

EFFECT OF INDIVIDUAL DESIRES ON THE
SEGREGATION OF NETWORK SOCIETY

BURCU GÜNEŞ

Submitted to the Graduate School of Engineering and Natural Sciences
in partial fulfilment of the requirements for the degree of
Master of Science

Sabancı University

December, 2015

© Burcu Güneş 2015

All Rights Reserved

EFFECT OF INDIVIDUAL DESIRES ON THE SEGREGATION OF NETWORK SOCIETY

Burcu Güneş

Master of Science in Industrial Engineering

Thesis Supervisor: Ali Rana Atılgan

Abstract

In this thesis, effect of individual desires on the segregation of society is investigated and new methods are defined to distinguish the outcomes. According to Schelling's segregation model, agents have tolerance and when the ratio of dissimilar neighbors exceeds the tolerance level, agents change their place. However, some sociological studies show that people can have an incentive to come together with dissimilar ones. Thus, a desire for diversity can be defined in simulations. According to simulation results, even if the desire for diversity is low, existence of this desire decreases the segregation level of society. We then defined different utility functions which include tolerance and desire for diversity and saw under which utility functions agents segregate, integrate or form an inverse homophilic network. Moreover, we observed that there can be networks with the same traditional segregation indices, but different patterns. We distinguished these patterns with a new measure named neighborhood distribution function. Then the motifs that the agents form with their neighbors are investigated, and we showed that different motifs lie behind the differences in network patterns. Motifs with different clustering coefficients might ile policy makers offer regulations which can change individuals' utility functions, it should not be ignored that these changes might affect the segregation level of society as well as the neighborhood motifs which might give rise to new social dynamics.

Keywords: Segregation, Homophily, Neighborhood Distribution Function, Network Motifs

BİREYSEL İSTEKLERİN TOPLUMUN AYRIŞMASI ÜZERİNDEKİ ETKİSİ

Burcu Güneş

Endüstri Mühendisliği Yüksek Lisansı

Tez Danışmanı: Ali Rana Atılgan

Özet

Bu çalışmada bireysel isteklerin toplumun ayrışmasına olan etkileri incelenmiş ve elde edilen sonuçların ayırt edilebilmesi için yeni yöntemler geliştirilmiştir. Schelling'in ayrışma modeline göre, bireylerin belirli bir toleransı vardır ve kendilerine benzemeyenlerin oranı toleransın izin verdiği miktarı geçince bireyler yer değiştirmektedir. Halbuki bazı sosyoloji çalışmaları gösteriyor ki, bireylerin kendine benzemeyenlerle bir araya gelmek için de sebepleri olabilir. O nedenle simülasyonlara toleransın yanısıra ötekiyle bir arada olma isteği de eklenebilir. Simülasyon sonuçlarımıza göre, bu istek düşük seviyede bile olsa böyle isteğinin oluşu ayrışma seviyesini düşürüyor. Daha sonra toleransı ve ötekiyle olma isteğini içeren farklı fayda fonksiyonları tanımlayarak, bireylerin hangi durumda entegre olduğu, ayrıştığı ya da ters homofilik bir düzen oluşturduklarını inceledik. Bu arada geleneksel ayrışma indekslerine göre aynı sonuçları verdiği halde farklı desenler veren ağlar olduğunu gördük. Bunların komşuluk dağılım fonksiyonu adını verdiğimiz bir yöntemle ayırt edilebilmesini sağladıktan sonra altta yatan motifleri inceledik ve bireylerin komşularıyla oluşturdukları motiflerin ağın desenini belirlediğini gördük. Kümelenme katsayısı farklı olan motifler, yayılım ve ortaklaşa davranış gibi durumlarında farklı sonuçlara neden olacaktır. Bu nedenle politika oluşturucular bireylerin fayda fonksiyonlarını değiştirecek düzenlemeler yaparken, toplumun ayrışma seviyesini etkilemenin yanısıra bireylerin komşularıyla oluşturacakları motifleri etkileyeceklerini ve bunun da farklı sosyal dinamikler doğurabileceğini gözardı etmemelidirler.

Anahtar Kelimeler: Ayrışma, Homofili, Komşuluk Dağılım Fonksiyonu, Ağ Motifleri

To my dear mother, and segregated country

Acknowledgements

I am grateful to my thesis supervisor Prof. Dr. Ali Rana Atılgan for his guidance, encouragement and patience during the research and writing of this thesis.

I would like to thank to Assoc. Prof. Dr. Güvenç Şahin and Asst. Prof. Dr. Ahmet Onur Durahim for being part of my thesis jury.

I am also thankful to my friends from IE graduate office, they were always supportive and I had great time with them during my masters. I am happy to know all of them.

I am lucky to have a mother who always supports my ideas and decisions and lucky to have my beloved Ahmet whose simile is enough to make my day better. I am grateful to them for everything.

Contents

1	Introduction	1
2	Segregation Under Different Utility Functions	4
2.1	The Effect of an Upper Bound on Segregation	7
2.2	Symmetric Utility Functions	9
2.3	Asymmetric Utility Functions	13
2.4	Segregation When There is a Minority Group in Network	17
2.5	The Effect of the Randomized Utility Functions	19
3	Segregation in Different Networks	21
3.1	Segregation in Ring Lattice and Small World Networks	24
4	Neighborhood Distribution Function	28
4.1	Comparing Networks with the Neighborhood Distribution Function	30
4.2	Neighborhood Motifs	33
5	Conclusions and Future Research	37
6	Appendices	43
6.1	Appendix A: The Effect of Percentage of Vacant Cells	43
6.2	Appendix B: Results of Different Movement Algorithms	44
6.3	Appendix C: Results of Symmetric Utility Functions for Regular Lattice with Degree 12	45
6.4	Appendix D: Comparison of the Ring Lattice and Small World Networks for Different Average Degrees	46

List of Figures

1.1	Simulation of Schelling's segregation model with different lower bound values	2
2.1	Different utility functions from the literature	6
2.2	The effect of an upper bound on segregation	8
2.3	Symmetric utility functions	11
2.4	Mixity and FSI values of different symmetric utility functions	12
2.5	Asymmetric utility functions	13
2.6	Simulation outcomes of networks with asymmetric utilities	15
2.7	Mixity and FSI values of different asymmetric utility functions	16
2.8	Simulation outcomes of networks which have a minority group	18
2.9	The resulting mixity values when some fraction of the agents have a higher tolerance and desire for diversity.	19
3.1	Change in the clustering coefficient and average shortest path length with increasing rewiring probability	22
3.2	Regular, random and small world networks	23
3.3	Degree distributions of the ring lattice and small world network	23
3.4	Segregation results of the ring lattice and small world network when there is just a lower bound	25
3.5	Segregation results of the ring lattice and small world network for different symmetric utility functions	26
3.6	Segregation results of the ring lattice and small world network for different asymmetric utility functions	27
4.1	Two networks with different utility functions, but same segregation indices. .	28
4.2	An illustration to explain how neighborhood distribution function works. . . .	29
4.3	The neighborhood distributions functions of black cells of two networks with the same segregation indices.	30
4.4	Two networks with different symmetric utility functions, but same segregation indices.	31

4.5	The neighborhood distribution functions of black cells of two regular networks with symmetric utility functions.	32
4.6	The neighborhood distribution functions of the red agents of the ring lattice and small world networks for 75% lower bound.	33
4.7	Nodes with different motifs	34
5.1	Distribution of moving cost for different settings	39
6.1	Results of different movement algorithms	44
6.2	Results of symmetric utility functions for regular lattice with degree 12	45
6.3	Segregation results of the ring lattice and small world network with degree 4 and 12 for different lower bounds.	46

List of Tables

2.1	Change in segregation measures when there is an upper bound.	9
2.2	Effect of an upper bound for different initial conditions of networks	9
2.3	Resulting mixity and FSI values and ratio of the unhappy agents when there is minority group in network.	17
2.4	The resulting segregation indices when there is a minority group and random- ized utility functions.	20
4.1	The distribution of the motifs in two networks with the same segregation indices	34
4.2	The clustering coefficients of different networks	35
6.1	Number of time steps until equilibrium for different percentages of vacant cells and different lower bounds and lower-upper bound couples.	43

1 Introduction

As Kleinberg and Easley explain in their book, people are inclined to be friends with people like themselves. Friends usually share common attributes such as age, gender, and race and characteristics such as interest, income, and occupation. This phenomenon is known as *homophily* [1].

Studies on homophily indicate that divisions in race, gender, ethnicity, age or education level may cause homophily. For instance, Moody's study with the national sample of middle and high school students shows the existence of homophily at racial and grade level [2]. Another national study demonstrates that only 8% of the respondents discuss important matters with people from other race or ethnicities [3]. Shrum et al.'s questionnaire which was conducted with more than 2000 students from 3rd to 12th grade indicates racial and gender homophily. As grade level increases, gender homophily decreases, whereas racial homophily increases [4]. Another study which includes 20 high schools in Netherlands is done by Baerveldt et al., and results show that most of the support relationships are intra-ethnic for Dutch and Turkish students [5]. Suitor and Keeton investigate the data of the women who returned to school in their midlife and see that educational attainment is an important factor to form support network for study or work [6].

Homophily generates a segregated spatial pattern as people prefer to move near others similar to themselves [1]. Segregation may not be anyone's intention, but it is the aggregate result [7]. Schelling explains this on a square lattice; cells are either occupied by one of the two types of agents or vacant, and each agent's neighborhood consists of surrounding eight cells which is known as Moore Neighborhood in the literature. Each agent has a tolerance which

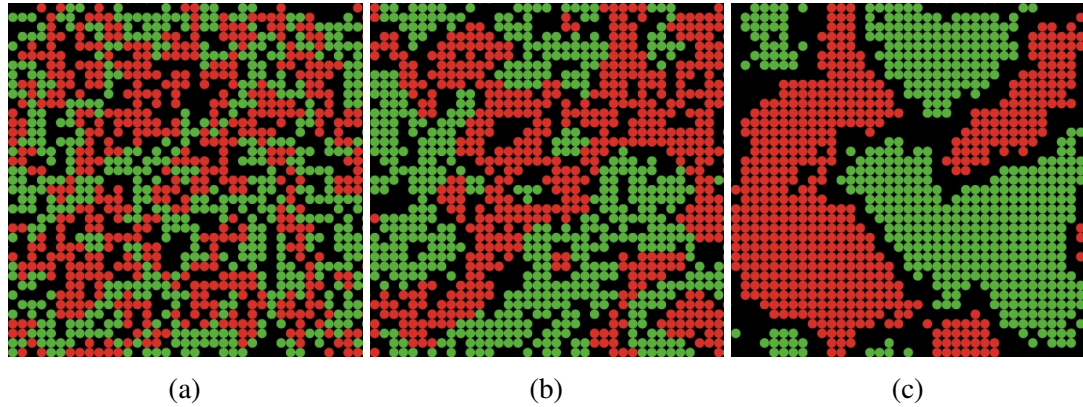


Figure 1.1: Simulation of Schelling’s segregation model with different lower bound values. Lower bound is 25% in (a), 50% and in (b) and 75% in (c). Increasing lower bound causes more segregated environment.

is represented with a lower bound. An agent is happy if the ratio of agent’s neighbors who are similar to agent is greater than or equal to the lower bound. Unhappy agents change their places by moving to vacant cells, this continues until each agents is happy with its place (which may never happen depending on the lower bound value) [8]. Schelling’s main concern was segregation by color in US, but his analysis is useful for segregations caused by other attributes than color [8].

Schelling’s segregation model is repeated with the agent-based simulation program NetLogo [9]. There is an additional rule at our model; if an agent’s neighborhood consist of vacant cells, this agent is unhappy. As nobody lives alone in society, with that rule agents are forced to have at least one neighbor. Pancs and Vriend make the same assumption in their study as well [10].

Figure 1.1 demonstrates the outcomes of 3 different settings. In each of the settings there are 37x37 cells and 1000 randomly distributed agents, half of whom are green and half of whom are red. Being red or green indicates being different with regard to one attribute. 369 black cells are the vacant cells. There is periodic boundary condition, so each agent, including the ones on the edges, can have eight neighbors. The difference between the settings is the lower bound. In setting (a) the lower bound is 25% while in (b) it is 50% and in (c) it is 75%.

As expected, higher lower bound causes more segregated environment. Although 25% of similar neighbors are enough for agents to be happy in setting (a), replicating the simulation

for 100 times gives that on average 61.8% of the neighborhood is between similar agents. This ratio is 88.5% and 99.8% for settings (b) and (c) respectively. However, in perfectly integrated network, one would expect on average 50% of the neighborhood to be between similar agents and other 50% to be between different agents.

Computerized data of Chicago residents at years 1940, 1950 and 1960 show that city becomes ethnically more segregated over the years [11]. As Schelling mentions, segregation is like a chain reaction. When an unhappy agent moves to another place, its similar neighbors lose one neighbor that is similar to themselves and maybe become unhappy and need to move, so at the end segregated pattern is obtained [8].

In the following chapter, we will introduce an upper bound to simulation to see the effect of it on segregation. Having a lower and upper bound couple make a step utility function, but we will then change this step utility function with different symmetric and asymmetric utility functions and test them. In the third chapter, we will compare the ring lattice and small world network to see if they segregate differently. Following this chapter, a tool named neighborhood distribution function will be defined. By using this tool and checking the network motifs, we will be able to catch network patterns, and in the last chapter we will summarise our findings.

2 Segregation Under Different Utility Functions

There can be many indices to measure and compare segregation in networks, and different indices may lead to different conclusions. One of the measures that is used in this study is *mixity*. If we call neighborhoods between similar agents homogeneous neighborhoods and neighborhoods between dissimilar agents heterogeneous, mixity is obtained by dividing the number of heterogeneous neighborhoods into the sum of homogeneous and heterogeneous neighborhoods, so it is very easy to understand and use. This indice can be also named as dissimilarity index. Schelling divides the number of homogeneous neighborhoods into the sum of homogeneous and heterogeneous neighborhoods, which is basically $1 - \text{mixity}$ [7]. Mixity is also used by Pancs and Vriend and Banos [10] [12]. Clearly, mixity takes values between 0 and 1, and under a strong segregation, one might expect mixity to be close to 0.

On the other hand, if different type of agents are not distributed evenly, in other words, if there is a minority group in network, mixity might be misleading. In that case, low values of mixity may not be because of segregation, but because of having less chance to observe neighborhoods between dissimilar agents. Thus, there should be a second measure. Lazarsfeld and Merton say that for homophily, a proportion of expected and observed friendships should be obtained and compared[13]. In the literature this measure is known as *Freeman's segregation index*, which is defined by Linton Freeman [14] and widely used in the literature [2] [4] [15]. Freeman's segregation index will be shortened as *FSI* along the study.

At any given network, if the fraction of first type of the agents is p and the fraction of second

type of agents is q , then the expected number of heterogeneous neighborhoods would be: $2pq$ * total number of neighborhoods. If we show the number of expected heterogeneous neighborhoods with $E[N]$ and the number of observed heterogeneous neighborhoods with N , FSI is found as:

$$FSI = \frac{E[N] - N}{E[N]} \quad (1)$$

Freeman considers the situations in which $E[N]$ is greater than N , so FSI is between 0 and 1 [14]. In our study, $E[N]$ can be over or under N , which means FSI takes values between -1 and 1. At a non-segregated environment, the number of observed heterogeneous neighborhoods would be around the number of expected heterogeneous neighborhoods, so FSI is expected to be around 0. If there are just a few heterogeneous neighborhoods, FSI would be around 1, which indicates segregation. If most of the neighborhoods are heterogeneous, FSI would be closer to -1^1 , it can be said that these kind of networks show *inverse homophily* which is the case especially at romance and sexual relationship networks [1].

