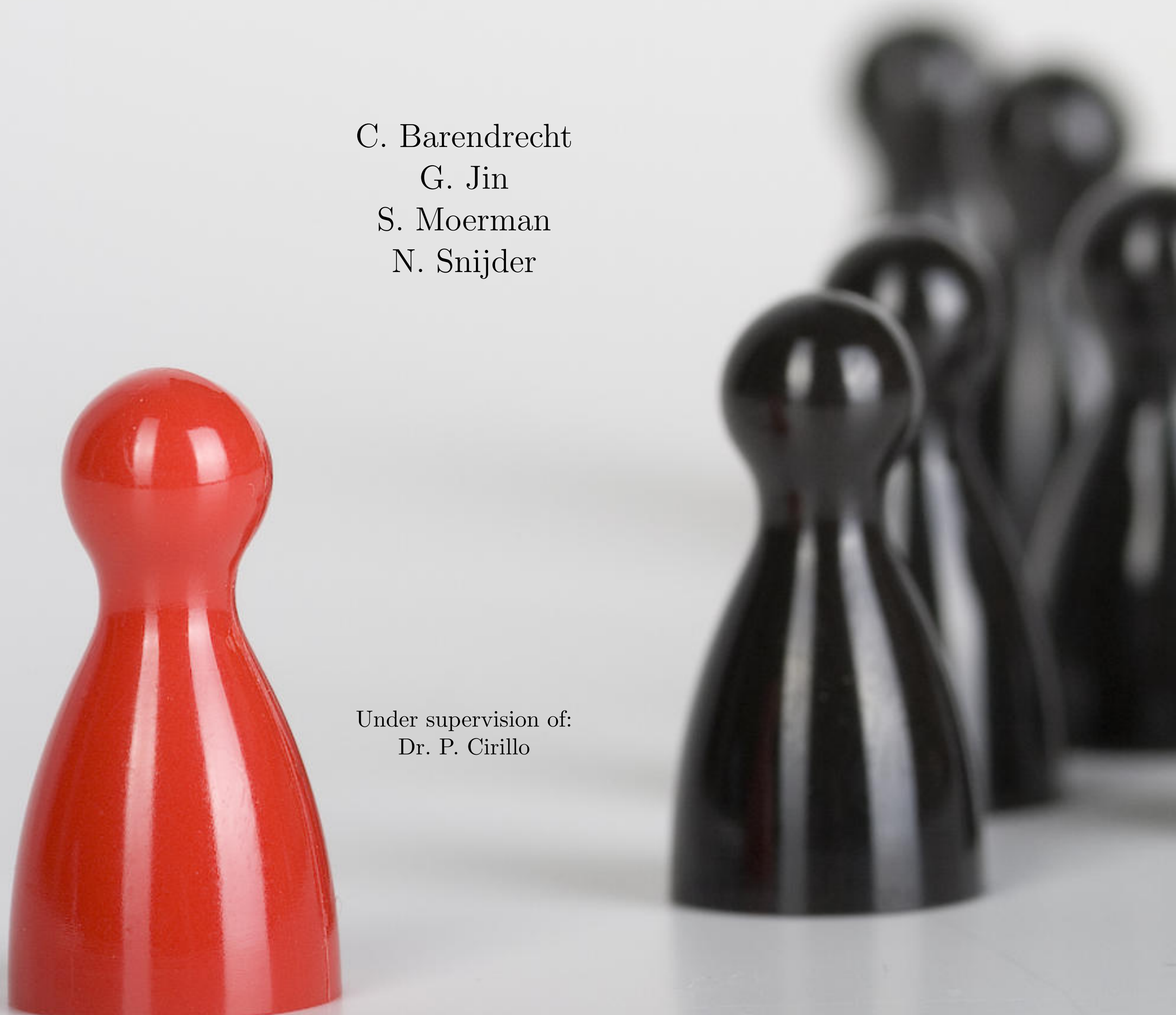
Technische Universiteit Delft

Modelleren 2A

Segregation

C. Barendrecht
G. Jin
S. Moerman
N. Snijder

Under supervision of:
Dr. P. Cirillo



This page is left intentionally blank.

Abstract

An extended version of the 'Shelling Tipping Model' is built...

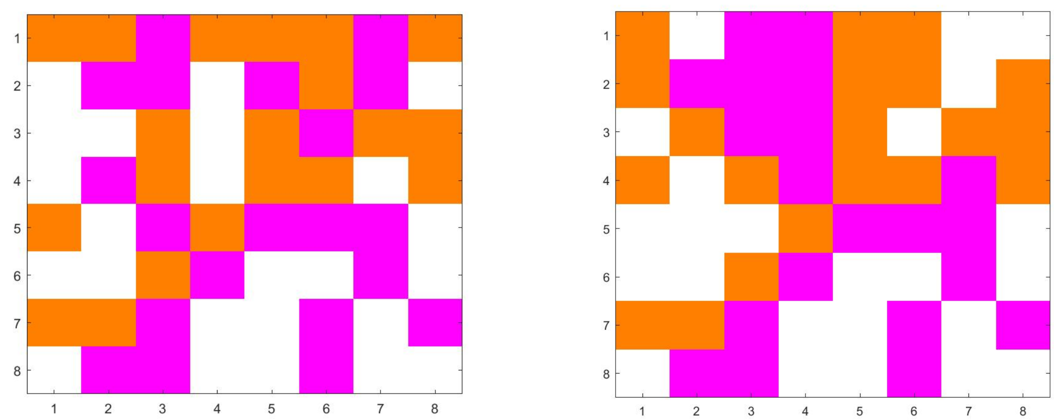
Casper Barendrecht, Guanyu Jin, Stijn Moerman en Nand Snijder. Delft, April 2017

Contents

1 Introduction

In 1978, Thomas C. Schelling developed his tipping model by placing pennies and dimes on a chess board and moved them according to various rules. By viewing the pennies and dimes as two types of people, the rule of moving as a preference for the individuals, and the chess board as a city, he soon discovered that segregation takes place on the board, even when the preference of an individual is very subtle.

Based on this idea, we created a basic model which consists of an 8×8 board with 40 individuals that are divided evenly into two types. The individuals are moved according to their 'Happiness' in the current place. For the basic model, an individual is considered happy if $\frac{1}{3}$ of his/her second order neighbours (a person has at most 8 neighbours) shares his type. Otherwise, an individual is considered unhappy and will move to the nearest place where his/her happiness is strictly higher. This will be referred to as the 'Happiness Rule'. An example of how this model goes in effect is given below. It is easily observed that not every individual in figure 1a is happy. After segregation the individuals have relocated themselves as in figure 1b, there every individual is happy.



(a) Situation before segragation

(b) Situation after segregation

Figure 1: An example of segregation in our basic model: type 1 is orange and type 2 is pink

After that, we extended the basic model by changing the parameters such as the size of the population, the boardsize and the number of types. We also included an option for random displacement: an individual which is not happy will be moved to a empty location chosen at random. (That is to say, that individual will be placed to any empty spot with equal probability.)

For both the basic and the extended model, we ran 500 simulations multiple times and investigated to what extent certain parameters affected the segregation pattern. In order to formulate our research goals precisely, the following definitions are introduced:
1. **Generation**: A population is said to have entered the next generation if the happiness of every individual has been checked once.
2. **Equilibrium after g generations**: The population is said to have reached an equilibrium after g generations if no individual has moved during the $g+1$-th generation.
3. **Homogeneity**: A person is said to live homogenous if all of his/her neighbours share his type.
4. **Segregation time at $n\%$**: The segregation time at $n\%$ is defined as the number of generations after which $n\%$ of the population is homogenous.

For this project, we focussed on the following main questions:

1. How do the parameters affect the equilibrium? Does the population always reach an equilibrium? How many generations on average does it take to reach an equilibrium? What's the probability distribution of the number of generations to reach an equilibrium?

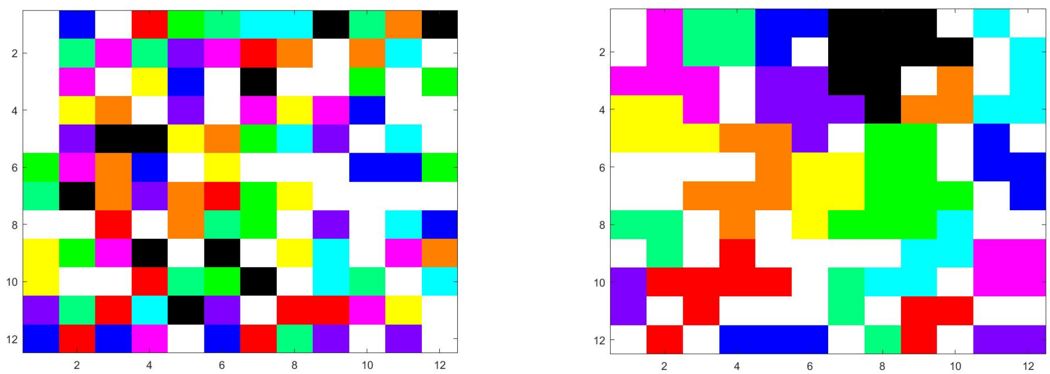
2. What fraction of the individuals is happy after the equilibrium? Can this be optimised by changing the boardsize?

3. What is the average segregation time at 60%, 80% and 100%. How is this segreggation time distributed with Happiness Rule 1 and how does the Happiness Rule affect the average segregation time?

2 Our Model

As mentioned in the introduction, we will mostly be investigating the basic model: segregation of 20 individuals of type 1 and 20 individuals of type 2 on an 8×8 board. In this model an individual is considered unhappy if less than $\frac{1}{3}$ of his/her neighbours shares his/her type. If an individual is unhappy, this individual will move to the nearest place where he/she would be happier. Any person without neighbours is by definition considered unhappy (so his/her happiness is 0), since people prefer living with others over living alone. Furthermore, the basic model uses the second order neighbourhood; the maximum number of neighbours a person can have is 8. The model however can be changed by preference in several ways.

First of all, the boardsize can be varied. If no other parameters are altered, this will only imply that the board becomes fuller or emptier. Secondly, we can change the number of types on the board. For instance, one can consider a 12×12 board with 10 types of individuals and 10 individuals per type: see figure 2. As one can see, far more individuals have been moved to another place on the board and segregation takes place.

