# Segregation and the Perception of the Minority*

Florent Dubois†, Christophe Muller‡

*Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS*

April 22, 2016

**Abstract**

In his seminal work, Schelling (1971) shows that even preferences for integration generate high levels of segregation. However, this theoretical prediction does not match with decreasing levels of segregation observed since the 1970s. We build a general equilibrium model in which preferences depends on the number of peers and unlike individuals, but also on the weight they attribute to living in the minority or along a sizable minority, which we call their perception of the minority. In this framework, there always exists a structure of the preferences for which integrated equilibria emerge and are stable. Even when individuals are racist, there is still a level of the perception of the minority for which integration is a stable outcome. We then propose an econometric method to derive the structural preference parameters of the model in the case of South Africa. Estimated preferences provide evidences toward more integration.

*We would like to thank David Albouy, Renaud Bourlès, Yann Bramoullé, Pierre-Philippe Combes, Joachim De Weerdt, Bruno Decreuse, Matthew Ellman, Maxime Gueuder, Nobuyuki Hanaki, Matthew Jackson, Alan Kirman, Eliana La Ferrara, Gibson Mudiriza, and Juliette Rouchier for their valuable comments. We particularly thank Reza Daniels for the many helpful discussions. We thank seminar participants at the 2nd International Conference on Sustainable Development in Dakar, at the Journée "Regards Croisés Economie Gestion" in Aix-en-Provence, and at the AMSE PhD Seminar. Part of this research was conducted while Florent Dubois was visiting the University of Cape Town, whose hospitality is gratefully acknowledged. We would like to express our greatest gratitude to Takwanisa Machemedze, Alex Montgomery, and Lynn Woolfrey at DataFirst for their precious help with the data. All remaining errors are ours. Any views expressed in this paper are only those of the authors and do not reflect the position of any institution.

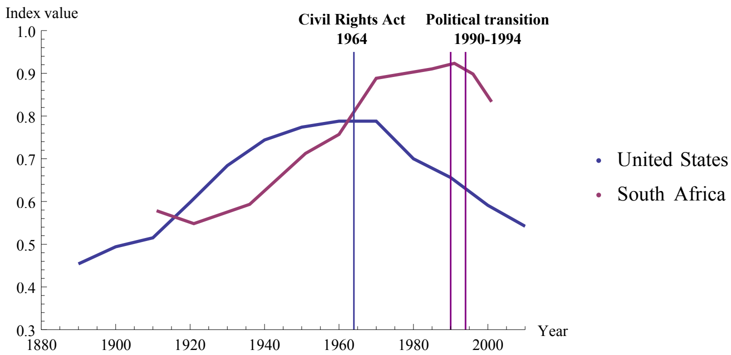
†corresponding author: florent.dubois@univ-amu.fr

‡email adress: christophe.muller@univ-amu.fr

# 1 Introduction

In his seminal contribution, Schelling[49] demonstrates that even if individuals exhibit preferences for living in a mixed neighborhood, complete segregation is the most likely outcome. However, the trend observed in the United States since the 1970s is declining and has almost returned to the levels observed before the Jim Crow Era.[1] In South Africa, we observe a similar trend since the end of the Apartheid in 1994 (see Figure 1). In this paper, we propose a theoretical model which can account for both segregationist and integrationist patterns. We then provide a structural econometric analysis of South Africa.

Figure 1: Evolution of segregation in the United States and South Africa




This graph plots the Black-White segregation for the United States and South Africa. It is measured by the dissimilarity index (Duncan and Duncan[27]). Data for the United States come from Glaeser and Vigdor[33]. Data for South Africa come from Christopher[18][19][20]. In both cases, segregation is an average of the main metropolitan areas.

We argue that people may perceive some benefits from living in integrated neighborhoods and choose to relocate in these newly desirable locations. Individuals are even willing to pay at least \$300 for less segregation in the United States (Zhang and Zheng[58]). Or they may simply value diversity (Aldrich et al.[2], and Wong[54]). These benefits can be of different orders. We might think for instance to complementarities in the job market as Blacks and Whites specialize in different tasks,[2] or to the improvement of risk-sharing due to the different assets held by

[1]Even if Black-White segregation is declining, Hispanics and Asians are still segregated in the United States (Charles[17], De la Roca et al.[23])

[2]We draw a parallel with the migration literature here. By specializing in different tasks, migrants and natives first avoid competition (Peri and Sparber[46]). Then they also increase wages and production due to complementarities in the tasks performed. Borjas[9] estimates the gain between \$7 and \$25 billion in the United States while Ottaviano and Peri[44] find an increase

Blacks.[3] These kind of effects depends explicitly on the size of the minority in the neighborhood which is why we talk about perception of the minority.

We build a location choice model in which we directly implement this complementary externality in the utility function of the individuals next to racial attitude effects. We assume that the utility function depends on the number of individuals of the different groups living in the neighborhood as Schelling and the following literature did.[4] We then depart from the literature by adding a term measuring the size of the minority in the utility function which aims to capture their perception of the minority.

We find that integration can emerge as a stable outcome although individuals have homophillic preferences. In this case, the premium generated by the presence of a minority needs to be sufficiently high to circumvent the effect of homophillic preferences. But this situation is dependent from the initial conditions. In fact, if the minority is too small in both locations, then each minority prefers to relocate in the location where their own group is the majority rather than keep on living in the minority. As a consequence, segregation emerges despite a benefit of integration positive enough. But segregation is not robust anymore, contrary to the literature. A shock strong enough will now displace the economy into a stable integrated state. The situation described by Schelling appears to be a special case of our model. We find also cases of non convergent dynamics.

Previous theoretical papers mainly try to explain the persistence of segregation.[5] Individuals look at their direct neighborhood on a grid and moves when dissatisfied by the racial mix. In this context, even strict preferences for integration lead to complete segregation due to space constraints. Our model is related to the bounded-neighborhood model (Schelling[49]) in which individuals look at the racial mix of the broad area where they live.[6] However, except for Miyao[42], the model is a partial equilibrium model accounting only for changes in one precise location. Thus, integration can be stable in one location while segregation increases as a whole. We follow Miyao[42] by having a general equilibrium framework. Note however that stable integrated equilibria in Miyao's work arise because preferences are sufficiently weak so that individuals move in fact at random in the absence of another mechanism of location selection. We also depart from the threshold utility function assumed after Schelling[49], because the linear form has some empirical

of 0.6% on average of the wage of natives.

[3]Bramoullé and Kranton[12] show that if there are risk-sharing relationships across communities, those who are linked (directly or indirectly) across neighborhoods have a higher welfare while those who are not have a lower condition but the aggregate welfare is higher.

[4]See Schelling[48][49][50], Pancs and Vriend[45], Zhang[57][56], or Grauwin et al.[35]

[5]See the previous footnote.

[6]See Schelling[48][49], Miyao[42], Granovetter and Soong[34], and Dokumaci and Sandholm[25].

supports (Bruch and Mare[14], and Easterly[28]). Moreover, it can also generate higher level of segregation than the threshold utility function in some cases (Van de Rijt et al.[52], and Bruch and Mare[15]).

We then propose an empirical strategy to recover structural estimates of the perception of the minority and other preference parameters. We use the asymmetric impact of the perception of the minority between locations dominated by Whites and those dominated by Blacks to identify our parameters of interest. We then recover the structural parameters in a two-stages procedure. First, we estimate the dynamics suggested by the theoretical model. Then, by identification between the theoretical and the econometric models, we implement a least-square solution of the overidentified system.

We adress the concerns of heteroscedasticity by using robust covariance matrices, and of endogeneity of the racial mix with instrumental variables. We exploit the spatial dependence between group's locations. If a location is inhabited by a large number of Whites, neighboring locations are likely to be inhabited by Whites too, and so on for neighboring locations of neighboring locations. At the same time, unobserved factors are likely to be correlated only locally. Therefore, there is a certain distance from which the number of individuals in the neighboring locations is still correlated with populations in the origin but no longer with the unobserved factors.[7] We thus use the average number of group members in the neighboring rings at order 2 and 3.

We use data from the South African censuses of 1996, 2001, and 2011 to avoid the Apartheid era. We harmonize the data at the 2001 subplace level to get a consistent and stable geography throughout the three waves. Our different estimations provide similar results. In all specifications, we find that the perception of the minority exceeds racists preferences in absolute value. While there are studies estimating preferences for the United States[8] or Singapore[9], we are, to the best of our knowledge, the first to study it for South Africa. Moreover, the type of data used is really important for segregation as different aggregation levels may produces contrasting pictures. A city may look integrated while all its neighborhoods may be completely segregated. However rich disaggregated data are difficult to find. Thus, researchers have employed public use micro areas (PUMAs) for city level analyses.[10] On the other hand, others have employed detailed data, usually of a particular city, for neighborhood based analyses.[11] Our study uses data of the

[7]Kasy[37] use a similar argument by using the average number of group members in neighboring units that are 3kms away from the origin.

[8]See Bajari and Kahn[5], Bayer et al.[7], Kasy[37], and Zhang and Zheng[58].

[9]See Wong[54].

[10]PUMAs are areas constructed to have at least 100 000 individuals. See Bajari and Kahn[5], and Zhang and Zheng[58].

[11]See Bayer et al.[7], Wong[54], and Kasy[37].

latter kind.

Our paper contributes to the literature on the estimation of preferences for neighborhoods.[12] However, our methodology is different. While they estimate preferences from the equilibrium relationship of a location choice model, we use instead the dynamics of such a model to estimate preferences. Because it takes time for individuals to adjust their location choices, estimating an equilibrium relationship at a particular time may not reflect the true preferences. Only two papers estimate a preference for diversity (Wong[54], and Zhang and Zheng[58]), but we all use a different identification strategy.[13] Wong[54] interprete the significantly negative sign of the squared number of Chineses, Indians, or Malays as a taste for other-group members. But the turning points from which other-groups are desired are not precisely estimated. Zhang and Zheng[58] insert segregation directly in the utility function. They are then able to estimate the willingness-to-pay for a decrease in segregation as a taste for diversity. However, they cannot distinguish the effect of own-group and other-group preferences. Instead, we specify a structure for the perception of the minority which allow us to explicitly distinguish between all preference parameters. The asymmetricity of the perception of the minority allows us to identify the effect. Finally, our goal is different from previous paper as we do not only want to estimate preferences. But we also want to assess if such estimated preferences are compatible with stable integration.

The paper is organized as follows. Section 2 first describes the model. We then characterize the necessary conditions for uniqueness and stability of an equilibrium. We analyse two relevent structures of preferences in the following subsection. Finally, we provide an extension of our framework, giving the basis for our identification strategy at the end of the section. In section 3, we describe the data and our empirical methodology. Then we present our reduced form and structural estimates at the end of the section. We discuss our results in the final section. All proofs are given in the appendix section.

# 2 A model of racial integration

## 2.1 The model

Think of a city (or a country) divided into two identical locations indexed on $i \in I = \{1; 2\}$. Two groups live in this city. Each individual is directly identifiable

[12]See Bajari and Kahn[5], Bayer et al.[7], Wong[54], Kasy[37], and Zhang and Zheng[58].

[13]Bajari and Kahn[5] find some preference for integration for Whites using a three-stage approach. But they restrict their sample only to working migrants which is likely to bias the estimates of such preferences.

by his type $k \in K = \{W; B\}$ which he can not hide.[14] The number of members of group $k$ living in location $i$ is thus denoted by $N_i^k$ and the total number of individuals of type $k$ in the city is represented by $L^k$. We consider that any number of individuals can live in the two locations. There are no spatial arrangements inside the neighborhoods. Individuals are either in or out. Thus, each location constitutes a bounded neighborhood, *i.e.* all individuals inside the location are neighbors with everyone else inside (as in Schelling[49][50]).

Individuals of type $k$ have an utility function $U^{ki}$ for living in location $i$. $U^{ki}$ is composed by a deterministic part[15] $u^{ki}$ and a stochastic part $\varepsilon^{ki}$, which can represent unobserved idiosyncratic characteristics, such that :

$$U^{ki} = u^{ki} + \varepsilon^{ki}. \tag{1}$$

As Schelling did, we assume that individuals care only about the racial mix of the location where they reside. Contrary to previous works, individuals also care about the presence of a sizable minority. Recent empirical works have shown the existence of such effect.[16] This takes the following form :

$$\begin{cases} u^{Wi}(N_i^W, N_i^B) = aN_i^W + bN_i^B + \gamma_W Min[N_i^W; N_i^B] \\ u^{Bi}(N_i^W, N_i^B) = cN_i^W + dN_i^B + \gamma_B Min[N_i^W; N_i^B] \end{cases}, \tag{2}$$

with $a, b, c, d, \gamma_W, \gamma_B$ real parameters expressing respectively the White taste for Whites, the White taste for Blacks, the Black taste for Whites, the Black taste for Blacks, and the White and Black perceptions of the minority. The choice of a linear form is debated in the literature.[17] Card et al.[16] show that tipping points occurs locally when the minority share exceeds a threshold ranging from 5% to 20%. However, globally they do not seem to be any tipping behavior (Bruch and Mare[14], Easterly[28]).

