Immigrant residential segregation in the inner-city of Johannesburg, South Africa

Gayatri Singh, Brown University
Michael J White, Brown University
Seth Spielman, Brown University

Motivation and Aim

An inquiry into the lives and experiences of migrant populations compels us to confront the relationship between distribution of individuals in space and the spatial dimensions of social processes. Spatial configurations of groups or individuals, particularly within residential locations, can be reflective of the underlying social interactions and hierarchies among groups. However, studies of residential segregation have predominantly used tract-based measures. Analysis at the scale of the tract is unable to take into account geographic proximity at the level of the individual household or person. In contrast, an explicitly spatial approach can be provided by an egocentric measure of residential segregation that takes into account questions of proximity as well as scale. The development of such approaches has been slow due to the lack of availability of the requisite data sources and the computational power required. While the data availability still remains an issue, advances in spatial software have opened up new avenues for understanding associations between uneven spatial distributions of groups and socio-economic dimensions such as income, education, age, race, immigrant nationality etc. across a range of social science disciplines.

But even though the analysis of racial and ethnic segregation has become methodologically more refined, the bulk of this work has remained focused on the United States (examples include, among several others, South et al., 2005; Quillian, 2002; Cutler et al., 1999; Borjas, 1998; South and Crowder, 1998; Massey et al., 1994; White et al 1986). Comparative analyses, where available, have been predominantly restricted to urban systems of immigrant receiving western nations (see Johnston et al, 2007 for a recent comparative analysis of Australia, Canada, New Zealand, the United Kingdom and United States). Little attention has been paid to the urban ecology of ethnic residential segregation in developing countries. This is not only an empirical limitation in the cumulative research on segregation but also reflects a level of theoretical impoverishment within the immigrant incorporation and urban segregation literature. Given that the origin of the theories on ethnic segregation lies in the experience of Western nations (particularly the United States) at the time their economies were experiencing unprecedented industrialization and growth (Zorlu and Mulder, 2007), it is legitimate to question their applicability to the experience of developing countries. This is especially pertinent in the case of African nations, like South Africa, with a legacy of late independence from colonial rule and widespread geographical segregation of majority Black populations in their own nations, making the selection of 'host group to assimilate to' even less straightforward than in countries like United States.

This study proposes to re-examine and broaden existing sociological theory of immigrant incorporation by undertaking a study of multi-group migrant settlement patterns in the inner city area of Johannesburg, South Africa. Our study raises important substantive questions about the applicability of existing sociological theories of immigrant incorporation in non-Western contexts. We situate our findings within the framework of sociological theories speaking primarily to the work on immigrant residential segregation that can be classified under two overarching theoretical models, namely *spatial assimilation* and *place stratification* and examine their applicability beyond the Western context.

Data

This study uses a unique dataset with spatial point data collected under Migration and Health survey (2007) in Johannesburg, South Africa under the auspices of the University of the Witwatersrand. This survey was funded as one of the stages of a mixed methods study (qualitative data collection followed by a survey phase) supported by a grant from Mellon Foundation's Joint Mellon Node on Migration and HIV. The survey data contains detailed information on a range of variables seen to influence migrant populations' HIV risk, including limited migration histories, contact with host and origin communities, information about social networks and social support in Johannesburg and detailed socioeconomic indicators over and above behavioral questions on health seeking. The survey took a random sample of four migrant communities in the inner city Johannesburg, namely Somali, Congolese, Zimbabwean and South African (rural-urban migrants).

This is one of the only available survey data that was collected using a spatially based, random sampling strategy in the inner city Johannesburg area, and to our knowledge, in South Africa. The sampling strategy exploited a relatively unique geo-database developed by a Johannesburg based urban consultancy that will enabled spatial sampling and subsequently provided x-y coordinates to be attached to each individual in the dataset, thus making a point based measurement of residential segregation possible for this paper. Information on 754 individuals is utilized in this paper from South African, Congolese and Zimbabwean migrants. Somali group has been excluded from this analyses as a different sampling approach (census of all Somali households in neighborhoods identified by prior community mapping) was used for this group due to their highly clustered nature to get an unbiased sample. The intricacies of the sampling strategy will be provided in greater detail in the full paper. *Figure 1* (below) provides a map of the sampled locations and the sampling frames within study area.

