Segregation and the Perception of the Minority*

Florent Dubois†, Christophe Muller‡

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

February 1, 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, Joachim De Weerdt, Matthew Ellman, Maxime Gueuder, Nobuyuki Hanaki, Alan Kirman, Eliana La Ferrara, Gibson Mudiriza, and Juliette Rouchier for their valuable comments. 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 the 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 kind 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

Since the abolition of Slavery, the cohabitation between Blacks and Whites is regularly a hot topic in the public debate of modern society.[1] A large fraction of the White population lost a lot during this transition.[2] Consequently, Whites (especially in rural areas) held Black people responsible for their economic difficulties in a racial rhetoric known as the "White-supremacy" complex.[3] The political debate were then reoriented[4] on racial grounds which ended in the enforcement of legal segregation in the United States with the *Separate but equal* paradigm during the Jim Crow period or the Apartheid regime in South Africa. After the Fair Housing Act in 1968, which put an end to the Jim Crow era, segregation were expected to decrease in the United States and it was effectively the case. After reaching a peak in 1970, segregation declined to its actual level equivalent to the level of segregation in 1910.[5] Segregation in South Africa follows a similar pattern. From the enforcement of the Apartheid policy in 1948, segregation increased until its maximum at the beginning of the nineties. After the political transition, segregation started to decline. However, the different explanations advocated do not fit with empirical evidences.

First, segregation is believed to be the result of discriminations in the housing market.[6] But if we look at several indicators of such discriminations such as the Black-White homeownership gap on one hand, it is quite puzzling to see that this gap is quite stable over a pretty long period of time (1930-2010, Figure 1a).[7] Likewise, although homeownership has increased for both Blacks and Whites, it did so during a period of increasing segregation between 1940 and 1980. Moreover, when segregation starts to decrease in the seventies, the homeownership gap appears to be stable for the subsequent period. On the other hand, if we look at the Black-White price differential, the evidences are quite unclear despite recurrent findings of a Black premium. Indeed, previous attempts to estimate this additional cost

[1] The Black-White relationship is not the only racial tension but probably one of the most prominent. You can think for instance to the persecution of Jews (see for example Voïgtlander and Voth[83] or Glaeser[42]).

[2] Piketty and Zucman[65] show that the stock of slaves represented up to 3 years of national income in the United States during the 1770-1860 period. Thomson[79] depicts a similar picture of South Africa during the same period.

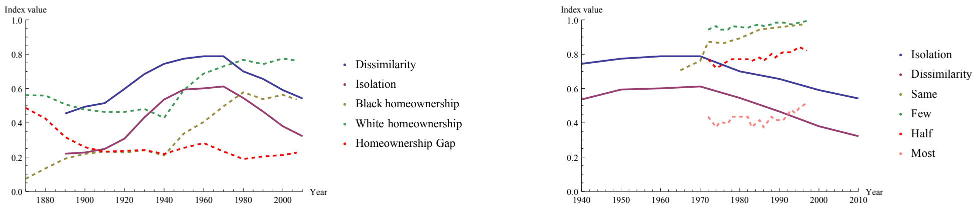
[3] See Ackiss[1] or Heer[47] for more details.

[4] See Glaeser[42] for a description of the mechanism and evidence about lynchings and hate-creating stories during this period.

[5] See Cutler *et al.*[22], or Glaeser and Vigdor[41]. See also Figure 1.

[6] See Taeuber[78], or King and Mieszkowski[53]. See Dawkins[25] and Charles[16] for detailed surveys.

[7] The data of homeownership are taken from Collins and Margo[21] while the data of segregation are taken from Glaeser and Vigdor[41]. They report dissimilarity and isolation indices for Blacks against non-Blacks. Both have used the U.S. census data to compute their indicators.



(a) Homeownership. (b) School preferences.

Figure 1: Segregation in the United States.

are often limited to local housing markets and do not study the dynamics of the prices. Thus, results differ according to the market considered and the period studied. For instance, King and Mieszkowski[53] find evidence of a Black premium in New Haven in 1969,[8] while Lapham[55] does not for Dallas in 1960.[9] More recently Cutler *et al.*[23] find that in 1990 Whites were paying a premium for living into homogeneous neighborhoods. Finally, Ihlanfeldt and Mayock[48] and Bayer *et al.*[4] provide recent evidences of price discriminations against Blacks between 1990 and 2008.[10] Lastly, discrimination practices conducted by real estate brokers are more subtle but have been proven to be resilient. Thus, steering and redlining are still used to discriminate between Blacks and Whites.[11]

Then, a branch of the literature[12] advocated that segregation occurs because Blacks and Whites have different racial preferences and that homophily and racism can explain the trends observed. Thus the decline in segregation would occur because individuals were becoming more tolerant. Although we can observe some compliance to the collective goals of integration and tolerance, individuals seems to be in practice as reluctant as before to integration. For instance, while they agree on the principle of integrated schools, the share of people agreeing to send their kids to schools with a specific share of black kids is pretty stable over time (See Figure

[8]Yinger[86] for St Louis in 1967 or Schafer[69] for Boston in 1970 find also evidence of a Black premium.

[9]Kiel and Zabel[51] document the evolution of the Black-White price differential over the period 1978-1991. They find a reduction for Chicago and Denver (while they find that prejudice increases contrary to Chicago) but a worsening for Philadelphia.

[10]Ihlanfeldt and Mayock[48] study the period 2003-2005 in Florida and find evidences of discriminations against Asians as well. Bayer *et al.*[4] document the situation in four major cities from 1990 to 2008.

[11]Munnell *et al.*[60] provide evidences of robustness of redlining practices in the mortgage market. Yinger[88] documents racial steering in 1981 while Galster and Godfrey[40] brings recent evidences of the persistence of such practices. Finally, Massey and Lundy[57] show that discrimination can occur even before seeing the real estate brokers.

[12]See Becker[8], Schelling[70][71][72], or Farley *et al.*[33][34]. Dawkins[25] and Charles[16] review extensively this point.

1b).[13] Even for neighborhood preferences, Whites still prefer to live in overwhelmingly White neighborhoods while Blacks favor equally populated neighborhoods.[14] However, these kinds of preferences have been proven to generate high levels of segregation as both Blacks and Whites preferences cannot be satisfied simultaneously.[15] An abundant literature in social psychology has demonstrated that these changes reflect the transformation of racial attitudes from the old-fashioned racism to a subtler form of racism.[16] This shift has led a majority of Whites to believe that Blacks are no longer discriminated and generated a profound disillusion inside the Black community. Moreover, they have also shown that negative racial stereotypes still remain prevalent in the United States and mirror much more the Whites' true concerns about integration.[17] As last questioning example, Glaeser[42] has documented the postbellum anti-Blacks hatred. He shows that both lynchings and stories about "negro murder" reach a top in the 1890s and decline afterwards while segregation continuously increases over the same period. Moreover, this particularly violent events occured during the least segregated years. All these evidences about racial preferences are particularly puzzling as the theoretical literature predicts that segregation should emerge and persist in the long run even if individuals have integrationist preferences.[18]

Finally, it is often argued that individuals choose where to live according to the provision of public services in the neighborhood, predicting a complete sorting by income of the neighborhoods.[19] Thus, segregation by race would be a side effect of segregation by income. However, if individuals have different ethnic preferences for public goods, this could hardly promote integration as Blacks would favor different locations than Whites according to the public good provided. For instance, Dawkins[26] estimates with data over the period 1980-2000 that an increase of 10% of this selection effect could increase segregation across jurisdictions by between 4% and 7%. Moreover, empirical studies show that individuals in integrated neigh-

[13] The data of school preferences are taken from Bobo[10], originally derived from the General Social Survey. Data on segregation again come from Glaeser and Vigdor[41]. The curve labelled "Same" represents the share of individuals responding favorably to the question "Do you think White students and Black students should go to the same schools or to separate schools?" while the curves labelled "Few", "Half", "Most" represent the share of individuals responding favorably to the question "Would you have any objection to sending your children to a school where [number] of the children are Blacks?" in a successive order.