The interpretation of the last term is twofold. First, the min term can be seen as an explicit modelling of what Schelling[48][49][50] calls the minority status, the fact that individuals have a preference on whether they live in the minority or not. Depending on the sign of the $\gamma$ coefficients, it directly expresses the taste for living in the minority if I belong to the minority and it reflects my perception of the minority if I belong to the majority which is also equivalent to minority status.

[14]We refer to W and B as Whites and Blacks as the racial dimension of segregation is one of the most salient feature in the United States or in South Africa, but we could have chosen any other dichotomy such as the young and the old, the rich and the poor, girls and boys as Schelling explains[49][50].

[15]Which can be interpreted as the representative utility of the group $k$ for the location $i$ (McFadden[40] and Miyao[42]).

[16]See Aldrich et al.[2], Wong[54], or Zhang and Zheng[58].

[17]Grauwin *et al.*[35] also provide an analytical solution of the Schelling model with potential functions in the case of similar linear utility functions.

Second, this min function expresses an idea of economic complementarity between the groups. For instance, think about a rich White community which needs a certain amount of poor Blacks in order to do some jobs that they do not want to do themselves like cleaning the sewers or picking up the trashes.[18] This argument is supported by the literature on the impact of migration on natives as developped in the introduction. The underlying idea is that if someone wants something that the other group has (a specific good or service, an insurance against shocks ...), he should tolerate at least some of their members.

Individuals choose where they want to live according to a best response rule. Consequently, individuals of type $k$ select location $i$ with probability $P^{ki}$ such that the location they have chosen is the one that maximizes their utility :

$$P^{ki} = Pr(U^{ki} > U^{kj}, \;\; \forall j \neq i \text{ and } i, j \in I). \tag{3}$$

We assume that individuals do not move if they are indifferent as the inequality is strict. Consequently, the Nash equilibrium of the game is an allocation of individuals accros locations such that all players live in the location which maximizes their utility :

$$N_i^{k^*} = P^{ki} L^k, \quad \forall i \in I, k \in K \tag{4}$$

with

$$\sum_{i \in I} N_i^k = L^k > 0. \tag{5}$$

## 2.2 Existence

**Proposition 1.** *Under the model expressed above, a Nash equilibrium exists.*

As our model has only two localities, it is possible to only study the situation in one location, say location 1, as the situation in the other location is complementary. This allows us to define intuitively the different states in which the system may end.

**Definition 1.** *An equilibrium is said to be **integrated** if all locations match the racial mix of the society (i.e. if it lies on the line which equation is $N_1^B = \frac{L^W}{L^B} N_1^W$) and **segregated** otherwise. Moreover, it is said to be **completely integrated** if each location is equally populated (i.e. if $N_1^k + N_1^{-k} = N_2^k + N_2^{-k} = \frac{L^k + L^{-k}}{2}$ $\forall k \in K$). Segregation is **complete** if the two groups live entirely in a separate location. Finally, we say that a group **deserts** one location if this group lives entirely in only one location.*

[18] Alan Kirman told me this example.

The definition of the different states can be generalized to the case of n groups. Moreover, this definition is more general than what the literature is using. Actually, they say that a state is integrated if all locations are equally populated by the two groups. Nevertheless, this definition does not take into account the possible disequilibrium in the size of the two groups, especially when one group is clearly assumed to be the minority (and therefore the other one is the majority). In our case, an integrated state reflects the relative size of the two groups in the society[19] and we go back to the standard case used in the literature if both groups are of equal size. This discrepancy in the literature is again a matter of local versus global integration. If we are interested in only one location, there is not point in considering different mixtures than the 50-50 racial mix as integrated. But if we want to evaluate the level of segregation of the whole city, we should then acknowledge that an integrated state should reflect the city relative sizes of the groups in each location.

## 2.3 Uniqueness and stability

Before analyzing the properties of an equilibrium in this model, we have to made another assumption about the distribution followed by the stochastic part. In order to keep the model simple, we assume that $\varepsilon^{kj} - \varepsilon^{ki}$ follows an uniform distribution on the interval $[\alpha;\beta]$ with $\alpha < 0$ and $\beta > 0$. This assumption reduces the model to a linear probability model well-known in the discrete choice theory (Anderson et al.[3]). Because we are focusing only on location 1 we can simplify the notations by replacing $N_1^W$ by $W$ and $N_1^B$ by $B$. We then get the following system:

$$\begin{cases} W = \dfrac{\Delta u^{W1} - \alpha}{\beta - \alpha} L^W \\[2ex] B = \dfrac{\Delta u^{B1} - \alpha}{\beta - \alpha} L^B \end{cases} \tag{6}$$

with

$$\Delta u^{ki}(N_i^k, N_i^{-k}) = u^{ki}((N_i^k, N_i^{-k})) - u^{kj}(N_i^k, N_i^{-k}) \tag{7}$$

where $-k$ denotes the type different from $k$. For the sake of simplicity, let denote the size of the support of the uniform distribution $\theta \equiv \beta - \alpha$.

At this stage, it is convenient to assume equal population sizes $L^W = L^B \equiv L$ to alleviate computations. In this case we know explicitly the behavior of the min term in the utility function.[20] We can then compute the equilibrium as a function

[19]As mentioned by Fossett[30], and Clark and Fossett[21]

[20]See the proof of proposition 1 for explicit details of the behavior of the min function in this case.

of the parameters of the model :

$$\left\{\begin{array}{l} W^* = \dfrac{L[L(2b+\gamma_W)(\alpha+L(c+d+\gamma_B))-(\alpha+L(a+b+\gamma_W))(2dL+L\gamma_B-\theta)]}{L^2(2b+\gamma_W)(2c+\gamma_B)-(2aL+L\gamma_W-\theta)(2dL+L\gamma_B-\theta)} \\ \\ B^* = \dfrac{L[L(2c+\gamma_B)(\alpha+L(a+b+\gamma_W))-(\alpha+L(c+d+\gamma_B))(2aL+L\gamma_W-\theta)]}{L^2(2b+\gamma_W)(2c+\gamma_B)-(2aL+L\gamma_W-\theta)(2dL+L\gamma_B-\theta)} \end{array}\right. . \tag{8}$$

We have now to specify how the dynamic adjustment takes place if the city is out of equilibrium. We assume that individuals update their behavior according to the configuration of the society in the previous period. Deriving from the system (6), we have the following dynamic adjustment process :

$$\left\{\begin{array}{l} \dot{W} = \dfrac{((2a+\gamma_W)L-\theta)W_t+(2b+\gamma_W)LB_t-L^2(a+b+\gamma_W)-\alpha L}{\theta} \\ \\ \dot{B} = \dfrac{(2c+\gamma_B)LW_t+((2d+\gamma_B)L-\theta)B_t-L^2(c+d+\gamma_B)-\alpha L}{\theta} \end{array}\right. \tag{9}$$

with $\dot{W} = \dfrac{\partial W_t}{\partial t}$ and $\dot{B} = \dfrac{\partial B_t}{\partial t}$.[21] At this point, we may note that the perception of the minority plays a role of correction of the racial preferences. We then have the following properties :

**Proposition 2.** *Under the dynamic adjustment process (9), an integrated equilibrium is unique if and only if it is stable.*

The link between the uniqueness and the stability is due to the linearity of this specification. Let us define what is a structure of preferences in this framework.

**Definition 2.** *A* ***structure of preferences*** *is a set of constraints on the preference parameters for the own group and the other groups, a,b,c, and d.*

**Proposition 3.** $\forall a, b, c, d, \theta, L$ *such that* $\frac{\theta}{L} \neq 2(d-c)$, $a \neq b$, *and* $c \neq d$, *there always exists a combination of perception of the minority* $(\gamma_W^*; \gamma_B^*)$ *such that the city will be integrated (possibly) without any policy intervention no matter the initial configuration. Moreover, there also always exists a combination of perception of the minority* $(\gamma_W^P; \gamma_B^P)$ *for which the city can become integrated through a relocation policy if it were segregated in the first place. These statements are true for any structure of preferences.*

[21] The system can be solved analytically which is done in appendix.

The "(possibly)" term in the proposition comes from the fact that the system will behave as a (spiral) sink. Thus all trajectories should converge to the integrated equilibrium. However, as there cannot be a negative number of individuals in a location, some trajectories are constrained to converge toward a segregated state. Hence, the "(possibly)" term. We will illustrate this property in the "mutual reject" case.

## 2.4 Exploring different structures of preferences

Despite a large number of possible structures of preferences, we restrict our analyses to three cases that we think to be relevant. All these structures have in common to keep the attitude of Whites constant. In each situation, Whites will have a tendency to segregate by seeking peers and rejecting the other group. Although there are signs of greater tolerance from Whites, we believe that this is still the main attitude in both the United States (Krysan et al.[39]) and South Africa (Duckitt et al.[26], and Dixon et al.[24]). So, we want to explore the dynamics of segregation given the attitude of Whites with respect to three possible attitudes that the discriminated groups may adopt.

We will further make a difference between an equilibrium and a state in the following sense:

**Definition 3.** *A **state** is a dynamically instable racial mix which is constant over time.*

With this definition, an equilibrium is stable in the Newtonian sense of the term as at this point the sum of the pushing and pulling forces will be null. On the contrary, a state will be stable because the city is constrained to have only positive numbers of individuals in each location. Thus, there could be stable configurations in which the city is trapped. We call such situations *states*, they are dynamically unstable as the sum of forces at these particular points are not null. But the city will converge to these states as time goes to infinity. They are basically corner solutions. Hence the definition above. We will consider these states as (at least locally) stable.

### 2.4.1 Mutual reject

When both groups are racist ($b$ and $c < 0$), we say that both groups are rejecting each other. Moreover, as explained in the proof of the proposition 3, antisocial behaviors are very unlikely to be representative of a large share of the population. Thus, we also restrain individuals to have homophillic preferences as well ($a$ and $d > 0$). Sakoda[47] describes a similar situation in his "segregation" case. It also portrays the relationship between Afrikaaners and Africans (Duckitt et al.[26]).

We use the following parameters set: $a = 10$, $b = -8$, $c = -2$, $d = 6$.[22] Figure 2 depicts the bifurcation diagram and the associated dynamics.

From these graphics, we can see that integration can arise and be stable both with and without any intervention. This directly comes from proposition 3. For a couple $(\gamma_B^*; \gamma_W^*)$ lying in area 1, the integrated equilibrium is stable. This occurs actually for a positive value of the Black perception of the minority and a negative value of the White perception of the minority.

With such values of perceptions, if the city starts in a configuration in which Whites dominate the most populated location. Then if the number of Blacks in this location is sufficiently large, Whites will move rapidly in the location dominated by Blacks. Whites in the location they dominate benefit from a large number of peers but endure a consequent Black minority as well. So the negative impact of the Black presence is amplified by the negative perception of the minority Whites have. In the other location, Whites suffer from living in the minority but this effect is attenuated by living with some peers and by the small number that constitute the White minority. They also suffer from the size of the local Black majority. Whites face then a trade-off between a location with a lot of peers but with a too high Black minority and a location with few peers but with a Black majority. Since they strongly reject the association with Blacks through their strong negative perception of the minority, the utility differential is in favor of the location dominated by Blacks as soon as the Black minority is large enough.

The fact that Whites would abandon a location where they dominate for a location where they constitute the minority in the first place seems to be counter-intuitive. However, consider a situtation in which at least some Whites would have (in addition to their preferences for the racial mix) eugenic preferences concerning mating and marriages. Then, as the number of Whites in the location grows, it would be more difficult for these White eugenicists to influence their peers toward homogamous relationships. Thus, even if they hate Blacks, they might move in a location where they are the minority in order to insure their eugenic goal. Punishment of defectors would be easier and would increase the average utility of the White minority (Boyd et al.[11]). Bisin and Verdier[8] also describe such segregation norms in the marriage markets of French aristocrats, or Orthodox Jews.

On the other hand, Blacks will be more reluctant to enter a location dominated by Whites. In this kind of location, they will benefit from a small share of peers but this effect is positively reinforced by the complementarity with Whites. However, they will suffer from a large number of Whites. In the location dominated by Blacks, they will enjoy a large number of peers and the few Whites will cause them only minor discomfort. But they will benefit only moderately from the complementarity with Whites. This might be the case if Whites possess the

[22] For the sake of simplicity, we fix $\theta = 6$, $\alpha = -3$ and $L = 1$ for the rest of the section.

factors of production while Blacks mainly compose the labor force, which was the case during the Apartheid. Still today, this situation persists despite some redistribution policies. As Blacks value strongly the complementarity with Whites, the location dominated by Blacks is not as desirable as it seems compared to the one dominated by Whites. So the utility differential between the two locations is lower than for Whites. It is also in favor of the location dominated by Whites as Blacks want to live with the biggest minority possible.

