

# Department of Economics

## Schelling's Spatial Proximity Model of Segregation Revisited

Romans Pancs and Nicolaas J. Vriend

Working Paper No. 487

January 2003

ISSN 1473-0278



Queen Mary  
University of London

# Schelling's Spatial Proximity Model of Segregation Revisited\*

Romans Pans, London School of Economics  
Nicolaas J. Vriend,<sup>§</sup> Queen Mary, University of London

January 2003

## Abstract:

Schelling [1969, 1971a, 1971b, 1978] presented a microeconomic model showing how an integrated city could unravel to a rather segregated city, notwithstanding relatively mild assumptions concerning the individual agents' preferences, i.e., no agent preferring the resulting segregation. We examine the robustness of Schelling's model, focusing in particular on its driving force: the individual preferences. We show that even if all individual agents have a *strict preference for perfect integration*, best-response dynamics will lead to segregation. What is more, we argue that the one-dimensional and two-dimensional versions of Schelling's spatial proximity model are in fact two qualitatively very different models of segregation.

*J.E.L. classification codes:* C72, C73, D62.

*Keywords:* Neighborhood segregation, Myopic Nash Equilibria, Best-response dynamics, Markov chain, Limit-behavior.

---

<sup>§</sup> Corresponding author: N.J. Vriend, Queen Mary, University of London, Dept. of Economics, Mile End Road, London, E1 4NS, United Kingdom, <n.vriend@qmul.ac.uk>, <http://www.qmul.ac.uk/~ugte173/>.

\* We would like to thank conference participants of WEHIA (Trieste) and CEF (Aix-en-Provence), seminar audiences at Queen Mary (London) and GREQAM (Marseille), and Herbert Dawid, Joe Harrington, and Alan Kirman for helpful comments. The usual disclaimers apply.

## 1. Introduction

Schelling [1969, 1971a, 1971b, 1978] presented a microeconomic model of neighborhood segregation that Schelling [1971a] called a “*spatial proximity model*” (p. 149), as it specifies a spatial setup in which the individual agents care only about the composition of their own local neighborhood. More specifically, distinguishing two types of agents, every agent is assumed to be equally happy with any configuration of his neighborhood *except* that he does not tolerate more than a certain fraction of it to be populated by agents of the other type. Unsatisfied agents get the chance (in some arbitrary order) to move to a satisfactory position, until nobody wants to move anymore. Low and behold, what you will see is an unraveling process, starting from a more or less integrated city into a rather segregated city.

Schelling's neighborhood segregation model seems to have become one of the most widely cited and acclaimed models in economics.<sup>1</sup> There may be three main reasons for this. A first kind of reasons is intellectual. The emergence of segregation seems intriguing. Not only is it remarkable that it is only micromotives at the local level that give rise to macrobehavior at the aggregate (global) level, but what makes it even more interesting is that this emerging macrobehavior does not simply correspond to the underlying micromotives, i.e., segregation occurs although no individual agent strictly prefers this. Moreover, the model appears to be one of the very first *models* of complexity and self-organization in economics. The second reason for the fame of Schelling's model is educational. The model is unusually simple. Combined with its intellectual appeal, this makes it a convenient means to explain to the unacquainted the importance of the idea of unintended consequences resulting from the interaction between individuals. What is more, the model is a '*do-it-yourself*' model of self-organization, as it can be easily verified by anyone who has a typewriter or a checkersboard. Finally, a third reason why Schelling's model is so widely cited is empirical. Neighborhood segregation and ghetto formation are a serious social and political issue in the USA and elsewhere.<sup>2</sup>

Notwithstanding the general recognition given to Schelling's model nowadays, there may be some reason to be concerned about the assumptions made by Schelling [1969, 1971a, 1971b,

---

<sup>1</sup> See, e.g., Akerlof [1997], Arrow [1998], Barkley Rosser Jr. [1999], Binmore [1992], Blume [1997], Brock & Durlauf [2001], Dixit & Nalebuff [1991], Glaeser & Scheinkman [2002], Ioannides & Seslen [2002], Krugman [1996], Lindbeck, Nyberg & Weibull [1999], Manski [2000], Skyrms & Pemantle [2000], or Young [1998].

<sup>2</sup> As Mr. Lott discovered in December 2002, when he had to resign as Republican leader in the US Senate after having made some remarks that appeared to support racial segregation. But segregation occurs not only in a racial context. It can also be found between men and women in an office canteen, between tourists and locals at a city square, between faculty and students in a seminar room, between different nationalities at a conference dinner, or between different species occupying their own territory.

1978] with respect to the individual agents' preferences. That is, although these assumptions are *relatively mild* in the sense that no agent prefers segregation, it is also true that no agent is against it. In other words, while agents in Schelling's model are content to live together in a ratio of 50-50, they are *equally content* to live in a completely segregated city, as long as they can live in a ghetto of like agents. Since these preferences are the driving force in Schelling's model of segregation, it seems worth investigating whether this is essential. Therefore, we provide an extensive formal as well as numerical analysis of the dynamic and equilibrium properties of Schelling's model of segregation.

Our research strategy is to focus on the driving force behind the dynamic behavior of Schelling's model, the individual preferences, while keeping all other details as general and simple as possible. To provide further robustness checks, we also consider a number of variations, such as a board vs. a torus, a line vs. a circle, sequential vs. simultaneous moves, and inertia vs. no inertia.

The main insight from our analysis can be summarized as follows. Schelling's model of segregation is very robust to a whole range of variations in its specification. Most interestingly, and perhaps surprisingly, Schelling's *mild* assumptions concerning the individual preferences can be made considerably more *extreme* in the following sense. Strict preferences for perfect integration by all individual agents will still lead to neighborhood segregation.

Apart from this main insight, our analysis leads to a better understanding of Schelling's model in the following three senses. First, the robustness checks allow us to narrow down the set of elements of the model that appear to be essential to explain segregation. Second, we find that there is a significant difference between the two versions of Schelling's model. We show that the version on a line and the variant on a lattice are qualitatively very different models of segregation. In the two-dimensional model the dynamics are characterized by a progressive deterioration of the level of integration of the available choice locations. Related to this, the key element of the two-dimensional model driving the segregation is the asymmetry in the utility function, i.e., the fact that agents favor a large-majority status over a small-minority status. In the one-dimensional model, however, perfectly integrated locations remain available for choice indefinitely, with newly arriving agents merely pushing incumbents away from such locations. As a result, in the one-dimensional model, segregation occurs even if the individuals strictly prefer perfect integration with no bias whatsoever in favor of the agent's own type. We present a proposition proving that complete segregation is the unique long-term outcome. Finally, we characterize (numerically on a lattice and formally on a line) the set of Myopic Nash Equilibria. This shows that the dynamics in the model are *not* simply an equilibrium selection issue. On the one hand, in most cases considered the set of Myopic Nash Equilibria itself is biased (segregated) relative to a random allocation. On the other hand, when this is not the case, i.e., in

the one-dimensional model with a strict preference for integration, the best-response dynamics do not select an equilibrium at all. In fact, as we will show formally, the completely segregated best-response outcome is the extreme opposite of any Myopic Nash Equilibrium in the sense that the latter are all perfectly integrated from the individual agents' point of view.

While we focus on a positive analysis of Schelling's model of segregation, a better understanding of the model has some normative relevance as well. First, in the one-dimensional model we derive unambiguous social welfare implications. In Schelling's formulation of the model, segregation was not a problem as, from a social welfare point of view, a completely segregated city may be just as good as a perfectly integrated one, whereas with strict preferences for integration this is no longer true. Hence, our analysis not only emphasizes the limits as to what education at the individual level can achieve, but also shows that its effect on welfare might be adverse. Second, suppose a social planner wanted to maximize integration. Then, what preferences would he like his citizens to have? Our analysis provides a benchmark (within the framework of Schelling's model) as to what the education of preferences for integration could achieve. Third, presuming that there is a social welfare case for integration (independent from the specification of the individual preferences), another normative issue concerns the question how a moving subsidy or tax system should look to prevent segregation and implement integration. Although this question is beyond the scope of the current paper, one lesson from our analysis is that a simple reward system for integrating moves or taxation of segregating moves would not work.

The remainder of this paper is structured as follows. In Section 2 we will briefly recapitulate Schelling's [1969, 1971a, 1971b, 1978] spatial proximity model of segregation, and we will outline the features of the model that we are going to modify and analyze in detail. Section 3 presents some analytical tools and benchmark allocations. Schelling's two-dimensional model will be analyzed in Section 4, while his one-dimensional, linear model will be considered in Section 5.

## **2. Schelling's Spatial Proximity Model**

### *2.1 Recapitulation of Schelling [1969, 1971a, 1971b, 1978]*

There are two basic variants of Schelling's spatial proximity model. The first version, presented in Schelling [1969], is a one-dimensional model. Besides this linear model, Schelling [1971a] presents a two-dimensional version as well, which is also the version appearing in Schelling [1971b, 1978].

Schelling [1969, 1971a] considers a number of individual agents, distinguishing two types of individuals (O and X), distributed along a line, i.e., in one dimension (1D).<sup>3</sup> Figure 1a gives an example. An agent's position is defined relative to his neighbors only, and there are no absolute positions. A given individual's neighborhood is defined as the four nearest neighbors on either side of him.<sup>4</sup> Agents towards the end of the line will have less than eight neighbors. Each individual is concerned only with the number of like and unlike neighbors. More specifically, each agent wants at most 50% unlike neighbors; otherwise agents are indifferent. The starting configuration in Schelling [1969, 1971a] is created by distributing equal numbers of agents of each type in random order. The dynamics, then, are an iterative process. At each stage all agents that are not satisfied are put in some arbitrary order. When an agent's turn comes, he moves to the nearest satisfactory position. Since all positions are relative only, he simply inserts himself between two agents (or at either end of the line). Similarly, his own departure does not lead to an empty position. This process continues until no agent wants to move anymore. The typical outcome is a highly segregated state, although nobody actually prefers segregation to integration.

**Fig. 1a.** 1D example

OOXXXOXXOOXOXOOOXXXO

**Fig. 1b.** 2D example

X	0	0	0	
0		X		X
	X	X	X	0
0	0		X	X
X	0	X	0	0

Schelling [1971a, 1971b, 1978] considers a regular lattice with bounds, such as a checkersboard. There are again two types of individual agents, who can each use one cell on the board to locate themselves. But now there are also some free cells left, as in a 5x5 example in Figure 1b. The neighborhood of an individual agent is defined to be the so-called Moore neighborhood. For an agent in the interior of the board this consists of the eight cells directly surrounding his own location, with less neighbors for agents at the boundary. Absolute rather

---

<sup>3</sup> The number of individuals can be finite, but Schelling [1971a] also refers to the possibility of an infinitely continuing line or a line closing in a ring (p. 152).

<sup>4</sup> Notice that the spatial proximity model differs from the so-called 'bounded neighborhood' or 'tipping' model of segregation (see Schelling [1969, 1971a, 1972, 1978]) in that each individual has his own, locally defined neighborhood.

than relative positions characterize this two dimensional (2D) setup, and agents can only move to empty positions.<sup>5</sup>

The preferences considered in Schelling [1971a] are the same as for the one-dimensional model (each agent accepts up to 50% of unlike neighbors), but Schelling [1971b, 1978] also considers the possibility that agents are happy to accept up to 2/3 of unlike neighbors. Schelling [1971a] starts with a random initial distribution of agents, while Schelling [1971b, 1978] creates the starting configuration by reshuffling a perfectly integrated board. In both cases the resulting starting state is typically highly integrated. The BR dynamics, then, work as follows. All unsatisfied agents are put on a list in some arbitrary order. When an agent's turn comes, according to the list, he moves to the nearest available satisfactory position. At the next stage a new list is compiled, and so on. This process continues until no agent wants to move anymore. The typical outcome is a highly segregated state, although nobody actually prefers segregation to integration, or may even be happy to live with some majority of unlike neighbors.

## 2.2 Schelling's model revisited

Many details of Schelling [1969, 1971a, 1971b, 1978] can be varied, but our research strategy will be to focus on the driving force behind the dynamic process, the individual preferences, while keeping the other details as simple and general as possible. More precisely, we will consider a number of utility functions that imply a *strict preference for perfect integration*. That is, all individual agents with such preferences strictly *dislike* living in a segregated neighborhood, even if they would be part of the majority.

Let the utility function of an individual agent be  $u$ , denote the percentage of his neighbors consisting of the other type as  $x$  ( $0 \leq x \leq 100$ ), and the maximum tolerable percentage of unlike neighbors as  $c$ . Then, the class of individual preferences considered can be represented as follows:

$$u(x) = \begin{array}{ll} a + d(50 - |x - 50|) & \text{for } x < 50 \text{ and } 50 < x \leq c \\ a + d(50 - |x - 50|) + b & \text{for } x = 50 \\ 0 & \text{for } x > c \text{ and if } x \text{ not defined (i.e., no neighbors),} \end{array}^6$$

---

<sup>5</sup> The reason is that transferring the moving technique used in 1D to 2D leads to some complications. It is not clear in which dimension one should create space or close empty spaces on a lattice. Further, when creating or closing space in one direction, *all* other agents on that row would see their neighborhood altered.

<sup>6</sup> For what the two-dimensional model concerns, Schelling [1971a, 1971b, 1978] distinguishes preferences expressed in absolute terms (number of like or unlike agents within a neighborhood) or relative terms (ratio of like to unlike neighbors within a neighborhood), but in neither case specifies preferences over completely empty neighborhoods. We assume that empty neighborhoods are the least preferred, on which none of our results depend.

with  $a > 0$ ,  $50 \leq c \leq 100$ ,  $d \geq 0$ , and  $b \geq 0$ .

The first utility function that we consider is based on Schelling [1969, 1971a, 1971b, 1978], setting  $b=d=0$  and  $c=50$ . This utility function consists of two entirely flat pieces with a discrete drop in utility at the cut-off point  $c$  (see Figure 2a). Thus, an agent is indifferent between a neighborhood without any unlike agents and any neighborhood with up to 50% unlike neighbors. In particular, perfect integration is no better than complete segregation as long as an agent can live among his own type.

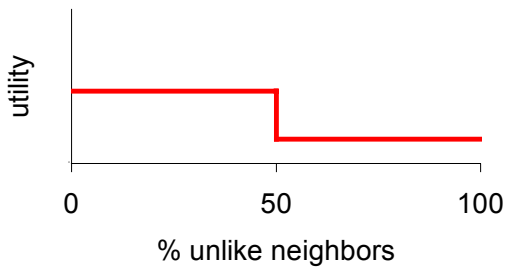
The first essential change we introduce concerns the parameter  $d$ , considering  $d > 0$ . This leads to a peaked utility function that has a peak at a 50-50 neighborhood, while retaining the flat part beyond the cut-off point at  $c=50$  (see the p50 utility function in Figure 2b). While such an agent still has an aversion against being in a minority, he now has a strict preference for perfect integration, preferring this to any majority of like agents.

We also consider a single-peaked utility function without any cut-off point ( $c=100$ ). That is, such an agent has no bias in favor of like agents at all. He strictly prefers living in a perfectly integrated neighborhood, but any neighborhood with  $x\%$  like or  $x\%$  unlike neighbors is equally good. This is represented by the perfectly symmetric p100 utility function in Figure 2c.

