Modified Schelling Model

The Effects of Infrastructure on Neighborhood Construction

The inclusion of urban geographical restraints into the Schelling Model will provide greater insight into the effects of certain spatial limitations on the formation of self-segregating neighborhoods. This paper seeks to understand the impact of urban blocks on these neighborhoods and how the area of such blocks relates to the number of neighborhoods created.

Geoffrey Stevens

Table of Contents

I. Introduction

II. Motivation

III. Simulation Setup and Parameters

IV. Analysis Variables

V. One Block Model

VI. Four Block Model

VII. Nine Block Model

VIII. Twenty-Five Block Model

IX. General Conclusions

X. Further Research

I. Introduction

The Schelling Model is an idea put forth by Thomas Schelling 1969 and 1971 that dealt with the concept of racial segregation and its motivations. Through a mathematical simulation, Schelling revealed that even a slight preference for similar neighbors or a willingness to accept diversity up to a certain degree can lead to total segregation of populations. In essence, Schelling showed the potential for individual decisions, harmless on a micro-level, to greatly and surprisingly affect macro-level behavior. In 1978, Schelling published *Micromotives and Macrobehavior* to articulate the mechanisms behind the Schelling Model and provide a compelling use of agent-based simulations within the field of economics.

The Schelling Model, as an agent-based simulation, is a grid composed of two types of actors: (1) empty spaces and (2) members of different populations. Each agent is assigned a membership into a certain population and is randomly dispersed throughout a grid, alongside empty spaces that are not occupied. At each iteration, each agent will decide to move to a new location or remain in its current location depending on whether or not the number of neighbors surrounding it reach a certain threshold of similarity.

II. Motivation

One component, however, that the Schelling Model lacks is the inclusion of geographical mechanisms that could potentially affect the movement of populations within the grid. In modern cities, critical infrastructure can dramatically shape the demographics of a given location. Often, train tracks or large highways appear to divide neighborhoods, sometimes along racial or socioeconomic lines, in a way that seems to suggest that there are additional nuances to the Schelling Model that have not been considered. Most urban residents are familiar with the phrase “wrong side of the tracks” or “wrong side of the street.” Can the Schelling Model explain these colloquialisms?

III. Simulation Setup and Parameters

To explore the influence of infrastructure on neighborhood development, a model was built that adds additional agents to the Schelling Model. These new agents, called BORDER, are encoded as black units and cannot be occupied or moved. EMPTY spaces are encoded in the grid as white units and will represent 10% of the available grid space (e.g. the entire grid not including BORDER units). For the purposes of this simulation, two populations, RED and BLUE, occupy 45% of the available grid space each. The grid is 24x24 units and contains 176 BORDER units.

The relocation algorithm allows an agent to randomly select an EMPTY space to move to if more than 50% of its neighbors are dissimilar. A neighbor is defined as an agent that is located either to the North, South, East, West, Northeast, Southeast, Northwest, or Southwest (e.g. the neighbor function includes both the cardinal directions and the diagonals.

The different trials that will be conducted will focus on exploring the size and shape of different infrastructure configurations. These simulations will be referred to as One Block Model, Four Block Model, Nine Block Model, and Twenty-Five Block Model. Each of these models will build upon the last model to subdivide the grid into city-like blocks in order to analyze the effects of infrastructure on segregation patterns. The simulator can be viewed at https://geoffstevens8.github.io/modified-schelling-model/.

IV. Analysis Variables

In the process of analyzing each different setup, there are certain variables that will be referenced later that are important to define:

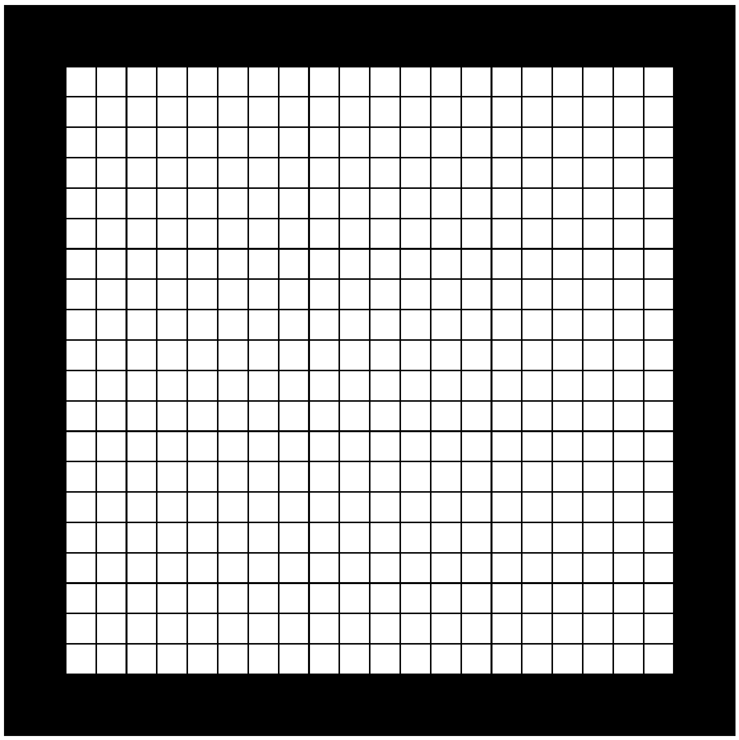
Block:
A block refers to a segment of the grid that is created by surrounding the location with BORDER units in a fashion that is similar to blocks in modern day cities

Neighborhood
A neighborhood is a contiguous mass of agents from the same population. Within a neighborhood, it is possible to reach all other members of the neighborhood by jumping between neighbors of a single population, either RED or BLUE.

Occupation
Occupation will be used in two analyses for the following trials. The first measurement will refer to the percent of blocks occupied, which represents the fraction of blocks that have a given population present. The second measurement will refer to the percent occupation of a block, which reveals how many agents within a block belong to a given population divided by the total number of agents in that block.

