Ethnic Preferences and Residential Segregation: A Simulation Study

Miruna Petrescu-Prahova[1]

Abstract

Physical characteristics of the urban environment, individual and aggregate socioeconomic characteristics, and individual preferences have all been identified as playing a major role in determining the spatial distribution of ethnic groups within modern cities. However, very few studies have attempted to explicitly identify the manner in which spatial patterns emerge from the interaction of these elements in field settings. In this paper, we use a novel statistical framework based on discrete exponential family models to focus on the role of ethnic preferences in determining spatial residential patterns. We simulate a simple scenario with 1000 households and 400 neighborhoods, and analyze the consequences that xenophobia (i.e., a preference not to reside in the same neighborhood as dissimilar alters) and homophily (i.e., a preference to reside in the same neighborhood as similar alters) combine with other factors to influence the spatial distribution of households to neighborhoods. We conclude that the presence of xenophobia almost always leads to segregation, whereas the effect of homophily depends on its interaction with other factors. These results show that making a distinction between these two types of preferences can provide important insights into the process of residential segregation.

1 Introduction

Ethnic residential segregation has been a visible and salient aspect of urban life in the U.S., especially after the country experienced massive waves of immigration during the 19th and early 20th century. Empirical studies conducted at the beginning of the 20th century noted the existence of ethnic neighborhoods in metropolitan areas with large immigrant populations, such as Chicago and New York (Thomas, 1921). Post-1965 immigrants, although hailing from different origins than their predecessors, have exhibited the same tendency to form ethnic communities in which institutions and services are tailored to the characteristic needs of the ethnic groups (Zhou, 1992; Foner, 2000; Waldinger, 2001). Apart

[1] *Department of Sociology, University of California, Irvine*
Phone: 949-824-6800 Fax: 949-824-4717 mpetresc@uci.edu

from ethnic neighborhoods that formed as a result of immigration, cities in the US are home to a large African American population, which is, and has consistently been, residentially segregated from the native-born white population (Taeuber and Taeuber, 1965; Massey and Denton, 1993; Gottdiener and Hutchinson, 2000).

Previous residential segregation studies have sought to identify the factors that determine residential settlement patterns (Clark, 1992; Zubrinsky Charles, 2001; Wilson and Hammer, 2001; Alba and Nee, 2003). They suggest several factors, which can be classified into three main categories: physical characteristics of the urban environment, individual and aggregate socioeconomic characteristics, and individual preferences for neighborhood composition. Important strides have recently been made in the direction of studying the interactions among these factors by researchers using agent-based and cellular automata models (Epstein and Axtell, 1996; Mare and Bruch, 2003; Benenson, 2004; Zhang, 2005; Fossett, 2006, Bruch and Mare, 2006), based on early work by Schelling (1969, 1971) and Sakoda (1971). Starting from Schelling's initial result that the outcome of a multitude of interrelated individual choices, where unorganized segregation is concerned, is a complex system with collective results that bear no close relation to individual behaviors each taken by themselves (Schelling, 1969), agent-based model research has furthered our understanding of how factors such as neighborhood composition preferences and socioeconomic characteristics influence spatial residential patterns. Fossett (2006) shows that

"ethnic preferences and social distance dynamics can, when combined with status preferences, status dynamics, and demographic and urban-structural settings common in American cities, produce highly stable patterns of multi-group segregation and hyper-segregation (i.e., high levels of ethnic segregation on multiple dimensions) of minority populations" (p. 185).

However, these results have largely been based on simulations of "toy" worlds, and the efforts to extend the analyses to real cases have been hampered by a lack of inferential tools to connect theoretical models with extant data.

In this study we use a novel statistical framework based on discrete exponential family models, which bridges this "inferential gap," allowing the researcher both to simulate simple scenarios in order to understand basic mechanisms, and to make inference based on existing data in order to identify mechanisms in real settings. Here we present results based on simulations of a simple scenario that will allow us to enhance our understanding of the behavior of the model and to build intuitions that will guide empirical data analysis that makes the subject of future research. [2].

2 Potential Determinants of Ethnic Residential Segregation

The physical characteristics of the urban environment are a set of factors that were many times emphasized by the classic urban sociologists of the Chicago School but are often overlooked in more recent studies. Modern cities have certain man-made features, which are intrinsic to their structure and to some extent independent of their resident population, as well as natural features, all of which may be conducive to certain patterns of land use (McKenzie, 1924, Hawley, 1950). Fixed infrastructure (e.g., roads, factories), the spatial distribution of land available for residential use (as opposed to economic use), and the number of housing units, combined with natural barriers such as rivers or hills can influence settlement patterns, since locations which present spatially isolated clusters of housing units may be more prone to segregation than locations with minimal barriers between units. (One of the expressions through which the urban vernacular has captured this situation is “the wrong side of the tracks”, which reflects the fact that oftentimes the borders of segregated neighborhoods are determined by such barriers as railroad tracks (Massey and Denton, 1993).) Foner (2000) notes that in the early years of the Jewish and Italian influx into New York, most immigrants settled in the downtown neighborhoods situated below Fourteenth Street, which ensured that they were living close to the sources of jobs – docks, warehouses, factories, and business streets (p. 39). They were able to move out of these neighborhoods only after the infrastructure of public transportation, roads, and bridges eased the access to new destinations such as Harlem, Brooklyn and Queens. However, even in the extremely densely populated area below Fourteenth Street, Italians and Jews were rarely close neighbors. The grid structure of the streets provided the barriers, and “most blocks were heavily dominated, if not exclusively populated, by one or the other immigrant group” (Foner, 2000, p. 41).

Another set of factors are individual and aggregate socioeconomic characteristics, especially personal income and rent levels. The relationship

[2]Although these initial analyses are based on a simulated landscape, covariates and geographic information from real cases can be used in this model in order either to simulate alternative configurations of real locations or to make inferences about parameter values.

between rent and personal income is a hard constraint on residential choice, especially for low-income households. As a consequence, households with comparable incomes seek locations with similar and affordable rent levels and consequently cluster together in certain parts of the metropolis (Hawley, 1950). If, in addition, we take into account the fact that poverty disproportionately affects members of the minority ethnic groups, we have the premises of ethnic residential segregation through income levels alone (Clark, 1986b; Gottdiener and Hutchinson, 2000). On the other hand, settlement patterns of ethnic groups in urban areas are determined partly by social networks of kinship, friendship, and co-ethnicity. To a large extent, these networks offer support to new immigrants, who are unfamiliar with American society and frequently lack proficiency in English. This leads to geographic concentration of ethnic or even national origin groups (Thomas, 1921; MacDonald and MacDonald, 1970; Massey et al. 1998; Menjivar, 2000). This phenomenon is not restricted to immigrants, however; human geography studies suggest that internal migrants also make settlement decisions based on the geographic location of friends and relatives (Clark, 1986a).

One of the most influential theories for the interpretation of ethnic population distribution across metropolitan space is the spatial assimilation framework, developed by Massey (1985), on the basis of the work of members of the Chicago School such as Robert Park and Louis Wirth. According to this framework, which is related to the normative view of immigrant assimilation in the host societies (as presented by Gordon, 1964), immigrant groups initially settle in enclaves located in the inner city, mainly in economically disadvantaged areas. As their members experience social mobility and acculturation, they usually leave these areas and move to “better” neighborhoods, namely areas that do not have such a high concentration of ethnic minorities, leading to a reduction in ethnic residential segregation levels.

