Schelling Model of Segregration

Petra Mengistu*

(Dated: May 12, 2023)

Segregation occurs widely in residential areas at notable levels and in seemingly erratic ways. Thomas Schelling, however, proposed a deterministic way in which areas populated by people of varying demographics affiliate themselves into unique locations that result in members of the same population grouping together. In this agent-based model, individuals are allowed to move to empty locations in cities until they reach a stable configuration where everyone is satisfied. We reproduce simulations of Schelling's original model for various compositions of cities that can host 10,000 inhabitants and explore a variation of this model with an addition of our own feature of swapping. With this feature, we find that the cities can reach stable configurations within a fewer number of rounds provided that there are enough vacant spaces in the system.

1. INTRODUCTION

One of the fundamental laws of nature is that like attracts like. Despite our increased cognitive capabilities of logic and reasoning, this law governs base human behavior to an extraordinary extent. One of the distinctly noticeable ways in which our affinity for similarity manifests is through the process of segregation. People of similar backgrounds, cultures/ethnicities, and identities are driven to associate with one another to the point of exclusion of those of a lower perceived similarity. Racial segregation can occur in a variety of ways, extending up to the gentrification of entire cities that form neighborhoods divided into areas populated by different races.

In 1969, Thomas Schelling developed a model that demonstrates the human behavior of discrimination for areas involving several populations [1]. The underlying principle of this model is that segregation occurs due to individual efforts to maximize personal satisfaction. Individual satisfaction is thus determined by perceived similarity of other neighboring people. In his original model, Schelling proposed a one dimensional depiction of a city organized along a line comprised of two racially distinct populations [1]. By evaluating the identity of each of their respective neighbors, individuals could either be satisfied or dissatisfied with their current location. Individual satisfaction is dependent upon the percentage of neighbors that are of the opposite population as the given individual: if this percentage is above a specific threshold capacity, which Schelling refers to as "tolerance," then the person is considered to be dissatisfied. In contrast, when the ratio of neighbors from a different population is less than an individual's tolerance, the individual is considered dissatisfied.

According to Schelling, individuals who are dissatisfied evacuate their current locations in search of other potential areas in which they may be satisfied. When an individual moves to a new location, Schelling identifies two changes in the satisfaction trends that can occur in the given city.

- The agent that moved may become satisfied or stay dissatisfied, depending on the neighbors surrounding the position to which they have moved.
- Other agents in the neighboring locations may become dissatisfied due to the incorporation of the displaced individual into their proximal neighbors.

The resulting arrangement of populations tends towards to a layout that Schelling terms "complete segregation." [1] In this phenomenon, the final composition of the populated grid exhibits "clusters" composed of people of the same population that clearly divide a city into neighborhoods of different populations – in other words, segregation. Since Schelling describes small residential cities, he thus predicts final states of has only two major clusters consisting of each aggregate populations [2]. However, Singh et al. [2] determine a quantitative measure of aggregation to show that global aggregation does not occur for larger systems; rather, systems develop multiple "clusters" that are composed of smaller portions of the total populations. To determine a quantifiable measure of aggregation, [2] measure the size of the clusters found within the final states of the system in equilibrium.

The final composition of the populated areas after the re-locations are considered equilibrium states of individual satisfaction, since motion within a city ceases when all agents are satisfied [1]. While all equilibrium states share the similarity that all agents are satisfied, the arrangements of populations in the city vary based on several factors, namely the initial configurations of agents in the residential area and the individual tolerance of neighboring diversity. Schelling identifies that a higher tolerance will generally result in a relatively more "mixed" equilibrium.

However, the impact of tolerance can become ambiguous when the population of one demographic group heavily outweighs the other. In this scenario, there are a greater number of clusters for the majority group; when

*Electronic address: pmengistu@haverford.edu

combined with the impact of individual tolerances, these clusters can grow to large sizes to reflect pronounced areas dominated by a majority with smaller pockets consisting of the minority population [2]. For such residential demographics, Schelling shows that the population which currently a minority is more likely to be displaced from their positions as individuals of those populations are more likely to be dissatisfied given the limited number of similar neighbors they can have in their vicinity [1]. Hence, segregation can occur to various extents depending on the initial composition of a given area, the tolerance of different populations, and the ratio of one population size to another.

Capturing the time evolution of populations, Schelling's model provides a simple, yet elegant depiction of human behavior. In terms of realistic applications, however, this model has several shortcomings that prevent it from portraying holistic residential conditions and the process of segregation. One of the questions that has been raised against the model is whether dissatisfied individuals would truly move to another position at random if they may end up in another location in which they are also dissatisfied. Extensive research has produced alternate implementations of Schelling's original model in an attempt to simulate feasible and comprehensive housing markets, Dall'Asta et al. [3] discuss the properties of two distinct classes of the Schelling model that have arisen due to differences in addressing the freedom given to the motion of agents across the grid: an unconstrained, or liquid, model that corresponds to Schelling's original theory and a constrained (or solid) model. In the following paper, we explore the properties of these two versions of the model and introduce our own variations of the Schelling model that encompass wider heterogeneity. As we initially replicate the unconstrained versions of the Schelling model and add onto the constrained version, we explain the distinctions between the two models and their associated notable properties in Sections 1.1 and 1.2 below.