Whites want to live with the lowest minority possible while Blacks desire the largest minority possible. So both groups are initially drifting apart as they prefer different locations. In the location dominated by Whites, a lot of Whites move out while at the same time a few more Blacks move in. While in the location dominated by Blacks, a lot of Whites move in and few Blacks move out. These two dynamics increase the utility differentials in the same direction. The location dominated by Whites becomes more attractive for Blacks as the minority increases and the number of Whites diminishes. While in the location dominated by Blacks, the minority increases but is still lower than the minority in the other location. Also, the Black majority decreases, which makes the Black dominated location more attractive for Whites. At some point, the number of Whites will be equal between the two locations. But as long as the number of Blacks is different in the two locations, Whites will continue to choose the location with the lowest minority while Blacks will choose the one with the largest minority even if the minority changed its color. This dynamic motion can continue to a completely segregated equilibrium.

However, after the minority changed color in both locations, there is a point from which Blacks switch their prefered location. This change occurs because the Black minority in the initially Black dominated location is sufficiently large compared to the White minority in the other location. Thus, the positive effect for Blacks of being in a Black minority offsets the negative impact of the large White majority and the benefit they would get from peers and a small White minority in the other location. Then Blacks will move in the White dominated location. Still, Whites are very sensitive to Black variations in their number. So, the Black minority will increase so much that even if Whites are all in the same location, they will evacuate this location. This process will move as described before until the city will reach the perfectly integrated equilibrium.

At this point, both Blacks and Whites do not want to move anymore because deviation will be harmful for both groups. Imagine that 1% of the Whites deviates from the equilibrium, then one location will have a 51% White majority but a 50% Black minority. As Whites do not like Blacks and Black minorities even more, they will face a strong penalty to live in this location. In the other location, 49% Whites will reside with 50% Black. However, the minority is first lower than in the location dominated by Whites and, as it is a White minority, homophilly weakens the negative perception of the minority. Thus, 1% of Whites will move back to the Black dominated location. If 1% of Blacks deviates, the equilibrium will be restored because homophilly strengthens the Black positive perception of the minority.

As mentioned before, even if for some trajectories the city will converge alone toward the perfectly integrated equilibrium,[23] there are some initial configurations for which the city ends in a completely segregated state. In this kind of situation, a government willing to build an integrated city could achieve its goal by implementing a relocation policy. However, the number of people displaced has to be sufficiently large in a one shot movement to bring back the city on an integration path. If the policy is not properly calibrated then people will move back to their previous location(s), or others will compensate their departure. This sort of big push policy are similar in the spirit to programs like the Moving To Opportunity experiment in the United States. Our analysis could help to calibrate the proper size of the program if it would be generalized [24].

Now, if a government would like to get an integrated city, he could also design a policy aiming at promoting positive perceptions of the minority for both groups. Thus, we would be in the first quadrant of the $(\gamma_B; \gamma_W)$-plan, and more generally in area 2 of the figure 2. The city behaves as a source, and can end in locally stable segregated or integrated state depending on the initial conditions. However, the government will be able to restore integration if the city was segregated in the first place by a sufficiently large relocation policy. Once again, the type of policy advocated is a kind of big push story.

If the city starts in a configuration where both locations are not too unbalanced, then the city will converge toward an integrated state where all individuals gather in the initially most populated location. If both groups perceive positively the minority, their utility will be larger in the location where the minority is the largest. If Whites dominates this location, they will also benefit from the large White majority through their homophillic preferences. If they constitute the minority, their positive perception of the minority will amplify their homophillic preferences. This effect will more than compensate the large Black majority if the White minority is large enough. In the least populated location, Whites will benefit less from both their majority status and from the Black minority. Thus the utility differential will be in favor of the most populated location for both groups. Thus both the minority and the majority groups will be larger as people moves from the least to the most populated location which widen further the utility differential.

If the city is initially unbalanced enough, the positive perceptions of the minor-

[23] One of them being the one starting from the big circle in the Sink picture of Figure 2.

[24] See Katz et al.[38] for details about the experiment

ity are not sufficiently large to compensate homophillic effects. Complete segregation would then occur. Whites in the location they dominate benefit largely from a large majority and benefit only a little from the small Black minority. In the other location, Whites will suffer from the large Black majority and gain only few from their minority status. The situation is symmetric for Blacks. So both groups will have a utility differential in favor of the location dominated by their peers. Thus, Whites will move in the White dominated location, while Black will move out, which broaden the utility differential. However, in this example, relocating more than 30% of Blacks (or more than roughly 45% of Whites) at once will bring back the city on an integration path.

Finally, if the social context is such that both groups perceive negatively the minority, then we are almost completely in the third quadrant of the $(\gamma_B; \gamma_W)$-plan. More generally, the area 3 of the figure 2 exhibits the same type of saddle dynamics, only segregated states will emerge. In this situation, the perceptions of the minority reinforce the segregation forces of homophilly and racism. Both groups want to live with the smallest minority possible. Whites living with a Black majority suffer both from being with a large number of Blacks and to live in the minority. While in the other location, their disutility of being associated with a Black minority is soften by the large peer presence they enjoy. So as long as the unlike minority is not too large, the utility differential for each group will be in favor of the location dominated by their peers. Thus, the two groups will drift apart and a completely segregated city will emerge.

There are also initial configurations in which the city may at first sight converge to the perfectly integrated equilibrium but finally collapse in a segregated state. If Blacks dominate the most populated location, the large White minority generates a lot of discomfort for them. In the other location, Whites constitute the majority but suffer moderately from the small Black majority. Then at first, both Blacks and Whites will prefer the location dominated by Whites and start moving in, which pulls the city toward more integration. However, at some point for Blacks, the racism effect of the growing White majority overcomes the homophillic effect of the growing Black minority (which is also counterbalanced by their negative perception of the minority). Then, the utility differential between the two locations changes. Whites still prefer the White dominated location whereas Blacks move in the Black dominated one. Finally, the two groups will drift apart until the city will be completely segregated as described earlier.

Once again we were in a situation in which a group, despite being largely dominating one location, will choose to leave this location. As a way to rationalize the Blacks' behavior, imagine that the White dominated location is the most affluent location. Then Blacks would move in to benefit from the higher standard of life in this affluent location. However, acculturation and status concerns may bring back Blacks in their previous location.

### 2.4.2 White segregation versus Black integration

Blacks may oppose to the White segregationist preference structure a desire to live with both Blacks and Whites. This translate into a positive taste both Blacks and Whites. We also choose to study the case of a stronger homophillic preference than the taste for the other group due to the homophillic bias already discussed earlier. Farley et al.[29] find such structure of preferences in the Detroit area in the late 1970s. In South Africa, Blacks have such preferences toward English speaking Whites, but English speaking Whites have also a friendly attitude toward Blacks (Duckitt et al.[26]). Our parameterization is set as follows: $a = 10$, $b = -8$, $c = 5$, $d = 7$. Figure 3 represents the associated bifurcation diagram.

In this context, a completely integrated equilibrium can emerge from two negative perceptions of the minority. More generally, integration can be generated by any combination of perceptions that lies inside the area 1 of Figure 3. However, the mechanism is similar as the one described in the "mutual reject" case. The main difference is that a relocation policy is not necessary in this case as the smallest level of integration is always prefered by Blacks. Assume that the city is completely segregated. Then, all Whites live in their most desirable location as they do not like Blacks and minority status. If 1% of Whites move out from this location, they will suffer from a number of Blacks, and constituting a small White minority. Compared to the all Whites location, they will choose to move back to their previous location, restoring the completely segregated state. On the other hand, if 1% of Blacks move in the all Whites location, they will benefit both from the maximum number of Whites, from the small presence of their peers, and from being in the minority. Compared to the all Black location, they will prefer to stay where they are. Thus the utility differential for Blacks shifts in favor of the location dominated by Whites. Integration will increase as both groups will favor the same location. Then, when the Black minority will be large enough, Whites will prefer to move out of this location. A completely integrated equilibrium will thus emerge as a stable outcome as described earlier.

Source dynamics can also arise in this context, but with this structure of preferences, we want to emphasize that all source dynamics might not be desirable. For perceptions of the minority lying in area 2 of Figure 3, the source dynamics is spiraling. When the city is in this situation, neither the completely segregated states nor the integrated ones are stable anymore as one group will chase the other indefinitely in a limit cycle. Imagine that the city starts in the integrated state in which all individuals live in location 2. Then, Whites will move in the deserted location because they actually suffer a lot from the largest minority possible (no matter the color of the minority), and from the largest group of Blacks possible.

On the other hand, they will enjoy living alone in the other location as there is no Black minority at all despite the limited peer presence. Thus the utility differential for Whites is in favor of the deserted location. For Blacks, the deserted location is not interesting as they would be alone only benefiting from the small number of peers while complementarity with a large number of Whites, and a large Black majority is strongly beneficial for them. Thus, the utility differential for Blacks is in favor of the most populated location. At some point, the White minority becomes too small compared to the benefit of a small Black minority living in the all White location. Then both groups will favor the same location until the point where the Black minority will grow too large. The dynamics will continue as described previously but for a different location, hence completing the cycle.

Finally, the saddle dynamics that can occur lead to a segregated state as previously. However, segregation is not complete in this case as a small number of Blacks will live within the location dominated by Whites.

#### 2.4.3 Acting White

Some individuals of the discriminated group may reject their own group and embrace the culture of the dominant group. This phenomenon is known as "acting White" or more generally as the oppositional culture hypothesis. It has been used to explain the different performances of Blacks and Whites in school tests (Ainsworth-Darnell and Downey[1], Austen-Smith and Fryer[4], Fryer and Torelli[31]) or on the job market (Battu et al.[6]) in the United States. This behavior is much less studied in South Africa, only partial evidences could be found in the literature (McKinney[41]). In our framework, this behavior translates into the following parameters set: $a = 10$, $b = -8$, $c = 5$, $d = -7$. Figure 4 portrays the associated bifurcation diagram.

Compared to the previous structures of preferences, a stable completely integrated equilibrium can emerge for a larger subspace of the $(\gamma_B; \gamma_W)$-plan (both with and without spiraling). Usually, completely integrated equilibria will not require a policy intervention in this case. Integrated states will occur also more frequently despite a reduced subspace for source dynamics because some saddle dynamics will lead to integrated states rather than completely segregated states. Moreover saddle dynamics will usually not produce completely segregated states when deviating from an integrated one.

In the first quadrant of the $(\gamma_B; \gamma_W)$-plan, saddle dynamics have almost replaced the source dynamics. For some combinations of perceptions of the minority, the saddle dynamics will lead only to integrated states as for the example illustrated in Figure 4. If the city starts in a highly segregated configuration, Whites benefit a lot from their large majority and their racism is compensated by their positive perception of the minority in the location they dominate. In the other location, they benefit only a little from their small peer presence and consequently gain only little from their complementarity with Blacks, but they suffer a lot from the large Black majority. So the utility differential is in favor of the location dominated by Whites. Besides, Blacks in the location dominated by Whites are benefiting from a large White majority, and both the small number of Blacks and living in the minority weaken the negative effect of a peer presence. In the other location, Blacks endure a large peer presence and benefit only a little from the small White minority. Hence, the utility differential is also in favor of the White dominated location. Then both Blacks and Whites will gather in the same location pushing the city into an integrated state.

Imagine now that Whites are initially close to an equal distribution in the two locations while Blacks are not. As Whites perceive positively the minority, they will move in the location dominated by Blacks because the White minority already living in the location is sufficiently large to compensate the large Black majority. On the other hand, Blacks will prefer to move out of this location at first as they can join a White majority and a small peer presence in the other location. As both groups move, the utility differential increases for Whites but diminishes for Blacks. The White majority is decreasing while the Black minority is increasing while in the other location, the White minority is increasing and the Black majority is shrinking. Thus when the White minority has grown sufficiently large, Blacks come back in their initially dominated location.

From the description of the above example, we can see that saddle dynamics are sensitive to the perceptions of the minority. For instance, let us assume that Blacks have a negative percetion of the minority and the city starts in an integrated state in which location 1 is deserted. Blacks suffer a lot in this location as there is the largest Black community possible. No matter the identity of the minority, Blacks have a negative pay-off in this location while it is actually null in the other location as there is nobody. So Blacks have an incentive to move out of this neighborhood. On the other hand, living with all their peers ensures Whites to have positive pay-off in this location, as the complementarity with Blacks compensate sufficiently their racist preference for the homophillic effect to dominate. Thus, Blacks destroy the integrated state of the city. However, Blacks do not want to live with peers. Consider a completely segregated state, Blacks would have an incentive to move in the location dominated by Whites as they can benefit a lot from the large number of Whites whereas they endure the largest Black community. So there should be a point in between the completely segregated state and the integrated one deserting location 1. The city reaches a stable state when the benefit of the White majority minus the cost of living in a Black minority is equal to the cost of living only with Blacks.