The underlying assumptions of the framework are that neighborhood location and housing are largely determined by market processes and that individuals are motivated to improve their residential status once they have acculturated and made some socioeconomic gains. In this context, residential exposure to the majority group is hypothesized to improve as a result of gains in an ethnic family's socioeconomic standing, acculturation (as measured, for instance, by its members' proficiency in speaking English), and generational status or, in the case of first generation immigrants, length of residence in the country of destination. Residence in the suburbs is also taken into account in the model because it is seen as a sign of enhanced residential assimilation. A series of studies of spatial assimilation for some of the main metropolitan regions, summarized by Alba and Nee (2003), focus especially on the median household income of the census tract of residence and the percent of non-Hispanic whites, the majority group, among residents, as indicators of spatial assimilation. For Asians and Hispanics, the most powerful determinant of living in a high income, high percent white neighborhood is their own socioeconomic position: the greater their income and the higher their educational status, the larger, for instance, the percentage of non-Hispanic whites in the population of the neighborhood where they reside.

The spatial assimilation framework does not apply, however, to African American communities and to immigrant groups that have mixed African ancestry (Haitians, West Indians), because of racial discrimination by the white population (Massey and Denton, 1993). Apart from this shortcoming, the spatial assimilation model, which was built primarily on the experience of the mainly Southern and Eastern European immigrant flows in the early 20th century, fails to account for the experience of new immigrant groups. Responding to these concerns, Portes and Zhou (1993) propose the theory of segmented assimilation, prompted by various research that showed different assimilation outcomes for ethnic groups in the post 1965 wave, which stands in contrast with the classic view of immigrant assimilation as a straight-line process. One of the assimilation trajectories is characterized by upward social and economic mobility in the context of the preservation of ethnic identity and culture, and strong ties with the ethnic community. The achievement of social mobility is no longer linked with the exit from the ethnic community – especially for those groups that have financial capital when they arrive in the U.S. – and remaining in the ethnic community represents a choice rather than a constraint for members of some national-origin groups such as Cubans (South et al, 2005).

The final set of factors suggested by previous research as a potential determinant of ethnic residential segregation are individual preferences for neighborhood composition (Clark, 1992; Zubrinsky Charles, 2001), which can vary according to the reference combination of ethnic groups. One of the first factors is the preference for homogeneity, which can be understood either as a desire to be close to co-ethnics (homophily), or a desire to be apart from ethnic "others" (xenophobia). This type of preference is mostly exhibited by the non-Hispanic white population, who prefers neighborhoods that are 70% or more white, when viewed as combinations of non-Hispanic white and black households (Clark, 1992). In contrast, blacks appear to want a sizable population of coethnics and substantial integration at the same time, leading to a preference for 50%-50% neighborhoods (Zubrinsky Charles, 2001). Hispanics tend to approximate the preferences of blacks, when the reference composition is Hispanic/non-Hispanic white, but approach a preference for neighborhoods that are 75% Hispanic when the potential neighbors are black. In turn, Asian respondents are much more open to integration with non-Hispanic whites than with other groups and find integration with blacks least appealing, while at the same time showing strong preferences for co-ethnic neighbors (Zubrinsky Charles, 2001).

Apart from influencing personal residential choices, neighborhood composition preferences are important because they can lead to discrimination in the housing market, for instance through restrictive covenants signed by neighborhood associations, which limited the choices available to minority groups and led to the creation of segregated neighborhoods (Massey and Denton, 1993). Although some of these extreme, formally implemented measures are now illegal, personal discrimination by real estate agents is harder to identify and eradicate, and its global effects are not well known (Clark, 1992).

Despite the wealth of empirical studies that analyze the potential determinants of residential segregation, very few of them have attempted to compare the relative impact of these factors in generating residential settlement patterns, or to explicitly identify the manner in which such patterns emerge from the interaction of these elements in field settings (Clark, 1986b). It is to the latter issue that we hope to contribute with this study.

3 Research Methodology

The assumption on which the present approach is built is that at any point in time, we can interpret the spatial residential pattern as an equilibrium state of a system of households and areal units, in which households are located in the areal units. However, this system contains various kinds of dependencies: people are tied to one another by kin or friendship relations, and geographic locations are related by virtue of being contiguous or being a certain distance apart from one another. As such, a traditional regression framework is not going to be very reliable in explaining outcomes, and it will fail to represent the complex dependencies within the system. One area of sociology that has seen tremendous advances toward developing stochastic models for social systems with complex dependence structures is social network analysis, where researchers have drawn on earlier results in other scientific fields such as spatial statistics and statistical physics (Robins and Pattison, 2005; Butts, 2005). Building further on these developments, Butts (2005) has proposed "a family of models for social phenomena which can be described in terms of the arrangement of various (possibly related) objects with respect to a set of (again, possibly related) locations" (p. 2). These "generalized location systems" can be used to characterize a range of social processes such as occupational segregation, stratification and settlement patterns. In the case of residential settlement patterns, households represent objects, areal units such as census tracts or block groups represent locations, and we model the probability of observing a particular assignment (i.e. the observed distribution of households across areal units) as resulting from the interaction of factors such as availability of housing, wealth, and preferences for neighborhood composition.

The advantages of this framework are that it can be readily simulated, allowing for the testing of simple scenarios, it is specifiable in terms of directly measurable properties, and supports likelihood-based inference (using Markov Chain Monte Carlo methods). Another set of characteristics that recommends the use of this framework for the study of residential settlement patterns is the ability to include as covariates a range of factors such as population density, inter-household ties, and areal unit characteristics, and examine the effect of their interactions in determining residential patterns.

The generalized location system model is defined as a stochastic model for the equilibrium state of a generalized location system, which represents the assignment of objects (persons, organizations, etc.) to locations (places, jobs, etc.). Given a set of possible configurations (C), the system will be found to occupy any particular configuration with some specified probability. The equilibrium probability of observing a given configuration can be written as

$$\Pr(S = l) = I_C(l) \frac{\exp(P(l))}{\sum_{l' \in C} \exp(P(l'))} \qquad (1)$$

where S is the random state, l is a particular configuration, and P is the quantity we are most interested in, the *social potential*. In this model, the location system is more likely to be found in areas of high potential (which, in turn, are areas of high probability). We need, therefore, to specify a functional form for the social potential that allows us to incorporate as many substantively meaningful effects as possible.

To start with, we can take into account the fact that both objects and locations have features, which can be attributes or relations among objects or among locations. Table 1 gives us a range of effects that we may include in the social potential function:

[TABLE 1 ABOUT HERE]

We now consider these four classes of effects and some examples, without paying attention, at this moment, to their functional form. After reviewing these we present the functional form of the social potential as a linear combination of the four types of effects.

Attraction/repulsion (frequently called push/pull) effects, are based on object and location attributes. Locations (neighborhoods, for example), have attributes that make them attractive (or undesirable) to objects (e.g., households) with particular attributes. High-income neighborhoods attract individuals with high income, and at the same time repel individuals with low incomes. Another important case of this type of effect is discrimination. In this framework, discrimination may be understood as a conditional tendency for households with certain features to be found in (or denied access to) certain locations.