1.1. Unconstrained (Liquid) Model

[3] describe the unconstrained model to mimic situations in which dissatisfied individuals are allowed to move to any location, regardless of whether the new location they move to increases their satisfaction. Hence, the unconstrained model is also referred to as "liquid" given that individuals are allowed to "flow" freely across empty locations in a grid until the system reaches a natural equilibrium.

As long as the city/grid has empty locations, individuals always have positions to which they can relocate; however, Dall'Asta et al. note that a property of Schelling model is that this continuous motion may never cease if there is no configuration of individuals on a grid for which all agents can be satisfied. One move might result in a dissatisfied agent becoming satisfied but simultaneously aggravates the satisfaction other neighboring individuals in their new location. In this case, these neighbors would decide to move to different locations, but those positions could result in the previously relocated agent becoming dissatisfied once more. Therefore, the liquid model can exhibit a tendency to have a fixed total number of dissatisfied agents, but with continuous internal motion of the agents within the grid such that the agents that are dissatisfied are shifting in a cyclical loop.

Dall'Asta identify that the amount of vacant slots influences the extent of motion of the liquid model. As the number of vacancies increases, the more likely it is that the system will converge to an equilibrium state where every individual is satisfied. However, for an area wherein the number of vacancies is below a certain threshold for the system, the agents continue to flow on the grid as the number of dissatisfied agents reaches a non-zero asymptotic value.

1.2. Constrained (Solid) Model

Another version of Schelling's model proposes a different mechanism by which agents are allowed to move to different locations. In the constrained model, a dissatisfied agent is only allowed to move to a location that has a higher satisfication than the position in which they currently reside; hence, each individual is "constrained" to make moves that increase their satisfaction. The resulting movements of the individuals occur in a more orderly, structured fashion, motivating the naming of this model as "solid" in contrast to the more erratic motion of a fluid.

However, in the constrained model, dissatisfied individuals can encounter the situation in which there are simply no empty locations that would result in an increase in their satisfaction. When the system reaches such a state, it is considered to be in "myopic/Nash equilibrium" as the dissatisfied agents are frozen in place with no alternative relocation options; hence, all motion in the system ceases and the configuration of positions remain solid with time. Just as in the liquid model, the number of vacant slots available can impact whether systems progress towards final states of zero-dissatisifaction equilibria or myopic equilibria. For an increasing number of vacancies, a given arrangement of individuals on a grid is more likely to reach zero-dissatisfaction states as individuals have more options to select as possible locations; conversely, Nash equilibria states are more likely to occur when the number of vacancies available are lower than a given threshold, when one population dominates another, or when the tolerance of individuals to neighboring diversity is low.

2. METHODS

To explore the properties of the Schelling model, we code a version of the original Schelling model that provides visual animations of the time evolution of the populations in a given residential area as well as a version of the constrained model that incorporates an additional feature through which agents can move. For all models, we consider a time step of the system as one of the rounds in which dissatisfied agents are allowed to move. In addition, we plot the number of dissatisfied agents as a function of the number of rounds for systems with different vacancies but similar thresholds.

2.1. Implementing Original Schelling Model

For the liquid model, we start developing the initial configuration of the populations on a grid that corresponds to an $n \times m$ matrix. We represent different populations by different numbers (population 1 and 2 by those numbers respectively) and empty slots by the number zero. Then, we assign members of one population to random positions on the grid and randomly assign members of the second population to the remaining available positions. The total number of agents from both populations should not exceed the housing capacity, which corresponds to the size of the matrix. Depending on the number of each population we select to place on the grid, the amount of vacancies available on the grid can be adjusted to reflect different initial conditions. Once the initial configurations of all agents has been set, we determine the "neighbors" for each occupied slot on the grid. A neighbor describes a cell adjacent to a given agent and can constitute an empty or occupied cell. For any grid larger than a 2×2 city, there are three numbers of possible neighbors that can be identified for a given agent based on their locations as listed below.

- If the agent is located on a corner of the grid, they have three neighbors
- If the agent is located on the top/bottom rows or the leftmost/rightmost column of the grid, they have five neighbors
- If the agent is located in a cell that can not be classified as one of these edges, they have eight neighbors

For each agent, we then determine a satisfaction factor σ based on the population they belong to and the composition of their neighbors. We define the satisfaction factor σ using the expression shown in Equation 1, where the number of similar people n_{sim} consists of the neighbors of a given agent that represent either an empty cell or a member of the same population.

$$\sigma = \frac{n_{sim}}{n_{neighbors}} \tag{1}$$

As one of the initial conditions of the system is the threshold level of similarity that individuals require to be satisfied, we compare the satisfaction factor of each agent to this threshold, t, using the procedure shown in [4]. If $\sigma < t$, the agent is marked as dissatisfied and will be identified as an agent that will relocate in that round. However, if $\sigma > t$, the agent is considered to be satisfied and can remain in their designated slot.