[14] See Farley *et al.*[33][34].

[15] See Schelling[71], Vigdor[82] and Clark and Fosset[20] among others has demonstrated this impossibility.

[16] See Gaertner and Dovidio[38], Kinder and Sears[52], or Bobo *et al.*[11]. See Bobo[10] for a more detailed review.

[17] See Bobo and Zubrinsky[9], Farley *et al.*[35], or Timberlake[81]. Bobo[10] and Charles[16] review precisely this literature.

[18] See Schelling[70][71][72], Pancs and Vriend[63], or Zhang[89] among others.

[19] See Tiebout[80].

borhoods vote for a lower provision of public goods.[20] In this case, individuals are likely to move to homogeneous neighborhoods if they want to counterbalance this effect, again predicting an increase of segregation. In the same spirit, Glaeser *et al.*[43] demonstrates that poors tends to concentrate in central cities due to the supply of public transportation in 2001. But again, this is an argument in favor of more segregation as reducing commuting costs allows individuals to more easily reside in homogeneous neighborhoods.[21] All these evidences are puzzling as segregation declines at the same period. Finally, several researchers have looked directly at the relation between inequality and segregation. They have found that segregation tends to rise when inequalities decline due to the formation of new Black affluent neighborhoods after the Black middle-class has reached a certain critical level.[22] However, the difference in income explains only a small fraction of the Black-White segregation.[23]

South Africa have been much less studied but the few evidences reported by the literature suggest a similar pattern. The trends in segregation have been documented by Christopher[17][19][18]. His work is only descriptive and stopped in 2001, only seven years after the political transition but shows that segregation rises and falls as in the United States while the timing is specific to South Africa (see Figure 2).[24] Concerning the discrimination in the housing market, no study really explores this issue in South Africa. Most of them describe the living conditions in the townships. Although a lot have been done to improve them, extreme poverty is still a common feature, high criminality and violence continue as well, and intolerance is still pervasive in the townships.[25] On the other hand, racial attitudes have been more studied. The evidences point to an improvement of the acceptation of the general principle of racial equality but when the question of its implementation comes, racial cleavages reappear.[26] However, the prevalence of the old-fashioned racism is a specifity of the South African case, maybe due to the particular timing in the country. While in the United States old-fashioned racism has mutated and is only slightly present in its original form, it is much

[20]Alesina *et al.*[2], Luttmer[56], Poterba[66], and La Ferrara and Mele[54] are good examples of this effect. See Alesina and La Ferrara[3] and Stichnoth and Van der Straeten[77] for detailed surveys.

[21]See Bayer and McMillan[6].

[22]See Sethi and Somanathan[75] and Bayer *et al.*[5]

[23]See Bayer *et al.*[7] and Sethi and Somanathan.[76]

[24]The data on segregation are taken from Christopher[17][19][18] who reports pairwise dissimilarity indices at the enumeration area level. We simply take the mean of each pairwise dissimilarity index in order to provide a general picture of segregation in South Africa which is represented by the curve labelled "dissimilarity" in Figure 2.

[25]See Jürgens *et al.*[49] or Darkey and Visagie[24].

[26]See Durrheim *et al.*[30], Dixon *et al.*[27]

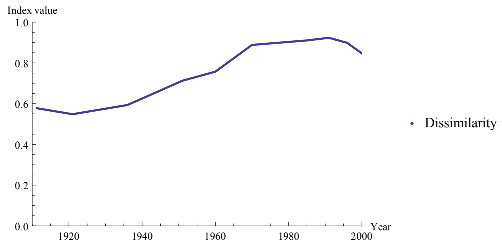


Figure 2: Segregation in South Africa.

more resilient in South Africa.[27] Moreover, racial attitudes in South Africa are also driven by stereotypes and ethnocentrism is present among all racial groups.[28] This is consistent with homophillic behaviors such as the Black diamonds.[29] It is difficult however to believe that desegregation should have occured by the action of the mentioned factors.

As the main explanations advocated seems uncorrelated to the rise and decline of segregation, what could explain the dynamics observed? In this paper, we provide a new explanation for the evolution of segregation. We argue that people may perceive some benefits from living in integrated neighborhoods and choose to relocate in these newly desirable locations. 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,[30] or to the improvement of risk-sharing due to the different assets held by Blacks.[31] Basically, these effects depend explicitly on the size of the minority in the neighborhood which is why we talk about perception of the minority. Moreover, they are supposed to be color-blind in the sense that a White minority should also provide the same kind of benefits in a Black neighborhood. Thus we build a location choice model in which we directly implement this complementary externality in the utility function of the individuals

[27] See Durrheim[31]

[28] See Duckitt *et al.*[29], Gordijn *et al.*[44], and Finchilescu[36] for examples.

[29] The term *Black diamonds* refers to Africans who prefer to keep living in the townships despite they can afford living in former Whites suburbs for status or networks concerns. See Donaldson *et al.*[28] for more details.

[30] Borjas[12] shows that natives benefit from immigration essentially by complementarities in the production. He estimates the gain from immigration between $7 and $25 billion in the United States. Ottaviano and Peri[62] find that immigration has raised the wage of natives with no high school degree by between 0.6% and 1.7% and on average the wage of natives by 0.6%. Finally, Peri and Sparber[64] demonstrate that immigrants and natives specialize in different tasks therefore avoiding competition.

[31] For instance, Bramoullé and Kranton[13] 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.

next to racial attitude effects. We assume that racial attitude is a function of the number of individuals of the different groups living in the neighborhood as Schelling and the following literature did.[32] However, our specification is more in the spirit of Sakoda[68] as all individuals have a specific taste for each racial group. This approach gives us more freedom in the situations studied and is more realistic than the fear of the minority status Schelling assumed.[33] Our other assumptions are standard in the literature as we want to concentrate on the interplay between the perception of the minority, our innovation, and the racial attitudes of the individuals. Thus we do not consider housing market effects as the literature on prejudice has shown that discrimination in this market can arise from racial preferences.[34] After documenting the evolution of segregation in South Africa, we then confront the predictions of our model to census data conducted in South Africa between 1970 and 2001.

We focus our theoretical study on two sets of racial preferences: mutual reject and White segregation versus Black integration. We find that integration can occur and be stable 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 emerge despite a benefit of integration positive enough. But one of the main difference between this situation and the one described by the literature is the robustness of the segregated state. In this case, a shock strong enough can displace the economy into a stable integrated state. As a matter of fact, the situation described by Schelling appears to be a special case of our model. We find also cases of non convergent dynamics. In the empirical part, we provide empirical evidences deriving from structural estimates using Census data of South Africa which give credits to a push toward integration generated by the preferences of the individuals.

