Urban Segregation with Housing Value Dynamics

Extending on Schelling's model of urban segregation

Minerva University
CS166 - Prof. Tambasco

Introduction of Model

This model extends upon the simple model for urban segregation proposed by Schelling (1969) to explore the relationship between urban migration patterns and the distribution of wealth over time. Schelling's model is a simple but effective way of showing how micro-scale preference for similar neighbors leads to the emergence of population segregation. However, the assumptions made in this model oversimplify the real social dynamics as there are factors beyond racial preference that determine urban migration patterns. Developing Schelling's findings of racially segregated equilibrium states that arise from racial preference, my model introduces a monetary variable into residential movement patterns. In integrating land value and house owner wealth information, this paper explores how racial similarity preferences and monetary variables influence each other. The guiding question of this study is how wealth dynamics are affected by social preference.

Model Structure

The model's macroscopic structure is a 2D nxn cellular automata grid containing two main pieces of information: housing info and house owner info. The housing information contains the following items: the state of the house (vacant/occupied) and its land value. The house owner's information contains information about the entity living in the house, including the following: their race (either pink or green) and their net value (as a random variable taken from a Pareto distribution). Since this model involves numerous discrete agents with individual attribute values (Sayama, 2015), this model can be described as an Agent Based Model (ABM).

Model Parameters

There are a number of parameters that inform the set-up of the system.

- *Racial preference intensity:* The preference for neighbors to be of the same race (0 = no preference, 1 = all neighbors must be of the same race).
- *Diffusion of land value:* How much land value of a house will be diffused to neighboring houses of less economic value (0 = no land value, 1 = entire land value).
- *Density:* The proportion of total cells that will be occupied by house owners
- *Utility threshold:* The acceptability threshold for an individual to occupy a certain cell before it desires to move.
- *Moving threshold:* The threshold for Monte-Carlo simulation to accept the possibility of a move.
- *Wealth difference:* The initialized difference in wealth between the different races (0 = no wealth difference, 1 = 100% wealth difference such that one race has on average, 200% of the other race).

Model Initialization:

The model is initialized with the parameters described above. Apart from these modifiable parameters, the simulation also initializes the grid assuming certain information about house owners and their homes.

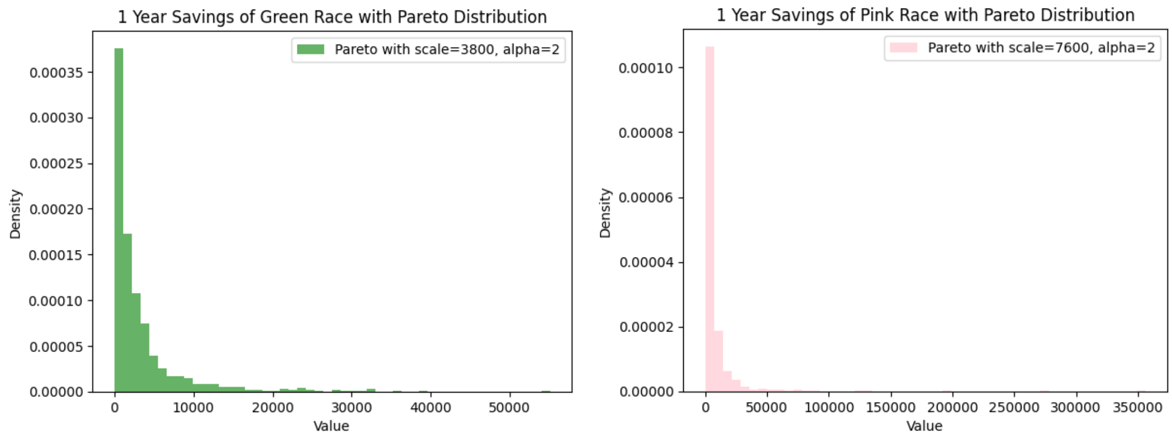

Figure 1. Wealth distribution from 1 year of income (assuming all income goes into savings) via a Pareto distribution. The green race, with a scale of 3800 has a mean income of $3277 and a median income of $1602. The pink race with a scale of 7600, has a mean income of $7171 and a median of $2728. These parameters have been initialized to best fit US census data from 1952 (Sawyer and Peel, 1952).

Homeowners will be defined in the grid with a density (default=0.7), and an equal likelihood of being of a green or a pink race. These imaginary races are defined to emphasize that this model is not validated and representative of the true dynamics of race. However, the attributes of these homeowners are, in fact, initialized with real data. The net wealth of homeowners is initialized via a Pareto distribution–a popular distribution for modeling economics (Tarascio, 1973), following the assumption that 80% of the population has 20% of the population–and historical US census data from the 1950s. From Table 3 of this 1952 census, we use info claims that the median net income of white families was ~$3100, and of blacks ~$1600. With the assumption that in order to own a home, a homeowner must have had to save for around 20 years, I initialize every individual's net wealth to the sum of 20 random variables from a Pareto distribution that follows the census data on net income (with pink=white, green=black). The home value is assumed to be 25-40% of the homeowner's 20 years of savings (DWA, 2021). If the house is initialized as a vacant home, it is assumed that the value of the home is between that of a green and pink homeowner's.