In Schelling's model, utility function is a step function like the function in figure 2.1a; lower bound is given as 50%, so agents are not happy, which means their utility is 0, when they have lower than 50% similar neighbors. Unhappy agents change their place and starting from 50% similar neighbors agents are happy and their utility is 1. Zhang uses the utility function in figure 2.1b, so each agent prefers integrated (50%similar 50% other neighbors) neighborhood most. In his model, agents change their place if there is a place with a higher utility for them and Zhang concludes that in spite of integration is wanted, segregation emerges [16]. Pancs and Vriend also emphasise that in Schelling's model no one prefers segregation, on the other hand no one is against it. Thus, they introduce different utility functions like figure 2.1c, figure 2.1d and spiked utility function all of which give the highest utility to 50% similar neighbors and they find that there is still segregation. [10].

In Zhang's and Pancs and Vriend's utility functions, integration leads to the highest utility, however functions in figures 2.1b and 2.1c are asymmetric; in case of not having 50% similar - 50% other neighbors, similar neighbors are preferred. For instance, 70% similar - 30% other

¹When two types of the agents are not distributed equally, and most of the neighborhoods are heterogeneous, FSI can be lower than -1, and the lower FSI values indicate the more inverse homophilic networks.

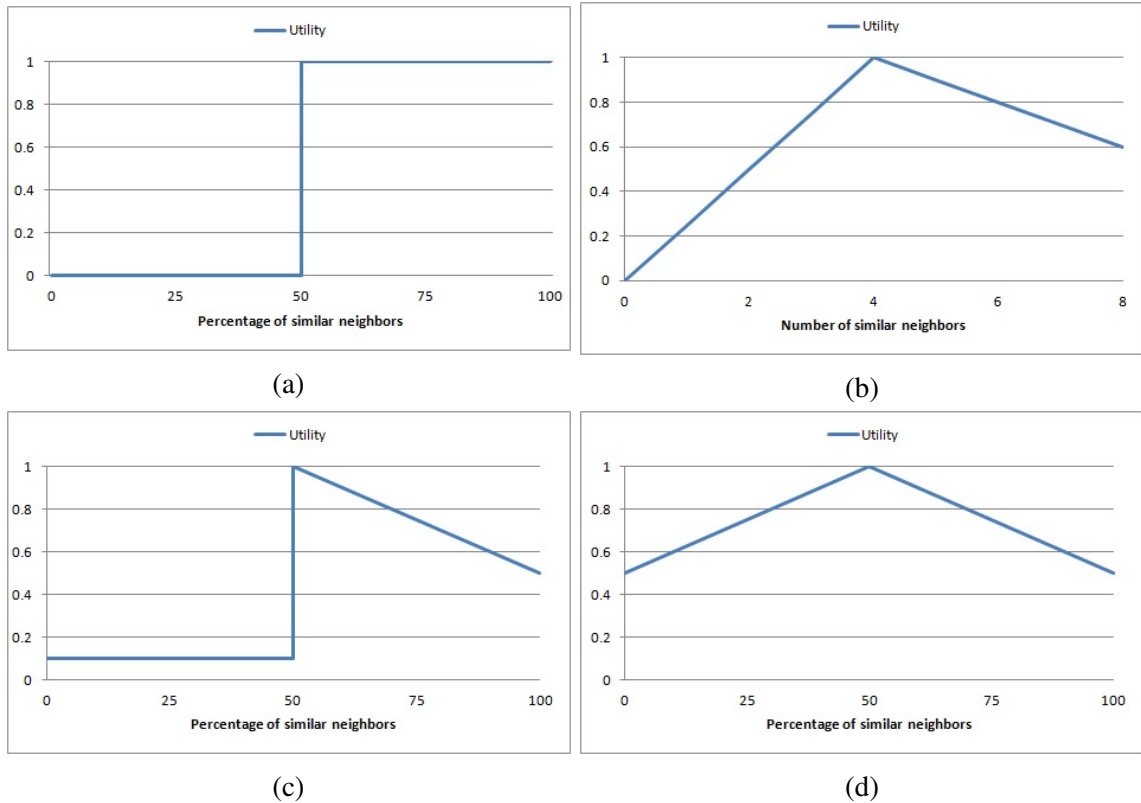


Figure 2.1: Different utility functions from the literature. (a) represents Schelling's model, Zhang uses the function in (b) and Pancs and Vriend use functions in (c) and (d) as well as Schelling's function and a spiked utility function [7] [16] [10].

neighbors have higher utility than 70% other 30% similar neighbors in Zhang's model. This might be the reason for segregation even if the integration produces the highest utility. So in the next sub-section different symmetric and asymmetric utility functions will be used.

Throughout the simulations, there will be 37×37 lattice which makes 1369 nodes. Larger environments could cause computational difficulties with long run time. 369 nodes are vacant places which correspond to 27% of total nodes. In Zhang's work, all nodes are occupied, an agent switches its place with another agent if their total utility increases [16]. Fagiolo et al. work with a network that 30% of the nodes are vacant [15]. In their two dimensional setup, Pancs and Vriend have 20% empty cells [10]. In Banos's work, this ratio is 20% too and he also emphasises that too many vacant nodes would limit the interaction between agents while too few vacant nodes may freeze the system, which means agents are not able to find suitable place to move [12]. Schelling also concludes that 25% to 30% vacancies let nodes move freely

without having a too empty lattice [7]. As can be seen in appendix A, different percentages for vacant cells are simulated. When the percentage of vacant cells is low, equilibrium takes more time with increasing lower bound. However, when there is also an upper bound, settings with too many vacant cells start to give results in longer time. It seems convenient to have 20% to 30% vacancies.

Moreover, there is perfect information; each node is aware of the attributes of others and places of vacant cells. Agents' attributes do not change during the simulation which means the attribute that makes the division between agents is immutable such as race or ethnicity. At each time step, agents who are not satisfied with their neighborhood move to vacant cells in random order. There is not a moving cost, so an unhappy agent can move to any vacant place. If triangular or circular utility functions, which will be defined in the section 2.2, are used or if step function is compared with other utility functions, agents move to a vacant cell which gives the highest utility and if there are more than one vacancies with the highest utility, agents choose one of them randomly. Moving continues until either each node is satisfied with its place or there is no vacant cell that give a better utility for any of the nodes. In some symmetric and asymmetric utility cases moving does not stop, small number of agents continue to circulate, but segregation indices still converge to some point.

2.1 The Effect of an Upper Bound on Segregation

In Schelling's research and following studies, it is assumed that agents are looking for others who are similar to themselves. However, at the same time agents may look for diversity for different reasons. For instance, in Farley et al.'s study, most of the black respondents say that they would rather live in mixed neighborhoods than mostly black neighborhoods so that their children know different kinds of people and learn to get along with others [17]. A very recent study conducted in London indicates that ethnic diversity of the neighborhood has positive effect on perceived social cohesion of residents [18]. Moreover, soon to be published report of Istanbul based think tank Public Policy and Democracy Studies (PODEM) reveals that educated young people who live in segregated districts would like to move to the mixed neighborhoods to live more freely [19]. So, in simulations an upper bound might be defined

and it would force agents to be neighbors with the ones who are not similar with themselves. For instance, if an agent's lower bound is 25% and upper bound is 95%, that agent will be happy only if its similar neighbor ratio is more than or equal to 25% and less than or equal to 95%.

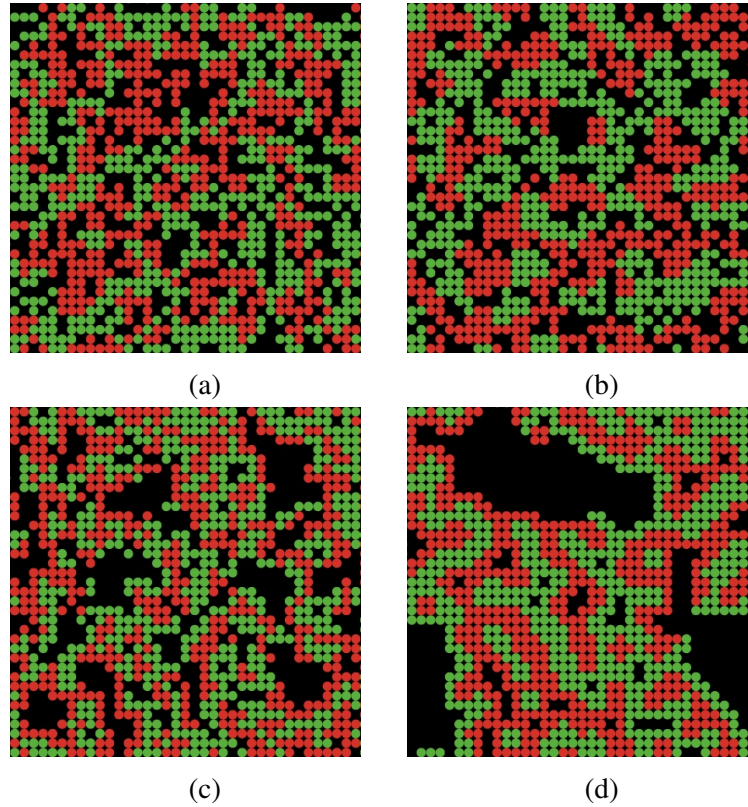


Figure 2.2: In (a) and (b) there are only lower bounds which are 25% and 35% respectively. In (c) and (d) 95% upper bound is added to simulations. If we compare (a) with (c) and (b) with (d), it is possible to say that networks become more integrated with the existence of an upper bound, even if the upper bound is very high.

Even if the upper bound is high, existence of an upper bound might reduce segregation. In figures 2.2a and 2.2b lower bounds are 25% and 35% respectively, and there is not an upper bound. Figures 2.2c and 2.2d are the outcomes when there is also a 95% upper bound. It is observed that agents come together with dissimilar agents. Segregation measures support this conclusion, as can be seen in table 2.1, network with 25% lower bound is slightly segregated, when an upper bound is added it becomes more integrated. Network with 35% lower bound is segregated, and addition of the upper bound makes it more integrated as well. As a result, we can say that introducing an upper bound supports integration.

Table 2.1: Change in segregation measures when there is an upper bound. Mixity and FSI values indicate that existence of an upper bound reduces segregation.

	25% no upper	25%-95%	35% no upper	35%-95%
mixity	0.38	0.46	0.20	0.36
FSI	0.23	0.07	0.59	0.29

If a network is already segregated, it will remain segregated, because none of the agents have an incentive to move near dissimilar agents. However, if an upper bound is defined for the agents of a segregated network, segregation diminishes as well. Table 2.2 shows that 25%-95% lower-upper bound couple gives the same result when it is applied to a network in which agents are randomly distributed and when it is applied to a network which is already segregated by 25% lower bound. Thus, even if the network is already segregated, if we can make agents look for diversity, segregation is not the ultimate result.

Table 2.2: Effect of an upper bound for different initial conditions of networks. 25%-95% lower-upper bound couple gives the same results when it is applied to a random initial condition and to a segregated initial condition. Even if the network is segregated, an upper bound decreases the segregation.

	Random + 25%-95%	25% + 25%-95%	Random + 35%-95%	35% + 35%-95%
mixity	0.46	0.46	0.36	0.35
FSI	0.07	0.08	0.29	0.29

2.2 Symmetric Utility Functions

The functions which will be investigated in this section are symmetric with respect to 50 percent of similar neighbors. If agents are equally happy when the percentage of their similar neighborhood is between the lower and upper bounds, then utility function is a step function, an example can be seen in figure 2.3a. Figures 2.3b and 2.3c show symmetric utility functions that take the maximum value with 50% similar 50% dissimilar neighbors, which means integration leads to the highest utility, and we will call them triangular and circular functions

respectively. Mathematical representation of these functions as follows.

$$\text{The percentage of similar neighborhoods} = S \quad (2)$$

$$\text{Lower bound} = L \quad (3)$$

$$\text{Upper bound} = U \quad (4)$$

$$\text{Utility of the step function} = \begin{cases} 0, & \text{if } S \leq L \\ 1, & \text{if } S > L \\ 0, & \text{if } S \geq U \end{cases} \quad (5)$$

$$\text{Utility of the triangular function} = \begin{cases} 0, & \text{if } S < L \\ \frac{(S-L)}{(0.5-L)}, & \text{if } L \leq S < 0.5 \\ \frac{(U-S)}{(U-0.5)}, & \text{if } 0.5 \leq S \leq U \\ 0, & \text{if } S > U \end{cases} \quad (6)$$

$$\text{Utility of the circular function} = \begin{cases} 0, & \text{if } S < L \\ \sqrt{2} \sqrt{\left(\frac{U-L}{2}\right)^2 - (S - 0.5)^2}, & \text{if } L \leq S \leq U \\ 0, & \text{if } S > U \end{cases} \quad (7)$$

When utility functions are used, agents are moving to vacant places which give the highest utility, otherwise having a triangular or circular utility function would be meaningless. However, when we are not comparing the step utility function with triangular and circular utility functions, agents can move to any vacant cell. As can be seen in appendix B, outcomes are not different, but looking for a better place produces the results in longer time as at each time step agents calculate potentials of all vacant cells and compare this potential with their current utility. Moreover, when we compare the step utility function with other functions, we assumed agents' similar neighbors ratio should be strictly more than the lower bound and less than the upper bound, because at triangular and circular functions, lower and upper bounds result in