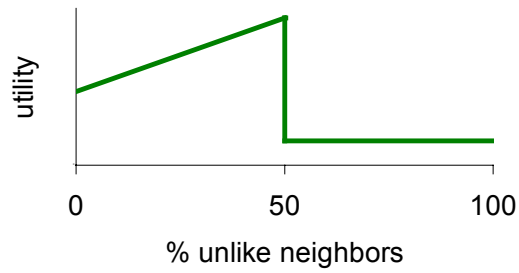
Finally, we consider parameter values of  $a=d=0$  and  $b > 0$ . The result is the spiked utility function shown in Figure 2d. These are strict preferences for perfect integration alongside indifference with respect to any not perfectly integrated neighborhood. In other words, these agents are driven exclusively by their obsession to live in a perfectly integrated neighborhood.



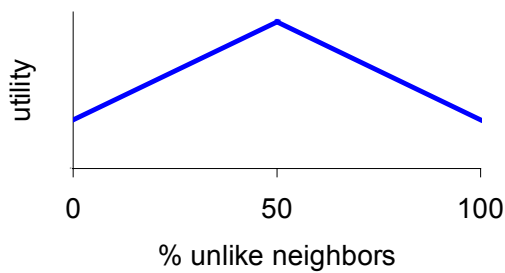
**Fig. 2a.** flat utility



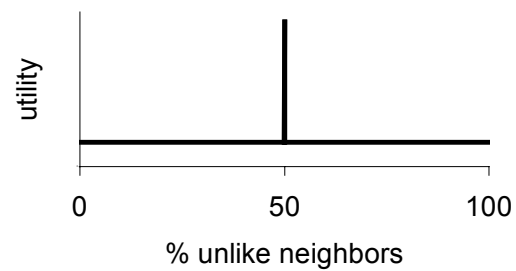
**Fig. 2b.** p50 utility



**Fig. 2c.** p100 utility



**Fig. 2d.** spiked utility



Our analysis is invariant to any positive monotonic transformation of these four utility functions, which form a logical order. The flat utility function is the weakest one with respect to preferences for integration. The p50 utility function retains the cut-off point, but introduces a strict preference for integration over living among like neighbors. The p100 utility function, in addition, removes the cut-off point. Finally, the spiked utility function emphasizes the strict preference for perfect integration by rendering an individual indifferent with respect to any other configuration.

Given the preferences, the behavioral assumption made by Schelling [1969, 1971a, 1971b, 1978] is that of myopic best-responses (BR). Schelling, however, also assumes inertia. That is, satisfied agents will always stay put, whereas it is not clear what happens with nonsatisfied agents who cannot find a satisfactory position. What is more, Schelling assumes that players move to the nearest satisfactory position. Both assumptions can be justified by the implicit assumption of a rather specific moving technology. The costs of moving would need to be strictly increasing with distance, but even for the largest possible distance they would need to be smaller than any possible strictly positive difference in utility between two locations. For a start, we abstract from this implicit assumption of moving costs, thus reducing the amount of structure imposed on the dynamic process, and focusing instead on the explicitly specified preferences. This means that we consider pure best-responses. In case of indifference between more than one position (possibly including his current position), a player chooses uniform randomly among all

positions offering the highest utility.<sup>7</sup> We will consider the issue of inertia separately at a later stage.

We also simplify the order of the moves. In Schelling [1969, 1971a, 1971b, 1978] all unsatisfied agents simultaneously put their name on a list, which is, then, processed sequentially in some arbitrary order, after which a new list is drawn, etc.<sup>8</sup> We simply select at each stage one agent uniform randomly, and ask him to choose a best-response. We will separately analyze the possibility that all agents simultaneously choose a best-response to the current state.<sup>9</sup>

Next, we need to specify the spatial setting. In 2D we focus on a standard board specification, a finite lattice, while in the 1D setup we concentrate on a ring. The reason to focus on a board instead of a torus is that tori do not seem to appear frequently in reality. The usual justification for considering a torus is that it is an approximation of an indefinitely extending two-dimensional space. It is unclear, though, that this is a meaningful approximation, especially when the underlying lattice is relatively small. Therefore, we prefer to stick to an ordinary board. No results depend crucially on this choice.

As to the one-dimensional setup, we focus on a ring for two reasons. First, a ring, unlike a torus, is relatively natural (e.g., in the form of a ring road in a city, chairs around a table, or the shoreline of an island or lake). Second, the positions near the boundary of a finite line have a huge impact on the existence of equilibria, which is an artefact caused by the shrinking size of a neighborhood for agents close to the edges. Schelling [1971a] explains that for  $k+k$  neighborhoods with odd  $k$ , alternating equilibria disappear. However, this is the case even for even  $k$ . What is more, with the peaked and spiked utility functions that we consider *all* (pure strategy) equilibria disappear for *any*  $k$ . Therefore, we prefer to consider a ring.<sup>10</sup>

For the sake of completeness, however, we will also consider a torus and a line, in the 2D and 1D setups respectively. Finally, in all setups analyzed, we consider neighborhoods defined in terms of the eight nearest surrounding positions (for agents in the interior).

---

<sup>7</sup> Hence, although we do not consider any noise per se, there are three sources of randomness in our model. These concern the initial allocations, the order of moves, and the way indifferences are solved.

<sup>8</sup> Notice that Schelling's specification does not seem very natural from a game-theoretic perspective: a currently satisfied agent might want to put his name on the list anyway, as he might no longer be satisfied at the moment his turn would come.

<sup>9</sup> In case of conflicting choices in the two-dimensional setup (two or more agents choosing the same location), we randomly allow one of these to be realized. We will, however, not consider the possibility of simultaneously choices in the one-dimensional setup, because of the conceptual difficulties arising from conflicting choices.

<sup>10</sup> Any *ad hoc* cures, e.g., modifying the utility function for agents near the edges, would rob the model of one of its major advantages: simplicity.

### 3. Analytical Tools and Benchmark Allocations

Given the model as specified in Section 2, we need to characterize the outcomes of the BR dynamics. To do so, in this subsection, we present two benchmark allocations, and define a number of measures to analyze the degree of segregation of these allocations.

#### 3.1 Benchmark allocations

In order to characterize the set of all possible steady-states of BR dynamics, we first introduce the concept of a Myopic Nash Equilibrium (MNE), describing all those configurations in which no agent can find a better location given the locations currently chosen by the other players. Notice that this equilibrium is myopic in the same sense as best-responses are myopic, as the dynamic structure of the game is disregarded.

*Definition 1:* Let  $Z$  be any configuration of location choices such that  $z_i \in Z$  is an element of the strategy space for each player  $i$ , and  $z_i \neq z_j$  for each  $i \neq j$ . Let  $V_i(Z)$  be the current payoff of agent  $i$ . Then  $Z^* = (z_i^*, Z_{-i}^*)$  is a Myopic Nash Equilibrium (MNE) iff  $V_i(z_i^*, Z_{-i}^*) \geq V_i(z_i, Z_{-i}^*)$  for each location  $z_i \notin Z_{-i}^*$  for each  $i$ . For strict inequality the MNE is strict, and it is weak otherwise.

The second benchmark is the set of all possible allocations, which in case of the 2D version we approximate by a sample of random allocations (i.e., agents scattered in an arbitrary order). This benchmark allows telling in which sense the MNE or the BR outcomes are out of the ordinary.

*Definition 2:* A random allocation is an allocation such that the probability that any given location is occupied by a particular type equals the ratio of the number of agents of this type to the number of possible locations.

#### 3.2 Segregation measures

Given these three types of allocations (BR outcomes, MNE, and random ones), we want to measure the degree of segregation in each of them. A number of slightly different measures will be used. These measures will be correlated to some extent, but they will each stress slightly different aspects of segregation. The emphasis is on segregation, rather than on some utility based measure, because this gives us an exogenous welfare criterion, allowing us to assess the consequences of different individual utility functions.<sup>11</sup>

*Clusters:* This measure counts the number of clusters that can be distinguished. Agent  $i$  and  $j$  belong to the same cluster if they are of the same type *and* they are, either directly or indirectly, linked laterally. Two agents are directly linked if they neighbor each other either horizontally or vertically. Moreover, if agent  $i$  belongs to the same cluster as agent  $j$ , and agent  $j$  to the same cluster as agent  $k$ , then agents  $i$  and  $k$  belong to the same cluster as well. An indirect lateral link goes through an uncontended zone of empty cells ('blanks'). Two blanks belong to the same zone if they are laterally linked, applying the same transitive relationship as above. Such a zone is contended if its neighbors, horizontally or vertically, are agents of different types, otherwise it is uncontended. Thus all blanks within a zone are either contended or not.<sup>12</sup> The one-dimensional version of the cluster measure is a straightforward simplification, as there are no blanks. The cluster measure, which is equivalent to the average cluster size, was used in Schelling [1969, 1971a, 1971b, 1978], without being formally defined. Notice that the cluster measure does not take into account how large individual clusters are, or how integrated (or not) agents within a cluster are.

*Switch rate:* Take the position of a given agent  $i$ , and make one full turn to observe all his neighbors. Let  $m_i$  be the number of  $i$ 's neighbors if it exceeds one, while it is zero otherwise. Let  $l_i$  be the number of switches, defined as the number of times that the type of a neighbor changes as we complete the turn, ignoring blanks. The switch rate, then, is  $\sum_i (l_i) / \sum_i (m_i)$ . The switch rate, unlike the cluster measure, cares about patterns. It measures how integrated patterns are, as seen by the individual agents.<sup>13</sup>

*Distance:* Let  $r_i$  be the minimal number of cells which need to be traveled by agent  $i$  (either laterally or diagonally) to reach an unlike agent, and  $t_i$  be the minimal number of cells to reach a like agent. The distance measure is:  $(1/N) \sum_i (r_i / t_i)$ , where  $N$  denotes the total number of agents.

*Mix deviation:* For a given agent  $i$ , let  $m_i$  be the absolute deviation from a 50-50 neighborhood:  $m_i = |0.50 - g_i / (g_i + f_i)|$ , where  $g_i$  is the number of like agents in agent  $i$ 's neighborhood (excluding the agent himself),  $f_i$  is the number of unlike agents, and  $m_i=0$  for agents with empty neighborhoods. The mix deviation measure, then, is:  $(1/N) \sum_i (m_i)$ . The mix deviation differs from the switch rate in two senses. First, for the mix deviation an ordered

---

<sup>11</sup> This approach would correspond to a government envisaging a desired outcome, such as multiculturalism or integration, and then implementing policies to induce particular individual attitudes.

<sup>12</sup> The extension of the measure to diagonal links is not straightforward, as one would need to define the concept of contended diagonal links, which can give rise to counterintuitive results. Moreover, considering diagonal links would only slightly decrease the number of clusters, as in many cases agents are already directly or indirectly linked laterally.

<sup>13</sup> Notice that agents with one or no neighbors are ignored, as no switches are possible for such neighborhoods. In the 1D setup the switch rate is equivalent to the clusters measure.

neighborhood such as XXXXOOOO would be the same as XOXOXOXO, whereas the switch rate would pick up the difference. Second, for the switch rate XXXOXXXX would be the same as XXOOOOXX, whereas the mix deviation would spot a difference.

*Share:* For a given agent  $i$ , let  $g_i$  and  $f_i$  be again the number of like and unlike agents respectively in his neighborhood. The share measure, then, is:  $\sum_i g_i / (\sum_i g_i + \sum_i f_i)$ , where agents with empty neighborhoods are ignored. The measure is based on Schelling [1969, 1971a]. A difference with the mix deviation is that it computes a weighted average of individual shares.

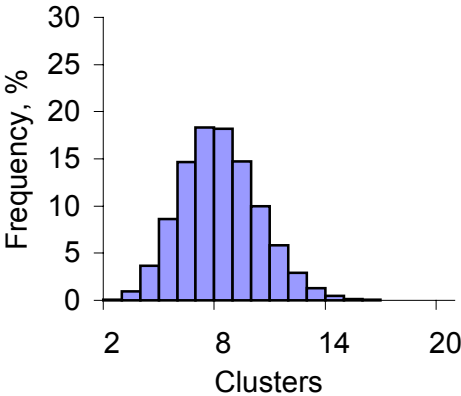
*Ghetto rate:* This measures the number of agents that lives in a neighborhood without any unlike neighbor. This measure, due to Schelling [1969], is a somewhat crude one as it treats having one unlike neighbor the same as having eight unlike neighbors.

### 4. Analysis of two-dimensional setup

#### 4.1 5x5 board

We start analyzing the model with a 5x5 board, with ten agents of each type, and five empty locations.<sup>14</sup> Figure 3 shows the relative frequency distribution of the clusters measure for one million random allocations. The distribution is rather symmetric, with the 20 agents forming on average 7.8 clusters.

Fig. 3. Random allocations, 5x5 board



<sup>14</sup> The reason to start with a 5x5 board is that we are able to characterize the complete set of MNE for this board. In the analysis of the 2D we always allocate 40% of the locations to each type, leaving 20% empty.

Table 1 shows the number of existing MNE for the flat and two peaked (p50 and p100) utility functions.<sup>15</sup> The number of MNE is of a similar order of magnitude for the flat and the p100 utility functions, whereas it is much lower for the p50 function.<sup>16</sup> Also, all MNE for the flat and p50 utility functions are non-strict, whereas almost 10% is strict with the p100 function.

**Table 1.** Number of existing MNE, 5x5 board

	<i>MNE</i>		
	<i>flat utility</i>	<i>p50</i>	<i>p100</i>
<i>non-strict MNE</i>	430,110	2880	351,472
<i>strict MNE</i>	0	0	36,482
<i>total MNE</i>	430,110	2880	387,954

Figures 4a to 4c show the frequency distribution of the cluster measure for the set of MNE for the flat and peaked utility functions. The MNE with the flat utility function are concentrated in the lower half of the range found for random allocations, with an average of 4.3 clusters per MNE. Notice that the set of MNE for the p50 utility function is a subset of the MNE for the flat utility function.<sup>17</sup> The question, then, is which MNE of Figure 4a will survive with the p50 function. Given the strict preference for perfect integration, one might conjecture that the subset of MNE will be more integrated. However, as Figure 4b shows, this turns out to be incorrect. The average of the 2880 MNE with the p50 utility function has 3.6 clusters, with a majority (64%) characterized by complete segregation. The distribution of the set of MNE for the p100 utility function looks similar to that for the flat utility function, with an average number of clusters of 4.0.

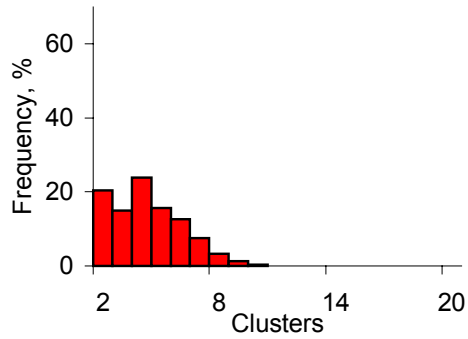
---

<sup>15</sup> Since for the 2D setup the findings for the spiked utility function are very similar to the p100 function, we omit them throughout. These results are available from the authors upon request.