The second category of effects deals with object homogeneity/heterogeneity based on location relations. In other words, this effect captures the tendency for associated locations to be occupied by objects with similar (or different) features. Xenophobia effects can be understood in this framework as the tendency for people of the same race or ethnic origin to reside in contiguous neighborhoods, based on their desire to reduce heterogeneity. This then leads to the formation of clusters of areas with high percentages of people from that group.

Effects of location homogeneity/heterogeneity through relations of objects capture the tendency for locations that are similar to be occupied by people who are associated in some way. An example of such effects is recruitment by entrepreneurs through networks of immigrants. The result is that similar types of jobs (supermarket assistants, for instance) are occupied by people from the same family or community. It is slightly more difficult to interpret this type of effect when locations are geographical units, and it will not be included in the simulations presented below.

Finally, alignment effects express the tendencies for objects that are related to occupy locations that are related in their turn. An example of such an effect is propinquity, the tendency for people who are linked (through kinship or friendship) to reside in neighboring locations. This category of effects is the most flexible, since this function uses matrices as inputs and many mechanisms can be expressed in terms of products of matrices (for instance, density avoidance or homophily).

The social potential is constructed as a linear function of these effects, and has the following expression:

$$P(l) = \sum_{i=1}^{a} \alpha_i t_i^{\alpha}(l) + \sum_{i=1}^{b} \beta_i t_i^{\beta}(l) + \sum_{i=1}^{c} \gamma_i t_i^{\gamma}(l) + \sum_{i=1}^{d} \delta_i t_i^{\delta}(l) \qquad (2)$$

where α, β, γ, and δ are the model parameter vectors, and t^{α}, t^{β}, t^{γ}, and t^{δ} are vectors of sufficient statistics.

We can also express the social potential in terms of the underlying covariates as

$$P(l) = \sum_{i=1}^{a} \alpha_i \sum_{j=1}^{n} \mathbf{Q}_{l_j i} \mathbf{X}_{ji} + \sum_{i=1}^{b} \beta_i \sum_{j=1}^{n} \sum_{k=1}^{n} \mathbf{B}_{i l_j l_k} \left| \mathbf{Y}_{ji} - \mathbf{Y}_{ki} \right| + \sum_{i=1}^{c} \gamma_i \sum_{j=1}^{n} \sum_{k=1}^{n} \mathbf{A}_{ijk} \left| \mathbf{R}_{l_j i} - \mathbf{R}_{l_k i} \right| + \sum_{i=1}^{d} \delta_i \sum_{j=1}^{n} \sum_{k=1}^{n} \mathbf{W}_{ijk} \mathbf{D}_{i l_j l_k} \qquad (3)$$

where **X** and **Y** are vectors of object (e.g.., household) attributes, **Q** and **R** are vectors of areal unit (e.g., census tract) attributes, **B** and **D** are arrays of areal unit relation adjacency matrices, and **A** and **W** are arrays of household relation adjacency matrices. By specifying these parameters in a simulated scenario, we can obtain assignments of households to locations that illustrate what the spatial patterns would be like if these particular mechanisms/effects were at play.

4 Simulation Results

The simulation of a simple scenario with a small population and few effects that are added successively allows us to better understand the behavior of the model and observe how the assignment of households to areal units (which from now on will be referred to as "neighborhoods") is affected by the incorporation of new effects. This scenario includes the following effects:

- Attraction: based on household income and neighborhood rent
- Xenophobia: object homogeneity effect for ethnicity based on the contiguity matrix of the neighborhoods, interpreted as the preference for being far from dissimilar alters
- Homophily: alignment effect between ethnicity and neighborhood contiguity, interpreted as a preference for being close to similar alters, without any preference toward members of the other group. In a scenario with two groups it can have two forms:
    - Single homophily, where only the members of one of the groups prefer to be close to similar alters
    - Double homophily, where members of both groups prefer to be close to similar alters

  This effect can be expressed as

$$\delta \sum_{k=1}^{n} \sum_{k=1}^{n} \mathbf{W}_{jk} \mathbf{D}_{l_j l_k} \qquad (4)$$

  where, for single homophily

$$\mathbf{W}_{jk} = \begin{cases} 1 & \textit{if} \quad \mathbf{Y}_j = \mathbf{Y}_k = 1 \ (\textit{or } 0) \\ 0 & \textit{otherwise} \end{cases} \qquad (5)$$

  for double homophily

$$\mathbf{W}_{jk} = \begin{cases} 1 & \textit{if} \quad \mathbf{Y}_j = \mathbf{Y}_k \\ 0 & \textit{otherwise} \end{cases} \qquad (6)$$

  and where **D** is a neighborhood contiguity matrix and **Y** is a vector of household characteristics, in this case, ethnicity

- Density: alignment effect based on total population counts in neighborhoods, prevents clustering in any one neighborhood, acting as an occupancy constraint
- Propinquity: alignment effect between the inter-household network and the matrix of Euclidean distances between neighborhood centroids

Our principal focus in this analysis is on homophily and xenophobia effects and so comparisons are drawn mainly between model specifications that do and do not include these effects.

In order to characterize and compare the assignments we use residential segregation indices. Researchers concerned with identifying and measuring residential segregation have developed a series of indices that reflect different ways of conceptualizing segregation (Duncan and Duncan, 1955; Lieberson, 1981; Massey and Denton, 1988; Grannis, 2002). Massey and Denton's (1988) classic analysis of segregation indices identifies 20 measures, classified according to five key dimensions of segregation: evenness (the differential distribution of the population), exposure (referring to potential contact between members of different groups), concentration (the relative amount of physical space occupied by groups), centralization (indicating the degree to which a group is located near the center of the city), and clustering (the degree to which minority group members live in contiguous areas). All of these indices measure the degree to which two or more groups live separately from one another, and their calculation is based on a division of the urban area into “neighborhoods”, which most often are Census tracts, and the percentages of the various group populations in the total and neighborhood population. In the present analysis we use two of these indices, the dissimilarity index and the spatial proximity index.

The dissimilarity index, *D*, is the most widely used evenness index, and one of the most widely used segregation indices overall. It measures departure from the even distribution of minority and majority population across areal units, and can be interpreted as the percentage of a group’s population that would have to change residence for each neighborhood to have the same percentage of that group as the urban area overall. For example, a value for *D* of 0.6 in an area where the minority group represents 20 percent of the whole population would mean that 60 percent of the members of the minority group population would have to move in order for all neighborhoods in the area to have a 20 percent minority population. The index ranges from 0 (complete integration) to 1 (complete segregation), and its formula is

$$D = \frac{\sum_{i=1}^{n} (t_i |p_i - P|)}{2TP(1 - P)} \quad (7)$$

where n is the number of neighborhoods (or tracts) in the urban area, T is the total population of the area, t_i is the total population of neighborhood i, P is the proportion minority in the total population, and p_i is the proportion minority in population in area i.