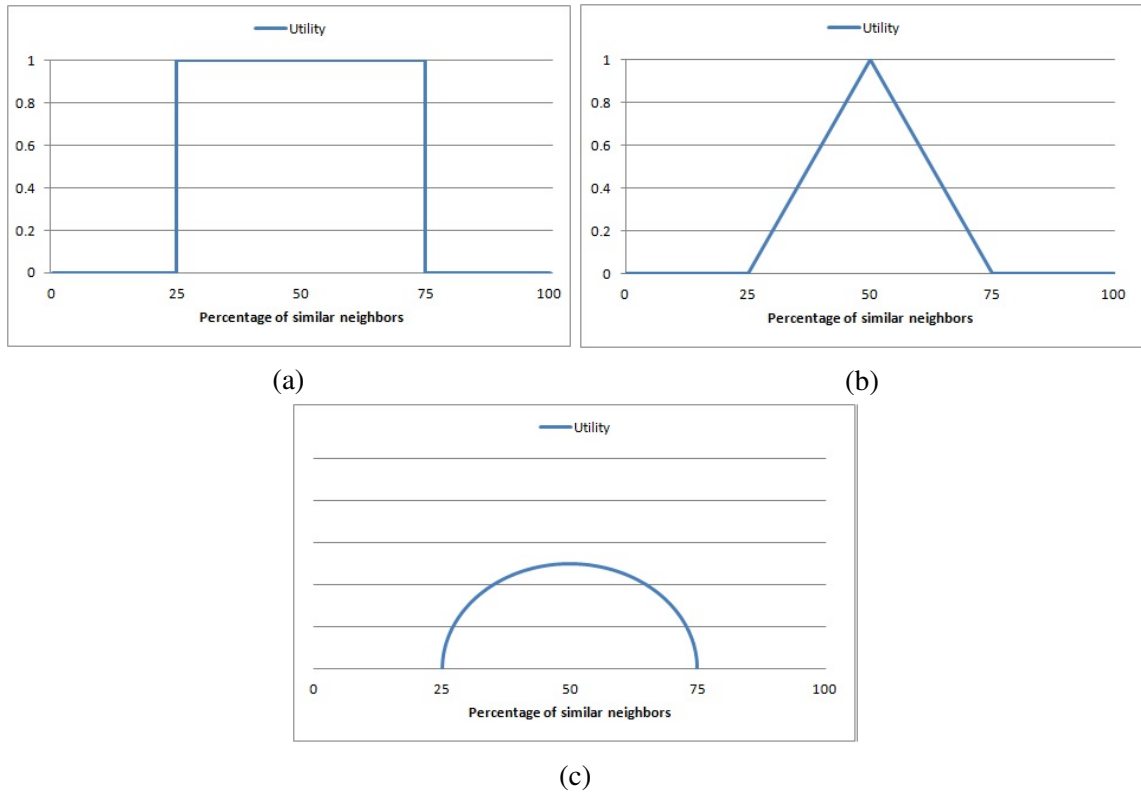


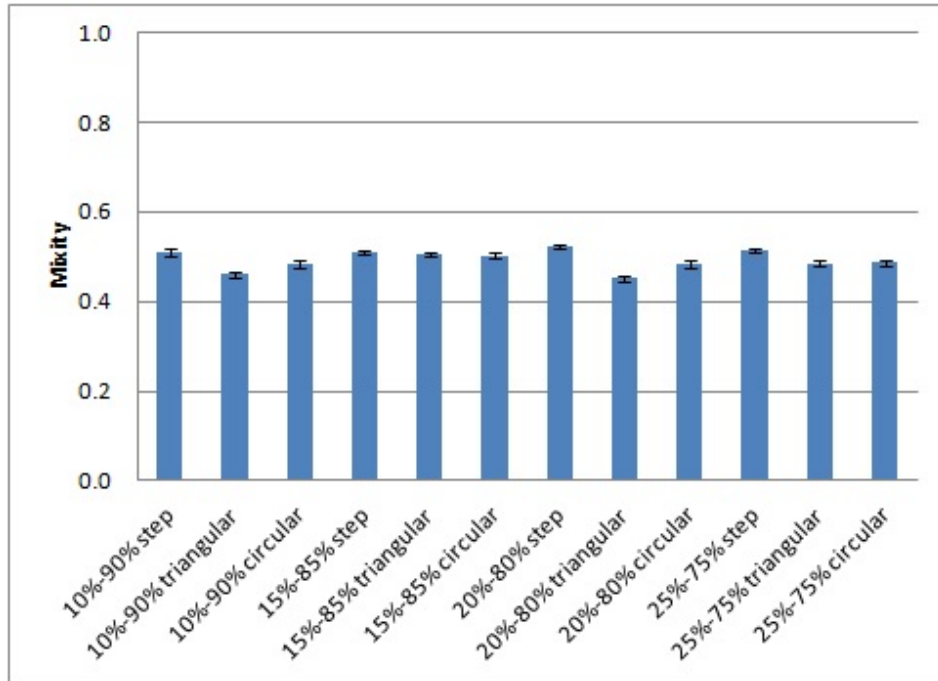
Figure 2.3: Symmetric utility functions which are investigated in this study. a is a step, b is a triangular and c is a circular utility function.

utility 0.

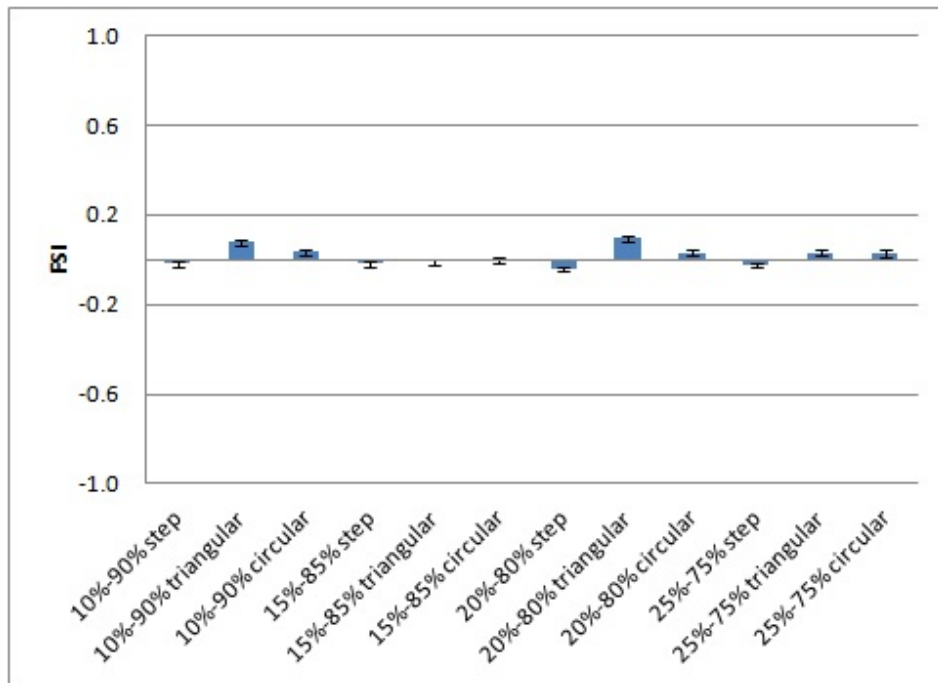
Three symmetric functions of four different lower-upper bound couples, which are 10%-90%, 15%-85%, 20%-80% and 25%-75%, are investigated here. Resulting mixity and FSI values can be seen in figure 2.4. In this and in the following figures error bars represent one standard deviation. In each of the 12 settings, mixity is around 0.5 and FSI is around 0, as expected from the integrated networks ².

As a result, different symmetric utility functions with different lower and upper bounds do not cause segregation, but provide integration, if function is symmetric with respect to 50 percent of similar neighbors. This finding holds when we change degree as 12, results are available in appendix C.

²20%-80% triangular function gives lower value of integration comparing to others, so bounds that are close to 20%-80% are also tried. From 19%-81% to 23%-77% same less integrated results are obtained. Moreover, as can be seen in the next section, same results are obtained for the ring lattice and small world networks. When the degree is increased to 12 in appendix C, the difference between the results of 20%-80% triangular function and other symmetric utility functions decreases.



(a)



(b)

Figure 2.4: Mixity and FSI values of different symmetric utility functions. In all of the settings, mixity is around 0.5, and FSI is around 0 which indicate integration.

2.3 Asymmetric Utility Functions

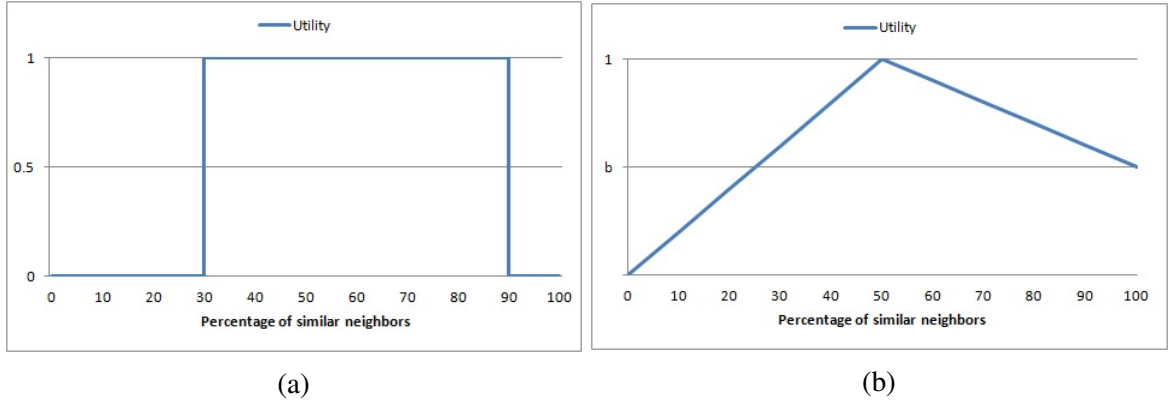


Figure 2.5: Two asymmetric utility function samples which favors similar neighbors. (a) is a step function with asymmetric lower-upper bounds, and function in (b) gives the highest utility to integration, but in case of not having 50% similar neighbors, similarity is favored.

As mentioned previously, asymmetric utility functions are studied by other researchers. However, as can be seen in figures 2.1a, 2.1b and 2.1c, those functions favor similar neighbors. We will also investigate what happens when asymmetry favors dissimilar neighbors.

We will use two types of asymmetric utility functions. First one is a step function with asymmetric lower and upper bounds. Figure 2.5a is an example with 30% lower bound and 90% upper bound. In symmetric utility functions, $L + U$ was equal to 1, but it is not equal to 1 anymore. If $L + U > 1$, then function favors similar neighbors, and it is the other way around if $L + U < 1$.

Second asymmetric utility function type will be the one shown in figure 2.5b, so integration produces the highest utility, but in case of not having 50% similar 50% dissimilar neighborhood, one side is favored. In figure 2.5b similarity is favored as utility is 0 when similar percentage is 0, and utility is b when similar percentage is 100. This function is obtained by equation 9. When dissimilarity is favored, equation 10 will be used, it gives utility b for 0% similar neighbors and utility 0 for 100%.

$$0 < b < 1 \quad (8)$$

$$\text{Utility of the asymmetric functions} \begin{cases} \frac{S}{0.5}, & \text{if } S \leq 0.5 \\ \frac{(b-1)S + 1 - 0.5b}{0.5}, & \text{if } 0.5 < S \end{cases} \quad (9)$$

$$\begin{aligned}
& \text{Utility of the asymmetric functions} \begin{cases} \frac{(1-b)S + 0.5b}{0.5}, & \text{if } S \leq 0.5 \\ \frac{1-S}{0.5}, & \text{if } 0.5 < S \end{cases} \\
& \text{which favors dissimilarity} = \begin{cases} \frac{(1-b)S + 0.5b}{0.5}, & \text{if } S \leq 0.5 \\ \frac{1-S}{0.5}, & \text{if } 0.5 < S \end{cases}
\end{aligned} \tag{10}$$

We simulated 4 step functions with asymmetric lower and upper bounds and 6 functions which favor one of the sides. Resulting mixity and FSI values are available in figure 2.7. 50%-100% and 30%-90% step functions favor similarity and cause segregated networks with mixity values 0.11 and 0.38 respectively. Other functions which favor similarity result in segregation as well despite giving the highest utility to integration. This result is consistent with the previous studies. Moreover, as Grauvin et al. indicate at their study, with increasing value of b , segregation level increases [20]. For instance, when b is 0.25, mixity is 0.4 whereas it is 0.3 when b is 0.75.

Step functions which favors dissimilarity give rise to inverse homophily, especially at 0% - 50% setting, which has a mixity of 0.7. On the other hand, other asymmetric functions favoring dissimilarity produce mixity values around 0.51-0.52, which means integration. They do not result in an inverse homophily, while their counterparts cause segregation. Because when dissimilar neighbors give higher utilities, agents try to move near others which are not similar to themselves, and for the very same reason, they will come across with the agents similar to themselves. However, when similarity is favored, agents move to places where they find similar neighbors, and maybe make some dissimilar agents unhappy and force them to move and segregation emerges easily.

Figure 2.6 shows simulation outcomes of four asymmetric utility functions. Lower and upper bounds are 50% - 100% in figure 2.6a and network is homophilic with just a few heterogeneous neighborhoods. In figure 2.6b lower and upper bounds are 0% - 50 % and high ratio of heterogeneous neighborhoods indicates inverse homophily. In figure 2.6c similarity and in figure 2.6d dissimilarity are favored and b is given as 0.75 at both of them. There is a segregation at 2.6c and integration is observed at 2.6d. Figures 2.6b and 2.6d may seem similar at the first sight, however at former case, red and green agents form straight lines next to each other, so an agent can have 2 neighbors similar to itself and 6 dissimilar neighbors. And in latter case there is not such an obvious pattern, aligning seems more random as expected from

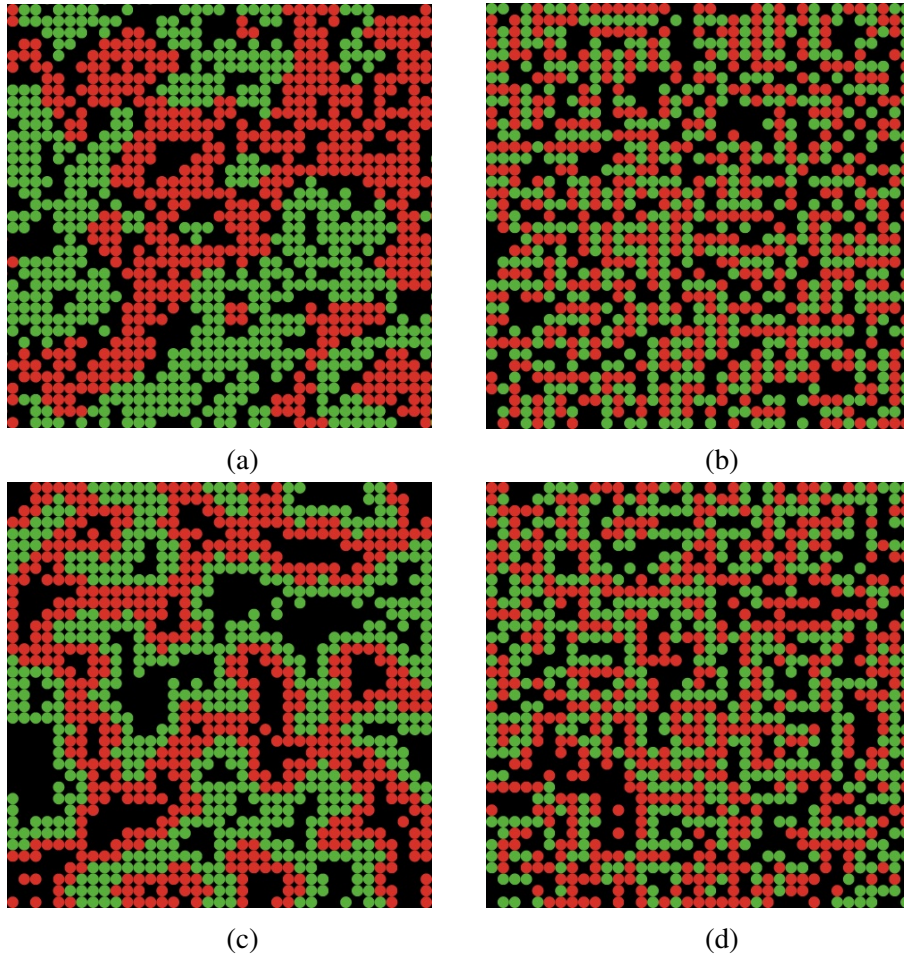
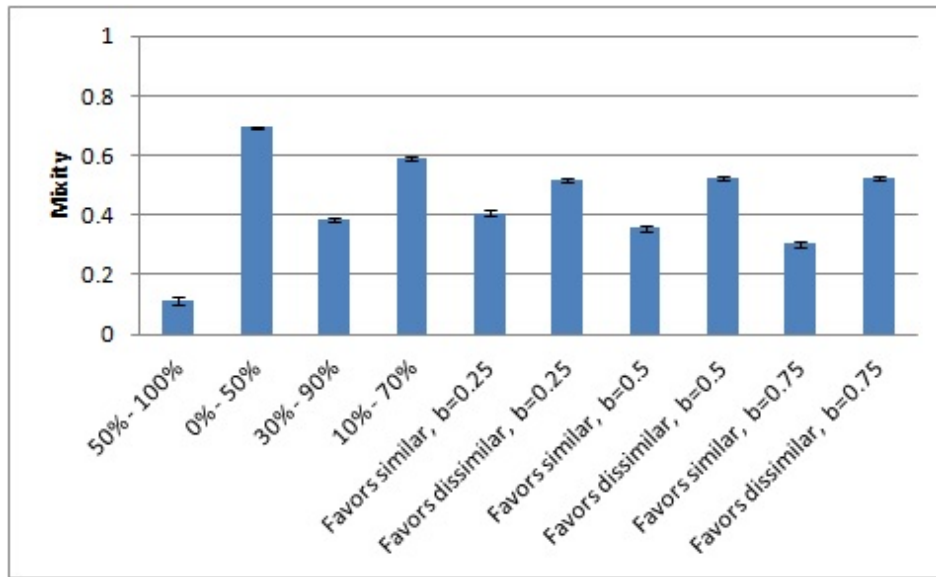