The paper is organized as follows. Section 2 first describes the model, then provides the existence result. We characterize the conditions necessary for uniqueness and stability of an equilibrium in the next subsection. We analyse the two relevent structures of preferences mentioned earlier in the following subsection. Finally, we provide an extension of our framework, giving the basis for our identification strat-

[32]See Schelling[70][71][72], Pancs and Vriend[63], or Zhang[89]

[33]Bruch and Mare[14][15] discuss the implication of the functional form used by Schelling and find that the treshold function produces substantially higher levels of segregation. Furthermore, they determine empirically that individuals react continuously to changes in the racial mix of their neighborhood. Easterly[32] also finds no evidence of tipping dynamics.

[34]See Rose-Ackerman[67], Schnare[73], or Yinger[87].

egy 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

The city is divided into two identical locations indexed on $i \in I = \{1; 2\}$. Two groups live in this city. Each individual is directly identifiable by his type $k \in K = \{W; B\}$ which he can not hide.[35] 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, as in the Schelling Bounded-Neighbourhood model[71][72], that any number of individuals can live in the two locations.[36] Moreover each location constitutes a bounded neighbourhood, *i.e.* all individuals inside the location are neighbors with everyone else inside. We assume, as in McFadden (1974)[58] or Miyao (1978)[59], that individuals of type k have an utility level U^{ki} for living in location i. U^{ki} is composed by a deterministic part[37] u^{ki} and a stochastic part ε^{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 about the racial mix of the location where they reside. Moreover, they also care about the presence of a sizable minority. This takes the following form :

$$\left\{ \begin{array}{l} u^{Wi}(N_i^W, N_i^B) = aN_i^W + bN_i^B + \gamma Min[N_i^W; N_i^B] \\ u^{Bi}(N_i^W, N_i^B) = cN_i^W + dN_i^B + \gamma Min[N_i^W; N_i^B] \end{array} \right. , \tag{2}$$

with a, b, c, d, γ 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 perception of the minority. The choice of a linear form is motivated

[35] 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 for instance, while 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[71][72].

[36] Or equivalently that each location is at least of size $L^W + L^B$ and, more generally, of infinite size.

[37] Which can be interpreted as the representative utility of the group k for the location i according to McFadden[58] and Miyao[59].

by the work of Bruch and Mare[14] essentially.[38] They show with Vignette data that agents tend to react continuously to the change in the racial mix of their neighbourhood.

The interpretation of the last term is twofold. First, the min term can be seen as an explicit modelling of what Schelling[70][71][72] 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 γ coefficient, 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.

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 like cleaning the sewers or picking up the trashes.[39] Finally other researchers[40] have provided argument supporting the idea that more diversity inside a zone could be profitable for both groups. To simplify, we assume that the parameter γ is common to both groups. Indeed, if there is an externality due to some complementarity, it is likely that both groups benefit equally from the presence of the other type. Going back to the exemple previously presented of Blacks cleaning the Whites' offices, it is reasonable to assume that the wages offered to the Blacks are here to pay the cleaning of the offices, and thus the benefits of both groups are equivalent. From the minority status perspective, it is reasonable to think that if one group value negatively living in the minority (e.g. due to criminality or poor public goods provision), it is very unlikely that the other group would appreciate to live with a sizable minority associated to deleterious effects.

Individuals choose where they want to live according to a best response rule. Consequently, individuals of type k select location i with probability[41] 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 :

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

[38] Grauwin *et al.*[45] also provide an analytical solution of the Schelling model with potential functions in the case of similar linear utility functions.

[39] Alan Kirman told me this example.

[40] See Semyonov and Glikman[74] among others

[41] Moreover, P^{ki}, as a probability, satisfies $P^{ki} \geq 0$ and $\sum\limits_{i \in I} P^{ki} = 1$

with

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

2.2 Existence

Then we can state the existence of an equilibrium in our framework.

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.*

The definition of the different states is easily generalizable 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[42] and we go back to the standard case used in the literature if both groups are of equal size.

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

[42] As mentioned by Fossett[37], and Clark and Fossett[20]

the model to a linear probability model well-known in the discrete choice theory.[43] Consequently, we can rewrite our equilibrium system and 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 :

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

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 behaviour of the min term in the utility function.[44] We can then compute the equilibrium as a function of the parameters of the model :

$$\begin{cases} W^* = \dfrac{L^2[(c+d+\gamma)(2b+\gamma)L + (a+b+\gamma)(\theta - (2d+\gamma)L) + \alpha(2b+\gamma)] + \alpha L(\theta - (2d+\gamma)L))}{(2b+\gamma)(2c+\gamma)L^2 - (\theta - (2d+\gamma)L)(\theta - (2a+\gamma)L)} \\ \\ B^* = \dfrac{L^2[(a+b+\gamma)(2c+\gamma)L + (c+d+\gamma)(\theta - (2a+\gamma)L) + \alpha(2c+\gamma)] + \alpha L(\theta - (2a+\gamma)L))}{(2b+\gamma)(2c+\gamma)L^2 - (\theta - (2d+\gamma)L)(\theta - (2a+\gamma)L)} \end{cases} \tag{8}$$

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

$$\begin{cases} \dot{W} = \dfrac{((2a+\gamma)L - \theta)W_t + (2b+\gamma)LB_t - L^2(a+b+\gamma) - \alpha L}{\theta} \\ \\ \dot{B} = \dfrac{(2c+\gamma)LW_t + ((2d+\gamma)L - \theta)B_t - L^2(c+d+\gamma) - \alpha L}{\theta} \end{cases} \tag{9}$$

[43] Anderson, de Palma, Thisse (1992)

[44] see the proof of proposition 2 for explicit details of the behaviour of the min function in this case.

with $\dot{W} = \dfrac{\partial W_t}{\partial t}$ and $\dot{B} = \dfrac{\partial B_t}{\partial t}$.[45] 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 equilibrium is unique and asymptotically stable if both :*

- *the corrected intragroup effects are strictly smaller than the standard deviation of the error terms,*
- *the interaction of the corrected intragroup effects is strictly larger than the interaction of the corrected intergroup effects.*

The link between the uniqueness and the stability is due to the linearity of this specification. Moreover, we can characterize the level of the perception of the minority for which the dynamics changes.

Proposition 3. *$\forall a, b, c, d, \theta, L$, there exists a γ^T for which the trace of the Jacobian matrix changes sign, a γ^D for which the determinant of the Jacobian matrix changes sign, and at most possibly two $\overline{\gamma}$ and $\underline{\gamma}$ for which the discriminant of the characteristic polynomial of the Jacobian matrix changes sign.*

2.4 Dynamics

In this section, we explore the relevant situations that can arise from different preference structures.

