EXPANDING THE SCHELLING AGENT-BASED MODEL: *DE JURE* SEGREGATION DYNAMICS ON A NETWORK

A Thesis

Submitted to the Faculty

in partial fulfillment of the requirements for the

degree of

Bachelor of Arts

in

Mathematics

by

Love Weiai Tsai

DARTMOUTH COLLEGE

Hanover, New Hampshire

May 2023

Abstract

Racial segregation is a persistent problem in the United States. Therefore, understanding how it arises (and how it might be ameliorated) is of top priority. The 1971 Schelling segregation model was a key breakthrough in agent-based modeling for this sociological phenomenon. Unfortunately, it only accounts for self-segregation due to individual in-group preference, making it a strictly *de facto* segregation model. Research into recent American history reveals that this approach is inaccurate and ineffectual at capturing segregation dynamics because of institutional policies that affected how people moved in the 20th century. Additionally, the Schelling model is done on the grid whereas most metropolitan cities are better approximated as various types of networks; limiting an analysis on segregation to a grid-like graph is therefore unrealistic.

In this thesis, we expand on the *de facto* Schelling segregation model by adding a *de jure* component through the modeling of redlining as implemented by the Federal Housing Authority in the National Housing Act of 1934. We first explore such dynamics as a cellular automata process before moving to networks. In doing so, we not only evaluate the effects of institutional policies and structural racism on individual action but we also evaluate how the choice of residential layout changes steady-state segregation. We find that 1) urban layouts and densities themselves have important implications for both types of segregation, 2) the *de jure* segregation mechanism of redlining significantly affects segregation as it manifests both on the grid and the network, and 3) *de jure* redlining is sufficient to result in mass movement even without *de facto* preferences.

Acknowledgments

I would like to thank my thesis advisor, Professor Peter Mucha, for his teaching, advice, and support during the research and writing process. Without his enthusiasm for the subject and his encouragement, none of this would have been possible. I would also like to thank the Department of Mathematics at Dartmouth College for providing the environment that helped me discover and cultivate my curiosity for math even as a pre-health student. The supportive and inclusive environment was critical for me to see a place for myself in the discipline.

Lastly, I would like to thank my family and loved ones for all their support and confidence in me. I am forever grateful for the life I've been given and the people who have surrounded me with their kind words and encouragement. I will do my best to pay it forward in the future.

Table of Contents

Abstract.....

Acknowledgments.....

Chapter 1: Introduction..... **1**

1.1 Problem Description..... 1

1.2 Background..... 3

1.2.1 Mathematical Modeling, Agent-Based Models, and Cellular Automata..... 3

1.2.2 The 1971 Schelling Model..... 6

1.2.3 De Jure Redlining..... 7

1.2.4 Network and Graph Theory..... 9

1.2.5 Networks to Model Urban Spaces and Interpersonal Relationships..... 10

Chapter 2: Methods..... **15**

2.1 The Schelling Model..... 15

2.2 Operationalizing FHA Redlining..... 17

2.3 Methods of Simulation Analysis..... 20

2.3.1 Measuring Segregation..... 20

2.3.2 Segregation's Impact on Health Disparities..... 22

2.4 Cellular Automata Simulations..... 23

2.5 Network Simulations..... 23

2.6 Mean Field Approximations..... 26

Chapter 3: Results and Analysis..... **28**

3.1 Modified Schelling Model on the Grid..... 28

3.1.1 Dissimilarity Index..... 28

3.1.2 Isolation Index..... 30

3.1.3 Construction of Qualitative Phase Diagrams..... 31

3.2 Modified Schelling Model on Networks..... 33

3.2.1 Network Generation and Preliminary Visualizations..... 33
3.2.2 Network and Segregation Evolution Per Time Step: Initial Redlining.....38
3.2.3 Network and Segregation Evolution Per Time Step: Imposed Redlining..... 42
3.2.4 Regular DLA vs. Clustered DLA.....47
3.2.5 Average Mixity Index for Schelling vs. Redline Model..... 48
3.2.6 Role of Density in Redlined-Segregation..... 50
3.3 Mean-Field Approximation..... 51
3.3.1 Approximations Plotted Per Time.....51
3.3.2 Role of Density in Redlined Segregation: Mean-Field Version.....54
3.3.3 Role of Probability in Redlined Segregation..... 55
Chapter 4: Conclusion.....56
References.....62

Chapter 1: Introduction

1.1 Problem Description

Segregation has been a long-standing problem in the United States. The phenomenon of segregation as used in this thesis can be defined as the "separation of groups with differing characteristics" [1]. In practice, it exists whenever two or more groups, distinct with respect to certain identities, are separated in space. What qualifies as a “distinct” identity differs according to context: in ethnically homogenous areas, for example, segregation may occur between groups differentiated by socioeconomic status (SES) or religion. In many parts of the United States, a prominent historical differentiator is race, particularly between whites and Blacks. Today, racial segregation is slightly more complicated because of an increasingly multiracial society, although a similar distinction exists between whites and people of color.

For applied mathematicians, understanding social phenomena through modeling is important for not only advancing the discipline but also contributing to the discovery of key insights about human behavior. Due to its pervasive and persistent nature, race-based segregation deserves attention in math research today.

There are numerous explanations for why race-based segregation persists in the United States. One of the most prominent mathematical explanations is the Schelling model [4]. In 1971, Thomas Schelling revolutionized the mathematical modeling of segregation by creating an agent-based model (ABM) that sought to explain racial segregation on a grid with only a few rules pertaining to individual discriminatory behavior. Through his work, Schelling showed how even relatively mild in-group preference could contribute to large-scale racial isolation. As such,

this model is an example of *de facto* segregation—that is, segregation that exists "by fact" and not because of any policies or laws. In the four decades since then, this model has remained a prime example of how simple algorithms and ABMs can approximate complex social processes—today, researchers still commonly use his framework to investigate cases of real-world segregation [40, 41, 42].

However, there are limits to the Schelling model. Indeed, even Thomas Schelling himself noted this in his 1971 paper, "Dynamic Models of Segregation," when he states that his approach omits two, potentially more insidious contributors to segregation in favor of in-group preference: organized action and economic pressure [5]. So while it may be tempting to believe that most racial isolation comes from these individual prejudices, recent discourse reveals that the bulk of segregation shaping America is influenced by intentional, *de jure* ("by law") governmental policies, not personal preference [3]. Richard Rothstein, author of National Book Award finalist *The Color of Law: A Forgotten History of How Our Government Segregated America* even ventured so far as to call *de facto* segregation a complete myth [12]. No matter the extent of *de jure* segregation's role in today's America, inattention to structural racism and the dismissal of the federal government's hand in perpetuating racial segregation is one weakness in the Schelling model that should be addressed.

Why should one expand the Schelling model? It is important to note that mathematical models are useful insofar as they accurately represent real-life phenomena and allow researchers to make educated decisions based on the resulting simulations. Such an approach like Schelling's therefore poses a problem for those who hope to use the theory to propose new policy since its findings may not be 100 percent relevant to real-life situations. In this thesis, we expand the Schelling model to include a prominent *de jure* segregation mechanism: redlining as

implemented by the Federal Housing Authority (FHA) starting with the agency's creation in 1934. Furthermore, we move this so-called “redline-Schelling model” from the cellular automata case onto a variety of networks to better understand segregation as it propagates through areas with diverse residential layouts. After all, very few cities are laid out in uniform blocks like New York City; using network analysis allows us to see if spatial layout matters for segregation, particularly in the instances of cliques and other types of hierarchies that only arise from a network. Lastly, we investigate the solitary role of *de jure* mechanisms on global segregation through a mean field approximation of the entire system.

The purpose of this thesis is to model a more holistic and accurate version of segregation that takes into account both *de facto* and *de jure* mechanisms through operationalizing FHA redlining on both the grid and more realistic experimental spaces.

1.2 Background

1.2.1 Mathematical Modeling, Agent-Based Models, and Cellular Automata

The 1971 Schelling model is an agent-based, cellular automata model [4, 8]. In this thesis, we are concerned with expanding this mathematical model to approximate real-world phenomena, namely *de jure* racial segregation driven by redlining in the 20th century United States.

Mathematical modeling as a research approach allows one to reduce complex processes into their most important (and most simple) component parts. As such, all models are simpler characterizations of their subjects instead of mirror copies. There is often no one "right" answer for operationalization, although mathematicians should be prepared to explain their modeling choices [6, 7]. A certain type of ambiguity exists in applied math modeling of complex social

systems, then, because there are multiple ways to address the same problem and consequently multiple potential solutions. Nonetheless, mathematical models are useful because they allow researchers to understand how steady states come to be, running experiments that would otherwise be impossible either due to a lack of resources or ethical complications. In an ideal world, such research can lead to predictions that then inform decision-making.

In ABMs in particular, one creates a world with agents that are given guidelines for how they interact with one another and with the experimental space [7]. At every time step, each agent will independently assess its state and act according to the specified rules. Depending on the experiment, multiple layers of conditions can be added to the model, making it more or less complex. The researcher is often not interested in what individual agents do but what emergent/global behavior arises from the sum of all agent actions. For example, modeling traffic jams, a financial crash, or biological development are all readily accomplished with ABMs in which the agents are cars, people with bank accounts, and cells, respectively. In each case, every agent's independent action will contribute to the final result. ABMs therefore lend themselves very well to algorithmic implementation on computers because of their repetitive and sequential nature.

ABMs are useful because they allow for the capturing of emergent behaviors, multiple layers of complexity, and group behavior without directly coding for them. However, one limitation is that setting rules for ABMs implies that the agents will always act the same way, hold the same values, and act rationally. Such assumptions may not be accurate, especially when modeling social systems where humans are concerned. In these instances, stochasticity can be added back to the system by inducing randomness at various steps of the algorithm (such as the

assessment or decision-making stage). On the whole, ABMs are extremely useful for mathematical modeling.

The Schelling model is an ABM. It is also an example of cellular automata (CA) modeling. CA is a branch of mathematical modeling that is even older than ABMs, dating back to the 1940s [9]. In these models, an *n*-dimensional grid holds a set of cells that embody a certain state and simultaneously update these states at each time step. In this way, CA are similar to ABMs. However, CA "agents" are bound to the cells of the grid and normally only look at their immediate neighbors during the assessment stage whereas agents in ABMs may have more flexibility in movement and/or utility evaluation. Two common neighborhoods for utility evaluation are the von Neumann and the Moore neighborhoods. A von Neumann neighborhood is diamond-shaped with $2r(r + 1) + 1$ cells. Its range *r* is given by the equation

(1) $$N^{v}_{(x_0,y_0)} = \{(x, y) : |x - x_0| + |y - y_0| \leq r\}.$$

The Moore neighborhood, on the other hand, is square-shaped with $(2r + 1)^2$ cells and has a range *r* given by the equation

(2) $$N^{M}_{(x_0,y_0)} = \{(x, y) : |x - x_0| \leq r, |y - y_0| \leq r\}.$$

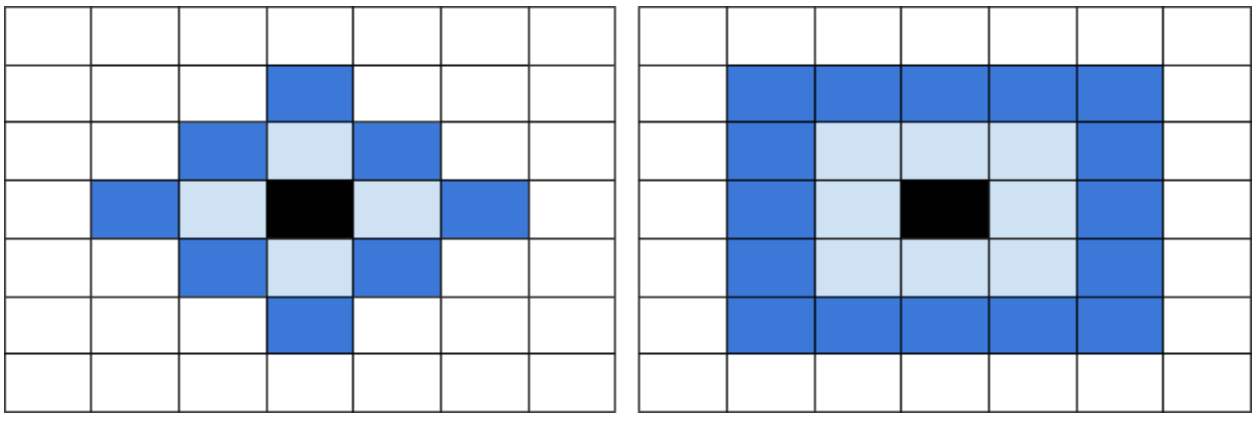


Figure 1. Example of von Neumann vs. Moore Neighborhood in 2-D.

It is the assessment of this local neighborhood that dictates what the "agent" in any given cell will do. Boundary conditions are usually either periodic or non-periodic; if periodic, the neighborhood will wrap around to the other side of the environment.

A popular two-dimensional CA example is Conway's Game of Life [10]. This simplistic model has cells that can either be alive or dead; therefore, at each time step, a cell can die, be born/come alive, or stay in the exact same state. The conditions for each action are distinguished by the quantity of living cells in the immediate neighborhood. For a cell to die, for example, one of two things must happen: either over- or under-population in the immediate neighborhood. In models run with the Moore neighborhood distinction, overpopulation occurs if more than four neighboring cells are alive and underpopulation occurs if less than two neighboring cells are alive. Conway's Game of Life is a deterministic-rule CA whereas the Schelling model has a stochastic rule in the evolution step, making dynamics from the Schelling model more random.

1.2.2 The 1971 Schelling Model

In 1971, Thomas Schelling became one of the first people to create a cellular automata ABM for any application, much less segregation. The Schelling model can be referred to as a "spatial proximity model" and attempts to mathematically replicate residential racial segregation, the details of which are outlined in the Methods section. Through it all, Schelling's work utilizes *de facto* ("by fact") segregation, which, as detailed in the introduction, is a legal term that describes when individuals self-segregate due to personal prejudice. There is no external factor impacting the agents' decisions. Schelling's model revealed that surprisingly significant amounts of racial isolation could happen even with comparably low levels of in-group preference. Additionally, more recent mathematical research has found that a certain level of segregation will always happen under this model, regardless of the initial state or agent preferences [11]. These

findings, combined with the simple set-up of an otherwise complex process, made his work a ground-breaking step in the field and a beginning answer to the question of why segregation can be such a hard problem to solve.

1.2.3 *De Jure* Redlining

The Schelling model is an optimal approach to modeling *de facto* segregation. However, segregation in the United States was not a strictly *de facto* phenomenon. In his book, *The Color of Law*, Richard Rothstein outlines the legacy of governmental segregation in the 20th century, making the claim that "today's residential segregation in the North, South, Midwest, and West is not the unintended consequence of individual choices and of otherwise well-meaning law or regulation but of unhidden public policy that explicitly segregated every metropolitan area in the United States" [12]. That is, *de jure* mechanisms played as much a role (if not more) in American segregation as *de facto* mechanisms did, rendering the Schelling model at best incomplete and at worst deceptive.

There are many ways that the U.S. government perpetuated segregation in the 20th century. Numerous scholarly sources have already identified federal redlining as a major component of *de jure* segregation [12]. In this thesis, we choose to highlight redlining as instituted by the Federal Housing Authority (FHA). The federal government also used another agency—the Home Owners' Loan Corporation (HOLC)—to encourage segregation, but we will focus on the FHA's implementation in our model because research by the Federal Reserve Bank of Chicago has shown that the FHA maps were of greater historical and practical import in individual movement than the HOLC's [13].

For some brief historical background, redlining was instituted in the United States by these two federal housing agencies—the FHA and the HOLC—in the early 1930s. The HOLC

was created by Congress in 1933 to help maintain housing stability by refinancing loans at risk of defaulting. The FHA was created one year later by the National Housing Act of 1934. Its purpose was to reform mortgage practices and encourage residential construction, which had largely slowed to a halt during the Great Depression, by making credit more accessible to homebuyers [13]. While the HOLC could write loans, the FHA could only insure loans. The HOLC was always meant to be temporary and was liquidated in 1955 whereas the FHA is permanent. Together, these two agencies made up much of the federal backbone for housing policy implementation in the 20th century.

In order to evaluate the financial solvency of these loans, both agencies created maps that divided neighborhoods in more than 200 cities into four categories: best, still desirable, definitely declining, and hazardous. The hazardous areas were color-coded red on these maps, giving rise to the name "redlining." Though the reported intention of these maps was to help creditors accurately assess monetary risk when granting home mortgages and small business loans, many of these neighborhood designations were not strictly based on economic stability. In fact, *The Underwriting Manual: Underwriting and Valuation Procedure Under Title II of the National Housing Act*, published by the FHA to help evaluators categorize neighborhoods, revealed that the FHA considered factors such as "social characteristics of neighborhood occupants," "stage and trend of neighborhood development," the presence of restrictive [racial] covenants, "natural physical protection" from "adverse influences" such as inharmonious racial groups, and "social attractiveness" as factors on par with financial solvency [14].

