Spatial patterns of residential segregation: A generative model

Flavia F. Feitosa[†,*], Joshua Reyes[‡], Walter Zesk[§]

[†] *Center for Development Research (ZEF), University of Bonn, Bonn, Germany*
[‡] *Department of Computer Science, University of Massachusetts Boston, 100 Morrissey Blvd, Boston, MA 02125*
[§] *Rhode Island School of Design and Architecture, Providence, RI*

Abstract. Residential segregation is an explicitly spatial phenomenon that emerges from the interaction of many individuals and displays markedly different global patterns depending on specific socioeconomic contexts. This paper presents a generative model of socioeconomic segregation that reproduces regular macro-level patterns of the phenomenon through the specification of a minimal set of parameters that drive the agents' behaviour. The purpose of these experiments is to provide insights about the emergence of certain types of segregation patterns that have been identified in many modern cities and measure the relation between these distinct outputs and the degree of segregation they produce.

Key words: urban planning; segregation; inequality; agent-based model.

1. Introduction

Residential segregation is a measure of social clumping in an urban environment. It has different meanings depending on the specific form and structure of the city, and its categories include income, class and race. The effects of segregation on cities are overwhelmingly negative. In particular, socioeconomic segregation limits access of disenfranchised population groups to infrastructure and job opportunities, while reinforcing racial and social prejudices. [Cal00] [Ft07] [St01]. The reintegration of urban populations has proved a complex and at times counterintuitive problem for policy makers. It is our hope that an understanding of the underlying mechanics of segregation, how it emerges and grows, could potentially inform and assist these efforts.

Residential segregation exhibits many of the characteristic hallmarks of complex adaptive systems, particularly emergence and non-linearity. Segregation is perceived as a large-scale urban phenomenon, but emerges from the interactions between individuals at a local level [EA96] [Sch78]. Positive feedback introduce a non-linearity into the system. As a result, small differences in socioeconomic and physical context and local behavior can generate large, unexpected and sometimes counter-intuitive outcomes that cannot be understood as the simple sum of the constituent parts [Gil08]. Segregation occurs in most large modern cities, including both the developed and the developing word. An examination of the patterns of segregation in various cities reveals the impact of specific socioeconomic environments. For example, most large cities in the United States have poor and non-white ghettos as a persistent feature of their central areas, whereas the wealthy and mostly white population prefers to live in suburbs [MD93]. In Latin America, wealthy families concentrate in areas that expand from the historical centre into a single geographical direction, while the poorest population mostly settles in the

[*]Corresponding author.
Email addresses: flavia@uni-bonn.de (F. Feitosa), jreyes@cs.umb.edu (J. Reyes), wzesk@risd.edu (W. Zesk).

roughly equipped peripheries [St01] [Vc01]. These two patterns are common and generally recognized, but the structural aspects that cause these distinct patterns remain unclear. What part do individual needs and preferences play in the mechanics of segregation? How do socio-demographic factors, such as inequality of income, impact the emergence of different spatial outcomes?

Generative models have been used to understand the structural characteristics of urban populations and many, similar complex social systems [Eps07] [EA96]. Most often, an agent-based computational model attempts to reproduce certain global regularities from minimally defined local interactions among individuals within the system. Such a model makes it possible to conduct experiments on a simplified population of agents in a spatial environment. Agents interact locally according to small set of rules, thereby generating global regularity from the bottom up. Simply, the motto of a generative social approach is: 'If you didn't grow it, you didn't explain it' [Eps07].

This paper presents a generative model of socioeconomic segregation that reproduces patterns much like those observed in real cities. The rules defining the interaction of agents are few and simple. The purpose of this experiment is to create a stylized, abstract model that is able to provide insight into the mechanics of emergence and the variety of segregation patterns identified in modern cities. In addition, this work understands each pattern in terms of the degree of segregation produced. Our simulation enables the quantitative comparison of different patterns based on defined segregation measures: exposure and clumpiness. The remainder of this paper is organized as follows: Section 2 gives a brief introduction to the classical types of segregation patterns that are considered in this work. Sections 3 and 4 describe the model's specification, its fundamental assumptions, and a discussion of results. Finally, Section 5 presents the concluding remarks and directions for future research.