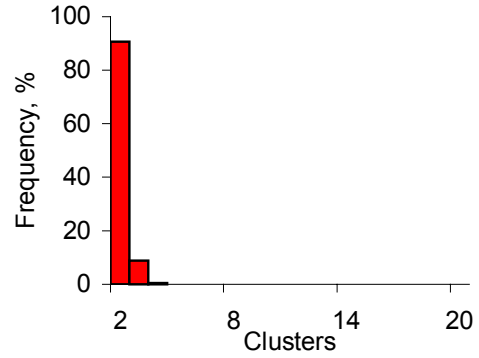
<sup>16</sup> When counting allocations, those obtained by swapping Os and Xs are regarded as being the same. Mirrored or rotated allocations are not grouped because, for example, the number of distinct rotations depends on the degree of symmetry of a particular allocation.

<sup>17</sup> Every MNE with the p50 function is also a MNE with the flat utility function. But the opposite is not true. Suppose an agent lives in a large-majority neighborhood while a perfectly integrated location is available. With a flat utility function this could be part of a MNE as such an agent would have no incentive to deviate. But an agent with the p50 utility function would deviate to that empty position.

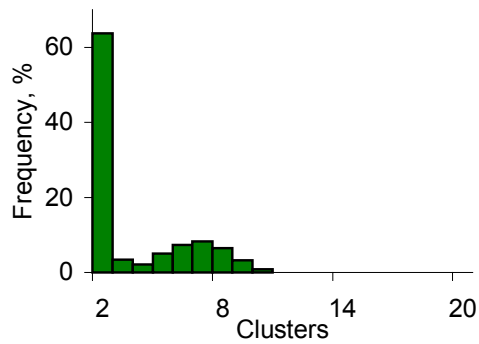
**Fig. 4a.** MNE, flat utility, 5x5 board



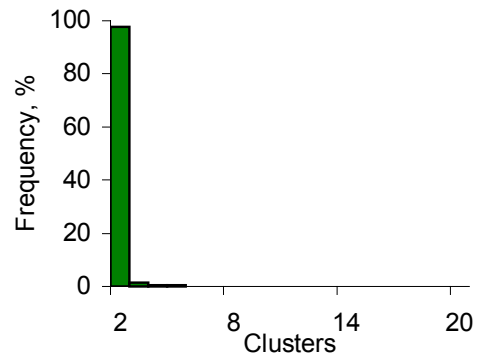
**Fig. 5a.** BR, 100,000 periods, flat utility, 5x5 board



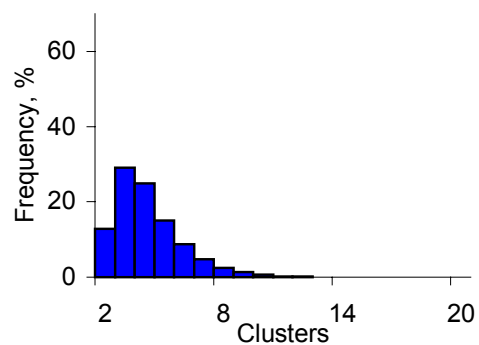
**Fig. 4b.** MNE, p50 utility, 5x5 board



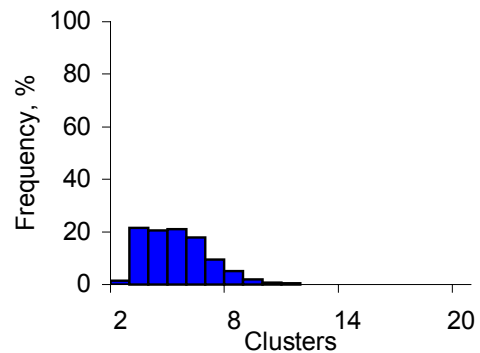
**Fig. 5b.** BR, 100,000 periods, p50 utility, 5x5 board



**Fig. 4c.** MNE, p100 utility, 5x5 board



**Fig. 5c.** BR, 100,000 periods, p100 utility, 5x5 board



Figures 5a to 5c show the outcomes of best-response (BR) dynamics for each of the three utility functions. Each diagram depicts the distribution of the cluster measure for 1000 runs after 100,000 periods.<sup>18</sup> In Figure 5a, for the flat utility function, 91% of the runs end in complete segregation, with 2.1 clusters on average. Comparing this with Figure 4a, we see that in as far as the BR dynamics reach a MNE, they favor a very specific subset of the set of potential steady states. As Figure 5b shows, with strict preferences for perfect integration, there is even more segregation. In 98% of the runs we observe complete segregation after 100,000 periods. The average is 2.0 clusters. Hence, again BR dynamics seem to home in on a biased subset of the set of potential steady states. This is not the case for the p100 utility function. As Figure 5c shows, the distribution of outcomes of the BR dynamics is not very different from the set of MNE. On average there are 5.0 clusters. Although complete segregation seems to be avoided, this still implies more segregation than with random allocations, where we observed on average 7.8 clusters.

Similar pictures can be plotted for other segregation measures. Table 2 summarizes this. For each of the measures there is hardly any difference between the outcomes of the BR dynamics with the flat and p50 function. While there are fewer clusters with the p50 than with the flat function, the other measures suggest there is slightly more segregation with the flat than with the p50 utility function.<sup>19</sup>

---

<sup>18</sup> A run corresponds to an independently executed BR sequence. A period is an instance when an agent is offered an option to move to a preferred location.

<sup>19</sup> The reason to emphasize the clusters measure is that it seems to capture the notion of segregation best at the intuitive level.



**Table 2.** Final distributions, 5x5 board

		<i>random</i>	<i>MNE</i>			<i>BR dynamics (100,000 periods)</i>		
			<i>flat</i>	<i>p50</i>	<i>p100</i>	<i>flat</i>	<i>p50</i>	<i>p100</i>
<i>Num. Obs.</i>		1,000,000	430,110	2880	387,954	1000	1000	1000
<b>Clusters</b>	<i>Average</i>	7.82	4.30	3.61	4.14	2.10	2.04	4.99
	<i>5%</i>	5	2	2	2	2	2	3
	<i>95%</i>	11	7	8	7	3	2	8
<b>Switch</b>	<i>Average</i>	0.53	0.35	0.31	0.43	0.21	0.23	0.51
	<i>5%</i>	0.40	0.19	0.19	0.28	0.16	0.19	0.42
	<i>95%</i>	0.65	0.52	0.52	0.60	0.27	0.28	0.62
<b>Distance</b>	<i>Average</i>	1.02	1.25	1.31	1.10	1.58	1.50	1.00
	<i>5%</i>	0.93	1.00	1.00	0.98	1.40	1.00	1.00
	<i>95%</i>	1.15	1.65	1.60	1.33	1.80	1.65	1.05
<b>Mix dev.</b>	<i>Average</i>	0.18	0.23	0.23	0.18	0.34	0.29	0.14
	<i>5%</i>	0.12	0.13	0.12	0.10	0.28	0.26	0.09
	<i>95%</i>	0.25	0.36	0.29	0.26	0.40	0.32	0.18
<b>Share</b>	<i>Average</i>	0.47	0.67	0.66	0.55	0.80	0.73	0.50
	<i>5%</i>	0.38	0.56	0.50	0.43	0.73	0.70	0.42
	<i>95%</i>	0.59	0.84	0.75	0.69	0.88	0.78	0.57
<b>Ghetto</b>	<i>Average</i>	1.06	4.89	5.76	2.07	10.49	8.76	0.08
	<i>5%</i>	0	1	1	0	8	6	0
	<i>95%</i>	4	12	10	6	13	10	1

Table 3 confirms the substantial number of MNE reached by the BR dynamics. Although the number of existing MNE is very small with the p50 utility function, BR dynamics lead to a MNE in 37% of cases. This is of the same order of magnitude as the number of MNE reached with the flat utility function (40%). With the p100 function we essentially always end up in a MNE, with most cases being a strict MNE. This shows that even non-strict MNE act as attractors.

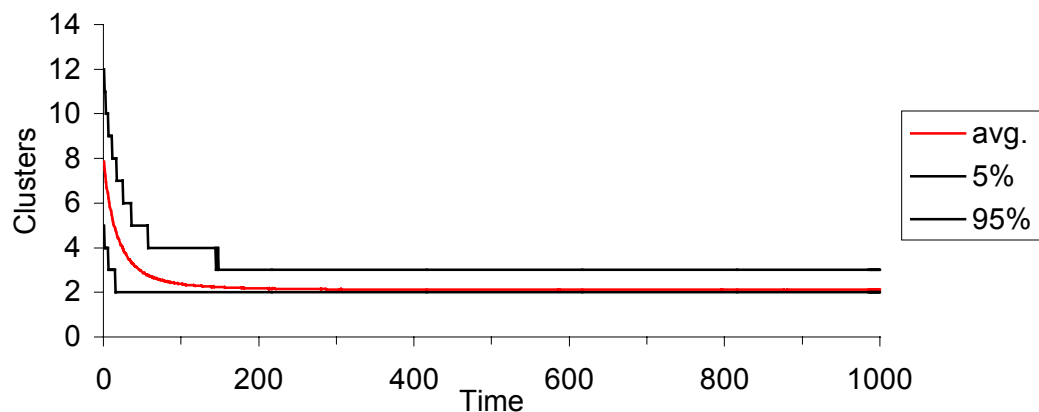
**Table 3.** Number of MNE reached, 5x5 board

	<i>BR dynamics (100,000 periods)</i>		
	<i>flat</i>	<i>p50</i>	<i>p100</i>
<i>Observations</i>	1000	1000	1000
<i>non-strict MNE</i>	404	370	145
<i>strict MNE</i>	n.a.	n.a.	854
<i>total MNE</i>	404	370	999

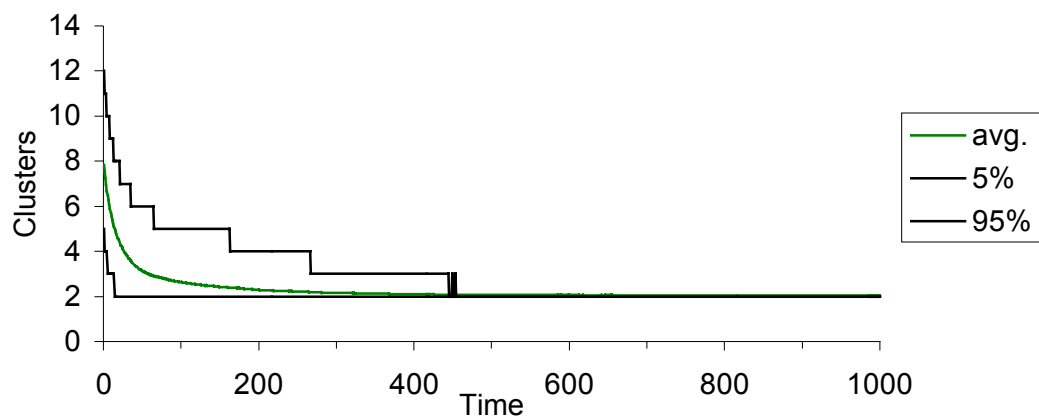
The numbers in Table 3 can be compared to the number of MNE one would expect in a sample of random allocations. Accounting for X/O symmetry, there are 4,908,043,140 possible allocations. This means that, given the number of MNE shown in Table 1, for the flat and p50 utility functions, where all MNE are non-strict, a sample of 1000 allocations most likely would contain no MNE. For the p100 function matters are slightly different, because some of the MNE are strict. Assuming that each of the 100,000 periods is a random draw, one would expect 524 MNE at the end of 1000 runs, with all of these MNE being strict. BR dynamics, however, give 854 strict and 145 non-strict MNE.

While looking at the last of 100,000 periods as evidence of the limiting behavior may be sensible, Figures 6a to 6c suggest that we could illustrate this by observing a considerably shorter spell of BR dynamics. The figures show the time series of the average cluster measure plus the 5th and 95th percentile for 10,000 runs. In each case the degree of segregation stabilizes within the first 1000 periods. With both the flat and the p50 utility functions average segregation is not only complete but also rather quick. With the perfectly symmetric p100 utility function segregation is substantial, but not extreme. Much of this segregation occurs relatively early on. In fact, initially the graph looks similar to those for the flat and c50 utility functions. Eventually, the 95th percentile is at 8 clusters, near the average of 7.8 clusters for random allocations. The corresponding graphs for the other segregation measures display a similar pattern.

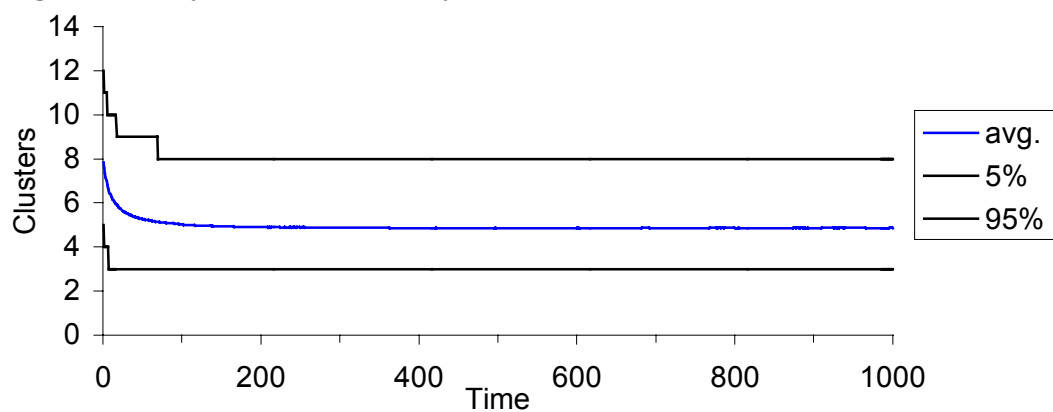
**Fig. 6a.** BR dynamics, flat utility, 5x5 board, 10,000 runs



**Fig. 6b.** BR dynamics, p50 utility, 5x5 board, 10,000 runs



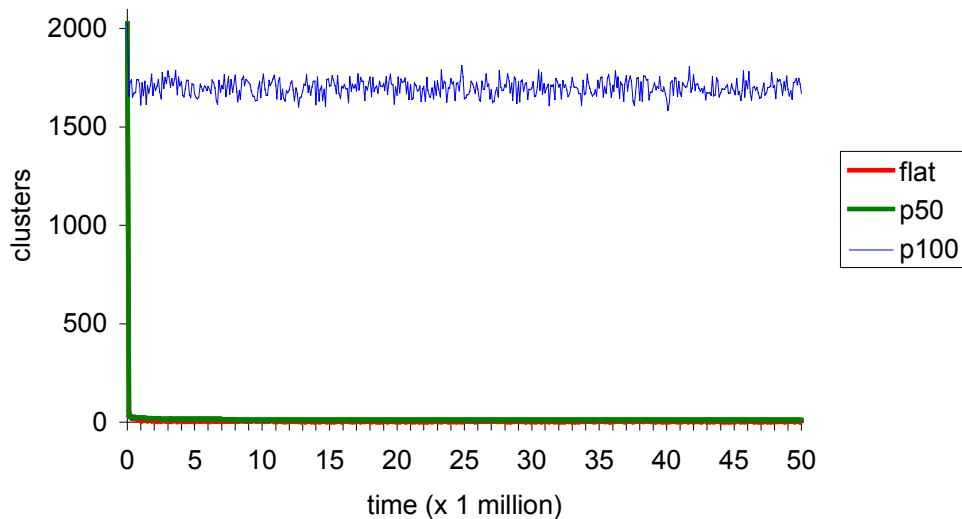
**Fig. 6c.** BR dynamics, p100 utility, 5x5 board, 10,000 runs



#### 4.2 100x100 board

We now turn to a larger board, with the same proportion of agents of each type as before. On a 100x100 board we have 4000 agents of each type, and 2000 empty locations. The distinction between a board and a torus is blurred as a board becomes bigger, because the share of agents at borders declines rapidly. Figure 7 plots the clusters segregation measure against time under BR moves for a 100x100 board. The three runs presented start from identical allocations with just over 2000 clusters, and differ only in the specification of the utility function. Each sequence is executed for 50 million periods, plotting every 100,000th data point.