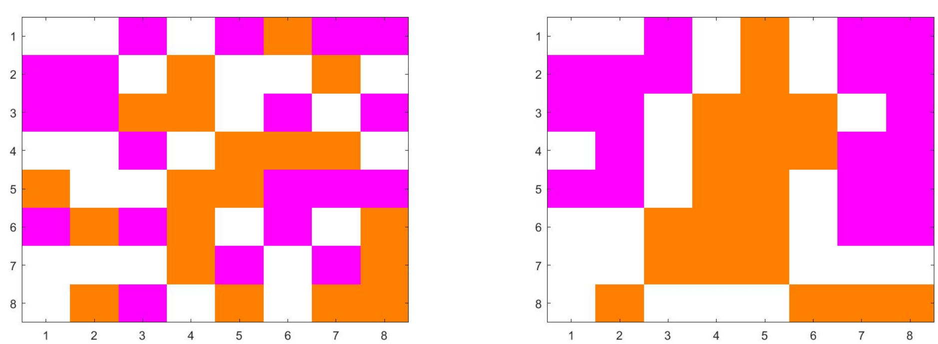


(a) Situation before segregation

(b) Situation after segregation

Figure 2: An example of segregation in a model on a 12×12 board with 100 individual of 10 different types

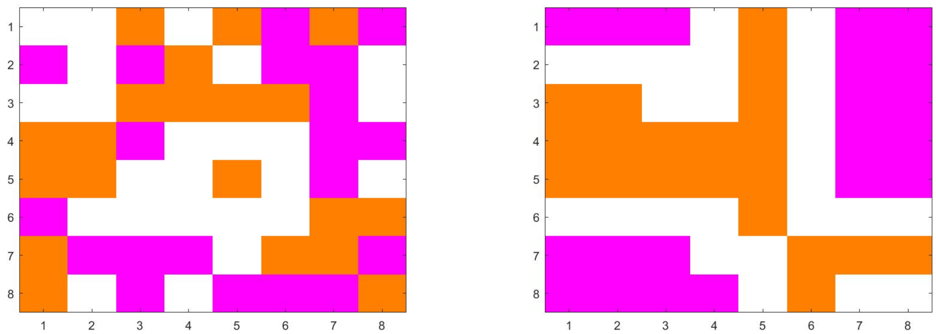
Subsesquently, we can vary our 'Happiness Rule': the minimal fraction of neighbours that has to share an individuals type, in order to be considered happy. The expectation is that the time to reach equilibrium increases as the Happiness Rule is strengthened. This makes sense as every individual needs to be surrounded by relatively more neighbours of his/her own type. In consequence, larger homogenous groups will be formed after segregation with a higher Happiness Rule. In figure 3 and 4 the effect of a Happiness Rule of respectively $\frac{2}{3}$ and 1 on the board after segregation is illustrated.



(a) Situation before segregation

(b) Situation after segregation

Figure 3: An example of segregation in a model with happinessrule $\frac{2}{3}$



(a) Situation before segregation (b) Situation after segregation

Figure 4: An example of segregation in a model with happinessrule 1

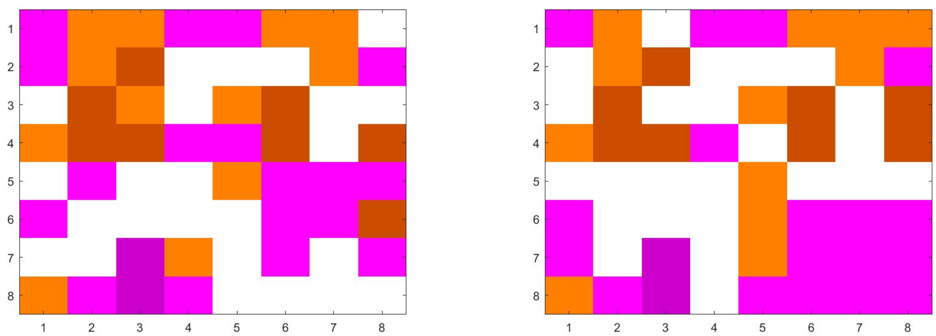
5	4	3	4	5
4	2	1	2	4
3	1	0	1	3
4	2	1	2	4
5	4	3	4	5
8	7	6	7	8

Figure 5: Neighbourhoodorder

In addition, the type of neighbourhood can be changed. The neighbourhood is by default a second order neighbourhood, in which both the adjacent and diagonal people are considered neighbours. There exists only one smaller type of neighbourhood: the first order neighbourhood, in which only adjacent neighbours are accounted for: A person has up to four neighbours. The way the different types of neighbourhood are defined is illustrated in figure 5. Here 0 represents the individual himself and every other place displays which order of of neighbourhood it belongs to. We will not research the effect of the type of neighbourhood on the segregation in our model in this report. However, some interesting graphs are included in the appendix.

Moreover, we included the possibility for individuals to move to a random place on the board if he/she is not happy, rather than moving to the closest. In the basic model he/she will move to the nearest place where he/she will become happier, With random displacement, we expect that it takes longer to reach an equilibrium. Again, we will not discuss the impact of the random displacement in our report, but we added some nice results in our appendix.

Furthermore, we expanded the model with the so-called 'criminals'. One can choose to have c out of the n individuals behave as if they were a criminal. Any person who is not a criminal himself, does not like criminals in his neighbourhood, and thus his/her happiness will be 0, overruling any positive effect of friendly neighbours, if a criminal lives in his/her neighbourhood. As a result, all criminals should end up living together. In figure 6, one can see an example of the basic model with 10 criminals among the 40 individuals. It is visible that our expectation is not always the outcome. In this case, equilibrium is reached, because every unhappy individual cannot move to another place, since every empty location is in the neighbourhood of a criminal. So the unhappy individuals will not become happier on these places.



(a) Situation before segregation

(b) Situation after segregation

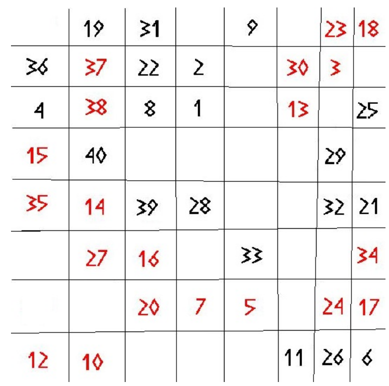
Figure 6: An example of segregation in the basic model with 10 criminals (the slightly darker places)

Last but not least, we made it possible for individuals to switch to another type. In our model there are two types of switching. The first is only possible if the only individuals are of type 1 and type 2. In this version of switching, every person has in every generation a chance p of switching the other type. In the other version, every unhappy individual will change type under the influence of his/her neighbours. This means that the chance to become of type t is the fraction of neighbours of type t. This is a correct probability distribution, because this sums up to 1. Individuals without neighbours will not switch type, because they have no neighbours who influence them. We will investigate the influence of the second type of switching further in section 7.

3 Equilibrium

As mentioned in the introduction, a board has reached quilibrium after g generations if no person has moved in the $g+1$-th generation. In this parapgraph a portion of the first research question will be answered. Namely, will a board always reach an equilibrium in finite time? Based on intuition, this is expected to be true. An individual who moves, will move to a place where relatively more neighbours share his/her type, intuitively, this should imply that the average happiness of the population increases during each generation. In the basic model this is not always true however. Even though equilibrium is nearly alays reached in finite time, there exists cases in which allow a periodic solution. One example of any such scenario is given below:

(a) The start situation. Individual 37 will move.



(b) Individual 37 is moved to the nearest location that better meets his/her desires.

(c) Accordingly, 38 will move next and is moved to the nearest location that better meets his/her desires.

(d) A generation has passed, 37 is next to move and will do so accordingly. The roles of 37 and 38 have interchanged, which leads to a periodic solution.

Figure 7: A possible configuration of the standard board in which the equilibrium will never be reached. The numbers indicate the turn order of an individuals. The two different colour indicates the two types. The equilibrium is not reached due to the periodic movements of individuals 37 and 38. The movements of inviduals 37 and 38 are tracked and shown in subfigure a to d in a chronological order.

In figure 7, a possible configuration in which equilibrium will not be reached is shown. The numbers represent the turn order (1 is selected first followed by 2, etc.). Red and black indicate the two different types. It is easily observed that that individuals 1 to 36 satisfy the happiness condition. 37 however does not. During 37's turn, we note that his/her happiness equals 0. The closest empty location has a happiness of $\frac{1}{7} > 0$. And thus 37 will move to the given location, which is indicated in figure 7b. Next in line to move is 38. Individual 38 has a happiness of $\frac{2}{7}$, which is less than the required $\frac{1}{3}$ and will thus have to move.

The closest spot with greater happiness is the nearby corner spot with happiness $\frac{1}{3}$. So 38 will move to that location, which is shown in figure 7c.

The indviduals 1 to 36 will remain pleased and will thus remain in place. Thus 37 is the first candidate to be moved. 37 has a happiness $\frac{1}{7} < \frac{1}{3}$. The closest empty spot has happiness $\frac{1}{6} > \frac{1}{7}$, thus 37 will move to this location. Since this location was te former spot of 38, we conclude that 37 and 38 have swapped positions, as is shown in figure 7d.
Since 37 and 38 are of the same type, these 3 moves will continue periodically and thus we find a periodic solution.

Now what went wrong?
After the first move, both 37 and 38 will gain happiness, but as 37 moves away, 38 will lose all of his/her happiness. An endless loop is formed. Note that, the same scenario can be constructed for larger boards. Since we can implement this exact board in a larger board and fill out the remaining locations to satisfy an equilibrium.

4 Average number of generations until equilibrium

There are numerous parameters that may affect the average number of generations until equilibrium is reached. We chose to focus on the effect of the Happiness Rule(HR) on the average number of generations. For each HR ranging from 0 to 1 with increment 0.01, we ran 500 simulations and calculated the average, maximal and minimal generations it took to reach equilibrium. The results are shown in figure 8.

From figure 8, we conclude that the average number of generations (blue graph) is increasing with the Happiness Rule, which is to be expected since a higher HR implies a higher need for neighbours of the same type, and thus a lower probability that a selected individual is happy, making it more likely that he/she will move. We also note that starting at a HR of aproximately 0.7, the average number of generations appears to be nearly constant. This can be explained noting that the required number of neighbours will practically be the same. For example, if the HR were 0.8, a person with 3, 5 or 8 neighbours would require respectively 3, 5 or at least 7 of the same type. This requirement hardly changes if a HR of 0.9 is applied. The same argument explains why the average number of generations is constant at very low HR or if we raise the HR with a sufficiently small amount like 0.01.

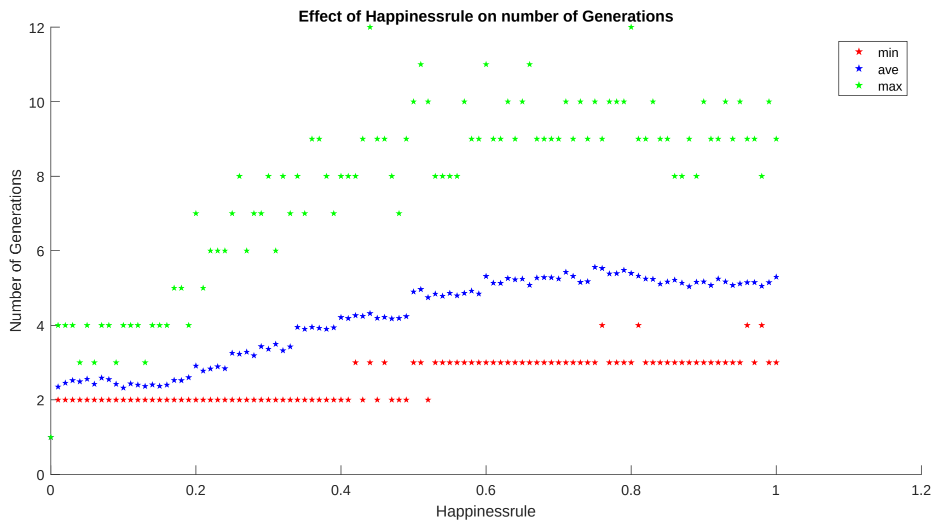
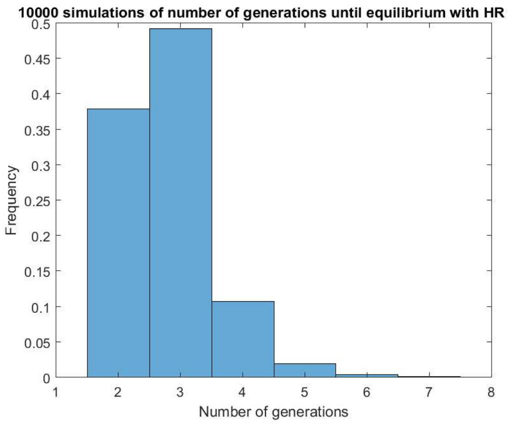
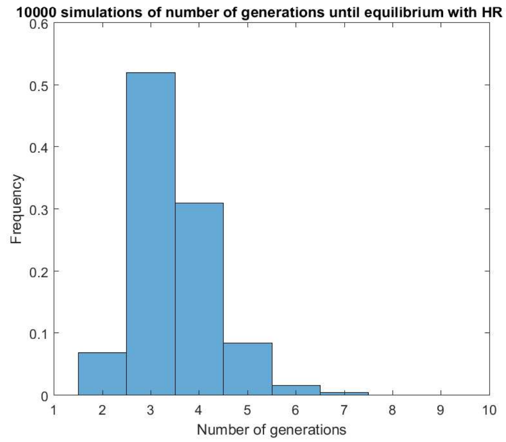