2. Classical Patterns of Residential Segregation

The Chicago School refers to set of urban studies that emerged in Chicago during the first half of the 20th century. It became famous for its systematic, formal approach, focused on the city as a social laboratory. The school's effort to understand the spatial organization of human activities resulted in classical urban models that attempt to distinguish and categorize patterns of residential segregation. The concentric model, proposed by Burgess [Bur24], argues that a city grows outward from a central point in a series of rings. Burgess observed that there was a correlation between the distance from this central point (the central business district or CBD) and the wealth of residential areas. Since Burgess' studies were based on Chicago, he observed that wealthier families tended to live further away from the CBD. Later examples also followed the Burgess model, but with an inverse correlation between the CBD and the wealth of neighborhoods. Some Latin-American cities exhibit this center-periphery pattern, especially until the 1980's, where wealthier families tend to concentrate in central areas, while poorer families occupy the outskirts of the city.

Hoyt [Hoy39] proposed a second model, known as the sector model, which assumes that cities develop in sectors rather than rings. According to this model, if a district is set up for high income residences, any new development in that district will expand from the outer edge and, therefore, the sector shape emerges.

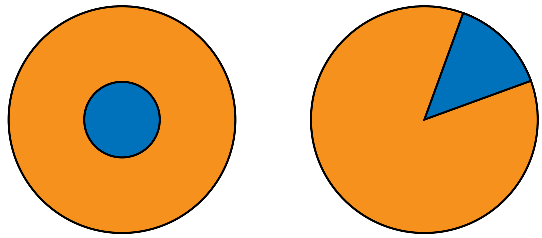

Figure 1. Patterns of residential segregation: concentric or center-periphery model of Burgess (*left*) and the sector model of Hoyt (*right*).

Figure 1 illustrates patterns of residential segregation based on Burgess' concentric model and Hoyt's sector model. Burgess' and Hoyt's models describe cities at a very high level, disregarding specific geography, culture or history. Moreover they impose several other stringent constraints as both models refer exclusively to monocentric cities built upon static landscapes. Nevertheless, the basic features of the center-periphery and sector models are still commonly recognized in cities around the world, and are often accepted as the basis for segregation analyses. This work adopts both models as reference for regular macro-level patterns of segregation that are expected to emerge from simple interaction rules among individuals.

3. The Model Specification

Our model relies on the premise that socioeconomic segregation is the outcome of a contest for the most convenient locations in the city. In this model, only two attributes define how convenient a location is: the quality, which acts as a general proxy for the public goods available at that location, and its proximity to the CBD. The total environment consists of a two-dimensional 45×45 grid. Each cell within the grid corresponds to a dwelling unit. At the beginning of each run, each cell receives an equal quality rating. The initial price distribution falls off as the inverse of distance to the CBD. This condition merely prices those locations closest to the center of the city highest before any individuals move in. We normalized income and prices between 20 and 100 units of currency. Households with incomes that exceed 100 units have no more purchasing power and political influence than those with 100.

Since the aim of our model is to 'grow' a city, the grid is initially populated with a minimum amount of agents. Each agent corresponds to a single household, which is characterized by its location and income. Incomes are drawn according to a Pareto distribution, the skewness of which can be controlled by its exponent, known as Pareto index. The Pareto index regulates the level of income disparity in the city and can be chosen by the user—higher values describe distributions with high disparities between the very rich and the very poor, while lower index values result in more uniform distributions of wealth.