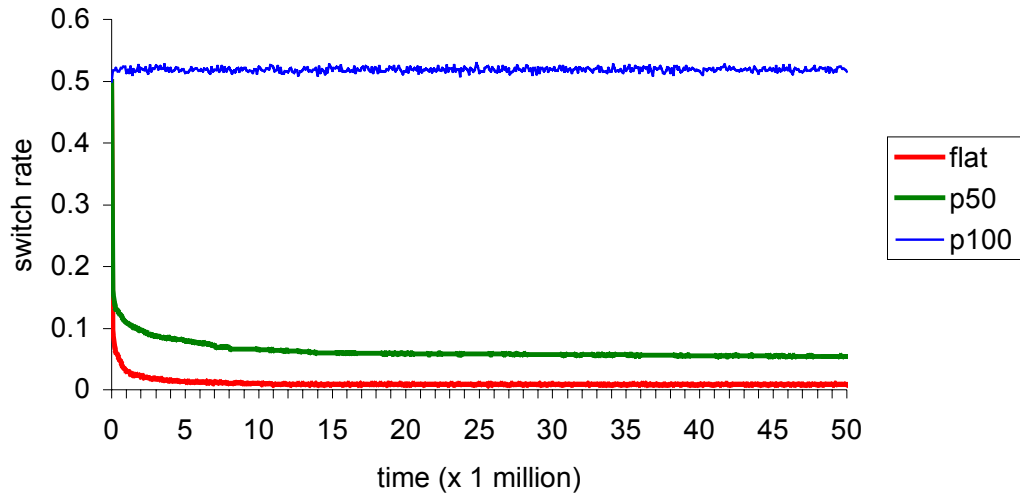
**Fig. 7.** BR dynamics, flat, p50 and p100 utility, 100x100 board, 1 run



While integration initially declines rapidly for each utility function, with the p100 function it reaches a level of about 1700, whereas with the flat and p50 utility functions almost complete segregation obtains. A close-up of the BR dynamics of the latter two functions would show that there are slight differences between the two. The average number of clusters is slightly higher with the p50 than with the flat utility function: 12 against 3.6 clusters respectively. Moreover, the number of clusters with the c50 function is rather steady, whereas with the flat function it fluctuates continuously, mainly between 2 and 6 clusters. These fluctuations with the flat utility function are due to the fact that, although there is basically complete segregation, all the time there are some very small clusters (usually less than a handful of agents) that happen to come temporarily loose at the border of the two ghettos, like water drops on the crest of a wave get loose from the sea, as their connecting neighbors move on. The switch measure is not sensitive to

such small clusters, and in addition it shows the difference between the flat and p50 function immediately. Figure 8 shows the switch rates for the same rates.

**Fig. 8.** BR dynamics, flat, p50 and p100 utility, 100x100 board, 1 run

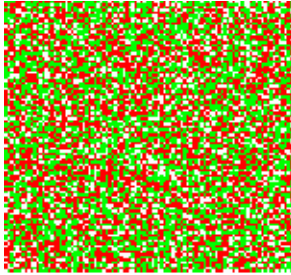


The switch rate seems virtually unchanged for the p100 function, but it drops rapidly for the flat and p50 functions, almost equally fast, until it hits a bend, somewhat earlier for the p50 function. The switch rate for the flat utility function becomes stationary quite early, whereas it continues to fall for the c50 function. Since the degree of segregation in the long run, according to the above diagrams, is rather stable, a look at a particular final allocation could be informative.

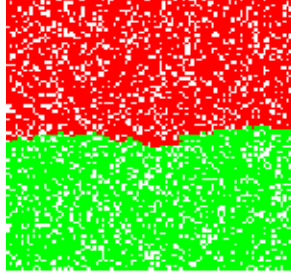
Figures 9a to 9c show a random board and two final allocations (after 50 million periods) for the flat and the c50 utility functions.<sup>20</sup> With the flat utility function, the boundary between the two colonies of distinct types is relatively straight and smooth, unlike that for the p50 function. This difference is due to the fact that with the flat utility function agents near the border of the two ghettos tend to drift to vacant positions inside their own ghetto, exposing their like neighbors left behind to a less familiar environment. With the p50 utility function, however, individuals are driven towards the borders by virtue of their strict preference for integration, rendering small clusters and highly curved boundaries more stable.

<sup>20</sup> The board for the p100 function looks very similar to a random one, just as Figures 7 and 8 predict, and is omitted here.

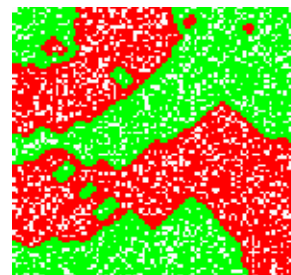
**Fig. 9a.** random allocation, 100x100 board



**Fig. 9b.** BR, 50m periods, flat utility, 100x100 board



**Fig. 9c.** BR, 50m periods, p50, 100x100 board



■ type X    ■ type O    □ blank

#### 4.3 Inertia, torus and simultaneous moves

As explained in Section 2, we decided to focus our analysis of the 2D version of Schelling's model on a board, with the players moving sequentially, without inertia. In this subsection we establish the robustness of the obtained results by introducing in turn inertia, a torus, and simultaneous moves.<sup>21</sup> For each of these three cases, the following two questions are addressed. First, was the assumption made by Schelling [1969, 1971a, 1971b, 1978] in this respect an essential element of his model of segregation? Second, does the introduction of inertia, simultaneous moves, or a torus change our findings concerning the relative effect of our utility functions with a strict preference for integration relative to Schelling's flat utility?

We first consider the effect of inertia. Schelling assumed inertia, but we removed it to avoid making implicit assumptions about further unspecified moving costs. The results presented above suggest that inertia was not essential in Schelling's model, as we still obtain segregation, also with Schelling's flat utility function. In fact, as can be seen in Table 4, with inertia one gets about the same degree of segregation on a 5x5 board as without inertia. The table also confirms that the introduction of inertia does not upset our findings thus far concerning peaked utility functions relative to the flat utility function. With inertia, the p50 function still leads to almost complete segregation (in more than 95% of the cases there are only 2 clusters), just as with the flat utility function, whereas the p100 function leads to significantly less segregation. This is consistent with our findings without inertia. In general, all measures, as well as the differences between the various utility functions, look very similar to those in Table 2.

---

<sup>21</sup> We also considered all combinations of these designs, but this did not lead to noteworthy new insights.

**Table 4.** Final distributions with inertia, 5x5 board

		<b>BR dynamics (100,000 periods)</b>		
		<b>flat</b>	<b>p50</b>	<b>p100</b>
	<i>Num. Obs.</i>	1000	1000	1000
<b>Clusters</b>	<i>Average</i>	2.36	2.08	4.97
	<i>5%</i>	2	2	3
	<i>95%</i>	4	2	8
<b>Switch</b>	<i>Average</i>	0.21	0.23	0.50
	<i>5%</i>	0.13	0.18	0.39
	<i>95%</i>	0.33	0.31	0.63
<b>Distance</b>	<i>Average</i>	1.59	1.44	1.02
	<i>5%</i>	1.28	1.20	1.00
	<i>95%</i>	1.85	1.60	1.10
<b>Mix dev.</b>	<i>Average</i>	0.34	0.28	0.14
	<i>5%</i>	0.24	0.26	0.09
	<i>95%</i>	0.41	0.30	0.21
<b>Share</b>	<i>Average</i>	0.80	0.73	0.50
	<i>5%</i>	0.73	0.71	0.42
	<i>95%</i>	0.88	0.76	0.60
<b>Ghetto</b>	<i>Average</i>	10.50	8.07	0.08
	<i>5%</i>	6	4	0
	<i>95%</i>	14	10	1

Table 5 shows that most BR dynamics ended up in a MNE after 100,000 periods. With the flat and p50 function agents always get stuck in a weak MNE, notwithstanding the relatively small number of MNE with the p50 function (see Table 1). Whereas with the p100 function we still reach a weak MNE in almost half of the times. While it is expected that BR would result in MNE in the presence of inertia (unless there are cycles with orbits bereft of MNE), our earlier finding that MNE can act as attractors even without inertia seems noteworthy.

**Table 5.** Number of MNE reached with inertia, 5x5 board

	<b>BR dynamics (100,000 periods)</b>		
	<b>flat</b>	<b>p50</b>	<b>p100</b>
<i>Observations</i>	1000	1000	1000
<i>non-strict MNE</i>	1000	1000	555
<i>strict MNE</i>	n.a.	n.a.	442
<i>total MNE</i>	1000	1000	997

As mentioned in Section 2, Schelling [1969, 1971a, 1971b, 1978] uses a somewhat special mechanism to determine the order of moves, combining simultaneous and sequential moves. That is, unsatisfied agents put their name simultaneously on a list, but this list is processed sequentially, with the agents making their location choice only when their turn comes. Up to now we focused on the simplest sequential move variant, but we now consider the opposite extreme: *simultaneous moves*. Each period, all agents simultaneously choose a best-response to the current

configuration. Then, whenever possible, these choices are carried out, and the procedure is repeated. In case of conflicting choices (one or more agents choosing the same location), one of them is selected randomly, with the other agent(s) staying put. The two questions to be considered are, first, whether the special order of moves is essential to Schelling's results, and second, whether our findings concerning the various utility functions considered would be upset by assuming a different way to determine the order of moves. Table 6 provides the answers to both questions.

**Table 6.** Final distributions with simultaneous moves, 5x5 board

		<i>BR dynamics (100,000 periods)</i>		
		<i>flat</i>	<i>p50</i>	<i>p100</i>
	<i>Num. Obs.</i>	1000	1000	1000
<b>Clusters</b>	<i>Average</i>	2.12	2.05	4.86
	<i>5%</i>	2	2	3
	<i>95%</i>	3	2	8
<b>Switch</b>	<i>Average</i>	0.21	0.22	0.50
	<i>5%</i>	0.16	0.17	0.42
	<i>95%</i>	0.27	0.26	0.61
<b>Distance</b>	<i>Average</i>	1.58	1.53	1.00
	<i>5%</i>	1.40	1.40	1.00
	<i>95%</i>	1.80	1.75	1.05
<b>Mix dev.</b>	<i>Average</i>	0.34	0.30	0.14
	<i>5%</i>	0.28	0.27	0.10
	<i>95%</i>	0.40	0.36	0.18
<b>Share</b>	<i>Average</i>	0.80	0.75	0.50
	<i>5%</i>	0.72	0.70	0.42
	<i>95%</i>	0.88	0.82	0.57
<b>Ghetto</b>	<i>Average</i>	10.50	9.35	0.10
	<i>5%</i>	8	8	0
	<i>95%</i>	13	12	1

Focusing on the flat utility function used by Schelling, the simultaneous moves setup yields the same degree of segregation as with sequential moves, and in both cases the outcome is similar to the results found by Schelling with his special combination of the two.<sup>22</sup> Furthermore, Table 6 also confirms that switching from sequential to simultaneous moves does not change the conclusions pertaining to the relative effect of the three utility functions. With simultaneous moves, p50 still leads to approximately the same degree of segregation as the flat utility function, whereas the p100 function still does not lead to much segregation.

<sup>22</sup> This seems to confirm Schelling's [1971a, 1971b, 1978] claim that while no exact rule for the order of moves was followed (he would typically simply proceed from left to right), such an order makes little difference for the general character of the outcome.



Table 7 shows that with the flat or p50 utility functions there are fewer MNE reached with simultaneous moves than with sequential moves (see Table 3). As Table 1 showed, almost all MNE are non-strict. With simultaneous moves every next allocation differs more from the previous one than with sequential moves, making it easier to break a weak MNE.

**Table 7.** Number of MNE reached with simultaneous moves, 5x5 board

	<i>BR dynamics (100,000 periods)</i>		
	<i>flat</i>	<i>p50</i>	<i>p100</i>
<i>Observations</i>	1000	1000	1000
<i>non-strict MNE</i>	328	171	149
<i>strict MNE</i>	n.a.	n.a.	851
<i>total MNE</i>	328	171	1000

Finally, we take a look at the effect of considering a torus instead of a board. Table 8 shows that this has a substantial effect on the set of MNE. In the case of a strict preference for integration there are fewer MNE on a torus than on a board. In particular, with the p50 utility function there are few MNE left, although ten strict ones appear.

**Table 8.** Number of existing MNE, 5x5 torus

	<i>MNE</i>		
	<i>flat utility</i>	<i>p50</i>	<i>p100</i>
<i>non-strict MNE</i>	463,730	200	75,550
<i>strict MNE</i>	0	10	20,170
<i>total MNE</i>	463,730	210	95,720

**Table 9.** Final distributions, 5x5 torus

		<i>random</i>	<i>MNE</i>			<i>BR dynamics (100,000 periods)</i>		
			<i>flat</i>	<i>p50</i>	<i>p100</i>	<i>flat</i>	<i>p50</i>	<i>p100</i>
<i>Num. Obs.</i>		1,000,000	463,730	210	95,720	1000	1000	1000
<b>Clusters</b>	<i>Average</i>	5.64	2.31	3.90	4.12	2.06	2.15	3.11
	<i>5%</i>	3	2	2	2	2	2	2
	<i>95%</i>	9	4	4	7	3	3	6
<b>Switch</b>	<i>Average</i>	0.53	0.32	0.38	0.51	0.30	0.32	0.46
	<i>5%</i>	0.41	0.27	0.37	0.41	0.26	0.29	0.36
	<i>95%</i>	0.64	0.44	0.37	0.62	0.35	0.38	0.57
<b>Distance</b>	<i>Average</i>	1.00	1.07	1.05	1.00	1.08	1.02	1.00
	<i>5%</i>	0.98	1.00	1.00	1.00	1.00	1.00	1.00
	<i>95%</i>	1.05	1.15	1.05	1.00	1.15	1.10	1.00
<b>Mix dev.</b>	<i>Average</i>	0.13	0.17	0.10	0.09	0.19	0.13	0.09
	<i>5%</i>	0.10	0.10	0.04	0.07	0.14	0.08	0.06
	<i>95%</i>	0.17	0.23	0.10	0.12	0.24	0.17	0.12
<b>Share</b>	<i>Average</i>	0.47	0.66	0.56	0.48	0.68	0.60	0.51
	<i>5%</i>	0.39	0.59	0.54	0.43	0.63	0.56	0.45
	<i>95%</i>	0.57	0.72	0.56	0.53	0.73	0.65	0.56
<b>Ghetto</b>	<i>Average</i>	0.08	1.31	0.95	0.00	1.62	0.50	0.00
	<i>5%</i>	0	0	0	0	0	0	0
	<i>95%</i>	1	3	1	0	3	2	0

Table 9 shows the segregation on a torus for the random and MNE allocations, as well as the outcomes of BR dynamics. Some measures appear sensitive to the change from a board to a torus. This applies in particular to the clusters and ghetto measures. Nevertheless, if we look at the outcomes of the BR dynamics with the flat utility function, we see that Schelling's assumption of a board was not essential for his results: populations on a torus become similarly segregated. Also our findings concerning the various utility functions are robust to the introduction of a torus. BR dynamics with the flat and p50 utility functions lead to segregation, but less so with the p100 function.