V. One Block Model

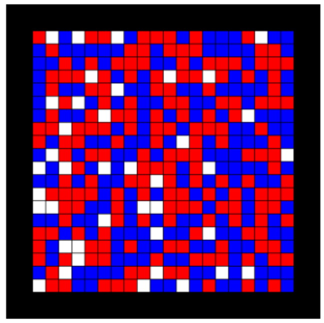
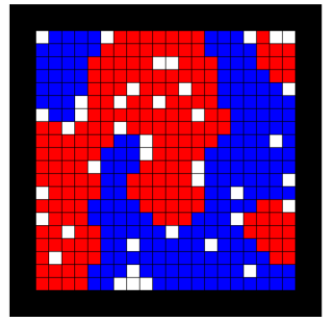
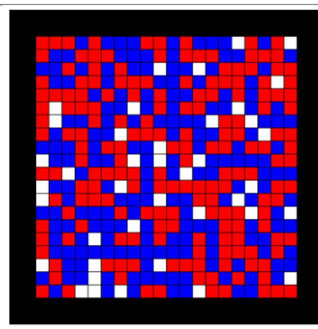
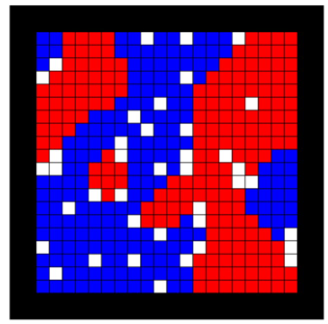
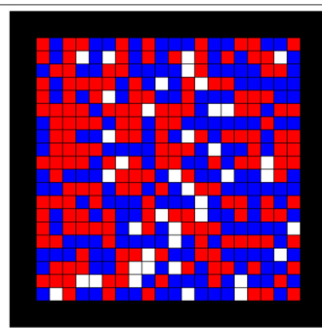
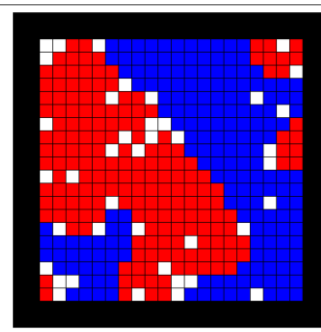
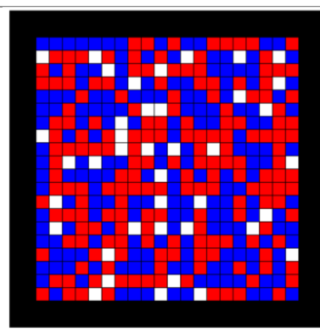
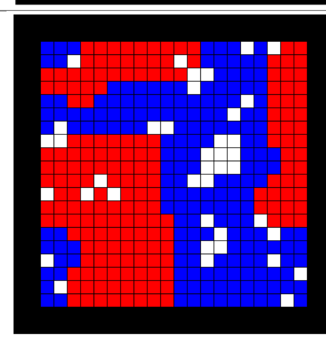
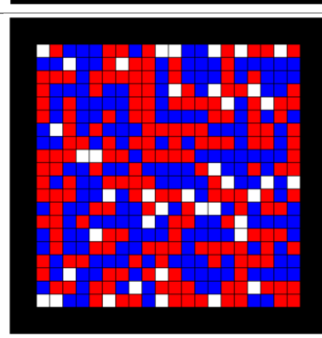
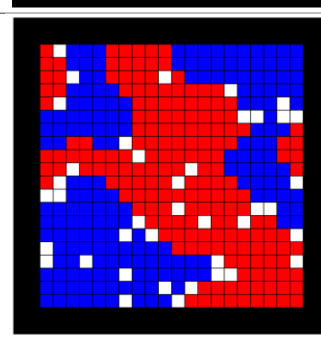
In order to provide some form of control group to which later simulations with infrastructure can be compared, it is important to generate quantifiable characteristics in a grid without any infrastructure (i.e. a grid composed of one large block). Through these simulations, a baseline for agent behavior in the Schelling Model can be produced to reveal the number of neighborhoods created and the percent of the population that has at least one dissimilar neighbor. The setup of the grid appears as follows:



The grid above contains 400 units in a 20x20 pattern. The grid will be filled with 180 RED agents, 180 BLUE agents, and 40 EMPTY units. An equivalent construction in a modern city would most likely be an area without any large infrastructure projects bisecting the residential neighborhood.

Simulations

The simulations were run for five different trials, involving a randomized initial positioning. Each trial continues iterating until a sufficient percentage, above 99%, of agents within the grid are satisfied. The trials generated the following results:

	START	END
Trial 1		
Trial 2		
Trial 3		
Trial 4		
Trial 5		

Analysis

From these simulations, the number of neighborhoods and the percent of each population with dissimilar neighbors can be calculated, as shown in the charts below:

Number of Neighborhoods		
	BLUE	RED
Trial 1	2	3
Trial 2	3	3
Trial 3	2	4
Trial 4	3	3
Trial 5	3	3
Average	2.6	3.2

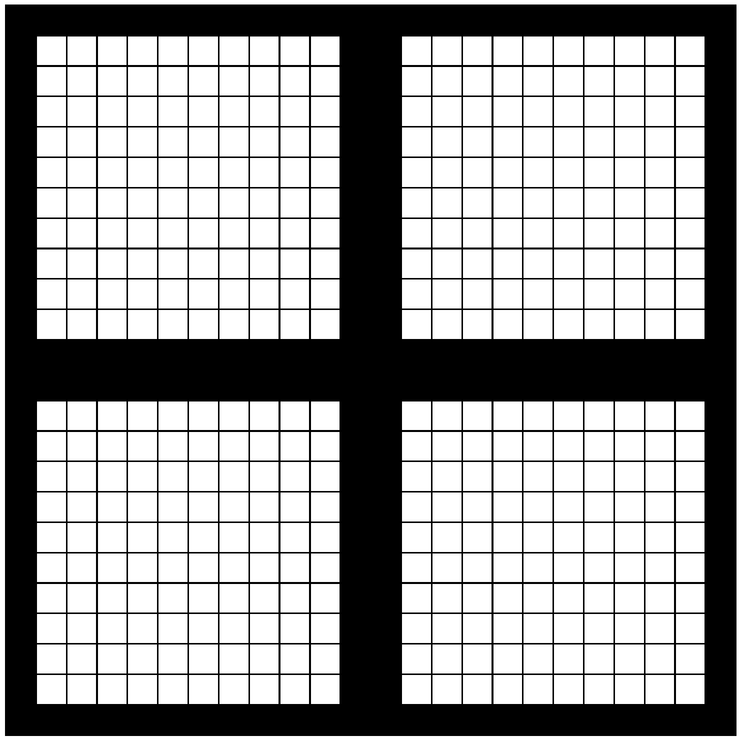
The number of neighborhoods per trial averaged around six, meaning that during each simulation there would be approximately three RED neighborhoods and three BLUE neighborhoods. Looking at the percent of the population that has a dissimilar neighbor can reveal important information about the One Block Model as well:

Percent of Population with Dissimilar Neighbors		
	BLUE	RED
Trial 1	43.3%	38.9%
Trial 2	38.3%	37.8%
Trial 3	32.8%	29.4%
Trial 4	37.2%	38.9%
Trial 5	39.4%	39.4%
Average	38.2%	36.9%

These results show the expected behavior of the Schelling Model, with slight preferences for similar neighbors leading to a largely segregated grid. These trials will provide an important background for the discussion of infrastructure and its impact on the mechanics of the Schelling Model.

VI. Four Block Model

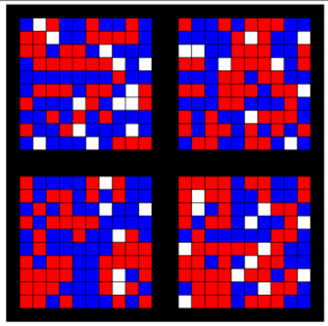
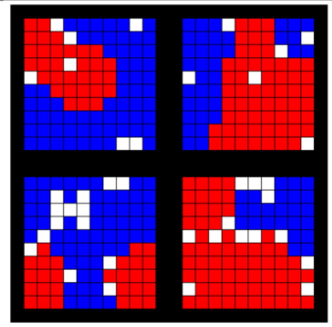
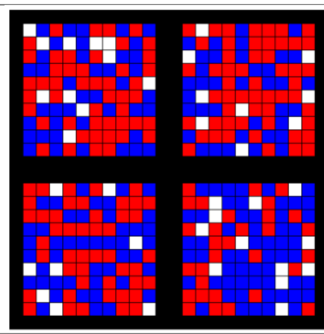
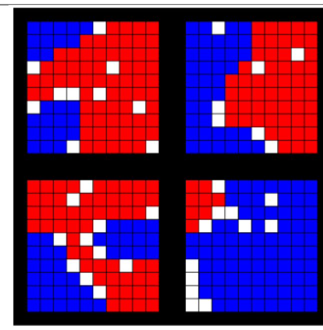
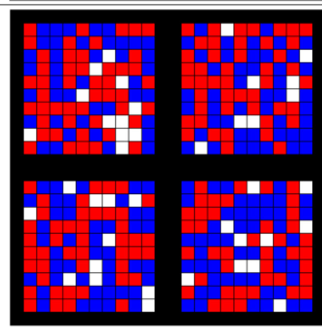
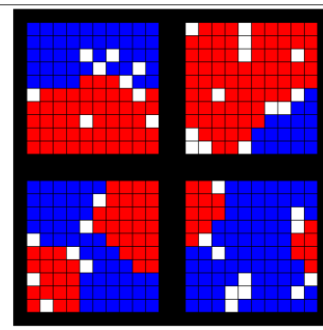
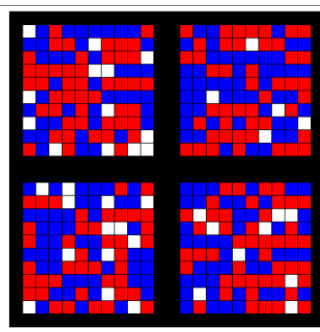
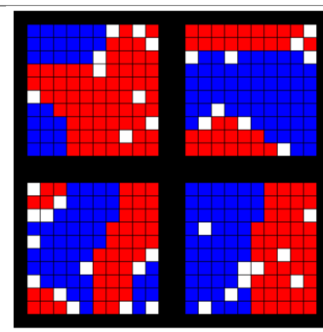
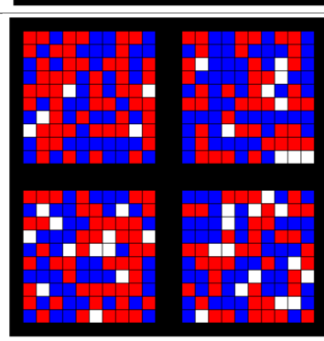
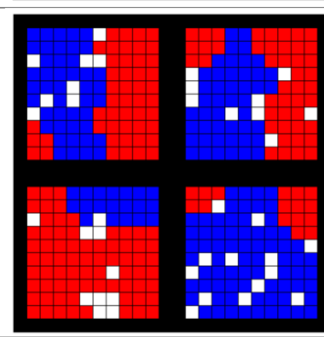
The next iteration of the Schelling Model will contain infrastructure that divides the grid into four equally-shaped quadrants with height and width of 10 units. The total number of units in the simulation remains at 400 in order to minimize any potentially confounding factors regarding number of agents. The setup of the four quadrants appears below:



The creation of four quadrants that are exactly one-fourth as large as the One Block Model reflects a modern-day pattern in infrastructure that could be found in suburbia, whereby only a few large infrastructure projects bisect the geography.

Simulations

The simulations were run for five different trials, involving a randomized initial positioning. Each trial continues iterating until a sufficient percentage, above 99%, of agents within the grid are satisfied. The trials generated the following results:

	START	END
Trial 1		
Trial 2		
Trial 3		
Trial 4		
Trial 5		

Analysis

From these simulations, like the pure Schelling Model in earlier sections, the number of neighborhoods and the percent of each population with dissimilar neighbors can be calculated. In addition, measurements for the percent of blocks occupied and the evolution of the percentage of each block occupied can be obtained. First, the number of neighborhoods per trial:

Number of Neighborhoods		
	BLUE	RED
Trial 1	5	5
Trial 2	6	4
Trial 3	4	6
Trial 4	6	7
Trial 5	4	7
Average	5	5.8

In comparison to the Schelling Model with no infrastructure, the number of neighborhoods has nearly doubled on average, from approximately 3 to slightly less than 6. Next, the percent of blocks occupied by each population:

Percent of Blocks Occupied		
	BLUE	RED
Trial 1	100%	100%
Trial 2	100%	100%
Trial 3	100%	100%
Trial 4	100%	100%
Trial 5	100%	100%
Average	100%	100%

Although quite trivial, each population has a presence in each of the four blocks, which also held true in the Schelling Model with no infrastructure. A slight change has occurred in dissimilar neighbors:

Percent of Population with Dissimilar Neighbors		
	BLUE	RED
Trial 1	33.3%	30.6%
Trial 2	25.0%	27.2%
Trial 3	27.2%	24.4%
Trial 4	30.0%	33.9%
Trial 5	29.4%	27.2%
Average	29.0%	28.7%

Quite interestingly, however, the percent of each population with at least one dissimilar neighbor has decreased by around 10 percentage points when shifting from the unmodified Schelling Model and the grid with four blocks. Thus, the creation of four blocks leads to around a 23% decrease in the percentage of the population with dissimilar neighbors. With the regards to each individual block, if we look at Trial 1, we can see evidence of entrenched populations:

Occupation of Block				
	Trial 1			
	BLUE		RED	
BLOCK	Start	End	Start	End
1	57.5%	66.7%	42.5%	33.3%
2	49.5%	38.3%	50.5%	61.7%
3	52.1%	67.4%	47.9%	32.6%
4	40.9%	27.6%	59.1%	72.4%

BLUE		END	
		< 50%	> 50%
START	< 50%	2	0
	> 50%	0	2

RED		END	
		< 50%	> 50%
START	< 50%	2	0
	> 50%	0	2

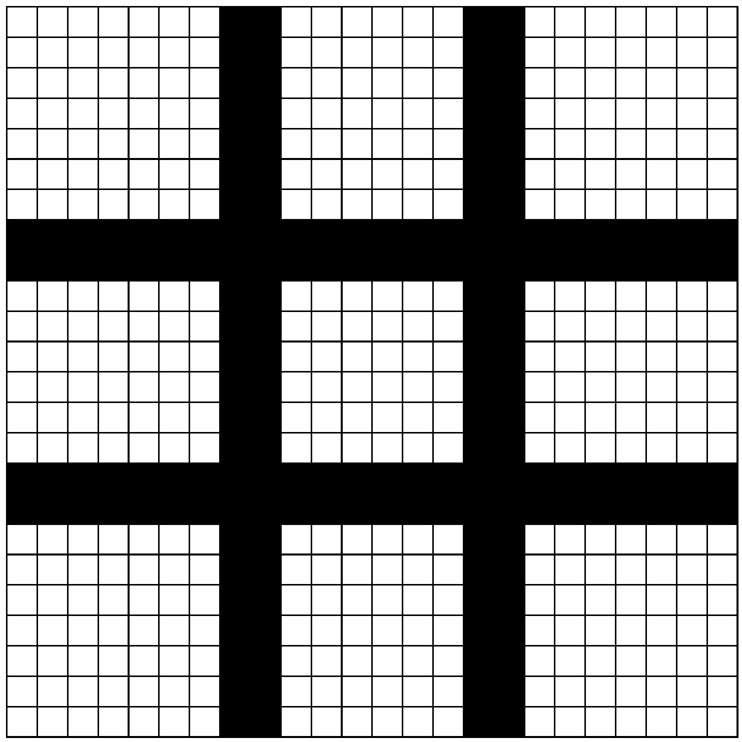
These characteristics reveal that, in this trial, if a population has a greater presence in a given block at the beginning of a simulation, they are more likely to hold the majority in that particular block.

For the Four Blocks simulation, we can reveal a few different conclusions, especially in relation to the One Block Model:

1. There is evidence of entrenched populations, whereby a population that has a majority in a given block is more likely to remain the majority in that block. Even as a majority, however, the majority population only ever achieves around 70% occupation of a block, which, as we will see in later simulations, is not as divisive or dramatic as other models.
2. Similar to the previous conclusion, the creation of large-enough subdivisions does not necessarily necessitate total segregation by block. Within grids with blocks of both size 10x10 and 20x20, the creation of exclusive blocks is extremely rare in the former case and impossible in the latter.
3. The addition of only two major segments of infrastructure leads to a fairly substantial decrease in the percent of each population that has at least one dissimilar neighbor.
4. The addition of only two major segments of infrastructure leads to a doubling of the number of neighborhoods created. The Schelling Model without infrastructure contains 84 BORDER units that can be abutted by an agent, while the Four Block Model contains 176 BORDER units. The doubling of BORDER units coincides with the doubling of neighborhoods.