To move dissatisfied agents into different locations, we identify all available empty slots in the grid. If the number of dissatisfied agents is less than or equal to the number of vacancies, we elect to move all the dissatisfied agents; otherwise, we choose a random sample of the dissatisfied agents that is equal to the number of empty slots as individuals that can move in the given round. These agents are then moved to randomly selected vacant cells, after which the dissatisfaction of all individuals on the grid is re-evaluated and the process is repeated for another round of re-locations until all agents are satisfied.

However, as described by [3], there is the possibility that the agents may be continuously shifting around the grid in a loop, as the number of moves in a given round remain constant so that the dissatisfaction fraction of the system is fixed. To prevent the code from running in an infinite loop, we insert an agent based stopping criterion that evaluates the difference in the number of moves made each round between two consecutive rounds. If this difference is zero for a significant number of iterations of the code, we break the loop and have the configuration of the system in this state represent the "pseudo-final" state of the system. The choice in the number of iterations to use as a benchmark was determined through trial and error of observing whether the system would converge to a state of zero dissatisfaction after various lengths of iterations.

2.2. Developing Solid Model

To develop the solid version of Schelling's model, we follow the same steps as that of the liquid model to identify neighbors and agent dissatisfaction at each round. However, before moving an agent to any grid cell, we determine whether moving to that vacant spot would result in their satisfaction by evaluating the σ factor for the agent in their potential new location. If the satisfaction factor in the new position is greater than the threshold ($\sigma > t$), then we allow the agents to move to this location. However, if σ remains less than t, we impose the restriction that the dissatisfied agent must remain in their location for that round. We then iterate through each round until all agents are satisfied.

2.2.1. Additional Feature: Swapping

As the constrained model exhibits a tendency towards entering states of Nash equilibria, we attempt to account for a restriction of Schelling model by allowing agents to swap positions with each other as long as the swap would result in both agents being satisfied. In the original Schelling model and the variations of it that have followed, dissatisfied agents generally only move to vacant locations. In reality, however, people can generally be open to trade-offs in locations if it will contribute to the long-term stability of a community. Hence, a more realistic housing market can be simulated by the inclusion of agent swapping into the Schelling model. As swapping is only allowed if agents become satisfied, this feature can be considered to be an additional constraint in the solid model.

To implement this feature, at each time step, we redefine a round to include two stages: a moving stage (to vacant slots) and a swapping stage, where agents can trade locations. We first implement the constrained model of the moving stage described in Section **??**; once all possible moves have been made, we allow the remaining dissatisfied agents to explore the possibility of swapping. To do so, we set up a function that identifies all the occupied slots on the instantaneous composition of the grid on a given time step. For each occupied slot, we then loop over all the remaining dissatisfied agents and determine the satisfaction factor of each individual if they were to complete the swap. If the σ of both agents involved in a swap is greater than the threshold t, we assign the individuals to each other's locations. However, if the σ of either one or both of the agents involved in a swap remains or becomes lower than the threshold, then the agents are not allowed to leave their current locations, which is established by requiring the value of the grid cells at the end of the round reflect the numbers corresponding to the original population of each respective agent.

2.3. Quantifying Segregation

Using the animation function in the Python's matplotlib library, we show the time evolution of the system through visual representations of the grid composition at each round. In addition to this qualitative visualization of how segregation proceeds in the system, we quantify the extent of segregation through a measure of the heterogeneity of the system using the Gini coefficient that we discussed in lectures. For a given area governed by two different populations, the Gini coefficient is determined using Equation 2 below, where $f(i)$ refers to the fraction of individuals of a given population.

$$G = 1 - \sum_i f(i)^2 \tag{2}$$

For our examination of the Schelling model, we identify the two populations on a grid as the two functions such that the Gini coefficient is determined by the expression in Equation 3. A low Gini coefficient represents a relatively homogenous group consisting of individuals of the same population which would imply a significant level of segregation; in contrast, a high Gini coefficient describes a relatively heterogenous group composed of individuals of different populations, indicating less extreme levels of segregation.

$$G = 1 - \left(\frac{n_1}{N_{tot}}\right)^2 - \left(\frac{n_2}{N_{tot}}\right)^2 \tag{3}$$

A property of Schelling model-like systems is that "clusters" of individuals from the same population form over time, exhibiting the levels of segregation that occur when individuals are able to shift in location. As the number of rounds increases, the clusters increasingly reflect regions of homogenous populations, which would have a Gini coefficient of approximately zero within that given cluster. Hence, to compare the changes in the Gini coefficient for the composition of the grid as the system evolves through time, we divide the grid into an appropriate number of sections and calculate the Gini coefficient in each sector.

We then determine the average Gini coefficient of the system at that point in time by taking the average of the G values for each sector. While the choice of the number of sectors is somewhat arbitrary, we select a means of dividing the grid in a way that does not result in too few grid cells within a cluster or the inclusion of numerous miniature clusters within clusters. To observe the occurence (or lack thereof) of segregation in the system, we plot the average Gini coefficient over the total grid as a function of time.