**Table 10.** Number of MNE reached, 5x5 torus

	<i>BR dynamics (100,000 periods)</i>		
	<i>flat</i>	<i>p50</i>	<i>p100</i>
<i>Observations</i>	1000	1000	1000
<i>non-strict MNE</i>	279	0	387
<i>strict MNE</i>	n.a.	4	610
<i>total MNE</i>	279	4	997

Table 10 describes how often a MNE was reached on a torus. The main difference with a board concerns the p50 utility function. As there are only a handful of MNE for that function, these MNE are very rarely reached.

#### 4.4 Discussion of findings in 2D setup

Our analysis of the 2D setup shows that Schelling's results are not only robust to a class of alternative specifications, but they can also be strengthened enormously. The simple model characterized by sequential moves (in a random order) in the absence of inertia exhibits rapid segregation, even when the flat utility function is strengthened to imply strict preferences for perfect segregation. The findings obtained by Schelling and the extensions presented in this section are preserved in eight possible designs regardless of the nature of the order of moves (sequential or simultaneous), the structure of the lattice (a board or a torus) or the presence of inertia. Thus Schelling's original assumptions on inertia and the specific way in which the order of moves was determined proved to be inessential.

While the strict preference for perfect integration (as with the p50 function) leads to approximately the same amount of segregation as the flat utility function, segregation for the p100 function is not as stark as for the flat and p50 utility functions. The essential difference is the asymmetry of the latter two. With Schelling's flat utility function there are two separate effects of this asymmetry. First, in case of indifference between a range of satisfactory positions, agents, on average, would choose a relatively segregated option. This implies a 'random drift' away from integration. Second, in the case of facing a choice between a small minority location and a large majority location (i.e., positions with either  $x\%$  unlike or  $x\%$  like, where  $x$  is greater than the cut-off point), agents favor the latter. With the p50 function, the flat part favoring the drift to segregation has been removed, while the cut-off point is retained, whereas with the p100 function the cut-off point has been removed as well. Since the p100 function does not induce substantial segregation, this further helps to pin down the essential element explaining segregation. It is the asymmetry related to the cut-off point, i.e., the fact that an agent favors his own ghetto over an unlike ghetto, that is the crucial element in the 2D setup.

The concept of a Myopic Nash Equilibrium (MNE) plays a key role in the analysis, as the BR dynamics often act as an equilibrium selection device, even in the absence of inertia. Best-responses eliminate attractive locations,<sup>23</sup> thus reducing incentives to deviate, and increasing chances to encounter a MNE. Further, MNE tend to be clustered together, as it is easy to obtain one MNE from another by moving indifferent agents. Consequently, notwithstanding their tiny share in the total number of possible allocations, MNE states tend to be quite persistent. While the set of MNE itself is already biased towards segregation, especially for the p50 function, the BR dynamics favor the most segregated among them.

---

<sup>23</sup> This begs the question whether the chosen population density (40% of each type of agents) is critical to the behavior of the system, as it influences the availability of choice locations. As the sensitivity analysis presented in Appendix A shows, this is not the case.

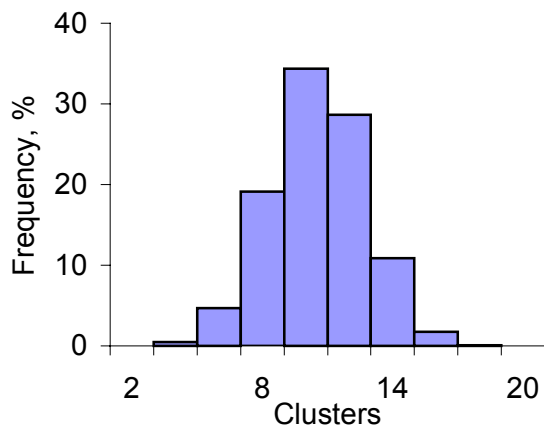
## 5. Analysis of one-dimensional setup

We now turn to an analysis of the 1D version of Schelling's spatial proximity model. As we saw above, in the 2D setup a considerable role was played by the MNE. Since the existence of MNE cannot be ensured on a line, in what follows the 1D space is assumed to be a ring.

### 5.1 10+10 ring

For a start, and comparison with the 2D model, we look at a ring with ten agents of each type, and a neighborhood formed by eight neighbors (four to the left, and four to the right). Figure 10 shows the distribution of the cluster measure for all 92,378 possible allocations.<sup>24</sup> On average the 20 agents are located in 10.5 clusters.

**Fig. 10.** All possible allocations,  
10+10 ring



The other benchmark we use is the set of MNE. Table 11 shows the number and type of MNE for each of the utility functions. The MNE are the same for the two peaked (p50 and p100) and spiked utility functions, forming a subset of the set of MNE for the flat utility function. The table also indicates that there are no strict MNE. That the number of MNE is much lower than in the 1D setup is, in part, due to the fact that there are fewer different allocations possible (since there are no empty spaces), and, in part, because the restrictions on equilibrium allocations are much more stringent. The formal analysis of these points is deferred until Section 5.2.

---

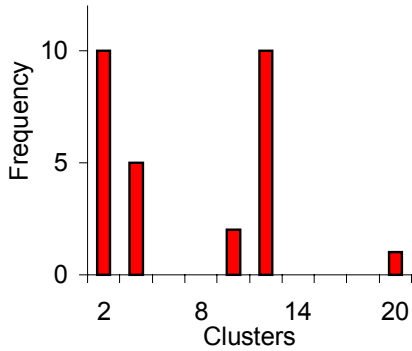
<sup>24</sup> Again, we neglect equivalent allocations obtained by swapping Os and Xs.

**Table 11.** Number of existing MNE, 10+10 ring

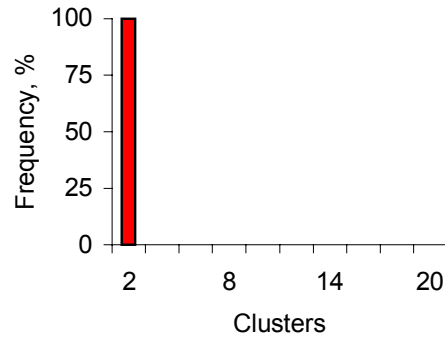
	<i>MNE</i>	
	<i>flat utility</i>	<i>peaked or spiked</i>
<i>non-strict MNE</i>	28	18
<i>strict MNE</i>	0	0
<i>total MNE</i>	28	18

Figures 11a and 11b show the distribution of the cluster measure for the MNE for each of the utility functions considered. MNE with the flat utility range from complete segregation (2 clusters) to perfect integration (20 clusters). This time, however, the intuition that the subset of MNE with a strict preference for perfect integration is characterized by more integration than the set of MNE with the flat utility function is correct. All completely segregated MNE of the flat utility function disappear with the peaked and spiked utility functions. The average number of clusters per MNE with the flat utility function is 7.1, and with the peaked or spiked utility functions it is 10.0.

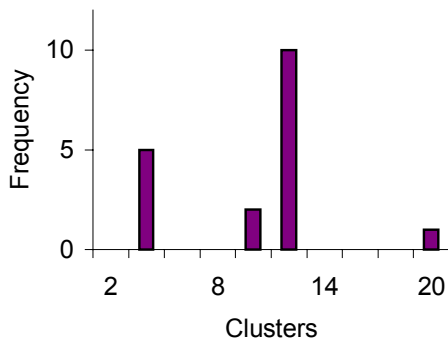
**Fig. 11a.** MNE, flat utility, 10+10 ring



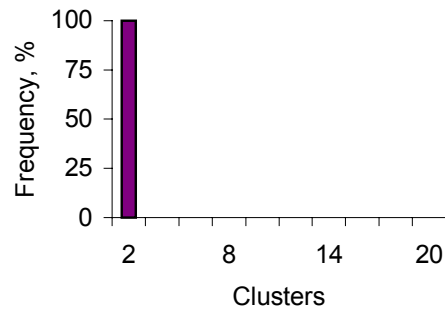
**Fig. 12a.** BR, 100,000 periods, flat utility, 10+10 ring



**Fig. 11b.** MNE, peaked or spiked utility, 10+10 ring



**Fig. 12b.** BR, 100,000 periods, peaked or spiked utility, 10+10 ring



Figures 12a and 12b show the outcomes for 1000 observations of the BR sequences run for 100,000 periods. For each of the utility functions considered (even for the spiked one), the BR dynamics invariably lead to complete segregation. For the flat utility function one could interpret this as the dynamics selecting a particular MNE. However, with any of the peaked or spiked utility functions there is no MNE corresponding to this outcome.<sup>25</sup> Hence, Schelling's [1971a] observation that "*(w)e could have surmised that our rules of movement would lead to equilibria*" (p. 151) is correct for the flat utility function, but it is not generally true for these rules of movement. In particular, the conjecture is not true for the class of peaked or spiked utility functions that we consider. Table 12 confirms this. Although this indicates that BR dynamics do

<sup>25</sup> As argued below, the state of complete segregation is almost the extreme opposite of any of the MNE. It is not true, however, that complete segregation is a state with the lowest possible utility. For example, with the spiked utility function, alternating clusters of size  $(k-1)$  would make all agents unsatisfied.

not simply act as an equilibrium selection device, the table also shows that BR dynamics with the flat utility function invariably end up in a (non-strict) MNE.

Since complete segregation is not a MNE with the peaked or spiked utility functions, it cannot be a steady-state allocation. Nevertheless, the pattern of complete segregation is stable. With complete segregation, and any of the peaked or spiked utility functions, only the agents at the border of their own ghetto will enjoy the ‘bliss’ level of utility. Each time an agent can make a move, he will locate himself exactly at such a boundary. Hence, almost all the time the agent whose turn has come moves to a better location, but the configuration of the two ghettos as such is stable. These ghettos only move around on the ring.

**Table 12.** Number of MNE reached, 10+10 ring

	<b>BR dynamics (100,000 periods)</b>	
	<i>flat</i>	<i>peaked or spiked</i>
<i>Observations</i>	1000	1000
<i>non-strict MNE</i>	1000	0
<i>strict MNE</i>	n.a.	n.a.
<i>total MNE</i>	1000	0

Table 13 summarizes the characterization all possible allocations as well as those of the MNE and the BR outcomes using the various segregation measures. MNE for the peaked and spiked utility functions are perfectly integrated according to the mix deviation, share and ghetto measures,<sup>26</sup> while according to the cluster and distance measures they are as integrated as the average random allocation. In each of the MNE with the peaked or spiked utility functions, all agents live in a 50-50 neighborhood, reaching maximum utility. This implies that the adjustment dynamics are an important concern from a social welfare point of view. In sharp contrast to the perfectly satisfactory MNE, BR dynamics always lead to complete segregation, where only four agents reach maximum utility.

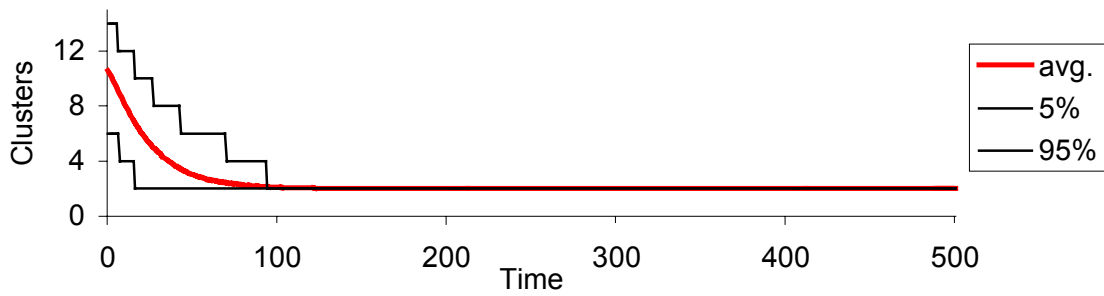
---

<sup>26</sup> Although this cannot be read from the table, this applies in fact to each MNE for the peaked and spiked utility functions.

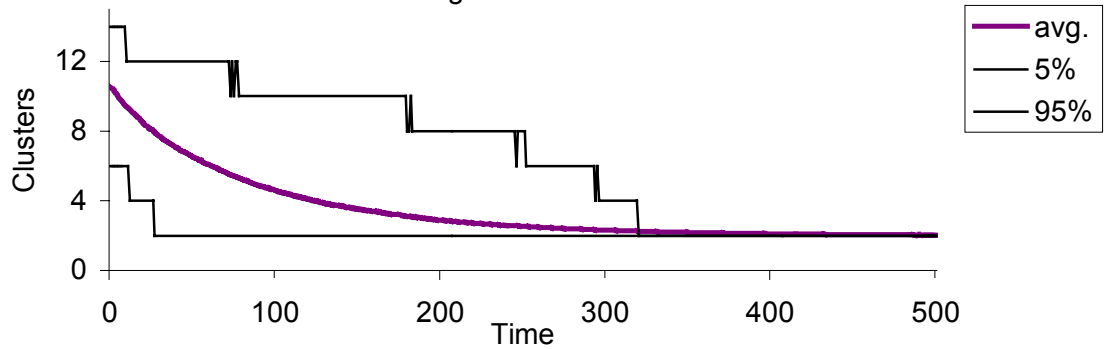
**Table 13.** Final distributions, 10+10 ring

		all	MNE		BR dynamics (100,000 periods)	
			flat	peaked or spiked	flat	peaked or spiked
Num. Obs.		92,378	28	18	1000	1000
<b>Clusters</b>	Average	10.53	7.14	10.00	2.00	2.00
	5%	6	2	4	2	2
	95%	14	12	12	2	2
<b>Distance</b>	Average	1.10	1.82	1.16	3.00	3.00
	5%	0.82	0.50	0.50	3.00	3.00
	95%	1.51	3.00	1.80	3.00	3.00
<b>Mix dev.</b>	Average	0.11	0.09	0.00	0.25	0.25
	5%	0.05	0.00	0.00	0.25	0.25
	95%	0.18	0.25	0.00	0.25	0.25
<b>Share</b>	Average	0.47	0.59	0.50	0.75	0.75
	5%	0.43	0.50	0.50	0.75	0.75
	95%	0.55	0.75	0.50	0.75	0.75
<b>Ghetto</b>	Average	0.00	1.43	0.00	4.00	4.00
	5%	0	0	0	4	4
	95%	0	4	0	4	4

**Fig. 13a.** BR dynamics, flat utility, 10,000 runs, 10+10 ring



**Fig. 13b.** BR dynamics, peaked or spiked utility, 10,000 runs, 10+10 ring





Figures 13a and 13b show how these final distributions tend to be approached over time. The average cluster measure is shown alongside the 5th and 95th percentiles for 10,000 runs for each of the utility functions, starting from identical initial configurations. The graphs for the peaked (p50 and p100) and spiked utility functions are identical. These series demonstrate that convergence to complete segregation occurs for each of the four utility specifications, but it happens faster with the flat utility function.