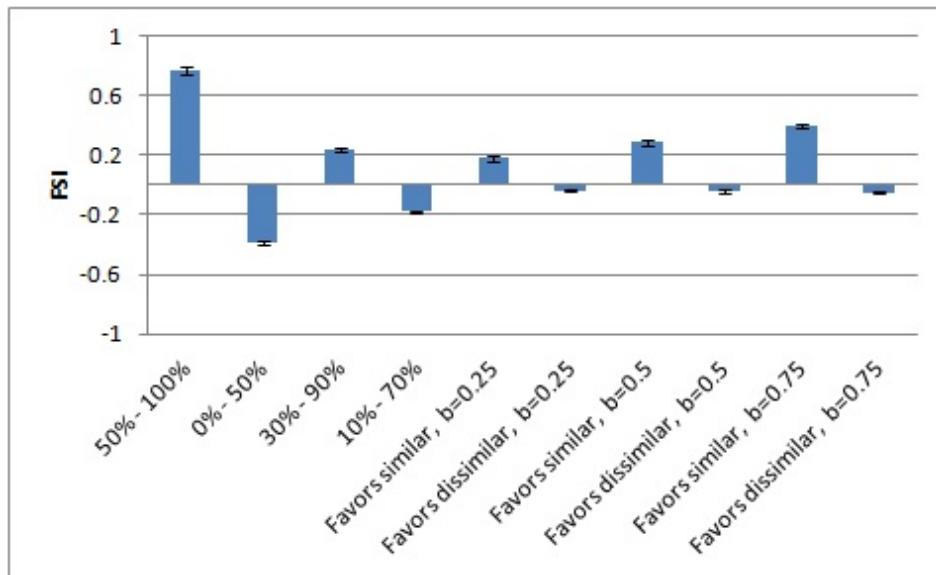
Figure 2.6: Simulation outcomes of networks with asymmetric utilities. Lower and upper bounds are 50% - 100% in setting (a) and 0% - 50 % in setting (b). (a) is homophilic with just a few heterogeneous neighborhoods and (b) is inverse homophilic with high ratio of heterogeneous neighborhoods. In setting (c) similarity and in setting (d) dissimilarity is favored and $b = 0.75$ at both of them. There is segregation in (c) and integration is observed in (d).

integration.

To sum up, when the utility function is asymmetric and favors similarity, segregation occurs. The more the utility favors similarity, the more the society segregates. If asymmetric utility functions favor dissimilarity, inverse homophily or integration might be obtained. Moreover, having an inverse homophilic network is harder than having a homophilic network.



(a)



(b)

Figure 2.7: Mixity and FSI values of different asymmetric utility functions. Utilities which favor similarity produce low level of mixity and high level of FSI which shows segregation. Step utility functions which favor dissimilarity result inverse homophilic networks, whereas other functions which favor dissimilarity produce integration.

2.4 Segregation When There is a Minority Group in Network

Race, ethnicity, political view, income level or other attributes and characteristics are not distributed in society evenly. The question here is how would be the segregation under different utility functions if there is a minority group in network.

In this sub-section, 20% of the agents will be one type and 80% of the agents will be the second type. Choosing minority percentage as 20% seemed convenient as 20%-25% of Turkey's population consists of ethnic minorities [21] [22].

Table 2.3: Resulting mixity and FSI values and ratio of the unhappy agents for different lower-upper bound couples when there is a minority group in network. If all agents or just the majority members have an upper bound, there are unhappy agents in the system. However, if just the minorities have an upper bound, segregation level decreases without causing unhappy agents.

	Mixity	FSI	Minority utility= 0	Majority utility= 0
All 50%	0.090	0.717	0.00	0.00
All 25%	0.243	0.242	0.00	0.00
All 25%-95%	0.333	-0.042	0.17	0.26
Majority 25%-95% minority 50%	0.349	-0.092	0.73	0.09
Minority 25%-95% majority 50%	0.187	0.417	0.00	0.00
All 25%-75%	0.369	-0.152	0.44	0.48
Majority 25%-75% minority 50%	0.360	-0.126	0.85	0.48
Minority 25%-75% majority 50%	0.198	0.382	0.00	0.00

As can be seen in the table 2.3, when each of the agents wants at least 50% of their neighbors to be like themselves, network is highly segregated. When each agent wants at least 25% of their neighbors to be like themselves, segregation decreases. Thus, if policy makers find ways for agents to be more tolerant, which indicates decrease in the lower bound, segregation diminishes. Feld emphasises that shared organizations, activities or interests increase the likelihood of two people to interact [23] and there are field studies which support that idea. For instance, PODEM's report indicates that people who have communication with other ethnicities are less prejudged and more tolerant towards others [19]. It is also emphasized in the London study; young people who grown up in more diverse regions are less likely to be prejudged toward other races, and social institutions supporting contact might be facilitating

to get to know dissimilars [18].

Addition to the decrease in the lower bound, if we make agents look for diversity with addition of an upper bound to simulation, there are unhappy agents in network. If we just make the majority members more tolerant and look for diversity, while minority members have 50% lower bound, most of the minorities (73%) are unhappy. However, if just minorities are more tolerant and look for diversity, each agent is happy, and comparing with the situation in which all nodes have 50% lower bound, segregation decreases even though the majority still has 50% lower bound. At that point, comparing different settings with the FSI is better as it accounts expected and observed heterogeneous neighborhoods. Mixity is already expected to be low at all cases because of the unequal distribution of the different types. Results are consistent when we apply 25% lower and 75% upper bounds, if we change only minorities' utility, there are no unhappy agents.

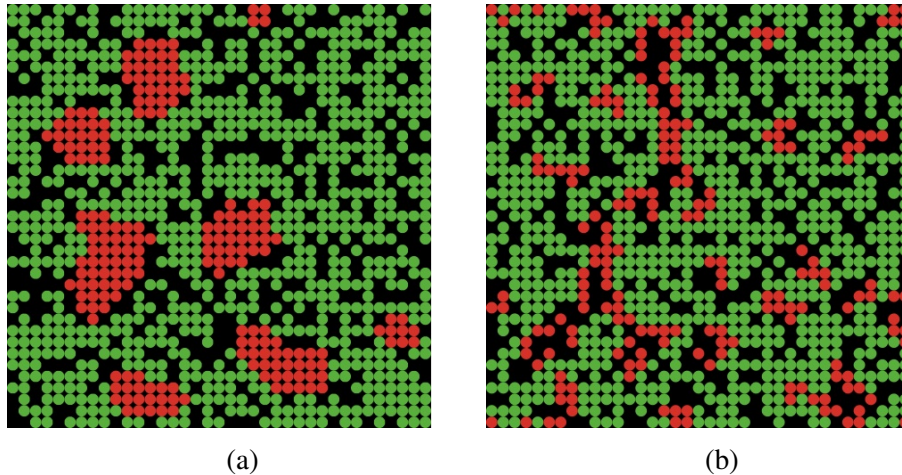


Figure 2.8: Simulation outcomes of networks which have a minority group. There is 50% lower bound in (a) and minorities form some ghettos. In (b) majorities still have 50% lower bound, but minorities have 25% lower and 75% upper bound and network is more integrated.

Figure 2.8 shows outcomes of 2 different utilities that minorities have and in both of them all agents are happy with their places. Figure 2.8a is the outcome when each agent has 50% lower bound, Society is highly segregated and minorities form some ghettos. In figure 2.8b majority members still have 50% lower bound, however minorities have 25% lower and 75% upper bound. Comparing with the left figure, society is more integrated.

In conclusion, to decrease the segregation level at a society which has a minority group, making

agents more tolerant, in other words decreasing the lower bound might be an option. Moreover, if some agents will look for diversity, these should be the minorities. Otherwise, minorities become surrounded by the majority members which makes minorities unhappy.

2.5 The Effect of the Randomized Utility Functions

Agents with the same race or ethnicity might have different attributes such as gender or age. Thus, even if two agents are same with regard to one of the attributes, they might have different utility functions. In this subsection, we would like to see the effect of having some agents which are more tolerant and looking for diversity.

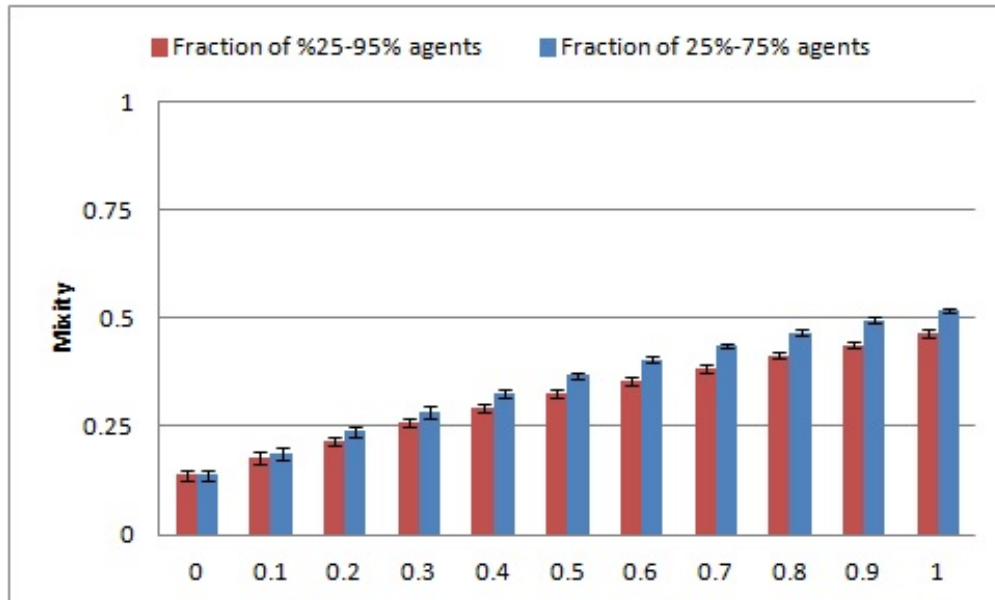


Figure 2.9: The resulting mixity values when some of the agents have a higher tolerance and desire for diversity. At the beginning, all agents have 50% lower bound, then some of them are changed as 25%-95% in one case and changed as 25%-75% in the second case. Even if just 20 percent of the agents have the 25%-95% or 25%-75% bounds, segregation level decreases. If most of the agents have these bounds, network is nearly integrated although there are still some agents which favor similarity.

At the beginning of the simulations, all agents have 50% lower bound. Then to make some agents have a higher tolerance and desire for diversity we changed some of the agents' bounds as 25%-95% in one case and changed as 25%-75% in another. As can be seen in the figure 2.9 even if just 20 percent of the agents have these bounds, segregation level decreases. If most

of the agents have the bounds 25%-95% or 25%-75%, network is nearly integrated although there are still some agents which favor similarity. Hence, if policy makers cannot change the utility function of everyone, changing it for some people is a step for integration.

We can look into the effect of agents with randomized utility functions when there is a minority group. As mentioned before, Sturgis et al.'s and PODEM's studies show young people might be more tolerant and might look for diversity [18] [19]. Thus, we can run the simulations taking half of the minority members and the half of the majority members as more tolerant and looking for diversity assuming the half of the population is young.

Table 2.4: The resulting segregation indices when there is a minority group and randomized utility functions. Making half of the minority and half of the majority members more tolerant and look for diversity decreases segregation.

	Mixity	FSI	Minority utility=0	Majority utility=0
All 50%	0.09	0.72	0	0
Half 50%, other half 25%-95%	0.27	0.15	0.19	0.10
All 25%	0.24	0.24	0	0
Half 25%, other half 25%-95%	0.31	0.04	0.05	0.10
Minority 25% majority 50%	0.17	0.47	0	0
Half of minority 25% other half 25%-75%				
Half of majority 50%, other half 25%	0.21	0.33	0	0

Table 2.4 shows making half of the minority and half of the majority members more tolerant and look for diversity decreases the segregation. As a result, policy makers can focus on youth as a beginning to reduce the segregation level of the society.

3 Segregation in Different Networks

In the literature not only the regular lattice, but also other type of networks are investigated in terms of segregation. Fagiolo et al. use Von-Neumann, Moore, regular, random, small world and scale free networks with the utility function like Schelling's and conclude that network type does not affect segregation [15]. Banos does not agree with this result; he tries regular lattice, random, scale free and fractal (Siempinski tree) networks and indicates that when the lower bound is between 30% - 70%, different networks segregate differently [12].

In this section, segregation of the regular ring lattice and small world networks will be compared. Small world networks have high clustering coefficients like regular lattices, but also have small average shortest path lengths between pair of nodes like random networks [24]. The clustering coefficient refers to the likelihood of two neighbors of an agent also being neighbors with each other [1]. For instance, if a node has 4 neighbors and only one pair of them know each other, this node's clustering coefficient is 1 over 6 as there are 6 possible relationships between 4 nodes. Watts and Strogatz further indicate that many real-life networks show small world network property and this network can be obtained by rewiring the links of a regular ring lattice with a small probability p . These random links create short cuts which decrease path lengths without decreasing the clustering coefficient much [24].

Figure 3.1 is obtained by simulating ring lattice which has 500 nodes, average node degree 8 and rewiring probabilities from 0.0001 to 1. Each data point is the average of 50 simulations and resulting figure is identical with the Watts and Strogatz's simulation which has 1000 nodes, average node degree 10 and 20 realizations [24]. For small values of p , both of the clustering coefficient (C) and average shortest path length (L) are high as they are in regular

lattice. When p is between 0.01-0.02, C is still high, but L is low, so this is the region that gives us small world network. If p continues to increase, C decreases, and network starts to be like a random network.

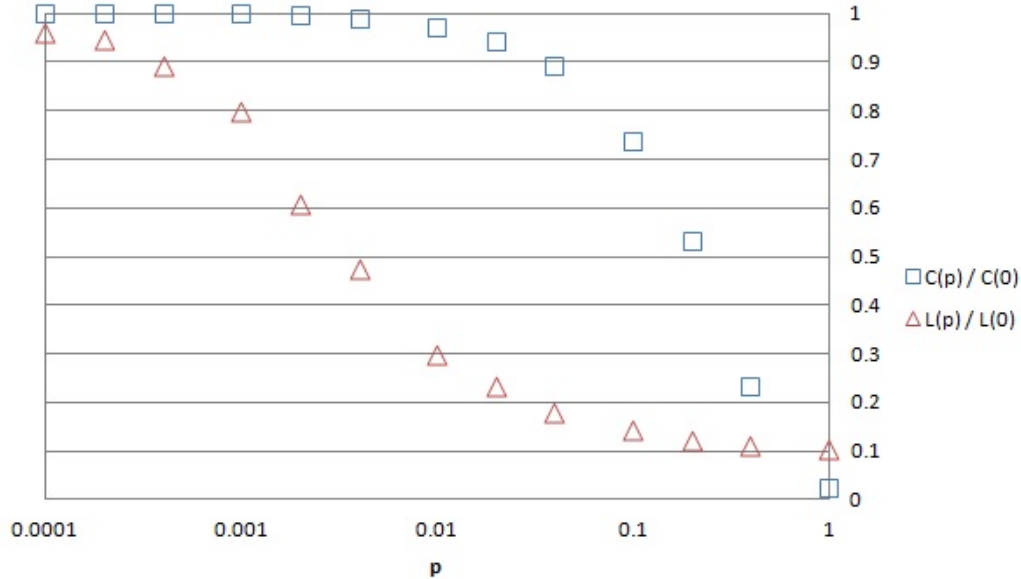


Figure 3.1: Change in clustering coefficient (C) and average shortest path length (L) with increasing rewiring probability (p). When p is around 0.01-0.02, the average shortest path length is low, but clustering coefficient is still high, this region is named small world network.