If the city follows a saddle dynamics, no relocation policy would restore an integrated state. The only possible policy would be to act on the perceptions of the minority to bring back the city either in a source dynamics (where a relocation policy can be implemented) or in sink dynamics. In the latter scenario, usually no intervention would be required. The sink dynamics is similar to the one described in the previous structure of preference. The source dynamics (without spiraling) differs only from before in the reduced number of trajectories leading to a segregated state (which will usually not be complete).

## 2.5 Different population sizes

Now, we relax the assumption of equal population sizes. At the city level one group overwhelms the other, say $L^B > L^W$ which is the South African case.[25] The particularity of this situation resides in the apparition of a subset of the set of racial mixes in which the minority group at the global level is also a minority at the local level for all the locations. When populations are equal this situation is impossible as a minority group in one location is, by complementarity, the majority in the other location.

Figure 5 gives a representation of the two types of subsets in which the society can be. The red diamond corresponds to a subset of the city configurations, let us call it $D$, where the minority group at the city level is also a minority at the local level.[26] The two white triangles, let us denote them $E$ and $F$, are the sets where the minority group in one location is the majority group in the other location.

The utility function, and therefore the difference in utility between two locations can be rewritten as:

$$\Delta u^{W1}(B,W) = \begin{cases} 2(a+\gamma_W)W + 2bB - (a+\gamma_W)L^W - bL^B & \forall (\text{B;W}) \in D \\ (2a+\gamma_W)W + (2b+\gamma_W)B - (a+\gamma_W)L^W - bL^B & \forall (\text{B;W}) \in E \\ (2a+\gamma_W)W + (2b+\gamma_W)B - aL^W - (b+\gamma_W)L^B & \forall (\text{B;W}) \in F \end{cases} \tag{10}$$

and

[25] Note that the reverse assumption $L^W > L^B$ is the American situation while the case of equality $L^W = L^B$ describes the Brazilian case.

[26] The red diamond is the set of all points where $min[W;B] = W$ and $min[L^W - W; L^B - B] = L^W - W$.

$$
\Delta u^{B1}(B,W) = \begin{cases} 2(c+\gamma_B)W + 2dB - (c+\gamma_B)L^W - dL^B & \forall \text{ (B;W)} \in D \\ \\ (2c+\gamma_B)W + (2d+\gamma_B)B - (c+\gamma_B)L^W - dL^B & \forall \text{ (B;W)} \in E \\ \\ (2c+\gamma_B)W + (2d+\gamma_B)B - cL^W - (d+\gamma_B)L^B & \forall \text{ (B;W)} \in F \end{cases} \tag{11}
$$

Then the corresponding dynamics is:

$$
W_{t+1} = \begin{cases} \dfrac{2(a+\gamma_W)W_tL^W + 2bB_tL^W - (a+\gamma_W)L^{W^2} - bL^BL^W - \alpha L^W}{\theta} & \forall \text{ (B;W)} \in D \\ \dfrac{(2a+\gamma_W)W_tL^W + (2b+\gamma_W)B_tL^W - (a+\gamma_W)L^{W^2} - bL^BL^W - \alpha L^W}{\theta} & \forall \text{ (B;W)} \in E \\ \dfrac{(2a+\gamma_W)W_tL^W + (2b+\gamma_W)B_tL^W - aL^{W^2} - (b+\gamma_W)L^BL^W - \alpha L^W}{\theta} & \forall \text{ (B;W)} \in F \end{cases} \tag{12}
$$

and

$$
B_{t+1} = \begin{cases} \dfrac{2(c+\gamma_B)W_tL^B + 2dB_tL^B - (c+\gamma_B)L^BL^W - dL^{B^2} - \alpha L^B}{\theta} & \forall \text{ (B;W)} \in D \\ \dfrac{(2c+\gamma_B)W_tL^B + (2d+\gamma_B)B_tL^B - (c+\gamma_B)L^WL^B - dL^{B^2} - \alpha L^B}{\theta} & \forall \text{ (B;W)} \in E \\ \dfrac{(2c+\gamma_B)W_tL^B + (2d+\gamma_B)B_tL^B - cL^WL^B - (d+\gamma_B)L^{B^2} - \alpha L^B}{\theta} & \forall \text{ (B;W)} \in F \end{cases} \tag{13}
$$

These last two sets of equations are important for the empirical part as they can be directly estimated and depending on the set in which the city is, we are able to identify directly all the coefficient and especially our parameters of interest $\gamma_W$ and $\gamma_B$.

# 3 Racial preferences in the Post Apartheid South Africa

## 3.1 Structural estimates

In order to approximate the different subsets of the previous section, we divide the sample in two subsamples. The first subsample is composed by districts in which Whites constitute the majority, whereas the second subsample is composed by districts in which Blacks constitute the majority. Thus, each district is an observation of a random variable corresponding to a location. Then regressing on a particular subsample will gives us the effect for this particular subset. For instance, if we regress on the White-dominated subsample, we are considering that location 1 is dominated by Whites. So location 2 is dominated by Blacks. Therefore, we are currently located in the subspace E. Similarly, if we regress on the Black-dominated subsample, we are considering that location 1 is dominated by Blacks. So location 2 is dominated by Whites. Therefore, we are currently located in the subspace F.[27] Thus, the empirical counterpart of the dynamic equation 12 (for Whites in E) is a linear autoregressive model:

$$\begin{aligned} W_i(t+1) = \delta &+ \beta_1 W_i(t) * L_i^W(t) + \beta_2 B_i(t) * L_i^W(t) + \beta_3 L_i^{W^2}(t) + \beta_4 L_i^W(t) * L_i^B(t) \\ &+ \beta_5 L_i^W(t) + \boldsymbol{\beta' X_i}(t) + \epsilon_{wi}(t+1) \end{aligned} \tag{14}$$

with $\delta$ a constant term, $\boldsymbol{\beta' X_i}(t)$ a set of location specific control variables with their associated coefficient expressing the attractiveness of the locations,[28] and $\epsilon_{wi}$ an idiosyncratic shock that is both location and group specific.

We estimate separately equation 14 for each group and each location by Ordinary Least-Squares. We thus get a set of four equations. To recover the structural parameters of our model, we have both a theoretical and an empirical description. By identification, we get a system of equations between the structural parameters and the reduced form parameters. Unfortunately, the system is overidentified and we have to turn to a least-squares solution.

Recalling equation 12 for Whites in the subspace E, we have that:

[27]Note that being in the subspace E is equivalent to being in the subspace F as E and F are symmetric. Thus it is just a matter of notations and how you define location 1 and location 2. Moreover, the construction of our subsamples insures that we cannot be in the subspace D as each subsample is dominated by a different group. For the rest of the paper, we will adopt the convention that the location 1 is the location dominated by Whites whereas location 2 is the location dominated by Blacks.

[28]Throughout the paper, boldface characters denote vectors and matrices.

$$W_1(t+1) = \frac{(2a+\gamma)W_tL^W + (2b+\gamma)B_tL^W - (a+\gamma)L^{W^2} - bL^BL^W - \alpha L^W}{\theta} \tag{15}$$

Then by identification with equation 14, we get the following system of five equations but four unknowns:

$$\begin{cases} \beta_1 = 2a+\gamma \\ \beta_2 = 2b+\gamma \\ \beta_3 = -(a+\gamma) \\ \beta_4 = -b \\ \beta_5 = -\alpha \end{cases} \tag{16}$$

The least-square solution results from the following minimization programme:

$$\min_{a,b,\gamma,\alpha} (\beta_1 - 2a - \gamma)^2 + (\beta_2 - 2b - \gamma)^2 + (\beta_3 + a + \gamma)^2 + (\beta_4 + b)^2 + (\beta_5 + \alpha)^2 \tag{17}$$

Then we derive analytically the expression of an estimator of our structural parameters using the first order conditions of the minimization programme. We obtain that:

$$\begin{cases} \hat{a} = \dfrac{7\beta_1 - 3\beta_2 + 4\beta_3 - 6\beta_4}{10} \\ \\ \hat{b} = \dfrac{2\beta_1 + 2\beta_2 + 4\beta_3 + -6\beta_4}{10} \\ \\ \hat{\gamma} = \dfrac{\beta_2 - \beta_1}{2} + \beta_4 - \beta_3 \\ \\ \hat{\alpha} = -\beta_5 \end{cases} \tag{18}$$

Finally, we recover the standard errors of all the estimators of the structural parameters by the delta method.

## 3.2 Data description

We exploit the data coming from the Community Profiles associated with the census waves conducted in South Africa between 1996 and 2011. Community Profiles are cross-tabulations of the full count aggregated by geographic areas. They aim to guide the action of local public authorities. They are available up to the enumeration area level for the 1996 and 2011 censuses, and to the subplace level for the 2001 Census. As our statistical unit is a geographic subdivision, we are facing two problems. First, we would like to have the largest sample size to conduct a statistical analysis. Second, as segregation measures are sensitive to changes in boundaries, we would like to have the most stable geographic layer. Unfortunately, over the 1996-2011 period, no geographic layer remained unchanged. Thus, we have chosen to work at the subplace level adjusted to the 2001 boundaries. To adjust the data, we use the freeze history approach.[29] The overlap between the "source" and the "target" polygons serves as weight to adjust the data of the "source" layer to the targeted layer. See Appendix "X" for more details. This procedure leads to a sample size of 21243 subplaces in the three Census waves. Moreover, the subplace has a concrete meaning for individuals as it is the broad area by which they locate their living place in a city. Real estate agencies also use this layer for theirs advertisements.

Our dependent variable should be the number of Whites[30] in a subplace at a particular census wave. However, we use instead the share of Whites in a subplace to avoid the effect of population size disparity between subplaces. We add 1 to this share and take the natural logarithm. The addition is made to avoid problems of existence of the natural logarithm for subplaces with one population group missing. The natural logarithm allow us to interprete the estimated coefficients as elasticities.

The main independent variables are the number of Whites and Blacks in a subplace at the previous census waves. We also apply the same transformations as the dependent variable. We interact them by a measure of the total size of the group at the province level. This measure is the natural logarithm of one plus the share of Whites at the province level. This variable also appears on his own and squared.

The set of control variables is composed by subplace specific variables of the basic socioeconomic variables such as the mean age, the mean income level, the unemployment rate, the mean number of years of education. They are either measured as the natural logarithm of these variables, or of one plus the share if their is a problem of existence as before. More details about the construction of these variables can be found in the appendices.

[29] See [43] for a more detailed description.

[30] We detail data construction only for Whites in the text for the sake of brevity. But the same transformations apply also for Blacks.

## 3.3 Endogeneity and heteroscedasticity issues

### 3.3.1 Heteroscedasticity issues

As we are interested in inference on structural parameters, we are concerned by heteroscedasticity issues. This is a very common feature of microeconomic databases. Our dataset does not avoid this problem.[31] We then use the heteroscedastic-robust variance-covariance matrix of White[53] as a correction.

### 3.3.2 Endogeneity issues

Our model is also subject to endogeneity issues for several reasons. First, in autoregressive models, if you have long memory in the error terms, then the error term at a period $t$ is correlated to the autoregressive regressor because the latter is correlated with the error term at the previous period $t-1$. In our context, forced displacements during the Apartheid can still possibly explain part of the current racial composition. Second, we might have omitted important factors. For instance, discriminatory practices in the housing or mortgage markets might have an impact on the racial composition. Thus discriminatory practices would be correlated with the number of Whites in the previous period and the error term at $t$ because of the autoregressive structure. Finally, there is a measurement error problem as censuses usually suffer from under- or over count. According to Statistics South Africa, undercount might be due to lack of accessibility which is correlated with race.

Following Kasy[37], we construct instruments using the spatial structure of the White population. We average the number of Whites in contiguous subplaces using queen contiguity of order 1,2, and 3. On the one hand, if a subplace is dominated by Whites, neighboring subplaces are more likely to be populated by Whites as individuals exert homophillic behaviors. On the other hand, discriminatory practices in a subplace are less likely to be correlated with the number of Whites in adjacent subplaces. Kasy[37] argues that neighboring areas 3 kms away from the origin are likely to be valid instruments. In this regard, we exclude the first ring of neighboring subplaces and use only 2nd and 3rd order contiguous subplaces. We estimate the model with GMM.