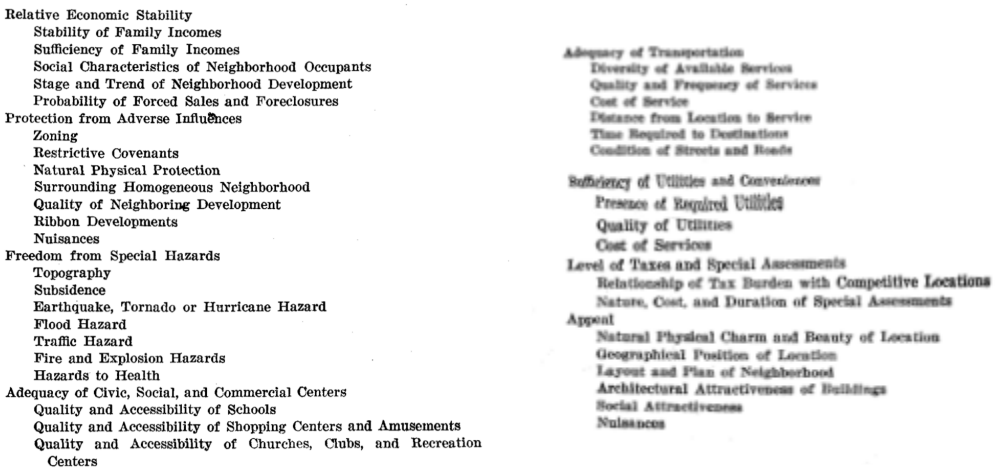


FIGURE 2. Principle considerations used in FHA neighborhood evaluation [14].

Considerable amounts of research since the creation of these maps have demonstrated that these redlining practices contributed to the *de jure* institutionalization of racism and the segregation seen in the 20th century. Even today, the racial demographics of many metropolitan neighborhoods still highly correlate with these categorizations [18]. For example, tracts with greater redlining were found to have higher percentages of minority residents and increased vacancies in 2020 [15]. Additionally, two-thirds of areas designated as "best" are still "overwhelmingly" white [16].

Redlining is an important component of American-style segregation but is unfortunately neglected in the Schelling model. In this thesis, we will expand the Schelling model to include this *de jure* segregation mechanism and compare it to strictly *de facto* segregation.

1.2.4 Network and Graph Theory

CA models like the Schelling model that we propose to expand in this thesis are composed of cells that are arranged in a uniform way. As such, all cells have the same relationship to their nearest neighbors. However, spatial relationships in the real world do not

look like this and this is another way that the 1971 model is lacking. As such, here is where network and graph theory come into play.

Graph theory is a branch of mathematics that deals with points and lines; the theory rests in how these points (often also referred to as nodes or vertices) and lines (also referred to as edges) are connected. Network science builds on the mathematics of graph theory, as well as techniques from other disciplines, to represent and model a variety of physical, biological, and social systems [17]. Nodes and edges are allowed to have various attributes in networks, which enables the analysis of both symmetric and asymmetric relationships, as well as more complicated influences of the attributes. For example, undirected graphs are an example of a graph having purely symmetric relationships whereas directed graphs may have asymmetric relationships in which edges connecting two nodes may only allow for movement in one direction. Node attributes may further influence the behavior of a model system. These variabilities within the experimental space make graph and network theory important components of applied math.

Due to their applicability, networks have been used in a variety of settings. Some notable advancements are their uses in sociology and social network analysis, medicine and epidemiology, and computer science and the Internet. In this paper, we will utilize network theory in relation to urban studies, spatial networks, and urban development.

1.2.5 Networks to Model Urban Spaces and Interpersonal Relationships

Cities are areas of high human density and overall population. The way residential, commercial, and communal spaces are laid out within these geographies can be represented by networks better than homogeneous grids due to the fact that, historically, most city growth was unregulated and/or not pre-planned. In fact, before the rise of network analysis, people often

referred to cities as sprawling metropolises with little rhyme or reason [19]. Only with the rise in mathematical modeling and spatial networks have urban centers been found to follow certain growth patterns in both size and shape. Michael Batty is one renowned spatial data scientist who, through network analysis, has found that cities are in fact fractal structures, with clusters arising from the need to centralize human activity and resources within a larger environment [20]. Though there is no universal definition for a fractal structure, researchers agree that one notable trait is "self-similarity," where zooming into one part of the structure can either completely (or sometimes merely approximately) resemble the entire structure [21]. Some famous fractals include the Mandelbrot set, the Cantor Set, and the Sierpiński triangle.

Since fractals are common in nature and have many potential applications, numerous mathematical models and algorithms exist to generate them. The Cantor Set, for example, is a regular geometric fractal generated by iteratively removing the open middle-third of a line segment an infinite number of times. On the other hand, most fractals that represent cities are "random fractals" that use stochastic rules on moving particles to generate networks in which each particle becomes a node and each connection is an edge. Since these methods are stochastic, each simulation will result in a unique network even if the initial conditions are the same. Brownian random walks, percolation clusters, and diffusion-limited aggregation (DLA) are all such methods to generate random fractals.

In this paper, following the developments of Batty and his colleagues at University College London's Bartlett Centre for Advanced Spatial Analysis, we use DLA to generate city structures and perform network analyses. Introduced by Tom Witten and Len Sander in 1981, DLA is a stochastic, particle-based method in which a seed is placed at an origin as a pixel and particles are added to the space iteratively. These additional particles are added far away from the

origin and/or at the border of the environment, and allowed to move via Brownian random motion indefinitely *or* until it meets a pixel—at which point the particle itself also stops moving and turns into a pixel. Brownian motion is a stochastic process that describes the movement of particles in a given medium [22]. If the particle meets a border, it can either bounce away from the edge or pass through to the other side as in a toroidally bound image. Over time, a dendritic fractal is created as particles are more likely to stick to the existing structure at the tips of clusters rather than at the centers [23].

Figure 3. Examples of DLA-Generated Networks with n = 500 nodes.

In 1989, Batty and his collaborators demonstrated that DLA could be used to approximate urban growth. The seed henceforth represented the first human structure in a developing area. New particles are spawned in "undeveloped" areas and migrate to the developed areas, sticking to create new buildings. The branching, clusters, and gaps that result from the process capture urban growth patterns as people aggregate to form communities where resources are more easily accessed [24].

Though it is not a perfect representation, pure DLA is a simple method that can approximate urban structure by linking urban growth and spatial development. In Batty, Fotheringham, and Longley's 1989 paper, they compared a DLA-generated city with the

mid-sized town of Taunton, Somerset in England wherein the seed was the town's first residential unit: a castle. Their findings affirmed that even in its simplest form, DLA is able to capture broad characteristics of urban growth. For example, the mean radius between a particle and the seed is 50 percent of the maximum radius in Taunton and also 50 percent in the generated city, and the ratio of the standard deviation to the mean radius was 0.225 for both cities [25]. Shared properties such as density gradients and ordered chaotic structures give credence to the idea that DLA is a suitable method to simulate urban growth [24].

Other stochastic, fractal-generating methods have been identified in the decades following this seminal 1989 paper. The running thread is that cities grow bottom-up rather than top-down [26]. While DLA-generated cities only allow for one generating seed, modern metropolises often are created via the combining of multiple smaller towns. A model published in 1995 used correlated percolation to address this problem [27]. Another model published in 2021 took advantage of modern computational power and increasingly comprehensive data sets to predict urban growth using a method based on documented human mobility behaviors rather than random particle processes [28]. In this paper, we will use simple DLA to run the modified Schelling model and also expand the network generation by allowing for multiple seeds; further work with more complex algorithms should be prioritized in the future.

Three other networks: grid-like/square lattices, Watts and Strogatz small-world, and Barabási–Albert scale-free networks are also used to evaluate redline-Schelling dynamics on networks in general. The grid-like graph is the network analogue to the grid used in two-dimensional cellular automata in the sense that each node has a similarly structured neighborhood, being connected to four other nodes with a square-lattice geometry. In the Watts-Strogatz model, regular graphs with degree k are rewired with probability p, avoiding

self-loops and link duplication [29]. As such, the resulting small-world models are highly clustered with short path lengths. While they may not spatially represent urban networks, they capture essential properties of *social* networks, which allows for social segregation to be investigated. In contrast, the Barabási–Albert model generates random, scale-free networks with degree distributions approximated by the power law $P(k) \sim k^{-3}$ [30]. A scale-free network is one where the degree distribution follows a power law and adds nodes via preferential attachment where new nodes are more likely to attach to nodes that are highly connected. Such networks are more reflective of many more real-world phenomena than random networks where the number of nodes is fixed and each node has an equal probability to be attached to any other node, regardless of degree.

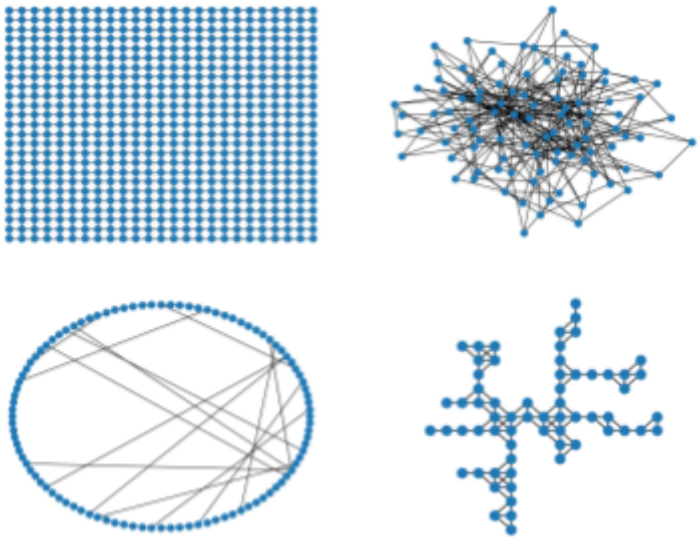


Figure 4. Examples of grid-like, Barabási-Albert/scale-free, Watts-Strogatz, and diffusion-limited aggregation networks.

Chapter 2: Methods

2.1 The Schelling Model

This paper uses the NetLogo implementation of the Schelling model for the grid version and a self-coded implementation in the NetworkX Python package for analysis on the network. Notably, this paper uses these implementations of the Schelling model both for comparison purposes between CA and networks, and as the basis for the expanded redlining model.

In the original model, there are two permanently distinct and same-sized groups within the population. Our two groups are blue and orange, with blue representing the minority group and orange representing the majority. The original experimental space is a grid, with each cell representing a one-person residential unit. On the various network types tested in this thesis, each *node* represents a one-person residential unit. Cells/nodes can either be inhabited or vacant and agents cannot forcibly "displace" others. That is, they can only move into vacant residences and economics plays no role in how agents move. The two distinct groups have commonly been understood as whites and Blacks in the American context, but they can stand for any polarized population where individuals are always aware of the status of both those around them and of themselves (e.g., majority/minority, male-presenting/female-presenting, liberal/conservative).

Apart from their group affiliation, agents have one other attribute known as "happiness" or satisfaction. Agents are "happy" when they are surrounded by a sufficient proportion of neighbors that are of the same type as them. This number, also referred to as a "similarity preference/threshold" λ, is a parameter the researcher defines and then applies uniformly to every single agent. Agents are “unhappy” when their neighborhood does not meet this similarity threshold. As such, a low similarity threshold means individual agents can be happy with greater

integration, whereas a higher number means individual agents have larger in-group preference and want greater segregation. At the extreme, a similarity threshold $\lambda = 0$ would imply that agents are always happy and never have to move whereas $\lambda = 1$ would imply that all agents require every single one of their neighbors to be of the same type.

Note that an agent's "neighbors" is defined by the experimental set-up. For the one-dimensional cellular automata case, for example, Schelling states that an agent's neighborhood consists only of himself and the two immediately adjacent cells. For the two-dimensional cellular automata case, one can either use the square-shaped Moore or diamond-shaped von Neumann neighborhoods; the Moore neighborhood is the more common one, giving each agent eight neighbors. On the network, the neighborhood is simply those agents that are directly connected to the node in question. The only other adjustable parameter in the Schelling model is the global environment's population density. Density plays a role because with a higher density, there are fewer open places for unhappy agents to move. Furthermore, with a low density, what it means for an agent to be “happy” can result in more interesting shapes (such as a lone majority agent staying in a minority-populated section because they possess no neighbors at all).

How does the algorithm work? In economics terms, happy agents have a utility of one and unhappy agents have a utility of zero:

(3) $$U_i = \begin{cases} 0 & : P_{ij} < \lambda \\ 1 & : P_{ij} \geq \lambda \end{cases}.$$

At each time step, all agents evaluate their utility at the same time. Those with zero utility will randomly pick a vacant cell to occupy in a random order, leaving their old cell. This process continues and the population will either reach a "steady state,”

(4) $$\sum_{i=1}^{m} U_i = |A| ,$$

where either nothing appreciably changes or the system continues oscillating forever. The code for the NetworkX implementation is available upon request; the code for the NetLogo implementation is freely available on the NetLogo website. The general algorithm is provided below.

At each time step, iterate through all nodes/cells. For each agent, perform the following:

- Evaluate the agent's utility by comparing the proportion of out-group agents to the designated similarity threshold λ.
- If the proportion does not meet the similarity threshold, the agent has a utility of 0 and is unhappy. Otherwise, the agent has a utility of 1 and is happy.
- Add the unhappy node to a list of unhappy agents for that time step *t*.
- After iterating through all agents, generate a list of vacant spots.
- Iterate through the list of unhappy agents, allowing them to randomly pick a vacant spot to move to and removing the vacant spot from the list when appropriate.

2.2 Operationalizing FHA Redlining

In our expanded Schelling model, we implement and operationalize a version of FHA redlining using guidelines from the *Underwriting* document that evaluators used when making reports [14].

On the grid, we do this with three key modifications; namely:

1) dividing the 50 x 50 grid into 25 10 x 10 spatially-equivalent neighborhoods,
2) assigning each neighborhood a status based on minority density, and
3) introducing stochasticity in terms of success of agent movement.

Since the *Underwriting Manual* shows that significant weight was given to protecting majority isolation (see Figure 4), we simplified redlining by assigning neighborhoods a status based on the percentage of blue minority agents in the neighborhood: best neighborhoods had less than 45 percent blue agents, desirable neighborhoods had between 45 (inclusive) and 50 percent blue agents, declining neighborhoods had between 50 (inclusive) and 55 percent blue agents, and the "hazardous" neighborhoods had more than or equal to 55 percent blue agents. Through redlining, majority agents can move anywhere, but minority agents have diminished chances of moving into the top three neighborhood designations. This stochasticity is approximated in the thesis and is based on how likely a minority agent could move into a given neighborhood with/without a loan or social support. For example, a minority agent might only have a 10 percent chance of successfully moving into a neighborhood with a "Best" rating whereas they have a 50 percent chance of successfully moving into a "Declining" neighborhood.

On the network, we kept the general same approach of three key modifications, except that the first modification of neighborhood establishments was modified. Due to the irregular nature of the node layout and edge connections, it was not simple to create an algorithm that resulted in geographically coherent neighborhoods with the same number of occupants that worked for all network types. Indeed, this question of what constitutes a neighborhood is heavily discussed in geography and sociology research; the persistently unsolved real-world application of this problem—drawing and defining census tracts—shows just how difficult it is. In our

model, we use an approach inspired by the breadth-first search (BFS) algorithm used in traversing a graph data structure. The algorithm for finding the neighborhood on the network is as follows.

Begin by setting all nodes in the graph with a status of "0" to indicate it is "unvisited."

- Pick a random node *N* in the network.
- If the node has not been visited, generate a subgraph of neighbors of *n*th degree.
- Make a neighborhood that now includes the node *N* and the nodes in the subgraph.
- Evaluate the proportion of minority agents in this neighborhood.
    - If there are no minority agents, set the status as "Best."
    - If minority agents make up less than 50 percent of agents, set the status as "Desirable."
    - If minority agents make up between 50 percent and 75 percent of agents, set the status as "Declining."
    - If minority agents make up more than 75 percent of agents, set the status as "Hazardous."
- Mark all nodes in the neighborhood as "visited" by giving it a new number in accordance with its redlined designation.
- Pick a new node and repeat the process.

There are many other ways to model redlining; we detail the limitations of our choices in the Discussion section.

2.3 Methods of Simulation Analysis

2.3.1 Measuring Segregation

The analyses in this paper use various segregation measures to evaluate the effects of operationalized redlining on both the grid and various networks.