VII. Nine Block Model

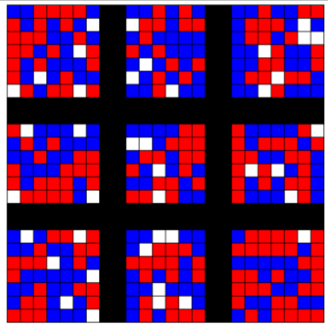
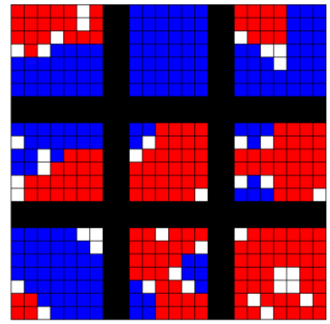
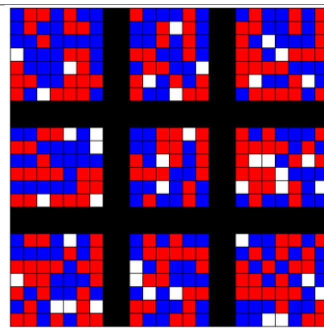
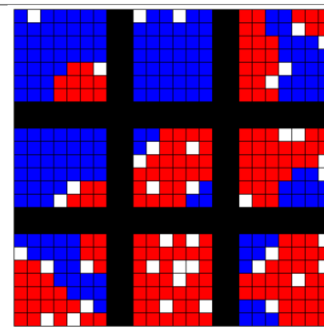
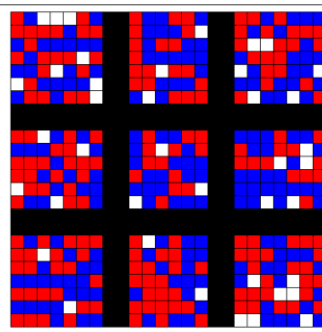
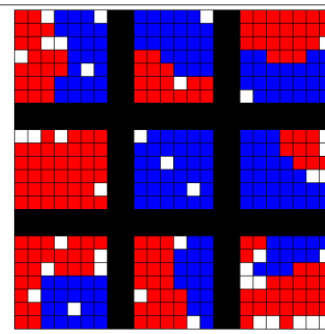
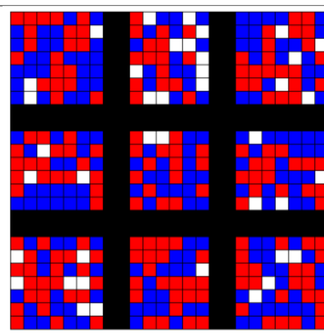
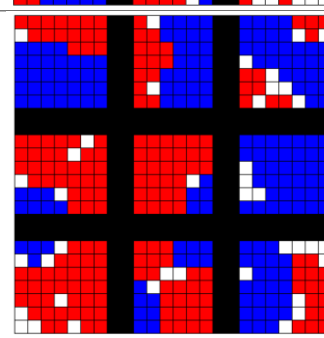
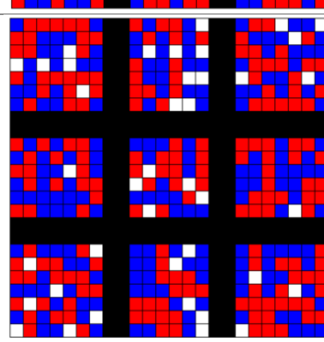
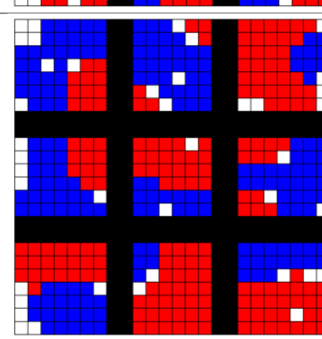
In order to further explore the effects of shrinking block dimensions, the next iteration of the Schelling Model contains nine blocks. Due to the properties of the grid and the desire to maintain a similar number of agents across models, the Nine Blocks Model contains four 7x7 blocks, three 7x6 blocks, and one 6x6 block. The setup is shown below:



As the subdivisions continue to get smaller in each model, the grid becomes more similar to the types of blocks that would typically be seen in an urban environment. The growing density of the infrastructure and the increasingly divisive nature of the infrastructure reveals a pattern common in many planned, modern cities.

Simulations

The simulations were run for five different trials, involving a randomized initial positioning. Each trial continues iterating until a sufficient percentage, above 99%, of agents within the grid are satisfied. The trials generated the following results:

	START	END
Trial 1		
Trial 2		
Trial 3		
Trial 4		
Trial 5		

Analysis

From these simulations, we can compare the number of neighborhoods, the percent of bocks occupied, the percent of each population with dissimilar neighbors, and the possibilities of entrenched populations. First, the number of neighborhoods:

Number of Neighborhoods		
	BLUE	RED
Trial 1	10	8
Trial 2	10	10
Trial 3	8	9
Trial 4	10	9
Trial 5	9	11
Average	9.4	9.4

The average number of neighborhoods created during each simulation hovers a slightly under twice that of the Four Block Model (with an average slightly above 5) and slightly under four times that of the One Block Model (with an average slightly above 2). Next, the percent of blocks occupied:

Percent of Blocks Occupied		
	BLUE	RED
Trial 1	89%	89%
Trial 2	89%	89%
Trial 3	89%	89%
Trial 4	100%	89%
Trial 5	100%	100%
Average	93%	91%

Also noticeable is the change between the Four Block Model, which had each population with a presence in each block during each simulation, and the Nine Block Model, in which we begin to see the creation of exclusionary blocks that are occupied by only one population. The percent of each population with dissimilar neighbors reveals another interesting pattern:

Percent of Population with Dissimilar Neighbors		
	BLUE	RED
Trial 1	19.4%	24.4%
Trial 2	27.8%	30.0%
Trial 3	29.4%	27.2%
Trial 4	26.1%	26.1%
Trial 5	33.3%	30.6%
Average	27.2%	27.7%

The percent of each population with dissimilar neighbors has surprisingly not changed much from the Four Block Model, shifting downward by only perhaps a percentage point. The transition from four to nine blocks leads to a minimal effect on the percentage of people that have a neighbor different from themselves.

Occupation of Block				
	Trial 1			
	BLUE		RED	
BLOCK	START	END	START	END
1	45.5%	56.8%	54.5%	43.2%
2	63.2%	100.0%	36.8%	0.0%
3	55.8%	75.6%	44.2%	24.4%
4	50.0%	48.6%	50.0%	51.4%
5	46.9%	9.1%	53.1%	90.9%
6	47.4%	19.0%	52.6%	81.0%
7	53.5%	90.9%	46.5%	9.1%
8	52.8%	27.0%	47.2%	73.0%
9	37.5%	0.0%	62.5%	100.0%

BLUE		END	
		< 50%	> 50%
START	< 50%	3	1
	> 50%	1	3

RED		END	
		< 50%	> 50%
START	< 50%	3	1
	> 50%	1	3

As with the previous models, there is a strong correlation between beginning location and final locations with respect to which population remains the majority in a given block. In the example trial above, if a block is composed of more than 50% of a given population, there is a 75% that the given population remains the majority.