2.4.1 Mutual reject

The first case, which can serve as a benchmark, could be a case where both groups exert homophilly and racism. If the individuals value negatively living in the minority then this situation is close to the one described by Schelling[71] where individuals have a preference for their peers but do not want to live in the minority. We are also close to the "Segregation" case proposed by Sakoda[68]. A clear example of this situation would be the forced cohabitation of Israeli and Palestinians. Obviously the intragroup parameters have to be positive and the intergroup parameters negative. We may refer later to this general form as the Schelling structure of the preferences or simply the Schelling structure. As an example, we set the parameters values to $a = 10, b = -8, c = -3, d = 7, \theta = 6, L = 1$, which leads to the following results :

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

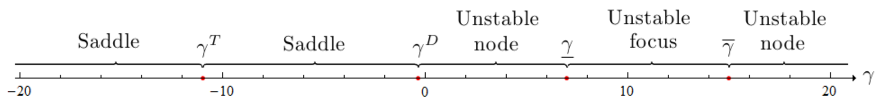


Figure 3: Bifurcation diagram of the parameter γ

For all the negative values of γ, we are explicitly in the situation described by Schelling. People do not want to live in the minority and the stronger the more segregated they are. We can understand precisely why Schelling and other authors[46] finds that the segregation is the only long-run equilibrium. Indeed, the only possible dynamics for a negative value of γ (except for a tiny zone between γ^D and 0) is a saddle (see figure 1), but the crucial point is that the saddle path is precisely the line of all the integrated equilibria. Hence the only way to get an integrated equilibrium in this situation is to start in an already integrated configuration. But any perturbation away from this trajectory leads to a completely segregated equilibrium. Therefore considering the spatial proximity framework he uses, the utility function and the neighborhood considered rule out the possibility to get an integrated equilibrium. Schelling assume that people wants to live in a neighborhood with at least one half of like individuals.

As a matter of fact, all the definitions of neighborhood are constructed the same way and take into account all the individuals located inside an area of any radius R constant for everyone. This is a basic notion of topology but in this case as only integer numbers of individuals can access live inside the neighborhood, you have an even number of people living around a central individual in an odd populated neighborhood. Thus, if you look for instance at the Moore neighborhood, you have at most 9 individuals living in the neighborhood which is already unbalanced (due to the fact that you have only two groups and there is no sum of even integers that can lead to an odd number). But if you remove a peer to the central individual in order to get a balanced configuration, then the central is dissatisfied because he has more unlike neighbors than peers and thus leaves the neighborhood which yields back an unbalanced configuration. Consequently, integrated equilibria cannot be reached and this is why the stochastically stable equilibria are those which are segregated.

Now if we consider positive values of γ, the unique completely integrated equilibrium is still unstable but the dynamics has changed for an unstable node. In this case, the basins of attraction of the integrated states no longer reduce to the previous saddle path. Now the set of the possible configurations is divided into four zones defining the basin of attraction of a corner state, two of them being

[46]See for example Zhang[89], Pancs and Vriend[63], or Young(1998)

integrated and the two others segregated. This configuration is particularly interesting, especially for public policies, because it is a parallel with the growth literature and the "Big push" story. Indeed, if a city is trapped into a segregated state then, by relocating enough people of one specific type, you can switch to an integrated city. However, this is at the cost of deserting the other. But we provide here theoretically some support for the design of public programs such as the well-known Moving To Opportunity program in the United States.

Are these results sensitive to a particular set of parameters? By keeping the same sign of the parameters as described above, applying a more extreme parameters set does not change the order of the different threshold values of γ but just shifts them upward. Thus, at a certain point γ has to be positive enough to trigger the switch between a saddle path and an unstable node. However, a less extreme parameters set can produce different dynamics. For instance, let choose the least extreme integer parameters respecting the Schelling structure of the preferences,[47] *i.e.* a=d=1, b=c=-1, the others remaining the same. We thus obtain the following bifurcation diagram :

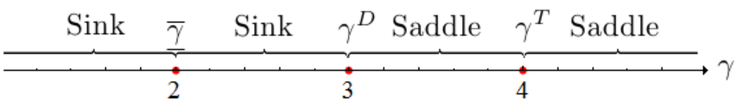


Figure 4: Bifurcation diagram of the parameter γ. Here, $\underline{\overline{\gamma}}$ represents the repeated root obtained by solving the equation $Tr^2 - 4Det = 0$ in γ. As the solution is unique, the discriminant does not change sign with respect to this root.

The two first regimes are in fact only one because of the repeated root which does not change the dynamics. Having a sink in this case is consistent with the results of Miyao[59], and in particular if we set γ equal to 0 we go back exactly to his work. However, we provide here a more general results because the sink dynamics holds for any value of γ belonging to the open interval $(-\infty; \gamma^D)$. Increasing again the value of γ yields a saddle dynamics. However, the saddle path in this case no longer matches the set of integrated states but links the two completely segregated states. The implications are important because segregation is no longer a stable state in the long run. If individuals deviate only slightly, then the only stable state are integrated and one location is deserted. This result is in line with the results of Bøg (2007) who shows that the only stochastically stable states are

[47] We could have chosen a structure of the preferences for which all individuals are indifferent to the presence of both type, *i.e.* a=b=c=d=0, the others being kept constant, the different threshold values of γ would have conserved the same order and thus the same dynamic properties ($\underline{\overline{\gamma}} = 0, \gamma^D = 3, \gamma^T = 6$).

integrated if there is at least one social activist (*i.e.* an individual who prefers to live in an integrated neighborhood if possible, and in an unlike ghetto rather than surrounded by peers otherwise).

Proposition 4. *$\forall a, b, c, d, \theta, L$, there always exists a level of the perception of the minority for which the only stable states are integrated.*

2.4.2 White segregation versus Black integration

Here, one group exerts both homophilly and heterophilly whereas the other has homophillic and racists preferences. This kind of preferences has been revealed empirically in the classical work of Farley *et al.*[33] in the Detroit area. Using data from the Multi-City Study of Urban Inequalities during the late Seventies, they find that Blacks prefer to live in an integrated neighborhood whereas Whites only value positively homogeneous areas despite that they have the same knowledge of the housing market. Obviously in this case, the Whites have the same preferences structure than in the Schelling structure but Blacks value positively both living with like and unlike individuals. We will refer to this general form as the Farley structure of the preferences. The parameter are set to $a = 10, b = -8, c = 5, d = 7, \theta = 6, L = 1$. The bifurcation diagram is thus :

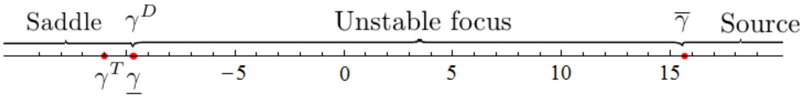


Figure 5: Bifurcation diagram of the parameter γ. Here, $\underline{\gamma}$ is represented on the same point than γ^D because at this scale the two points are confunded in an unique one due to the closeness of the two values ($\gamma^D \simeq -9.71$ and $\underline{\gamma} \simeq -9.65$). Moreover, for the sake of clarity, we represent only one interval between $(-\infty; \gamma^D)$ as the type of the dynamics does not change by crossing over γ^T.