On the grid, we use three segregation indices: the dissimilarity index, isolation index, and in-radius neighbor ratio. The dissimilarity index is a common measure of evenness used by the U.S. Census Bureau when studying populations and measures the percentage of a group's population that would have to move to another neighborhood to achieve perfect evenness [30]. "Evenness" refers to the spatial distributions of different groups; segregation is minimized with even distributions when each geographic unit (e.g., a census tract) has the same proportion of minority individuals. The dissimilarity index formula used in this paper is the most common one:

$$D = \frac{1}{2}\sum_{i=1}^{N}\left|\frac{O_i}{O_{tot}} - \frac{B_i}{B_{tot}}\right|, \tag{5}$$

where N is the set of all neighborhoods, O_i and B_i indicate the number of orange/majority or blue/minority agents in the i-th neighborhood respectively, and O_{tot} and B_{tot} represent the number of orange or blue agents in the entire experiment space.

Another group of measures of residential segregation concerns *exposure* instead of *evenness* [30]. As evidenced by the name, this group of metrics represents the probability for one group to be exposed to the other. Segregation has many facets which are all valid and important; this is why we thought it important to include more than one segregation index in the CA portion of the thesis. The isolation index is one such exposure index that we will use in this paper, which

measures the extent to which minority individuals are only exposed to each other and the potential for inter-group interaction is minimized,

$$I = \sum_{i=1}^{N} \left(\frac{B_i}{B_{tot}} \cdot \frac{B_i}{B_i + O_i}\right) . \quad (6)$$

For both the dissimilarity and isolation indices, a higher value indicates more segregation. By using two of the most common segregation indices, we are able to compare simulation results with real data.

The last measure of segregation we will use for the cellular automata case is a more informal measure that we will introduce in this paper. For a certain radius *r,* the in-radius neighbor ratio R_r gives the ratio of the proportion of blue or orange neighbors,

$$R_r = \frac{B_r}{N_r} : \frac{O_r}{N_r} \quad , \quad (7)$$

where B_r , O_r are the number of blue or orange neighbors, respectively, and N_r is the total number of neighbors within the radius *r*. A ratio of 1 would indicate no/little segregation at that scale of *r,* and evaluating the in-radius neighbor ratio *R* at different radii would allow one to analyze how pervasive the segregation mechanism is. Ratios that remain far from 1 even for larger and larger *r's* would indicate very powerful segregating forces.

For the network analyses, we will primarily lean on the social mixity index introduced in the Schelling paper and co-opted by other researchers who have studied network effects on his model [36]. This index reveals the mean proportion of contacts between unlike neighbors, omitting vacant nodes, and is calculated as

(8) $$M = \frac{1}{|A|}\sum_{i=1}^{|A|} P_{ij},$$

where A is the number of agents. As such, the social mixity index is bound by $M = 0$ and approximately $M = 0.5$, where $M = 0$ would indicate complete segregation and $M = 0.5$ would indicate complete mixing (i.e., 50 percent of a node's neighbors are unlike them in a world where there are equal numbers of majority and minority agents). However, note that there are numerous ways to achieve the same index. For example, in the example below, 0.5 mixity occurs when there is a random distribution of agents on the grid-like network but also when the network is laid out in a checkerboard pattern. If the network were instead a grid and we evaluated the center point's utility based on a Moore neighborhood vs. a von Neumann neighborhood, we would get very different results.

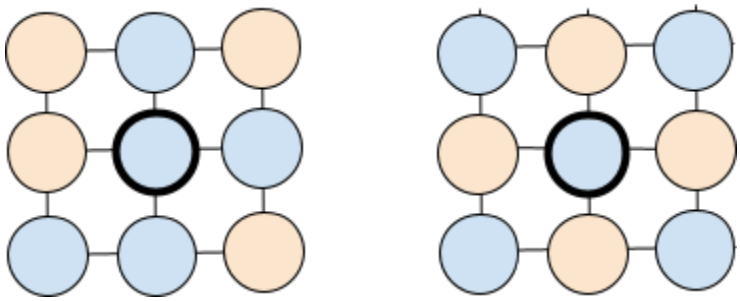


FIGURE 5. Two ways to get "complete integration" based on Schelling's social mixity index.

2.3.2 Segregation's Impact on Health Disparities

The role of racial segregation as investigated in this paper has numerous implications for other types of segregation as well. In particular, sociology and public health researchers have recently named racial segregation in the United States as a fundamental cause of health inequality [33, 34]. In order to preliminarily analyze how our modified version of the Schelling model of racial segregation impacts neighborhood-level health disparities, we imposed two

different probability distributions on majority and minority agents in the cellular automata case for physical activity levels based on CDC data: majority agents had a 30 percent chance of being physically inactive whereas minority agents had a 23 percent chance [35]. With this addition, we hope to analyze the impact that race-based segregation has on health disparities even with as small a difference in health activity as 7 percent.

2.4 Cellular Automata Simulations

In part one of our simulations on the grid, the densities ranged from five to 90 percent whereas the similarity threshold λ ranged from zero to 100 percent; each density-preference pairing was tested three times to get a total of 1153 total simulations across both models. The three segregation metrics were then averaged for each density-preference pairing to create heatmaps and bar graphs. Phase diagrams were also constructed from heatmap analysis.

In part two of our simulations on the grid, the density was kept at 50 percent and the similarity threshold λ at 30 percent. This simulation was run three times to find the average in-radius neighbor ratio for radii of 2, 4, 8, 15, and 50 cells, as well as the average activity rate of each neighborhood.

2.5 Network Simulations

In part one of our simulations on a network, we first did basic network generation of the five tested network types: grid-like, Barabási-Albert with attachment $m = 3$, Watts-Strogatz with attachment $m = 3$ and rewiring probability $p = 0.15$, DLA, and clustered DLA. The redline-Schelling model was run once on each network type to obtain visualizations at time steps $n = 0$, 5, and 15 for a density of $d = 0.8$ and a similarity threshold $\lambda = 0.6$. These time steps were

chosen after preliminary simulations revealed that steady state was generally reached after around 15 to 20 iterations; the density was chosen at 0.8 for a number of reasons.

Firstly, for interesting behavior to result from the *de facto* portion of the redline-Schelling model, agent density must neither be too high or too low. If density is too high, then the system would be frozen in place and unhappy agents would have very few vacant spots to move to. If density is too low, agents are much more likely to be spatially isolated simply as a result of the number of vacancies. The final spatial distribution, then, would be more random than expected by the similarity threshold. This limitation on density minima is even revealed in NetLogo's web implementation of the Schelling model, where users can only choose density d from a range of 0.5 to 1. Secondly, $d = 0.8$ was chosen for this portion of the thesis because previous research on network effects in Schelling-style segregation used this number. For ease of comparison and to affirm that the Python model was implemented correctly, we chose to use the same number. Lastly, FHA redlining in the 20th century was primarily done in metropolitan areas in the United States. Running visualizations on low density areas would be irrelevant in a historical and practical sense because of this fact.

In parts two and three of our simulations on a network, we were interested in plotting social mixity vs. time to compare how adding redlining to the Schelling model affects how segregation evolves over time. Because of limits on computational power, only a couple combinations of network types, densities, and similarity thresholds were chosen. Notably, to represent spatial segregation, we chose to investigate the grid-like and DLA networks. We ran both the original Schelling model and the redline-Schelling version with density and similarity threshold combinations of $d = 0.50, 0.75, 0.85$ and $\lambda = 0.25, 0.50, 0.75, 0.9$. There are therefore 12 possible density-similarity threshold combinations and two models. In part two of our

simulations on a network, the redlining model was implemented at the beginning of the experiment. In part three, redlining was implemented after 10 time steps. This was done to reflect the fact that, in most cases, the FHA was evaluating neighborhoods and imposing policies on places that had already been segregated through various *de facto* measures such as in-group preference and migration due to job security. By applying redlining to a segregated space, one could investigate an additional layer of model propagation on the network. Each experimental set-up was simulated five times and the results averaged to get a total of 240 simulations (120 simulations in part two and 120 simulations in part three).

In part four of our simulations on a network, we investigated the role of clusters in highly-clustered and fractal-like networks in segregation propagation as influenced by redlining at time step $t = 0$ and in-group preference. The density was held constant at $d = 0.80$ as before, with the threshold being one of the following $\lambda = 0, 0.2, 0.4, 0.6, 0.8, 1$. $\lambda = 0$ indicates that all agents are always happy, whereas $\lambda = 1$ indicates that all agents are always unhappy. Mild in-group preference occurs at $\lambda = 0.2$ and 0.4 whereas stronger in-group preference occurs at $\lambda = 0.6, 0.8$—that is, agents want a majority of their neighbors to be of the same type as them. All resulting evolutions for one model were plotted on one graph to allow for ease of comparison both within models and across models for social mixity at time step $n = 35$ relative to λ.

In parts two, three, and four of investigating segregation dynamics on a network, we were interested in seeing how segregation propagates through time. Now in parts five and six, we finally move to analyzing end-stage segregation. That is, how does the model choice and set-up impact the intensity of segregation at steady state? In part five, we kept density at $d = 0.8$ and iterated through the following similarity thresholds $\lambda = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1$. The evaluated networks include grid-like, DLA, and clustered DLA. Each density-similarity

pairing was run five times and the social mixity index at n = 35 averaged over the five simulations. In total, there are 330 simulations for part five (55 for the Schelling model and 55 for the redline-Schelling expansion across three networks). In part six, rather than varying λ, we vary density d to see how network vacancy might impact redline segregation dynamics. λ is held at 0.6 to indicate moderate in-group preference and d = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, or 0.9 for one of four networks: grid-like, Barabási-Albert, DLA, and clustered DLA. Simulations for each set-up were run five times to n = 35 for a total of 180 simulations.

2.6 Mean Field Approximations

In the thesis so far, many of our analyses have involved simulations in mathematical modeling. Because CA and network models, though simplistic in their implementation, can result in highly irregular and nonlinear behaviors, we now turn to mean field theory and approximation to investigate the *de jure* redlining dynamics analytically.

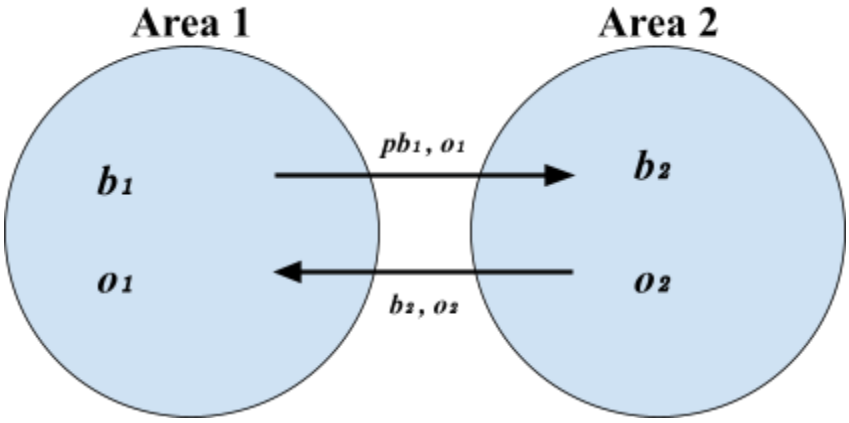


FIGURE 6. Graphical depiction of mean field simplification.

By using mean field approximation, we are able to simplify the research question considerably into a system of coupled, nonlinear ordinary differential equations. As in our simulations, blue agents (denoted here as b) are minority agents whereas orange agents (denoted

here as *o*) are the majority agents. Rather than having four neighborhood designations, we simplify the system into two neighborhood types: "undesirable" (denoted as area 1) or "best" (denoted as area 2). Therefore, b_2 gives the proportion of available sites in "best" neighborhoods occupied by minority agents. The sum of b_1, b_2, o_1, o_2 must be bound on the upper extreme by two and the sums $b_1 + o_1$ and $b_2 + o_2$ are both bound by one. In line with the Schelling assumption that the size of the two groups are equal, we also mandate that b_1, b_2, o_1, o_2 have the same value at $t = 0$. The probability of minority agents moving into the best neighborhoods is the parameter p.

This resulting system of equations was solved using the SciPy Python library:

$$b_1' = b_2(1 - (b_1 + o_1)) - pb_1(1 - (b_2 + o_2)), \tag{9}$$

$$b_2' = pb_1(1 - (b_2 + o_2)) - b_2(1 - (b_1 + o_1)), \tag{10}$$

$$o_1' = o(1 - (b_1 + o_1)) - o_1(1 - (b_2 + o_2)), \tag{11}$$

$$o_2' = o_1(1 - (b_2 + o_2)) - b_2(1 - (b_1 + o_1)). \tag{12}$$

In part one of the mean field approximations, the proportion of each population is initialized at $n = 0.3$ and the probability of success p drawn from one of the following: $p = 0$, 0.1, 0.25, 0.5, 0.75. The system was solved from $t = 0$ to $t = 20$ and we plot n across time for all four populations. In part two, we investigate the role of density on minority agent movement by varying $d = 0.01$, 0.05, 0.4, 0.2, 0.3, 0.4, 0.45, 0.49 and plotting time vs. b_2. Lastly, in part three, we solve the system of equations for various probability p's ($p = 0.2$, 0.4, 0.6, 0.8, 0.95) to more

directly compare how the probability of success affects the movement of minority agents into the protected areas.

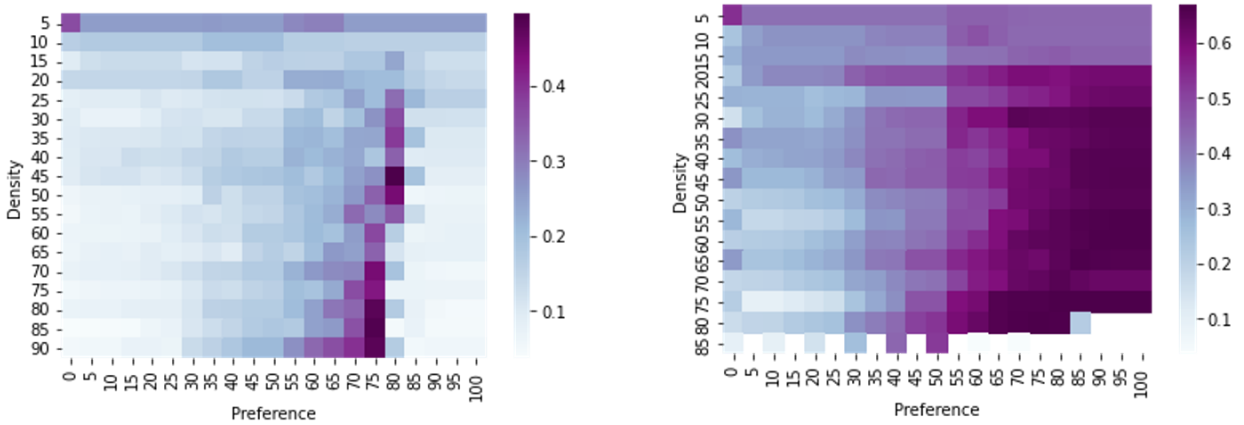
Chapter 3: Results and Analysis

3.1 Modified Schelling Model on the Grid

Through analysis of the three metrics outlined in the methodology, we find that our redlining modification to the Schelling model results in more intense and variable segregation than the original Schelling model.

3.1.1 Dissimilarity Index

Figure 7 shows the heatmaps for the dissimilarity indices from the steady state of each set of experiments. Should a simulation fail to run, the corresponding cell is left as pure white; this only occurred in the modified Schelling model (the heatmap on the right) with very high densities that prevented even one step from completing. It is probable that this occurred because the environment had fewer unoccupied locations, meaning that unhappy blue agents (and to some extent, orange agents) had fewer places to choose from when being forced to move.

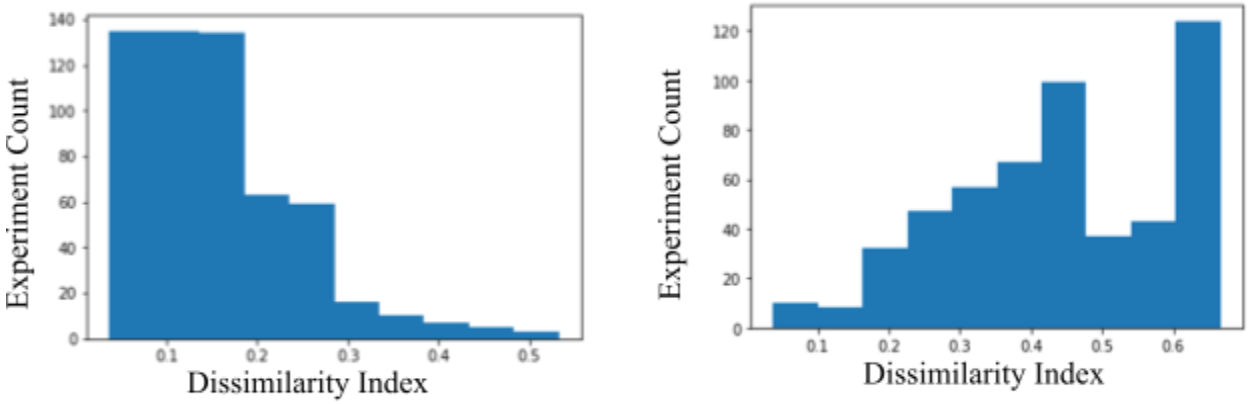


(a) Original Model (b) Redlining Model

FIGURE 7. Heatmaps showing average dissimilarity index for the steady state of various density-preference pairings.