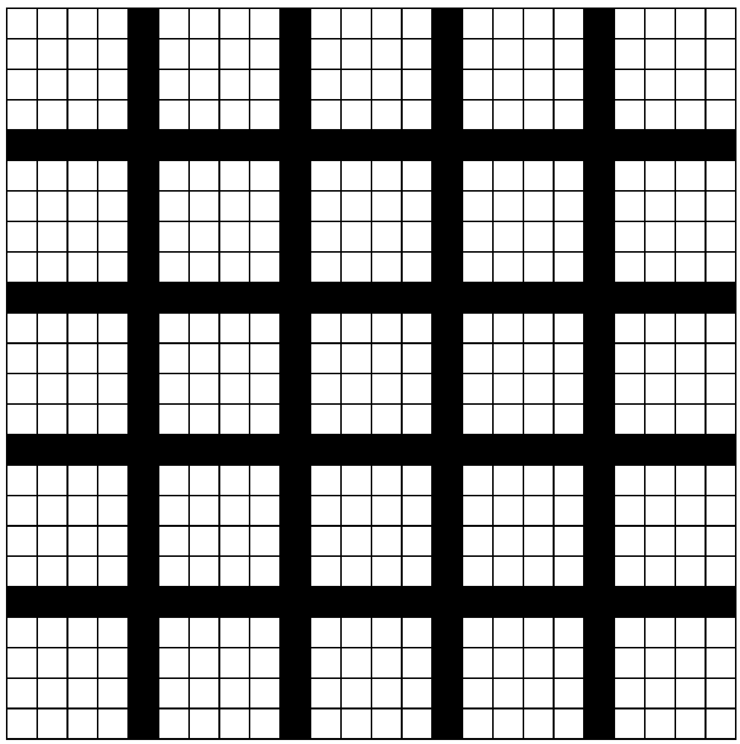
Based on the results above and with respect to the previous models, we can thus draw some temporary conclusions:

1. Entrenched populations continue to exist and the disparities in the final percentage breakdown of each block is becoming more extreme.
2. However, similar to the Four Block Model, complete segregation by block is surprisingly uncommon, but generally occurs at least once during each simulation, which is a dramatic change from the large block models.
3. The addition of four major segments of infrastructure leads to a fairly negligible decrease in the percent of each population that has at least one dissimilar neighbor when compared to the addition of two pieces of infrastructure.
4. The addition of four major segments of infrastructure leads to a near doubling of the number of neighborhoods created when compared to the Four Block Model. The Schelling Model without infrastructure contains 84 neutral neighbors (border units) that

can be abutted by an agent, the Four Blocks Model contains 176 neutral neighbors, and the Nine Block Model contains 276 neutral neighbors. Again, the doubling of border units roughly coincides with the doubling of neighborhoods.

VIII. Twenty-Five Block Model

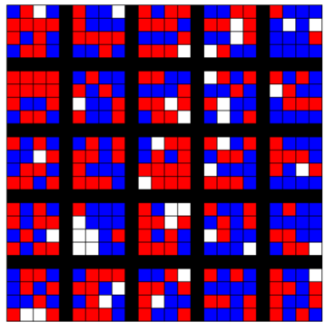
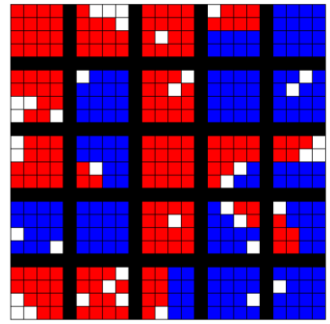
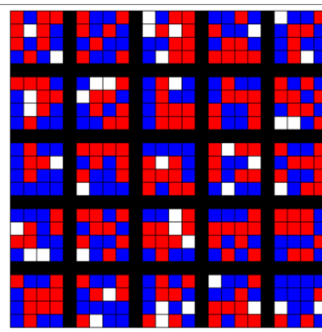
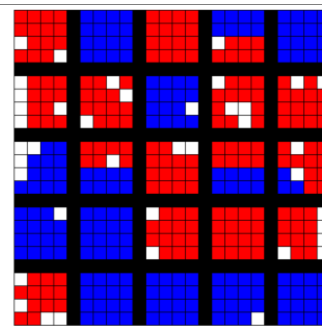
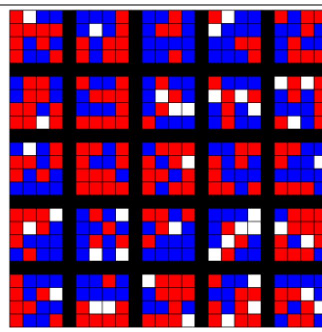
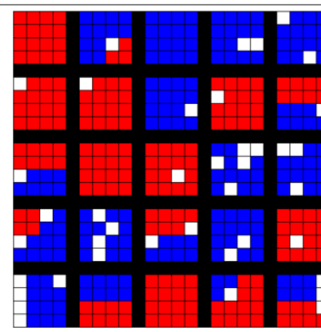
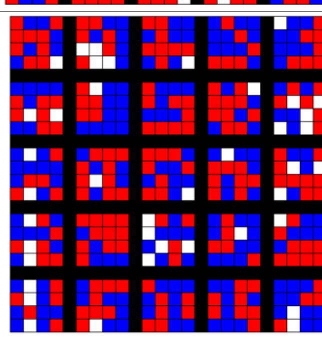
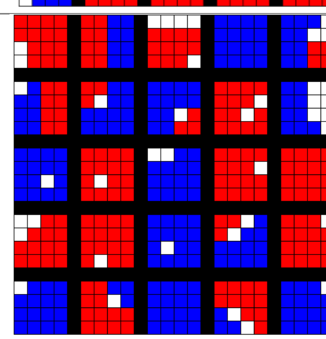
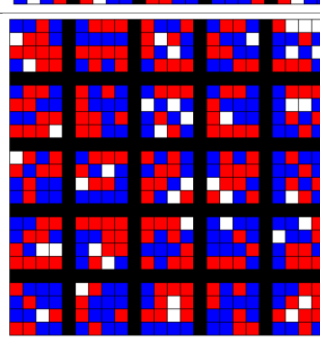
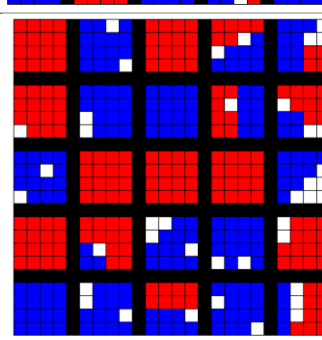
The last subdivision which will be explored is the creation of a grid divided into 25 equal blocks each with size 4x4. The number of available spaces remains at 200. The setup appears as follows:



This final simulation containing blocks most closely mirrors the densest parts of a planned, modern city that operates heavy infrastructure projects in an extremely organized way. Portland, Oregon, for example, has blocks of around 200ft in the downtown area. Such heavily gridded areas are essential for the smooth flow of pedestrians and vehicles.