We test for endogeneity using the augmented regression approach of Wu[55] as the assumption of homoscedasticity of the Hausman's test is not verified here.[32] Moreover, we test overidentifying conditions as we have four instruments for two endogenous variables. Results of the Hansen's test[36] is provided in table 6.

[31] We provide the results of the Breusch-Pagan test[13] in the appendix (table 7). The homoscedasticity assumption is rejected in all cases.

[32] See Table 7 for heteroscedasticity. See Table 5 for tests of endogeneity.

# 4 Discussion

The first observation we can make about our structural estimates is that we find strong homophillic tastes for both Blacks and Whites. It is always stronger for Whites than their taste for Blacks. It is almost always true for Blacks also. However, Whites always have a positive taste for Blacks which is more surprising considering the strong racist rhetoric of the White government during the Apartheid years. Nevertheless what we observe is a mean effect, and a possible explanation is that Whites racists may have been marginalized through time and the efforts of reconciliation made by Mandela. For Blacks, the evidences concerning their taste for Whites are more ambiguous. When we estimate separately each group and location (table 4, columns 1-4), they express some aversion for Whites as c is always negative. But the magnitude of this effect varies a lot. When we switch to joint estimations, the aversion is replaced by a positive taste for Whites with again a lot of variation. All these elements may indicate that the definitive effect is probably small and require more informations to be precisely estimated.

When we turn to the perception of the minority coefficient, again we have some differences between the separate estimations and the joint estimations. In the former case, Whites living in Whites dominated districts express a distaste for living in the minority while Blacks living in Black dominated districts tend to like the presence of a White minority. Table 18 provides more evidences in this direction. Whites generally reject living in the minority while accepting the presence of Blacks. They even dislike more living in the minority than they like the presence of Blacks which may explain their reluctance to integrate. In the same time, Blacks are prone to live in the minority, and it is almost always stronger than their taste for Whites. This should act as a strength of the integration of Blacks. When we look at the joint estimations, the evidences point out a distaste for diversity, even if in the Black dominated location minority is positively viewed. Whites even tends to integrate more than Blacks as they prefer more the presence of Blacks than they dislike living in the minority. In fact, separate and joint estimations go in the opposite direction. However, test about the equality of the gamma coefficients seems to give credits to the interpretations with group specific values of gamma. Thus we should turn to a model considering this difference. Nevertheless, the discrepancy between the two hypotheses about gamma does not alter the results of more integration. It just impact who is integrating with whom. When both group perceives the minority the same way, Whites tends to integrate with Blacks while the reverse occurs if each group has a specific perception of the minority.

# 5 Appendix

## 5.1 Proof of Proposition 1

*Proof.* Consider a simplex $S$ defined by $N_i^k \geq 0$ and $\sum_k \sum_i N_i^k = L > 0$. Let us first study the differentiability of the function $\Delta u^{ki}(W,B)$. As the two populations are equal, we can explicit the behavior of the min term. If $W < B$ then $min[W,B] = W$ and $min[L^W - W, L^B - B] = L^B - B$. If $W > B$ then $min[W,B] = B$ and $min[L^W - W, L^B - B] = L^W - W$. If $W = B$ then $min[W,B] = W$ and $min[L^W - W, L^B - B] = L^W - W$ by convention. Then by simple algebra, we can deduce that the function $\Delta u^{ki}$ can be expressed finally for Whites as :

$$\Delta u^{Wi}(W,B) = \begin{cases} (2a+\gamma_W)W + (2b+\gamma_W)B - L(a+b+\gamma_W) & \text{if } W \neq B \\ \\ (a+b+\gamma_W)(2W-L) & \text{otherwise} \end{cases} \tag{19}$$

Then the function $\Delta u^{ki}(W,B)$ is differentiable on a domain if both the partial derivatives exists and if it has a total differential in each point of its domain. First let us look at the case $W = B$. Thus, $\Delta u^{ki}(W,B)$ reduces to a function of a single variable and we can easily check that $\Delta u^{ki}(W,B)$ is effectively differentiable. When $W \neq B$ we can easily see that the two partial derivatives exists, and, with a bit of algebra, that for an arbitrary $(W_0, B_0)$ with $W \neq B, W_0 \neq B_0$ :

$$\lim_{\substack{(W,B)\to(W_0,B_0) \\ W \neq B}} \frac{\Delta u^{ki}(W,B) - \Delta u^{ki}(W_0,B_0) - [\frac{\partial \Delta u^{ki}}{\partial W}(W_0,B_0)](W-W_0) - [\frac{\partial \Delta u^{ki}}{\partial B}(W_0,B_0)](B-B_0)}{|(W-W_0)^2 + (B-B_0)^2|} = 0 \tag{20}$$

Then $\Delta u^{ki}(W,B)$ is differentiable for all $W \neq B$ and *in fine* differentiable for all $(W,B)$. Then $P^{ki}L^k$ is a continuous function which maps from $S$ (which is a convex and compact set) into itself. Hence, the existence of a fixed point $N_i^{k*} \geq 0$ such that $N_i^{k*} = P^{ki}(u^{ki}(N_i^{k*}, N_i^{-k*}), u^{kj}(N_j^{k*}, N_j^{-k*}))L^k, \quad \forall k \in K, \forall i,j \in I$ is ensured by Brouwer's fixed point theorem. $\square$

## 5.2 Proof of Proposition 2

*Proof.* Let us first provide the conditions for uniquenes and the ones for stability, then let us show that uniqueness implies stability, and finally that stability implies uniqueness. Define two vectors

$$N \equiv (N_1^W, N_2^W, N_1^B, N_2^B),$$
$$f \equiv (f^{1W}, f^{2W}, f^{1B}, f^{2B}) \tag{21}$$

with $f^{ki} \equiv N_i^k - P^{ki}(.)L^k \quad \forall k \in K$. Hence, solving the system $f(N) = 0$ gives us the equilibrium. As shown in the proof of proposition 1, the function $\Delta u^{ki}(W, B)$ is differentiable which implies that $f$ is a differentiable mapping from $\Omega$ into $\mathbb{R}^4$, with $\Omega$ a closed rectangular region $\Omega = \{N|0 \leq N_i^k \leq L^k\}$. The Jacobian matrix is thus :

$$J_f = \frac{1}{\theta}\begin{pmatrix} \theta - (2a+\gamma_W)L & -(2b+\gamma_W)L \\ -(2c+\gamma_B)L & \theta - (2d+\gamma_B)L \end{pmatrix} \tag{22}$$

if $W \neq B$, and is equal to diag$\{\theta - 2(a+b+\gamma_W)L,\ \theta - 2(d+c+\gamma_B)L\}$ otherwise. Then according to the theorem 4 of Gale and Nikaidô[32], the mapping $f$ is univalent if the Jacobian matrix is a P-matrix (*i.e.* a matrix with all its principal minors positive). Thus in our case, as $\frac{1}{\theta} > 0$,[33] we have the following sufficient conditions :

$$\begin{cases} \theta > (2a+\gamma_W)L, \\ \theta > (2d+\gamma_B)L, \\ (\theta - (2a+\gamma_W)L)(\theta - (2d+\gamma_B)L) > (2c+\gamma_B)(2b+\gamma_W)L^2. \end{cases} \tag{23}$$

If this conditions are satisfied, the uniqueness of the equilibrium is implied by the univalence of the mapping $f$.

Let us now examine the stability conditions considering the dynamic adjustment process in equation (9). We can remark that the right-hand side of the dynamic system is equal to $-f(N)$ leading to the same Jacobian matrix multiplied by $-1$ :

$$J_f = \frac{1}{\theta}\begin{pmatrix} (2a+\gamma_W)L - \theta & (2b+\gamma_W)L \\ (2c+\gamma_B)L & (2d+\gamma_B)L - \theta \end{pmatrix}. \tag{24}$$

Then by classical arguments, the equilibrium is stable if the two eigenvalues of our system have a negative real part which can be viewed by conditions on the trace and the determinant :

$$\begin{cases} 2(a+d)L + (\gamma_W + \gamma_B)L - 2\theta < 0, \\ ((2a+\gamma_W)L - \theta)((2d+\gamma_B)L - \theta) - (2c+\gamma_B)(2b+\gamma_W)L^2 > 0, \end{cases} \tag{25}$$

[33] Because of the assumption $\alpha < 0$ and $\beta > 0$.

which is easily seen by simple algebra to be the same conditions as for the uniqueness which completes the proof.

□

### 5.3 Proof of Proposition 3

*Proof.* We solve the equation $Tr_{J_f} = 0$ for $\gamma_W$ as a function of $\gamma_B$ which leads to

$$\gamma_W = \frac{\theta}{L} - 2(a+d) - \gamma_B \tag{26}$$

Then solving $|J_f| = 0$ for $\gamma_W$ as a function of $\gamma_B$ leads to

$$\gamma_W = \frac{\gamma_B[((2b-2a)L+\theta)L] + 4(cb-ad)L^2 + ((2a+2d)L-\theta)\theta}{((2d-2c)L-\theta)L} \tag{27}$$

These two lines are defined on $\mathbb{R}$. We can also characterize when the system will oscillate by solving $Tr^2_{J_f} - 4|J_f| = 0$ for $\gamma_W$ as a function of $\gamma_B$. This gives us the following solutions:

$$\gamma_W = -2a - 4c + 2d - 3\gamma_B \pm \sqrt{S} \tag{28}$$

with

$$\begin{aligned} S \equiv\ & (4c-2d)^2 + a(4+16c-8d) - (2(a+d)L-2\theta)^2 - 8aL\theta - 8dL\theta - 4\theta^2 \\ & + 16adL^2 - 16cdL^2 + \gamma_B(16a-8b+24c-16d) + 8\gamma_B^2 \end{aligned} \tag{29}$$

Then a sink will be characterized by a negative trace and a positive determinant. These conditions are fullfilled somewhere in $\mathbb{R}^2$ for a certain couple $(\gamma_B^*, \gamma_W^*)$ if the two lines (equations (26) and (27)) intersect. But they are also fullfilled if the two lines are parallel with different intercepts. They might also been fullfilled when the two lines are the same but this case is not interesting as it will become clear by the end of the proof.

Then the two lines are parallel if they have the same slope coefficient. Thus, let us solve the following equation:

$$\frac{((2b-2a)L+\theta)L}{((2d-2c)L-\theta)L} = -1 \tag{30}$$

Note that we need to have $2(d-c) \neq \frac{\theta}{L}$ in order for the line (27) to exist. After computations, we obtain:

$$b - a = c - d \tag{31}$$

Then if this condition is not satisfied, the two lines where the trace and the determinant are null are not parallel. Thus they intersect somewhere in the $(\gamma_B, \gamma_W)$ plan.

Now, if they are parallel, they are mingled if they have the same intercept. Thus, let us solve the following equation:

$$\frac{\theta}{L} - 2(a+d) = \frac{4(cb - ad)L^2 + ((2a+2d)L - \theta)\theta}{((2d-2c)L - \theta)L} \tag{32}$$

After computations, we get:

$$(2d - 2c)L\theta + 2c(2a - 2b)L^2 + 2d(2c - 2d)L^2 = 0 \tag{33}$$

From this equation, we see that if $a = b$ and $c = d$ then the two lines are the same as they will have the same slope and intercept coefficients. However this situation means that both groups like (or dislike) equally the two groups. This is not really credible as people have often a homophillic bias in their interactions (Currarini et al.[22], and Skvoretz[51]). Moreover, antisocial behaviors will be punished in environment where you have frequent or long lasting interactions(Bowles[10]). In the context of segregation, you are precisely in a situation where you are going to have these kinds of interactions with your neighbors. Thus antisocial behaviors will not characterize a substantive share of individuals in the population.

Now, if $a \neq b$ and $c \neq d$, we can rewrite the previous equation as:

$$\begin{aligned} & L(2d-2c)(\theta - 2dL) + 2c(2a-2b)L^2 = 0 \\ \Leftrightarrow & (2d-2c)(\theta - 2dL) = -2c(2a-2b)L \\ \Leftrightarrow & \theta - 2dL = \frac{-2c(2a-2b)L}{2d-2c} \\ \Leftrightarrow & \frac{\theta}{L} = 2(d-c) \end{aligned} \tag{34}$$

The last equation comes from the assumption that the two lines are parallel (*i.e.* $b - a = c - d$). However this condition implies that the line (27) does not exist.

So if we exclude unlikely structures of preferences (*i.e.* $a = b$ and $c = d$), and if the line (27) exists, then the line (26) and (27) either intersect once or are parallel with different intercepts. Then there exists parameters regions $(\gamma_B^*; \gamma_W^*)$ where the trace will be positive while the determinant will be negative. Thus, the equilibrium will be a sink, integrated and stable.