For the original Schelling model, dissimilarity indices for the steady state ran from 0.037 to 0.53 for the density-preference pairing. For the redlining model, these same indices ran from 0.037 to 0.66, indicating that the modified *de jure* model didn't affect the minimum *D* but increased the maximum. Throughout the whole set of simulations, the Schelling model had an average dissimilarity index of 0.16 whereas the proposed model had 0.44—the new model increased the dissimilarity index average by a factor of 2.75.

Figure 8 has bar graphs showing the distribution of dissimilarity indices in both models. The original model's distribution is somewhat right-skewed and follows a general curve that indicates a more predictable and consistent segregation outcome—in most cases, one can expect the dissimilarity index to be below 0.2. The redlining model, however, looks very different. Not only are most values to the right of 0.2, but the curve itself is not easily characterized, indicating a more volatile and sensitive segregation mechanism.



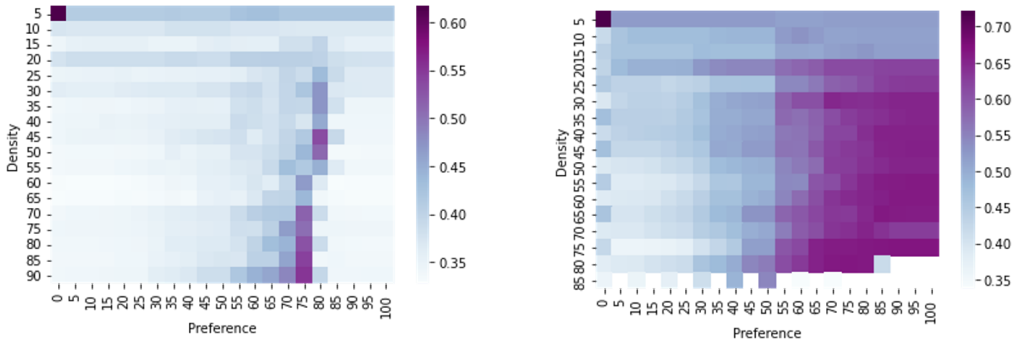
(a) Original Model (b) Redlining Model

FIGURE 8. Bar graphs showing distribution of average dissimilarity index for the steady state of various density-preference pairings.

Therefore, our redlining, *de jure* segregation model enhances unevenness and introduces more variability in neighborhood outcomes.

3.1.2 Isolation Index

Figure 9 shows the heatmaps for the isolation index of both simulations. Note that the white cells in these figures are from the same problem detailed in the beginning of the discussion of the dissimilarity index results.

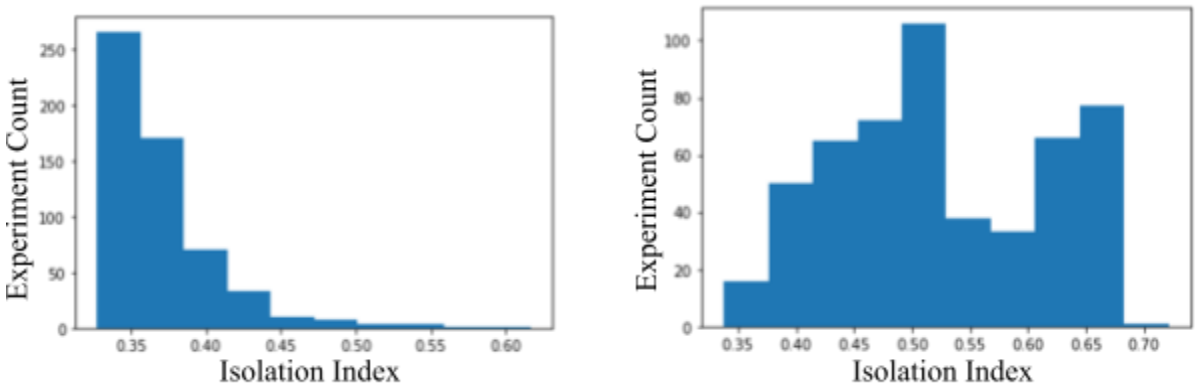


(a) Original Model (b) Redlining Model

FIGURE 9. Heatmaps showing average isolation index for the steady state of various density-preference pairings.

As stated before, the isolation index is a measure of the exposure that the minority group—in this case, the blue agents—has to the majority group. For the Schelling model, the isolation index had a range between 0.33 and 0.62, inclusive. The modified version had a range of 0.34 and 0.72, inclusive. Like with the dissimilarity index, the isolation index minimum is not affected but the maximum is increased. The mean for all simulations in the Schelling model was 0.37; the modified version had an impact 1.42 times larger than this, with an average isolation index of 0.53.

Figure 10 shows the distribution of isolation indices for the simulations that were run. Similar to the case with the dissimilarity index, the redlining model shifted the distribution to the right, resulting in higher isolation indices and more intense segregation. Again, we see that the overall distribution changes from a moderately smooth, right-skewed distribution to one with two peaks that is more left-skewed.



(a) Original Model (b) Redlining Model

FIGURE 10. Bar graphs showing distribution of average isolation index for the steady state of various density-preference pairings.

Therefore, the redlining *de jure* segregation model reduces exposure and introduces more variability in neighborhood outcomes.

3.1.3 Construction of Qualitative Phase Diagrams

Phase diagrams (Figure 11) were drawn from combining the results of both the dissimilarity and isolation index heatmaps. Phases are named in alphabetical order, with phase A having the least amount of segregation and phase E having the most (see Table 1 for details).

TABLE 1. Phases of segregation and their corresponding *D, I* index ranges.

Phase Designation	**Dissimilarity Index (*D*)**	**Isolation Index (*I*)**
A	$D < 0.15$	$I < 0.25$
B	$0.15 \leq D < 0.25$	$0.25 \leq I < 0.35$
C	$0.25 \leq D < 0.40$	$0.35 \leq I < 0.50$
D	$0.40 \leq D < 0.50$	$0.50 \leq I < 0.60$
E	$0.50 \leq D$	$0.60 \leq I$

While these phase designations are not necessarily grounded in empirical evidence, notice that the modified Schelling model completely loses the two phases with the least amount of segregation. This indicates that the redlining model is a much more powerful segregating force than the original model.

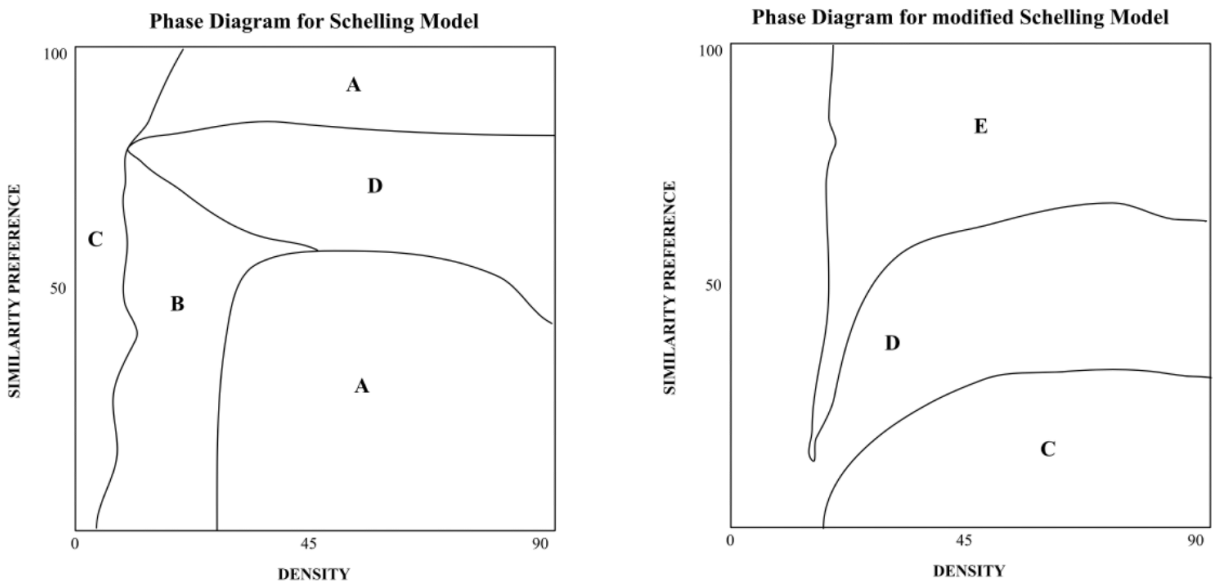


FIGURE 11. Phase diagrams for both the original Schelling model and the modified/redlining Schelling model.

3.2 Modified Schelling Model on Networks

Through analysis of the Schelling social mixity metric outlined in the methodology, we find that moving our redlining modification to the Schelling model to the network impacts segregation outcomes at specific densities and preferences (e.g., segregation traps uniquely occur in fractal-like urban structures only when densities are high).

3.2.1 Network Generation and Preliminary Visualizations

In this section of the thesis, we run the redlining model on five different networks to create preliminary, qualitative visualizations through which we can better understand the accompanying simulation results. Note these are qualitative because the networks cannot be directly compared as their degree distributions are not standardized. Particularly, the Schelling social mixity index as applied to the network redefines “neighbors” from the Moore/von

Neumann definition in the cellular automata case to adjacent nodes on the network. Each neighbor, then, has a weight of $\frac{1}{n}$ if a given node has n neighbors and networks where nodes have more neighbors are more resistant to in-group preference than nodes with smaller "neighborhoods."

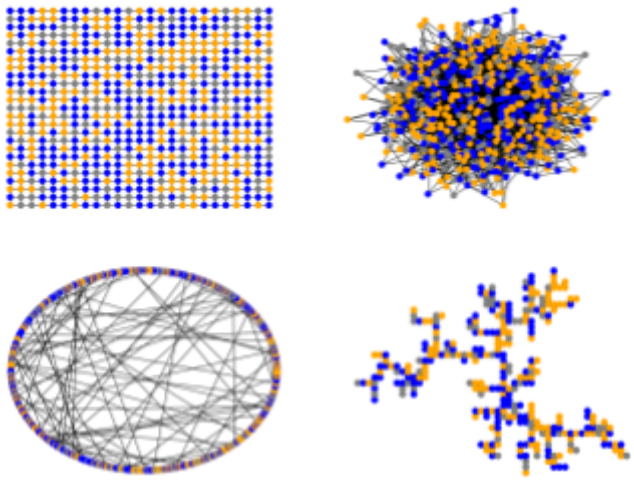


FIGURE 11. Initial states of grid-like, Barabási-Albert, Watts-Strogatz, and DLA networks.

In Figures 11 - 13, we display the grid-like, Barabási-Albert, Watts-Strogatz, and DLA networks side-by-side. Figure 11 shows the initial states when the density d is set at 0.8 for n = 625 nodes. The social mixity indices at this time step are as expected (close to 0.5 for a random distribution) for all four networks, affirming that the index does work as expected for multiple network types and node connections and that our Python implementation was successful. Figures 12 and 13 visualize the network at time steps t = 5 and 10, respectively, for the redline-Schelling model. The grid-like graph's set-up strongly resembles the cellular automata case as it indeed can also be viewed as a grid case where the von Neumann neighborhood is used with non-periodic boundary conditions to determine agent utility.

TABLE 2. Comparisons of social mixity index at various time steps for $d = 0.8$, $\lambda = 0.6$.

Time	**Regular**	**Barabási-Albert**	**Watts-Strogatz**	**DLA**	**Clustered DLA**
$t = 0$	0.498	0.476	0.432	0.482	0.506
$t = 5$	0.252	0.317	0.132	0.179	0.162
$t = 10$	0.155	0.278	0.0288	0.0902	0.110

Dynamics on the Barabási-Albert network appear to be the most resistant to segregation for these selected time steps, as evidenced by the visualizations and the social mixity calculations presented in Table 2. The preferential attachment modeled in this network's generation allows for this phenomenon to occur since agents can be added and connected to pre-existing agents without regard to their relative placement, status, or neighbor types. This type of node addition also means that very little clustering occurs so there are not many "segregation traps." Lastly, we should note that the Barabási-Albert network generation used in this thesis comes from the NetworkX Python Package and has no spatial component (i.e. the edge lengths and node placements have no meaning). Therefore, the resulting figures retain a lot of noise and the segregation dynamics aren't as prevalent as for other networks visually, even for similar social mixity indices.

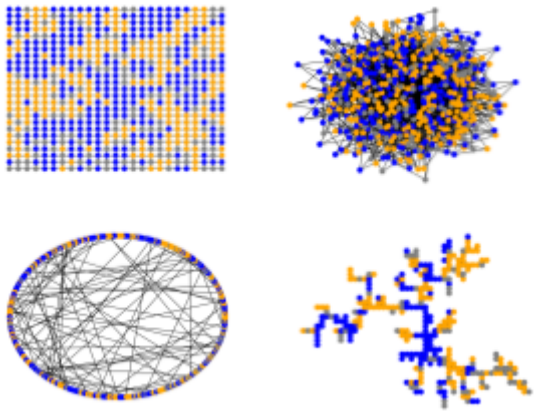


FIGURE 12. States of grid-like, Barabási-Albert, Watts-Strogatz, and DLA networks at $t = 5$.

Because of these facts, the Barabási-Albert model is not an accurate representation of urban layouts. Therefore, though we consider redline-Schelling segregation evolution dynamics on Barabási-Albert networks in this section of the thesis, it is important to note that the findings hold little relevance for the larger problem at hand: to model *de jure* and *de facto* segregation on geographically-relevant experiment spaces. In a similar vein, the Watts-Strogatz small world network also does not reflect city structure as well as the grid-like graph (for cases like New York City) or DLA-generated graphs (most other cities). Instead, the Watts-Strogatz network is a nice "toy" model of social networks, as explained in the Methods section. Interestingly, of the five networks visualized and tested here, this network results in the lowest social mixity index at both $t = 5$ and $t = 10$. This has numerous implications for further work in sociology and research into interpersonal relationships. For example, reframing the redline-Schelling model as a social dynamic model where node movement represents entering social cliques and redlining neighborhoods is redefined as designating the social and/or cultural capital necessary to enter certain communities effectively changes the research question and allows for the same model to be appropriated for a different purpose.

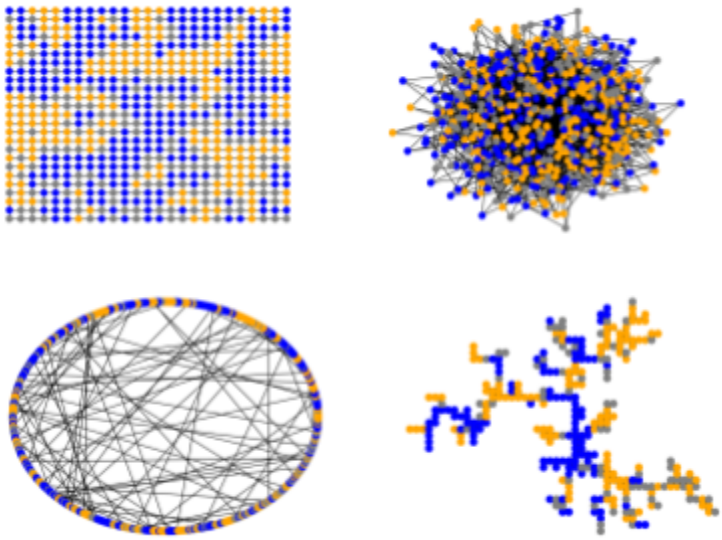


FIGURE 13. States of grid-like, Barabási-Albert, Watts-Strogatz, and DLA networks at $t = 10$.

The last network visualized in Figures 11 - 13 is created by DLA wherein one seed gives rise to a fractal-like network. This network is most like modern cities in that it allows for a variable degree of clustering and the formation of residential hubs vs. putting everyone in blocks next to each other like in the grid-like graph. As seen in Table 2, the social mixity index drops off rather steeply within the first five time steps, rivaled only by the Watts-Strogatz case. Furthermore, this network type allows for more spatial isolation that goes undetected by the social mixity index as implemented in this thesis (see Figure 14). These arise because of the clusters and branches that exist in fractal-like networks wherein an entire section of the graph could be inhabited by agents of one type *and* there are no *n*th neighbors the way they exist on the grid-like, Barabási-Albert, or Watts-Strogatz graph.

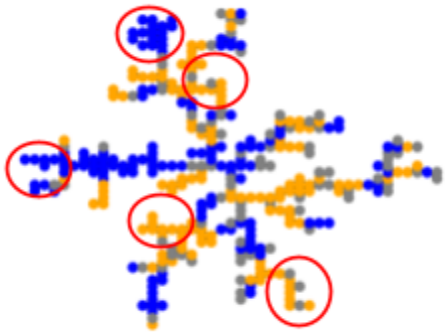


FIGURE 14. Segregation traps in the DLA network at $t = 25$.