Methods

Using unique point data from the aforementioned survey carried out in the inner city of Johannesburg, South Africa, we attempt to describe patterns of multi-group residential segregation between Black South African host population and immigrants from Democratic Republic of Congo (DRC) and Zimbabwe. We do this by adapting methods taken form spatial point process analysis, extending the K function (Ripley, 1988) to the case of multiple groups on one hand, and to survey data with incomplete residential information on the other hand (base on insights from Diggle, Lang, and Benes, 1991 and Diggle and Chetwynd, 1991).

Extension of the K Function

Based on information of the spatial location (x,y) of each observation, the spatial point process analysis techniques go beyond the well-known "nearest neighbor" approaches. They employ individual-level data, reflect spatial proximity in a way that is fairly robust to topography, and provide formal inferential tests of differences in spatial pattern across groups. The extension of K function that is most interesting for our purposes is the case of two groups, host and immigrant, where the calculation of Cross K function (K_{21}, with reference to the second group) is carried out. The resulting value can used to measure deviation from an expected distribution under independence i.e. the assumption that the two populations are randomly distributed with respect to each other, with suitable edge correction. It is important to note that in the case of human socio-behavioral processes (unlike the spread or clustering of disease), the comparison of the theoretical (or expected) distribution cannot be made under the assumption of complete spatial randomness. Rather, the expected distribution must be derived under the conditions of random labeling of a marked stationary point process defined by two (host and immigrant) qualitatively different 'events' (Diggle and Chetwynd, 1991) or in our case, residential locations labeled by the nationality of the survey respondent. The *K* values are calculated continuously

across distance. Statistics can be calculated so as to reveal negative spatial dependence (members of one group living far apart from members of the other) and the uncommon phenomenon of positive spatial dependence (members of one group being more likely to live near members of the other group). Descriptive results can also be graphically displayed and provide powerful insights into differential nature of residential neighboring with South Africans for Zimbabweans and South Africans as a starting point for the analysis. These are discussed briefly in the section on preliminary findings.

Multivariate analysis:

After describing the residential patterns of Zimbabweans and Congolese with respect to the black South African host population, we intend to engage in a multivariate analysis to estimate a model of residential assimilation that allows us to quantify the effect of immigrant nationality on neighboring with the South African host population, while controlling for important individual level socio-demographic characteristics. To do this, we first come up with a measure of 'neighboring with South Africans' for the immigrant sample. *Neighbor scores* (detail in full paper) have been developed to capture the extent to which an individual is surrounded by persons from the host population in the survey *across* the study area using an inverse relationship of distance from the reference individual. The idea is to determine whether an individual at a geographical location is, *on the whole*, surrounded by those belonging to the host nationality (*host nationality neighbor score maximum)*, within the boundaries of the study area. An individual who is residentially proximate to the South African group (i.e. highest neighbor score for the individual is South African) is given a value of 1 on a new binary variable generated as *HostMax*.

Variable HostMax is then used as the dependent variable in the logistic regression model to provide insight into the factors that predict residential neighboring with the host group by testing for the effect of the certain independent variables on ethnic residential isolation. The independent variables that will be examined are (a) time of first arrival in Johannesburg; (b) self reported nationality; (d) socio-economic status measures based on individual asset score; (e) years of education; (g) immigrant documentation type, (h) presence of children living with the individual in Johannesburg and (i) trust of South Africans. The regression model can be specified as follows:

$$\ln\left[\frac{p_i}{1-p_i}\right] = b_0 + \sum b_i X_i + e_i$$

where, p indicates the probability of having maximum neighbors score South Africans (HostMax=1); the X_i are the covariates and e is an error term.