The saddle paths still link the integrated equilibria but from now on, the segregated states that can occur are not complete. Even if Whites want to avoid the contact with Blacks, they have to live with a share of Blacks which size depends on the value of the perception of the minority. The less negatively the minority is perceived, the larger the share of Blacks living with Whites. The mechanism is quite simple here. If the perception of the minority is higher, Whites are more willing to tolerate more Blacks in their location but also more Blacks are willing to live in the minority with Whites. These two effects increase consequently the number of Blacks living with Whites. Moreover, the repartition is quite consistent with the patterns observed by Farley *et al.*[33]. They describe the city of Detroit

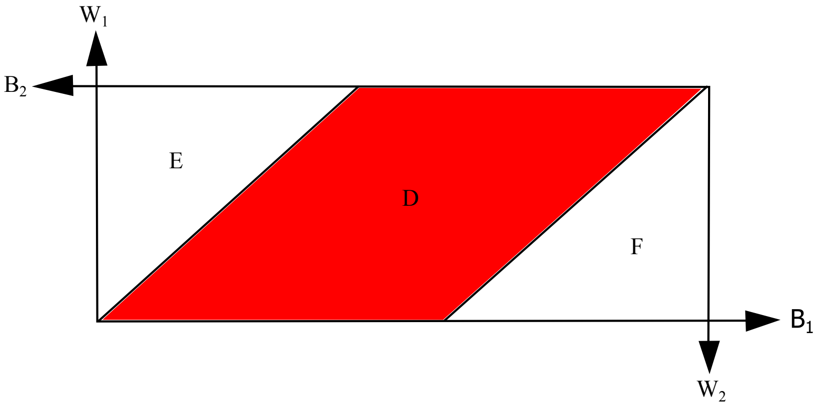


Figure 6: Representation of the different subsets of the city.

as mainly inhabited by Blacks in the central districts while Whites prefer to live in the suburbs, hence the title of their paper "Chocolate city, Vanilla suburb". Then the spiral source emerges as the perception of the minority increases. At this point, Blacks in the Whites-deserted location are attracted by both the maximal concentration of Whites and a growing Black minority in the other one. But as Blacks come in the Whites-occupied location, Whites move out to the previously Blacks-occupied one. Blacks are then attracted by their former predominant location and so on. This leads to a limit cycle where Blacks chase Whites and the system alternate between a completely segregated state and an integrated one with complete desertion. Then finally as the perception of the minority becomes large enough, a source dynamics appears again. The difference with the source dynamics with the Schelling structure resides in the fact that the segregated states which can arise are not completely segregated, a small share of Blacks still living with Whites.

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.[48] 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 group in minority in one location is, by complementarity, the majority in the other location.

[48]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.

Figure 6 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, let us call it D, where the minority group at the city level is also a minority at the local level.[49] 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.

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

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

and

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

Then the corresponding dynamics is:

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

and

[49] As a proof, the red diamond is the set of all points where $min[W;B] = W$ and $min[L^W - W; L^B - B] = L^W - W$ which concludes the proof.

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

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 society is, we are able to identify directly all the coefficient and especially our parameter of interest γ.

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.[50] Thus, the empirical counterpart of the dynamic equation 12 (for Whites in E) is a linear autoregressive model:

[50]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.

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

with δ a constant term, $\boldsymbol{\beta'}\mathbf{X_i}(t)$ a set of location specific control variables with their associated coefficient expressing the attractiveness of the locations[51], and ϵ_{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:

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

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:

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

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

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.[52] 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[53] 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

[52] See [61] for a more detailed description.

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

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.

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.[54] We then use the heteroscedastic-robust variance-covariance matrix of White[84] 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

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

Statistics South Africa, undercount might be due to lack of accessibility which is correlated with race.

Following Kasy[50], 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[50] 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[85] as the assumption of homoscedasticity of the Hausman's test is not verified here.[55] Moreover, we test overidentifying conditions as we have four instruments for two endogenous variables. Results of the Hansen's test[46] is provided in table 6.

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

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

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 behaviour 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 Δu^{ki} can be expressed finally as :

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

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 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

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

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 Ω into $\mathbb{R}^4$, with Ω a closed rectangular region $\Omega = \{N | 0 \leq N_i^k \leq L^k\}$. The Jacobian matrix is thus :

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

if $W \neq B$, and is equal to diag$\{\theta - 2(a+b+\gamma)L,\ \theta - 2(d+c+\gamma)L\}$ otherwise. Then according to the theorem 4 of Gale and Nikaidô[39], 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 $\frac{1}{\theta} > 0$,[56] we have the following sufficient conditions :

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

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

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 :

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

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 :

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

which is easily seen by simple algebra to be the same conditions as for the uniqueness. Moreover, we can easily rewrite the condition on the trace of the Jacobian matrix such that the standard deviation of the uniform error is :

$$\frac{\theta}{\sqrt{12}} > \frac{(a+d+\gamma)L}{\sqrt{12}} \tag{26}$$

which completes the proof.

□

5.3 Proof of Proposition 3

Proof. We solve the equation $Tr_{J_f}=0$ for γ which leads directly to

$$\gamma^T = \frac{\theta}{L} - a - d. \tag{27}$$

Then solving $|J_f|=0$ for γ leads to

$$\gamma^D = \frac{\theta^2 - 2a\theta L - 2d\theta L + L^2(4ad-4bc)}{2[bL+cL-aL-dL+\theta]L}. \tag{28}$$

Finally, solving $Tr^2_{J_f} - 4|J_f| = 0$ for γ is a bit more computationally intensive. First, let rewrite the equation as :

$$Tr^2_{J_f} - 4|J_f| = [2((a+d+\gamma)L-\theta)]^2 - 4[(\theta-(2a+\gamma)L)(\theta-(2d+\gamma)L)-(2c+\gamma)(2b+\gamma)L^2]. \tag{29}$$

We can thus rearrange with respect to γ which leads to a secondary degree equation in γ :

$$\gamma^2 + 2\gamma(c+d) + (a-d)^2 + 4bc = 0 \tag{30}$$

which gives us two possible solutions if the discriminant is positive :

$$\begin{cases} \underline{\gamma} = -(c+b) + \sqrt{(c-b)^2 - (a-d)^2} \\ \\ \overline{\gamma} = -(c+b) + \sqrt{(c-b)^2 - (a-d)^2} \end{cases} \tag{31}$$

as the only way of changing sign of the discriminant of the characteristic polynomial of the Jacobian matrix is to have two real solutions in function of γ, which then completes the proof.