Although they are based on proportions of minority/majority population in clearly defined neighborhoods, most residential segregation indices do not take into account the location of these spatial units of measurement relative to each other, thus ignoring important aspects of segregation such as the geographic distance between two group concentrations (White, 1983; Massey and Denton, 1988; Grannis, 2002). Clustering indices address this shortcoming and measure "the extent to which areal units inhabited by minority members adjoin one another, or cluster, in space" (Massey and Denton, 1988, p. 293). The spatial proximity index (*SP*) is a clustering index proposed by White (1986), which calculates the average of intragroup proximities for the minority and majority populations, weighted by the proportions each group represents of the total population.

$$SP = \frac{XP_{xx} + YP_{yy}}{TP_{tt}} \quad (8)$$

where

$$P_{xx} = \sum_{i=1}^{n} \sum_{j=1}^{n} \left[\frac{x_i x_j c_{ij}}{X^2} \right] \quad (9)$$

$$P_{yy} = \sum_{i=1}^{n} \sum_{j=1}^{n} \left[\frac{y_i y_j c_{ij}}{Y^2} \right] \quad (10)$$

$$P_{tt} = \sum_{i=1}^{n} \sum_{j=1}^{n} \left[\frac{t_i t_j c_{ij}}{T^2} \right] \quad (11)$$

and x_i is the minority population of neighborhood i, y_i is the majority population of neighborhood i, t_i is the total population of neighborhood i, X is the total minority population in the urban area, Y is the total majority population, T is the total population, and c_{ij} has a value of 1 if neighborhoods i and j are contiguous, and 0 otherwise (i.e., c_{ij} is the value of the ij-th cell in the contiguity matrix).

Spatial proximity equals 1 if there is no differential clustering between minority and majority group members. It is greater than 1 when members of each group live nearer to one another than to members of the other group, and is less than 1 if minority and majority members live nearer to members of the other group than to members of their own group.

We begin by specifying the covariates and parameter values used in this simulation scenario[3]. A number of 1000 households are allocated to 400 neighborhoods, represented by squares in a 20 x 20 grid. Each household has one of two types of ethnicity, which is randomly assigned in equal proportions (500 households belong to each type), and is given a random income (drawn independently from a log-normal distribution with parameters 10 and 1.5). Households are tied by social ties (kin or friendship, for example), which are modeled as a Bernoulli graph with mean degree of 1.5 (i.e., a graph in which each edge is an independent Bernoulli trial with probability approximately 0.0015). Each neighborhood is assigned a rent value which scales with the inverse of the distance between its centroid and the center of the grid. Neighborhoods have equal area and relationships among them are expressed in terms of either Euclidean distance between centroids or Queen's contiguity (i.e., two neighborhoods are considered contiguous if they share a border or a point).

The parameter values used in this analysis are listed in Table 2. They are constant across model specifications and have been selected to provide the best illustration of the effect they quantify.

[TABLE 2 ABOUT HERE]

Figures 1 through 11 illustrate simulated draws from various specifications the model. For each model specification the figures correspond to one Metropolis draw, which was sampled after a burn-in sample of 100,000 draws was taken and discarded. Households are represented by circles, with color indicating ethnicity and diameter scaling with income levels (the bigger the diameter, the larger the

[3] For details on the simulation process, which is based on the Metropolis algorithm, see Butts (2005).

household income). Network ties among households are represented by gray lines, neighborhood boundaries are given by the black dotted lines of the grid, and within-neighborhood household positions are jittered to prevent overlap. All models include attraction and density effects and we build on this base by adding various combinations of other effects. Values of the dissimilarity and spatial proximity indices for configurations determined by each models specification are listed in Table 3.

The first set of configurations we analyze is illustrated in Figures 1 through 4. We begin with the model specification that includes only attraction and density effects (Figure 1), and then add, in turn, the xenophobia (Figure 2), single homophily (Figure 3), and double homophily effects (Figure 4). When only attraction and density are present, evenness (as measured by *D*) has moderate levels: 48 percent of the population would have to move in order for all the neighborhoods in the grid to have the 50/50 distribution that characterizes the total population (*D* = 0.48). Clustering, on the other hand (as measured by *SP*, which in this case equals 1.02), is almost non-existent, with the exception, perhaps, of a tendency for higher income households of both ethnicities to congregate close to the center of the grid.

When we turn to the model specification in which xenophobia is added (Figure 2), we observe an assignment of households to neighborhoods that is highly, even completely, segregated, as measured by both indices (*D* = 1.00, *SP* = 1.99). (For this scenario with two groups of equal size, the maximum value of *SP* is 2). The areas occupied by the two groups are separated by an almost empty band, due to the fact that in this case xenophobia is based on the neighborhood contiguity matrix and therefore direct contact between the two groups is discouraged. (In contrast, using Euclidean distance in this case would push the two groups as far part as possible, in diagonal corners of the grid.)

Adding a single homophily effect to the initial attraction and density model generates a configuration that is less segregated than the one that includes xenophobia, but still has relatively high values on both indices (*D* = 0.87, *SP* = 1.64). By adding the single homophily effect (which in this case refers to an above chance tendency for red colored households to be found close to one another), a cluster of red colored households is formed around the center of the grid. This has two consequences. First, there are now many neighborhoods for which the red/black ratio departs from 50/50, leading to a high value for *D*. Second, since *SP* measures clustering directly, its value increases relative to the one obtained in the attraction and density assignment, but as black colored households are still mixed with red ones in some neighborhoods, it does not reach its maximum value as in the attraction, density and xenophobia case.

The double homophily effect that we add last leads to an even less segregated configuration ($D = 0.25$, $SP = 1.02$). Both groups are clustered around the center of the grid, and as they occupy roughly the same area, segregation is not present.

[FIGURES 1-4 ABOUT HERE]

The purpose of analyzing the next group of configurations is to enhance our understanding of the manner in which the simultaneous presence of xenophobia and homophily effects influences the assignment of households to neighborhoods. As can be gleaned from Figure 5 (xenophobia and single homophily, $D = 1.00$, $SP = 1.96$) and Figure 6 (xenophobia and double homophily, $D = 1.00$, $SP = 2.00$), the presence of xenophobia *and* homophily at the same time leads to (almost) complete segregation. These configurations and index values stand in stark contrast with the two configurations in Figures 3 and 4, in which homophily effects were present just by themselves. This result shows that the presence of homophily, not accompanied by xenophobia, is not sufficient to produce high levels of segregation, especially in the case of double homophily.

[FIGURES 5 AND 6 ABOUT HERE]

In the third set of configuration (Figures 7, 8, and 9) we focus on the consequences of adding the propinquity effect to the model. As we noted above, propinquity is an alignment effect which implies that households that are linked via social network ties tend to be found in neighborhoods that are close to each other. In this case, “closeness” is determined by Euclidean distance between neighborhood centroids rather than by contiguity.

When propinquity is added to the attraction, density, and xenophobia model, households belonging to the two groups cluster into ethnically homogeneous bands separated by empty regions. The big areas occupied by the two groups in the previous configuration determined by the attraction, density, and xenophobia effects is broken into smaller bands that are formed so that households that are tied can be found in neighborhoods that are close to each other in Euclidean space. However, the two groups remain highly segregated along ethnic lines (*D* = 0.99, *SP* = 1.98).

By comparing the three configurations in this set, we see again that the model that includes xenophobia leads to the most segregated configuration (*D* = 0.99, *SP* = 1.98, compared with *D* = 0.56, *SP* = 1.20 for single homophily and *D* = 0.21, *SP* = 1.01 for double homophily). An interesting consequence of adding the propinquity effect, for all three cases, is the fact that isolates and lone dyads now appear on the periphery of the grid; the combination of higher income and bigger number of ties has pulled the other households toward the center. This effect is more apparent in the model that includes a single homophily effect, since black colored households do not exhibit the tendency to be close to ethnically similar alters, thus suggesting a connection between low income and social and geographic isolation.