The degree distributions of the ring lattice and small world networks are illustrated in figure 3.3. In ring lattice, all nodes have exactly 8 neighbors, in small world network most of the nodes have 8 neighbors, but there are also a few nodes whose number of neighbors are not 8, but between 5 and 11.

In figure 3.2a there is a regular ring lattice in which each agent is neighbor with 4 agents on its right and 4 agents on its left. In figure 3.2b, the small world network which is obtained by rewiring the links of the regular ring lattice with a rewiring probability $p = 0.02$ can be seen. For bigger values of p , the network starts to be random. For instance, in figure 3.2c p is 1, so each link rewires and we obtain a random network.

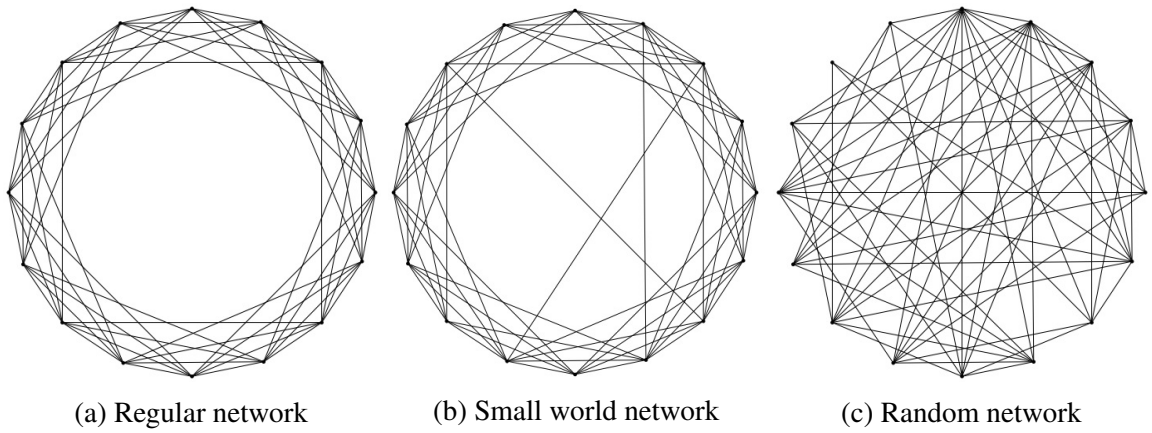


Figure 3.2: Regular network with node degree 8, small world network which is obtained from this regular network with rewiring probability 0.02, and the random network which is obtained with rewiring probability 1.

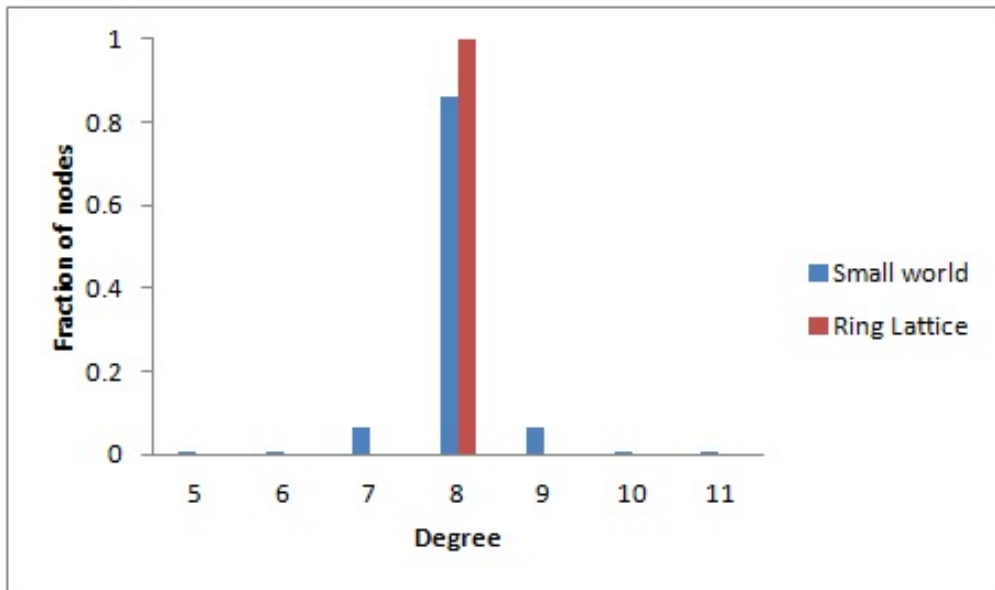


Figure 3.3: Degree distributions of the ring lattice and small world network. In ring lattice each node has 8 neighbors, but rewiring some of the links creates nodes which have 5 to 11 neighbors in the small world network.

3.1 Segregation in Ring Lattice and Small World Networks

Throughout the simulations, the ring lattice with degree 8 and the small world network which is obtained from this ring lattice with 0.02 rewiring probability are used. Rewiring probability is chosen as 0.02 to have more random links and observe differences between the ring lattice and the small world network better. Both of the networks have 1360 nodes, 500 of them are green, 500 of them are red and the remained 360 are the vacant cells. Moving algorithms are same with the algorithms used in the previous sections.

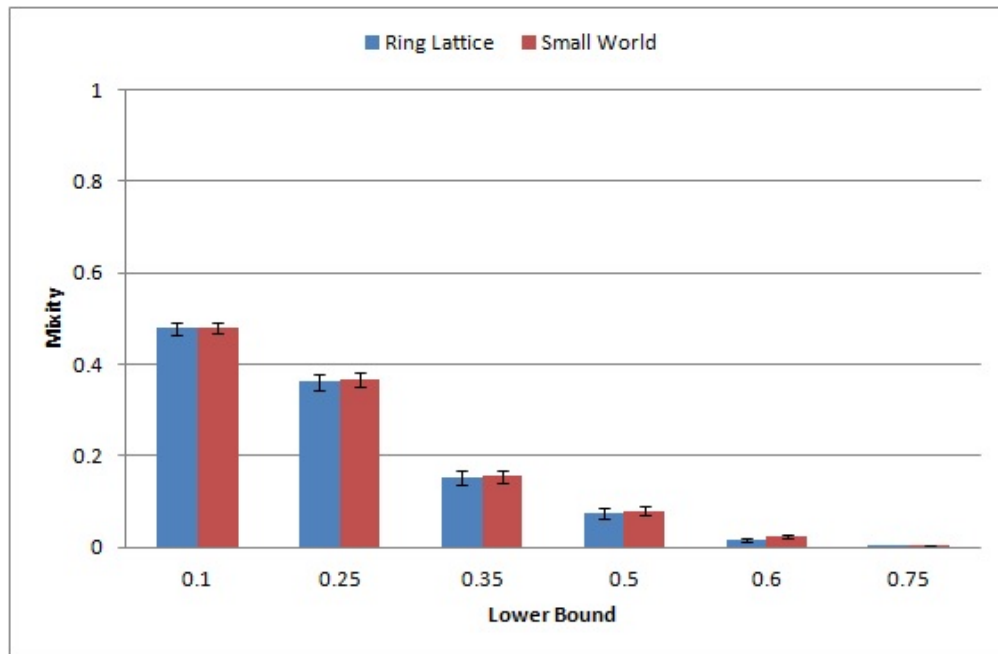
Firstly, we compared the ring lattice and small world network for different lower bound values. Resulting outcomes are represented in figure 3.4 and we see that these two networks have the same mixity and FSI values for variety of lower bounds. This result is consisted for different degrees. When the degree is 4 or 12, the ring lattice and the small world networks still produce same mixity and FSI values. Graphs are available in appendix D.

Fagialo and Banos compare different networks when there is just a lower bound. In addition to lower bound, here the ring lattice and small world networks compared with different symmetric utility functions. These functions are the ones which were used in sections 2.2 and 2.3.

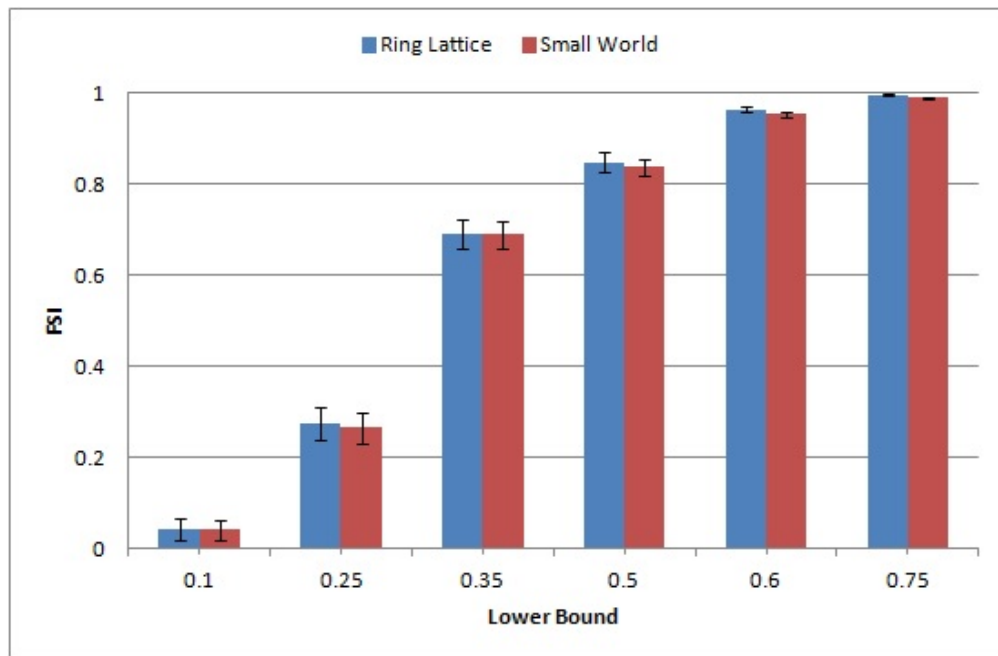
As can be seen in figure 3.5, mixity and FSI values of the ring and small world networks are the same and indicate integration as observed with the regular lattice.

Figure 3.6 shows the mixity and FSI values of ring lattice and small world network for different asymmetric utility functions. For asymmetric utility functions these networks give the same results as well. Moreover, when similar neighbors are favored, segregation emerges, and 0%-50% bounds, which favors dissimilarity, results in inverse homophily with 0.65 mixity value. However, other functions that favor dissimilarity produce integration, not an inverse homophily. These results are consistent with what we observed with the regular lattice.

As a result, figures 3.4, 3.5 and 3.6 show that if we compare the ring lattice and the small world networks with mixity and FSI, these networks do not segregate differently. Moreover, utility functions which are symmetric with respect to 50 percent of similar neighbors give rise to integration at ring lattice and small world network whereas asymmetric utility functions might result segregation, inverse homophily or integration as observed with the regular lattice.

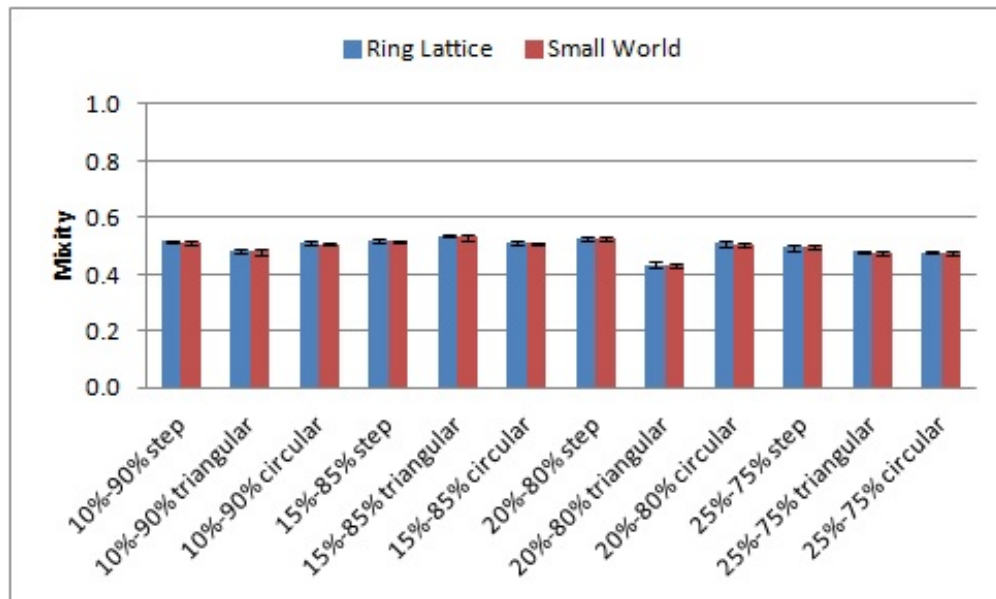


(a)

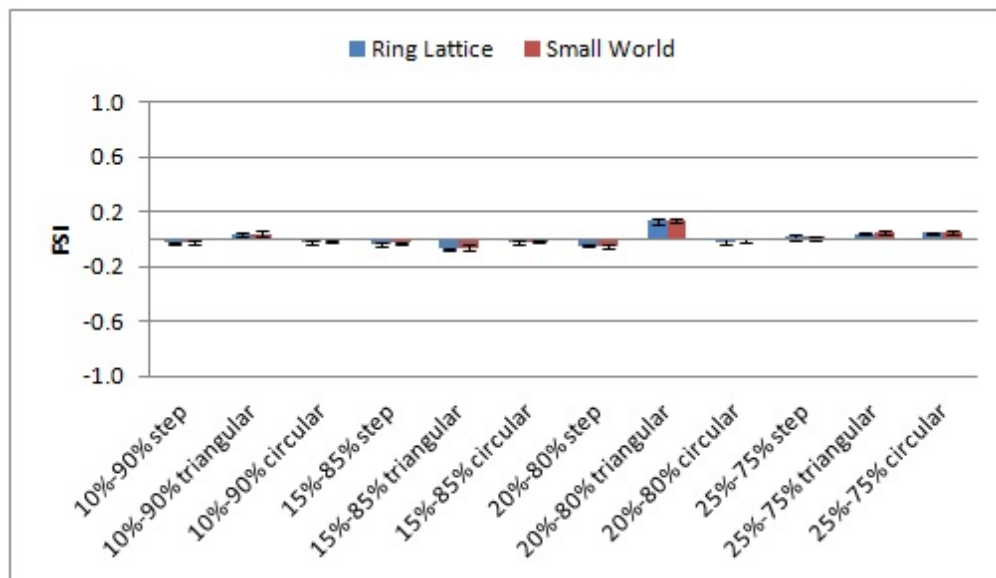


(b)

Figure 3.4: Mixture and FSI values of the ring lattice and small world network when there is just a lower bound. Resulting values are overlapping for variety of lower bounds.