Simulations

The simulations were run for five different trials, involving a randomized initial positioning. Each trial continues iterating until a sufficient percentage, above 99%, of agents within the grid are satisfied. The trials generated the following results:

	START	END
Trial 1		
Trial 2		
Trial 3		
Trial 4		
Trial 5		

Analysis

As with the other models previously, similar measurements can be made in the Twenty-Five Block Model. First, the number of neighborhoods:

Number of Neighborhoods		
	BLUE	RED
Trial 1	15	17
Trial 2	14	15
Trial 3	17	16
Trial 4	16	17
Trial 5	17	15
Average	15.8	16

The number of neighborhoods created for each population average around 16 over the course of five trials. This is approximately three to four times the average number of neighborhoods per population in the Four Block Model and slightly under two times the average number of neighborhoods per population in the Nine Block Model. Next, the percent of blocks occupied:

Percent of Blocks Occupied		
	BLUE	RED
Trial 1	60%	68%
Trial 2	56%	60%
Trial 3	68%	64%
Trial 4	64%	68%
Trial 5	68%	60%
Average	63%	64%

The percentage of blocks occupied has decreased dramatically, by nearly 30 percentage points, from the previous model. As the size of the block has decreased, the number of exclusive blocks has risen. A similar drop is seen in the percent of each population with dissimilar neighbors:

Percent of Population with Dissimilar Neighbors		
	BLUE	RED
Trial 1	13.9%	12.2%
Trial 2	7.8%	7.8%
Trial 3	13.9%	15.0%
Trial 4	14.4%	12.8%
Trial 5	11.1%	11.7%
Average	12.2%	11.9%

Unlike the difference between the Four Block Model and the Nine Block Model, the Twenty-Five Block Model has seen a dramatic decrease in the percentage of each population with dissimilar neighbors when compared to the Nine Block Model. With a decrease of around 25 percentage points, the number of agents that border an agent different from themselves has dropped significantly. Lastly, looking at the first trial, the starting and ending occupation percentage of each block by population:

Occupation of Block				
	Trial 1			
	BLUE		RED	
BLOCK	START	END	START	END
1	40.0%	0.0%	60.0%	100.0%
2	53.3%	0.0%	46.7%	100.0%
3	33.3%	0.0%	66.7%	100.0%
4	46.2%	46.7%	53.8%	53.3%
5	85.7%	100.0%	14.3%	0.0%
6	25.0%	0.0%	75.0%	100.0%
7	60.0%	100.0%	40.0%	0.0%
8	35.7%	0.0%	64.3%	100.0%
9	61.5%	100.0%	38.5%	0.0%
10	60.0%	100.0%	40.0%	0.0%
11	46.7%	0.0%	53.3%	100.0%
12	43.8%	80.0%	56.3%	20.0%
13	21.4%	0.0%	78.6%	100.0%
14	50.0%	21.4%	50.0%	78.6%
15	60.0%	61.5%	40.0%	38.5%
16	40.0%	100.0%	60.0%	0.0%
17	81.8%	100.0%	18.2%	0.0%
18	30.8%	0.0%	69.2%	100.0%
19	64.3%	76.9%	35.7%	23.1%
20	66.7%	66.7%	33.3%	33.3%
21	50.0%	0.0%	50.0%	100.0%
22	35.7%	0.0%	64.3%	100.0%
23	50.0%	53.3%	50.0%	46.7%
24	62.5%	100.0%	37.5%	0.0%
25	46.7%	100.0%	53.3%	0.0%

BLUE		END	
		< 50%	> 50%
START	< 50%	9	3
	> 50%	1	9

RED		END	
		< 50%	> 50%
START	< 50%	9	1
	> 50%	3	9

The concept of entrenched populations has become even more salient as the size of the blocks has decreased. From the data, even a slight majority in the beginning of the simulation will nearly guarantee either a majority or complete ownership of a block.

Based on the results above and with respect to the previous models, we can thus draw some temporary conclusions:

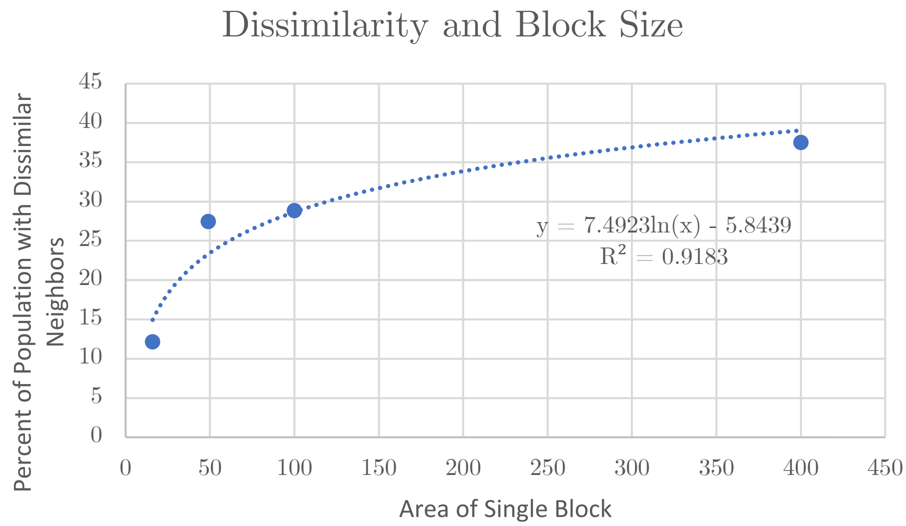
1. Entrenched populations continue to exist and the disparities in the final percentage breakdown of each block is becoming more extreme.
2. Unlike the previous models, complete segregation by block is fairly common, which is a dramatic change from the large block models. Around two thirds of the blocks can be exclusive to a single population.
3. The addition of eight major segments of infrastructure leads to a fairly large decrease in the percent of each population that has at least one dissimilar neighbor when compared to the addition of two or four pieces of infrastructure.
4. The addition of eight major segments of infrastructure leads to a near doubling of the number of neighborhoods created when compared to the Nine Block Model. The Schelling Model without infrastructure contains 84 neutral neighbors (border units) that can be abutted by an agent, the Four Blocks Model contains 176 border neighbors, the Nine Block Model contains 276 border neighbors, and the Twenty-Five Block Model contains 500 border neighbors. Again, the doubling of border units roughly coincides with the doubling of neighborhoods.