[FIGURES 7-9 ABOUT HERE]

The last set, which comprises Figures 10 and 11, presents draws from two models that include all effects we have considered so far. In these cases the main "structural signatures" observed so far for each of the effects are present: the buffer zone characteristic for xenophobia separates the areas occupied by the two groups, red colored households are clustered together, while black ones are either scattered (single homophily) or clustered (double homophily), and tighter clusters as well as isolates and lone dyads are present due to the propinquity effect. Both configurations are highly segregated, with segregation index values of *D* = 0.99, *SP* = 1.96, and *D* = 0.91, *SP* = 1.92, respectively.

[FIGURES 10 AND 11 ABOUT HERE]

[TABLE 3 ABOUT HERE]

5 Conclusion

The model proposed in this study differs from agent-based and cellular automata models based on ethnic preferences most importantly because it allows researchers to differentiate between the preference to be close to similar alters (homophily) and the preference to be far from dissimilar alters (xenophobia), in contrast with the percentage/threshold approach based on the ethnic composition of the neighborhood employed by previous studies. Several important conclusions can be drawn from the analysis of the simulation results presented here:

- Homophily and xenophobia are distinct processes: models that include a xenophobia effect always lead to segregated configurations, while those including homophily only under certain conditions
- Even within homophily, distinguishing between single and double homophily can provide useful insights: models that include a single homophily effect sometimes lead to moderately segregated configurations, while those including a double homophily effect almost always do not
- Homophily and xenophobia have different structural signatures in terms of spatial patterns of residential settlement, and interact in different and non-trivial ways with other effects.

We must emphasize here that these conclusions are based on the particular covariates and parameters used in simulating the model. Further research in which multiple covariate and parameter values are employed will help improve our understanding of model behavior and residential segregation processes.

References

Alba R, Nee V (2003) Remaking the American mainstream: Assimilation and contemporary America. Harvard University Press, Cambridge

Benenson I (2004) Agent-based modeling: From individual residential choice to urban residential dynamics. In Goodchild MF, Janelle DG (eds.) Spatially Integrated Social Science. Oxford University Press, New York, pp. 67-94

Bruch EE, Mare RD (2006) Neighborhood choice and neighborhood change. American Journal of Sociology 112:667-709

Butts CT (2005) Building inferentially tractable models of complex social systems: A generalized location framework. Institute of Mathematical and Behavioral Sciences
Technical Report MBS 05/08. University of California, Irvine

Clark WAV (1986a). Human Migration. Sage Publications, Beverly Hills, CA

Clark WAV (1986b) Residential segregation in American cities: A review and interpretation. Population Research and Policy Review 5:95-127

Clark WAV (1992) Residential preferences and residential choices in a multiethnic context. Demography 29:451-466

Denton NA, Massey DS (1989) Racial identity among Caribbean Hispanics: The effect of double minority status on residential segregation. American Sociological Review 54:790-808.

Duncan OD, Duncan B (1955) A methodological analysis of segregation indices. American Sociological Review 20:210-17.

Epstein, Joshua M. and Robert Axtell. 1996. Growing artificial societies: Social science from the bottom up The Brookings Institution Press, Washington, D.C.
Foner N (2000) From Ellis Island to JFK: New York's two great waves of immigration. Yale University Press.
Fossett M (2006) Ethnic preferences, social distance dynamics, and residential segregation: Theoretical explorations using simulation analysis. Journal of Mathematical Sociology 30:185–274
Frey WH, Farley R (1994) Latino, Asian, and Black segregation in U.S. metropolitan areas: Are multi-ethnic metros different? Demography 33:35-50
Gordon MM (1964) Assimilation in American life. Oxford University Press, New York
Gottdiener M, Hutchinson R (2000) The new urban sociology. McGraw-Hill, Boston
Grannis R (2002) Discussion: Segregation indices and their functional inputs. Sociological Methodology 32:69-84.
Hawley AH (1950) Human ecology: A theory of community structure. Ronald, New York
Iceland J, Weinberg D, Steinmetz E (2002) Racial and ethnic residential segregation in the United States: 1980-2000. U.S. Census Bureau, Washington, DC
Lieberson S (1981) An asymmetrical approach to segregation. In Peach C, Robinson V, Smith S (eds), Ethnic Segregation in Cities. Croom-Helm, pp. 61-82
MacDonald JS, MacDonald L (1974) Chain migration, ethnic neighborhood formation, and social networks. In Tilly C (ed) An urban world. Little and Brown, Boston, pp 226-235
Mare R, Bruch E (2003) Spatial inequality, neighborhood mobility, and residential segregation. California Center for Population Research Working Paper No. 003-03. University of California, Los Angeles
Massey DS (1985) Ethnic residential segregation: A theoretical synthesis and empirical review. Sociology and Social Research 69:315-350
Massey DS, Denton NA (1985) Spatial assimilation as a socioeconomic outcome. American Sociological Review 50:94-106
Massey DS, Denton NA (1987) Trends in the residential segregation of Blacks, Hispanics, and Asians: 1970-1980. American Sociological Review, 94:802-825
Massey DS, Denton NA (1988). The dimensions of residential segregation. Social Forces 67:281-315
Massey DS, Denton NA (1993). American Apartheid: Segregation and the making of the underclass. Harvard University Press, Cambridge, MA
Massey DS., White MJ and Phua V (1996) The dimensions of segregation revisited. Sociological Methods and Research 25:172-206
McKenzie R (1924) The ecological approach to the study of urban community. Reprinted in Short JF (ed) (1971). The social fabric of the metropolis: Contributions of the Chicago school of urban sociology. University of Chicago Press, Chicago, pp 17-32
Menjıvar C (2000) Fragmented ties: Salvadoran immigrant networks in America. University of California Press, Berkeley, CA

Portes A, Zhou M (1993) The new second generation: Segmented assimilation and its variants among post-1965 immigrant youth. Annals of the American Academy of Political and Social Science 530:74-98

Robins G, Pattison P (2005) Interdependencies and social processes: Dependence graphs and generalized dependence structures. In Carrington PJ, Scott J and Wasserman S (eds.) Models and methods in social network analysis, Cambridge University Press, Cambridge, MA, pp. 192-214

Sakoda JM (1971) The checkerboard model of social interaction. Journal of Mathematical Sociology 1:119-132

Schelling TC (1969) Models of segregation. American Economic Review 59:483-493

Schelling TC (1971) Dynamic models of segregation. Journal of Mathematical Sociology 1:143-186

South SJ, Crowder K, Chavez E (2005) Geographic mobility and spatial assimilation among U.S. Latino immigrants. International Migration Review 39:577-607

Tauber KE, Tauber AF (1965) Negroes in cities: Residential segregation and neighborhood change. Aldine, Chicago

Thomas WI (1921) The immigrant community. Reprinted in Short JF (ed) (1971). The social fabric of the metropolis: Contributions of the Chicago school of urban sociology. University of Chicago Press, Chicago, pp 120-130

Waldinger R (ed.) (2001). Strangers at the gates: New immigrants in urban America. University of California Press, Berkeley

Waters MC (1999) Black identities: West Indian immigrant dreams and American realities. Harvard University Press, Cambridge, MA

White MJ (1983) The measurement of residential segregation. American Journal of Sociology 88:1008-1019

White MJ (1986) Segregation and diversity: Measures in population distribution. Population Index 52:198-221