Moreover, there also exists parameters regions $(\gamma_B^P; \gamma_W^P)$ where the trace and the determinant will be both positive. Thus the system will be dynamically a source. In this case, the equilibrium will not be stable anymore but the system will end in one of the corner of the edgeworth box depicting our city. Then if the city is in the basin of attraction of an integrated state,[34] there is no need for a public policy as the system will converge by itself toward an integrated state. On the contrary, if the city is in the basin of attraction of a completely segregated state, then a relocation policy that will displace a sufficient number of members of the local majority to the other location will replace the city on an integration path. Hence, the proof is complete.

□

## 5.4 Analytic solution of the first-order differential system

*Proof.* Recalling the system (9):

$$\begin{cases} \dot{W} = \dfrac{((2a+\gamma)L-\theta)W_t + (2b+\gamma)LB_t - L^2(a+b+\gamma) - \alpha L}{\theta} \\ \\ \dot{B} = \dfrac{(2c+\gamma)LW_t + ((2d+\gamma)L-\theta)B_t - L^2(c+d+\gamma) - \alpha L}{\theta} \end{cases} \tag{35}$$

we can rewrite it in a more tractable form :

$$\begin{cases} \dot{W} = AW_t + PB_t - K_w \\ \\ \dot{B} = CW_t + DB_t - K_b \end{cases} \tag{36}$$

with $A \equiv \dfrac{(2a+\gamma)L-\theta}{\theta}$, $P \equiv \dfrac{(2b+\gamma)L}{\theta}$, $C \equiv \dfrac{(2c+\gamma)L}{\theta}$, $D \equiv \dfrac{(2d+\gamma)L-\theta}{\theta}$, $K_w \equiv \dfrac{-L^2(a+b+\gamma)-\alpha L}{\theta}$, $K_b \equiv \dfrac{-L^2(c+d+\gamma)-\alpha L}{\theta}$.

Then we can get the following system by expressing $B_t$ as a function of $W_t$ and its differentials :

$$\begin{cases} B_t = \dfrac{\dot{W} - AW_t - K_w}{P} \\ \\ \dot{B} = CW_t + DB_t - K_b \end{cases} \tag{37}$$

Then deriving an expression of $\dot{B}$ from this first equation :

[34] See definition 3 for the precise meaning of what we call a state.

$$\dot{B} = \frac{\ddot{W} - A\dot{W}}{P} \tag{38}$$

We can then rewrite the second equation of the system (25) as :

$$\frac{\ddot{W} - A\dot{W}}{P} = CW + D\frac{\dot{W} - AW_t - K_w}{P} + K_b \tag{39}$$

which can be rearranged as :

$$\ddot{W} - \frac{(A+D)}{P}\dot{W} + (A - \frac{C}{P})W_t = K_b - \frac{D}{P}K_w \tag{40}$$

Then we solve first the homogeneous equation related to equation (28):

$$\ddot{W} - \frac{(A+D)}{P}\dot{W} + (A - \frac{C}{P})W_t = 0 \tag{41}$$

which has the characteristic equation:

$$r^2 - \frac{(A+D)}{P}r + (A - \frac{C}{P}) = 0 \tag{42}$$

with $r$ a generic term. Then depending on the sign of the discriminant of equation (30), we get the following general solutions denoted by the superscript g :

$$\begin{cases} \text{If } \Delta > 0, & \text{Then } W_t^g = k_1 e^{r_1 t} + k_2 e^{r_2 t} & \text{with } r_1, r_2 = \dfrac{\frac{A+D}{P} \pm \sqrt{\Delta}}{2} \\ \text{If } \Delta = 0, & \text{Then } W_t^g = k_1 e^{rt} + k_2 t e^{rt} & \text{with } r = \dfrac{A+D}{2P} \\ \text{If } \Delta < 0, & \text{Then } W_t^g = e^{\xi t}(k_1 cos\varphi t + k_2 sin\varphi t) & \text{with } \xi = \dfrac{A+D}{2P} \text{ and } \varphi = \dfrac{\sqrt{\Delta}}{2} \end{cases} \tag{43}$$

Then for a particular solution (denoted by the superscript p), as the forcing term is a constant, let us assume that y(t) is a constant, then the first differential is null while the second differential does not exist which implies that a particular solution for $W_t$ is :

$$W_t^p = \frac{PK_b - DK_w}{PA - C} \tag{44}$$

Then as the solution of a differential equation is the sum of a general solution and a particular solution, we know the form of the solution for $W_t$. Therefore we can compute the solution for $B_t$. First, if $\Delta > 0$, then we have :

$$W_t^* = k_1 e^{r_1 t} + k_2 e^{r_2 t} + \frac{PK_b - DK_w}{PA - C} \Longleftrightarrow \dot{W}^* = r_1 k_1 e^{r_1 t} + r_2 k_2 e^{r_2 t} \tag{45}$$

Then substituting into the first equation of (25), we get:

$$B_t^* = (r_1 - A) k_1 e^{r_1 t} + (r_2 - A) k_2 e^{r_2 t} - A \frac{PK_b - DK_w}{PA - C} \tag{46}$$

$\square$

## 5.5 Details of the variables

Table 3: Reduced form estimates

- **N2L5** represents the interaction between N2 and L5 which are respectively $ln(1+B)$ and $ln(1+L^W)$, with $B$ the share of Blacks in a district in 1996, and $L^W$ is the share of Whites in a province in 1996. It is equivalent to the term $B_t L^W$ in the theoretical model.

- **N5L5** represents the interaction between N5 and L5 which are respectively $ln(1+W)$ and $ln(1+L^W)$, with $W$ the share of Whites in a district in 1996, and $L^W$ is the share of Whites in a province in 1996. It is equivalent to the term $W_t L^W$ in the theoretical model.

- **L2L5** represents the interaction between L2 and L5 which are respectively $ln(1+L^B)$ and $ln(1+L^W)$, with $L^B$ the share of Blacks in a province in 1996, and $L^W$ is the share of Whites in a province in 1996. It is equivalent to the term $L^B L^W$ in the theoretical model.

- **L5L5** represents the interaction between L5 and L5 which is $ln(1+L^W)$, with $L^W$ the share of Whites in a province in 1996. It is equivalent to the term $(L^W)^2$ in the theoretical model.

- **L5** represents $ln(1+L^W)$ with $L^W$ the share of Whites in a province in 1996. It is equivalent to the term $L^W$ in the theoretical model.

- **N2L2** represents the interaction between N2 and L2 which are respectively $ln(1+B)$ and $ln(1+L^B)$, with $B$ the share of Blacks in a district in 1996, and $L^B$ is the share of Blacks in a province in 1996. It is equivalent to the term $B_t L^B$ in the theoretical model.

- **N5L2** represents the interaction between N5 and L2 which are respectively $ln(1+W)$ and $ln(1+L^B)$, with $W$ the share of Whites in a district in 1996, and $L^B$ is the share of Blacks in a province in 1996. It is equivalent to the term $W_t L^B$ in the theoretical model.

- **L5L2** represents the interaction between L5 and L2 which are respectively $ln(1+L^W)$ and $ln(1+L^B)$, with $L^B$ the share of Blacks in a province in 1996, and $L^W$ is the share of Whites in a province in 1996. It is equivalent to the term $L^B L^W$ in the theoretical model.

- **L2L2** represents the interaction between L2 and L2 which is $ln(1+L^B)$, with $L^B$ the share of Blacks in a province in 1996. It is equivalent to the term $(L^B)^2$ in the theoretical model.

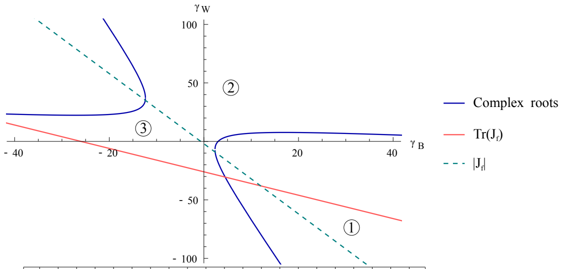
- **L2** represents $ln(1+L^B)$ with $L^B$ the share of Blacks in a province in 1996. It is equivalent to the term $L^B$ in the theoretical model.

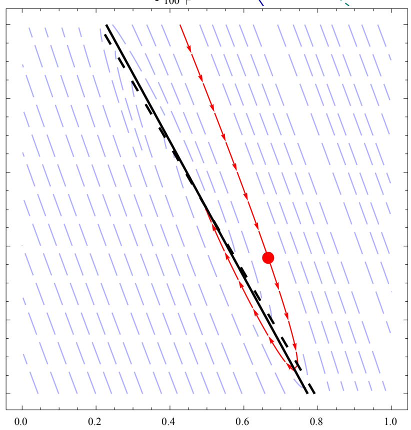
- **Mean years of education (1996)** represents the natural logarithm of the mean number of years of schooling in a subplace in 1996.

- **Mean age (1996)** represents the natural logarithm of the mean age in a subplace in 1996.

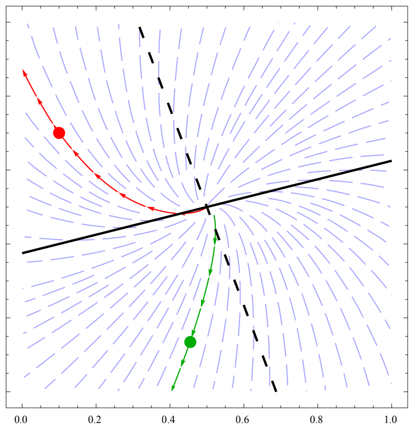
- **Mean income (1996)** represents the natural logarithm of the mean income in a subplace in 1996. The income is computed as the center of the class in which the individual has declared to be. Incomes of other year are deflated to 1996 Rands.

- **Unemployment rate (1996)** represents the natural logarithm of 1 plus the unemployment rate in a subplace in 1996. The unemployment rate is computed as the number of individuals aged 15 or older declaring that they are unemployed and looking for a job over the sum of the individuals aged 15 or older currently employed and of the individuals declaring being unemployed.