Lastly, we ran visualizations for a modified DLA algorithm wherein multiple seeds are generated at the start of the network formation. The social mixity indices for these experiments are comparable to those of the conventional DLA network, although they are slightly higher. This could happen because of the existence of multiple clusters, where more nodes have higher degree. As stated earlier, networks with larger neighborhoods are more resistant to segregation as measured by the network version of social mixity as implemented in this thesis; therefore, this result is somewhat expected in light of the measurements used. Defining a better index to

compare with the dissimilarity or isolation indices used by the U.S. Census Bureau may be a relevant endeavor in the future, to investigate exactly what type of segregation is happening in these networks and what index is more useful to inform policy.

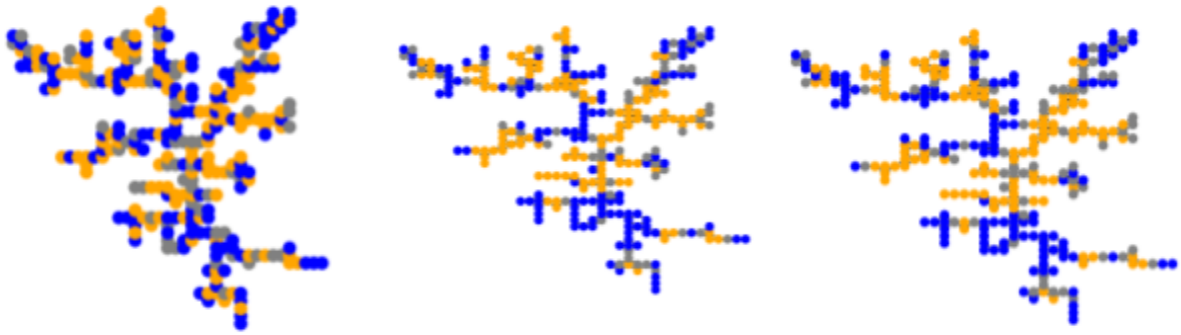


FIGURE 15. States of clustered DLA network at t = 0, 5, and 10.

3.2.2 Network and Segregation Evolution Per Time Step: Initial Redlining

In the following section of the thesis, we plot the social mixity index as a function of time on both the grid-like and DLA network. For low preference ($\lambda = 0.25$) and low densities, there is very little movement in both models and on both network types due to the fact that it is highly likely that agents are already relationally isolated. At higher densities, more movement must occur but there is still relatively little movement. The redlining modification at this point does not always guarantee more segregation/a lower social mixity index, although it does occur on the DLA network by t = 35 for all three tested densities whereas it does not for any density on the grid-like network. This could occur for a number of reasons. It is likely, for example, that redlining at high densities and low preferences moreso "enshrines" agents in their pre-existing locations rather than have them look elsewhere, preserving more of the random distribution that the population was initialized with.

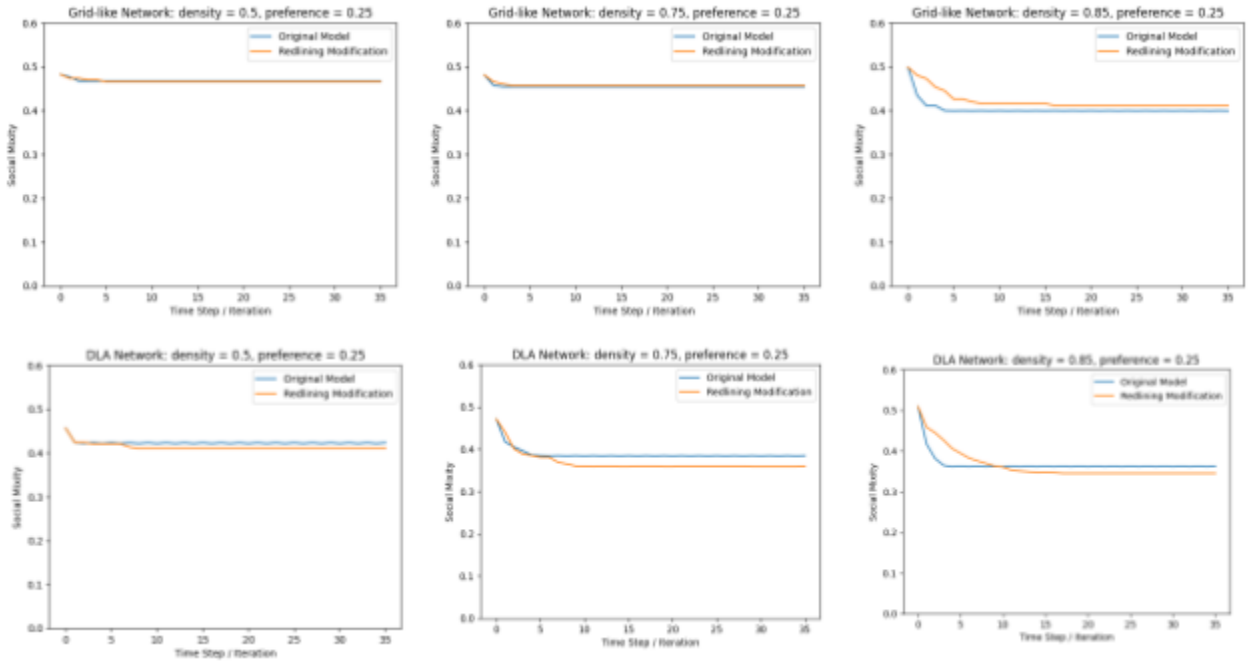


FIGURE 16. Social Mixity Index for $\lambda = 0.25$ and $d = 0.5, 0.75, 0.85$ on grid-like/grid-like and DLA networks: both Schelling model and redline-Schelling modification.

In Figure 16, we move to $\lambda = 0.5$. In practice, this is still a relatively low similarity threshold, as it means that agents just need half of their neighbors to be like them. However, even with this low in-group preference, we see dramatic changes in social mixity through time, particularly for higher densities such as $d = 0.75$ and 0.85. In addition, the DLA network has consistently more segregation than the grid-like network, which is in line with what one expects given the clustering that is present. Interestingly, redlining is seen to somewhat preserve integration in some cases, although the difference between the original and the redlined model is relatively small.

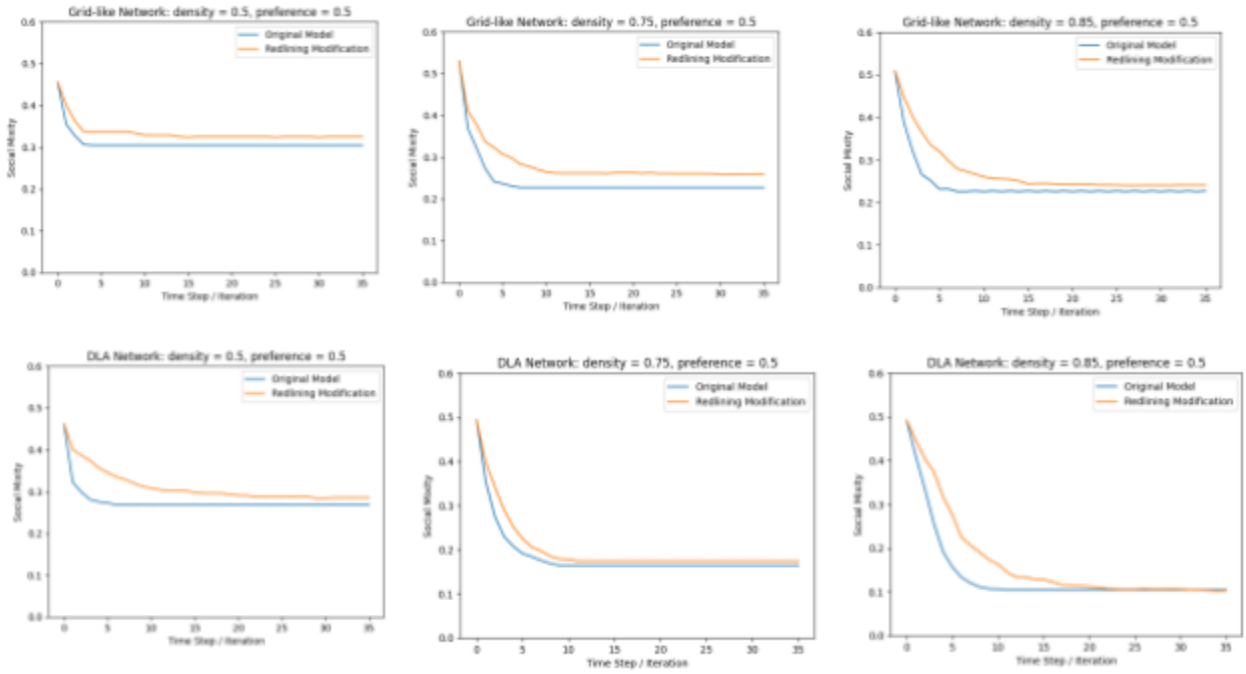


FIGURE 17. Social Mixity Index for $\lambda = 0.5$ and $d = 0.5, 0.75, 0.85$ on grid-like/grid-like and DLA networks: both Schelling model and redline-Schelling modification.

Figures 17 and 18 move to higher similarity thresholds, which induces more movement on the part of the agents and therefore more volatility in the social mixity index. Because of lack of computational resources, we keep the experiment bound by $t = 0$ and 35, although it would be appealing to also look at the social mixity index at steady state ($t = \infty$). In Figure 17, very similar patterns hold from earlier experiments. The DLA network, for example, results in more segregation than on the grid-like graph for both models, but moreso for the redlined-Schelling modification than the original. Judging by the lines themselves, it also appears as if there is more oscillation in segregation for the original model when compared to the redlined one: the curves are notably smoother for most of the graphs shown below. This may indicate that redlining stabilizes the evolution of global segregation on top of the directionality of agent movement: whereas pure *de facto* segregation may have more volatility as people move according to their

preferences, institutional policies and structural racism as depicted in *de jure* segregation and redlining drive segregation more uniformly through time.

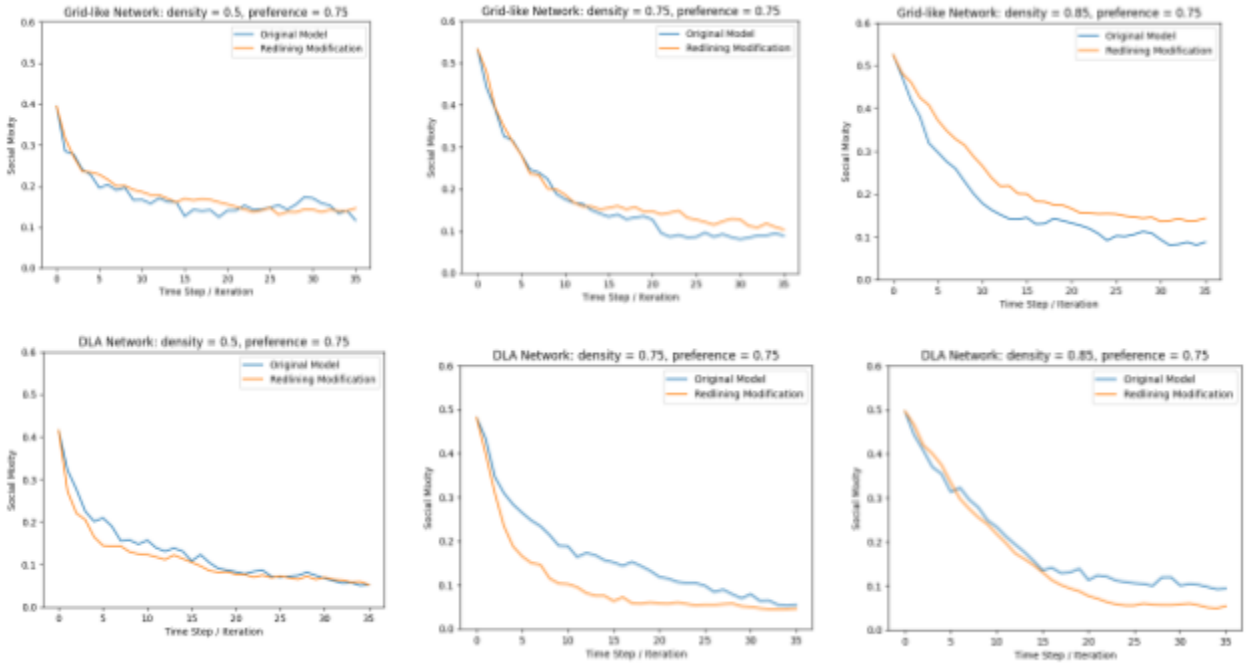


FIGURE 18. Social Mixity Index for $\lambda = 75$ and $d = 0.5, 0.75, 0.85$ on grid-like/grid-like and DLA networks: both Schelling model and redline-Schelling modification.

Lastly, we plot social mixity vs. time for the preference threshold $\lambda = 0.9$, which is relatively high. It is at this point that we see the redlining model make a drastic difference in driving segregation on both the grid-like and DLA networks. The largest differences come at lower densities, which can be rationalized in the sense that agents have more vacant nodes to choose from when unhappy and can further segregate throughout time. Furthermore, the original Schelling model at $\lambda = 0.9$ results in consistent oscillation for all experiments run and also higher overall social mixity indices than for $\lambda = 0.75$, meaning that having a high similarity threshold can actually be counter-productive in environments where only *de facto* mechanisms exist to power segregation. However, when *de jure* mechanisms are implemented, this phenomenon disappears, explaining the effects of segregation in the 20th century U.S.

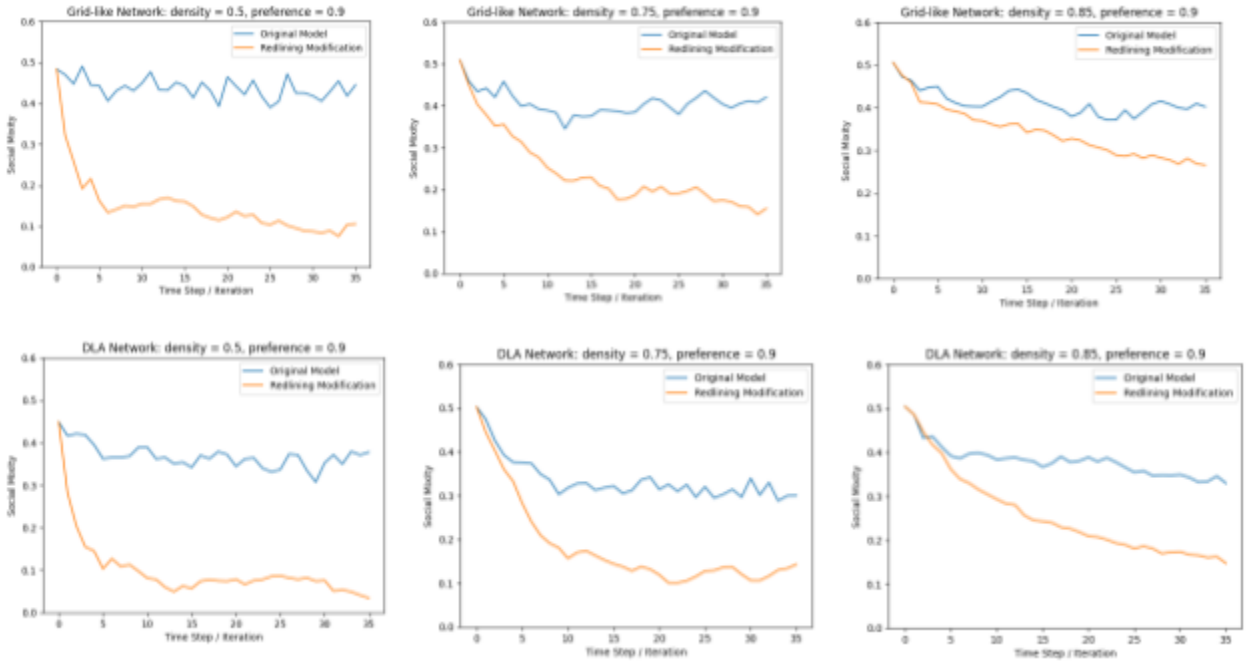


FIGURE 19. Social Mixity Index for $\lambda = 0.9$ and $d = 0.5, 0.75, 0.85$ on grid-like/grid-like and DLA networks: both Schelling model and redline-Schelling modification.

3.2.3 Network and Segregation Evolution Per Time Step: Imposed Redlining

In this section of the thesis, we note historical precedence and impose redlining only after the environment has been somewhat segregated due to in-group preference. More specifically, we impose redlining at $t = 10$, meaning that for the redlining modification, all movement before this is due only to *de facto* segregation mechanisms.