Figure 8: Effect of the happiness rule on the number of generations until reaching equilibrium. The green graph represents the maximal number of generations, the blue graph the average number of generations and the red graph displays the minimal number of generations

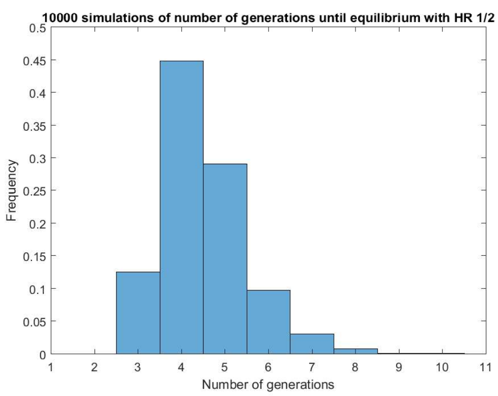
It is also interesting to investigate how the random variables Y_j, which we denote as the number of generations it takes to reach the equilibrium under HR j, is distributed, and how their distributions are effected by the HR. For this purpose, we ran 10000 simulations for with Happiness Rules of 1/4,1/3,1/2 and 1 and plotted several histograms (figure 9). We chose bin size 1 for each histogram given that Y_j is a random variable that only takes integer values. Thus a histogram of a bin of for example 0.5 will not give any additional information and there would be lots of loss of information if bin size 2 is chosen.



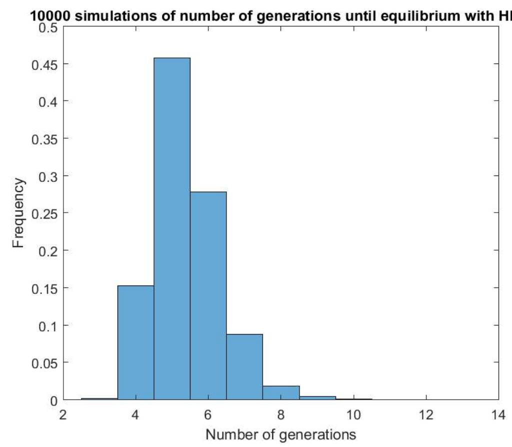
(a) 10000 simulations of number of generations until equilibrium with HR 1/4



(b) 10000 simulations of number of generations until equilibrium with HR 1/3



(c) 10000 simulations of number of generations until equilibrium with HR 1/2



(d) 10000 simulations of number of generations until equilibrium with HR 1

Figure 9: The probability distribution of Y (the number of generations until reaching the equilibrium) is approximated with a histogram of bin size 1. For each HR (1/4, 1/3, 1/2 and 1), 10000 simulations were ran.

Looking at the histograms, all Y_j's ($j \in \{1/2, 1/3, 1/4, 1\}$) might come from the same distribution, but with different parameters, since their sample means are obviously different.

For simplicity, we chose to only investigate the distribution of $Y_{1/3}$, since other Y_j's might follow the same family distribution as suggested by the histrograms. We suspected that $Y_{1/3}$ might be Poisson distributed. So with this assumption, we first used the 'fitdist'-function of Matlab to calculate the maximal likelihood estimator $\hat{\lambda}$ for the real Poisson parameter λ. Then, we performed the chi-sqaured test using the chi2gof function in Matlab, with null-hypothesis $H_0 : F_{1/3} = F_{P,\hat{\lambda}}$, for which F is the unknown cumulative distribution function(CDF) from the data and $F_{P,\hat{\lambda}}$, the CDF from Poisson($\hat{\lambda}$). The alternative hypothesis is $F_{1/3} \neq F_{P,\hat{\lambda}}$. Furthermore, we chose 8 bins for the chi-sqaured test(otherwise the test is not accurate due to low expected counts in some bins). The test has rejected the null-hypothesis and the test was of significance level $\alpha = 0.05$.

With the same method, we've also tested whether $Y_{1/3}$ is binomial or negative binomial distributed. But again, they have been rejected by the chi-sqaured test. So we can only conclude that $Y_{1/3}$ is not Poisson, binomial and negative binomial distributed.

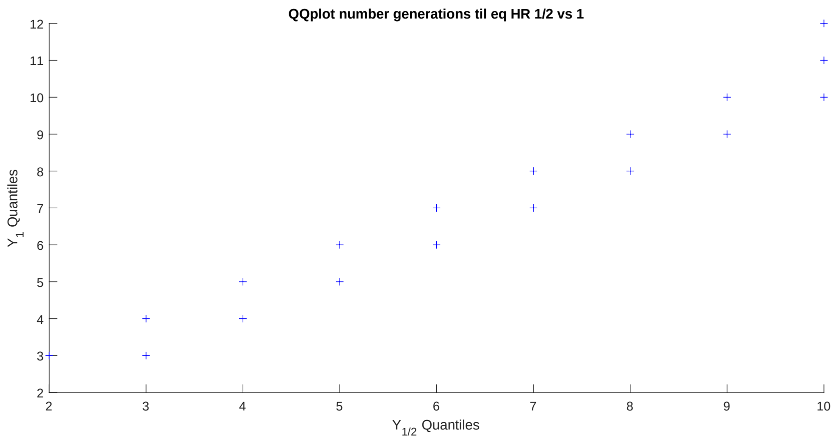
We've also performed this statistical analysis on other Y_j's(for instasnce Y_1), and we've obtained the same result as for $Y_{1/3}$.

In order to make a better comparison of the distribution between $Y_{1/3}, Y_{1/4}, Y_{1/2}$ and Y_1, we made some QQ-plots, which is shown in figure 10.

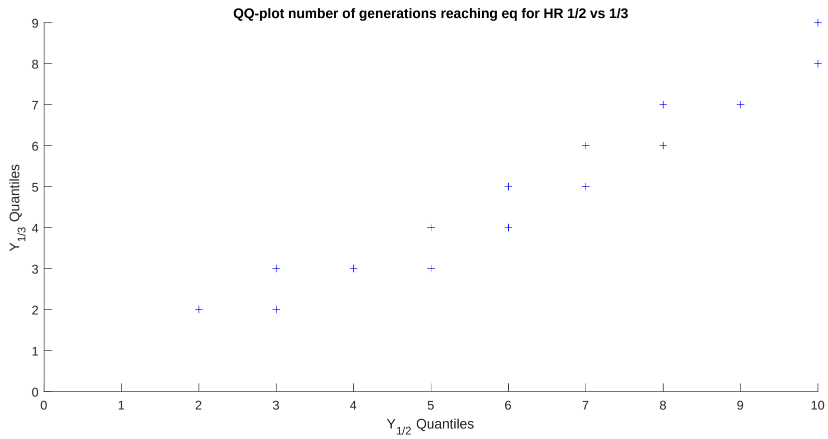
Looking at the histograms in figure 9, all Y_j's ($j \in \{1/2, 1/3, 1/4, 1\}$) might come from the same distribution, but with different parameters, since their sample means are obviously different. In an attempt to find a good (confidence level 95%) fit for the distribution of $Y_{1/3}$, we've used the chi-squared goodness of fit test(chi2gof) to see whether $Y_{1/3}$ is Poisson distributed. We first used the 'fitdist'-function of Matlab to calculate the maximal likelihood estimater for $\hat{\lambda}$ and then performed the chi-sqaured test. The test has rejected the null-hypothesis $H_0 : Y_{1/3} \sim \text{Poisson}(\hat{\lambda})$ with significance level $\alpha = 0.05$.

With the same method, we've also tested whether $Y_{1/3}$ is binomial or negative binomial distributed. But again, they have been rejected by the chi-squared test. So we can only conclude that $Y_{1/3}$ is not Poisson, binomial and negative binomial distributed.

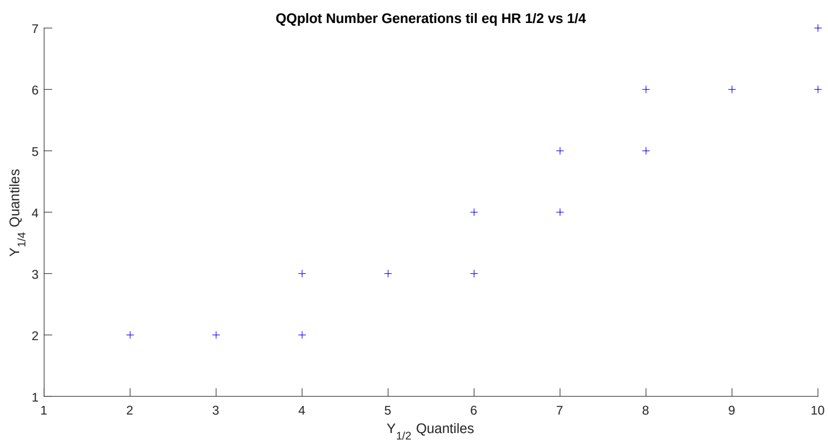
In order to make a better comparison of the distribution between $Y_{1/3}, Y_{1/4}, Y_{1/2}$ and Y_1, we made some QQ-plots, which is shown in figure 10.



(a) QQ-plot of Y_1 vs $Y_{1/2}$



(b) QQ-plot of $Y_{1/3}$ vs $Y_{1/2}$



(c) QQ-plot of $Y_{1/4}$ vs $Y_{1/2}$

Figure 10: QQ-plot of Y_1 vs $Y_{1/2}$, $Y_{1/3}$ vs $Y_{1/2}$ and $Y_{1/4}$ vs $Y_{1/2}$

As expected, the QQ-plots have a stairwise pattern. This is because our data came from a discrete random variable. We see that the plots appears linear, especially the plot of Y_1 vs $Y_{1/2}$. The plot indicates that these 4 Y_j's might belong to the location-scale family of some distribution.

With the two-sample kolomogorov-smirnov test (kstest2) of significance level $\alpha = 0.05$ in Matlab, which has null-hypothesis $H_0 : F(X) = G(Y)$ for F, G the CDF of X and Y, it can be concluded that these Y_j's don't come from the same distribution with the same parameters. Unfortunately, we weren't able to find a test that tests whether these Y_j's comes from the same family distribution (i.e. same distribution but different parameter values).

5 Is every individual happy after equilibrium?

As it is shown in the section 3, equilibrium is not always reached within 10000 generations, stronger even, equilibrium is occasionally never reached at all. But when it does, it is natural to question whether every individual is happy and how happy they are. To answer this question, we ran 500 simulations for each HR ranging from 0 to 1 with increments of 0.01 and we calculated what the average, maximal and minimal happiness of the population is for each happiness rule. We ran the simulations om the standard board and the results are shown in figure 11.