The city is gradually populated by households according to a fixed growth rate. While looking for their residential location, the new inhabitants attempt to maximize the utility derived by a location over some subset of the presently unoccupied cells according to the Cobb-Douglas function:

$$U = \left(\frac{1}{D}\right)^{1-\alpha} Q^{1+\alpha},$$

where Q is the quality index of the cell, D is its distance to the CBD, and $\alpha \in [-1, 1]$ is the quality-priority index. The quality-priority index introduces a cultural bias to the model. Households will prioritize locations with high quality indices if α is near 1, and prioritize cells close to the city's center for values of α close to -1. For values of α near zero, individuals try to strike a balance between quality and distance when searching for a place to inhabit. In this way the quality-priority index tries to capture the culture of the city. Below we explain how this parameter effects the geometry of the resultant city.

The individuals in this model exhibit boundedly rational behavior. Because search is costly, every household evaluates only a limited number of the total, possible locations to move to. This maximization is constrained by the household's income. Unlike other models of segregation, the households choose their new location without giving consideration their neighbors, their income, or any other attribute explicitly. The decision-making process relies entirely on the amount of public goods available to potential residence and its distance to the downtown. At each update, however, individuals compare utilities among others with similar level of income. In our model, households with utilities that are much lower than their financial neighbors grow dissatisfied and may repeat the search process and move to another location.

As individuals choose their residential location, they change the environment. This model assumes that individuals with high income are able to advocate for more public goods in their neighbourhood, and therefore slightly increase property value (price) and introduce more public goods, such as improved school systems, police, or public parks (quality). On the other hand, individuals with low income lack similar resources, and so property value and public goods in those neighbourhoods decrease slightly.

3.1. **Measurements.** The spatial arrangement of the population is constantly monitored through segregation indices. Since segregation is not a unidimensional phenomenon [MD88], we selected segregation indices that are able to capture the segregation dimensions defined by Reardon and O'Sullivan [RO04]: *spatial evenness/clustering* and *spatial exposure/isolation* (Figure 2). To measure each of these dimensions, we compute spatial indices of segregation proposed by Feitosa *et al.* [Ft07].

Spatial evenness/clustering refers to the balance of the distribution of population groups, and it is measured by the generalized neighbourhood sorting index GNSI, which calculates how much variances in income between different neighbourhoods contribute to the total income variance of the city. Generally, higher variance between neighbourhoods signifies higher levels of segregation along this axis. Spatial exposure/isolation refers to the likelihood for members from different groups to live side-by-side. We use spatial the exposure index $\check{P}_{(m,n)}$, which measures the average proportion of group n in the neighbourhoods of each member of group m, to measure exposure. Since both these indices are global quantities, we also compute local indices of spatial isolation $\check{q}_m$ that can be displayed as maps and

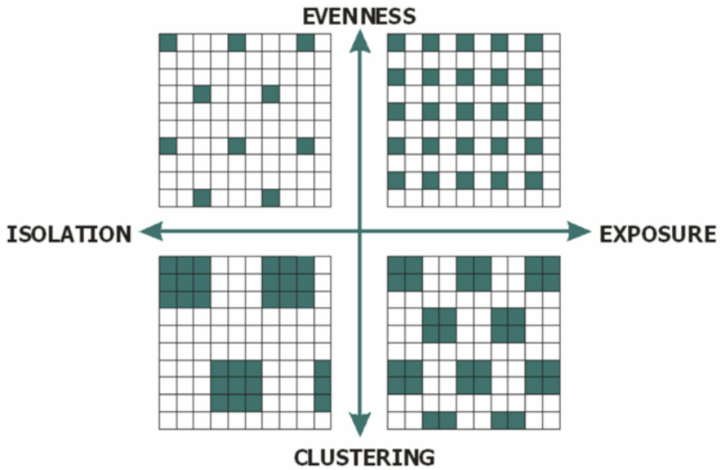

FIGURE 2. Spatial Dimensions of Segregation.

facilitate the visual interpretation of the simulations. For the computation of group-based measures $\check{P}_{(m,n)}$ and $\check{q}_m$, we split the population according to income. Those individuals with the income above the median value belong to the group "wealthy," while those with the income falling below the median belong to the group "poor." The size of the neighbourhoods considered in the calculation of the indices are established according to a user-defined radius.