IX. General Conclusions

Analyzing each of the models in context with each other provides valuable insight into the relationship between certain variables when constructing the Schelling Model with infrastructure. Based on the conclusions drawn previously, we can note the following observations:

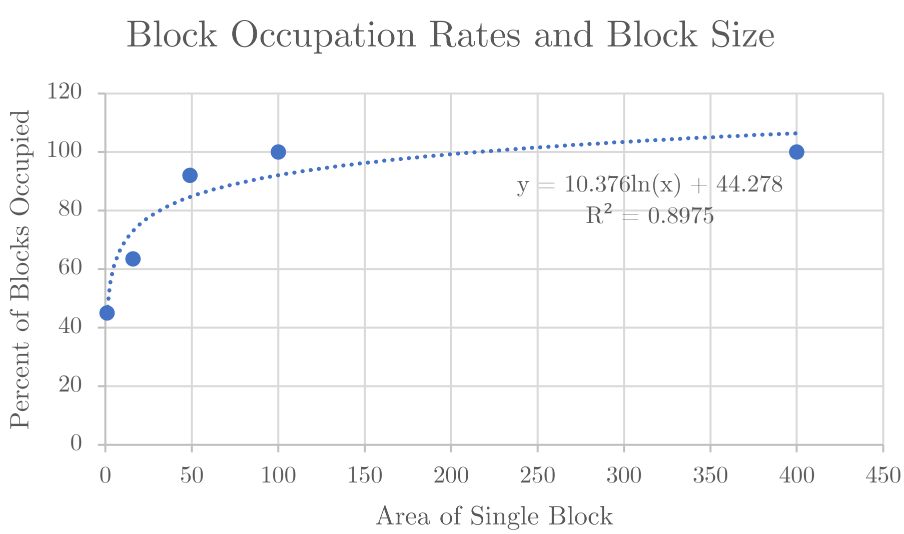
1. The smaller the grid size, the greater the impact on determining the shape of the neighborhoods. Adding infrastructure provides additional border surface area that can influence the shape of neighborhoods.

2. Grids decrease contact and increase segregation between populations. By providing natural boundaries, infrastructure can further enhance the segregating mechanisms present in the Schelling Model. The relationship between the percentage of a population's agents that have dissimilar neighbors and the area of a single block in a given grid setup reveals a logarithmic pattern, as shown below. From this relationship, we can deduce that increasing the area of a block has a much more dramatic impact when the block is already small, but the marginal gain when the blocks are already large is fairly negligible.

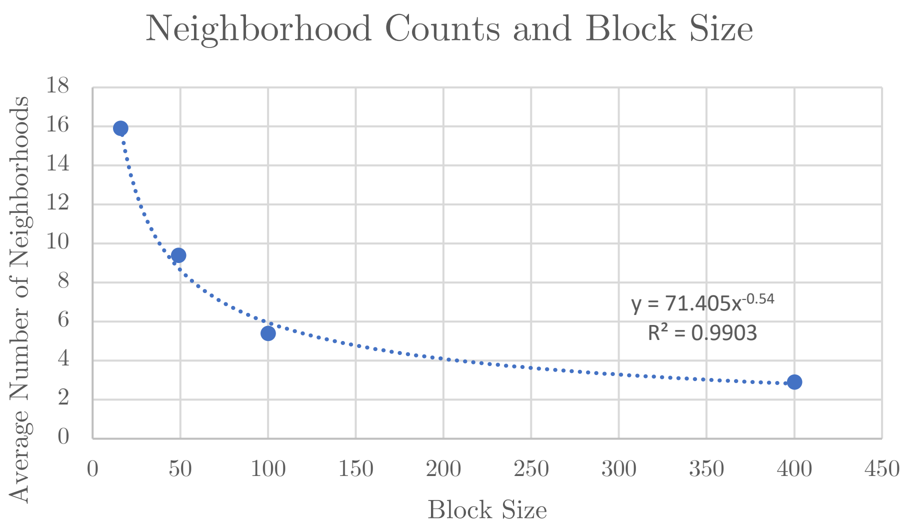


3. Entrenched populations become more important at smaller block sizes but are present at most levels. Smaller neighborhoods are far more conducive to exacerbating a block's majority population and population disparity. Starting location becomes crucial for determining ownership of a block for a population.

4. At a sufficiently small size, integration within a block becomes rare but still exists. Segregation under the Schelling Model does not necessarily mean that all of the blocks will become segregated. Accordingly, there appears to be another logarithmic relationship between the area of a single block in a grid and the percentage of blocks that a given population will have a presence in.



5. The relationship between the number of BORDER units and the number of neighborhoods is of a power nature, according to the data in the trials above.



Doubling the size of a block will halve the number of neighborhoods, which, due to properties of the model, will asymptotically approach 2.

6. Diverse blocks arise from the placement of empty spaces. Looking at the trials for the smaller block sizes show a correlation between abutting populations and the presence of an EMPTY space. Strategically placing an EMPTY space can entice certain populations to settle in an area where they may otherwise be a minority.

X. Further Research

While the trials above provide a reasonable start to a fairly complex question, there remain several important variations that were no explored that would be suitable for further research. The way in which the neighbor is defined, and whether or not that definition should be influenced by the creation of blocks, is a worthwhile modification to see the true effects of infrastructure. Rather than define neighbors as the agents directly abutting an agent, a utility function could be used that diminishes as the distance increases between agents, decreasing more rapidly across infrastructure. Secondly, the shape of the infrastructure and its construction on the grid can be explored further with greater variations that mimic other real-world scenarios, including parallel railroads or variable density.

The Schelling Model is an excellent example of the capabilities of agent-based computational methods in beginning to explain the rationale behind modern phenomena. The flexibility of the model allows us to alter and modify the model to more accurately imitate behavior in the real world. The addition of infrastructure into the Schelling Model provides us with insight into the relationship between the increasing divisiveness of infrastructure and the construction of self-segregating neighborhoods.