□

5.4 Analytic solution of the first-order differential system

Proof. Recalling the system (9):

$$\begin{cases} \dot{W} = \dfrac{((2a+\gamma)L-\theta)W_t + (2b+\gamma)LB_t - L^2(a+b+\gamma) - \alpha L}{\theta} \\ \\ \dot{B} = \dfrac{(2c+\gamma)LW_t + ((2d+\gamma)L-\theta)B_t - L^2(c+d+\gamma) - \alpha L}{\theta} \end{cases} \tag{32}$$

we can rewrite it in a more tractable form :

$$\begin{cases} \dot{W} = AW_t + PB_t - K_w \\ \\ \dot{B} = CW_t + DB_t - K_b \end{cases} \tag{33}$$

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 :

$$\begin{cases} B_t = \dfrac{\dot{W} - AW_t - K_w}{P} \\ \\ \dot{B} = CW_t + DB_t - K_b \end{cases} \tag{34}$$

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

$$\dot{B} = \frac{\ddot{W} - A\dot{W}}{P} \tag{35}$$

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

$$\frac{\ddot{W} - A\dot{W}}{P} = CW + D\frac{\dot{W} - AW_t - K_w}{P} + K_b \tag{36}$$

which can be rearranged as :

$$\ddot{W} - \frac{(A+D)}{P}\dot{W} + (A - \frac{C}{P})W_t = K_b - \frac{D}{P}K_w \tag{37}$$

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

$$\ddot{W} - \frac{(A+D)}{P}\dot{W} + (A - \frac{C}{P})W_t = 0 \tag{38}$$

which has the characteristic equation:

$$r^2 - \frac{(A+D)}{P}r + (A - \frac{C}{P}) = 0 \tag{39}$$

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 :

$$\begin{cases} \text{If } \Delta > 0, & \text{Then } W_t^g = k_1 e^{r_1 t} + k_2 e^{r_2 t} & \text{with } r_1, r_2 = \dfrac{\dfrac{A+D}{P} \pm \sqrt{\Delta}}{2} \\ \text{If } \Delta = 0, & \text{Then } W_t^g = k_1 e^{rt} + k_2 t e^{rt} & \text{with } r = \dfrac{A+D}{2P} \\ \text{If } \Delta < 0, & \text{Then } W_t^g = e^{\xi t}(k_1 cos\varphi t + k_2 sin\varphi t) & \text{with } \xi = \dfrac{A+D}{2P} \text{ and } \varphi = \dfrac{\sqrt{\Delta}}{2} \end{cases} \tag{40}$$

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 :

$$W_t^p = \frac{PK_b - DK_w}{PA - C} \tag{41}$$

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 :

$$W_t^* = k_1 e^{r_1 t} + k_2 e^{r_2 t} + \frac{PK_b - DK_w}{PA - C} \Longleftrightarrow \dot{W}^* = r_1 k_1 e^{r_1 t} + r_2 k_2 e^{r_2 t} \tag{42}$$

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

$$B_t^* = (r_1 - A) k_1 e^{r_1 t} + (r_2 - A) k_2 e^{r_2 t} - A \frac{PK_b - DK_w}{PA - C} \tag{43}$$

□

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.

- **Children (mean)** represents the mean number of living children per household in a district in 1996.

- **Schooling (mean)** represents the mean number of years of schooling in a district in 1996.

- **Age (mean)** represents the mean age in a district in 1996.

- **Income (mean)** represents the natural logarithm of the mean income in a district in 1996. The income is computed as the center of the class in which the individual has declared to be.

- **Unemployment** represents the natural logarithm of the unemployment rate in a district 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, of the individuals declaring being unemployed, and the individuals aged 15 or older declaring being retired or having a pension.

Table 1: OLS First stage estimation 2001

	OLS 2001			
	Blacks (E)	Whites (E)	Blacks (F)	Whites (F)
1996 L_2_	0.480** [0.210]		1.260*** [0.089]	
1996 N_2L_2_	1.764*** [0.071]		1.508*** [0.022]	
1996 L_2L_2_	-0.407 [0.261]		-1.865*** [0.087]	
1996 L_5L_2_	0.394*** [0.131]	-0.199 [0.135]	0.026 [0.049]	-0.324*** [0.081]
1996 N_5L_2_	-0.189** [0.082]		0.169*** [0.040]	
Mean age (1996)	-0.003*** [0.001]	0.003*** [0.001]	-0.001*** [0.000]	0.002*** [0.000]
Mean years of education (1996)	0.019*** [0.003]	0.000 [0.004]	0.002*** [0.000]	-0.003*** [0.000]
Mean income (1996)	-0.000*** [0.000]	0.000*** [0.000]	0.000 [0.000]	0.000*** [0.000]
Unemployment rate (1996)	0.006 [0.044]	-0.202*** [0.047]	0.010*** [0.003]	-0.034*** [0.003]
1996 L_5_		-1.915*** [0.404]		0.307*** [0.063]
1996 N_2L_5_		-1.712*** [0.219]		0.121*** [0.036]
1996 N_5L_5_		4.256*** [0.146]		4.148*** [0.170]
1996 L_5L_5_		-0.172 [1.498]		-1.652*** [0.237]
Observations	3143	3143	17266	17266
R^2	0.529	0.538	0.794	0.363

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_	0.151 [0.276]		0.793*** [0.099]	
2001 N_2L_2_	1.539*** [0.056]		1.448*** [0.017]	
2001 L_2L_2_	-0.225 [0.324]		-1.313*** [0.099]	
2001 L_5L_2_	0.042 [0.139]	0.058 [0.147]	-0.259*** [0.065]	-0.278*** [0.073]
2001 N_5L_2_	0.029 [0.062]		0.112*** [0.034]	
Mean age (2001)	-0.002*** [0.000]	0.004*** [0.000]	-0.001*** [0.000]	0.003*** [0.000]
Mean years of education (2001)	0.014*** [0.003]	-0.008*** [0.003]	0.003*** [0.001]	-0.003*** [0.000]
Mean income (2001)	-0.000*** [0.000]	0.000*** [0.000]	0.000 [0.000]	0.000 [0.000]
Unemployment rate (2001)	0.199*** [0.040]	-0.363*** [0.038]	0.030*** [0.003]	-0.054*** [0.003]
2001 L_5_		-3.469*** [0.471]		0.200*** [0.066]
2001 N_2L_5_		-0.637*** [0.217]		0.318*** [0.037]
2001 N_5L_5_		4.926*** [0.145]		5.381*** [0.158]
2001 L_5L_5_		4.303** [2.040]		-1.496*** [0.292]
Observations	2555	2555	15584	15584
R^2	0.552	0.620	0.710	0.459

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_	3.114*** [0.203]		1.584*** [0.030]	
1996 N_5L_2_	-0.920*** [0.312]		0.311** [0.124]	
1996 L_2L_2_	-0.955*** [0.339]		-1.905*** [0.088]	
1996 L_5L_2_	-0.155 [0.196]	-0.023 [0.290]	-0.032 [0.050]	-0.401*** [0.141]
1996 L_2_	1.062*** [0.271]		1.228*** [0.089]	
Mean age (1996)	-0.002*** [0.001]	0.002*** [0.001]	-0.001** [0.000]	0.002*** [0.000]
Mean years of education (1996)	0.050*** [0.009]	-0.039*** [0.010]	0.002*** [0.001]	-0.003*** [0.000]
Mean income (1996)	-0.000*** [0.000]	0.000*** [0.000]	-0.000 [0.000]	0.000*** [0.000]
Unemployment rate (1996)	-0.127* [0.075]	-0.036 [0.063]	0.010** [0.004]	-0.035*** [0.003]
1996 N_2L_5_		-3.684*** [0.704]		0.204** [0.104]
1996 N_5L_5_		7.112*** [0.681]		3.947*** [0.281]
1996 L_5_		-4.002*** [0.652]		0.319*** [0.064]
1996 L_5L_5_		2.956* [1.760]		-1.679*** [0.244]
Observations	3143	3143	17266	17266
R^2	0.311	0.445	0.792	0.362