From figure 11, we observe that the minimal happiness is almost always equal to or greater than the happiness rule, which means that every individual is in fact happy in the equilibrium. We also see that maximal happiness is always 1, which means there is always at least one invididual who lives homogenous. Furthermore, we see that the average happiness is always closer to the maximal happiness than to the minimal happiness. At a HR of approximately 0.8, we see that every individual has happiness 1. This implies that happiness 0.8 guarantees nearly full segregation. Complete segregation can only be assured at a HR greater than 87.5, Namely, assume a person is not homogenous, then at least 1 out of at most 8 neighbours does not share his/her type. In other words, the happiness of this person is at most equal to $1 - \frac{1}{8} = 87.5\%$.

It is not very surprising that every individual is happy after the equilibrium, since it would otherwise mean that one individual is unable to find a place that better meets his/her desire. On a board with 24 empty spaces, this is rather unlikely. Another remarkable result is that figure 11 shows that the minimal happiness is higher than the HR and that the average happiness is nearly 1 if the happiness rule higher than 0.8. At the first glance this might seems impossible because the heavy requirement of being happy, but on the other hand, it shows that the strong 'need' for segregation actually leads to greater happiness.

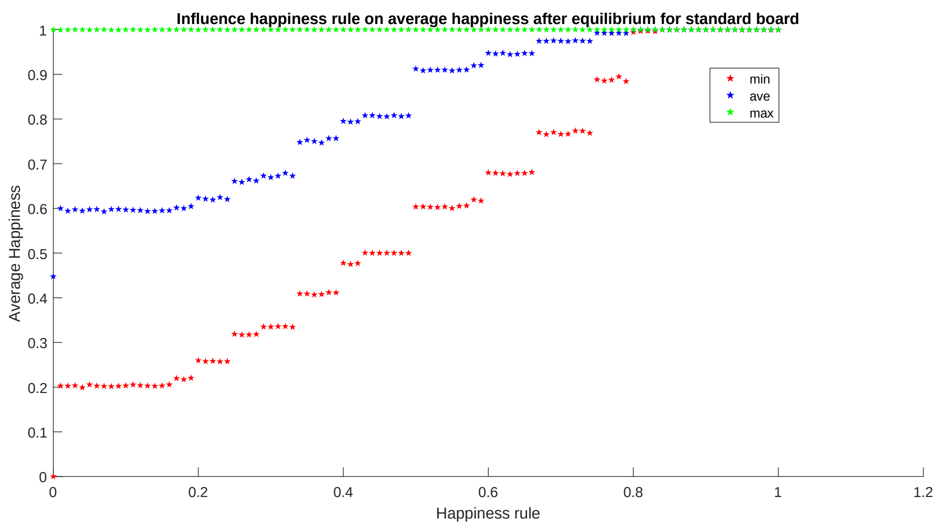


Figure 11: Effect of happiness rule on the happiness of the individuals. The green graph shows the maximal happiness, the blue graph shows the average happiness and the red graph shows the minimal happiness

To illustrate that the variance of every above calculated average happiness is very low, we selected Happiness Rule 1/3 as an example and made a histogram in figure 12. In figure 12, we see that the probability of every individual is happy is practically 1.

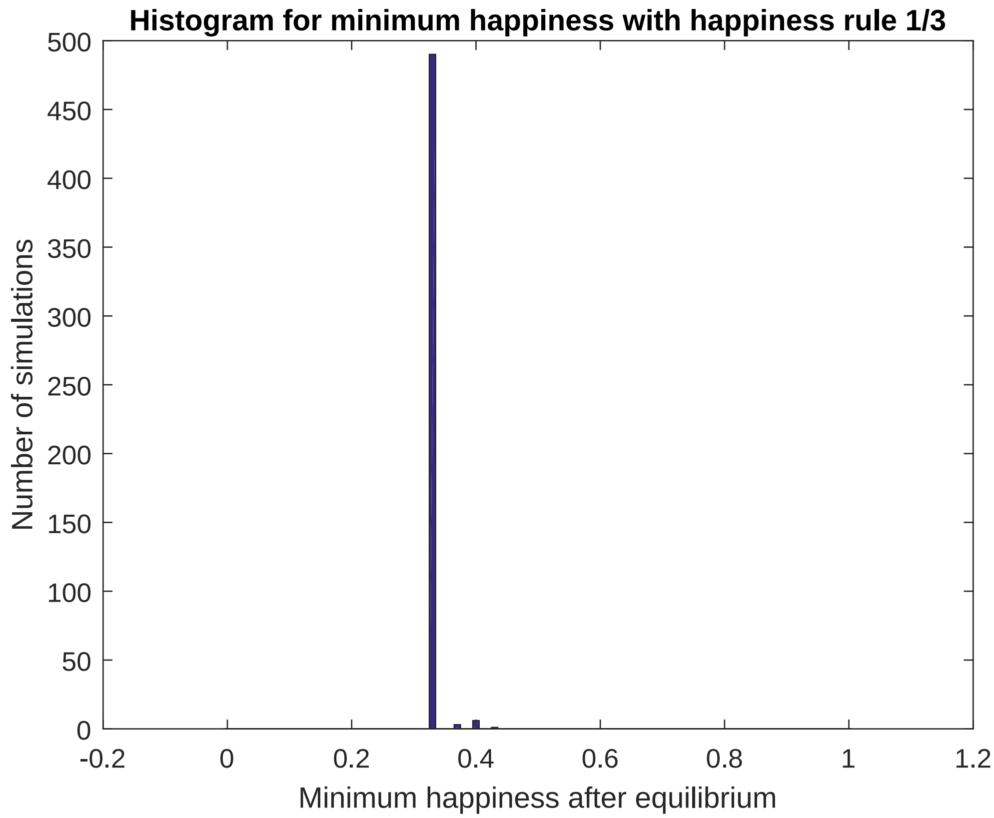


Figure 12: Histogram for minimum happiness after 500 simulations with HR=1/3

6 Average segregation time

As mentioned in the introduction, the segregation time at $n\%$ is defined as the number of generations until at least $n\%$ of the population on a board lives in homogenous groups. Where person i is said to live homogenous if for any neighbour j of i, we have $\text{Type}(j) = \text{Type}(i)$. This gives immediate rise to questions concerning the relation between the choice of n and the average segregation time at $n\%$. Furthermore, it is unclear if segregation at $n\%$ is guaranteed before a board reaches an equilibrium and what the effect is of the happiness boundary on the existence of a segregation time.

To research any of the given questions, we will first have to formalise our choices of board as well as the questions proposed.

6.1 Formalisations

Prior to starting any test or properly formalising our research questions however, we note that segratation at $n\%$ does not necessarily have to happen: If we consider $n = 100$ on the standard board with happiness 1/3. We will nearly never have total segregation before the board reaches an equilibrium. Therefore one might instead consider the average fraction of segregation at equilibrium,for any given happiness fraction.

Furthermore, the average segregation time as function of the segregation fraction should theoretically be a strictly increasing function since for any given board we have:

$$\begin{aligned} &n\% \text{ lives in homogenous groups after } k \text{ generations } \Rightarrow \\ &m\% \text{ lives in homogenous groups after } k \text{ generations, for any } 0 \leq m \leq n \end{aligned}$$

Having noted these facts, we can now properly formalise the research questions.
The following questions are proposed:

1. What is the relation between the average segregation time and n.
2. What is the average segregated fraction of the population after a board reaches equilibrium for given choices of happiness.
3. For any Happiness Rule ranging from 0 to 1, how often do we reach at least 60%, at least 80% or even full (100%) segregation?
4. For 60% and 80% segregation, what distribution do we get for the segregation time, if $HR = 1$? Are these comparable?

To establish results regarding these questions, we consider different setups in testings. We will be testing two different boards. The first board to be analysed is the standard board. The second board is a larger "4-Type" board. The details are specified below:

Table 1: Specs of the two considered boards

	Standard Board	4-Type Board
Number of types:	2	4
Length:	8	10
Width:	8	10
Happiness:	1	1
Population per type:	20	16

The 4-Type board is constructed to maintain the same ratio of inhabited and uninhabited spots as the standard board. The choice of happiness on these boards is 1 unlike the usual $\frac{1}{3}$. This guarantees that for any $n \leq 100$, segregation at $n\%$ takes place prior to the board reaching an equilibrium. To observe the average segregation time $n\%$ for any n, 500 simulations will be ran per board and averaged out in order to give an approximation for the average segregation time at $n\%$. Likewise the average segregated fraction will be estimate by the average of the segregated fraction of an equilibrium from 500 simulations with given happiness q.

6.2 Results

6.2.1 Question 1

The results regarding the first question are shown below:

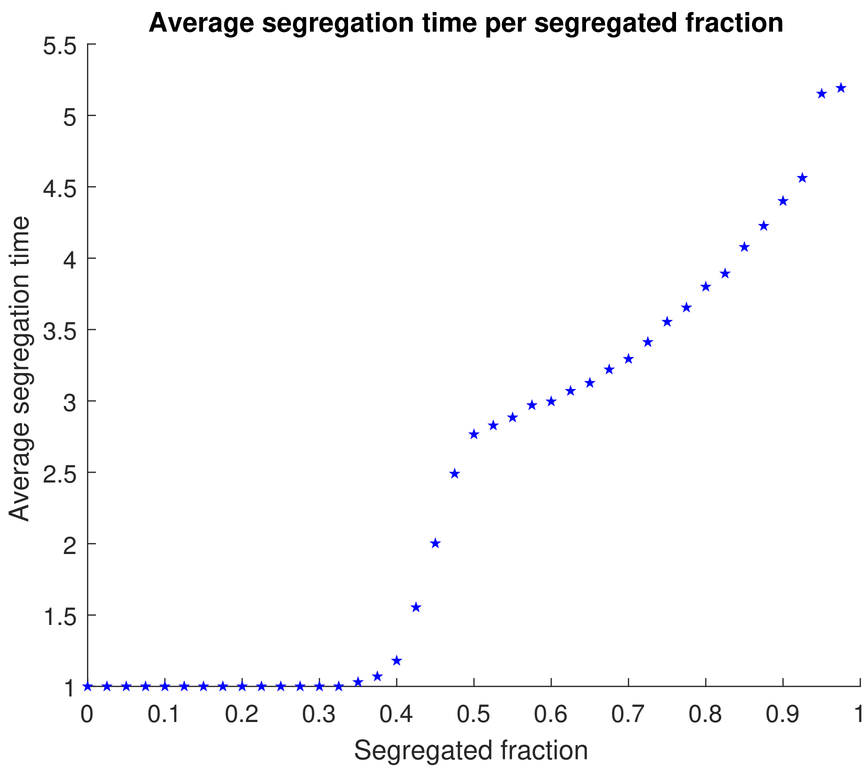


Figure 13: Average segregation time on the standard board

Most notable of figure 13 is that it is neither linear nor exponential, which is what one might initially expect. Instead, it appears to be partially exponential and partially lineair. From the figure, we note that the average segregation time increases fastest between 0.4 and 0.5. Which is to say that it takes three times longer for 50% of the population to live in homogenous groups than for 40%. Also note that the 'lift off' is approximately at $\frac{1}{3}$, but it is unlikely that this has anything to do with the standard happiness rule of $\frac{1}{3}$.