Model Update Rules:

The model rules are shown simplified in the flowchart in Figure 1. The simulation runs with the same update rules at each step (thought to be time intervals). For each step, 20 occupied cells in the grid are chosen randomly to check their utility and attempt a move according to its utility threshold–which is how much unhappiness the house owner can bear. To follow the principles of a Monte Carlo simulation, not every desire to move will result in an attempt to move, there must be a random draw, by which there is a 0.5 probability of attempting a move (according to the moving threshold). If approved to move, a random vacant house will be chosen and its utility compared against the house owner's current utility. The house owner will leave their current house for the new location if their utility will be higher in this move. After each move, the house values will be recalculated via rules of diffusion that follow home market assumptions that houses in richer areas will acquire value, while houses in less valuable neighborhoods will lose value. This simulation will be run for a sufficient amount of steps until the system stabilizes into equilibrium, as observed in Figure []. In this stable state, all house owners have satisfied their utility function to the best of their ability. Note that a house owner cannot move into a currently occupied home.

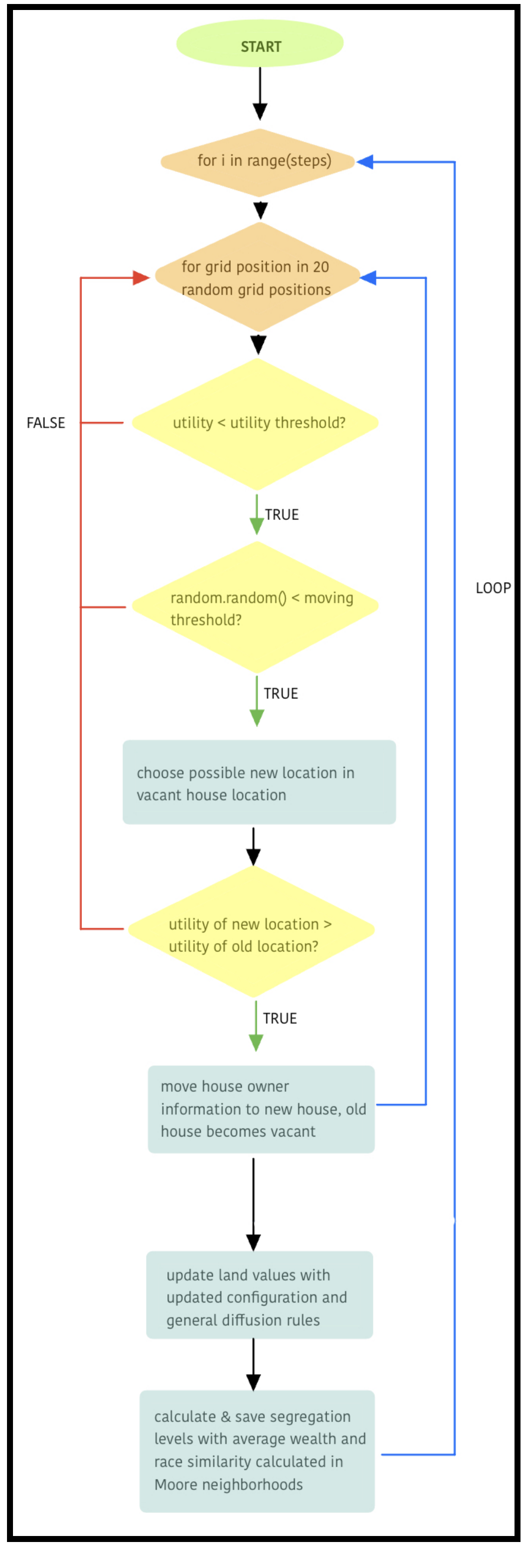

Figure 2. Flowchart of update rules

Utility Calculations

$$ideal\,land\,value\ =\ uniform(0.25,\ 0.4)\cdot net\,wealth \qquad \textbf{| EQ. 1}$$

$$race\,U\ =\ \begin{pmatrix} 1\,if\ \frac{\#\,of\,same\,race\,neighbors}{total\,\#\,of\,neighbors} \geq racial\,preference\,parameter \\ 0\,otherwise \end{pmatrix} \qquad \textbf{| EQ. 2}$$

$$land\,value\,U\ =\ \begin{pmatrix} 1\ -\ \frac{|average\,neighbor\,land\,value\,-\,ideal\,land\,value|}{ideal\,land\,value}\ if\ \frac{|average\,neighbor\,land\,value\,-\,ideal\,land\,value|}{ideal\,land\,value} < 0.5 \\ 0\,otherwise \end{pmatrix} \qquad \textbf{| EQ. 3}$$

$$U\ =\ race\,U\ +\ land\,value\,U \qquad \textbf{| EQ. 4}$$

Diffusion Calculation

$$\begin{aligned} updated\,land\,value &= current\,land\,value\ -\ diffusion\,rate\cdot(current\,land\,value\ -\ average\,neighbor\,land\,value) \\ &\quad if\,current\,land\,value\ \geq average\ neighbor\,land\,value \\ &= current\,land\,value\ +diffusion\,rate\cdot(average\,neighbor\,land\,value -\ current\,land\,value) \\ &\quad otherwise \end{aligned} \qquad \textbf{|EQ. 5}$$