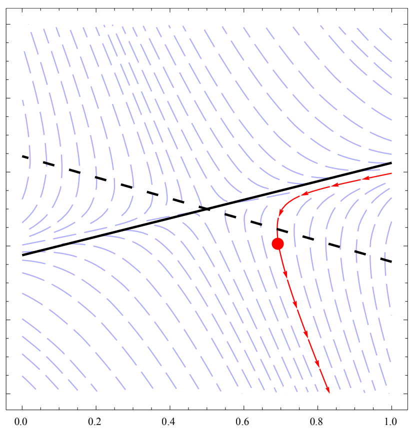


①: Sink

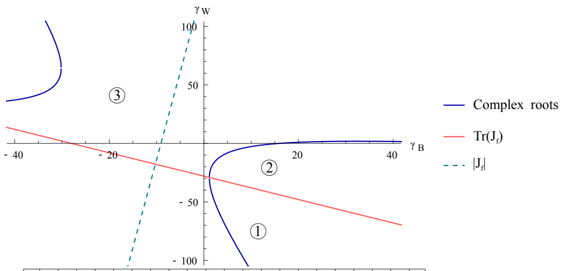


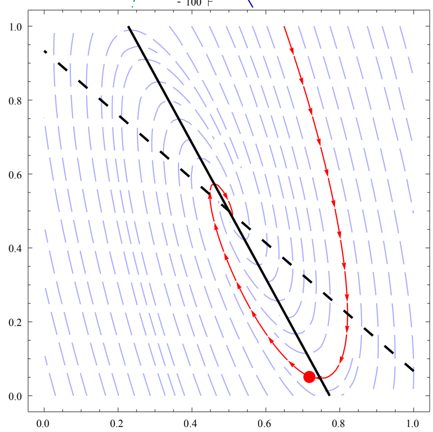
②: Source



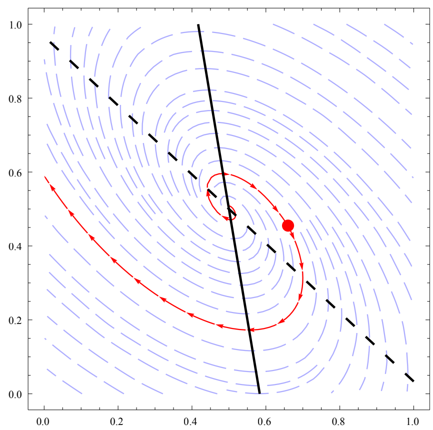
③: Saddle

Figure 2: Mutual reject

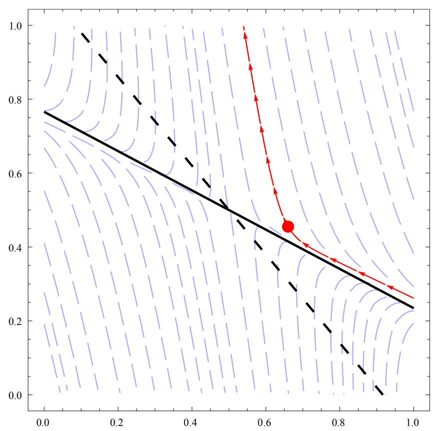


①: Sink



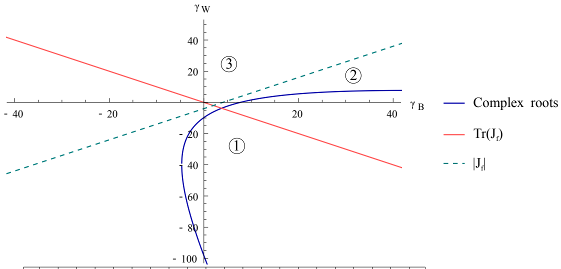
②: Source

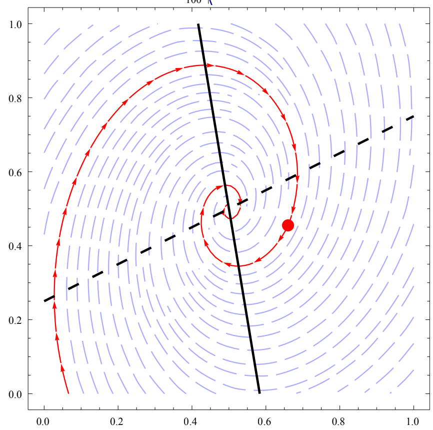


③: Saddle

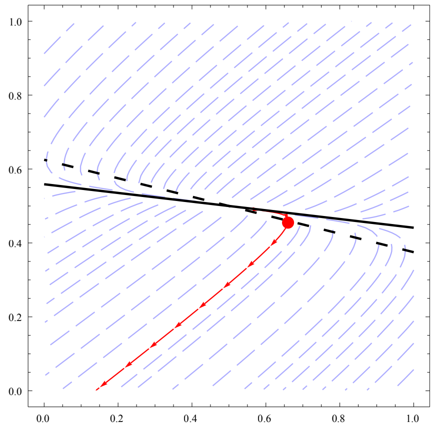


Figure 3: White segregation vs Black integration

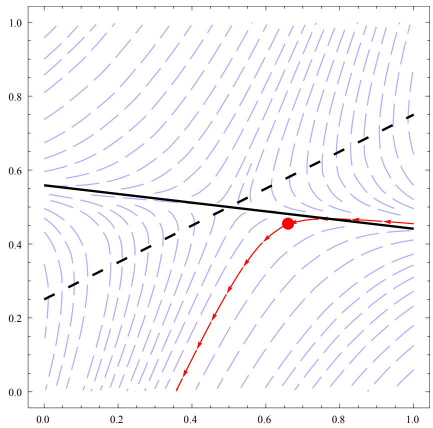


①: Sink



②: Source



③: Saddle

Figure 4: Acting White

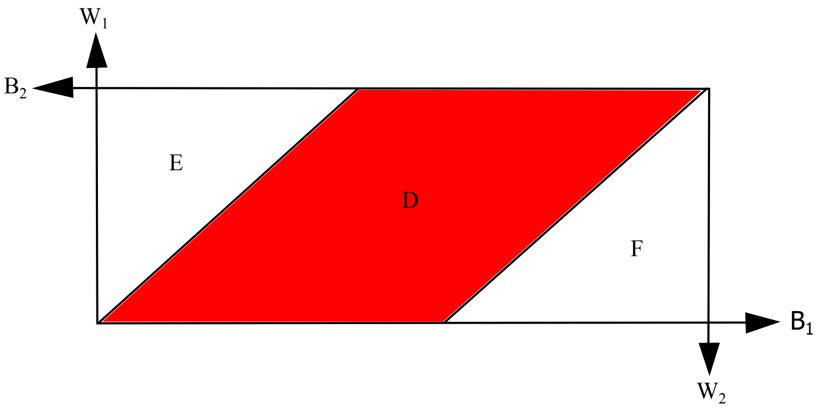



Figure 5: Representation of the different subsets of the city.

Table 1: OLS First stage estimation 2001

| | OLS 2001 | | | |
|---|---|---|---|---|
| | Blacks (E) | Whites (E) | Blacks (F) | Whites (F) |
| 1996 L_2_ | 0.583*** [0.213] | | 1.265*** [0.089] | |
| 1996 N_2L_2_ | 1.801*** [0.071] | | 1.504*** [0.022] | |
| 1996 L_2L_2_ | -0.571** [0.266] | | -1.868*** [0.087] | |
| 1996 L_5L_2_ | 0.296** [0.132] | -0.140 [0.133] | 0.026 [0.050] | -0.300*** [0.080] |
| 1996 N_5L_2_ | -0.134 [0.082] | | 0.173*** [0.039] | |
| Mean age (1996) | -0.120*** [0.017] | 0.128*** [0.018] | -0.019*** [0.006] | 0.045*** [0.006] |
| Mean years of education (1996) | 0.186*** [0.026] | -0.025 [0.026] | 0.004*** [0.001] | -0.008*** [0.001] |
| Unemployment rate (1996) | -0.093 [0.058] | -0.155** [0.063] | 0.021*** [0.005] | -0.054*** [0.004] |
| Mean income (1996) | -0.067*** [0.008] | 0.060*** [0.009] | 0.002*** [0.001] | 0.004*** [0.001] |
| 1996 L_5_ | | -1.709*** [0.407] | | 0.283*** [0.063] |
| 1996 N_2L_5_ | | -1.895*** [0.222] | | 0.117*** [0.035] |
| 1996 N_5L_5_ | | 4.015*** [0.145] | | 4.268*** [0.167] |
| 1996 L_5L_5_ | | -0.413 [1.503] | | -1.626*** [0.237] |
| Observations | 3143 | 3143 | 17258 | 17258 |
| $R^2$ | 0.531 | 0.540 | 0.794 | 0.361 |

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: OLS First stage estimation 2011

| | OLS 2001 | | | |
|---|---|---|---|---|
| | Blacks (E) | Whites (E) | Blacks (F) | Whites (F) |
| 2001 L_2_ | 0.330 [0.278] | | 0.793*** [0.099] | |
| 2001 N_2L_2_ | 1.574*** [0.057] | | 1.443*** [0.017] | |
| 2001 L_2L_2_ | -0.479 [0.326] | | -1.311*** [0.100] | |
| 2001 L_5L_2_ | -0.051 [0.139] | 0.064 [0.145] | -0.255*** [0.065] | -0.301*** [0.072] |
| 2001 N_5L_2_ | 0.098 [0.062] | | 0.119*** [0.033] | |
| Mean age (2001) | -0.089*** [0.018] | 0.140*** [0.017] | -0.018*** [0.007] | 0.060*** [0.006] |
| Mean years of education (2001) | 0.144*** [0.026] | -0.094*** [0.024] | 0.010*** [0.003] | -0.014*** [0.002] |
| Unemployment rate (2001) | 0.159*** [0.050] | -0.355*** [0.048] | 0.051*** [0.005] | -0.082*** [0.004] |
| Mean income (2001) | -0.040*** [0.006] | 0.032*** [0.006] | 0.001 [0.001] | 0.003*** [0.001] |
| 2001 L_5_ | | -3.288*** [0.470] | | 0.185*** [0.066] |
| 2001 N_2L_5_ | | -0.731*** [0.219] | | 0.358*** [0.037] |
| 2001 N_5L_5_ | | 4.759*** [0.147] | | 5.321*** [0.153] |
| 2001 L_5L_5_ | | 3.910* [2.026] | | -1.546*** [0.291] |
| Observations | 2555 | 2555 | 15580 | 15580 |
| $R^2$ | 0.557 | 0.626 | 0.710 | 0.462 |

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: GMM First stage estimation 2001

| | GMM 2001 | | | |
|---|---|---|---|---|
| | Blacks (E) | Whites (E) | Blacks (F) | Whites (F) |
| 1996 N_2L_2_ | 3.074*** [0.189] | | 1.582*** [0.030] | |
| 1996 N_5L_2_ | -0.775** [0.301] | | 0.287** [0.112] | |
| 1996 L_2L_2_ | -1.183*** [0.339] | | -1.912*** [0.089] | |
| 1996 L_5L_2_ | -0.259 [0.189] | 0.049 [0.284] | -0.029 [0.050] | -0.380*** [0.140] |
| 1996 L_2_ | 1.182*** [0.274] | | 1.235*** [0.090] | |
| Mean age (1996) | -0.084*** [0.024] | 0.064*** [0.024] | -0.020** [0.009] | 0.046*** [0.007] |
| Mean years of education (1996) | 0.402*** [0.064] | -0.271*** [0.065] | 0.005*** [0.002] | -0.009*** [0.001] |
| Unemployment rate (1996) | -0.344*** [0.104] | 0.045 [0.083] | 0.019** [0.008] | -0.057*** [0.005] |
| Mean income (1996) | -0.102*** [0.013] | 0.066*** [0.012] | 0.002*** [0.001] | 0.004*** [0.001] |
| 1996 N_2L_5_ | | -3.622*** [0.683] | | 0.204** [0.103] |
| 1996 N_5L_5_ | | 6.750*** [0.650] | | 4.076*** [0.269] |
| 1996 L_5_ | | -3.819*** [0.655] | | 0.291*** [0.065] |
| 1996 L_5L_5_ | | 2.939* [1.775] | | -1.635*** [0.246] |
| Observations | 3143 | 3143 | 17258 | 17258 |
| $R^2$ | 0.341 | 0.455 | 0.792 | 0.360 |

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: GMM First stage estimation 2011

| | GMM 2001 | | | |
|---|---|---|---|---|
| | Blacks (E) | Whites (E) | Blacks (F) | Whites (F) |
| 2001 N_2L_2_ | 2.409*** [0.132] | | 1.681*** [0.037] | |
| 2001 N_5L_2_ | -0.661*** [0.225] | | -0.302** [0.128] | |
| 2001 L_2L_2_ | -0.906** [0.388] | | -1.582*** [0.104] | |
| 2001 L_5L_2_ | -0.415** [0.174] | 0.531 [0.328] | -0.335*** [0.070] | -0.606*** [0.195] |
| 2001 L_2_ | 0.916*** [0.332] | | 0.858*** [0.103] | |
| Mean age (2001) | 0.012 [0.028] | 0.065*** [0.024] | 0.014 [0.009] | 0.048*** [0.006] |
| Mean years of education (2001) | 0.300*** [0.051] | -0.277*** [0.050] | 0.012*** [0.003] | -0.011*** [0.002] |
| Unemployment rate (2001) | -0.105 [0.076] | -0.229*** [0.063] | 0.002 [0.009] | -0.066*** [0.005] |
| Mean income (2001) | -0.056*** [0.008] | 0.036*** [0.007] | 0.007*** [0.001] | 0.001 [0.001] |
| 2001 N_2L_5_ | | -2.312*** [0.699] | | 0.404*** [0.141] |
| 2001 N_5L_5_ | | 7.145*** [0.510] | | 7.089*** [0.366] |
| 2001 L_5_ | | -5.030*** [0.628] | | 0.390*** [0.074] |
| 2001 L_5L_5_ | | 6.434*** [2.246] | | -2.428*** [0.311] |
| Observations | 2555 | 2555 | 15580 | 15580 |
| $R^2$ | 0.370 | 0.552 | 0.693 | 0.436 |

Standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Structural parameters after an OLS first stage

| | OLS 2001 | | | | OLS 2011 | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) Blacks (E) | (2) Whites (E) | (3) Blacks (F) | (4) Whites (F) | (5) Blacks (E) | (6) Whites (E) | (7) Blacks (F) | (8) Whites (F) |
| a | | 3.298*** | | 1.732*** | | 5.076*** | | 1.943*** |
| | | (0.659) | | (0.148) | | (0.861) | | (0.178) |
| b | | 0.343 | | -0.343** | | 2.331*** | | -0.538*** |
| | | (0.647) | | (0.143) | | (0.854) | | (0.172) |
| c | -0.173 | | -0.428*** | | -0.137 | | -0.059 | |
| | (0.135) | | (0.032) | | (0.156) | | (0.039) | |
| d | 0.794*** | | 0.238*** | | 0.601*** | | 0.603*** | |
| | (0.128) | | (0.033) | | (0.155) | | (0.037) | |
| gamma | 0.100 | -2.682* | 1.229*** | 0.750*** | 0.311 | -6.590*** | 0.394*** | 1.236*** |
| | (0.183) | (1.568) | (0.068) | (0.246) | (0.218) | (2.084) | (0.080) | (0.288) |