3. RESULTS

3.1. Original Schelling Model

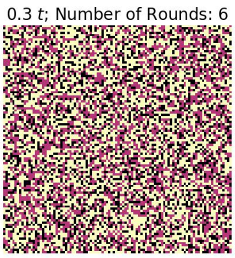
3.1.1. Visual Representation

For the original Schelling model, we are able to observe segregated systems evolve for a simulation of a city with 10,000 inhabitants. The initial conditions we use to observe these simulations are summarized in Table I. With these initial conditions, we observe the time evolution of the system for a random initial configuration of individuals on a 100x100 grid. We vary the similarity threshold t that individuals require to be surrounded by; the impact on the resulting final state of the system in which all agents are dissatisfied is shown in Figure 1. For a lower similarity threshold, the final state of the system shows a rather heteregeneous neighborhood whereas a large threshold of $t = 0.7$ evolves into distinctly segre-

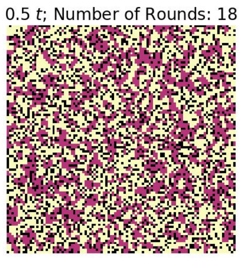
100x100 Liquid Model	
Population 1	30%
Population 2	45%
Vacancies	25%

TABLE I: Initial conditions used to observe animations.

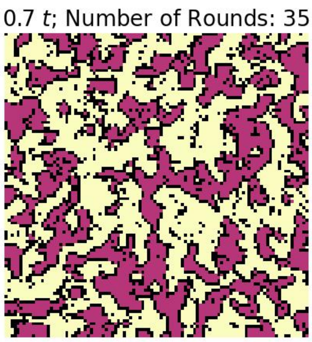
gated areas with large neighborhoods dominated by single populations.



(a) Heterogeneous system



(b) Mildly heterogeneous



(c) Segregated system

FIG. 1: Animations of the time evolution of cities with a capacity of 10,000 inhabitants. Panel (a) shows the simulation for a city in which inhabitants require that 30 % of their neighbors are of the same demographic population while panels (b) and (c) show the simulations for the same city but with individuals having 50 % and 70 % of similarity thresholds respectively.

When the similarity coefficient is sufficiently high, we observe the phenomenon described by [3] where the system

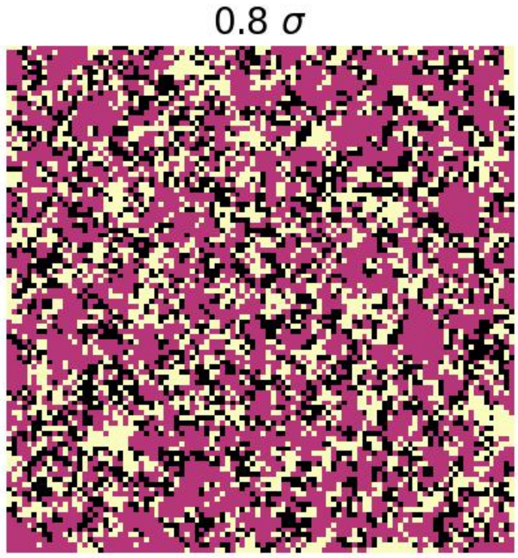
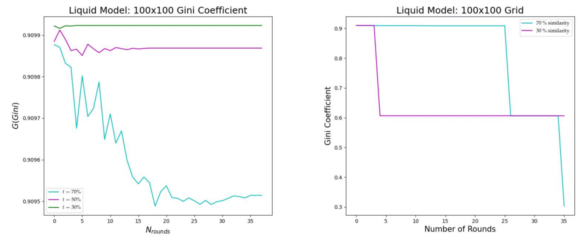


FIG. 2: A system that would be expected to reflect high levels of segregation but appears somewhat heterogeneous as it is in a pseudo-final state inserted by the stopping condition.

is unable to converge to a state of zero dissatisfaction; hence, the agents continuously move around the grid until the stopping condition is fulfilled. An example of the final state of such a system is shown in Figure 2, where the state of the system is still quite heterogeneous in spite of the high similarity threshold . This is because large portions of the overall population would continue to be in motion if the stopping condition had not been inserted to break the loop in the case of the system's inability to converge to a zero dissatisfaction state.

3.1.2. Gini Coefficient

While the qualitative images shown in Figure 1 provide a visualization of the segregation of neighborhoods, we confirm the occurrence and extent of segregation using a quantitative measure of the heterogeneity of the system through comparisons of the Gini coefficient of the system at various points in time. For the same initial conditions given in Table I, we plot the Gini coefficient of the system as a function of time for the various similarity thresholds of $t = 0.3, 0.5, 0.7$. The resulting plot in Figure 3 shows that systems with higher threshold for similarity undergo greater decays in the Gini coefficient over the range of time required for the system to stabilize. When the system reaches a satisfaction equilibrium, the Gini coefficient for systems with higher t are lower than those with a lower threshold, indicating that these systems have quantitatively segregated more over time. For easier visual comparison between the different trends, we plot the trendlines for the range of the longest number of rounds so that the Gini coefficient for systems that converge earlier flattens out after convergence.



(a) Continuous change over minimal range

(b) Discontinuous decreases

FIG. 3: Average Gini coefficient as a function of time for a 100x100 city. Panel (a) and (b) show that the Gini coefficient undergoes greater decrease for higher t.