Assumptions

Although this paper aims to provide a more realistic view of urban movement and wealth distribution, this model has a few key assumptions that must be addressed.

- *Simplification of land value dynamics:*
  The dynamics of land value diffusion are simplified to assume that if a house is surrounded by neighbors with a higher average land value, the house increases in value, depending on the average of the surrounding neighbors. For these houses of greater than average land values, their value will regress towards the mean.
- *Homogenous utility functions:*
  Each “house owner” will follow the same utility function to determine its happiness level with its current neighbors. In reality, different individuals will have different preferences when determining their movement.
- *Assuming general wealth differences in races:*
  One of the parameters in my model will account for racial differences in wealth. This difference in average wealth is informed by true historical information on racial economic discrimination. However, given that this is a parameter, we will be able to analyze the true effect of such an economic disadvantage.
- *Static net value/wealth levels:*

In the model, we assume that the wealth of an individual homeowner (measured as net wealth) is static. However, in real-world situations, we understand that homeowners can gain net value through events such as job promotions and inheritance, and lose through the loss of a job or bad spending habits.

- *Land usage assumptions:*
  We are assuming that all land in the grid represents a house, and that land can only take on the label of vacant or occupied when in reality, land can be used for many other reasons (eg. schools, roads, etc).
- *Proportion of Population Updates:*
  At each time step, 20% of total house owners will randomly be chosen to evaluate their happiness level and move accordingly. This is loosely following historical census data that shows that 20% of the population in the US between the 1950s-1960s had moved within the last year (Cohn & Morin, 2008).
- *Moore neighborhood with cut-off boundaries:*
  All calculations in this model will be working with a Moore neighborhood. Individuals will only be looking at their direct 8 (or less) possible neighbors and comparing their similarities with them.

Outputs / Measurements

This paper will focus on analyzing a feature of this model under empirical analysis and supported by theoretical analysis:

> *Q1: "How do individual racial preferences affect the macroscopic distribution of race and wealth over time and at stability?"*

To answer this question, the simulation will be run for different racial preference parameter values between 0.1 and 0.9. The evolution of the system will be measured in terms of the overall racial and wealth similarity of neighborhoods, which directly relates to the segregation of both. A higher average similarity of agents in a neighborhood indicates greater clustering and segregation. These behaviors will first be predicted by theoretical analysis and validated by empirical analysis.

Theoretical Analysis

Introduction

Before delving into a theoretical analysis, it is valuable to understand the limitations of analyzing an ABM mathematically. While Sayama discusses the merits of ABM in depth, including a discussion of Schelling's model of urban segregation, he explicitly writes that such complex models lose analytical tractability, therefore "it is generally not easy to conduct an elegant mathematical analysis of an ABM" (429). Additionally, Antosz et al., discuss the analysis and building of "theory" in the Schelling model to consist mainly of replicating previous study results, as well as validation through real-world situations (2023).

Although there are significant limitations in mathematically analyzing ABM, we can still analyze elements of the model to better understand its overall behavior and validate empirical results. We will predict the long-term or steady-state behavior of the system by extrapolating for the probabilities of an individual having certain utilities with different parameters and conditions.

Predictions for behavior according to parameter values

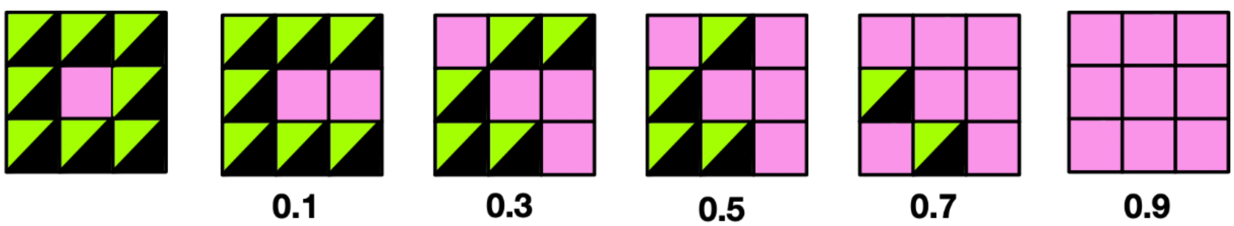

Figure 3. A visual representation of the states required to reach race_utility = 1 for each racial preference threshold in a Moore neighborhood with the center pink cell representing a house owner of the pink race. The far left represents this cell with no neighbors of the same race (green/black cell represents either a house occupied by green race homeowner, while black represents a vacant house), then subsequent grids show the minimum state of neighbors for race_utility to change from 0 to 1. For race_utility = 0.1, in an 8 neighbors model, only 1 neighbor must be of the same race, while for race_utility=0.9, all neighbors must be of the same race.