Partial explanation

The first part of this graph is easilyunderstood. When the board is initially formed, the odds are fairly high that several persons might already live homogenous, Due to the HR of 1, any non homgenous person will move, resulting in a further segregation. As the boundary of 10% is fairly low, it is likely that this will be achieved in less than 1 generation. The second interesting segment of this graph is the segment between 0.4 and 0.5, as mentioned, the figure shows that it takes considerably longer to reach a segregation of 50% than it takes to reach 40%. This can be argued from a sociological perspective as follows some people are in general unhappy with their lives and will thus move to a new location anyways, the process of segregation, is a fairly lenghty process this way. However once people discover that their close neighbours are moving, they in turn decide to move as well. This does not neccessarily result in more moves, however as one type leaves, the other type is to remain in place, since they will tend to be more homogenous when their opposite type leaves. Thus once 50% segregation has been reached, it becomes easier to segregate even further. Since the repositioning of one individual may result in three people becoming homogenous. This explains the part between 50% and 70%. It might appear reasonable to believe that this increase will continue and that it that it would take at most two more generations to reach equilibrium but after 70% segregation, it becomes harder for individuals to move to a homgenous location. This is also heavliy impacted by the laws defined in the model, a person will not move to the best location but to the closest better location, which might lead to a person moving from one disfavourable location to another.

Next, we consider the same question however this time on the 4-Type board. This gives following results:

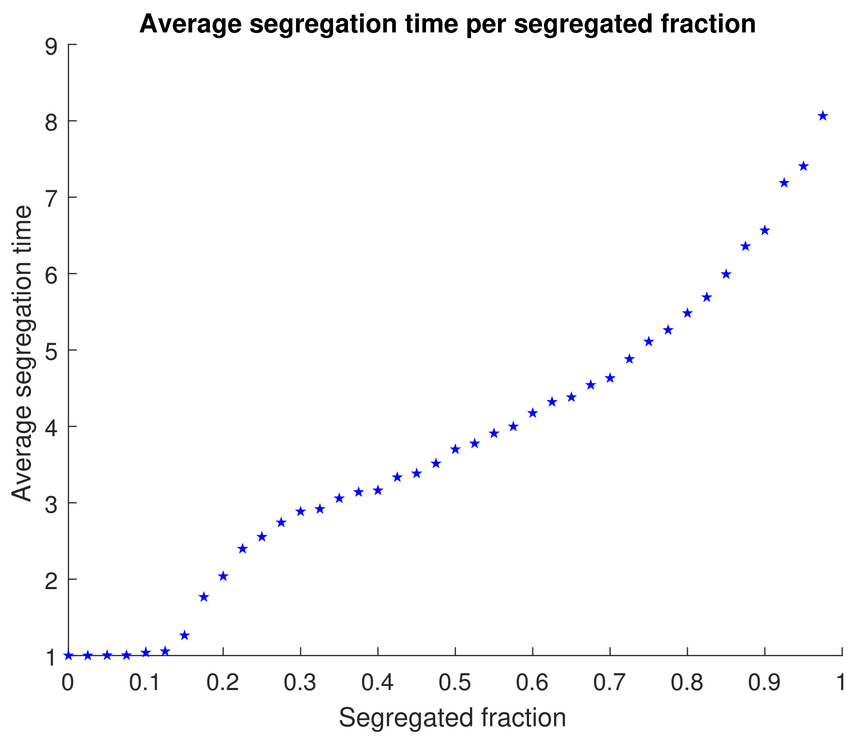


Figure 14: Average segregation time on the 4-Type board

The most notable difference between this figure and the figure 13 is that the 'lift off' lies happens much earlier than in figure 13. Although this difference feels intuitive, we will give a formal explanation. Consider the probability that any individual i has a neighbours of the same type, and no neighbours of another type, given $k - a$ empty spots, where k is the maximum number of possible neighbours for the location of i. In the basic model, this probability equals

$$p_b = \frac{19}{39}\frac{18}{38}...\frac{19-a}{39-a} = \frac{\frac{19!}{(19-a)!}}{\frac{39!}{(39-a)!}} = \frac{19!(39-a)!}{39!(19-a)!}$$

In the 4-Type model, this probability is

$$p_4 = \frac{15!(63-a)!}{63!(15-a)!}$$

which is less than half of p_b.

Also note that an individual is more likely to be homogenous if he/she is placed in a corner or on an edge. This is because a corner or edge spot has less possible neighbours than an interior spot, thus increasing the probability of only having neighbours of the same type. The 4-Type board has relatively more interior locations and fewer edge positions than the standard board. This effect also decreases the fraction of segregation prior to displacement of individuals. However the people/boardspace ratio is maintained in both models and thus should not have too much of an impact on the results.

Another effect is the rise of the number of generations, that is, the time it takes before the board reaches full segregation. However, this effect is less strong than one might initially expect. On the standard board, it takes on average 5.5 generations to reach full segregation. On the 4-Type board however, it takes 8.5.

The remaining part of figure 14 is quite similar to figure 13: quick rise, followed by a slightly decreasing ascending relation. The explanation is similar to that of the basic model.

6.2.2 Question 2

Figure 15: Average segregated fraction as a result of happiness on the standard board

Figure 15 displays three scattered values. The red and green dots represent the minimum and maximum segregrated fraction of 500 boards for the given happiness. Notable in this picture is that the average segregated value is greater than the happiness. Another thing worth noting is that the segregated fractions tend to appear in different groups seperated by relatively large percentages.

Figure 16: Average (blue) segregated fraction as a result of happiness on the 4-Type board. Also included max (green) and min (red).

Figure 16 shows that contrast to the standard board, this board does not respect the property that the average segregated fraction is always greater than or equal to the happiness. A thing of interest however, is that the averages of both figure 15 and 16 seem to cluster. This is because, as mentioned earlier, the effect of the Happiness Rule is quite discrete when using second order neighbourhood.

6.2.3 Question 3

In order to answer the third question, we ran 500 simulations for each HR ranging from 0 to 1 with increment 0.025, and for each segregation %, 60, 80 and 100. The results are shown below.

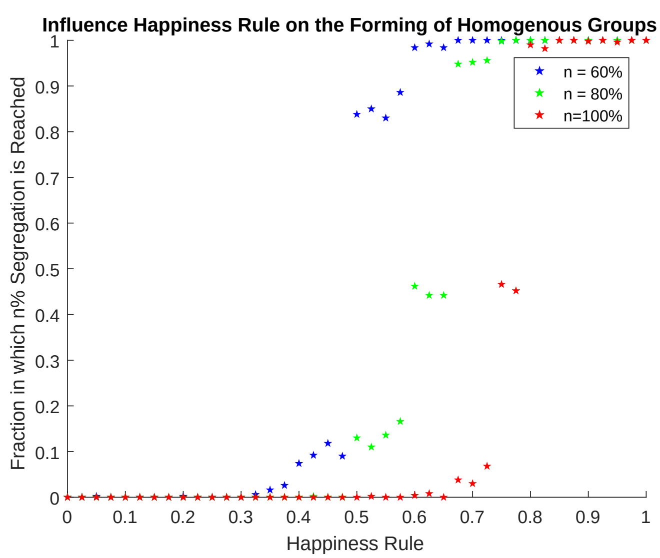


Figure 17: For each $HR = 0, 0.025, ..., 1$, the fraction of the simulations in which 60% (blue),80% (green) and 100% (red) segregation is reached.

6.2.4 Question 4

We will now focus on a more relevant question, namely the distribution of the 60% and 80% segregation time. The plots are shown below.

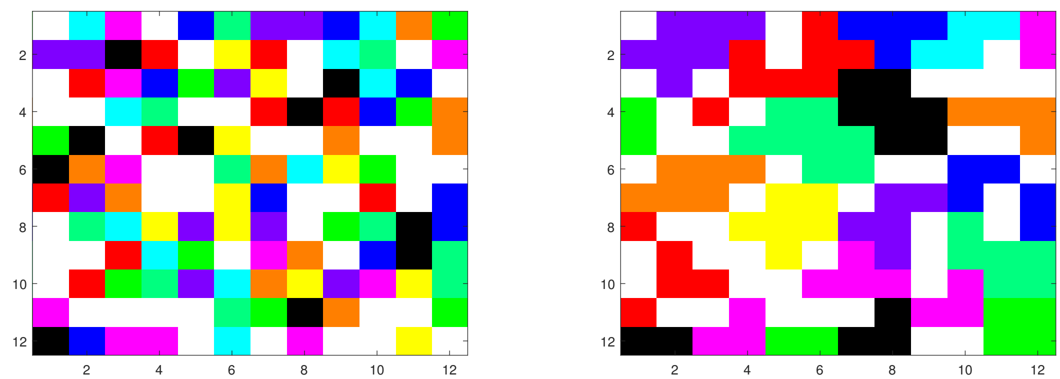
7 The effect of switching types

Segregation is a phenomenon taking many different forms. Rather than limiting the scope to racial segregation, we decided to broaden our view. For example, one might delve into the segregation in studies with respect to friendship. In this case, for any person k. The neighbours of k represent the k's friends or the persons with whom k spends most of his time. In this case, k's type, can represent either the classes he is currently taking, or the kind of sports person k is doing, or more likely, k's political beliefs. In any of these cases, the type of k is not necessarily set indefinitely.
In these cases, the type of k might switch, depending on the types of his friends. This is a sociological process known as conformation and plays a major role in the everyday behaviour of people.

This gives rise to a new modification of the model. For any person k, before moving to a new location, k has a catagorical probability with to switch to a different type, where

$$\mathbb{P}(\text{NewType}(k) = t) = \begin{cases} \frac{\#\text{Neighbours of type } t}{\text{Total number of neighbours}} & \text{if total number of neighbours} > 0 \\ \mathbb{1}_{\{t=\text{Type}(k)\}} & \text{if total number of neighbours} = 0 \end{cases}$$

Note that this is well defined, since if total number of neighbours > 0, we have $\mathbb{P}(\text{NewType}(k) = t) \geq 0$ and $\sum_{i=0}^{\text{types}} \mathbb{P}(\text{NewType}(k) = t) = 1$. An example of how this may effect the equilibrium, is illustrated below:



(a) Situation before segregation

(b) Situation after segregation

Figure 18: An illustration of the effect of switching types on the 12×12 Board

A crucial difference between equilibrium with and without switching, is that if switching is allowed, the concentration of the types may differ between the starting board and the equilibrium status. This can be seen in figure 18 as the number of people of the green type decreased from 10 to 8, the yellow type even decreased from 10 to 6. As a matter of fact, allowing switching might lead to the extinction of several types.

This is not generally the case for lower bounds on the happiness rule, but if the happiness rule is increased, several types might cease to exist. Despite this being quite interesting, the extinction of types is not included in this report. What will be studied in this report, is the effect of switching, on the given research questions. In other words, how will the results discussed in this report, vary when switching is allowed.

7.1 Average number of generations until equilibrium

The first graph to be discussed, will be the effect of switching on the average number of generations until equilibrium. From figure 8, we deduced that on average, a board reaches equilibrium in only a few generations. This however decreases even more drastically when switching goes in effect:

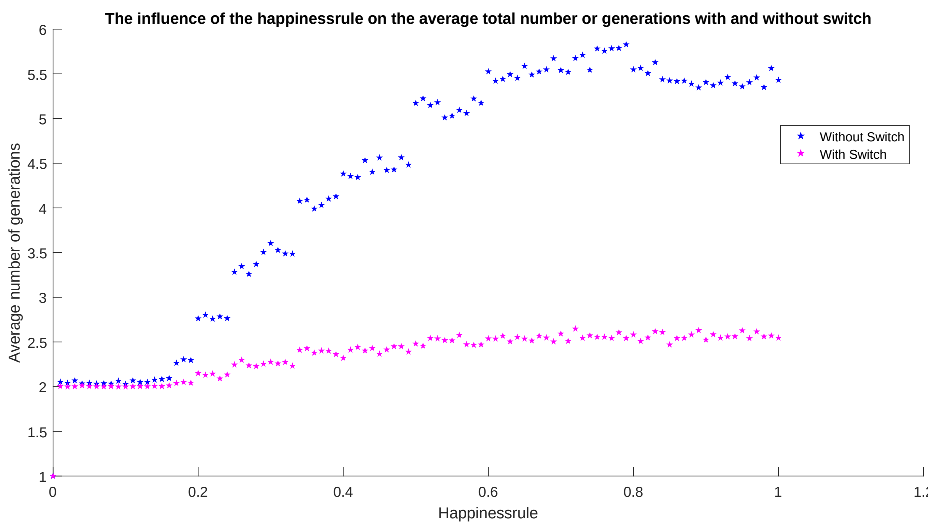


Figure 19: Number of generations until equilibrium on an 8x8 bord with standard setting, with and without type switching

In figure 19, we see a comparison of the average number of generations for with and without type switching. There is a clear difference that when type switching is applied, it takes much less generations to reach the equilibrium. This is expected, since the probability of becoming the same type as the neighbours is more favored in our model, which results in a increase of happiness. Or, in a sociological view, the person is willing to adapt to the enviroment rather than moving on.