However, a feature of the plots is that the changes in the Gini coefficients are either very steep and abrupt or occur over a very narrow range. For a smaller city of only 64 inhabitants, the changes in the Gini coefficient as a function of time occur over a larger range, as shown in Figure 4, which provides further information of how the Gini coefficient evolves with time. Hence, we conclude that the division of the grid into quadrants for the calculation of the Gini coefficient for larger grids (such as the 100x100 sized city) covers areas that are too broad to provide a truly informative measure of heterogeneity, as the Gini coefficients of different clusters average out to a uniform value.

In the case where the system does not converge to a zero dissatisfaction state for all agents, the Gini coefficient also provides a useful measure of the behavior of the system. As dissatisfied agents are still moving in the liquid model, the number of dissatisfied agents remain constant, but the areas in which dissatisfaction occurs are continuously shifting cyclically with the motion of the agents. Therefore, the heterogeneity of the system changes correspondingly, as shown in Figure 5 that shows the oscillation of the value of the Gini coefficient about a fixed value. We include the results for smaller city as the use of quadrants is more informative for a more representative depiction of the system's relation to

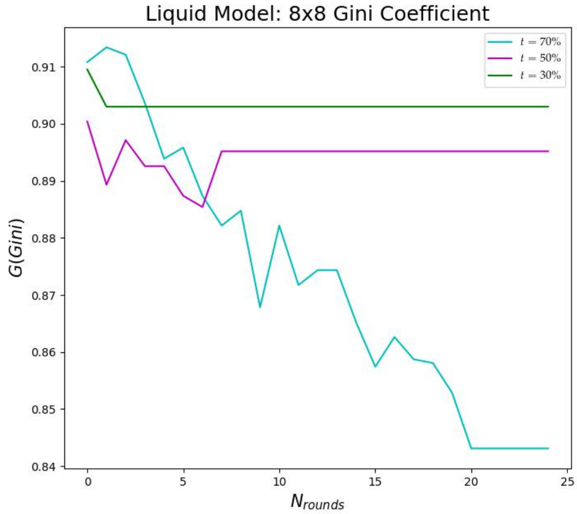


FIG. 4: Average Gini coefficient as a function of time for an 8x8 grid. Similar trend observed that for increasing t, G decreases.

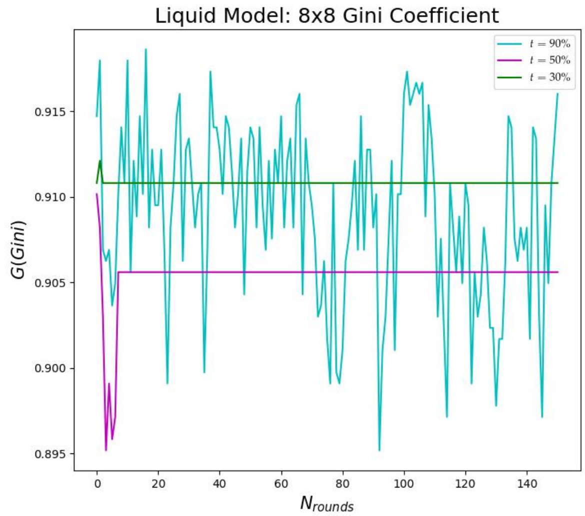


FIG. 5: Oscillation of average Gini Coefficient about a fixed value when system is in continuous loop of motion.

the average Gini coefficient.

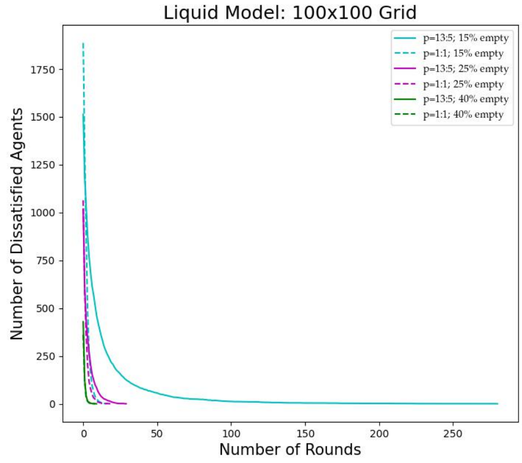
3.1.3. Dependency on Vacant Slots

For a city of the same number of inhabitants, we plot the number of dissatisfied agents as a function of the number of rounds required for the model to converge to a state at which each individual is satisfied or the number of dissatisfied agents is fixed. We compare this convergence rate for different number of vacancies available and

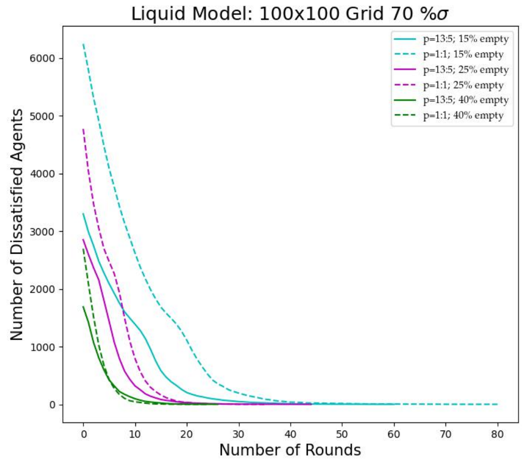
100x100 Liquid Model				
	Population 1	Population 2	Ratio	Vacancies
City 1	6500 (76%)	2500 (24%)	13:5	15%
City 2	4250 (50%)	4250 (50%)	1:1	15%
City 3	5550 (76%)	1950 (24%)	13:5	25%
City 4	3750 (50%)	3750 (50%)	1:1	25%
City 5	4440 (76%)	1560 (76%)	13:5	40%
City 6	3000 (50%)	3000 (50%)	1:1	40%