We begin by theoretically predicting the behavior of a cell depending on a racial preference parameter. Figure 3. displays the required conditions to satisfy a racial utility of 1, following the EQ.2. Currently, we ignore the utility function gained from the land value similarity and assume that any individual will be urged to leave their current cell state if the racial utility is equal to 0, and otherwise, it will stay in its current space. Knowing the number of neighbors required to have racial utility equal to 1 in an 8-neighbor model, we can calculate the probability of an individual agent desiring to move/star. Although the visualization in Figure 3.

shows the case for an individual belonging to the pink race, since these probabilities are symmetric–there is an equal probability of a cell being green or pink–we are calculating the probability of moving for both pink and green house owners alike. These probabilities are listed in Table 1, determined by a binomial distribution. The random variable X is represented as the number of neighbors of the same race.

Racial Preference Parameter	**Same Race Neighbors (8 Neighbors)**	**Expression of Probability that Racial Utility = 1 (with grid occupation density = 0.7)**	**Probability of racial utility = 1**
0.1	1	$P_{racial\ utility = 1} = P_{\geq 1\ same\ race\ neighbor}$ $P(X \geq 1) = 1 - P(X = 0)$ $P(X = 0) = \sum_{i=0}^{8} (0.5 \cdot 0.7)^i \cdot (1 - 0.7)^{8-i}$ $P(X \leq 1) = 1 - \sum_{i=0}^{8} (0.5 \cdot 0.7)^i \cdot (1 - 0.7)^{8-i}$	0.999
0.3	3	$P(X \geq 3) = \sum_{i=3}^{8} P(X = i) = \sum_{i=3}^{8} \binom{8}{i} (0.7 \cdot 0.5)^i (1 - (0.7 \cdot 0.5))^{8-i}$	0.572
0.5	4	$P(X \geq 4) = \sum_{i=4}^{8} P(X = i) = \sum_{i=4}^{8} \binom{8}{i} (0.7 \cdot 0.5)^i (1 - (0.7 \cdot 0.5))^{8-i}$	0.294
0.7	6	$P(X \geq 6) = \sum_{i=6}^{8} P(X = i) = \sum_{i=6}^{8} \binom{8}{i} (0.7 \cdot 0.5)^i (1 - (0.7 \cdot 0.5))^{8-i}$	0.025
0.9	8	$P(X = 8) = (0.7 \cdot 0.5)^8$	0.0002

Table 1. A theoretical analysis of an individual's racial utility with different racial preference parameters and an occupation density of 0.7.

From these probabilities, we can intuitively understand that as the probability of a racial utility value of 1 decreases, the probability of an individual cell attempting to change its current position will go up. This inverse relationship can be analyzed further as we conduct an empirical analysis of the system. However, even with such values, we can already begin to predict the behavior of the system. Since every agent is making decisions based on the utility of their current position, extremely low racial parameters such as 0.1 in a random distribution of neighbors will satisfy the utility 99% of the time. This indicates that very low racial preferences will lead individuals to stay put, and to lead to a stable state of heterogenous neighborhoods. Opposingly, slightly higher racial preference values such as 0.3 and 0.5 will be less likely to have neighborhood conditions leading to racial utility values of 1. This individual “unhappiness” will lead them to attempt moves more frequently, and optimize for their utility, leading to more segregated outcomes. Finally, high racial preferences such as 0.7 and 0.9 are much rarer, with the state of racial utility = 1 being highly improbable. This will lead individual agents to move frequently, without ensuring that “happy” conditions may even be met, because if every agent is extremely picky, and continues to move, the state may never reach a steady state. We can predict that this will lead to low racial segregation in the long term. Let’s now include a brief discussion of the role of land value utility and preferences.