## 5.2 Formal analysis

Having shown the emergence of complete segregation with BR dynamics for each of the utility functions considered in the case of a 10+10 ring with a 4+4 neighborhood, we now present some formal results applicable to BR dynamics for any ring size, and any neighborhood size. The results show that some of the striking features reported above are not coincidental. It is convenient to assume that the number of Xs equals the number of Os.

Proposition B1 shows that in the long run complete segregation is the only possible outcome of best-response dynamics with the spiked utility function. Corollary 1 explains that the same applies to the flat and peaked (p50 and 100) utility functions.

**PROPOSITION B1.** If a neighborhood on a circle of size  $2m$  is defined as  $k$  neighbors to the left and  $k$  neighbors to the right, then, if  $m$  is the number of each type,  $m > k$ , and the utility function is spiked, then the process of best-responses has a unique recurrent state consisting of all completely segregated states.

**DISCUSSION.** The complete proof can be found in Appendix B. Here is a sketch. To prove that complete segregation is the limit outcome of the BR dynamics it should be possible to reach complete segregation from each initial configuration with positive probability, while all completely segregated states constitute a single recurrent state. On a completely segregated ring, there are two borders between ghettos. These borders are the only locations that offer positive utility, and this protects integrity of the ghettos. Thus, it is sufficient to offer an algorithm showing just one possible path leading from each allocation to complete segregation. In the proof this is done in two steps. A ‘seed’ is defined as a segment of a ring formed by  $k$  agents of one type followed by  $k$  agents of the other type. If an allocation contains a seed, then construction of complete segregation by means of positive probability moves is trivial, as this seed can always grow by adding agents at the border inside the seed. If no seed is present initially, it will eventually emerge whatever the initial allocation.

**COROLLARY 1.** Proposition B1 also holds for the flat and peaked (p50 and p100) utility functions.

PROOF. Any best-response move according to the spiked utility function would also be a best-response according to the other utility functions. This claim, in turn, relies on the fact that there always exists a location offering the bliss level of utility. That it is indeed so is established in the course of the proof of Proposition B1 (see Appendix B).  $\square$

We now turn to a characterization of the set of MNE for any ring size. For the flow of exposition the number of a proposition is indicated in parentheses after a claim. Its formal statement and a proof are in the appendices.

On a ring there exist at least two positions where an agent can insert itself and enjoy perfect integration (Lemmas B2 and C1). Then it follows that in a MNE all agents necessarily enjoy the highest possible utility (Proposition C1), for otherwise someone would be willing to migrate to one of the perfectly integrated locations. Thus MNE are the only Pareto efficient outcomes. This is so for any of our utility specifications, because they all assign the highest possible utility to living in a perfectly integrated neighborhood. Thus the welfare implication drawn for the special case above on the basis of a numerical analysis is a general feature of the 1D model. For all utility functions considered other than the flat one, BR dynamics lead to a Pareto inferior outcome. Pareto efficiency of MNE immediately explains why, with a strict preference for integration, according to the mix, share and ghetto measures, equilibria were characterized as perfectly integrated.

The absence of strict equilibria is also a generic feature (Proposition C2), and it follows from the multiplicity of bliss locations. This claim can also be inferred from the fact that complete segregation is the only feasible long run outcome of BR dynamics. Otherwise the BR dynamics could have become stuck at one of the strict MNE, which would have destroyed the result.

In the 2D case, the set of MNE was biased towards segregation. The opposite is true in the 1D case: in any MNE, no cluster can exceed  $k + 1$  agents in size with the peaked or spiked utility functions, where  $k$  is the size of a neighborhood in either direction (Proposition C3). For a bigger cluster size, agents in the middle of the cluster would enjoy less than the bliss level of utility. But then, this cluster would not be a part of a MNE allocation, because an equilibrium requires that all agents enjoy the highest possible utility. Thus the BR outcome has nothing to do with MNE but for the flat function, when complete segregation happens to be an equilibrium.

As already mentioned, no MNE exist on a line for the peaked and spiked functions (Proposition D1). Equilibria on a line, however, do exist for the flat utility function. Additionally, with the peaked or spiked utility functions, no MNE exist if the number of agents of each type is different (Proposition C6). This follows from the requirement that all agents need to be satisfied.

The sets of MNE for the spiked and two peaked utility functions coincide (Proposition C7). This set is a subset of MNE with the flat utility function (Proposition C8). Extra equilibria in the case of the flat function come from the absence of the upper bound on an admissible cluster size, so that even complete segregation becomes an equilibrium. Logic similar to that of Proposition C7 shows that BR moves for all utility specifications with a strict preference for integration are identical.

The following property allows to construct the set of MNE for a spiked (and hence peaked) utility function. If the neighborhood parameter  $k$  is odd, then, in and only in a MNE, agents at locations  $i$  and  $i+k+1$  are of opposite types (Proposition C4). If  $k$  is even, then either agents  $i$  and  $i+k+1$  are of opposite types or agents  $i$  and  $i+k$  are of the same type and  $a_i + a_{i+1} + \dots + a_{i+k-2} + a_{i+k-1} = 0$ , where  $a_j=1$  if an agent in the  $j$ th position is X and  $a_j=-1$  otherwise (Proposition C5). The above properties of MNE imply that many rings, depending on their length, will not have the full set of potential MNE given  $k$ . The perfectly integrated MNE are robust to the length of a ring.<sup>27</sup>

If  $k=4$ , as it is throughout this Section 5, then all equilibria described in Proposition C5, can be summarized by the following five patterns: XXXXXOOOOO, XXXOXOOOXO, XXOXXOOXOO, XXOO and XO. These different patterns cannot be combined; only shifting and concatenation of single patterns are allowed. Accounting for X/O symmetry, the above patterns give rise to 5, 5, 5, 2 and 1 MNE respectively. These 18 equilibria are the only ones possible for the spiked and peaked utility functions, whatever the size of a ring, whereas the number of MNE increases with the size of a ring for the flat utility function.<sup>28</sup>

### 5.3 100+100 ring

The formal analysis in Section 5.2 allows us to characterize the set of MNE, and we also know that complete segregation is the only possible long-run outcome.<sup>29</sup> Proposition B1 does not, however, say how soon this happens given the size of a ring, nor does it imply anything about the relative speed of convergence for the various utility specifications. Therefore, we now

---

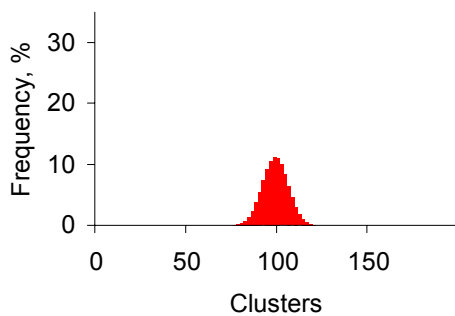
<sup>27</sup> For instance, if we took an 11+11 ring instead of a 10+10 one, there would be only one MNE, characterized by perfect integration, with 22 clusters, which would make the BR outcome of complete segregation even more striking.

<sup>28</sup> For  $m=100$  agents of each type and  $k=4$ , there are only 18 MNE for the peaked or spiked utility functions, but 60,575,676,973,999,910,976,213 for the flat utility function. The latter comprises  $1.34 \cdot 10^{-34}$ % of all possible allocations. The share declines rapidly in  $m$ : for  $m=10, 20, 30, 40$  and  $50$  the corresponding shares are  $0.03, 6 \cdot 10^{-6}, 1.5 \cdot 10^{-9}, 4 \cdot 10^{-13}$  and  $10^{-16}$ %.

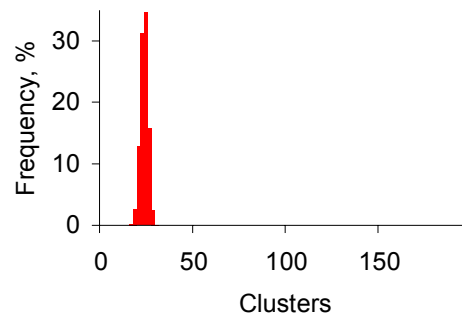
<sup>29</sup> For a modification of Schelling's 1D model, with agents swapping locations and in the presence of noise, Young [1998] also argues that complete segregation (which is a MNE) is the only viable outcome of the BR dynamics, as long as the agents' preferences are biased in favor of their own type.

examine the dynamics on a ring with 100 agents of each type ( $m=100$ ) and a neighborhood defined by eight neighbors ( $k=4$ ). The formal analysis of the previous section allows to characterize the set of MNE. By Proposition C3, in any MNE with the peaked or spiked utility functions there are at least  $2m/(k+1) = 40$  clusters. With the flat utility function, in addition, there is also a relatively segregated range of MNE with numbers of clusters between 2 and 40. In fact, the eighteen integrated MNE inherited from the peaked and spiked utility functions are negligible compared to the over 60 sextillion relatively segregated MNE inherent to the flat utility function. Figures 14a and 14b depict the distribution of the cluster measure for the set of  $4.5 \cdot 10^{58}$  possible allocations and its subset of  $6 \cdot 10^{22}$  MNE for the flat utility function respectively. The distribution for eighteen MNE with strict preferences is identical to that in Figure 10b once the horizontal axis is scaled up by a factor 10, and is therefore not shown here.

**Fig. 14a.** All possible allocations, 100+100 ring



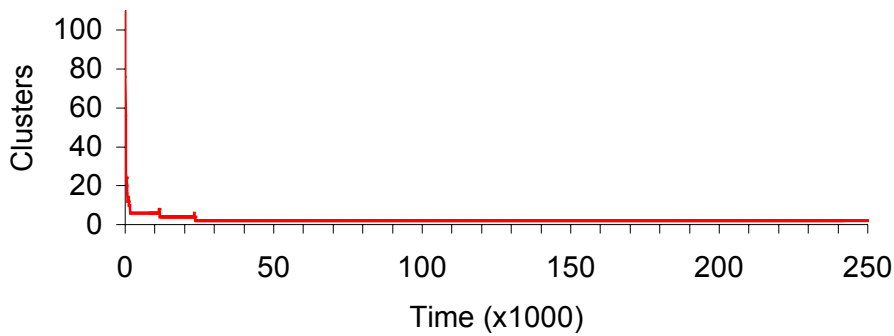
**Fig. 14b.** MNE, flat utility, 100+100 ring



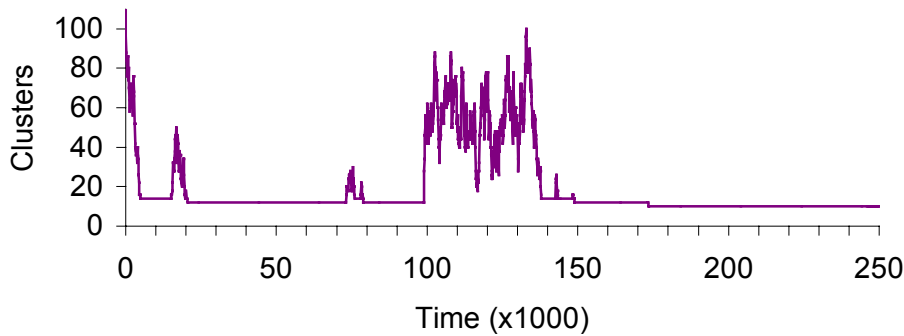
Since complete segregation is the only possible long-run outcome, it is instructive to look at the number of clusters over time under BR dynamics with the various utility functions. Figure 15a shows the first 250,000 periods of a single run of 50 million periods, while Figure 15b depicts the same for the peaked and spiked utility functions.<sup>30</sup> The random initial allocation, the same for each utility specification, has 110 clusters. After 50 million periods with the peaked or spiked utility function, a level of six clusters is reached, whereas complete segregation occurs within 25 thousand periods with the flat utility function. Moreover, while the flat utility function leads to a steady decline in the number of clusters, the degree of segregation is much more erratic for the peaked and spiked preferences.

<sup>30</sup> As argued in the previous section, the BR dynamics are identical for each of the peaked or spiked functions considered in this paper.

**Fig. 15a.** BR dynamics, flat utility, 1 run,  
100+100 ring



**Fig. 15b.** BR dynamics, peaked or spiked utility,  
1 run, 100+100 ring



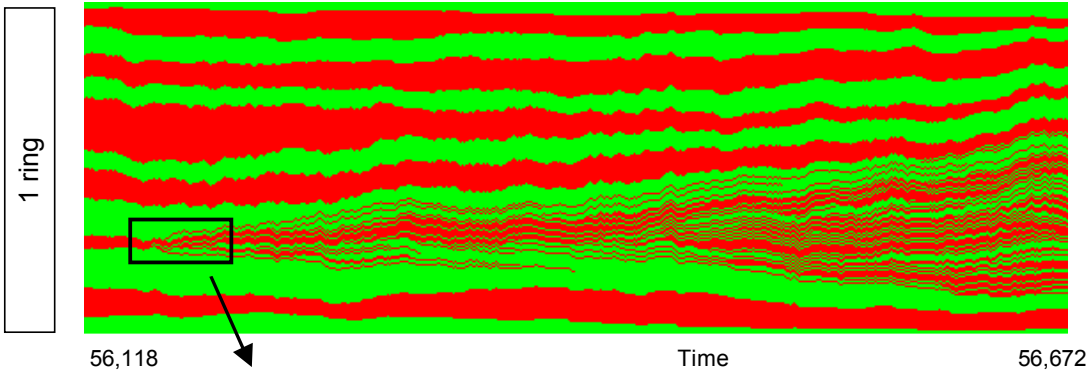
With the peaked utility functions, for the first few hundred periods there are around 80 clusters, then for 3000 periods the number of clusters oscillates between 60 and 70. Suddenly, a decline to around 35 clusters follows. It stays at that level for a thousand periods, followed by a steep fall to 14 clusters – all this in less than 5000 periods. Then a long spell of relative stability follows, with just two brief medium-sized outbursts of integration. Subsequently, after almost 100,000 periods with only twelve clusters, the number of clusters soars abruptly to 75. The number of clusters is volatile for some time, and even reaches a level of 100, which corresponds to the level of integration for a random allocation. That is, all endogenous segregation has disappeared. However, as sudden as this integration had burst on to the scene, it disappears again. The higher the level of segregation, the less likely are further outbursts. These sudden outbursts of integration do not occur in all runs. But they do occur with peaked utility functions, whereas they never seem to occur with the flat utility function.

Thus, the predicted limit result is obtained with the flat utility function, whereas the peaked and spiked utility functions lead to remarkably extreme segregation in finite time. This shows that strict preferences for integration ensure not only that, if segregation is reached, it is

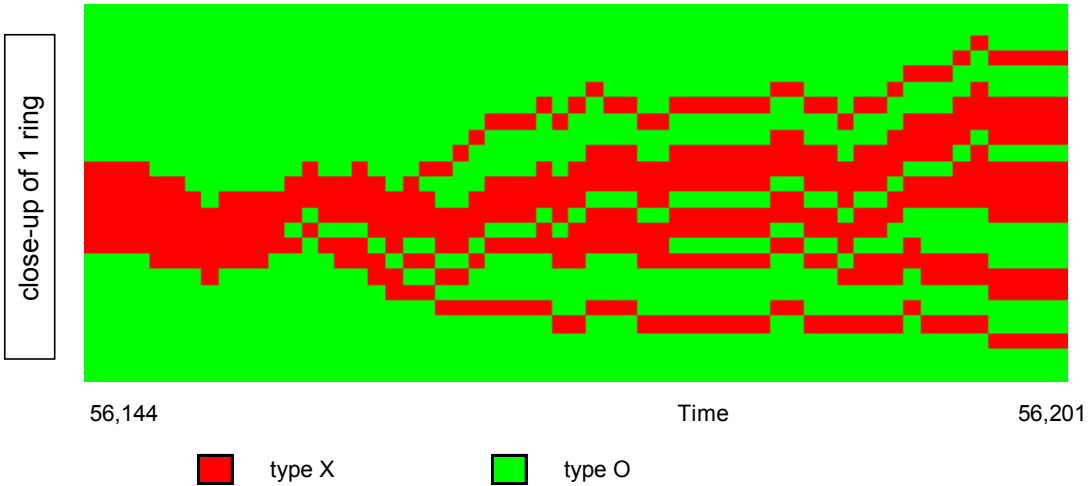
permanent, but also that they promote a consistent (though possibly volatile) drive towards segregation.