TABLE II: Initial conditions for a 10,000 inhabitant capacity city with different numbers of empty slots and population demographics.

for a different ratio of populations to each other. The full set of initial conditions are shown in Table II. At a similarity threshold of $t = 0.5$, we observe that for a lower percentage of available vacancies, the system requires a larger number of rounds to converge than the a city with more vacancies, as shown in Figure 6. For a similarity threshold of $t = 0.7$, this same trend is observed for all six cities. Interestingly, however, the ratio of the population sizes also impacts the rate of convergence: for each pairs of cities with an equal number of inhabitants shown in Figure 6, when there are equal numbers of individuals from each population (1:1 ratio), a greater number of rounds are required for the city to reach a steady state. We propose that this is due to the property that a greater number of individuals are initially dissatisfied for such demographic compositions, as no single population is in a clear majority.



(a) More vacancies, faster convergence



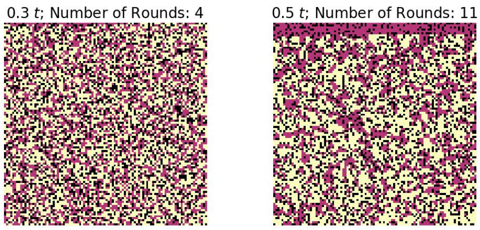
(b) Greater difference in population sizes has shallower decay.

FIG. 6: Number of dissatisfied agents as a function of time.

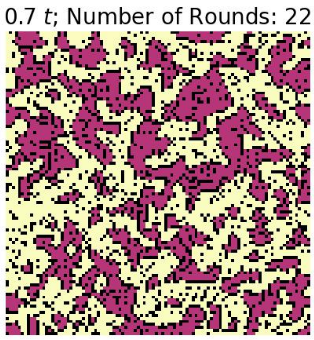
3.2. Revised Housing Market: Swapping

3.2.1. Visual Representation

For a city of 10,000 inhabitants, the solid model with swapping takes a long time to converge. This is because the code that implements swapping is a nested loop that runs through every dissatisfied agent for every occupied slot, giving rise to a computation that thus scales as n^2. For the first several iterations of the system, the number of dissatisfied agents is high, particularly for a high threshold of similarity. Hence, the code has to check significantly more cases than in later iterations where the number of dissatisfied agents has reasonably decreased. For high similarity thresholds, the computation is particularly slowed down as these systems require more rounds to converge, which indicates that the level of dissatisfaction remains at higher values for a longer periods of time. Consequently, the code would take roughly 12 hours to converge for the plots of dissatisfaction as a function of vacancy as the average of several runs of the codes needs to be taken for each plot. However, we are able to observe the animations of the time evolution for a city of magnitude 100x100 with the selection of 25 % vacancies to reduce the number of occupied slots to loop through. The results of these simulations for various thresholds of similarity as we observed for the liquid model are displayed in Figure 7. Similar segregation effects are observed for the constrained model with swapping; however, we note that the number of rounds required to achieve a steady state is less with the addition of the swapping feature in the housing marked than the original Schelling model. Hence, the inclusion of swapping has the benefit of incorporating the aspect of the constrained model that only allows moves that increase satisfaction to occur and completes the process within a fewer rounds of re-locations.



(a) Heterogeneous (b) Some homogeneity



(c) Segregated system

FIG. 7: Animations of a 100x100 grid including swapping.

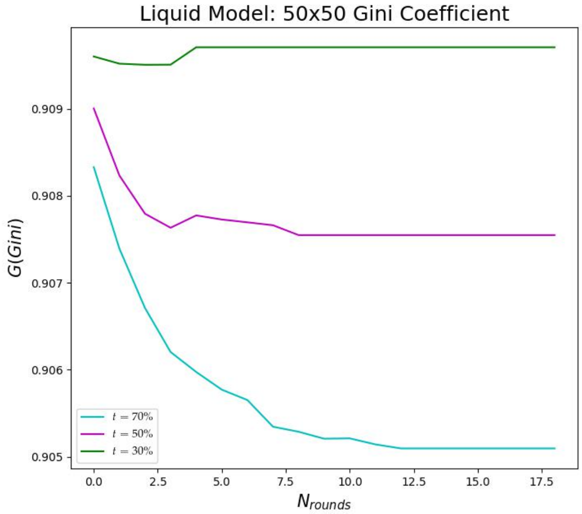


FIG. 8: Reduction in Gini Coefficient for Schelling model inclusinve of Swapping

50x50 Solid Model				
	Population 1	Population 2	Ratio	Vacancies
City 1	553 (76%)	1572 (24%)	13:5	15%
City 2	1063 (50%)	1062 (50%)	1:1	15%
City 3	488 (76%)	1387 (24%)	13:5	25%
City 4	938 (50%)	937 (50%)	1:1	25%
City 5	390 (76%)	1110 (76%)	13:5	40%
City 6	750 (50%)	750 (50%)	1:1	40%

TABLE III: Initial conditions for revised Schelling Model of 50x50 grid.