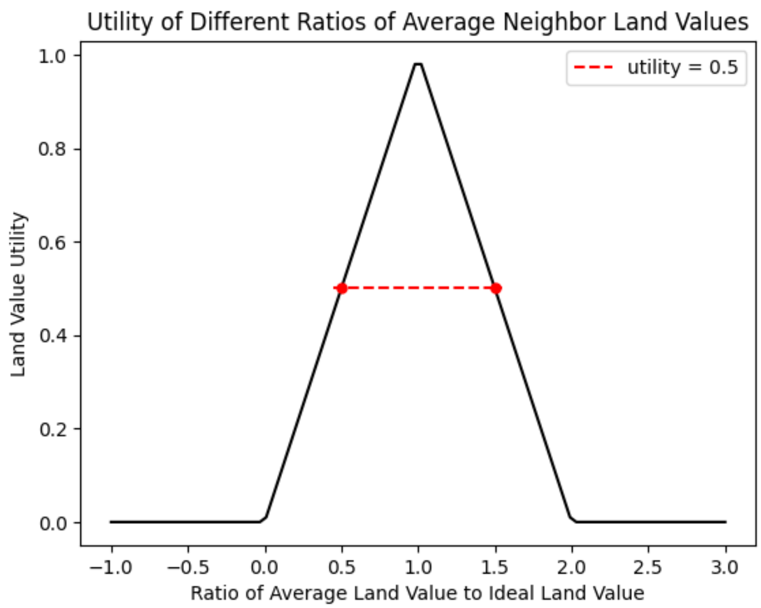

Figure 4. Land value utility for different ratio values of the average neighborhood land values compared to ideal land value. Observe that in order for land value utility to be equal to, or greater than 0.5, the average neighborhood land value must be between 0.5-1.5 times the ideal land value. Average neighborhood land values of over a ratio of 2 will lead to land value utilities of 0.

From Figure 4, we can see the relationship between the ratio between average neighborhood land values and the ideal land value and land value utility. With a utility threshold of 1.5, we understand from the previous discussion that the land utility value must be above 0.5 and the racial neighborhood preferences must be met. For the land utility value to be above 0.5, the average neighborhood land value must not be over 0.5 times greater or less than the ideal land value.

Simulation and Results

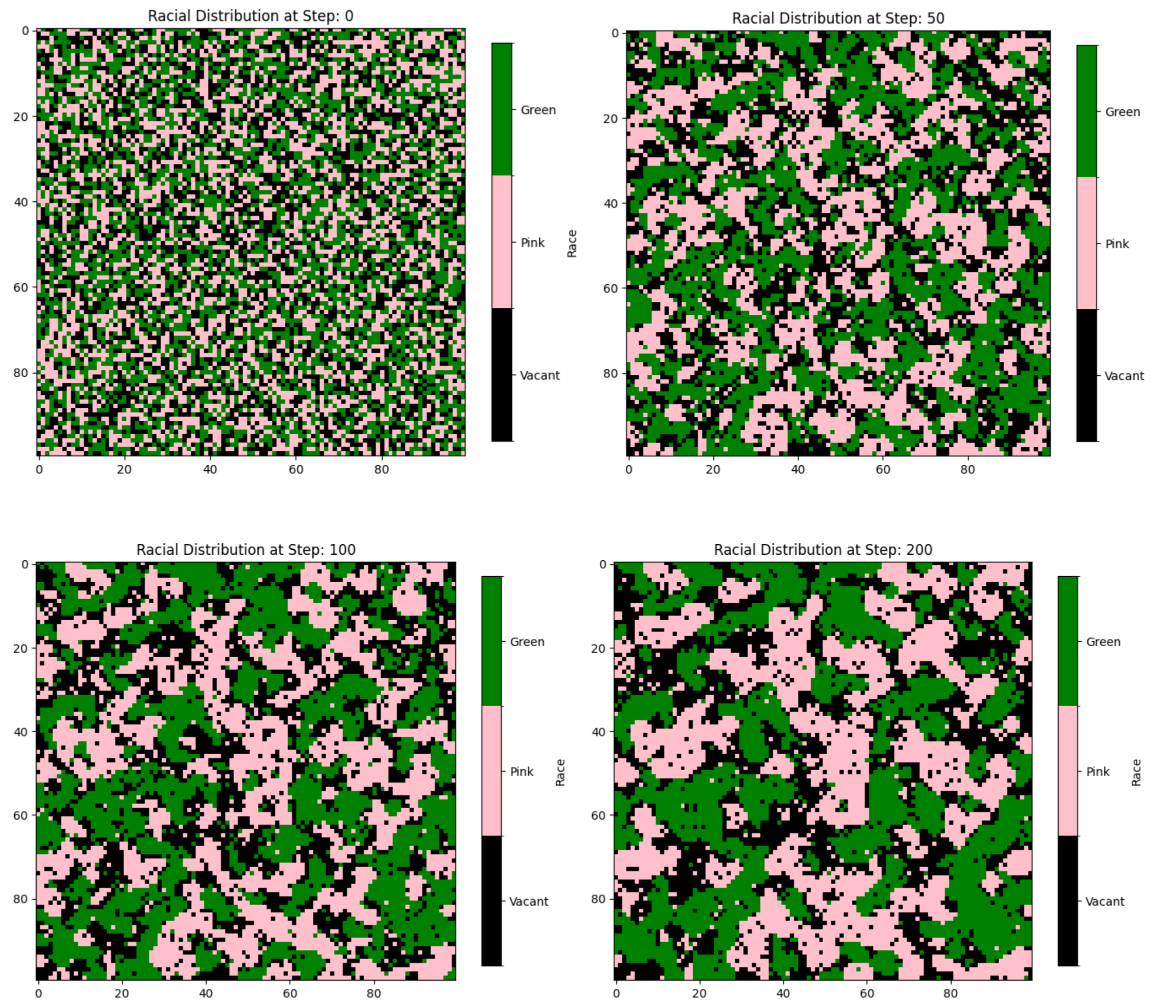

Figure 5. Visualization of the distribution of race on a 100X100 grid every 50 steps with a racial preference of 0.5, wealth difference of 2, and utility threshold of 1.5. Observe that clustering behavior occurs as the system develops.