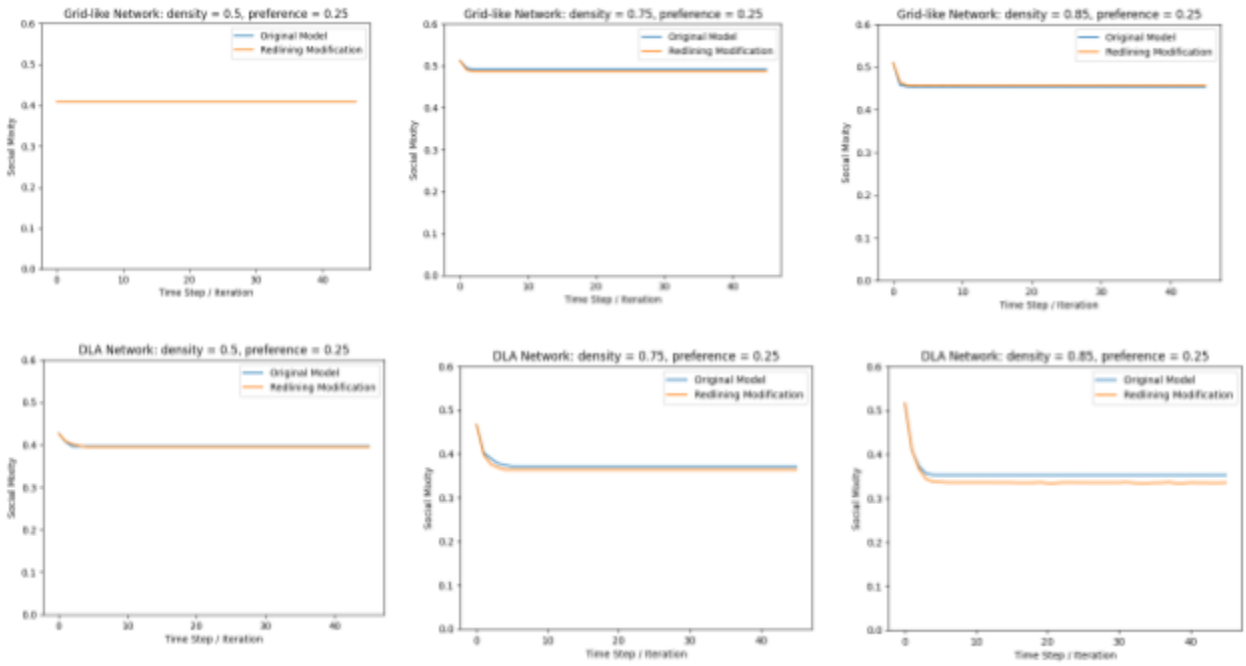


FIGURE 20. Social Mixity Index for $\lambda = 0.25$ and $d = 0.5, 0.75, 0.85$ on grid-like/grid-like and DLA networks: both Schelling model and redline-Schelling modification imposed at $t = 10$.

Similarly to before, there is very little difference between the two evaluated networks and the two models at $\lambda = 0.25$ simply because there is very little movement happening in general at such a low similarity preference. Predictably, more movement and subsequent segregation happens at higher densities as more nodes are in proximity to each other. Interestingly, the redlining modification appears to result in more segregation at this level compared to the original model and the previous simulations. This suggests that exactly what type of network *any* segregation model is imposed on may be important for analysis; research and policy decisions inspired by such research must therefore take this into account when designing experiments, as

most segregation models are run on randomly distributed networks or cellular automata grids.

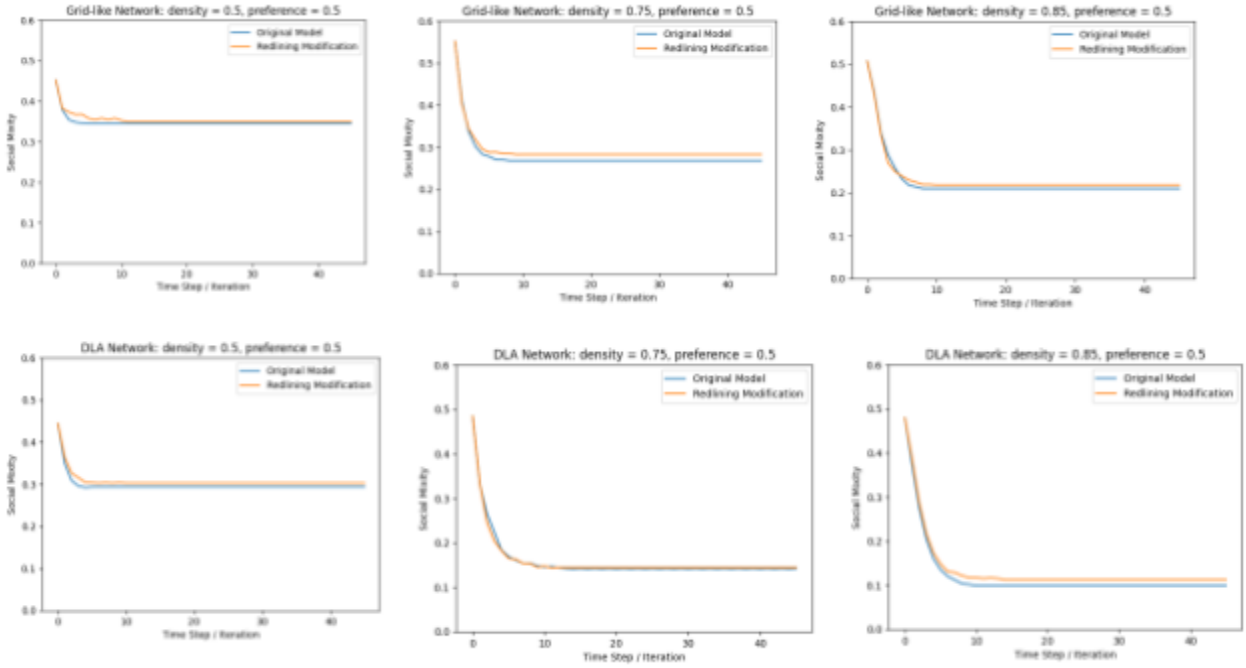


FIGURE 21. Social Mixity Index for $\lambda = 0.5$ and $d = 0.5, 0.75, 0.85$ on grid-like/grid-like and DLA networks: both Schelling model and redline-Schelling modification imposed at $t = 10$.

In the previous set of simulations where *de jure* redlining was implemented at the beginning of the *de facto*-driven movement, the redlining modification resulted in social mixity scores higher than in the original model across all densities at $\lambda = 0.5$. Here, we see that this difference is more or less negated when redlining is implemented at $t = 10$ instead of $t = 0$, although there doesn't appear to be more segregation from the redlining. Exactly why this occurs is unclear, although it may be connected to the fact that $\lambda = 0.5$ is still a relatively mild preference in that agents can be happy when merely half of their neighbors (in an evenly divided population) are of the same type as them. At this point, perhaps redlining doesn't convince the necessary number of agents to move to exert effects on global segregation.

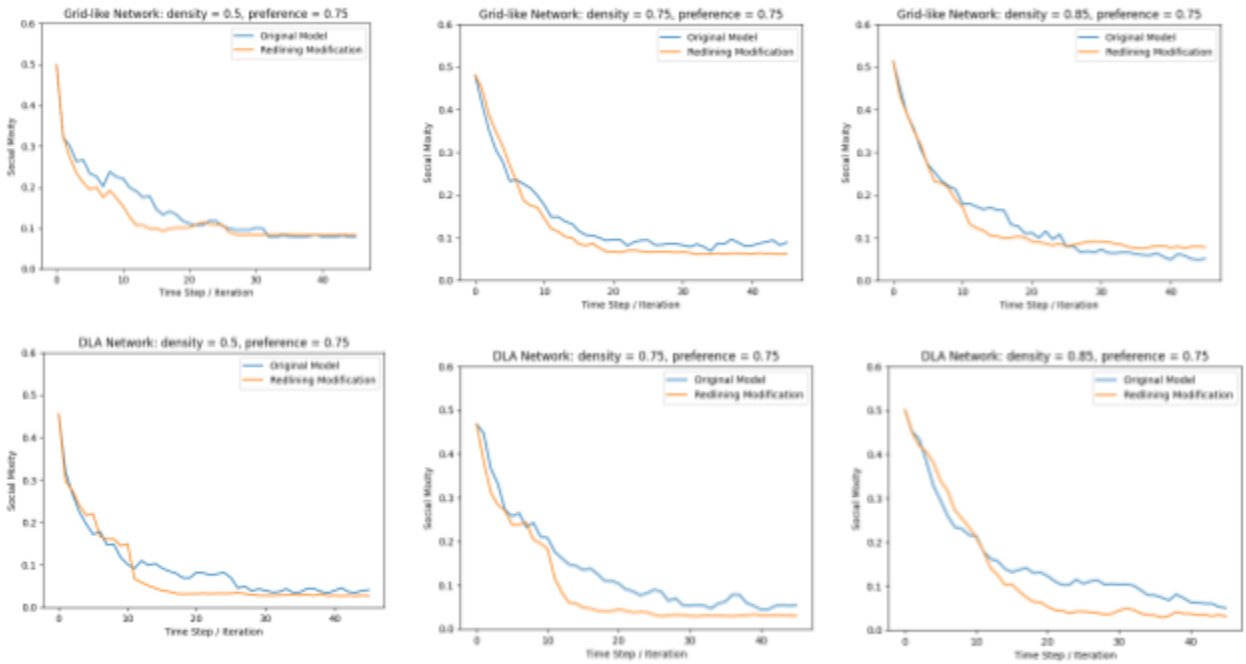


FIGURE 22. Social Mixity Index for $\lambda = 0.75$ and $d = 0.5, 0.75, 0.85$ on grid-like/grid-like and DLA networks: both Schelling model and redline-Schelling modification imposed at $t = 10$.

For higher preference thresholds, both the Schelling model and the redline-Schelling modification begin to oscillate more and take longer to reach a steady state. As in 3.2.2, the simulations run in Figure 21 for $\lambda = 0.75$ generally do not reach a steady state; we keep the same time-scale as previous figures to allow for standardized comparisons of how segregation evolves through time. Whereas the slopes of Figure 20 are steepest near the beginning of the simulation and quickly decrease to zero, the slopes in Figure 21 do not exhibit this type of behavior and instead have a much more drawn out decay that is also more prone to oscillation. Accordingly, the differences in the social mixity index prior to $t = 10$ most likely arise from experimental noise as only five trials were run for each density; with many more experiments, one would expect the two curves from $t = 0$ to $t = 10$ to be the same. It is interesting to note that at this preference threshold, the redlining modification begins to result in more segregation and a lower social mixity index than the original Schelling model, particularly for the DLA network.

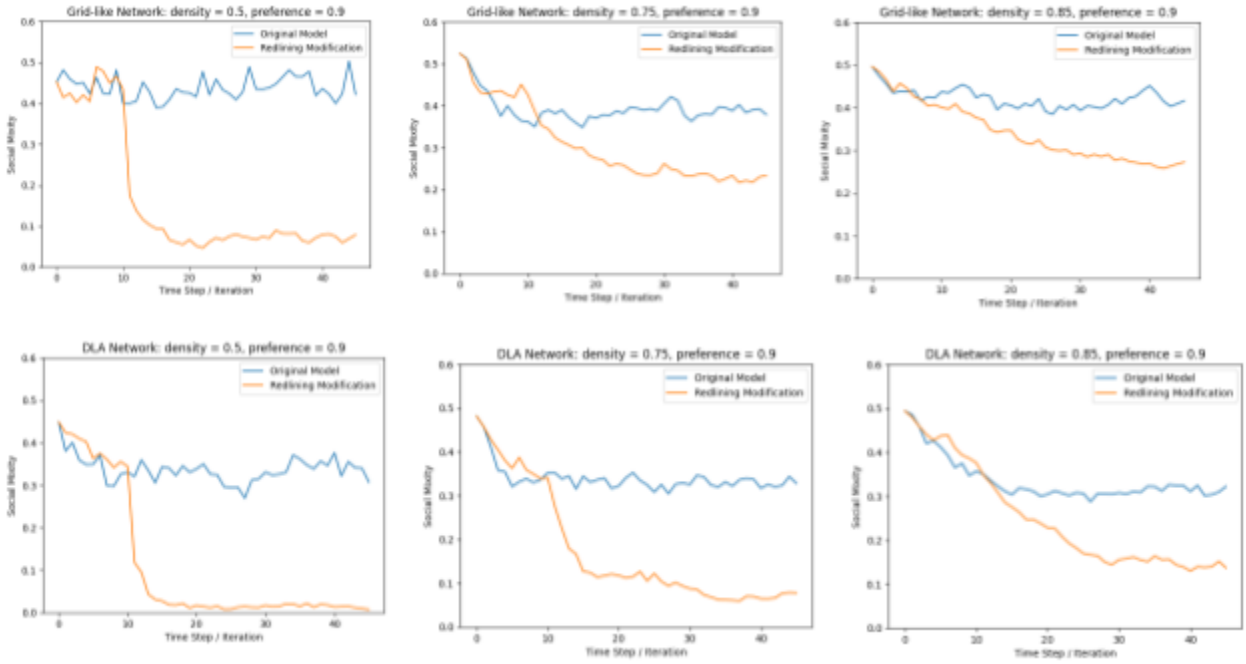


FIGURE 23. Social Mixity Index for $\lambda = 0.9$ and $d = 0.5, 0.75, 0.85$ on grid-like/grid-like and DLA networks: both Schelling model and redline-Schelling modification imposed at $t = 10$.

Lastly, we test social mixity vs. time for $\lambda = 0.9$, which is a relatively high threshold for agent happiness. As stated above, in normal situations, such a high threshold is actually counter-intuitive for agents since they will move so often as a result of never being satisfied that the global environment will remain relatively integrated when compared to situations where agents have lower preferences. However, we see here, even moreso than in the previous section, that implementing redline-Schelling segregation mechanisms forces segregation to still occur even at these high thresholds. The result is most stark in the DLA networks and/or at lower densities.

3.2.4 Regular DLA vs. Clustered DLA

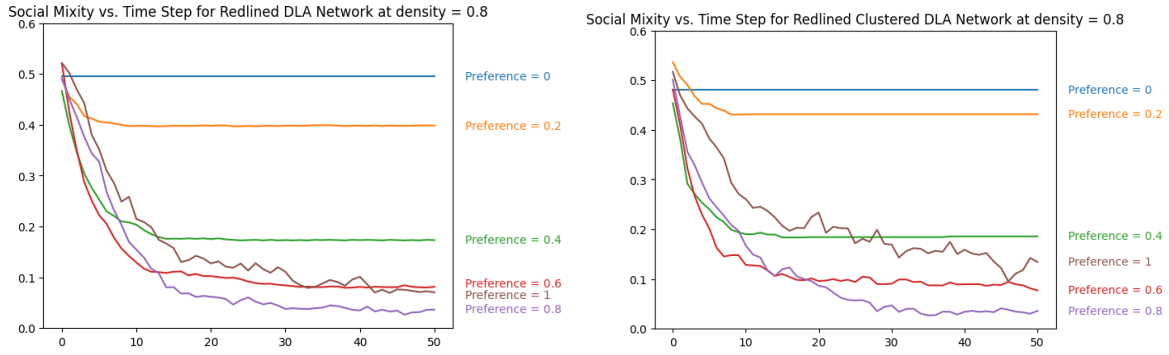


FIGURE 24. Social Mixity vs. Time for DLA and Clustered DLA networks at $d = 0.8$ and $\lambda = 0.2, 0.4, 0.6, 0.8, 1$.

In part four of our simulations on the network, we investigate the difference between two geographically relevant simulations of urban environments: the DLA and clustered DLA-generated networks. As a reminder, DLA is a conventional form of fractal generation in which one seed is placed at the origin and new particles added to the increasing network of pixels by undergoing random Brownian motion and stopping when they "hit" a pre-existing part of the graph. In this thesis, the "clustered DLA" algorithm refers to a slightly modified version of this wherein multiple seeds are generated a fixed distance away from the origin. As one can see, the difference between the two networks for $\lambda = 0, 0.2, 0.4$ and 0.8 is relatively minimal, if existent at all. Interestingly, $\lambda = 1$ has more oscillatory behavior and retains more integration in the clustered DLA network than in the normal one. More investigation is required to investigate why the additional origins in this network don't result in "clique traps" and clusters as expected at this density; we should also evaluate the cluster coefficient to see if there actually is more clustering in this form of multiple-seed DLA.

3.2.5 Average Mixity Index for Schelling vs. Redline Model

In part five of our simulations on a network, we run both the Schelling and redline-Schelling model on grid-like, DLA, and clustered DLA networks with $d = 0.8$ for similarity thresholds ranging from $\lambda = 0$ to $\lambda = 1$. The results for the grid-like graph reflect those of previous studies and confirm the validity of our self-coded model implemented in Python. Notably, the greatest segregation happens at moderate similarity thresholds (e.g., $\lambda = 0.6$ and 0.7) whereas both minimum and maximum extremes preserve integration either because very few agents feel like moving or too many agents keep moving, negating any segregating effect.