4. Growing Spatial Segregation Patterns

We ran our simulation several times for varying preferences and economic background. During these runs the quality-priority ranged over the values $\alpha = -1$, -0.5, 0, and $+0.5$ for each of scenario captured by the Pareto inequality indices $\beta = 0.5$, 1, and 1.5—cities with low, moderate, and extreme disparities in wealth. We excluded the extreme case $\alpha = 1$ because in this culture households do not give any importance to living near the center of the city whatsoever. When proximity to the CDB is completely disregarded, the resultant 'cities' are not cities at all but expanding and uncoordinated spreads of individuals seeking out untouched land. Each simulation ran until its population exceeded one thousand households. We stopped the runs short to limit unreasonable amounts of noise from the artificial, rectangular boundary of our model. Generally, real cities do not have a hard edge beyond which they cannot develop. Imposing an early termination nods to this fact.

Figure 3 presents the distribution of households for two runs in our experiment. The rightmost column depicts the physical distribution of households, colored by income, across the model's landscape. In this picture, red, orange, and blue squares correspond to households with low income, moderate, and high incomes respectively. In the middle and leftmost columns, we have plotted the spatial isolation of the poor $\check{q}_{\text{poor}}$ and of the wealthy $\check{q}_{\text{wealthy}}$. Dark patches signify areas that are exceptionally isolated. It is worth remembering that because the isolation index averages over neighbourhoods, a cell need not be inhabited for it to be colored in these plots.

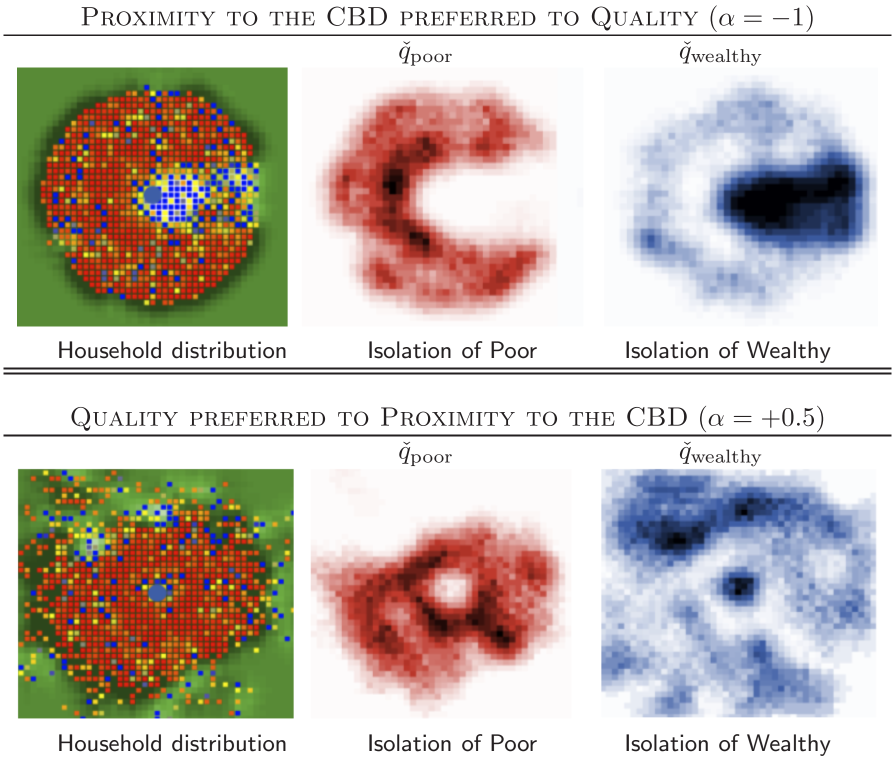

Figure 3. Simulation outcomes based on different values for the quality-index α.