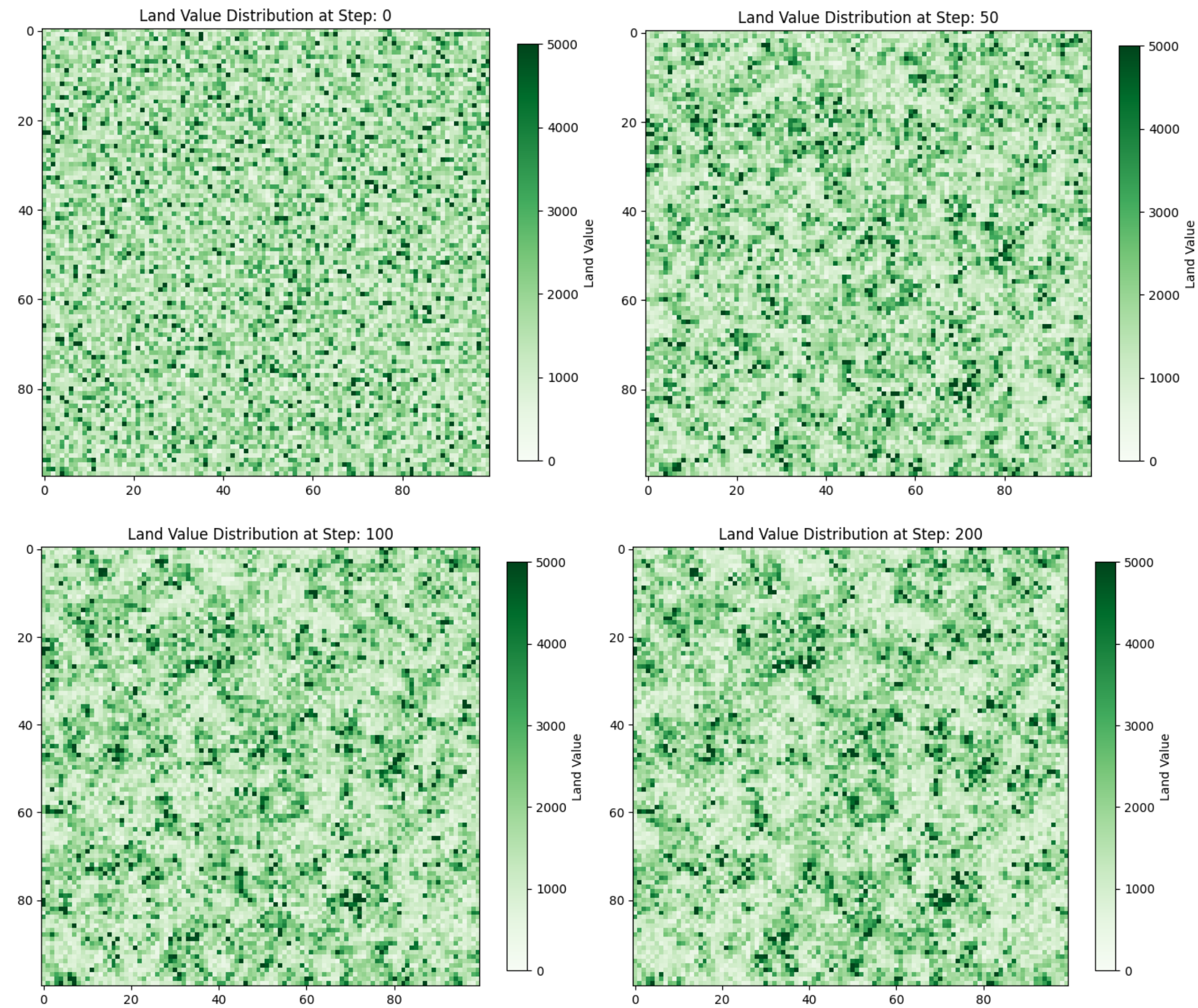

Figure 6. Visualization of the distribution of land values on a 100X100 grid every 50 steps with a racial preference of 0.5, wealth difference of 2, and utility threshold of 1.5. Darker green values represent higher land values. Observe that as the land values diffuse, compared with figure 8, areas occupied by pink populations are generally more prosperous.