Preliminary Findings

Figures 2 and 3 below show the plots of Cross K function with simulated envelopes calculated by 99 runs of random labeling of residential locations in the survey data, first for South Africans and Zimbabweans taken together and then for South Africans and Congolese. Bailey and Gatrell (1995) describe the plots Cross K Function quite intuitively in terms of spatial interactions (attraction or repulsion) in the case of two distinct populations. Where there is no spatial interaction between two types of 'events', the empirical Cross K function (marked 'obs' in the figures below) would lie in between the simulated envelopes (marked by 'hi' and 'lo' curves) in figures 2 and 3. When there is 'repulsion' between two types of 'events' we should observe the empirical function lying outside the envelope and below the mean of the simulation runs (marked 'mmean' in the figures 2 and 3). In the case of 'attraction', the empirical function would also lie outside the envelope but above the mean of the simulation runs. We find that in the case of Zimbabweans (Figure 2) there is 'attraction' towards neighboring with South Africans whereas in the case of the Congolese there is 'repulsion' against neighboring with South

Africans in the sample. This is interesting as most Zimbabweans in the inner city are relatively new arrivals as compared to the Congolese, face higher levels of xenophobia by South Africans but seem to settle in close proximity of South Africans. It is likely that several other mediating factors cause these settlement patterns for Zimbabweans. Unlike the Congolese, Ndebele speaking Zimbabweans (majority among the immigrants fleeing Zimbabwe in the recent years) are well versed in Zulu, making it easier for them to mingle, as this is the language spoken by the dominant Black South African ethnic group in Johannesburg. The qualitative phase of Migration and Health study also found evidence of deep mistrust of Zimbabweans towards each other arising out of the climate of paranoia that the President Mugabe's regime has cultivated. Suspicions regarding the political affiliations and motives of fellow compatriots seem to prevent the formation of diaspora-based social bonds of mutual trust. If this is the case, the likelihood of an inclination towards a preference for co-ethnic residence with fellow Zimbabweans may be low. The influence of some of these potential dynamics will be tested in the multivariate analysis.

Figure 1: Map of the sampled locations and the sampling frames within Study Area utilized in this study. (Inset: Inner City Johannesburg Urban Development Zone (UDZ) with the Study Area polygon highlighted.)

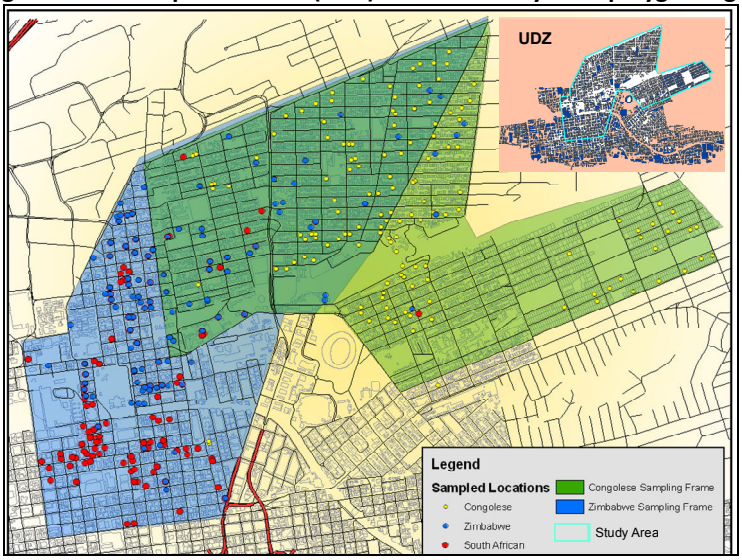

Figure 2

K-Cross plot: South Africans vs. Zimbabweans

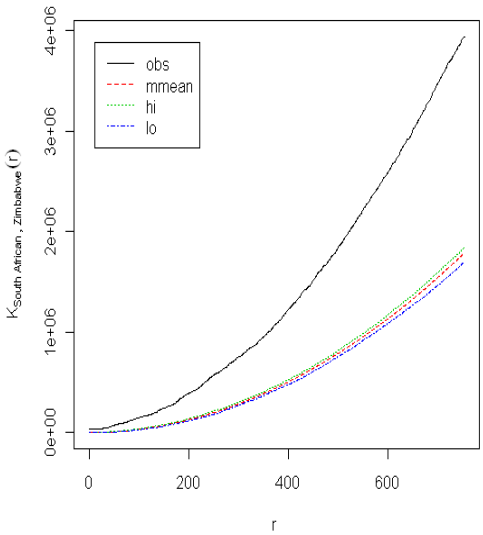

Figure 3

K-Cross plot: South Africans vs. Congolese