We saw that the two parameters in our model, the quality-priority α and Pareto inequality index β, effect two different properties of the emergent city, namely, its shape and its fill. By varying them appropriately, we were able to recover both of the classic patterns observed in the Chicago school as well as other commonly observed phenomenon such as wealthy suburbs. The quality-priority index steers the exterior shape of the city. It determines whether the city is compact and circular or irregular and expansive. This effect is not completely unexpected. For large, negative values of α, inhabitants strongly discount cells far from the CBD. Consequently all households will seek out the available homes closest to downtown regardless of their income. Figure 3 shows examples of both compact and sprawling cities.

The inequality index controls the interior fill, roughly understood as whether the the city grows by rings or by sectors. Both are special cases of a pinching process. Initially very few households inhabit the city. Because all individuals prefer to some extent to live near the CBD, a ring of households circles the city's downtown. Because households either increase or decrease the quality and price of their neighbourhood proportionally to their income, areas with rich inhabitants will increase in price and quality, while those with poor inhabitants will decrease in price and quality without changing the cell's distance to the center of the city. These adjustments render cells adjacent to wealthy households more desirable for other

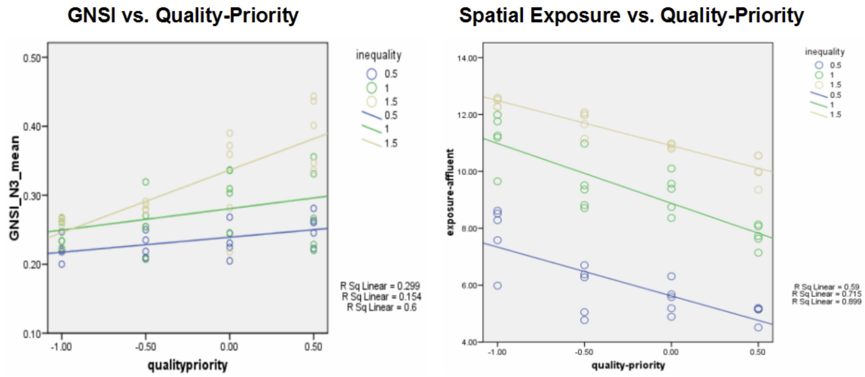

FIGURE 4. Segregation indices computed for simulation outcomes with different values for the quality-index α.

wealthy households, while at the same time making them prohibitively expensive for poor households. If the inequality β is sufficiently low, then there will be many wealthy households—enough to form a barrier to prevent being surrounded by poorer neighbors. The top row in Figure 3 illustrates the case in point. The group isolation plots sure the clear emergence of the sector pattern commonly seen in Latin American cities where wealthy families do indeed concentrate in areas that expand from the center in a single direction. If, however, inequality is very high the adjustment effects incurred by the overwhelming majority of households with lower than median incomes permit the wealthy downtown area to be enclosed entirely by poor individuals. Cities that fall in this part of the parameter space tend to resemble the Burgess model of concentric segregation. When inequality is high and preference is given to quality rather than distance, we witness the spontaneous emergence of wealthy suburbs. The bottom row in Figure 3 gives an example of this kind of city. In North American cities, where lower transportation costs decrease the importance of proximity to the CBD, we commonly experience precisely this pattern of urban development.

Figure 4 presents plots of segregation indices measured for simulations conducted with four different values for the quality-index α. In order to obtain more robust results, we repeated these measurements using five different random-seeds and three different degrees of inequality (presented in different colors). The first graphic shows how the index GNSI, which measures the segregation dimension evenness/clustering, presents a positive correlation with the quality-priority index. The second graphic shows how the exposure of affluent families to poor families decreases with the increase in the quality-priority index. Both results indicate that as soon as households attribute less value to proximity and more value to quality, the segregation level increased in both dimensions. Therefore, the 'disperse city' dominated by wealthy suburbs seem to be the segregation pattern that produces the highest levels of segregation. Regarding inequality, the graphics of Figure 4 reveal an unexpected result: while higher income inequality increases the segregation in the dimension evenness/clustering (GNSI), it increases the exposure amongst different income groups and therefore decreases the segregation in the dimension exposure/isolation.