Checking for stabilization

We first determine the general steps until convergence for different utility thresholds. This tells us how many steps are generally required until the system is generally stable. From Figure 7. we see that for most utility thresholds, the system stabilizes after 200 moves to under 0.5, apart from a utility threshold of 2. I have chosen to conduct my further analysis with a utility threshold of 1.5 in this case.

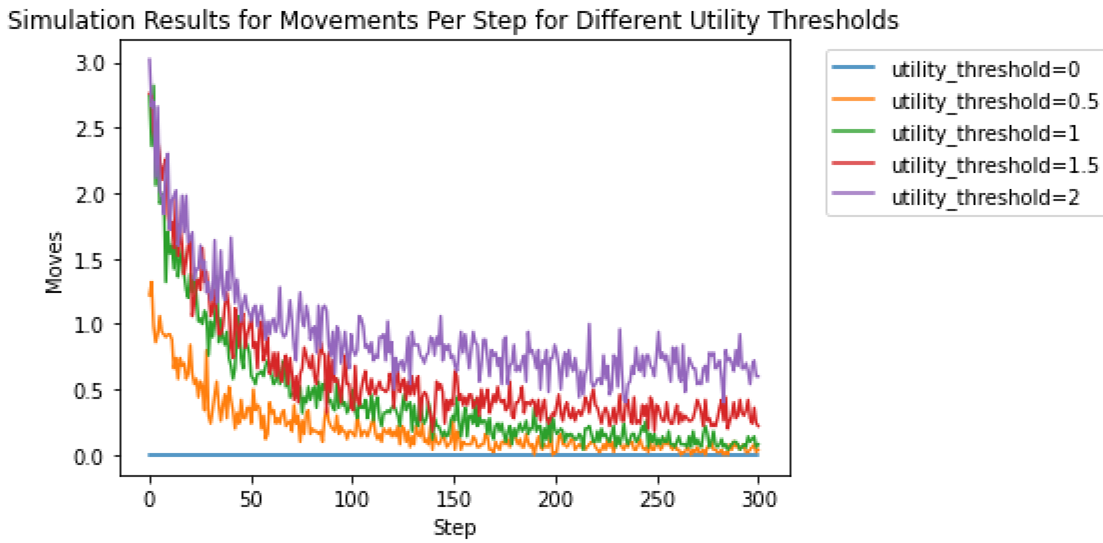

Figure 7. Parameterized with a wealth difference of 0.5, and racial preference of 0.5, a grid of 10x10, 50 trials, and 300 steps.

Checking for equilibrium states with differing racial preferences

Figure 8 displays the average neighborhood race and land value similarities for different racial preference parameters over the course of 300 steps.

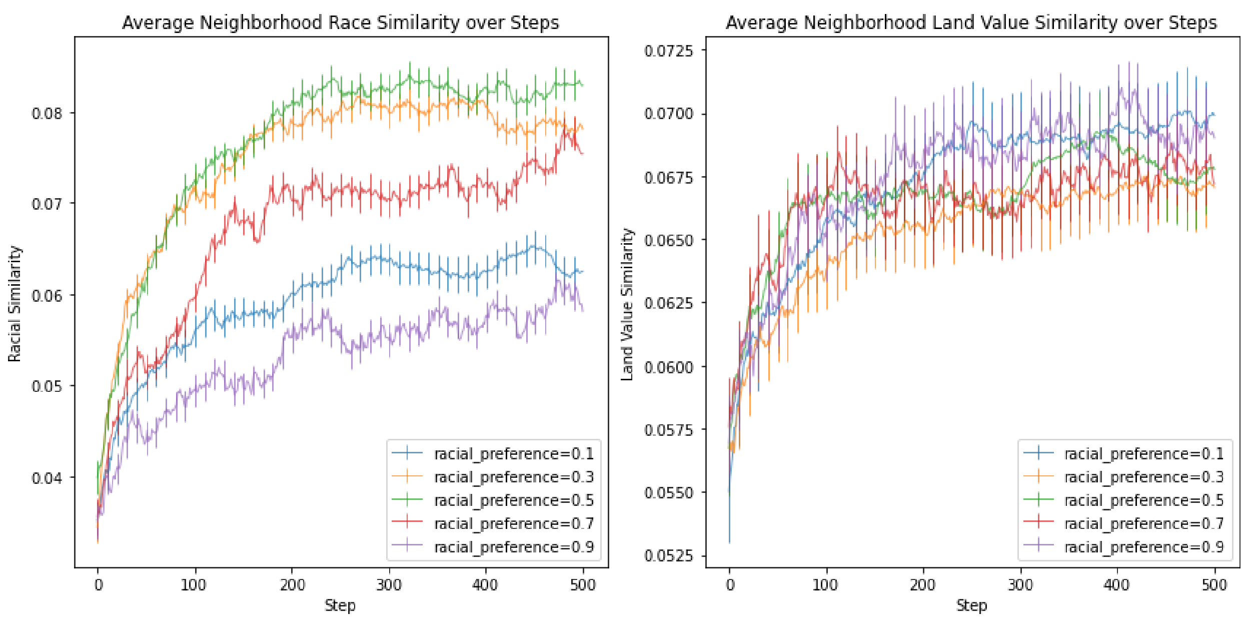

Figure 8. Graph of the racial and land value similarity over 300 steps for different racial preference parameters. The graphs include a 95% confidence interval. In the average neighbor race similarity, a racial preference of 0.5 leads to the greatest racial similarity (leading to the

greatest racial segregation), while a racial preference of 0.9 leads to the least racial similarity. Opposingly, the wealth similarity is greatest for a racial preference of 0.9 and lowest for a racial similarity of 0.5.