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_	2.408*** [0.134]		1.682*** [0.037]	
2001 N_5L_2_	-0.800*** [0.233]		-0.340** [0.134]	
2001 L_2L_2_	-0.597 [0.385]		-1.543*** [0.101]	
2001 L_5L_2_	-0.302* [0.175]	0.611* [0.337]	-0.329*** [0.071]	-0.545*** [0.195]
2001 L_2_	0.719** [0.328]		0.822*** [0.101]	
Mean age (2001)	0.001 [0.001]	0.002** [0.001]	0.000 [0.000]	0.002*** [0.000]
Mean years of education (2001)	0.035*** [0.006]	-0.034*** [0.006]	0.004*** [0.001]	-0.002*** [0.000]
Mean income (2001)	-0.000*** [0.000]	0.000*** [0.000]	0.000*** [0.000]	-0.000** [0.000]
Unemployment rate (2001)	0.007 [0.058]	-0.252*** [0.049]	-0.002 [0.006]	-0.045*** [0.003]
2001 N_2L_5_		-2.554*** [0.719]		0.338** [0.141]
2001 N_5L_5_		7.364*** [0.521]		7.231*** [0.384]
2001 L_5_		-5.186*** [0.630]		0.391*** [0.073]
2001 L_5L_5_		6.627*** [2.263]		-2.352*** [0.312]
Observations	2555	2555	15584	15584
R^2	0.341	0.536	0.692	0.432

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		3.543***		1.715***		5.325***		1.926***
		(0.659)		(0.148)		(0.867)		(0.180)
b		0.559		-0.298**		2.544***		-0.605***
		(0.644)		(0.145)		(0.861)		(0.172)
c	-0.259**		-0.426***		-0.290*		-0.0582	
	(0.132)		(0.0318)		(0.156)		(0.0388)	
d	0.717***		0.243***		0.466***		0.610***	
	(0.126)		(0.0341)		(0.155)		(0.0379)	
gamma	0.175	-3.011*	1.221***	0.686***	0.488**	-7.026***	0.387***	1.314***
	(0.179)	(1.568)	(0.0689)	(0.247)	(0.218)	(2.101)	(0.0809)	(0.288)

Standard errors in parentheses

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

Table 6: Structural parameters after an IV 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		7.280***		1.677***		8.205***		2.707***
		(1.118)		(0.164)		(1.122)		(0.207)
b		1.881**		-0.194		3.246***		-0.739***
		(0.809)		(0.149)		(0.972)		(0.191)
c	-1.067***		-0.364***		-1.045***		-0.151**	
	(0.300)		(0.0443)		(0.266)		(0.0598)	
d	0.950***		0.272***		0.559***		0.860***	
	(0.178)		(0.0459)		(0.193)		(0.0515)	
gamma	1.217***	-8.377***	1.237***	0.594**	1.309***	-10.98***	0.202**	1.639***
	(0.313)	(2.092)	(0.0794)	(0.258)	(0.310)	(2.431)	(0.0878)	(0.316)

Standard errors in parentheses

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

Table 7: Test 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: Test for overidentifying conditions

	2001				2011			
	Blacks (E)	Whites (E)	Blacks (F)	Whites (F)	Blacks (E)	Whites (E)	Blacks (F)	Whites (F)
Hansen's 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: Test for heteroscedasticity

	2001				2011			
	Blacks (E)	Whites (E)	Blacks (F)	Whites (F)	Blacks (E)	Whites (E)	Blacks (F)	Whites (F)
Breusch-Pagan 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

References

[1] Thelma D. Ackiss. Sociopsychological Implications of the "White-Supremacy" Complex. *American Journal of Sociology*, 51(2):142, 1945.

[2] Alberto Alesina, Reza Baqir, and William Easterly. Public goods and ethnic divisions. *Quarterly Journal of Economics*, (November), 1999.

[3] Alberto Alesina and Eliana La Ferrara. Ethnic Diversity and Economic Performance. *Journal of Economic Literature*, XLIII(September):762–800, 2005.

[4] Patrick Bayer, Marcus D. Casey, Fernando Ferreira, and Robert McMillan. Estimating racial price differentials in the housing market. 2012.

[5] Patrick Bayer, Hanming Fang, and Robert McMillan. Separate When Equal? Racial Inequality and Residential Segregation. *Journal of Urban Economics*, (May), may 2014.

[6] Patrick Bayer and Robert McMillan. Tiebout sorting and neighborhood stratification. *Journal of Public Economics*, 96(11-12):1129–1143, dec 2012.

[7] Patrick Bayer, Robert McMillan, and Kim S. Rueben. What drives racial segregation? New evidence using Census microdata. *Journal of Urban Economics*, 56(3):514–535, nov 2004.

[8] Gary Becker. *The Economics of Discrimination.* Chicago: University of Chicago Press, 1957.

[9] Lawrence Bobo and Camille L. Zubrinsky. Attitudes on Residential Integration: Perceived Status Differences, Mere In-Group Preference, or Racial Prejudice ? *Social Forces*, 74(3):883–909, 1996.

[10] Lawrence D. Bobo. Racial Attitudes and Relations at the Close of the Twentieth Century. In Neil J. Smelser, William Julius Wilson, and Faith Mitchell, editors, *America Becoming: Racial Trends and Their Consequences, Volume 1*, chapter 9, pages 264–301. Washington, DC: The National Academies Press, 2001.

[11] Lawrence D. Bobo, James R. Kluegel, and Ryan A. Smith. Laissez-Faire Racism: The Crystallization of a Kinder, Gentler, Anti-Black Ideology. In S. Tuch and J. Martin, editors, *Racial Attitudes in the 1990s: Continuity and Change*, pages 15–42. Westport, Connecticut: Praeger, 1997.

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

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

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

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

[18] 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.

[19] Anthony J. Christopher. First steps in the desegregation of South African towns and cities, 1991-6. *Development Southern Africa*, 18(4), 2001.

[20] 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.

[21] William J Collins and Robert a Margo. Race and Home Ownership from the End of the Civil War to the Present. *American Economic Review Papers and Proceedings*, 101(3):355–359, may 2011.

[22] David M. Cutler, Edward L. Glaeser, and Jacob L. Vigdor. The Rise and Decline of the American Ghetto. *Journal of Political Economy*, 107(3):455–506, 1999.

[23] David M. Cutler, Edward L. Glaeser, and Jacob L. Vigdor. The Rise and Decline of the American Ghetto. *Journal of Political Economy*, 107(3):455, jun 1999.

[24] Dan Darkey and John Visagie. The more things change the more they remain the same: A study on the quality of life in an informal township in Tshwane. *Habitat International*, 39:302–309, jul 2013.

[25] Casey J. Dawkins. Recent Evidence on the Continuing Causes of Black-White Residential Segregation. *Journal of Urban Affairs*, 26(3):379–400, aug 2004.

[26] Casey J. Dawkins. Tiebout choice and residential segregation by race in US metropolitan areas, 1980–2000. *Regional Science and Urban Economics*, 35(6):734–755, nov 2005.

[27] 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.

[28] Ronnie Donaldson, Thobeka Mehlomakhulu, Dan Darkey, Michael Dyssel, and Pakama Siyongwana. Relocation: To be or not to be a black diamond in a South African township? *Habitat International*, 39:114–118, jul 2013.

[29] 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.