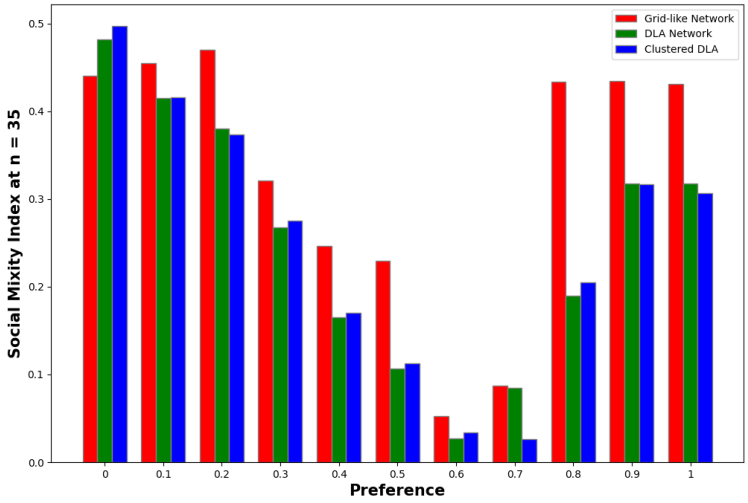


FIGURE 25. Social mixity index per similarity threshold ("preference") for Schelling model on grid-like, DLA, and clustered DLA networks.

Previous research has identified two general phrases that correspond with the social mixity index on the grid-like graph and while there are still two general phases for the DLA and clustered DLA network, they have different properties as outlined in Table 3 [36].

TABLE 3. Comparisons of social mixity index at various time steps for $d = 0.8$, $\lambda = 0.6$.

Network	**Phase Descriptions**	**λ Ranges**
Grid-like	**Phase 1:** convergence towards integration **Phase 2:** increasing convergence towards segregation	$\lambda < 0.3$ and $\lambda > 0.7$ $0.3 \leq \lambda \leq 0.7$
DLA and Clustered DLA	**Phase 1:** increasing convergence towards segregation **Phase 2:** convergence to mixed states	$\lambda < 0.8$ $\lambda \geq 0.8$

There are notable effects on end-stage segregation in the redline-Schelling model, perhaps the most robust seen in this thesis. Most importantly, there are no longer any phases in any networks that converge towards integration. Instead, the general trend is decreasing integration from $\lambda = 0$ to 0.7. The higher similarity threshold values result in a general mixed state that is still lower in the social mixity index than most general mixed states that can arise from the original Schelling model. As expected, the grid-like graph has the most integration out of the three networks because it lacks the clusters that form racial cliques and "segregation traps."

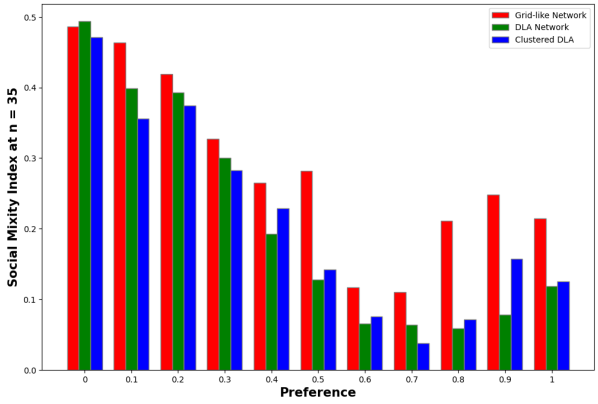


FIGURE 26. Social mixity index per similarity threshold ("preference") for redline-Schelling model on grid-like, DLA, and clustered DLA networks.

3.2.6 Role of Density in Redlined-Segregation

In our last simulation on the network, we investigate the role of density in segregation as propagated by the redline-Schelling model. As mentioned before, extremely low or high densities hold little value for the original model and frequently lead to an absence of self-organization or oscillatory behavior *but* the redline-Schelling model is able to drive segregation even at these extrema and we see that reflected in Figure 26 as the highest social mixity index achieved under this set of simulations is still under 0.35.

The scale-free Barabási-Albert network retains more integration than any other model which is in line with results from previous scholarly work [34]. The behavior for the DLA and clustered DLA networks is also mostly similar to what we expect and have seen in other simulations in this thesis. Interestingly, the grid-like graph has notably higher social mixity indices at larger densities, indicating that this fractal and clustering characteristic of DLA networks may be important in denser cities whereas modeling segregation in less populated areas may be more independent of network-type.

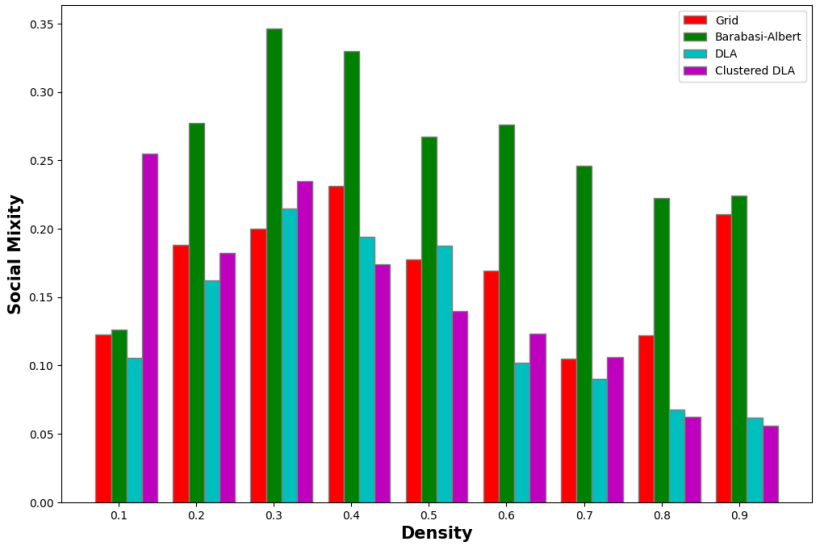


FIGURE 27. Social mixity index per density *d* for redline-Schelling model on grid-like, Barabási-Albert, DLA, and clustered DLA networks.

3.3 Mean-Field Approximation

In the last part of this thesis, we move to analysis using mean field theory. By doing so, we reduce the phenomenon into an even simpler formulation of two area designations ("best" and "undesirable") and investigate the effect of the purely *de jure* segregation mechanisms through a system of coupled, nonlinear ordinary differential equations which we will list again here for convenience:

(9) $$b_1' = b_2(1 - (b_1 + o_1)) - pb_1(1 - (b_2 + o_2))$$

(10) $$b_2' = pb_1(1 - (b_2 + o_2)) - b_2(1 - (b_1 + o_1))$$

(11) $$o_1' = o(1 - (b_1 + o_1)) - o_1(1 - (b_2 + o_2))$$

(12) $$o_2' = o_1(1 - (b_2 + o_2)) - b_2(1 - (b_1 + o_1)).$$

For more information, please refer to Section 2.6 in the Methods chapter.

3.3.1 Approximations Plotted Per Time

First, we looked at how the four populations changed with respect to time for a given beginning proportion $n = 0.30$ and a variety of probabilities $p = 0$, 0.1, 0.25, 0.5, 0.75, and 1. As a reminder, p represents the probability that a minority agent can move into any "best" neighborhood (Area B). A world with no redlining would be $p = 0$ and a world with complete *de jure* racial segregation would have $p = 1$.

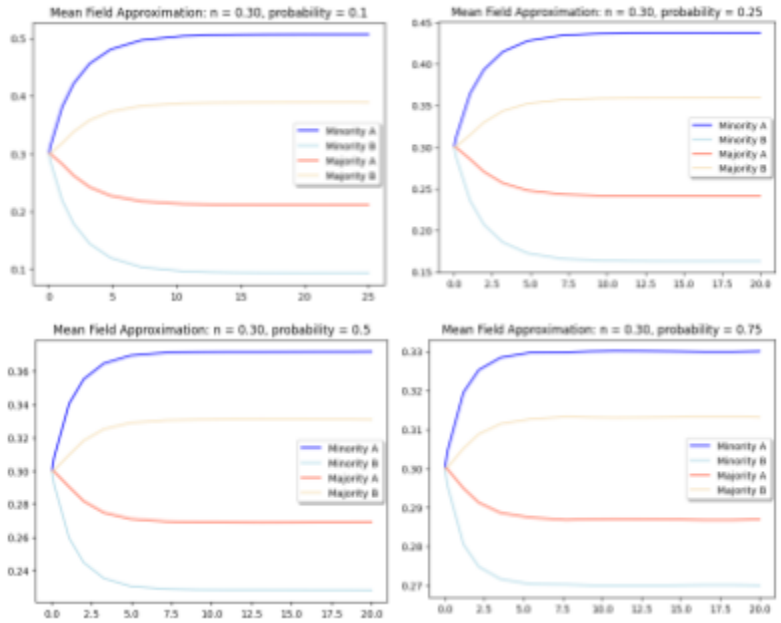


FIGURE 28. Mean field approximation per time with initial agent proportion $n = 0.30$ and success probability $p = 0.1$, 0.25, 0.5, and 0.75.

In Figure 27, we see that the simulations follow a general pattern in which the proportion of minorities in undesirable neighborhoods (Minority A) increases with time drastically while the proportion of agents who qualify as the majority in the best neighborhoods (Majority B) increases over time but not by the same margin. This clarifies that 1) segregation can happen with strictly *de jure* forces and 2) that redlining's main effect isn't so much urging majority agents into good neighborhoods or out of bad ones, but denying minorities the right of entry. Furthermore, with higher probabilities of success, the slope of the Minority A curves right before steady state gets increasingly steeper (i.e., the slope from $t = 0$ to $t = 5.0$ is steepest for $p = 0.75$ out of all simulations in Figure 27). This indicates that greater redlining forces will create environments where it takes longer for agents to reach steady state and stop moving around, indicating potential new areas of research for housing and neighborhood stability.

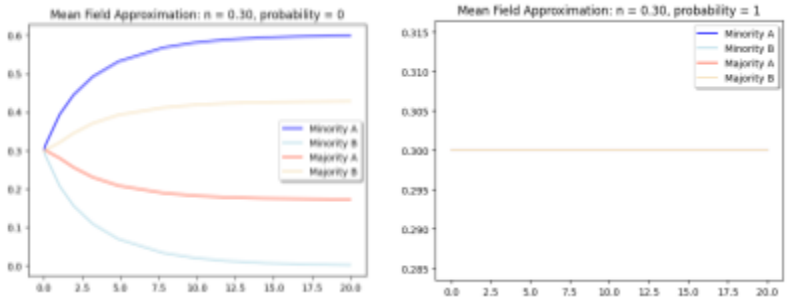


FIGURE 29. Mean field approximation per time with initial agent proportion $n = 0.30$ and success probability $p = 0$ and 1.

We also solve the system of ordinary differential equations for the complete and zero redlined worlds discussed at the beginning of this section. As expected, a world in which minority agents have no chance of moving into the best neighborhoods quickly becomes extremely segregated, where all the minority agents are in one neighborhood and not the other. There are also fewer majority agents in the "undesirable" neighborhood than in all the other simulations, indicating that not only are minority agents not evenly distributed across space, they are also not exposed to agents unlike them. If one were to simulate this using methods from Section 3.1, one would probably find extremely high dissimilarity and isolation indices. Lastly, also as expected, a world in which the probability of success p is 100 percent results in no net movement, with complete racial parity and integration preserved for all time.

3.3.2 Role of Density in Redlined Segregation: Mean-Field Version

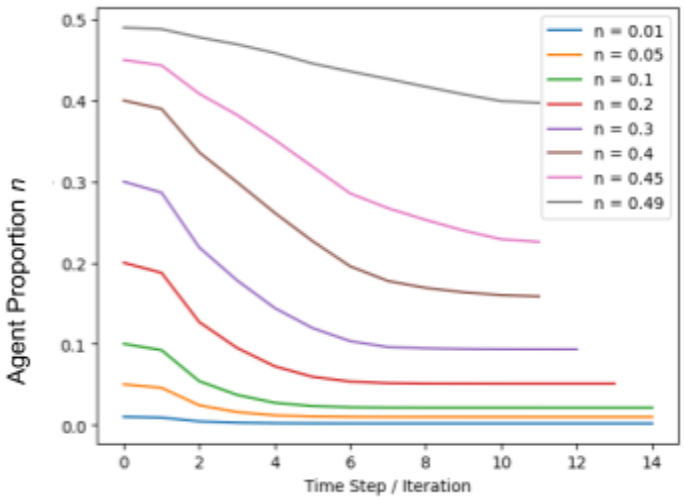


FIGURE 30. Proportion of Minority B agents per time for different proportions $n = 0.01$, 0.05, 0.1, 0.2, 0.3, 0.4, 0.45, and 0.49.

We discovered in earlier sections that density played a big role in the Schelling model in driving segregation, but not as much in the redline-Schelling version. In this mean field experiment, we tested the role of density with set probability $p = 0.2$ by solving the system of ordinary differential equations and plotting the proportion of minority agents in the best neighborhood (Minority B) against time. With extremely low densities to begin with, not much movement happens (e.g., $n = 0.01$). However, as the initial proportion of Minority B agents increases, the slope of the solution from time steps $t = 1$ to $t = 6$ gets steeper and steeper for $n = 0.05$, 0.1, 0.2, 0.3, 0.4, and 0.45. Steeper slopes indicate that movement is happening faster for higher densities, highlighting the idea that denser cities may see segregation happen faster in less dense cities. This effect does diminish at the upper extreme of n (e.g., $n = 0.49$) because there are not as many empty spots to move into for either neighborhood. Both phenomena reflect results from the simulations, confirming our previous simulations.

3.3.3 Role of Probability in Redlined Segregation

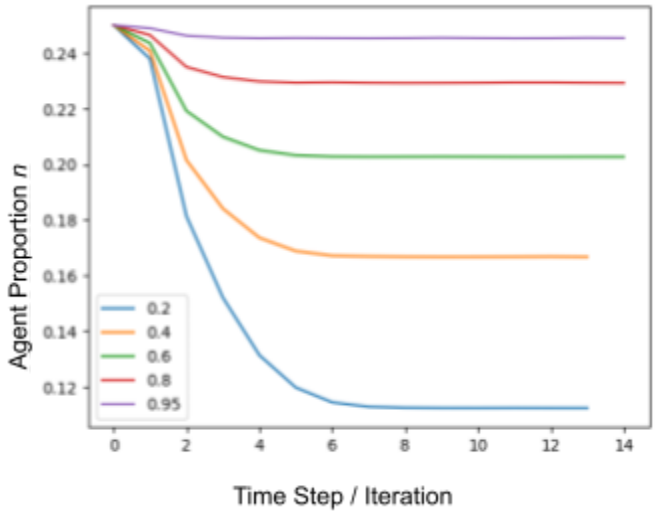


FIGURE 31. Proportion of Minority B agents per time for different success probabilities $p = 0.2$, 0.4, 0.6, 0.8, and 0.95 with $n = 0.3$ at $t = 0$.

In our network and cellular automata simulations, we did not test different probabilities to see how that would drive segregation. Here, we continue to investigate the role of probability of success in segregation through our mean field calculations. Rather than plotting the proportion per time for all four populations like in Section 3.3.1, we now plot only the agent proportion n of minority agents in the "best" neighborhoods. As one can see through Figure 30, $p < 0.5$ yields the steepest slopes and largest decreases in n whereas $p > 0.5$ results in progressively less segregation. Interestingly, all five experiments reach steady state around the same time ($t = 4$). This suggests that, without *de facto* segregation mechanisms like the in-group preference tested in the Schelling model, a sort of equilibrium can be reached in the same time frame regardless of the actual p tested—there is no situation where an oscillatory pattern would result from just *de jure* segregation mechanisms.