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Structural parameters after a GMM first stage

| | GMM 2001 | | | | GMM 2011 | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) Blacks (E) | (2) Whites (E) | (3) Blacks (F) | (4) Whites (F) | (5) Blacks (E) | (6) Whites (E) | (7) Blacks (F) | (8) Whites (F) |
| a | | 6.957*** | | 1.685*** | | 7.950*** | | 2.713*** |
| | | (1.096) | | (0.165) | | (1.110) | | (0.203) |
| b | | 1.771** | | -0.251* | | 3.221*** | | -0.629*** |
| | | (0.825) | | (0.145) | | (0.974) | | (0.188) |
| c | -0.859*** | | -0.374*** | | -0.807*** | | -0.156*** | |
| | (0.284) | | (0.044) | | (0.257) | | (0.060) | |
| d | 1.066*** | | 0.274*** | | 0.728*** | | 0.835*** | |
| | (0.179) | | (0.041) | | (0.193) | | (0.049) | |
| gamma | 1.001*** | -8.075*** | 1.235*** | 0.681*** | 1.043*** | -10.631*** | 0.256*** | 1.520*** |
| | (0.294) | (2.081) | (0.073) | (0.255) | (0.301) | (2.408) | (0.084) | (0.313) |

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Augmented regression tests for endogeneity

| | 2001 | | | | 2011 | | | |
|---|---|---|---|---|---|---|---|---|
| | Blacks (E) | Whites (E) | Blacks (F) | Whites (F) | Blacks (E) | Whites (E) | Blacks (F) | Whites (F) |
| GMM C statistics | 66.84 | 25.11 | 9.63 | 2.27 | 68.48 | 33.53 | 59.22 | 31.34 |
| P-value | 0.00 | 0.00 | 0.01 | 0.32 | 0.00 | 0.00 | 0.00 | 0.00 |

Table 8: Hansen's test for overidentifying conditions

| | 2001 | | | | 2011 | | | |
|---|---|---|---|---|---|---|---|---|
| | Blacks (E) | Whites (E) | Blacks (F) | Whites (F) | Blacks (E) | Whites (E) | Blacks (F) | Whites (F) |
| J statistics | 1.44 | 0.49 | 8.94 | 7.94 | 1.50 | 6.82 | 1.65 | 5.40 |
| P-value | 0.49 | 0.78 | 0.01 | 0.02 | 0.47 | 0.03 | 0.44 | 0.07 |

Table 9: Breusch-Pagan's test for heteroscedasticity

| | 2001 | | | | 2011 | | | |
|---|---|---|---|---|---|---|---|---|
| | Blacks (E) | Whites (E) | Blacks (F) | Whites (F) | Blacks (E) | Whites (E) | Blacks (F) | Whites (F) |
| F-statistics | 47.67 | 34.62 | 2124.90 | 1194.54 | 16.26 | 24.72 | 202.52 | 882.67 |
| P-value | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

Assumption of the normality of the residuals is not used to compute these tests. Arguments against the normality of the residuals can be obtained from authors upon request.

## References


[1] James W. Ainsworth-Darnell and Douglas B. Downey. Assessing the Oppositional Culture Explanation for Racial/Ethnic Differences in School Performance. *American Sociological Review*, 63(4):536, 1998.

[2] Eric M. Aldrich, Peter S. Arcidiacono, and Jacob L. Vigdor. Do People Value Racial Diversity? Evidence from Nielsen Ratings. *Topics in Economic Analysis & Policy*, 5(1), 2005.

[3] Simon P. Anderson, André de Palma, and Jacques-François Thisse. *Discrete Choice Theory of Product Differentiation.* MIT Press, 1992.

[4] David Austen-smith and Roland G. Fryer. An Economic Analysis of "Acting White ". *Quarterly Journal of Economics*, 120(2):551–583, 2005.

[5] Patrick Bajari and Matthew E Kahn. Estimating Housing Demand With an Application to Explaining Racial Segregation in Cities. *Journal of Business & Economic Statistics*, 23(1):20–33, jan 2005.

[6] Harminder Battu, McDonald Mwale, and Yves Zenou. Oppositional identities and the labor market. *Journal of Population Economics*, 20(3):643–667, 2007.

[7] Patrick Bayer, Fernando Ferreira, and Robert McMillan. A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy*, 115(4):588–638, 2007.

[8] Alberto Bisin and Thierry Verdier. "beyond the melting pot": cultural transmission, marriage, and the evolution of ethnic and religious traits. *Quarterly Journal of Economics*, (August):955–989, 2000.

[9] George J. Borjas. The Economic Benefits from Immigration. *Journal of Economic Perspectives*, 9(2):3–22, 1995.

[10] Samuel Bowles. Endogenous Preferences : The Cultural Consequences of Markets and Other Economic Institutions. *Journal of Economic Literature*, 36(1):75–111, 1998.

[11] Robert Boyd, Herbert Gintis, and Samuel Bowles. Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rare. *Science*, 328(5978):617–620, 2010.

[12] Yann Bramoullé and Rachel Kranton. Risk Sharing across Communities. *American Economic Review Papers and Proceedings*, 97(2):70–74, 2007.

[13] T. S. Breusch and A. R. Pagan. A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, 47(5):1287–1294, 1979.

[14] Elizabeth E. Bruch and Robert D. Mare. Neighborhood Choice and Neighborhood Change. *American Journal of Sociology*, 112(3):667–709, nov 2006.

[15] Elizabeth E. Bruch and Robert D. Mare. Preferences and Pathways to Segregation: Reply to Van de Rijt, Siegel, and Macy. *American Journal of Sociology*, 114(4), 2009.

[16] David Card, Alexandre Mas, and Jesse Rothstein. Tipping and the Dynamics of Segregation. *Quarterly Journal of Economics*, (February), 2008.

[17] Camille Z. Charles. The Dynamics of Racial Residential Segregation. *Annual Review of Sociology*, 29(1):167–207, aug 2003.

[18] A. J. Christopher. Segregation Levels in South African Cities, 1911-1985. *International Journal of African Historical Studies*, 25(3):561–582, 1992.

[19] A. J. Christopher. Urban Segregation in Post-apartheid South Africa. *Urban Studies*, 38(3):449–466, 2001.

[20] A. J. Christopher. Further progress in the desegregation of South African towns and cities, 1996–2001. *Development Southern Africa*, 22(2):267–276, jun 2005.

[21] William A. V. Clark and Mark Fossett. Understanding the Social Context of the Schelling Segregation Model. *Proceedings of the National Academy of Sciences of the United States of America*, 105(11):4109–14, mar 2008.

[22] Sergio Currarini, Matthew O Jackson, and Paolo Pin. An Economic Model of Friendship: Homophily, Minorities, and Segregation. *Econometrica*, 77(4):1003–1045, 2009.

[23] Jorge De la Roca, Ingrid Gould Ellen, and Katherine M. O'Regan. Race and Neighborhoods in the 21st Century: What Does Segregation Mean Today? *Regional Science and Urban Economics*, doi: 10.10, sep 2013.

[24] John Dixon, Kevin Durrheim, and Colin Tredoux. Intergroup contact and attitudes toward the principle and practice of racial equality. *Psychological science*, 18(10):867–72, oct 2007.

[25] Emin Dokumaci and William H Sandholm. Schelling Redux : An Evolutionary Dynamic Model of Residential Segregation. *Unpublished manuscript, University of Winsconsin-Madison*, 2007.

[26] John Duckitt, Jane Callaghan, and Claire Wagner. Group identification and outgroup attitudes in four South African ethnic groups: a multidimensional approach. *Personality & social psychology bulletin*, 31(5):633–46, may 2005.

[27] Otis D. Duncan and Beverly Duncan. A methodological analysis of segregation indexes. *American sociological review*, pages 210–217, 1955.

[28] William Easterly. Empirics of Strategic Interdependence: The Case of the Racial Tipping Point. *The B.E. Journal of Macroeconomics*, 9(1), jan 2009.

[29] Reynolds Farley, Howard Schuman, Suzanne Bianchi, Diane Colasanto, and Shirley Hatchett. "Chocolate city, vanilla suburbs:" Will the trend toward racially separate communities continue? *Social Science Research*, 7(4):319–344, dec 1978.

[30] Mark Fossett. Ethnic Preferences, Social Distance Dynamics, and Residential Segregation: Theoretical Explorations Using Simulation Analysis. *The Journal of Mathematical Sociology*, 30(3-4):185–273, jun 2006.

[31] Roland G. Fryer and Paul Torelli. An empirical analysis of 'acting white'. *Journal of Public Economics*, 94(5-6):380–396, 2010.

[32] David Gale and Hukukane Nikaidô. The Jacobian matrix and global univalence of mappings. *Mathematische Annalen*, 159(2):81–93, apr 1965.

[33] Edward Glaeser and Jacob Vigdor. The End of the Segregated Century: Racial Separation in America's Neighborhoods, 1890-2010. Technical report, 2012.

[34] Mark Granovetter and Roland Soong. Threshold Models of Diversity: Chinese Restaurants, Residential Segregation, and the Spiral of Silence. *Sociological Methodology*, 18:69–104, 1988.

[35] Sébastian Grauwin, Florence Goffette-Nagot, and Pablo Jensen. Dynamic Models of Residential Segregation: An Analytical Solution. *Journal of Public Economics*, 96(1-2):124–141, feb 2012.

[36] Lars Peter Hansen. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50(4):1029–1054, 1982.

[37] Maximilian Kasy. Identification in a model of sorting with social externalities and the causes of urban segregation. *Journal of Urban Economics*, (October), oct 2014.

[38] Lawrence F. Katz, Jeffrey R. Kling, and Jeffrey B. Liebman. Moving to Opportunity in Boston : Early Results of a Randomized Mobility Experiment. *Quarterly Journal of Economics*, 116(2):607–654, 2001.

[39] Maria Krysan, Mick P. Couper, Reynolds Farley, and Tyrone A. Forman. Does race matter in neighborhood preferences? Results from a video experiment. *American journal of sociology*, 115(2):527–59, sep 2009.

[40] Daniel McFadden. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, pages 105–142. 1974.

[41] Carolyn Mckinney. 'If I speak English, does it make me less black anyway?''Race' and English in South African desegregated schools. *English Academy Review*, 24(2):6–24, 2007.

[42] Takahiro Miyao. A Probabilistic Model of Location Choice with Neighborhood Effects. *Journal of Economic Theory*, 19:347–358, 1978.

[43] P. Norman, P. Rees, and P. Boyle. Achieving Data Compatibility Over Space and Time: Creating Consistent Geographical Zones. 386:365–386, 2001.

[44] Gianmarco I. P. Ottaviano and Giovanni Peri. Rethinking the Effect of Immigration on Wages. *Journal of the European Economic Association*, 10(1):152–197, feb 2012.

[45] Romans Pancs and Nicolaas J. Vriend. Schelling's Spatial Proximity Model of Segregation Revisited. *Journal of Public Economics*, 91(1-2):1–24, feb 2007.

[46] Giovanni Peri and Chad Sparber. Task Specialization, Immigration, and Wages. *American Economic Journal: Applied Economics*, 1(3):135–169, 2009.

[47] James M. Sakoda. The Checkerboard Model of Social Interaction. *Journal of Mathematical Sociology*, 1(May):119–132, 1971.

[48] Thomas C. Schelling. Models of Segregation. *American Economic Review Papers and Proceedings*, 1969.

[49] Thomas C. Schelling. Dynamic Models of Segregation. *Journal of Mathematical Sociology*, 1(May):143–186, 1971.

[50] Thomas C. Schelling. Sorting and Mixing. In *Micromotives and Macrobehavior*. 1978.

[51] John Skvoretz. Diversity, Integration, and Social Ties: Attraction versus Repulsion as Drivers of Intra- and Intergroup Relations. *American Journal of Sociology*, 119(2):486–517, 2013.

[52] Arnout Van De Rijt, David Siegel, and Michael W. Macy. Neighborhood Chance and Neighborhood Change: A Comment on Bruch and Mare. *American Journal of Sociology*, 114(4):1166–1181, 2009.

[53] H. White. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 48(4):817–838, 1980.

[54] M. Wong. Estimating Ethnic Preferences Using Ethnic Housing Quotas in Singapore. *The Review of Economic Studies*, 80(3):1178–1214, jan 2013.

[55] De-Min Wu. Alternative Tests of Independence between Stochastic Regressors and Disturbances. *Econometrica*, 41(4):733–750, 1973.

[56] Junfu Zhang. A Dynamic Model of Residential Segregation. *The Journal of Mathematical Sociology*, 28(3):147–170, jul 2004.

[57] Junfu Zhang. Residential segregation in an all-integrationist world. *Journal of Economic Behavior & Organization*, 54(4):533–550, aug 2004.

[58] Junfu Zhang and Liang Zheng. *Are People Willing to Pay for Less Segregation? Evidence from U.S. Internal Migration.* Elsevier B.V., 2015.