3.2.2. Gini Coefficient

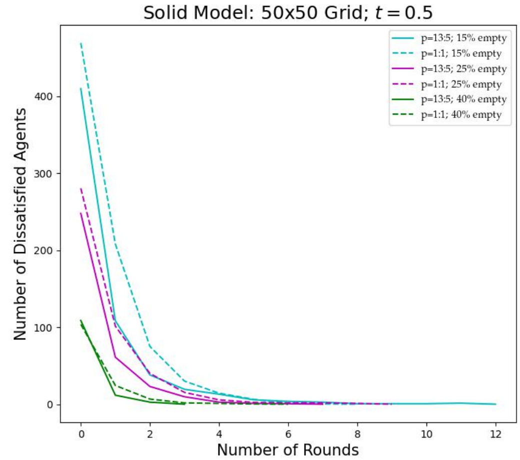
We obtain a quantitative measure of the segregation shown in the animations using the change in the average Gini coefficient. As shown in Figure 8, the Gini coefficient decreases the most for the system with highest similarity threshold, which indicates a higher level of segregation in the final state of the system. However, we note that the same limitation of using a four quadrant division of the grid to evaluate the Gini coefficient is observed in the very minimal overall change of the Gini coefficient over the course of all the rounds.

3.2.3. Dependence on Vacancies

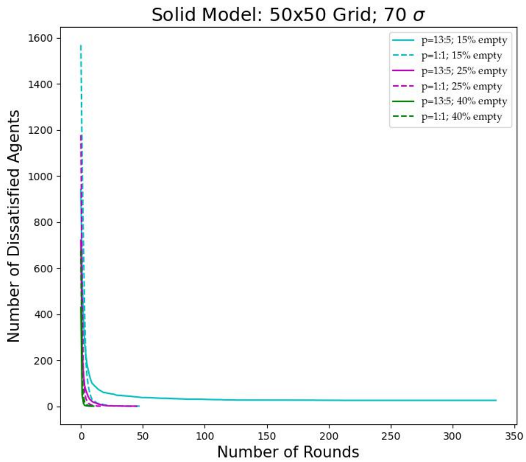
With the inclusion of swapping, we explore whether the number of available empty slots impacts the rate of convergence of the system to a state where the number of moves that agents in the system undergo remains constant. Figure 9 shows the dependency of the convergence rate of dissatisfaction in the system for the sets of initial conditions describing different cities given in Table III.

We note that the trend observed in the liquid model is also observed for the constrained swapping model, as cities with less vacant spots require longer times to converge for both the thresholds of 50 % and 70 % similarity, although the higher threshold takes significantly longer to converge compared to the lower threshold.

Another notable feature of the system with higher similarity thresholds is the decay of the fraction of dissatisfied agents with time to a non-zero asymptote. As the system is constrained, this system exhibits the "myopic" equilibrium condition more readily than that of the unconstrained model, which is able to converge to a state with no dissatisfied agents for grid four times the size of the constrained model. This "freezing" of the system shows a dependency on the amount of vacancies available in the system as only one of the cities with the lowest percentage of available empty locations has a non-zero number of dissatisfied agents.



(a) Higher threshold and low vacancies takes longest to converge.

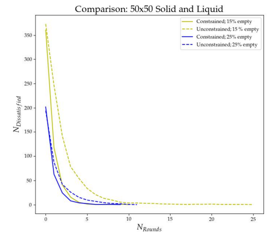


(b) Trend for vacancy dependence and population ratio on convergence is observed.

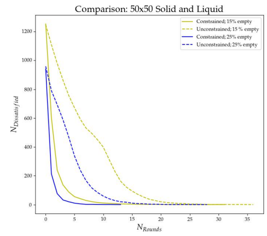
FIG. 9: Number of dissatisfied agents as a function of the rounds to reach convergence.

To make quantitative comparisons between the original Schelling model and the incorporation of the swapping feature, we plot the number of dissatisfied agents as a function of the number of rounds for both systems in Figure 12. We observe that for both a $t = 0.5$ and $t = 0.7$ similarity threshold, the constrained Schelling model with swaps undergoes a more pronounced exponential decay and thus requires fewer rounds to have the number of dissatisfied agents asymptotically reach zero, just as the number of rounds required in the generation of the animations indicate. Thus, the inclusion of swapping, although computationally more expensive, does have overall benefits for the construction of more streamlined and efficient housing markets.

