overall higher sensitivity to neighbourhood composition.

Overall, the results of this study suggest that ethnic preferences should be accounted for when developing policy measures sustaining a greater residential integration. Indeed, as previous experience has shown, focusing on the accessibility aspect only could lead to the housing measures being underutilised or inefficient in promoting residential integration.

Given the high level of heterogeneity in ethnic preferences, further analysis should investigate the nonlinearities the preferences structure, potentially revealing thresholds or tipping points for the shares of specific ethnic groups. Finally, the resulting preferences for ethnic and non-ethnic location characteristics could be used as input in Agent-Based simulation models describing the rules of agents’ behaviour over their residential location decisions. Such models could then be used to test different policy measures, developing different scenarios and predicting the impacts on future ethnic segregation dynamics. Substantial additional work is also needed in terms of data collection, with scope for longitudinal data on real world residential location choices in an ethnically diverse context.

**Acknowledgements**

The authors wish to acknowledge ITS Engineers South Africa and Mr. Storm Wright for assisting in the data collection. This project was carried out with the financial support by the European Research Council through the consolidator grant 615596-DECISIONS and by the Swiss National Science Foundation through the Early Postdoc.Mobility fellowship n. P2TIP1_168560.

## References

Alonso, W. (1964). *Location and Land Use. Toward a General Theory of Land Rent*. Cambridge: Harward University Press. http://dx.doi.org/10.4159/harvard.9780674730854

Atkinson, R., Blandy, S. (2005). Introduction: International perspectives on the new enclavism and the rise of gated communities. *Housing Studies*, 20, 177–186.

Atkinson, R., Flint, J. (2004). Fortress UK? Gated communities, the spatial revolt of the elites and time–space trajectories of segregation. *Housing Studies*, 19, 875–892.

Boraine, A. et al. (2006). The state of South African cities a decade after democracy. *Urban Studies*, 43 (2), 259–284.

Crime Stats South Africa (2017). https://www.crimestatssa.com/toptenbyprovince.php?ShowProvince=Western%20Cape

Dumont, J., Keller, J., Carpenter, C., 2015. RSGHB: Functions for hierarchical Bayesian estimation: A flexible approach. R package version 1.1.2, URL http://CRAN.Rproject.org/package=RSGHB.

Frempong, G., Reddy, V., Kanjee, A. (2011). Exploring equity and quality education in South Africa using multilevel models. Compare: A Journal of Comparative and International Education, 41: 6.

Ibraimovic, T. (2013). Investigating the role of ethnic preferences in residential location decisions: Choice analysis on Stated Preferences data. Ph.D. Thesis, USI, Switzerland. doc.rero.ch/record/32276

Ibraimovic, T., 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.

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

Ibraimovic, T., Hess, S. (2017). A latent class model of residential choice behaviour and ethnic segregation preferences. *Housing Studies* (forthcoming).

Klug, N., Rubin, M., Todes, A. (2013). Inclusionary housing policy: a tool for re-shaping South Africa's spatial legacy? *Journal of Housing and the Built Environment*, 28 (1), pp. 667-678.

Landman, K. (2006). Privatising public space in post-apartheid South African cities through neighbourhood enclosures. *GeoJournal*, 66 (1–2), 133–146.

Lemanski, C. (2004). A new apartheid? The spatial implications of fear of crime in Cape Town, South Africa. Environment and Urbanization, 16, pp. 101-112.

Lemanski, C. (2006a). Residential responses to fear (of crime plus) in two Cape Town suburbs: implications for the post-apartheid city. Journal of International Development, 18, pp. 787-802.

Lemanski, C. (2006b). The impact of residential desegregation on social integration: Evidence from a South African neighbourhood. Geoforum, 37:417-435.

Low, S. M. (2008). Incorporation and gated communities in the greatermetro-Los Angeles region as a model of privatization of residential communities. *Home Cultures*, 5, 85–108.

Motala, S. (2009). Privatising public schooling in post-apartheid South Africa – equity considerations. Compare: A Journal of Comparative and International Education, 39: 2.

Newton, C., Schuermans, N. (2013). More than twenty years after the repeal of the Group Areas Act: Housing, spatial planning and urban development in postapartheid South Africa. *Journal of Housing and the Built Environment*, 28, 579–587.

News24 (2016). 2016 Matric Results. http://matricresults.news24.com/. Accessed: 10 February 2016.

Rose, J.M. and Bliemer, M.C.J. (2014), "Stated choice experimental design theory: the who, the what and the why", in Hess, S., Daly, A. (Eds.), Handbook of Choice Modelling. Edward Elgar, Cheltenham, pp. 152–177.

Sayed, Y., Ahmed, R. (2011) Education quality in post-apartheid South African policy: balancing equity, diversity, rights and participation. Comparative Education, 47: 1.

Statistics South Africa (2013). Population Census 2011. Pretoria, Statistics South Africa, 2013. http://www.statssa.gov.za/. Accessed: February 2016.

Statistics South Africa (2014). Poverty Trends in South Africa: An examination of absolute poverty between 2006 and 2011. Pretoria: Statistics South Africa, 2014. Technical Report No.: 03-10-06.

Statistics South Africa (2015). Measuring household expenditure on public transport, In-depth analysis of National Household Travel Survey 2013 data. Pretoria: Statistics South Africa, 2015. Technical Report No.: 03-20-11.

TDA (2017). Comprehensive Integrated Transport Plan 2017 – 2022. Cape Town: The City of Cape Town's Transport and Urban Development Authority, 2017. Draft report.

Train, K. (2009). *Discrete Choice Methods with Simulation (2nd edn)*. Cambridge University Press, Cambridge.

Turok, I. (2001). Persistent polarisation post-apartheid? Progress towards urban integration in Cape Town. *Urban Studies,* 38 (13): 2349–2377.