Comparing these results with our theoretical predictions, indeed, we observe that racial similarity is higher for medium racial preferences of 0.3 and 0.5, as we predicted that these values will prompt an agent to move until satisfaction, until a racial preference of 0.1, which is likely for an individual to be already satisfied. Also, as predicted in the probability of a utility value of 1, a racial preference parameter of 0.9 will result in the lowest racial similarities. This is likely because of the difficulty in reaching a satisfactory state, so all agents chosen will move with every chance they have, leading to an overall unstable state. This validates our model further.

Analysis of Uncertainty

Racial Preference	**Average 95% CI for Race Similarity**	**Average 95% CI for Land Value Similarity**
0.1	0.0113	0.0017
0.3	0.0142	0.0018
0.5	0.0171	0.0019
0.7	0.0145	0.0017
0.9	0.0115	0.0018

Table 2. Table of 95% Confidence Intervals for the racial and land value similarity for varying racial preferences. These are the results for 100 trials of 300 steps.

Table 2 represents the level of uncertainty between trials of determining race and land value similarity with different levels of racial preference. Analyzing these values, we see that for both values are relatively small, indicating low uncertainty. The average CI for land value and race similarity remains relatively stable between different parameter values for racial preference. In order to narrow the CI further, it is suggested to run the simulation for more than 300 steps, as not all systems may have reached equilibrium

Data-Driven Advice for Urban Policymakers

Looking at many urban landscapes and their respective demographics, it may be surprising to notice that many such environments are incredibly racially segregated. While urban policies, such as redlining in the United States, have contributed to such segregation, it is important to consider the self-emerging properties of such a phenomenon and its consequences.

Microscope individual preferences for neighbors of the same race as you can lead to large-scale segregation, as can be observed in Figures 5 and 8. Besides racial segregation, this has economic impacts on communities, as races that have historically been economically disadvantaged can remain without proper economic integration in their urban environments. As can be observed from Figure 6, communities of great wealth, or poverty arise. With this knowledge, how might policies be targetted at improving the overall economic health of the subcommunities of your urban environment, as well as the entire urban environment?

One such suggestion is to use spatial data to identify these underserved communities and begin balancing the wealth and land value gap by introducing further economic activity and engagement. Rather than simply pouring money into such communities, all policies for the better integration of races and economies must focus on decreasing the degree of segregation preferences by making all sections of the city affordable and attractive to all. This may include providing subsidies for business, transportation, and education. It is important to understand, however, the complexity involved in predicting behaviors in an urban environment, therefore, understanding the cultural relevancies will improve your chance at successfully producing beneficial urban change.

AI Usage Statement:

I used ChatGPT in my assignment when determining how to code and display CIs, error bars, and cut-off boundaries.

WORD COUNT: 3500 words

References

Antosz, P., Birks, D., Edmonds, B., Heppenstall, A., Meyer, R., Polhill, G., O'Sullivan, D., & Wijermans, N. (2023). What do you want theory for? - A pragmatic analysis of the roles of "theory" in agent-based modelling. *Environmental Modelling and Software*, *168*, 105802. https://doi.org/10.1016/j.envsoft.2023.105802

Cohn, D. (2020, May 30). *Who moves? who stays put? Where's home? | Pew Research Center*. Pew Research Center's Social & Demographic Trends Project. https://www.pewresearch.org/social-trends/2008/12/17/who-moves-who-stays-put-wheres-home

DWA. *What percentage of net worth should be in real estate?* (2021, June 2). Delta Wealth Advisors. https://deltawealthadvisors.com/blog/what-percentage-of-net-worth-should-be-in-real-estate

Sawyer, C., & Peel, R. V. (1952). Current Population Reports. *U.S. Department of Commerce*. https://www2.census.gov/library/publications/1952/demographics/p60-09.pdf

Sayama, H. *Introduction to the Modeling and Analysis of Complex Systems - Open Textbook Library*. (2015). Open Textbook Library. https://open.umn.edu/opentextbooks/textbooks/233

Schelling, T. "Dynamic models of segregation,". (1971). Journal of mathematical sociology, vol. 1, no. 2, pp. 143-86, 1971

Tarascio, V. J. THE PARETO LAW OF INCOME DISTRIBUTION on JSTOR. (1973). *www.jstor.org*. https://www.jstor.org/stable/42859204