However, due to the constraint of satisfaction increasing



(a) $t = 0.5$



(b) $t = 0.7$

FIG. 10: The fraction of dissatisfied agents at each round for both the constrained model, shown using solid lines, and the unconstrained model using dashed lines. The time evolution for two cities with different numbers of vacant slots available with each model are shown in yellow and blue.

motion, the swapping solid model does have the limitation of "freezing" for grid sizes that are much smaller than the liquid model. Figure 11 shows a case where the inability of the constrained model to reach the a state where all agents are satisfied is emphasized in contrast to the liquid model. The same given city with a 15% fraction of available empty slots converges to a state where all agents are satisfied for the liquid model but can only reach a minimum of 50 dissatisfied agents for the constrained model. In the case where the constrained model can converge, which is given by the higher fraction of empty slots, the trend that the constrained model converges in a fewer number of rounds is still observed.

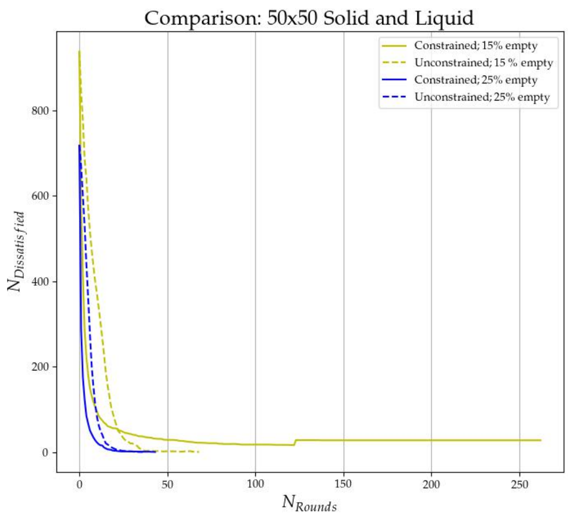
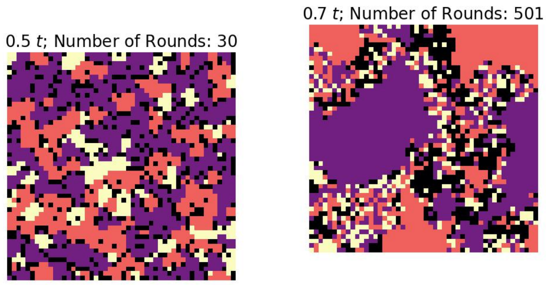


FIG. 11: Caption

3.3. Additional Expansions of Schelling Model

3.3.1. *Application to Three Populations*

We expand the Schelling model to apply to systems with three populations, and run simulations for cities of a 50x50 grid assuming each individual has similarity thresholds of 50 % and 70 % as regarding members of the opposite population. Segregation is observed in these systems, although we note that the system with the 0.7 % threshold has a rather heterogeneous final state. This is due to its inability to converge to a steady state of zero dissatisfaction; hence, Figure **??** shows the pseudo-final state imposed by the stopping condition. As there are three different populations whose satisfaction needs need to be met, the liquid model cannot reach an equilibrium state where every agent is satisfied for the 50x50 grid whereas the original model with two populations could reach such a state for a 100x100 grid in far fewer rounds. This indicates that greater diversity of populations limits the ability to achieve a steady state system.



(a) Somewhat segregated

(b) Pseudo-final state (stopped)

FIG. 12: Application of liquid Schelling model to encompass three populations.

4. CONCLUSION

Therefore, we observe that the Schelling model is a viable way of simulating residential segregation. This model can simulate large scale cities reasonably well and replicate observed phenomena of gentrication through the formation of very homogeneous neighborhoods. We confirm the occurence of segregation in Schelling model systems through visual animations and quantitative measures of the Gini coefficient. Ultimately, we conclude that the dependence of the average Gini coefficient on the division of grids is not an ideal way to measure heterogeneity as some information is lost in these divisions. For larger grids in particular, smaller grid structures are likely necessary to understand the true behavior of the system. However, we do note that the Gini coefficient decreases, albeit minimally, for higher similarity thresholds in both the cases of the original Schelling model and revised model with swapping. This provides some quantitative support to the observed visual segregation.
In addition, we compare the original Schelling model to the revised model we develop and find that the inclusion of swapping allows the system to reach a steady state of zero dissatisfaction in a fewer number of rounds. The tradeoffs are that the revised model with swapping is more computationally expensive, and thus each round initially lasts longer for larger grid sizes and the tendency of the system to enter states of myopic equilibria where the dissatisfied agents are stuck in place due to a lack of a relocation or swap that would improve their satisfaction. Nevertheless, the practicality of swapping could be a useful feature to be explored, and future work could integrate the benefits of both models to achieve a comprehensive approach that is applicable to a wide range of grids. comp dist and swap to just dist

[1] T. C. Schelling, The American Economic Review **59**, 488 (1969), ISSN 00028282, URL http://www.jstor.org/stable/1823701.
[2] A. Singh, D. Vainchtein, and H. Weiss, Demographic Research **21**, 341 (2009), ISSN 14359871, 23637064, URL http://www.jstor.org/stable/26349348.
[3] L. Dall'Asta, C. Castellano, and M. Marsili, Journal of Statistical Mechanics: Theory and Experiment **2008** (2008).
[4] F. McCown, *Schelling's model of segregation*, URL http://nifty.stanford.edu/2014/mccown-schelling-model-segregation/.