Wilson FD, Hammer RB (2001) The causes and consequences of racial residential segregation. In O'Connor A, Tilly C, Bobo L (eds.) Urban inequality in the United States: Evidence from four cities. Russell Sage Foundation, New York, pp 272-303

Zhang J (2004) A dynamic model of residential segregation. Journal of Mathematical Sociology 28:147-170

Zhou M (1992) Chinatown: The socioeconomic potential of an urban enclave. Temple University Press, Philadelphia, PA

Zubrinsky Charles C (2001). Processes of racial residential segregation. In O'Connor A, Tilly C, Bobo L (eds.) Urban inequality in the United States: Evidence from four cities. Russell Sage Foundation, New York, pp 217-271

Tables

Table 1 – Elements in the Social Potential Function

	Location Attributes	Location Relations
Object Attributes	Attraction/Repulsion Effects	Object Homogeneity/ Heterogeneity Effects
Object Relations	Location Homogeneity/ Heterogeneity Effects	Alignment Effects

Table 2 – Parameter values in the simulation scenario

Parameter	**Effect**	**Value**
α	Attraction	0.00075
β	Xenophobia	-0.5
δ_1	Single Homophily	0.05
	Double Homophily	0.03
δ_2	Propinquity	-1
δ_3	Density	-0.01

Table 3 – Residential segregation index values for different model specifications

Effects included in the model	Dissimilarity index	Spatial proximity index
Attraction, Density	0.48	1.02
Attraction, Density, Xenophobia	1.00	1.99
Attraction, Density, Single Homophily	0.87	1.64
Attraction, Density, Double Homophily	0.25	1.02
Attraction, Density, Xenophobia, Single Homophily	1.00	1.96
Attraction, Density, Xenophobia, Double Homophily	1.00	2.00
Attraction, Density, Xenophobia, Propinquity	0.99	1.98
Attraction, Density, Single Homophily, Propinquity	0.56	1.20
Attraction, Density, Double Homophily, Propinquity	0.21	1.01
Attraction, Density, Xenophobia, Single Homophily, Propinquity	0.99	1.96
Attraction, Density, Xenophobia, Double Homophily, Propinquity	0.91	1.92