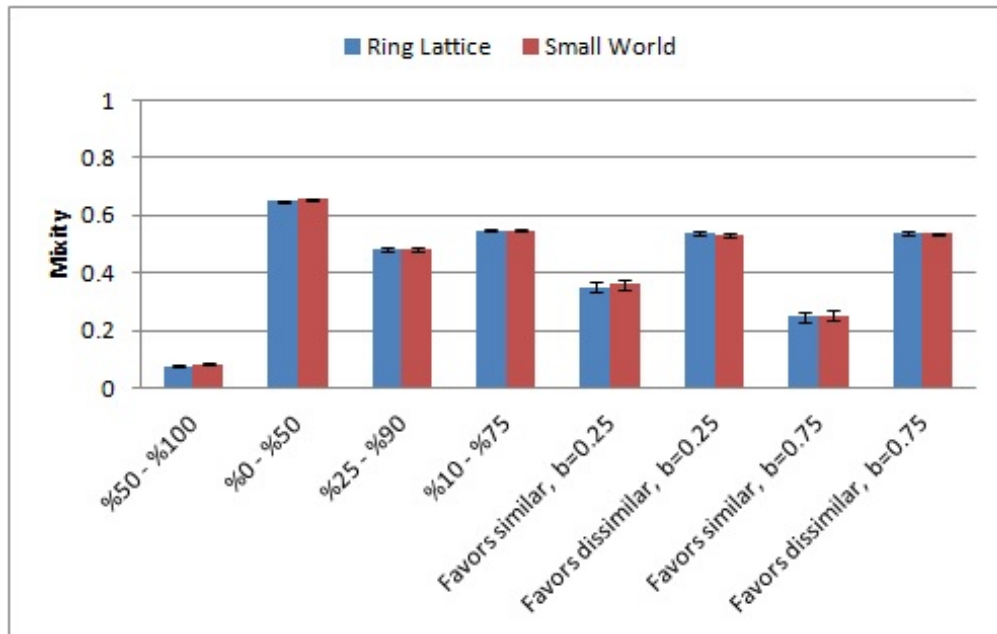


(a)

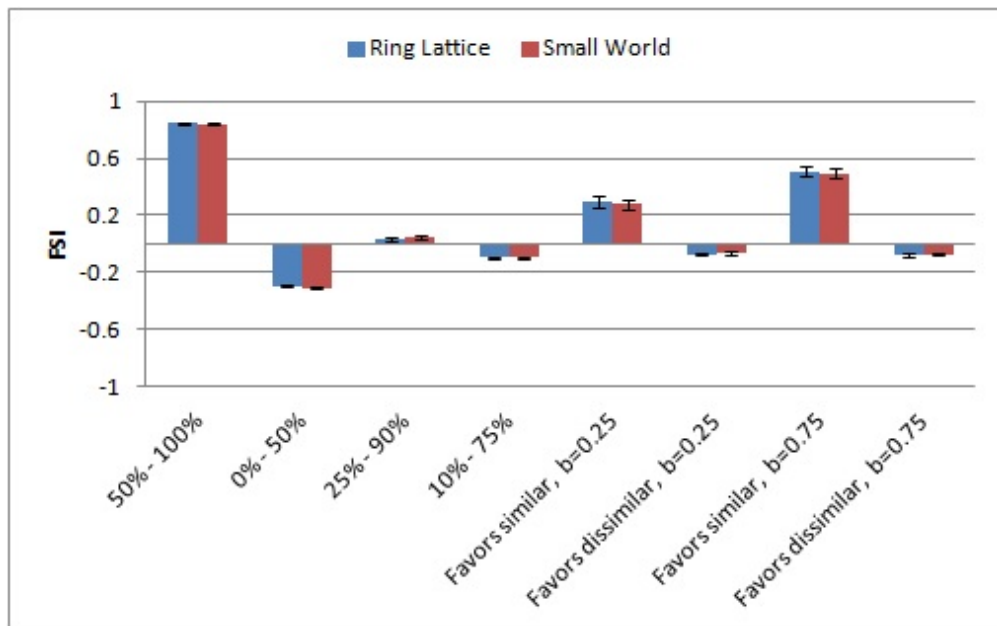


(b)

Figure 3.5: Mixity and FSI values of the ring lattice and small world network for different symmetric utility functions. As a result, both of the networks' segregation indices show integration for symmetric utility functions.



(a)



(b)

Figure 3.6: The mixity and FSI values of the ring lattice and small world network for different asymmetric utility functions. These networks give the same results for different asymmetric utility functions as well. Moreover, when similar neighbors are favored, segregation emerges, and 0%-50% bounds which favors dissimilarity results in inverse homophily with 0.65 mixity value. However, other functions which favor dissimilarity produce integration, not an inverse homophily. These results are consistent with what we observed with the regular lattice.

4 Neighborhood Distribution Function

Mixity and FSI indices indicate networks' segregation level, however they may be not sufficient to catch the network patterns. This also means even if mixity and FSI indices of different networks are similar, resulting network patterns might be different. For instance, the network in figure 4.1a has 25% lower bound, and the network in figure 4.1b has 30% lower bound with 95% upper bound. Both of these networks have 0.38 mixity and 0.23 FSI, but there are bigger islands of vacant cells in figure 4.1b.

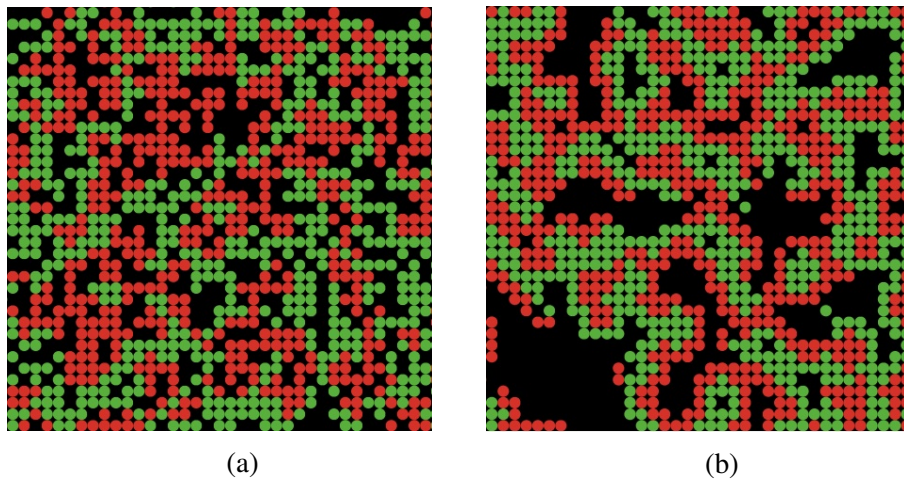


Figure 4.1: Two networks with different utility functions, but same segregation indices. Even if the mixity and FSI values are the same, outputs might be different, so we need to define another measurement to compare the networks.

Moreover, if there are such differences in the ring lattice or small world networks, it is not possible to catch them by looking in the networks as neighbors are not nearby each other, but they are connected via links. Hence, we need to define another measurement to estimate and compare the patterns.

We will define a new measurement which we call *neighborhood distribution function*. Basically, we obtain the color distribution of the neighbors of the red and black agents. When there are the same number of red and green agents and when they have the same utilities, the neighborhood distribution of green agents will be symmetric to that of the red ones, so there will be no need to find the green agents' neighborhood distribution. To be able to distinguish the islands, not only direct neighbors (we call them distance-1 at the graphs), but also neighbors of neighbors (distance-2) and neighbors of neighbors' neighbors (distance-3) are obtained. This can be continued for further neighbors.

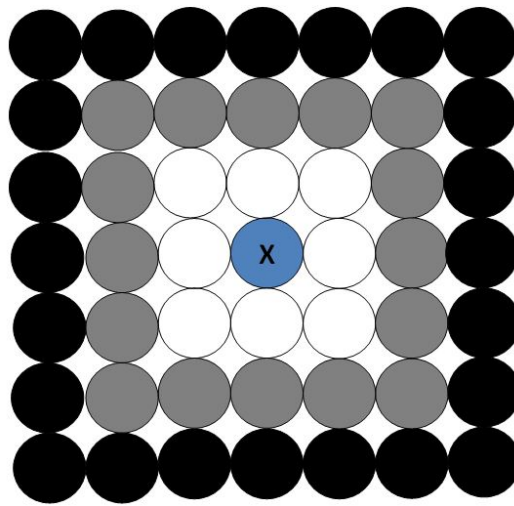


Figure 4.2: An illustration on regular lattice to explain how neighborhood distribution function works. White nodes are distance-1 neighbors of node X, grey nodes are distance-2 and black nodes are distance-3 neighbors of node X.

Let us explain how neighborhood distribution function works on a regular lattice. Consider each node is neighbor with 8 surrounding nodes and we want to find the neighborhood distribution of the node X which is in the middle of the figure 4.2. Its direct neighbors are the distance-1 neighbors, which are represented as white nodes in the figure. Its distance-2 neighbors are the nodes which are the neighbors of node X's neighbors, but not the neighbors of node X's itself. Distance-2 neighbors are represented with color grey. Node X's distance-3 neighbors are the nodes which are neighbors of neighbors' neighbors of node X and which are not already stated as distance-1 or distance-2 neighbors of node X. In figure, the distance-3 neighbors of node X are the black nodes. With mixity and FSI indices, we were focusing on direct neighbors, but now we start to consider the distribution of vacant cells, and expand our

vision to the nodes and vacancies which are in 3 steps range.

After the neighborhood distribution of a red agent is found, it will be divided into the number of neighbors at this distance for normalization, and after this is found for each red agent, the average will be taken. We will follow the similar steps for black agents as well, so that we can catch the islands of vacant cells.

4.1 Comparing Networks with the Neighborhood Distribution Function

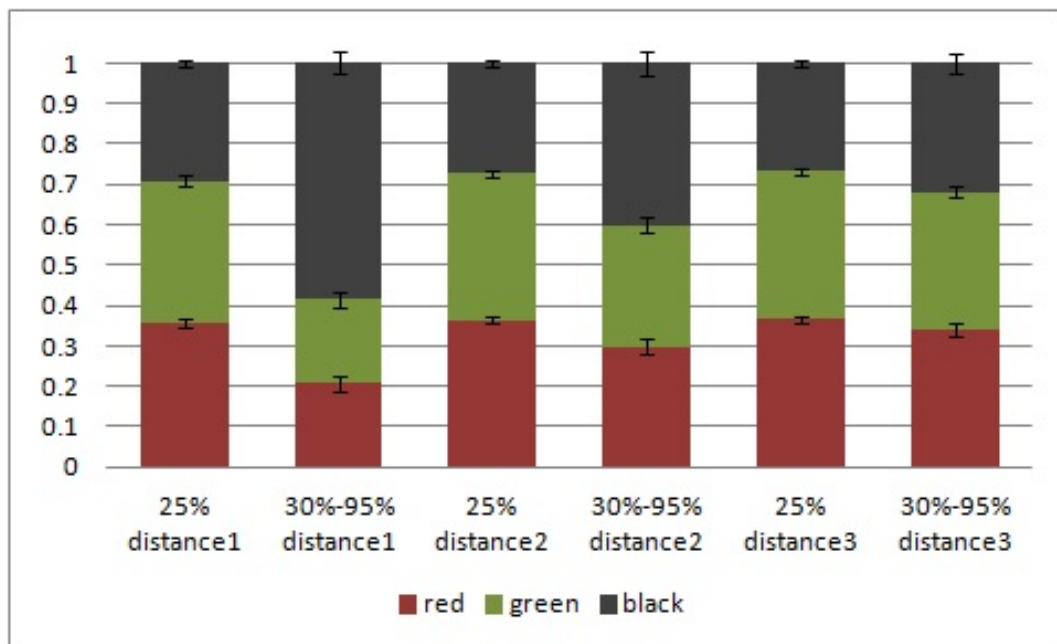


Figure 4.3: The neighborhood distributions functions of black cells of two networks with the same segregation indices. In 30%-95% setting, most of the black cells are neighbors with other black cells as distance-1 column indicates. Moreover, in all distances black cells of 30%-95% setting have more black neighbors than 25% lower bound setting. This result would make someone think 30%-95% setting has bigger islands of black cells.

Figure 4.3 shows the neighborhood distributions functions of the black cells of two networks shown in figure 4.1. One of the networks has 25% lower bound and the other one has 30% lower and 95% upper bound, but they have the same segregation indices. Distance-1 column of the neighborhood distribution function represents that in 30%-95% setting, most of the black cells are neighbors with other black cells. Moreover, in all distances black cells of 30%-95% setting have more black neighbors than 25% lower bound setting. This result would

make someone think of bigger islands of black cells at 30%-95% setting without seeing visual outcomes of the networks which was given in figure 4.1.

This finding suggests that, neighborhood distribution function might be useful for comparison of the networks which have the same segregation indices. In chapter 2.2, we saw symmetric utility functions produce the same indices and in chapter 3 we saw the ring lattice and small world networks give the same results. We can check these situations with neighborhood distribution function to see if there are differences.

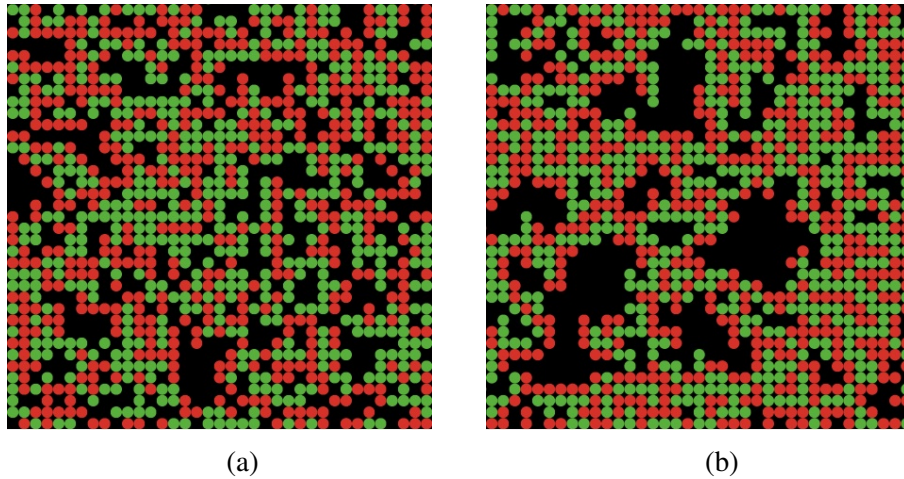


Figure 4.4: Two networks with different symmetric utility functions, but same segregation indices. Despite having the same mixity and FSI values are the same, visual outcomes are different.

Figure 4.4a belongs to the regular network with bounds 10%-90% and figure 4.4b belongs to the regular network with bounds 25%-75%. Both of the settings have 0.51 mixity and -0.02 FSI, but again we observe different outcomes. Figure 4.5 shows the neighborhood distributions functions of black cells of these networks. In 25%-75% setting, black cells are neighbors with other black cells more than they are in 10%-90% setting which indicates the bigger black islands that 25%-75% setting has.

If there are differences in the patterns of the ring lattice and small world networks, they would not be as observable as they are at the regular lattice because of these networks' layouts. At that point, neighborhood distribution function is useful to check if these kind of networks have the same patterns.