From figure 19, we observed that the average generations is nearly constant for any HR in the interval $(0.6, 1)$. Just to get an idea whether this statement is justified, we selected only three HR's, namely 0.6, 0.8 and 1 and performed the two sample t-test(ttest2) in Matlab. This t-test tests the null-hypothesis $H_0 : \mu = \nu$ with μ, ν respectively the mean of the distribution of X and Y, against $H_1 : \mu \neq \nu$. It uses the test statistics:

$$T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_X^2}{m} - \frac{S_Y^2}{n}}}$$

This test assumes that the mean difference $\mu - \nu$ is normally distributed, and since our data are from the discrete random variables, this assumption would not make sense. However, according to the central limit theorems, we may make this assumption for large enough sample size. To avoid difficult calculations of how big the sample size is needed, we simply chose a size of 50000. Also, in our case, we made sure $m = n$, and the test can also be adjusted to not assume equal variance.

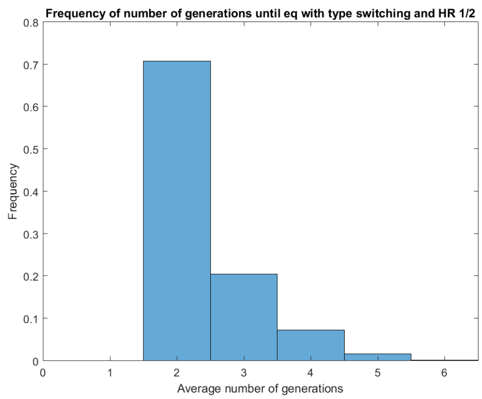
After the simulations, we obtained the following sample means: We applied the t-test twice. First time

Table 2: Sample means of number of generations simulated with HR 0.6, 0.8 and 1, with 50000 sample size for each HR.

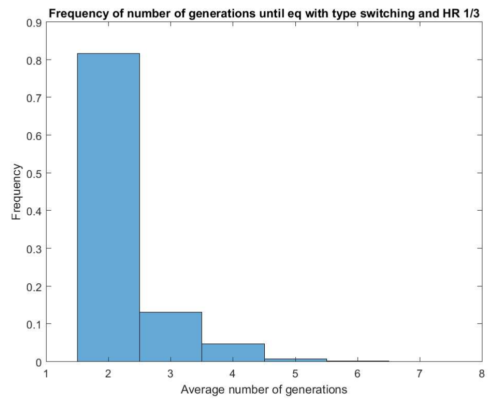
HR	0.6	0.8	1
Sample Mean	2.5213	2.5616	2.5711

comparing the mean of HR 0.6 and 0.8. Second time for comparing the mean of 0.8 and 1. The first test rejected the null-hypothesis, the second did not. So with 95% confidence level, we may conclude that the average generations are unequal for HR 0.6 vs 0.8, but equal for HR 0.8 vs 1.

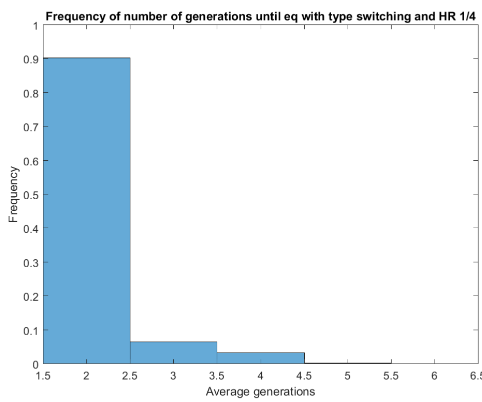
Just like in section 4, in order to get an idea about the distribution of the random variable of $Y_{j,s}$(referring to the notation in section 4, with s meaning type switching), we plotted again histograms for HR 1/4, 1/3, 1/2 and 1. This is shown in figure 20.



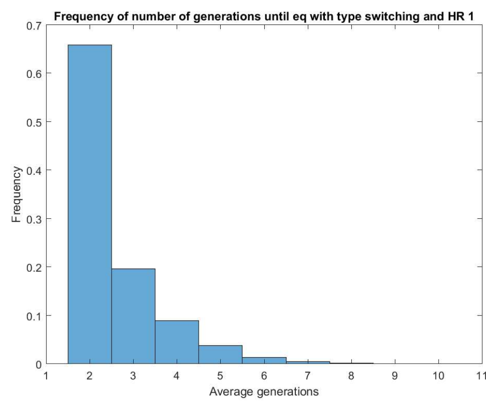
(a) 5000 simulations of number of generations until equilibrium with HR 1/4 with switch



(b) 5000 simulations of number of generations until equilibrium with HR 1/3 with switch



(c) 5000 simulations of number of generations until equilibrium with HR 1/2 with switch



(d) 5000 simulations of number of generations until equilibrium with HR 1 with switch

Figure 20: The probability distribution of Y (the number of generations until reaching the equilibrium) is approximated with a histogram of bin size 1. For each HR (1/4, 1/3, 1/2 and 1), 5000 simulations were ran.

Again, all $Y_{s,j}$'s appears to be following the same distribution just as without switch(section 4). The histograms suggest that the $Y_{j,s}$'s might follow a geometrical distribution, which is a Negative Binomial with $r = 1$. But we could not perform the same statistical analysis as in section 4, because the maximum likelihood estimators for the Negative Binomial could not be calculated through the fitdist or other functions. This is due to the fact that the sample means exceeded the sample variance. So from this we could conclude that $Y_{j,s}$'s cannot be geometrically(or NB) distributed.

It is much more interesting to compare the distribution of $Y_{j,s}$ with Y_j. A remarkable observation is that the histograms of $Y_{j,s}$ completely lost it's symmetry compared to the histograms of Y_j. Apparently, type switching stimulates segregation so well that it only takes 1 generation to reach equilibrium(remark: we always start with 1 generation, so the fastest segregation is reached at the 2nd generation). To further examine how type swithcing affects generations until equilirbium, we chose to compare $Y_{1,s}$ with Y_1 in a QQ-plot:

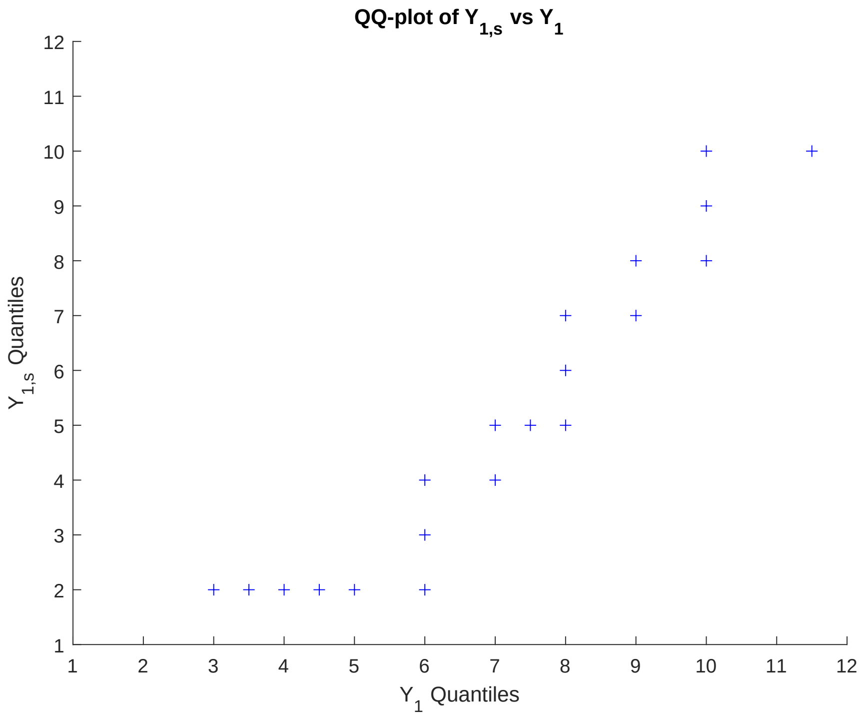


Figure 21: QQ-plot of $Y_{1,s}$ vs Y_1

We see from figure 21 that starting from quantile 3 of Y_1, the line is flat. This is expected because the most quantile of $Y_{1,s}$ is concentrated on 2. After the flat line we see a linear pattern, which suggests that the 'tail' part of $Y_{1,s}$ and Y_1 are not too differently distributed. This can be also observed in the histograms.

7.2 Average segregation time

The final topic to be discussed in this report, will be the effect of conformation on the (average) segregation time. For this purpose, we will compare the results obtained in section 6 with the results if switching were allowed. We will only consider the results for the standard board.
For any $n \in [0, 100]$. We will consider the time until which $n\%$ of the population lives in homogenous groups. With a chosen HR of 1, we obtain:

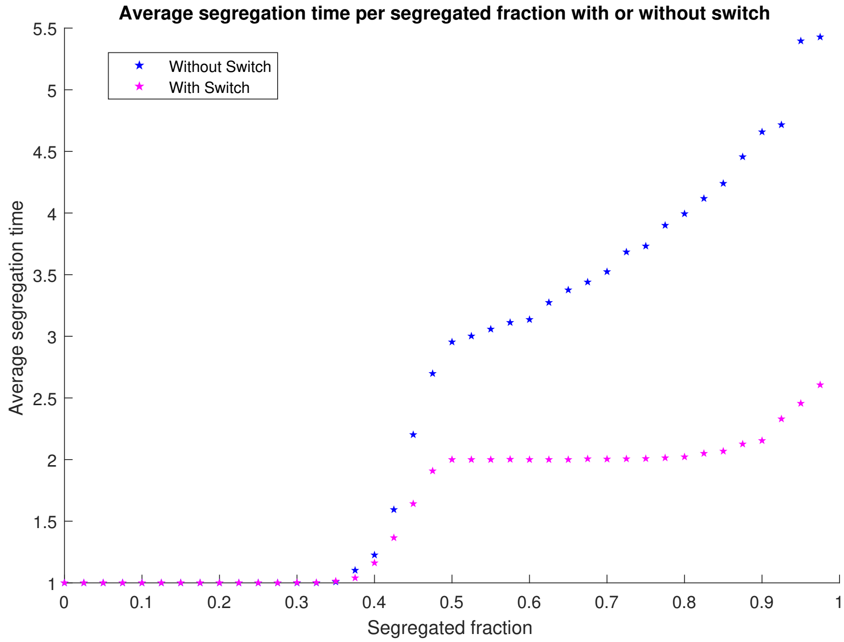


Figure 22: Average (blue) segregated fraction as a result of happiness on the 4-Type board. Also included max (green) and min (red).