5. Concluding Remarks

This paper presents an agent-based model that generates patterns of segregation much like those identified in real cities. These patterns emerge despite the extreme simplicity of the model and its parameters. It is an example of how complexity can emerge from just a few parameters and suggests that these parameters are crucial to the underlying dynamics of segregation.

The model demonstrates how the variation of a single parameter, such as quality-priority was able to generate classical patterns of segregation. This is an indication of how a cultural and economic environment influences segregation patterns. Low transportation costs in North American cities minimize the importance of proximity to the center of the city and promote the emergence of dispersed cities with wealthy suburbs. It will be interesting to see the effect of the current energy crisis on segregation in these cities.

This work includes a quantitative comparison of segregation patterns through the application of global segregation indices. This is a particular advantage provided by generative models, since indices computed for real cities cannot be compared due to their susceptibility to the modifiable areal unit problem (MAUP). Our findings indicate that disperse cities produce higher degree of segregation in both dimensions, while unequal cities present higher degree of segregation in the dimension evenness/clustering, but not in the dimension exposure/isolation.

Segregation remains a remarkably common, persistent and difficult problem for cities of all types. Like many social phenomena, it grows out of simple interactions between individuals, but manifests as a complex resilient phenomena at the scale of the city. Agent-based models provide a unique laboratory for experimentation and examination of these phemonema, stripped down to their essential mechanics. We view this work as a simple kind of null model upon which urban planners and policy makers may introduce and isolate the effects of legislation, e.g., the social-mixed housing policies that have been recently instated in some North American cities.

Acknowledgements

This work was partially supported by the Santa Fe Institute whose research and education programs are supported by core funding from the National Science Foundation and by gifts and grants from individuals, corporations, other foundations, and members of the Institute's Business Network for Complex Systems Research.

References

[Bur24] E. W. Burgess. The growth of the city: an introduction to a research pro ject. *Publications of the American Sociological Review*, 18:85–97, 1924.

[Cal00] T. Caldeira. *City of walls: crime, segregation and citizenship in São Paulo*. University of California Press, Berkeley, 2000.

[EA96] J. M. Epstein and R. Axtell. *Growing Artificial Societies: Social Science from the Bottom Up*. Brookings Institution Press, Washington, DC, 1996.

[Eps07] J. M. Epstein. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton University Press, 2007.

[Ft07] F. d. F. Feitosa and *et al.* Global and local indicators of urban segregation. *International Journal of Geographical Information Science*, 21:299–323, 2007.

[Gil08] N. Gilbert. *Agent-Based Models*. Quantitative Applications in the Social Sciences. Sage Publications, 2008.

[Hoy39] H. Hoyt. *The Structure and Growth of Residential Neighborhoods in American Cities*. Federal Housing Administration, Washington, DC, 1939.

[MD88] D. S. Massey and N. A. Denton. The dimensions of residential segregation. *Social Forces*, 67:281–315, 1988.

[MD93] D. S. Massey and N. A. Denton. *American apartheid: segregation and the making of the underclass*. Harvard University Press, Cambridge, MA, 1993.

[RO04] S. Reardon and D. O'Sullivan. Measures of spatial segregation. *Sociological Methodology*, 34:121–162, 2004.

[Sch78] T. C. Schelling. *Micromotives And Macrobehavior*. W. W. Norton, London & New York, 1978.

[St01] F. Sabatini and *et al*. Segregación residencial en las principales ciudades chilenas: tendencias de las tres últimas décadas y posibles cursos de acción. *EURE (Santiago)*, 27:21–42, 2001.

[Vc01] F. Villa¸ca. Segregation in the brazilian metropolis. paper presented at International Seminar on Segregation in the City, Lincoln Institute of Land Policy, Cambridge, 2001.