Figures

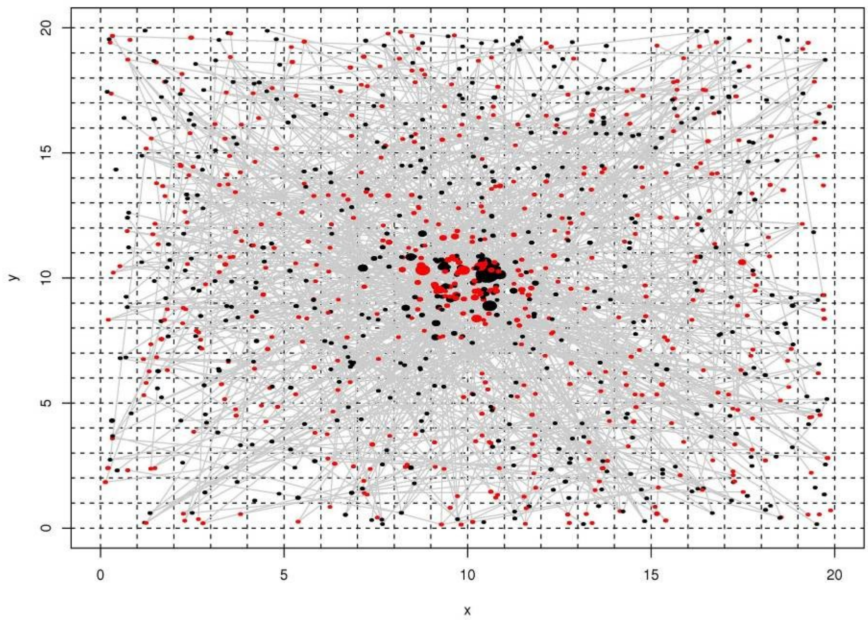

Figure 1 Attraction, Density Effects

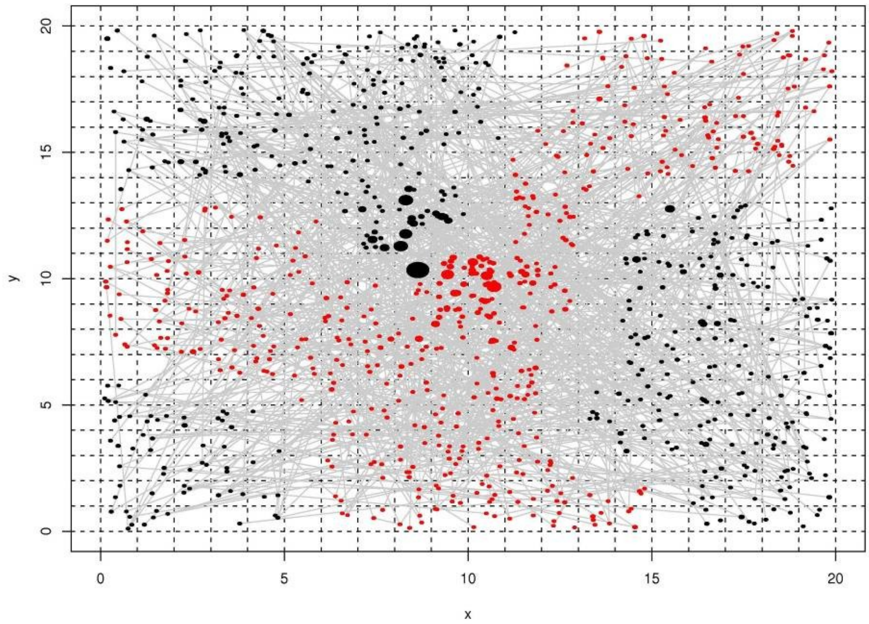

Figure 2 Attraction, Density, Xenophobia Effects

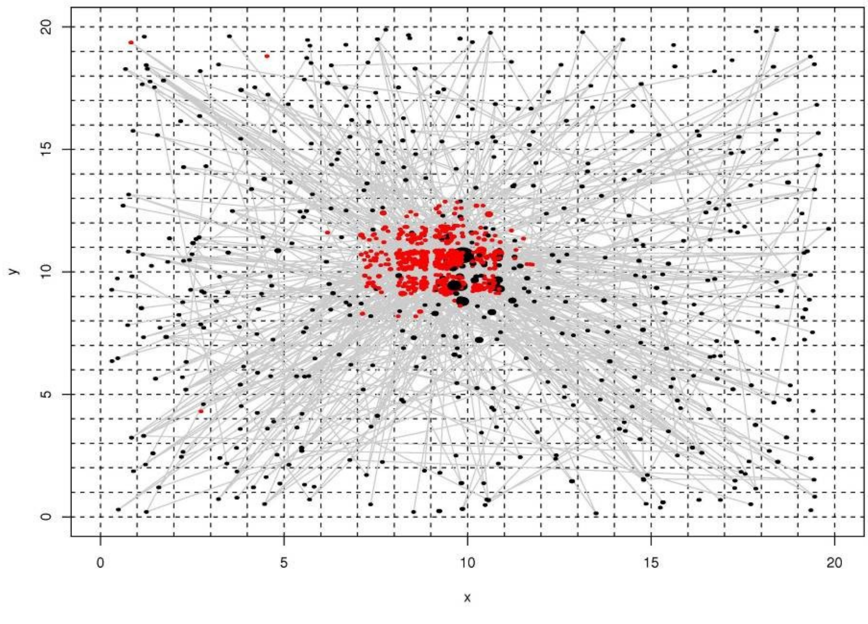

Figure 3 Attraction, Density, Single Homophily Effects

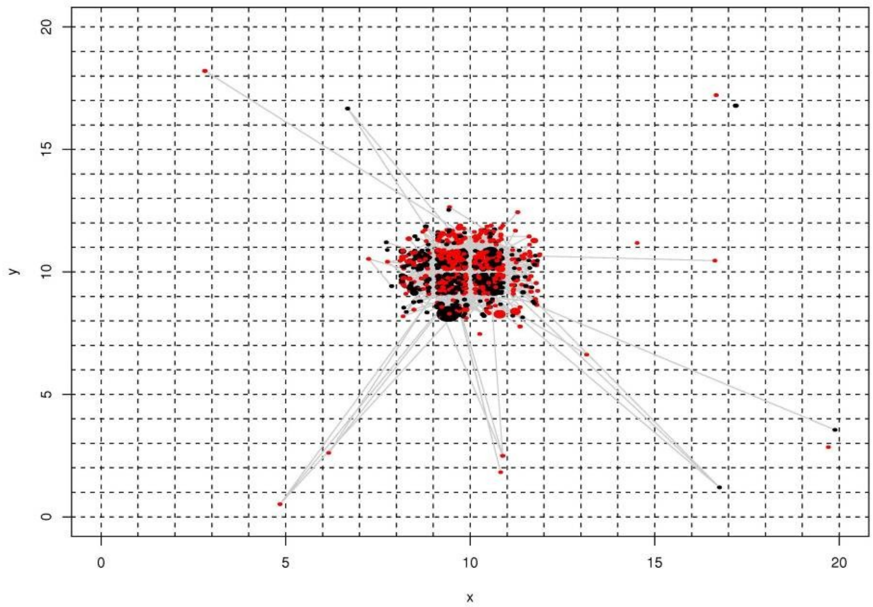

Figure 4 Attraction, Density, Double Homophily Effects

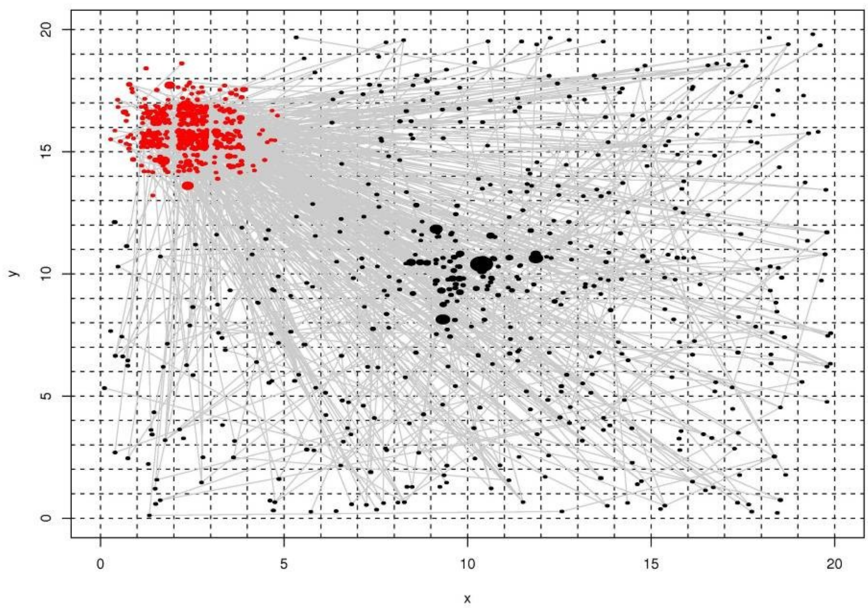

Figure 5 Attraction, Density, Xenophobia, Single Homophily Effects

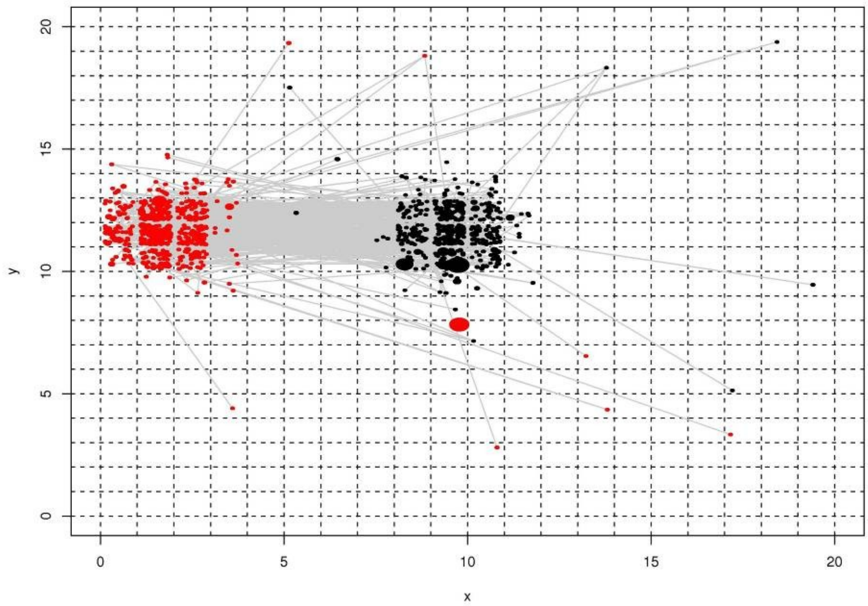

Figure 6 Attraction, Density, Xenophobia, Double Homophily Effects