Appendix

Here are some of our results we did not use in our report, because we did not investigate them. However, these graphs may be really interesting to investigate further.

First, we show our results on the effect of the Happiness Rule on the total number of moves.

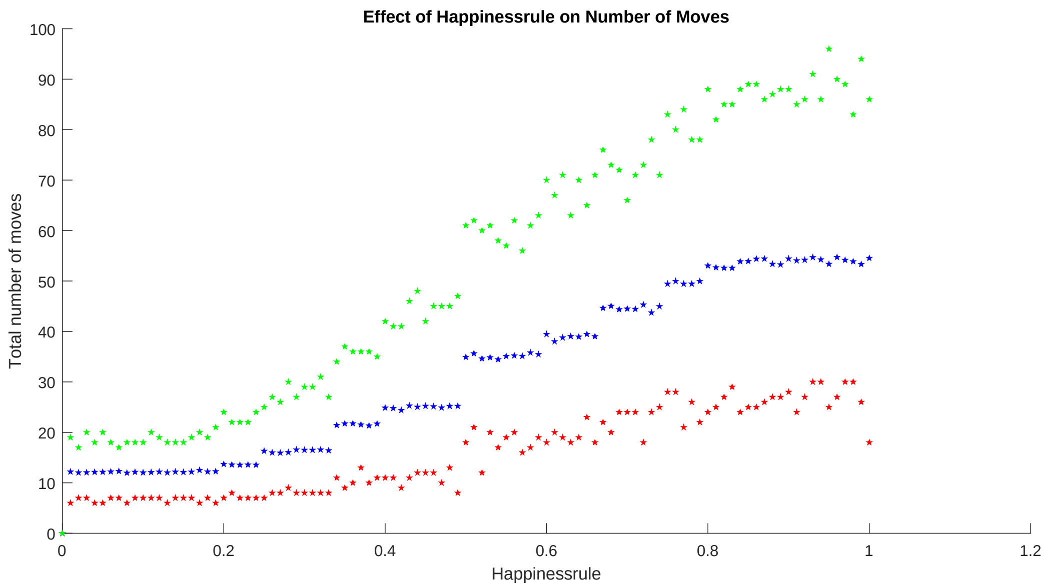


Figure 23: Effect of the Happiness Rule on the total number of moves with the standard settings

We also have some results of the effect of the random move on the number of generations.

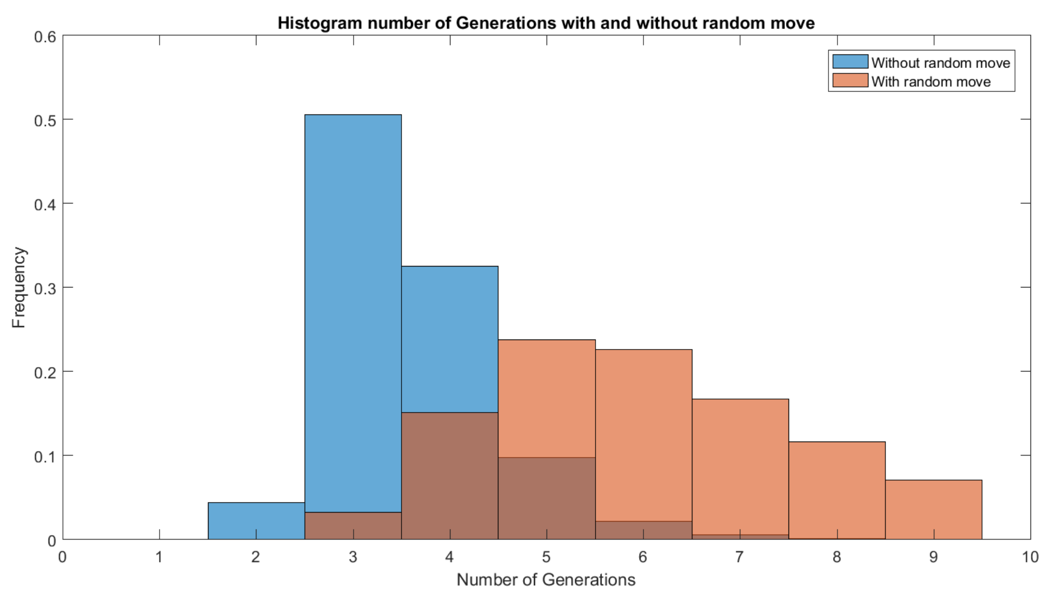


Figure 24: Effect of the random move on the number of generations with the standard settings

Furthermore, we researched the effect of the order of neighbourhood on the number of generations, the number of moves, the segregration fraction and the average happiness of all individuals in equilibrium.

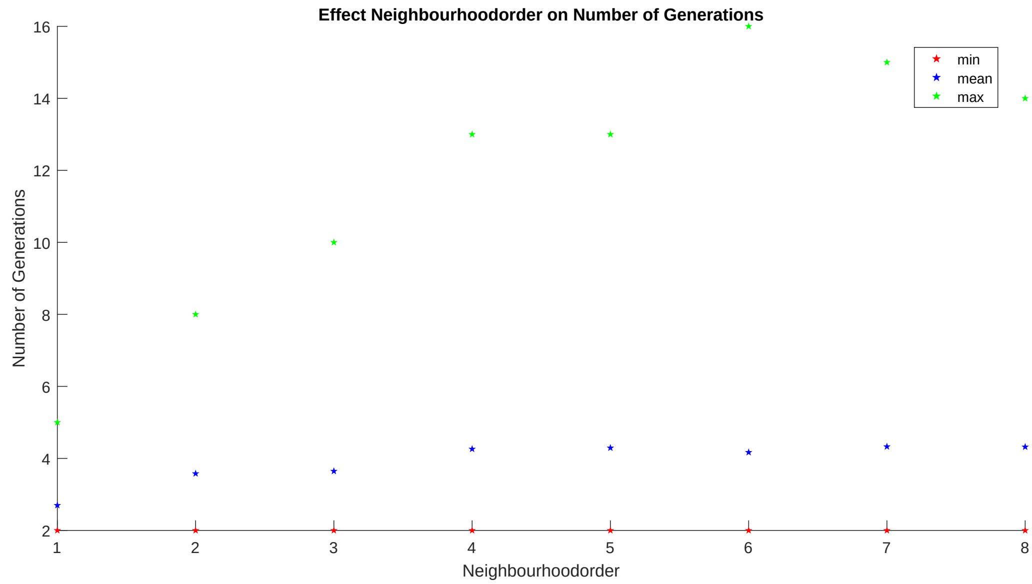


Figure 25: Effect of the Neighbourhoodorder on the number of generations with the standard settings

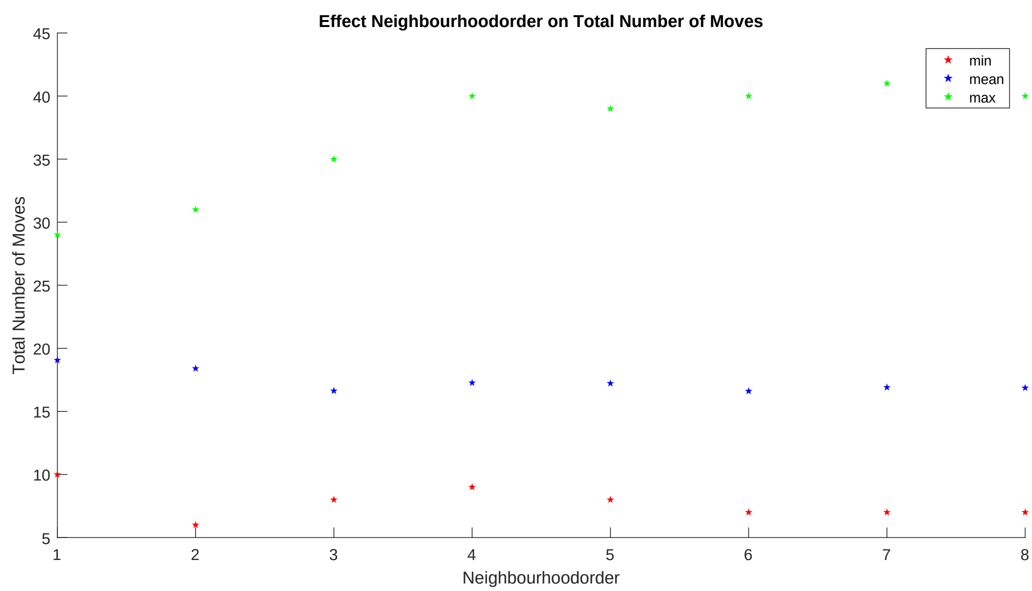


Figure 26: Effect of the Neighbourhoodorder on the total number of moves with the standard settings

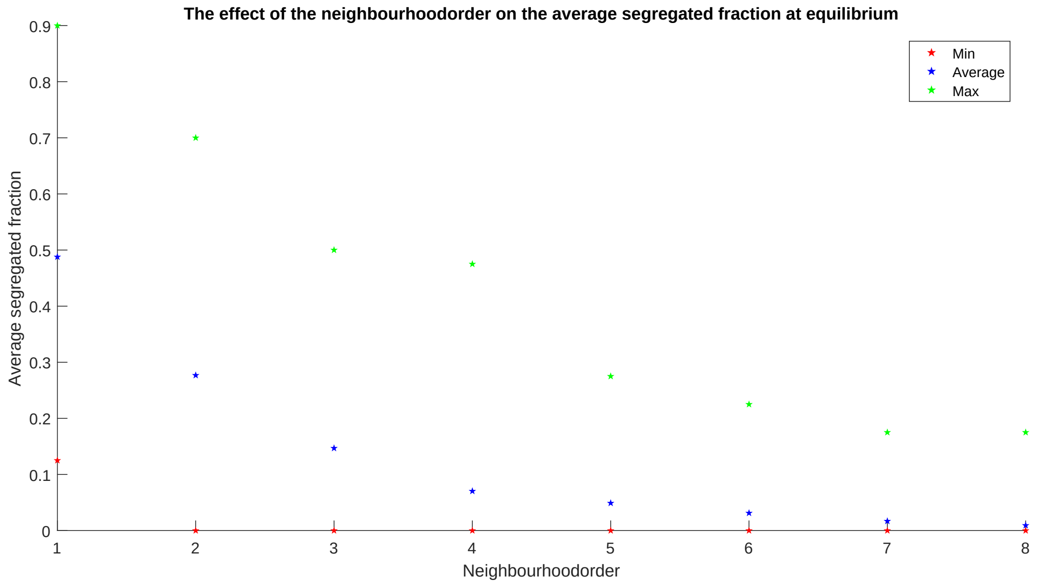


Figure 27: Effect of the Neighbourhoodorder on the segregated fraction in equilibrium with the standard settings

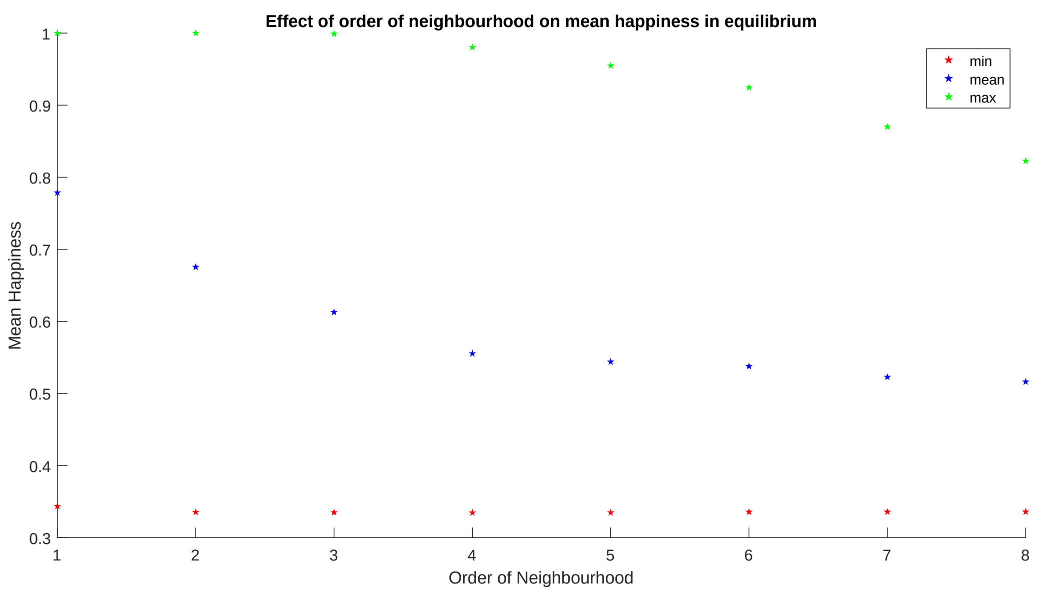


Figure 28: Effect of the Neighbourhoodorder on the mean happiness in equilibrium with the standard settings

Besides, the effect of the criminals is really interesting. Here are some effects of the number of criminals in the basic board (so the maximum number of criminals is 40).