Chapter 4: Conclusion

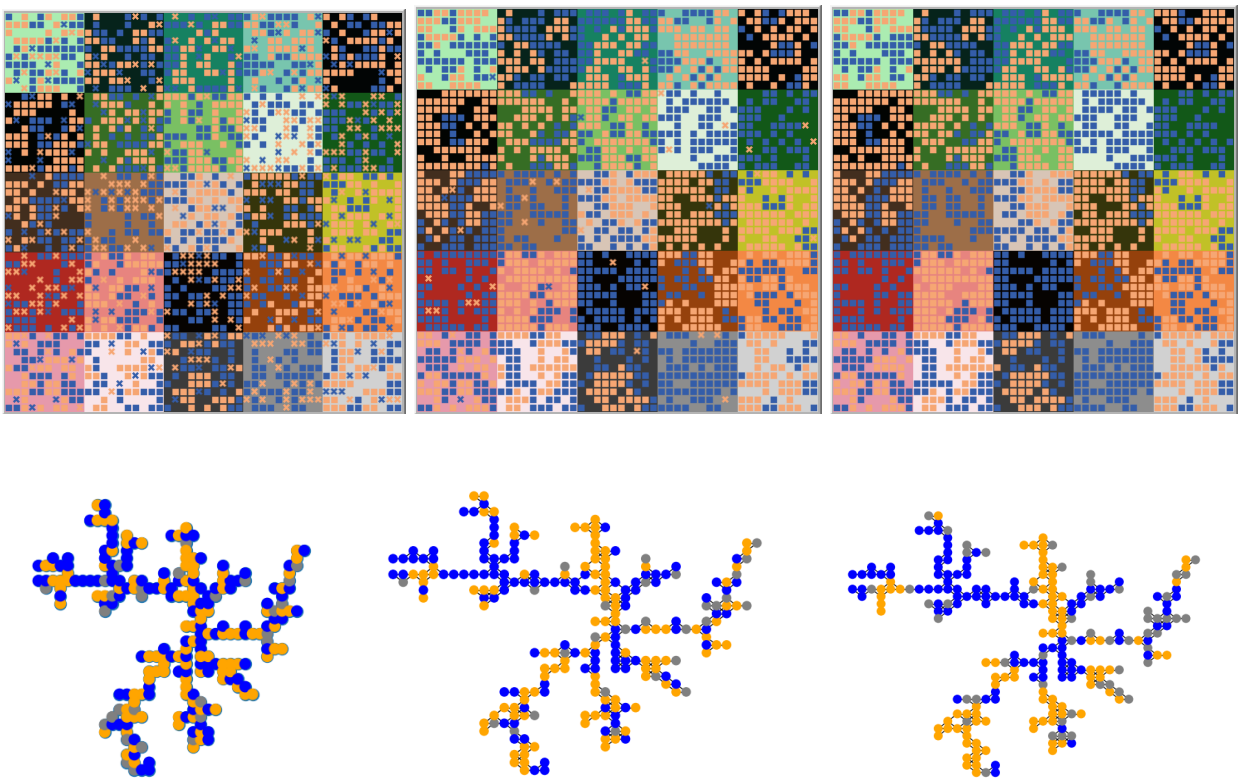


FIGURE 32. Redline-Schelling model simulations as cellular automata and on the DLA network.

The 1971 agent-based Schelling model for segregation was a seminal work in social system modeling. However, it plays into the myth of *de facto* segregation and includes only segregation driven by individual discriminatory prejudice. This makes it an insufficient model for modeling segregation in the U.S., as a large proportion of racial isolation is driven instead by *de jure* mechanisms and government policy.

In this paper, we modify the original Schelling model to account for one historically significant instance of segregation—redlining. In particular, redlining as it occurred in the United States has both *de jure* and *de facto* elements, which made it an optimal choice for modifying the Schelling model. Through the quantitative analysis of two Census Bureau segregation measures,

we found that adding *de jure* mechanisms on the grid enhances *de facto* segregation and results in much higher dissimilarity and isolation across neighborhoods.

Table 4 has the in-radius neighbor ratio results from a simulation with 50 percent density and 30 percent similarity preference, further showing how segregation is more localized near each individual agent in the modified redlining version, and that these effects are still more persistent even with increasing *r*. This same simulation set-up resulted in neighborhood activity standard deviations that were 1.56 greater in the redlining model than the original one, demonstrating that *de jure* racial segregation does profoundly impact health disparities.

TABLE 4. Average in-radius neighbor ratios from three simulations with 50 percent density and 30 percent similarity preference.

r	Original	Modified
2	2.96	3.89
4	2.02	2.60
8	1.61	2.00
15	1.43	1.66
30	1.33	1.50
50	1.30	1.40

Furthermore, *de jure* segregation as implemented by our cellular automata redlining model not only resulted in more intense racial/in-kind isolation than predicted by the most commonly cited model, it also resulted in more volatile and sensitive environments, as demonstrated by the histograms in Figures 3 and 5. This is important, as neighborhood designations and the subsequent segregation have been reported to explain up to 51 percent of opportunity in equality between neighborhoods; as mathematicians continue their pursuit of

more accurate models and legislators use their findings to generate and substantiate policy, they should ensure that their foundational models are relevant to the American context.

We not only modify the Schelling model to include redlining on the grid in this paper, but we also move it to the network to explore how segregation dynamics shift in response to different residential spatial layouts. This portion of the thesis was extremely important because most cities are not built on grids and so investigating segregation as a cellular automata model has inherent limitations. To do this portion of the thesis, we researched methods to generate both conventional/toy networks and networks that reflect realistic urban development. We ended up with testing five different networks through the course of this thesis: grid-like, Barabasi-Albert, Watts-Strogatz, diffusion-limited aggregation (DLA), and clustered/multiple seed DLA. Through a series of simulations involving the Schelling social mixity index as applied to the network, we discovered that segregation is intensified by *de jure* redlining when 1) there are segregation traps in the form of clusters or cliques, 2) the similarity threshold λ is high, and 3) when redlining is implemented on already semi-segregated networks. Perhaps most importantly for this section of the paper, redlining on the network pushed segregation in new extremes that were not seen with only *de facto* mechanisms.

Lastly, we ran several mean field calculations to confirm the validity of our model and simulations with some analytical data. These calculations, done with the assistance of a numerical solver coded with SciPy, showed that purely *de facto* segregation mechanisms still have the power to result in extreme segregation and confirm that there is a “sweet spot” of densities in which segregating forces have the most power: extremely high or low densities will result in either mixed states or integration.

There are several limitations to our work, the majority of which connect back to the choices we made in operationalizing redlining. In short, both the implementation of FHA evaluation rules and the subsequent private actions of banks, realtors, and other interested parties was extremely complex. To implement HOLC rules, we used the *Underwriting Manual* published by the FHA and specifically focused on simplistic, race-specific criteria. Given the more random nature of our starting environment on the grid, we were forced to use larger numbers for the allowed percentage of minority agents in each category, whereas these restrictions were much more stringent in real life—even one Black resident was enough to exclude a neighborhood from being labeled "best" [14]. Additionally, we did not account for "border conditions" listed in the *Underwriting Manual,* wherein the proportion of Black individuals surrounding a given neighborhood could also bring down its value. Numerous other factors were also left unaccounted for, such as immigration status or family size. These modeling limitations had measurable effects on our models, as modern data shows that dissimilarity indices in American cities reached a high of 87.90 percent, but the highest dissimilarity index in the modified model is only around 65 percent [37]. The universal preference rules applied to all agents is another simplification that we make for the sake of the model. Research in the social sciences has suggested however that this is not the case in reality. Instead, white people typically prefer to live in neighborhoods with very few Black neighbors (around 20 percent) whereas some Black people in some settings have been shown to prefer "50-50" neighborhoods [41].

Lastly, it is important to note that overt redlining and *de jure* residential segregation was outlawed by the Fair Housing Act of 1968 [36]. However, modern segregation still persists through other manifestations of systemic racism and there are indeed examples of "modern redlining" in which people of color are being denied bank loans or mortgages with very little

rhyme or reason [38]. This research is therefore but a first step in altering traditional forms of modeling to better suit reality in the 21st-century; immediate next steps may include attempting to model gentrification or other unfair housing practices that result in displacement. Additionally, it would be interesting to validate the simulations using real data. Though the results in this thesis are interesting and introduce the idea that the original Schelling model is not enough to represent segregation as it happened in the 20th century United States, more research must be done to more confidently assert that the findings of this thesis have practical significance.

In summary, in this thesis, we learned about state-of-the-art mathematical models, researched historical deviations from the assumptions made in these models, implemented and operationalized redlining as a mathematical model on both the grid and on various networks, ran a variety of simulations to test the effects of density, network type, preference levels, etc. between the two models, and tested pure *de jure* mechanisms by running analytics using mean field theory. Through our research, we learned that *de facto* segregation is intensified by *de jure*/redlining mechanisms, redlining drives segregation at similarity thresholds and densities that normally result in integration, network choice matters at specific densities and preferences (e.g., segregation traps uniquely occur in fractal-like urban structures only when densities are high), and that *de jure*/redlining mechanisms on their own are able to result in segregation.

Historical relevance, practical significance, and a desire to understand and account for society's short-comings is still a main motivation of this study. As Richard Rothstein wrote, "Little has been done to desegregate neighborhoods because we think it is the result of private prejudice instead of explicit government policy" [37]. Though investigating mathematical theory and running simulations is interesting, applied math must be pursued in relation to a real-world problem in order to deliver its full value. This pursuit involves not only finding and defining a

problem but also improving methods previously introduced and re-evaluating them in light of new information. We began this thesis because we noticed that it neglected to tell the story of a marginalized population in the United States whose livelihoods were negatively impacted by governmental policy. Applied math is an extremely useful tool to not only bring awareness to the issue, but also to work on potential solutions. We therefore urge mathematicians to continue studying mathematical models and using them to describe complex social phenomena, to better the world and the discipline simultaneously.

References

[1] Encyclopædia Britannica, inc. (n.d.). Segregation. *Encyclopædia Britannica*. https://www.britannica.com/topic/segregation-sociology

[2] Encyclopædia Britannica, inc. (n.d.). Minority. *Encyclopædia Britannica*. https://www.britannica.com/topic/minority

[3] Townsley, J., Andres, U. M., Nowlin, M. (2021). The Lasting Impacts of Segregation and Redlining. *SAVI.* https://www.savi.org/2021/06/24/lasting-impacts-of-segregation/.

[4] McCown, F. (n.d.). *Schelling's Model of Segregation*. Nifty Assignments. http://nifty.stanford.edu/2014/mccown-schelling-model-segregation/

[5] Schelling, T. (1971). Dynamic models of segregation. *The Journal of Mathematical Sociology*, 1(2), 143–186.

[6] Fowler, A. C. (1997). Mathematical Models in the Applied Sciences. United Kingdom: Cambridge University Press.

[7] Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, *99*(suppl_3), 7280–7287. https://doi.org/10.1073/pnas.082080899

[8] Conte, R., & Paolucci, M. (2014). On agent-based modeling and Computational Social Science. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00668

[9] Kari, J. (2005). Theory of cellular automata: A survey. *Theoretical Computer Science*, *334*(1–3), 3–33. https://doi.org/10.1016/j.tcs.2004.11.021

[10] Shiffman, D. (2012). *Chapter 7: Cellular Automata*. The Nature of Code. https://natureofcode.com/book/chapter-7-cellular-automata/

[11] Zhang, Junfu. "A dynamic model of residential segregation." Journal of Mathematical

Sociology 28.3 (2004): 147-170.

[12] Rothstein, R. (2018). *The color of law: A forgotten history of how our government segregated America*. Liveright Publishing Corporation, a division of W.W. Norton & Company.

[13] Rose, J. (2012). *Revisiting How Two Federal Housing Agencies Propagated Redlining In The 1930s*. Federal Reserve Bank of Chicago. https://www.chicagofed.org/research/mobility/policy-brief-federal-housing-programs-redlining

[14] *Underwriting Manual: Underwriting Valuation and Procedure Under Title II of the National Housing Act*. (1938). Federal Housing Administration.

[15] *Tracing the legacy of redlining: A new method for tracking the origins of housing segregation* . NCRC. (2022, July 8). https://ncrc.org/redlining-score/

[16] Best, R., & Mejía, E. (2022, February 9). *The Lasting Legacy Of Redlining*. FiveThirtyEight. https://projects.fivethirtyeight.com/redlining/#:~:text=Thepercent20legacypercent20ofpercent20redliningpercent20extendspercent20farpercent20beyondpercent20housing,arepercent20morepercent20pronepercent20topercent20floodingpercent20andpercent20extremepercent20heat.

[17] Committee on Network Science for Future Army Applications (2006). Network Science. National Research Council. doi:10.17226/11516

[18] Capps, K. (2018, April 11). *50 years after the Fair Housing Act, redlining persists*. Bloomberg.com.

https://www.bloomberg.com/news/articles/2018-04-11/50-years-after-the-fair-housing-act-redlining-persists

[19] Morris, W. (1902). *Architecture, industry, and wealth: Collected papers*. Garland Pub.

[20] Batty, M. (2008). The size, scale, and shape of Cities. *Science*, *319*(5864), 769–771. https://doi.org/10.1126/science.1151419

[21] Encyclopædia Britannica, inc. (n.d.). Fractal. *Encyclopædia Britannica*. https://www.britannica.com/science/fractal

[22] Encyclopædia Britannica, inc. (n.d.). Brownian motion. *Encyclopædia Britannica*. https://www.britannica.com/science/Brownian-motion

[23] Halsey, T.C. (2000) Diffusion-Limited Aggregation: A Model for Pattern Formation. *Physics Today, 53*(11), 36-41. https://doi.org/10.1063/1.1333284

[24] Fotheringham, S., Batty, M., & Longley, P. (1989). Diffusion-limited aggregation and the fractal nature of urban growth. *Papers of the Regional Science Association*, *67*(1), 55–69. https://doi.org/10.1007/bf01934667

[25] Batty, M., Longley, P., & Fotheringham, S. (1989). Urban growth and form: Scaling, fractal geometry, and diffusion-limited aggregation. *Environment and Planning A: Economy and Space*, *21*(11), 1447–1472. https://doi.org/10.1068/a211447

[26] Makse, H., Havlin, S. & Stanley, H. (1995). Modelling urban growth patterns. *Nature* 377, 608–612. https://doi.org/10.1038/377608a0

[27] Hsu, C.-L. (2020). *Watts-Strogatz model of small-worlds*. An Explorer of Things. https://chih-ling-hsu.github.io/2020/05/15/watts-strogatz

[28] U.S. Census Bureau. (2021, November 21). *Housing patterns: Appendix B: Measures of*

residential segregation. Census.gov. https://www.census.gov/topics/housing/housing-patterns/guidance/appendix-b.html

[29] Newman, M. E. (2003). Mixing patterns in networks. *Physical Review E*, *67*(2). https://doi.org/10.1103/physreve.67.026126

[30] Barabási, A.L. (2014). *Network Science: The Barabási-Albert Model.*

[31] Riley, A. R. (2018). Neighborhood disadvantage, residential segregation, and beyond—lessons for studying structural racism and health. *Journal of racial and ethnic health disparities*, 5(2), 357-365.

[32] Williams, D. R., & Collins, C. (2001). Racial residential segregation: a fundamental cause of racial disparities in health. *Public health reports*, 116(5), 404.

[33] Centers for Disease Control and Prevention. (2022). Adult Physical Inactivity Prevalence Maps by Race/Ethnicity. *Centers for Disease Control and Prevention*. https://www.cdc.gov/physicalactivity/data/inactivity-prevalence-maps/index.html.

[34] Arnaud Banos. (2012) Network effects in Schelling's model of segregation: new evidences from agent-based simulation. *Environment and Planning B: Planning and Design*, 39 (2), 393-405. https://doi.org/10.1068/b37068ff.

[35] Segregation: Dissimilarity Indices. (n.d.) *CensusScope*. https://censusscope.org/us/rank_dissimilarity_white_black.html.

[36] The Fair Housing Act, 42 U.S.C. § 3601 (1968). https://www.justice.gov/crt/fair-housing-act-1.

[37] Rothstein, R. (2017, May 16). America is still segregated. We need to be honest about why. *The Guardian*.

https://www.theguardian.com/commentisfree/2017/may/16/segregation-us-neighborhoods-reasons.

[38] Martinez, E., & Glantz, A. (2021, June 30). *Modern-day redlining: Banks discriminate in lending*. Reveal. https://revealnews.org/article/for-people-of-color-banks-are-shutting-the-door-to-homeownership/

[39] Krysan, M., Couper, M. P., Farley, R., & Forman, T. A. (2009). Does race matter in neighborhood preferences? Results from a video experiment. *AJS; American journal of sociology*, *115*(2), 527–559. https://doi.org/10.1086/599248

[40] Silver, D., Byrne, U., & Adler, P. (2021). Venues and segregation: A revised Schelling model. *PloS One*, 16(1), https://doi.org/10.1371/journal.pone.0242611.

[41] Zhang, J. (2011). Tipping and residential segregation: a unified Schelling model. *Journal of Regional Science*, 51(1), 167-193.

[41] Benard, S., & Willer, R. (2007). A wealth and status-based model of residential segregation. *Mathematical Sociology*, 31(2), 149-174.

[42] Ray, R., Perry, A., Harshbarger, D., Elizondo, S., & Gibbons, Al. (2021). Homeownership, racial segregation, and policy solutions to racial wealth equity. *Brookings Institution.* https://www.brookings.edu/essay/homeownership-racial-segregation-and-policies-for-racial-wealth-equity/.

[43] The Myth of "De Facto Segregation": How Land-Use Policies Created a Segregated America. (2018). *Center for Planning Excellence*. https://www.cpex.org/blog/the-myth-of-de-facto-segregation-how-land-use-policies-created-a-segregated-america.