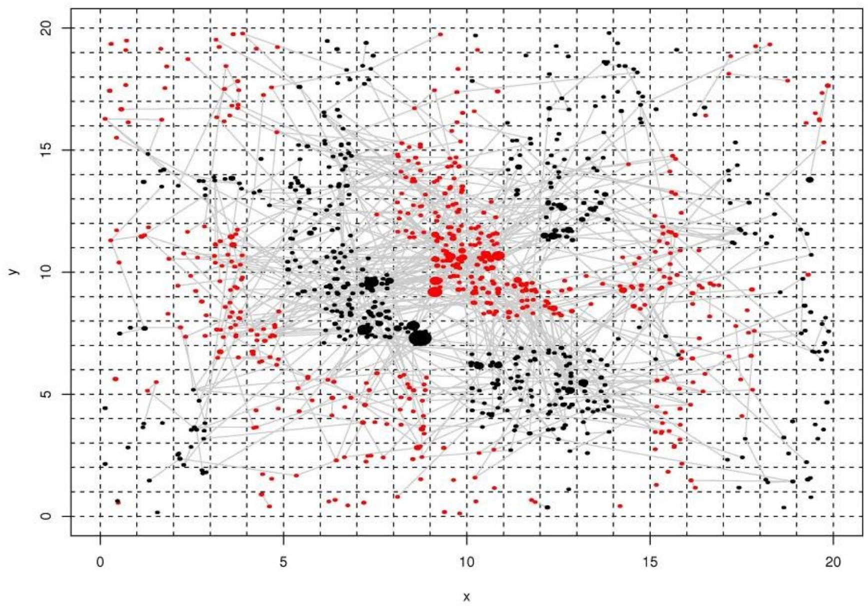

Figure 7 Attraction, Density, Xenophobia, Propinquity Effects

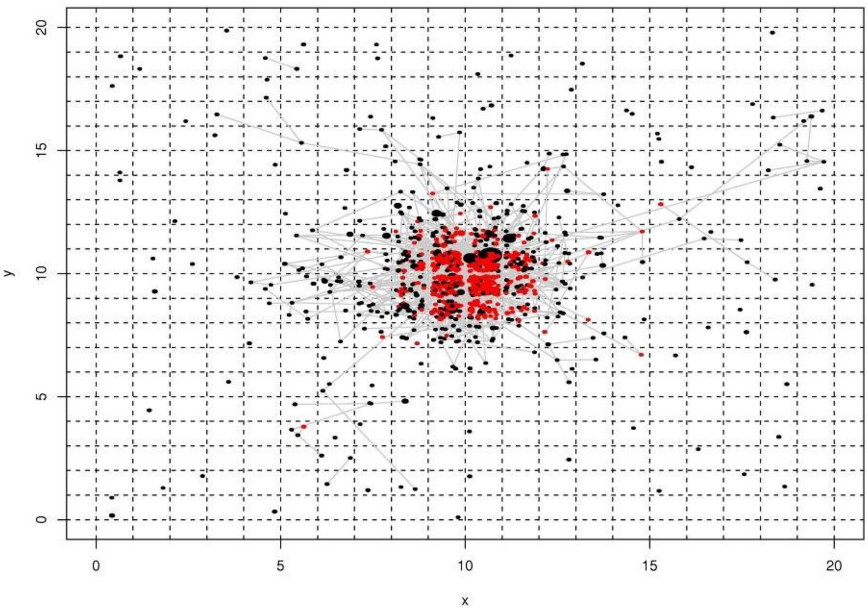

Figure 8 Attraction, Density, Single Homophily, Propinquity Effects

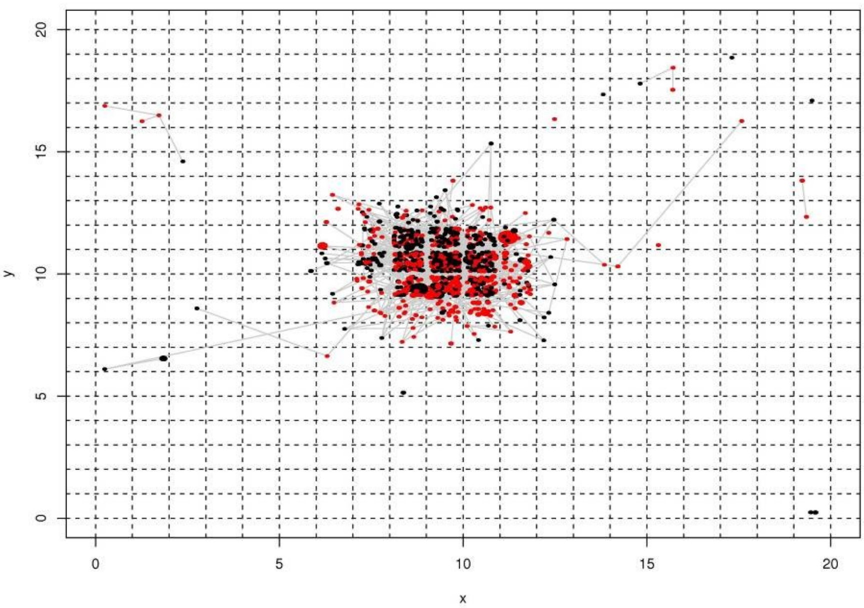

Figure 9 Attraction, Density, Double Homophily, Propinquity Effects

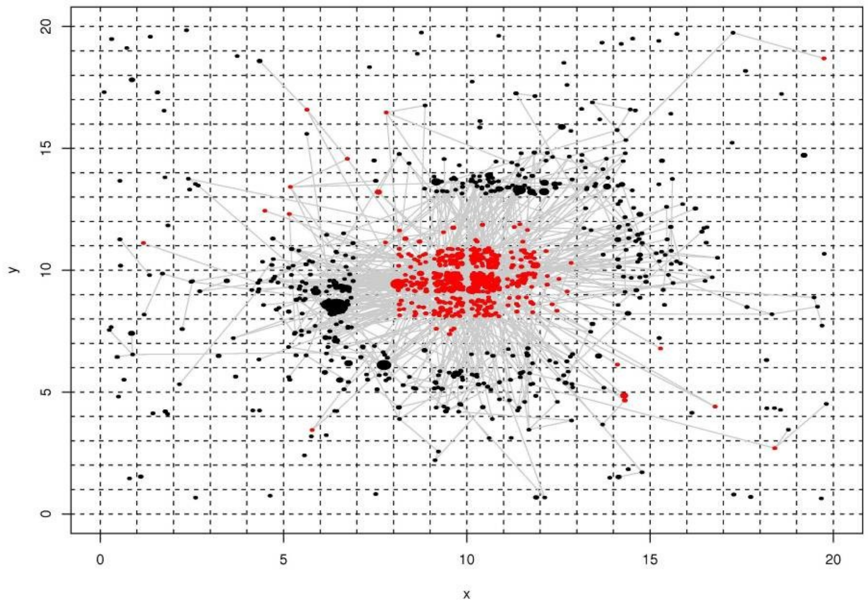

Figure 10 Attraction, Density, Xenophobia, Single Homophily, Propinquity Effects

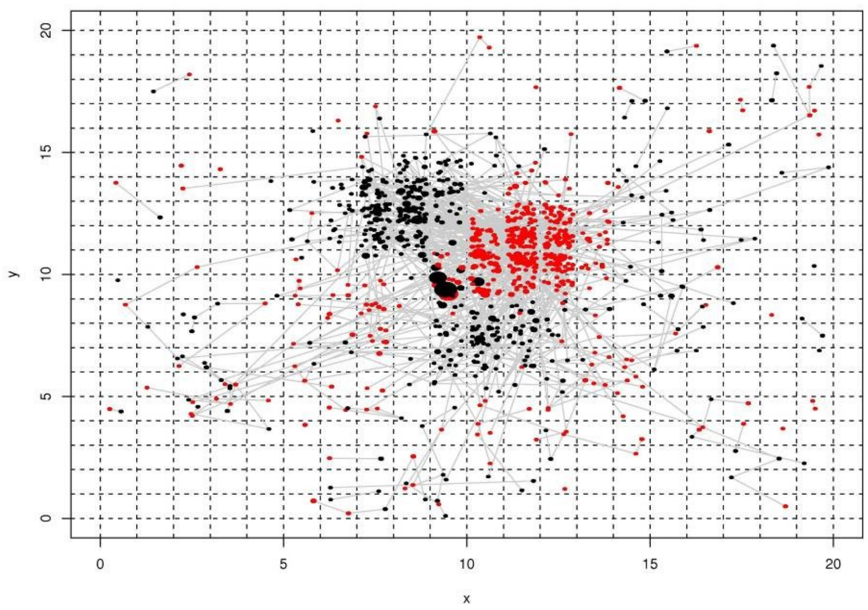

Figure 11 Attraction, Density, Xenophobia, Double Homophily, Propinquity Effects