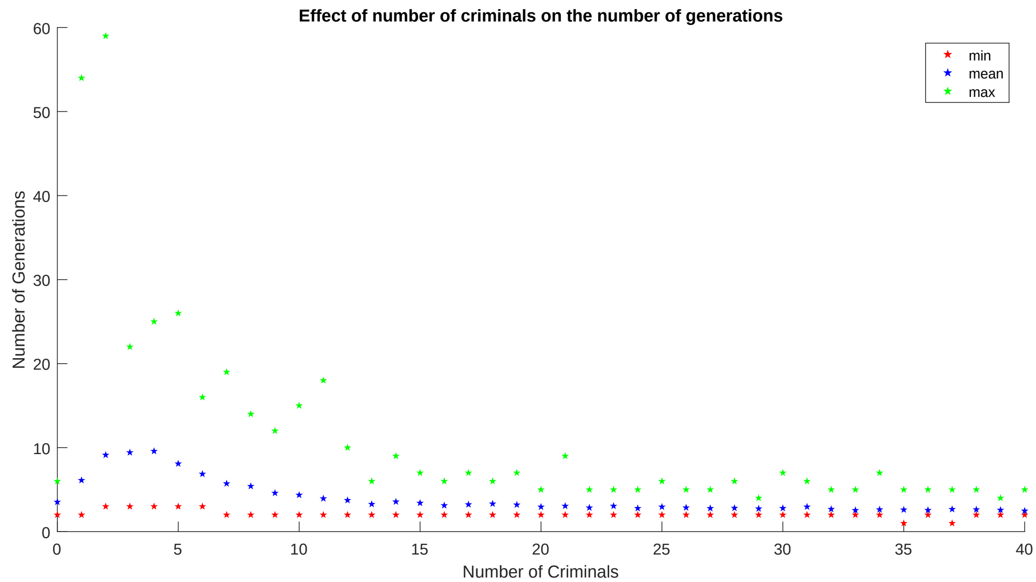


Figure 29: Effect of the number of criminals on the number of generations on the basic board

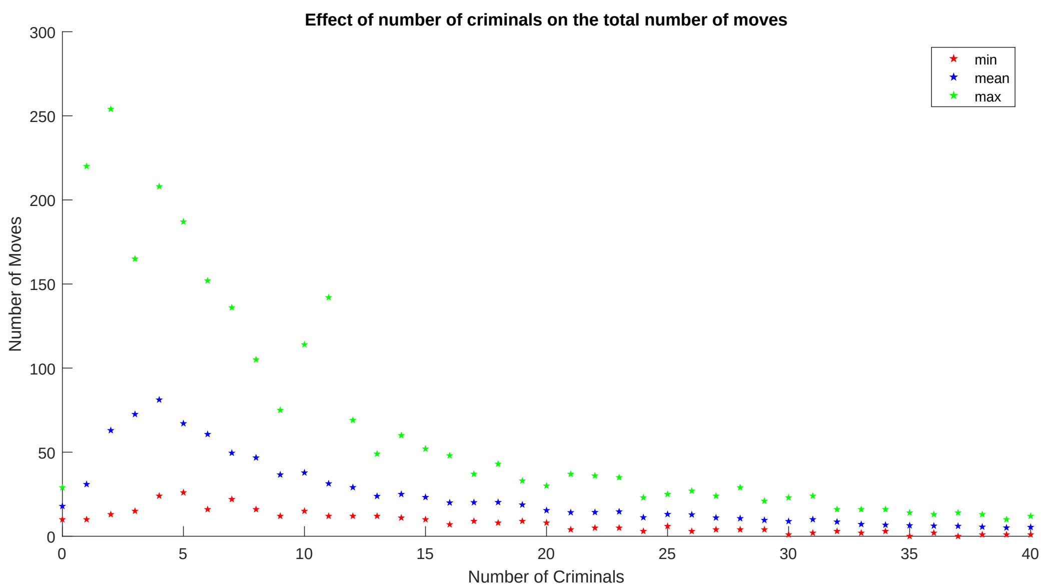


Figure 30: Effect of the number of criminals on the total number of moves on the basic board

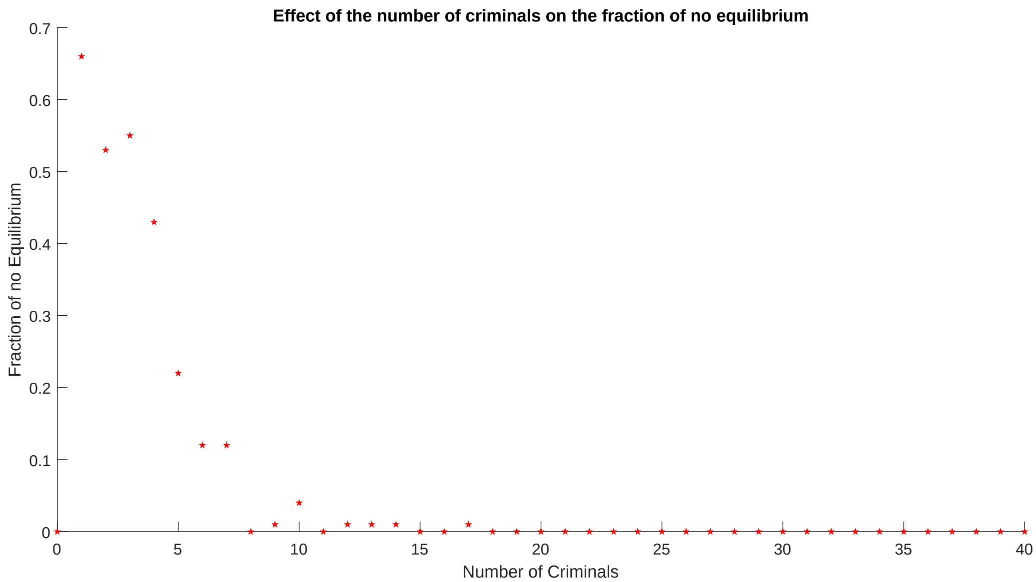


Figure 31: Effect of the number of criminals on the relative number of times with no equilibrium on the basic board

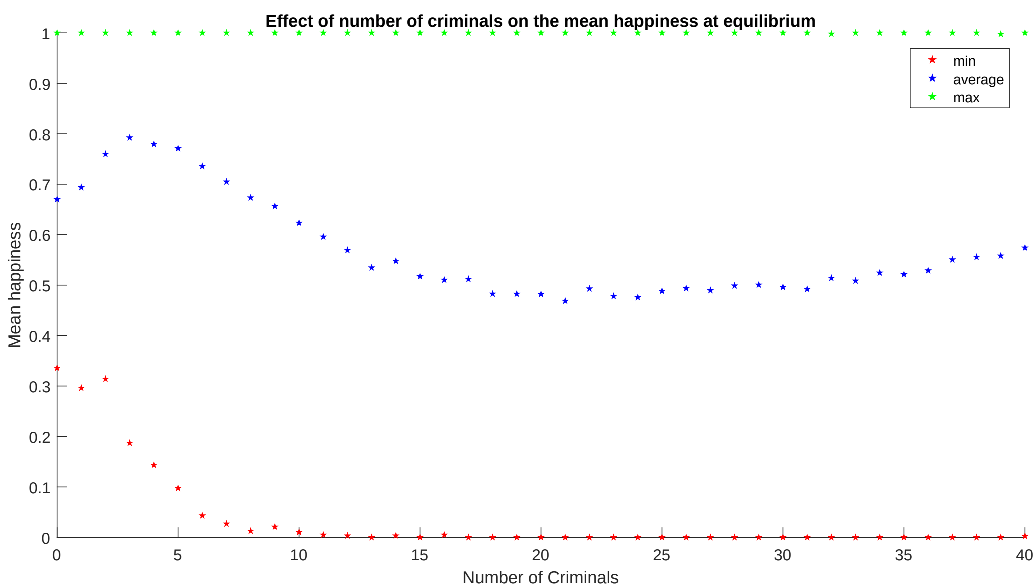


Figure 32: Effect of the number of criminals on the mean happiness at equilibrium on the basic board

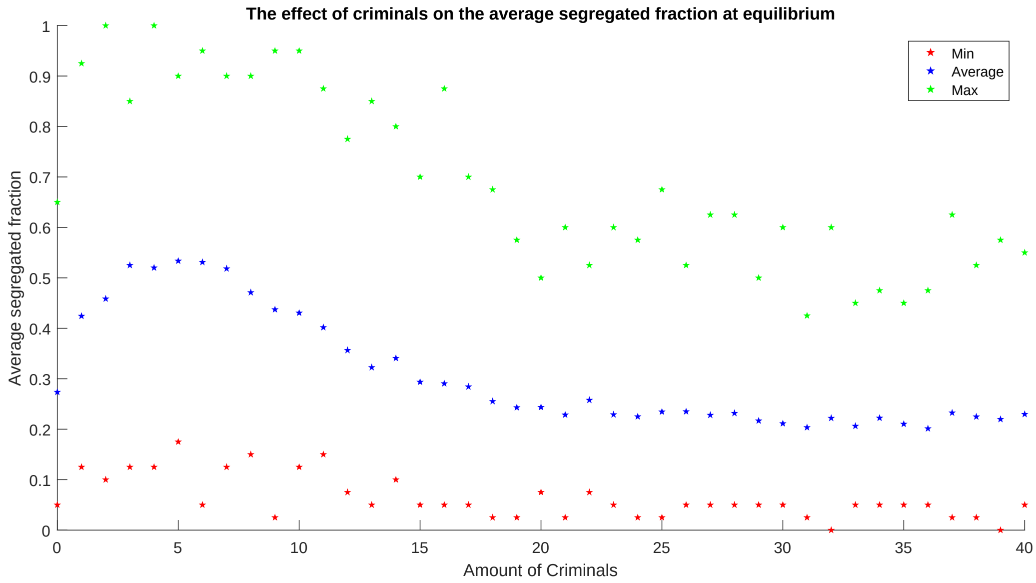


Figure 33: Effect of the number of criminals on the mean segregation on the basic board