[30] K. Durrheim, C. Tredoux, D. Foster, and J. Dixon. Historical Trends in South African Race Attitudes. *South African Journal of Psychology*, 41(3):263–278, sep 2011.

[31] Kevin Durrheim. White Opposition to Racial Transformation. Is It Racism? *South African Journal of Psychology*, 33(4):241–249, 2003.

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

[33] 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.

[34] Reynolds Farley, Charlotte Steeh, Tara Jackson, Maria Krysan, and Keith Reeves. Continued Racial Residential Segregation in Detroit : " Chocolate City , Vanilla Suburbs " Revisited. *Journal of Housing Research*, 4(1):1–38, 1993.

[35] Reynolds Farley, Charlotte Steeh, Maria Krysan, Tara Jackson, and Keith Reeves. Stereotypes and Segregation: Neighborhoods in the Detroit Area. *American Journal of Sociology*, 100(3):750–780, 1994.

[36] Gillian Finchilescu. Intergroup Anxiety in Interracial Interaction: The Role of Prejudice and Metastereotypes. *Journal of Social Issues*, 66(2):334–351, jun 2010.

[37] 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.

[38] Samuel L. Gaertner and John F. Dovidio. The Aversive Form of Racism. In Samuel L. Gaertner and John F. Dovidio, editors, *Prejudice, Discrimination, and Racism*, pages 61–90. New York: Academic Press, 1986.

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

[40] George Galster and Erin Godfrey. By Words and Deeds: Racial Steering by Real Estate Agents in the U.S. in 2000. *Journal of the American Planning Association*, 71(3):251–268, 2004.

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

[42] Edward L. Glaeser. The Political Economy of Hatred. *Quarterly Journal of Economics*, 120(1):45–86, 2005.

[43] Edward L. Glaeser, Matthew E. Kahn, and Jordan Rappaport. Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics*, 63(1):1–24, jan 2008.

[44] Ernestine Gordijn, Gillian Finchilescu, Louise Brix, Nienke Wijnants, and Willem Koomen. The influence of prejudice and stereotypes on anticipated affect : feelings about a potentially negative interaction with another ethnic group. *South African Journal of Psychology*, 38(4):589–601, 2006.

[45] 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.

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

[47] David M. Heer. The Sentiment of White Supremacy, an Ecological Study. *American Journal of Sociology*, 64(6):592–598, 1959.

[48] Keith Ihlanfeldt and Tom Mayock. Price discrimination in the housing market. *Journal of Urban Economics*, 66(2):125–140, sep 2009.

[49] Ulrich Jürgens, Ronnie Donaldson, Stephen Rule, and Jürgen Bähr. Townships in South African cities – Literature review and research perspectives. *Habitat International*, 39:256–260, jul 2013.

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

[51] Katherine A. Kiel and Jeffrey E. Zabel. House Price Differentials in U.S. Cities: Household and Neighborhood Racial Effects. *Journal of Housing Economics*, 5(2):143–165, 1996.

[52] Donald R. Kinder and David O. Sears. Prejudice and politics: Symbolic racism versus racial threats to the good life. *Journal of Personality and Social Psychology*, 40(3):414–431, 1981.

[53] Thomas A. King and Peter Mieszkowski. Racial Discrimination, Segregation and the Price of Housing. *Journal of Political Economy*, 81:590–605, 1973.

[54] Eliana La Ferrara and Angelo Mele. Racial Segregation and Public School Expenditure. 2006.

[55] Victoria Lapham. Do Blacks Pay More for Housing ? *Journal of Political Economy*, 79(6):1244–1257, 1971.

[56] Erzo F.P. Luttmer. Group Loyalty and the Taste for Redistribution. *Journal of Political Economy*, 109(3):500–528, 2001.

[57] D. S. Massey and G. Lundy. Use of Black English and Racial Discrimination in Urban Housing Markets: New Methods and Findings. *Urban Affairs Review*, 36(4):452–469, mar 2001.

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

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

[60] Alicia H. Munnell, Geoffrey M. B. Tootell, Lynn E. Browne, and James McEneaney. Mortgage Lending in Boston : Interpreting HMDA Data. *American Economic Review*, 86(1):25–53, 1996.

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

[62] 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.

[63] 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.

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

[65] Thomas Piketty and Gabriel Zucman. Capital is Back : Wealth-Income Ratios in Rich Countries 1700-2010. *CEPR Working papers series*, (August), 2013.

[66] James M. Poterba. Demographic Structure and the Political Economy of Public Education. *Journal of Policy Analysis and Management*, 16(1):48–66, 1997.

[67] Susan Rose-Ackerman. Racism and Urban Structure. *Journal of Urban Economics*, 2:85–103, 1975.

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

[69] Robert Schafer. Racial discrimination in the Boston housing market. *Journal of Urban Economics*, 6(2):176–196, apr 1979.

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

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

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

[73] Ann B. Schnare. Racial and Ethnic Price Differentials in an Urban Housing Market. *Urban Studies*, 13:107–120, 1976.

[74] M. Semyonov and A. Glikman. Ethnic Residential Segregation, Social Contacts, and Anti-Minority Attitudes in European Societies. *European Sociological Review*, 25(6):693–708, dec 2009.

[75] Rajiv Sethi and Rohini Somanathan. Inequality and Segregation. *Journal of Political Economy*, 112(6):1296–1321, 2004.

[76] Rajiv Sethi and Rohini Somanathan. Racial Inequality and Segregation Measures: Some Evidence from the 2000 Census. *The Review of Black Political Economy*, 36(2):79–91, may 2009.

[77] Holger Stichnoth and Karine Van der Straeten. Ethnic Diversity, Public Spending, and Individual Support for the Welfare State: a Review of the Empirical Literature. *Journal of Economic Surveys*, 27(2):364–389, apr 2013.

[78] Karl E. Taeuber. Racial Segregation : The Persisting Dilemma. *Annals of the American Academy of Political and Social Science*, 422(November):87–96, 1975.

[79] Leonard Thomson. *A History of South Africa.* Yale University Press, 3rd revise edition, 2001.

[80] Charles M. Tiebout. A Pure Theory of Local Expenditures. *Journal of Political Economy*, 64(5):416–424, 1956.

[81] Jeffrey M. Timberlake. Still Life in Black and White: Effects of Racial and Class Attitudes on Prospects for Residential Integration in Atlanta. *Sociological Inquiry*, 70(4):420–445, 2000.

[82] Jacob L. Vigdor. Residential segregation and preference misalignment. *Journal of Urban Economics*, 54(3):587–609, nov 2003.

[83] Nico Voigtländer and Hans-joachim Voth. Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany. *Quarterly Journal of Economics*, 127(3):1339–1392, 2012.

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

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

[86] John Yinger. The Black-White Price Differential in Housing: Some Further Evidence. *Land Economics*, 54(2):187–206, 1978.

[87] John Yinger. Prejudice and Discrimination in the Urban Housing Market. In Peter Mieszkowski and Mahlon R. Straszheim, editors, *Current Issues in Urban Economics*, pages 430–468. Baltimore: Johns Hopkins University Press, 1979.

[88] John Yinger. Measuring Racial Discrimination with Fair Housing Audits: Caught in the Act. *American Economic Review*, 76(5):881–893, 1986.

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