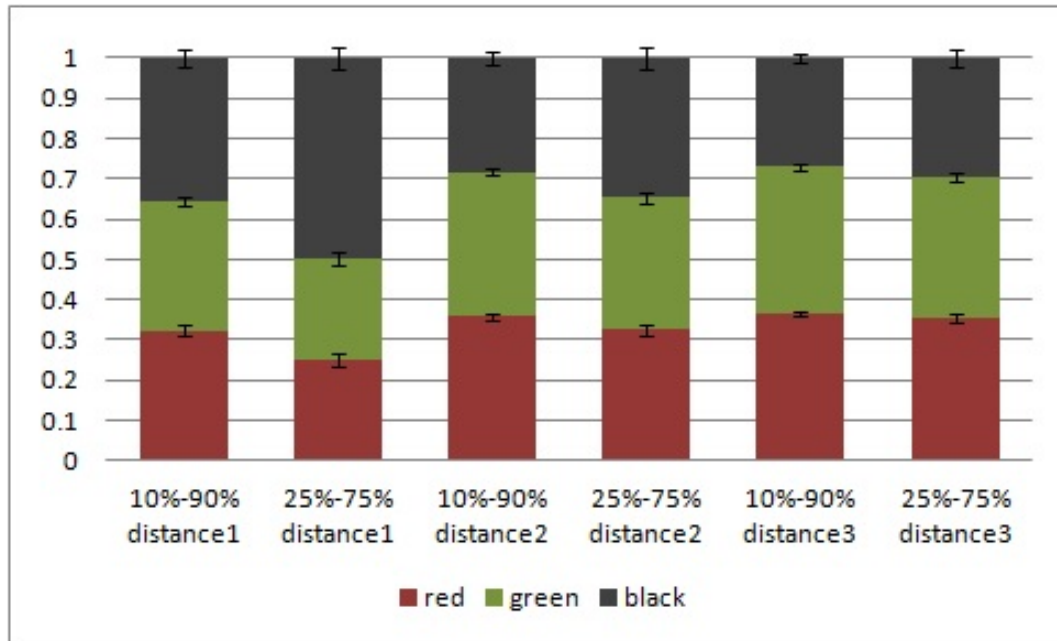


Figure 4.5: The neighborhood distribution functions of black cells of two regular networks with symmetric utility functions. In 25%-75% setting, black cells are neighbors with other black cells more than they are in 10%-90% setting which shows the bigger black islands that 25%-75% setting has.

For instance, figure 4.6 shows the neighborhood distribution functions of the red agents when the lower bound is 75%. In all distance levels, red agents of the small world network have more green agents in their neighborhood which means the ring lattice should have larger communities of red agents and as green agents would give a symmetric result, the ring lattice has larger communities of greens as well.

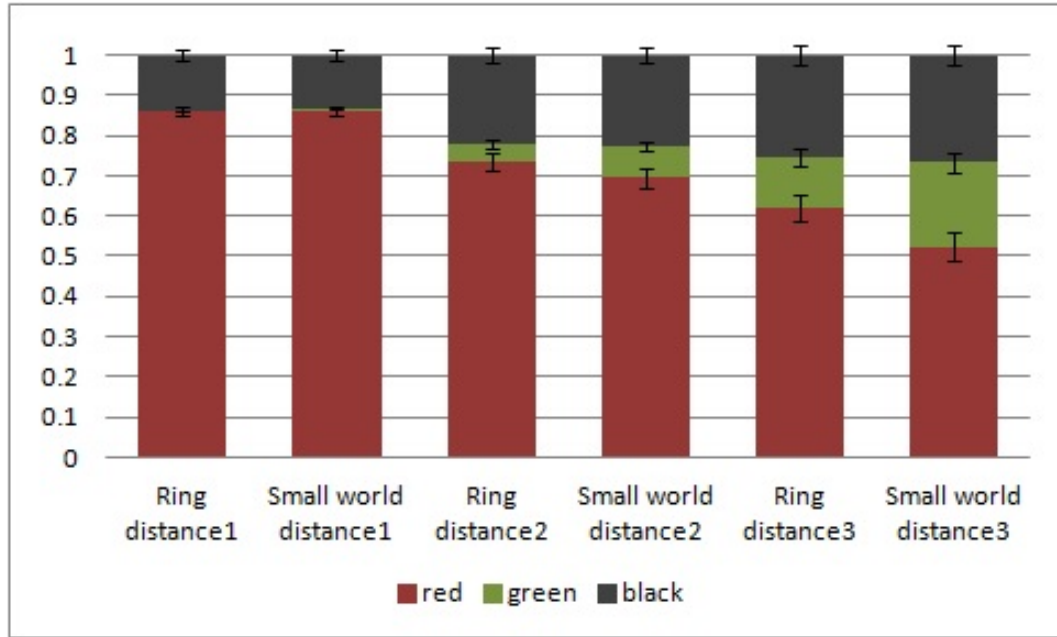


Figure 4.6: The neighborhood distribution functions of the red agents of the ring lattice and small world networks for 75% lower bound. In all distance levels, red agents of the small world network have more green agents in their neighborhood. We can conclude that, there are larger communities of red agents at the ring lattice.

4.2 Neighborhood Motifs

If we observe differences in network level by neighborhood distribution function, we should be able to observe differences in agents' neighborhood motifs, so in this section, we will focus on the motifs.

In both of the figures 4.7a and 4.7b, the agents in the middle have 3 similar neighbors, but their motifs are different. Lets name the motif in figure 4.7a motif-1 and motif in figure 4.7b motif-2. One interpretation can be this; at motif-1 similar neighbors of the agent in the middle know each other which indicates triadic closures, and in the literature, triadic closure is associated with building trust among triadic closure members [1]. Whereas at motif-2 none of the similar neighbors of the agent in the middle know each other. Moreover, clustering coefficient of the agent in the middle of motif-1 is 1, while it is 0 for the agent in the middle of motif-2, and higher clustering coefficients are the signs for tightly knit communities.

If we call the motif in figure 4.7c motif-3 and the motif in figure 4.7d motif-4, we can make

similar interpretations for motif-3 and motif-4. In both of them the agent in the middle has 4 similar neighbors, but in motif-3 clustering coefficient of the agent is high ($2/3$) whereas it is 0 at motif-4.

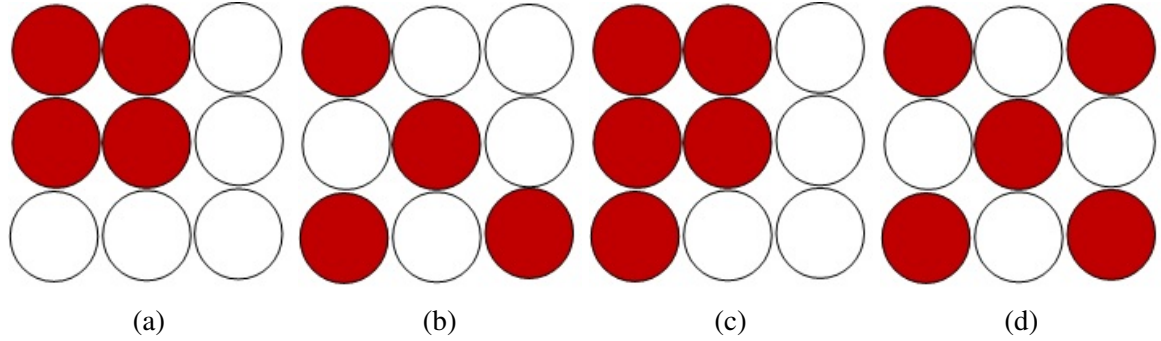


Figure 4.7: Nodes with different motifs. In (a) and (b) both of the agents in the middle have 3 similar neighbors, but in former case these 3 neighbors know each other which builds trust among them and in the latter case, they do not know each other. There is a similar situation in (c) and (d); in both of them the agent in the middle has 4 similar neighbors, but the motifs are different.

Apparently, because of the symmetry, there are 4 versions of motif-1. If we put upper left agent of motif-2 to upper middle, 3 neighbors of the agent in the middle will continue to not to know each other, so there are 8 symmetric versions of motif-2. In the other orientations of 3 similar neighbors, either 1 or 2 couple of them know each other. We consider motif-1 and motif-2 as they are more significant with having clustering coefficient 1 and 0 respectively. There are 8 symmetric versions of motif-3 too, but motif-4 is unique. At our calculations we consider all versions of these motifs.

Table 4.1: The distribution of the motifs in two networks with the same segregation indices. The network with a 30%-95% lower-upper bound couple has more motif-1 and motif-3 and less motif-2 and motif-4 than the network with 25% lower bound. Thus, we can conclude that similar neighbor communities are more tightly knit in the 30%-95% network.

Bounds	Similar neighbors:3	Motif-1	Motif-2	Similar neighbors:4	Motif-3	Motif-4
25%	271.48	26.48	18.74	263.3	31.06	1.84
30%-95%	250.24	42.82	2.16	325.7	70.2	0.08

Table 4.1 shows the distribution of the motifs of the networks with 25% lower bound and 30%-95% lower-upper bound couple. The network with 30%-95% bounds has more motif-1

Table 4.2: The clustering coefficient of different networks. Left column shows the average clustering coefficient of the agents, right column shows the average clustering coefficient of the agents with their similar neighbors. Results proof that the agents in the network with 30%-95% bounds have more tightly-knit communities than the agents in the network with 25% lower bound.

Bounds	Clustering coefficient	Clustering coefficient with similar neighbors
25%	0.43	0.48
30%-95%	0.46	0.54
10%-90%	0.44	0.42
25%-75%	0.47	0.41
Small world 75%	0.66	0.67
Ring 75%	0.69	0.70

and and motif-3 and less motif-2 and motif-4 than the network with 25% lower bound. Thus, we can conclude that similar neighbor communities are more tightly knit in the 30%-95% network. If we look at figure 4.1 again, we can notice motif-1 and motif-3s are the motifs that form carpet pattern like paths at this network.

If the network with 30%-95% bounds has more motif-1 and motif-3 than the network with 25% lower bound has, then its average clustering coefficient expected to be higher. Table 4.2 shows the average clustering coefficient and the average clustering coefficient with the similar neighbors of different networks. Results proof that agents in the network with 30%-95% bounds have more tightly-knit communities than the agents in the network with 25% lower bound.

Dense clusters might be barriers for the diffusion of behaviors and innovations, because people adopt a behavior or an innovation if some of their friends adopt it and if a person's friends are in the same community, nobody can be an initial adopter, so it is hard to diffuse to agents in a tightly knit cluster [1]. From that point of view, one might conclude that diffusion would be easier in the network with 25% lower bound than the network with 30%-95% bounds. On the other hand, the situation is the other way around for collective actions. Because at collective actions, such as demonstrations, people consider others' thoughts about the action, so dense clusters support collective actions [1]. One might expect a larger collective action from the network with 30%-95% bounds than the network with 25% lower bound.

The distribution of the motifs of the networks with 10%-90% and 25%-75% bounds are not different, but neighborhood distribution function showed bigger islands of vacant cells at 25%-75% case. If vacant cells are together, other agents should also come together. Table 4.2 shows this is true as 25%-75% situation have a higher clustering coefficient than 10%-90% situation has. It is not possible to observe motifs in the ring lattice and small world networks, but the clustering coefficient of the ring lattice is higher than the small world's.

As a result, different utility functions may cause agents to have different motifs and clustering coefficients, so by changing the tolerance and desire for diversity of the agents, not only segregation level but also trustworthiness and reactions to social dynamics, such as diffusion and collective action, might change.

To sum up, there might be networks with the same mixity and FSI values, but different patterns. Thus, traditional segregation indices are not sufficient to catch differences. Neighborhood distribution function let us observe islands of one type of agents and islands of vacancies. This can be useful especially when studied network is not a regular network and relationships are defined by links rather than location. Moreover, different utility functions make agents form different motifs with their neighbors and these motifs are the reasons that lie behind the different network patterns. Lastly, different motifs may cause changes in social dynamics other than segregation, so policy makers should be aware of them.

5 Conclusions and Future Research

In this thesis, first of all we investigated Schelling's segregation model for variety of utility functions. We showed that existence of desire for diversity, which means addition of an upper bound to simulations, decreases segregation. Moreover, simulation results indicate that if utility functions are symmetric with respect to 50% similar neighbors, integration is obtained. However, when utility function is not symmetric and favors similarity, even if having 50% similar neighbors results in the highest utility, segregation emerges. We also tested the asymmetric utility functions which favor dissimilarity and saw that inverse homophily or integration might emerge, but obtaining an inverse homophilic network is not as easy as obtaining a segregated network.

We further considered the situations where two types of agents are not distributed evenly. 20% of the agents were one type and 80% of the agents were the other type, so there were minority and majority groups in the network. In that case, increasing the tolerance by decreasing the lower bound decreases the segregation level. Moreover, if there will be an upper bound, minority members should have it, otherwise a large number of majority members surround minorities and make minorities unhappy with their places.

Later, we compared the segregation level of the regular ring lattice and small world networks for different lower bounds and utility functions. Resulting mixity and FSI indices of these networks were overlapping, so we concluded that these networks do not segregate differently. Functions which are symmetric with respect to 50% similar neighbors produce integration, asymmetric functions which favor similarity result in segregation and asymmetric functions which favor dissimilarity produce inverse homophily or integration like regular lattice.

Lastly, we defined a new measure named neighborhood distribution function. This measurement not only considers nodes' direct neighbors, but also neighbors which are reachable in 2 and 3 steps, so it allows us to follow the islands in networks. We showed that there might be two networks with the same segregation indices, but different patterns, and neighborhood distribution function is able to catch them. This measurement might be useful especially when network patterns are not observable because of the layout of the network. For instance, with neighborhood distribution function we presented that the ring lattice and small world network with the same utility function and same segregation indices might have different patterns. Moreover, we investigated the motifs that agents form with their neighbors, and saw that distribution of the motifs change from network to network and these motifs generate different patterns at network level. Different motifs have different social structures and clustering coefficients which might affect the social dynamics like trustworthiness, diffusion and collective action.

Main purpose of the study was to understand the effects of individual desires on the segregation of the networked society. Simulation results suggest that higher tolerance level and desire for diversity decrease segregation. Simulating different utility functions result in segregation, integration or inverse homophily, so segregation does not have to be the ultimate result. After observing networks with the same traditional segregation indices but different patterns, we defined neighborhood distribution function and become able to distinguish the patterns. Lastly, when we investigated the neighborhood motifs of the agents, we saw differences among different utilities. Thus, the change in individual desires might affect not only the segregation level, but also other social dynamics.

Throughout the study, moving cost is accepted as zero. So when there was just a lower bound, any unhappy agent could move to any vacant cell. And when agents have a utility function, they were able to move to any vacant place which gives the highest utility. However, for more realistic simulations cost of moving can be taken into consideration. We assign one unit cost to each movement, regardless of the distance, assuming the fixed cost of moving is high. Then the total cost would be the sum of the costs that each agent faced while changing its place. Figure 5.1 shows the cost distribution of the agents for different settings. As we showed in the

previous sections, 25% lower bound produces slightly segregated environment and 25%-75% lower-upper bound couple results in integration. One might think that integration is easier to obtain than segregation, because agents are distributed randomly at the beginning of the simulation and a random distribution is closer to be integrated than segregated. However, the figure shows that it costs more to have integration. Thus, possible costs of different utility functions might be a subject for future research.