To understand the sudden outbursts of integration with peaked or spiked preferences, Figure 16a shows 554 periods, taken after about 56 thousand periods of another arbitrary run. One vertical slice of the graph represents a single ring with two types of agents. The same ring next period is placed directly to its right. At the start of these plotted ‘panel data’, the size of all clusters exceeds  $k=4$ . As clusters fluctuate in size in a random manner, sooner or later one of the clusters reaches the size of  $k$ . From that moment on, an agent of the other type can jump into the middle, effectively creating two extra clusters, and offering other agents as well the opportunity to insert themselves. Figure 16b offers a close-up of this process. As they do so, this increases the number of optimal locations within the growing ‘mushroom’, draining other clusters. Eventually this integration disappears.

**Fig. 16a.** BR dynamics, peaked or spiked utility function, 1 run, 100+100 ring

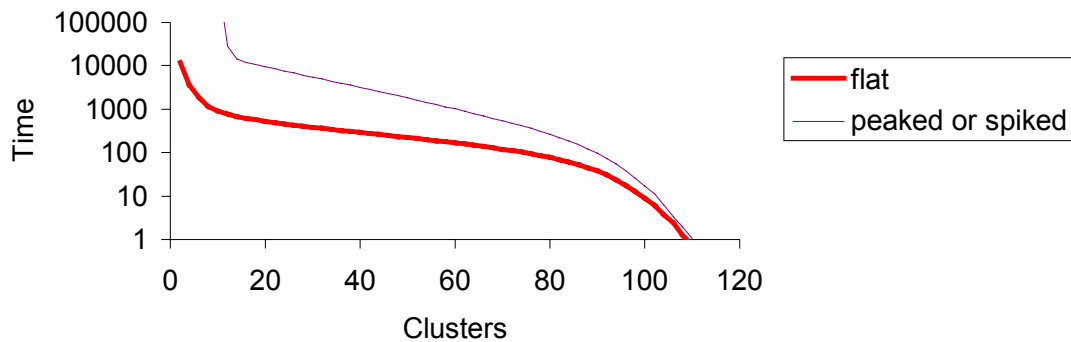


**Fig. 16b.** Close-up of BR dynamics, peaked or spiked utility function, 1 run, 100+100 ring



The more integrated a ring, the less likely it is that such a ‘mushroom’ phenomenon occurs, because it is less likely that any cluster shrinks to size  $k$ . But if one of the clusters reaches that size, its impact will be bigger the more segregated a ring is, because with few clusters, locations offering the highest utility are scarce. So the newly created optimal location is more likely to be chosen, and the alternative scenario of the cluster shrinking further to size  $k-1$  and subsequently disappearing becomes less likely. This argument explains why similar outbursts of integration do not occur with the flat utility function, as with that function all agents inside a ghetto enjoy the highest possible utility as well. Hence, the emergence of an extra optimal location is more likely to remain unnoticed and disappear.

**Fig. 17.** BR dynamics, average ‘first passage times’, 1000 runs, 100+100 ring



To further characterize the BR dynamics, we compute the average ‘first passage times’ for the dynamics. The first passage time for  $x$ , the number of clusters, is the first period in which there are not more than  $x$  clusters in that run.<sup>31</sup> An analysis of first passage times abstracts from the outbursts shown above. Figure 17 plots the average first passage times for the flat utility function and for the peaked or spiked utility functions for 1000 runs for a 100+100 ring. The sets of random initial configurations are the same for the flat and for the peaked and spiked utility functions. As envisaged, average passages are significantly quicker for the flat utility specification than for the peaked or spiked ones. Also for both types of utility function, as the number of clusters decreases, the first passage times increase sub-exponentially and exponentially for a wide range of values, and soar super-exponentially for extreme levels of segregation only.

<sup>31</sup> For instance, if a random starting allocation has 104 clusters, then the first passage time for each  $x \geq 104$  is 0. If there are 106 clusters in period 1, and 102 clusters in period 2, then the first passage time for  $x=102$  is 2.

#### 5.4 Inertia and line

In this section we consider BR dynamics with inertia, and BR dynamics on a line (instead of a ring) to answer the same two questions as in Section 4.3. First, are inertia and the line essential for Schelling's [1969, 1971a] results? Second, would inertia or a line change our findings concerning the peaked and spiked utility functions?

**Table 14.** Final distributions with inertia, 10+10 ring

		<b>BR dynamics (100,000 periods)</b>	
		<b>flat</b>	<b>peaked or spiked</b>
	<i>Num. Obs.</i>	1000	1000
<b>Clusters</b>	<i>Average</i>	2.01	2.01
	<i>5%</i>	2	2
	<i>95%</i>	2	2
<b>Distance</b>	<i>Average</i>	2.99	3.00
	<i>5%</i>	3.00	3.00
	<i>95%</i>	3.00	3.00
<b>Mix dev.</b>	<i>Average</i>	0.25	0.25
	<i>5%</i>	0.25	0.25
	<i>95%</i>	0.25	0.25
<b>Share</b>	<i>Average</i>	0.75	0.75
	<i>5%</i>	0.75	0.75
	<i>95%</i>	0.75	0.75
<b>Ghetto</b>	<i>Average</i>	3.98	3.99
	<i>5%</i>	4	4
	<i>95%</i>	4	4

The results presented in the preceding sections indicate that inertia was not essential for Schelling [1969, 1971a]. Without inertia we got segregation as well and, as Table 14 shows, even slightly more segregation than we did with inertia. Table 14 also shows that inertia would not upset our own findings. With inertia, there still is the same degree of segregation for the flat and the peaked or spiked utility functions. Remarkably, the outcomes are almost the same as without inertia. For the flat utility function, there are four clusters in five cases out of one thousand, and complete segregation otherwise. For the peaked or spiked utility function, 997 runs end in complete segregation, while three runs have four clusters (all of size five), constituting one of the eighteen MNE. Table 15 summarizes this, showing that, even with inertia, the existence of a number of MNE (all of them integrated), is hardly an obstacle for the BR dynamics on the way towards complete segregation.



**Table 15.** Number of MNE reached with inertia, 10+10 ring

	<b>BR dynamics (100,000 periods)</b>	
	<b>flat</b>	<b>peaked or spiked</b>
<i>Observations</i>	1000	1000
<i>non-strict MNE</i>	1000	3
<i>strict MNE</i>	n.a.	n.a.
<i>total MNE</i>	1000	3

We now turn to a line. Was the fact that Schelling [1969, 1971a] used a line instead of a ring essential to his results? The answer to this question is clearly ‘no’, as our results concerning a ring show that segregation does occur also there. Next, would considering a line change our findings concerning the peaked and spiked utility functions? The answer is again negative.

**Table 16.** Number of existing MNE, 10+10 line

	<b>MNE</b>	
	<b>flat utility</b>	<b>peaked or spiked</b>
<i>non-strict MNE</i>	3	0
<i>strict MNE</i>	0	0
<i>total MNE</i>	3	0

The main change that comes with a line is the non-existence of MNE for the peaked and spiked utility functions (see Proposition D1 plus Corollary D1 in Appendix D), whereas the flat utility function is left with only three MNE.<sup>32</sup> This follows from the fact that the set of MNE on a line with flat utilities consists of all possible combinations of clusters, each populated by at least  $k+1$  agents (Proposition D2). Table 16 summarizes these observations.

**Table 17.** Number of MNE reached, 10+10 line

	<b>BR dynamics (100,000 periods)</b>	
	<b>flat</b>	<b>peaked or spiked</b>
<i>Observations</i>	1000	1000
<i>non-strict MNE</i>	1000	n.a.
<i>strict MNE</i>	n.a.	n.a.
<i>total MNE</i>	1000	n.a.

<sup>32</sup> These equilibria are: XXXXXO0000000000000XXXXX, XXXXXO00000XXXXXO0000, and XXXXXXXXXXXXO00000000000.

The question, then, is whether this has any effect on the outcomes of the BR dynamics. As Table 17 shows, BR dynamics lead invariably to a MNE for the flat utility function. But, according to Table 18, there is no effect on segregation. Notwithstanding the fact that the set of MNE is very different, for each of the utility functions considered we always end up in a state of complete segregation (just as on a ring). Table 18 also shows that the segregation measures as such are slightly different for a line, also for the distribution of all possible allocations and for the set of MNE, due to the fact that the neighborhoods for agents towards the ends of the line are slightly different.

**Table 18.** Final distributions, 10+10 line

		<i>all</i>	<i>MNE</i>		<i>BR dynamics</i>	
			<i>flat</i>	<i>peaked or spiked</i>	<i>flat</i>	<i>peaked or spiked</i>
<i>Num. Obs.</i>		92,378	3	0	1000	1000
<b>Clusters</b>	<i>Average</i>	11.00	3.00	n.a	2.00	2.00
	<i>5%</i>	7	2	n.a	2	2
	<i>95%</i>	15	4	n.a	2	2
<b>Distance</b>	<i>Average</i>	1.15	3.63	n.a	5.50	5.50
	<i>5%</i>	0.80	2.40	n.a	5.50	5.50
	<i>95%</i>	1.68	5.50	n.a	5.50	5.50
<b>Mix dev.</b>	<i>Average</i>	0.12	0.24	n.a	0.38	0.38
	<i>5%</i>	0.07	0.10	n.a	0.38	0.38
	<i>95%</i>	0.20	0.38	n.a	0.38	0.38
<b>Share</b>	<i>Average</i>	0.47	0.71	n.a	0.86	0.86
	<i>5%</i>	0.41	0.57	n.a	0.86	0.86
	<i>95%</i>	0.57	0.86	n.a	0.86	0.86
<b>Ghetto</b>	<i>Average</i>	0.10	6.00	n.a	12.00	12.00
	<i>5%</i>	0	2	n.a	12	12
	<i>95%</i>	1	12	n.a	12	12

### 5.5 Discussion of findings in one-dimensional setup

Schelling's result of segregation in a 1D setup proved to be robust to assumptions on inertia and the configuration of the space. This allowed us to abstract from superfluous details on the order of moves, assume away inertia and work with a ring, to construct the Schelling's model in its purest form, and to concentrate on the fundamental question posed by the model: what is it about the preferences that leads to the striking outcome of segregation?

The answer is different from the one offered in the 2D case. The assumptions on the preferences can be extremely mild. In particular, the asymmetry of the utility function (favoring a majority over a minority neighborhood) plays no role. A sufficient condition on the utility

function is that it implies a strict preference for perfect integration. Given that, the utility function may have multiple peaks, and it may even describe the preference for living in any minority neighborhood rather than in any majority neighborhood.<sup>33</sup>

Particular specifications of the preferences, however, matter when it comes to the speed of convergence to the limit result. The flat utility function provides the strongest impetus towards quick segregation compared to the family of peaked functions. But in contrast to the 2D model, in its 1D counterpart, myopic equilibria do not act as attractors. In other words, in the 1D setup, the BR dynamics do not act as an equilibrium selection device.

It is the arrangement of the space and the definition of moves that matter. In particular, it is the fact that perfectly integrated positions always exist that ensures that extreme segregation is favored by the BR dynamics.<sup>34</sup> Agents move to a bliss location only to be displaced by a new entrant inserting himself at the border between two ghettos. Moves in the 2D setup, on the other hand, are characterized by a quick disappearance of ideal locations (due to the urge to avoid a minority status), and agents having to move to less satisfactory ones. Consequently, the two spatial proximity models proposed by Schelling offer two very different explanations for segregation.

---

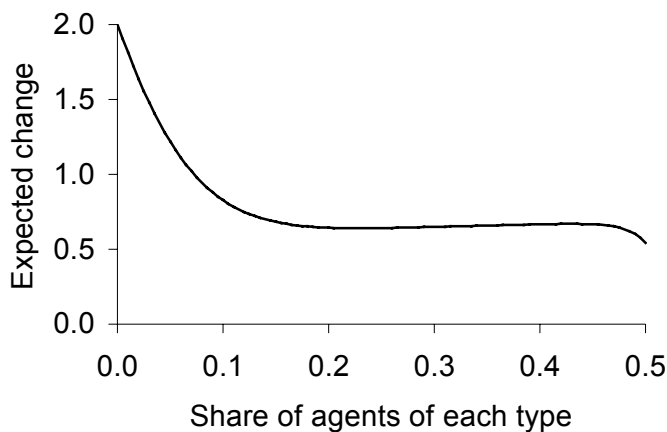
<sup>33</sup> An example would be the horizontal mirror image of the peaked p50 utility function. Such features seem more relevant from a biological perspective, where species might want to avoid living with too many like competitors.

<sup>34</sup> To check that it is not the dimensionality as such that is the essential difference between the 1D and 2D models, we considered the following one-dimensional version of the 2D model. Take a 1x25 board with ten agents of each type and five empty locations, and connect the first and the last cell of the board. BR dynamics in this variant lead to segregation only for the flat and the p50 utility function, just as for the other 2D models analyzed. This suggests that it is the technical nature of the moves as such that constitutes the essential difference.

## Appendix A. Some sensitivity analysis of 2D setup

The analysis in Section 4 assumes that 20% of the board or torus remains unoccupied. Since BR dynamics eliminate attractive choice locations, the question arises whether the chosen population density is critical to the behavior of the system. For the flat utility function, this issue can be tackled analytically and is discussed here in brief.<sup>35</sup> Figure A1 summarizes the analysis. A random board of an infinite size is taken as a starting point, so the results apply asymptotically to both a torus and a board. The share of agents of each type (relative to the number of locations) is measured on the horizontal axis. The parameter value used in the analysis described above was 0.4. The limit is taken for the values of 0 (an ‘empty’ board) and 0.5 (a ‘full’ board) of the share parameter. The vertical axis shows the expected number of agents who switch from being dissatisfied to being satisfied as a result of a BR move. This includes all agents, whose neighborhood is affected by the move.

**Fig. A1.** Total number of satisfied agents



The peak at 0 is caused by the assumption that an agent with no neighbors is dissatisfied. Thus, when a BR move occurs, a lonely agent joins a lonely agent with probability one, which yields two satisfied agents. The almost horizontal part of the diagram is insensitive to the assumption about satisfaction of a lonely agent. It also indicates that the results are largely insensitive to the assumption about the population density.