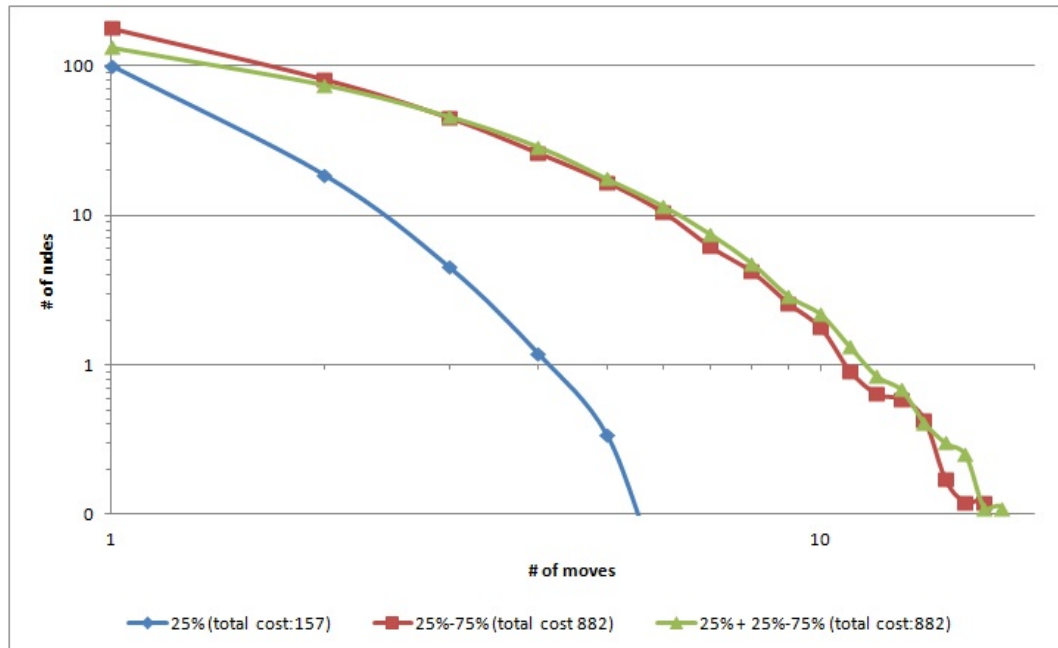


Figure 5.1: Distribution of moving costs for different settings. In all of the three settings most of the nodes find a suitable place in 1 or 2 movements, but there are some nodes which change their place several times. Moreover, 25%-75% setting is much more costly than 25% lower bound setting. However, the addition of a 75% upper bound to a network that is in equilibrium state with 25% lower bound does not result in a different cost than the initially random 25%-75% setting. Costs of different utility functions and networks might be interesting to analyse.

References

- [1] D. Easley and J. Kleinberg, *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press, 2010.
- [2] J. Moody, “Race, school integration, and friendship segregation in america,” *American journal of Sociology*, vol. 107, no. 3, pp. 679–716, 2001.
- [3] P. V. Marsden, “Core discussion networks of americans,” *American sociological review*, pp. 122–131, 1987.
- [4] W. Shrum, N. H. Cheek Jr, and S. MacD, “Friendship in school: Gender and racial homophily,” *Sociology of Education*, pp. 227–239, 1988.
- [5] C. Baerveldt, M. A. Van Duijn, L. Vermeij, and D. A. Van Hemert, “Ethnic boundaries and personal choice. assessing the influence of individual inclinations to choose intra-ethnic relationships on pupils’ networks,” *Social Networks*, vol. 26, no. 1, pp. 55–74, 2004.
- [6] J. Suitor and S. Keeton, “Once a friend, always a friend? effects of homophily on women’s support networks across a decade,” *Social Networks*, vol. 19, no. 1, pp. 51–62, 1997.
- [7] T. C. Schelling, “Dynamic models of segregation,” *Journal of mathematical sociology*, vol. 1, no. 2, pp. 143–186, 1971.
- [8] T. C. Schelling, *Micromotives and Macrobehavior*. Norton, 1978.
- [9] U. Wilensky, “Netlogo itself,” *Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston*, <http://ccl.northwestern.edu/netlogo>, 1999.
- [10] R. Pans and N. J. Vriend, “Schelling’s spatial proximity model of segregation revisited,” *Journal of Public Economics*, vol. 91, no. 1, pp. 1–24, 2007.
- [11] M. M. Möbius and T. S. Rosenblat, “The process of ghetto formation: Evidence from chicago,” *Unpublished manuscript*.

- <http://www.tanyarosenblat.org/sites/default/files/tanyarosenblat/files/chicago-dec7.pdf>, 2001.
- [12] A. Banos, “Network effects in schelling’s model of segregation: new evidences from agent-based simulation,” *Environment and Planning B: Planning and Design*, vol. 39, no. 2, pp. 393–405, 2012.
- [13] P. F. Lazarsfeld, R. K. Merton, *et al.*, “Friendship as a social process: A substantive and methodological analysis,” *Freedom and control in modern society*, vol. 18, no. 1, pp. 18–66, 1954.
- [14] L. C. Freeman, “Segregation in social networks,” *Sociological Methods & Research*, vol. 6, no. 4, pp. 411–429, 1978.
- [15] G. Fagiolo, M. Valente, and N. J. Vriend, “Segregation in networks,” *Journal of Economic Behavior & Organization*, vol. 64, no. 3, pp. 316–336, 2007.
- [16] J. Zhang, “Residential segregation in an all-integrationist world,” *Journal of Economic Behavior & Organization*, vol. 54, no. 4, pp. 533–550, 2004.
- [17] R. Farley, H. Schuman, S. Bianchi, D. Colasanto, and S. Hatchett, “Chocolate city, vanilla suburbs: Will the trend toward racially separate communities continue?,” *Social Science Research*, vol. 7, no. 4, pp. 319–344, 1978.
- [18] P. Sturgis, I. Brunton-Smith, J. Kuha, and J. Jackson, “Ethnic diversity, segregation and the social cohesion of neighbourhoods in london,” *Ethnic and Racial Studies*, vol. 37, no. 8, pp. 1286–1309, 2014.
- [19] PODEM, *Personal Interview on the Manuscript of Youth Study Report*. <http://podem.org.tr/en/topics/youth-studies/>, 2015.
- [20] S. Grauwin, F. Goffette-Nagot, and P. Jensen, “Dynamic models of residential segregation: An analytical solution,” *Journal of Public Economics*, vol. 96, no. 1, pp. 124–141, 2012.
- [21] CIA, *The world factbook*. <https://www.cia.gov/library/publications/the-world-factbook>, 2010.

- [22] Konda Research and Consultancy, “Toplumsal yapı araştırması - biz kimiz,” <http://www.konda.com.tr/tr/raporlar.php>, 2006.
- [23] S. L. Feld, “The focused organization of social ties,” *American journal of sociology*, pp. 1015–1035, 1981.
- [24] D. J. Watts and S. H. Strogatz, “Collective dynamics of small-world networks,” *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

6 Appendices

6.1 Appendix A: The Effect of Percentage of Vacant Cells

Table 6.1: Number of time steps until equilibrium for different percentages of vacant cells and different lower bounds and lower-upper bound couples.

Percentage of vacant cells	Number of time steps until equilibrium				
	25%	50%	75%	10%-90%	25%-75%
10	6.2	17.3	316.9	3.3	33.3
20	6.2	16.0	184.5	4.2	31.4
27	6.3	14.0	116.1	4.9	34.9
30	6.7	14.5	97.1	5.1	37.7
40	6.6	13.6	65.2	7.3	48.7
50	6.2	12.5	54.2	10.5	62.2

Table 6.1 shows when percentage of vacant cells is low, equilibrium takes more time with increasing lower bound. However, when there is also an upper bound, settings with too many vacant cells start to give results in longer time. Thus, a mild percentage of vacant cells seems convenient and 27% is used along the study.

6.2 Appendix B: Results of Different Movement Algorithms

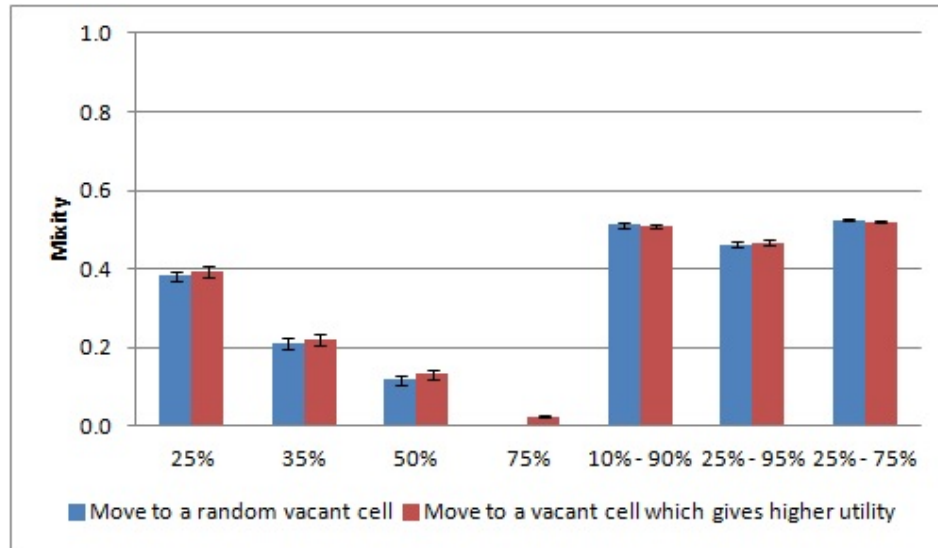


Figure 6.1: Results of different movement algorithms. The figure shows mixity values of a regular lattice with different lower-upper bound values and two different moving algorithms. At first algorithm, unhappy agents move to one of the vacant places, whereas at second algorithm they move to one of the vacant cells which gives a better utility. Resulting mixity values are not different, so when step function is used and there is not a comparison with other utility functions, first algorithm is used. Because the second algorithm takes more time as agents are finding the potentials of all vacant cells and compare it with their current utility value before moving.

6.3 Appendix C: Results of Symmetric Utility Functions for Regular Lattice with Degree 12

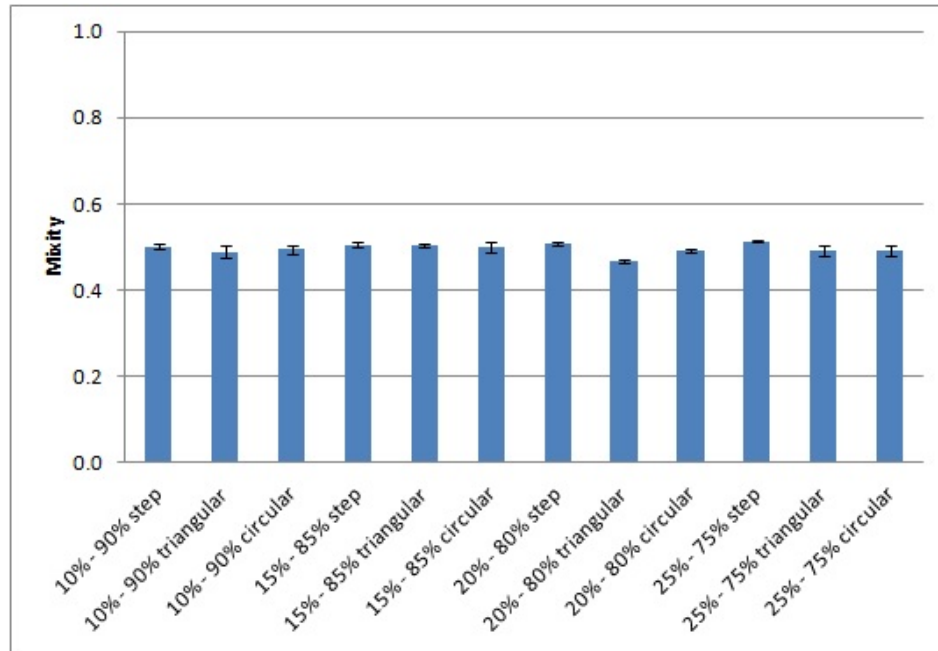
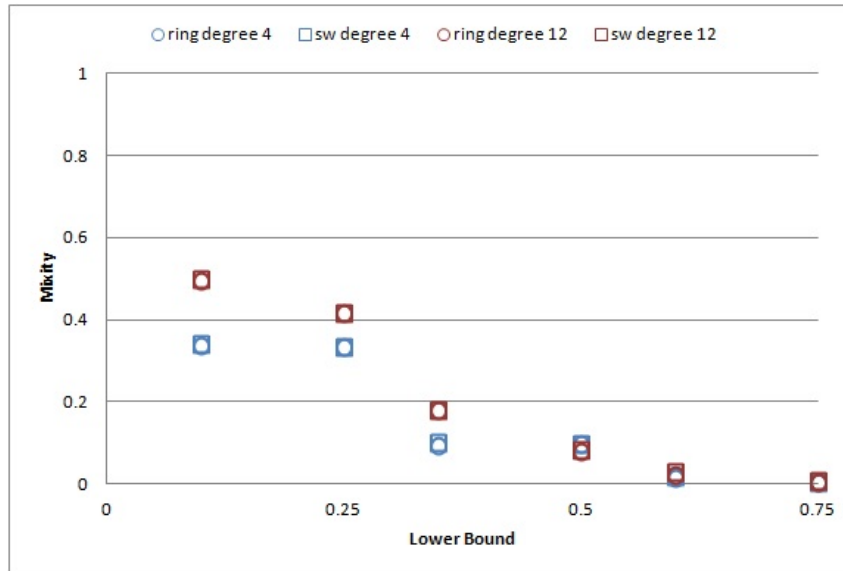
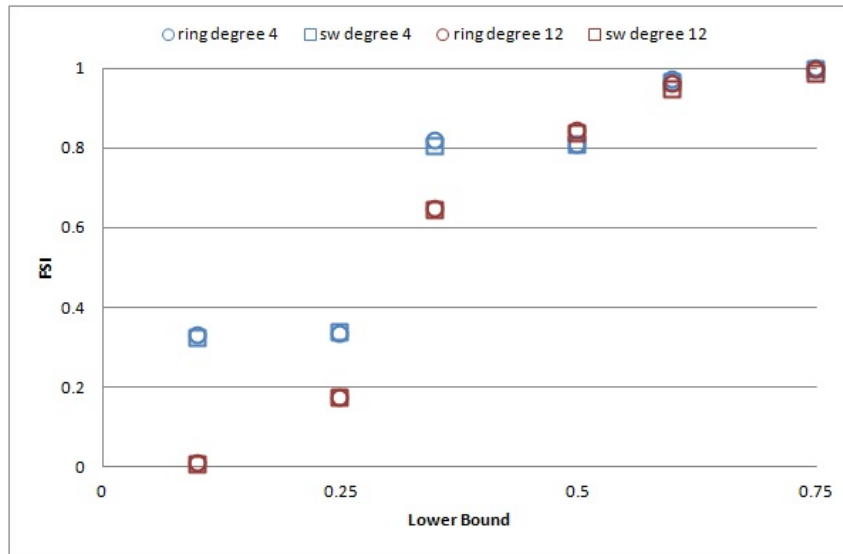


Figure 6.2: Mixture values of the symmetric utility functions for regular lattice with degree 12. Symmetric utility functions continue to give integration when degree is changed as 12. Outcomes are even smoother than the ones obtained with degree 8, because when the degree is 12 there are more options to supply given utility functions.

6.4 Appendix D: Comparison of the Ring Lattice and Small World Networks for Different Average Degrees



(a)



(b)

Figure 6.3: Mixity and FSI values of the ring lattice and small world network with degree 4 and 12 for different lower bounds.

For different node degrees ring lattice and small world networks still give the same indices. Also notice that when the degree is 4, different lower bounds can give the same results because of the limited number of options that an agent with 4 neighbors has.