The result in Figure A.1 is a composition of three effects: own effect (the change in utility of a migrant), and the effects on old and new neighborhoods. The migrant will be either better off or indifferent (there is no inertia) after the move occurs. The effect on the abandoned neighborhood is near zero at 0.4 and varies in sign with the population density as shown in Figure A2. The effect is zero for extremely sparse populations, because the abandoned neighborhood will most likely be empty. The following dip is due to the assumption that a lonely agent is dissatisfied. The sharp drop towards 50% of each type is due to the fact that an agent is more likely to be satisfied with eight neighbors rather than seven.

The logic behind Figure A3, describing the expected impact on the joined neighborhood, is similar to the above, after having accounted for characteristics which made this neighborhood a likely destination of a BR move. The expected effect on the migrant's new neighbors is positive

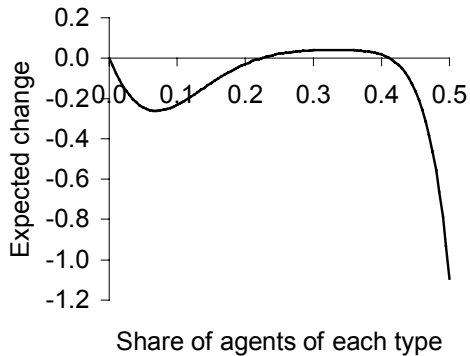
---

<sup>35</sup> The derivation of these results is available from the authors.

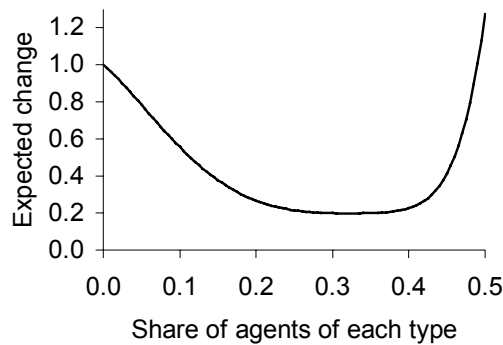
for any population density. Finally, the overall externality is positive, and the effect is reinforced by the positive own effect, which results in Figure A1.

This shows that neither the assumption on the utility of a lonely agent nor the chosen population density are critical to the obtained results.

**Fig. A2.** Impact on neighborhood left



**Fig. A3.** Impact on neighborhood joined



### Appendix B. Proof of complete segregation on a ring

Utility is one if the number of O-neighbors equals the number of X-neighbors and zero otherwise. Best response is assumed without inertia. Only dynamics on a circle is considered here. For these proofs it is assumed that the number of Xs equals the number of Os equals  $m$ .

To prove that complete segregation is the limit outcome of the best response dynamics it should be possible to reach complete segregation from each initial configuration with positive probability. This requires the set of all completely segregated outcomes to be the absorbing state while all other states be transient. Then in the long run all transient states are vacated for the absorbing state.

On a completely segregated circle there are two borders between ghettos, which are the only locations, which offer positive utility, which protects integrity of the ghettos. Thus, only transience of the remaining states needs to be proved. To prove this, it is sufficient to offer an algorithm showing just one possible path leading from each allocation to complete segregation.

In the proofs below this is done in two steps. A ‘seed’ is discovered or constructed through positive probability moves (PPMs) and then complete segregation is built from this seed. PPM can not only be used to insert an agent into a position with a utility of one, but also to drop out an agent from a position with a utility of zero, provided the existence of an appropriate destination is proven. PPMs with source and destination within a segment can also be made.

The following notation will be used:

- $2k$  is the size of a neighborhood:  $k$  neighbors to the left and to the right.
- $2m$  is the number of agents:  $m$  of each type.
- $\{X\}_n$  is a segment consisting of  $n$  Xs. Thus  $\{X\}_5$  is XXXXX,  $\{O\}_3$  is OOO.
- Square brackets [...] are used to indicate a seed.

·  $\{X, O\}_{l,n}$  is any segment consisting of  $l$  Xs and  $n$  Os in an arbitrary order.<sup>36</sup>

LEMMA B1. Any segment of the type  $\{X, O\}_{k-l,l} X \{X, O\}_{l,k-l}$ , where  $2k$  is the size of a neighborhood and  $0 < l < k$ , can be transformed into  $\{O\}_l \{X\}_{k+l} \{O\}_{k-l}$  by means of positive probability moves.

PROOF. The two braces of the given segment contain  $k$  Xs and  $k$  Os. The agent in the central position neighbors them all and thus enjoys the highest utility. It is possible to move successively every X from each brace into the central position and such move will be a positive probability move. For instance,

$$\{X, O\}_{k-l,l} X \{X, O\}_{l,k-l} \rightarrow \{X, O\}_{k-l-1,l} \{X\}_2 \{X, O\}_{l,k-l}.$$

The central position will always enjoy the highest utility, because whenever X is taken from the left brace the former central X shifts to the left and enters the brace from the right end. The removed X is always put in the central position. Similarly, when X is taken from the right brace the former central X shifts to the right and enters the right brace, while the removed X takes its place.  $\square$

LEMMA B2. On any circle of size  $2m$ , where  $m$  is the number of each type, and  $m \geq k$ , where  $k$  is the size of  $k+k$  neighborhood, there exist at least two perfectly integrated positions between two neighboring agents.

PROOF. Define  $\sigma_i$  as the number of Xs, which a segment of length  $2k$  contains, where  $2k$  is the size of a neighborhood and  $i$  is the index of a position between two consecutive agents in the middle of the segment. It follows that  $\sigma_i \in \{0, 1, \dots, 2k\}$  and  $i \in \{1, 2, \dots, 2m\}$ . Position  $i$  provides positive utility if and only if  $\sigma_i = k$ . If we go through all such possible segments once, then each agent will be covered exactly  $2k$  times – not more because it is given that  $m \geq k$ . Therefore

$$\sum_{i=1}^{2m} \sigma_i = 2km. \quad (*)$$

The mapping  $\sigma$  has a property for all  $i$  that

$$|\Delta\sigma_i| = |\sigma_i - \sigma_{i+1}| \leq 1, \quad (**)$$

where  $\sigma_{2m+1} = \sigma_0$ . If segment  $i$  is shifted by one position, it is possible that

- one X enters the segment, one X leaves, then  $|\Delta\sigma_i| = 0$
- one X enters, none leave, then  $|\Delta\sigma_i| = 1$
- none enter, none leave, then  $|\Delta\sigma_i| = 0$
- none enter, one leaves, then  $|\Delta\sigma_i| = 1$ .

---

<sup>36</sup> For instance,  $\{X, O\}_{2,3}$  could be XOOOX, as well as OOXOX or OXXOO.

Two cases are possible:

a)  $\exists i : \sigma_i > k$ . Then for (\*) to hold it follows that  $\exists j : \sigma_j < k$ . Without loss of generality, assume that  $i < j$ . From (\*\*) it follows that  $\exists l_1 : \sigma_{l_1} = k$  and  $\exists l_2 : \sigma_{l_2} = k$  such that  $i < l_1 < j < l_2$ .

b)  $\neg \exists i : \sigma_i > k \Rightarrow \neg \exists j : \sigma_j < k$ . Consequently  $\forall i : \sigma_i = k$ .

Thus there always exist at least two locations that can provide positive utility.  $\square$

LEMMA B3. Any segment of the type  $\{O\}_{l_1} \{X\}_{k+1} \{O\}_{k-l_2}$ , where  $2k$  is the size of a neighborhood,  $0 < l < k$ , and the number of each type on the circle is  $m > k$ , can be transformed into a seed  $[\{O\}_k \{X\}_k]$  by means of positive probability moves.

PROOF. The proof is done by induction.

Inductive base: segment  $\{O\}_{k-1} \{X\}_{k+1} O$  can be preceded by either O or X. The former case yields a seed  $[\{O\}_k \{X\}_k] XO$ .

Consider the latter case of  $X\{O\}_{k-1} \{X\}_{k+1} O$ . By Lemma B2 there exist at least two positions with positive utility. At the same time between any of the two Xs, which are inside the fragment  $\{X\}_{k+1}$  and immediately to the left of  $\{X\}_{k+1} : \sigma > k$ . Consequently,  $\sigma = k$  in at least two places outside this fragment.

a) If  $\sigma = k$  immediately to the right of  $\{X\}_{k+1}$ , then we have a seed.

b) If  $\sigma = k$  anywhere enclosed between horizontal bars  $X\{O\}_{k-1} \{X\}_{k+1} O$  (bars included), then by PPM we insert there O and obtain a seed  $X[\{O\}_k \{X\}_k] XO$ .

c) If  $\sigma = k$  is outside the segment  $X\{O\}_{k-1} \{X\}_{k+1} O$ , then the following action will be a positive probability move:  $X\{O\}_{k-1} X\{X\}_k O \rightarrow X\{O\}_{k-1} \{X\}_k O$ . Since the destination is outside the segment, the moved X could not have constituted a part of its neighborhood. Finally,  $X\{O\}_{k-1} \{X\}_k O \rightarrow X\{O\}_{k-1} XO\{X\}_{k-1} O \rightarrow X[\{O\}_k \{X\}_k] O$ .

Inductive step: assume that  $\{O\}_{l_1} \{X\}_{k+1} \{O\}_{k-l_2}$  (\*) can be transformed into a seed. It needs to be shown that segment  $\{O\}_{l_1-1} \{X\}_{k+1} \{O\}_{k-l_2+1}$  (\*\*) can also be transformed into a seed. The segment (\*\*) can be preceded by either O or X. In the former case we get:  $\{O\}_{l_1} \{X\}_{k+1} \{O\}_{k-l_2} O$ . The obtained segment contains (\*) and therefore can also be transformed into a seed by assumption.

Consider the latter case when X precedes (\*\*):  $X\{O\}_{l_1-1} \{X\}_{k+1} \{O\}_{k-l_2+1}$ . As with the inductive step, we know that inside  $\{X\}_{k+1}$  and immediately to the left of  $\{X\}_{k+1}$  we have  $\sigma > k$ .

a) If  $\sigma = k$  immediately to the right of  $\{X\}_{k+1}$ , then we already have a seed.

b) If  $\sigma = k$  anywhere enclosed between horizontal bars  $X\{O\}_{l_1-1} \{X\}_{k+1} \{O\}_{k-l_2+1}$  (bars included), then by PPM move there O from the right hand side so that the resulting segment  $X\{O\}_{l_1} \{X\}_{k+1} \{O\}_{k-l_2}$  contains (\*).

c) If  $\sigma = k$  anywhere within the rightmost bracket  $\{O\}$ , then insert there the free O. Three cases are possible

i) A seed has formed:  $X\{O\}_{l-1}X\{X\}_k\{O\}_k$ .

ii) If in the rightmost bracket  $\{O\}$  there still is a position with  $\sigma = k$ , then repeat c) again (i.e., insert there the O, and see what happens).

iii) If there is no position in the rightmost  $\{O\}$  with  $\sigma = k$ , then the following is possible:

$$\begin{aligned} X\{O\}_{l-1}\{X\}_{k-l}X\{X\}_l\{O\}_{k-l+1} &\rightarrow X\{O\}_{l-1}\{X\}_{k-l}\{X\}_l\{O\}_{k-l+1} \rightarrow \\ X\{O\}_{l-1}\{X\}_{k-l+1}O\{X\}_{l-1}\{O\}_{k-l+1} &\rightarrow \{O\}_{l-1}\{X\}_{k-l+1}O\{X\}_l\{O\}_{k-l+1} \rightarrow \\ \{O\}_{l-1}\{X\}_{k-l}O\{X\}_{l+1}\{O\}_{k-l+1} &\rightarrow \{O\}_{l-1}\{X\}_{k-l-1}O\{X\}_{l+2}\{O\}_{k-l+1} \rightarrow \dots \rightarrow \\ \{O\}_{l-1}XO\{X\}_k\{O\}_{k-l+1} &\rightarrow \{O\}_l\{X\}_{k+1}\{O\}_{k-l}O. \end{aligned}$$

The obtained segment contains (\*) and therefore can also be transformed into a seed.  $\square$

**PROPOSITION B1.** If a neighborhood on a circle of size  $2m$  is defined as  $k$  neighbors to the left and  $k$  neighbors to the right, then, if  $m$  is the number of each type,  $m > k$ , and the utility function is spiked, then the process of best-responses has a unique recurrent state consisting of all completely segregated states.

**PROOF.** If an allocation contains a seed, then construction of complete segregation by means of PPMs is trivial. Otherwise construct a seed. If no agent enjoys positive utility, go to a location that ensures one (it exists by Lemma C1). Now at least one agent has positive utility. Take one such agent and consider all possible configurations of his neighborhood. Without loss of generality we can assume that X is such an agent. Since its utility is positive, its neighborhood is the segment  $\{X, O\}_{k-l,l}X\{X, O\}_{l,k-l}$ . By Lemma B1 this can be transformed into  $\{O\}_l\{X\}_{k+1}\{O\}_{k-l}$  and then by Lemma B3 into  $\{O\}_k\{X\}_k$  with positive probability.

Thus a seed can always be constructed and complete segregation then built.  $\square$

### Appendix C. On MNE on a ring

Assume a ring with equal number of each type and a spiked utility function.<sup>37</sup> Now take any one agent out, say, without loss of generality, take agent X out. Then the following lemma applies.

**LEMMA C1.** On any circle of size  $2m-1$ , where  $m$  and  $m-1$  is the number of each type, and  $m > k$ , where  $k$  is the size of a  $k+k$  neighborhood, there exist at least two perfectly integrated positions between two neighboring agents.

**PROOF.** Borrow the definition of mapping  $\sigma_i$  and its properties from Lemma B2. Without loss of generality assume that there are  $m$  Os and  $m-1$  Xs. The following equality will hold:

$$\sum_{i=1}^{2m-1} \sigma_i = 2k(m-1) \quad (*)$$

<sup>37</sup> The focus on such rings is in part justified by Proposition C6.



The proof proceeds by contradiction. Assume there exist no two  $i$ , for which  $\sigma_i = k$ .

From (\*) it follows that  $\exists i : \sigma_i < k$  for otherwise

$$\sum_{i=1}^{2m-1} \sigma_i > k(2m-1) > 2k(m-1),$$

so that (\*) would not hold. Since for some  $i$   $\sigma_i < k$  and for no two  $i$   $\sigma_i = k$ , then  $\neg \exists i : \sigma_i > k$ , because  $\sigma_i$  can only change by increments of 1. Hence, for at most one  $i$   $\sigma_i = k$ , while for the remaining ones  $\sigma_i < k$ . Consequently,

$$\sum_{i=1}^{2m-1} \sigma_i \leq (2m-2)(k-1) + k.$$

The necessary condition for (\*) to be satisfied is

$$(2m-2)(k-1) + k \geq 2k(m-1) \text{ or } m \leq (k+2)/2.$$

This, however, contradicts the premise of the lemma requiring  $m > k$ . Therefore there exist at least two  $i$ , for which  $\sigma_i = k$ .  $\square$

**PROPOSITION C1.** For an allocation on a circle with  $m$  agents of each type, where  $m > k$ , and a spiked utility function, to be a MNE it is necessary and sufficient that all agents enjoy positive utility.

**PROOF.** The proof proceeds by contradiction. Assume a MNE and pick any unsatisfied agent; without the loss of generality let it be X. Take it out. Lemma C1 applies: there will always exist a place, where X can go to obtain positive utility. This contradicts the definition of MNE. Hence, it is necessary that there are no dissatisfied agents in a MNE. If all agents have positive utility, none will be strictly better off from moving elsewhere; hence such an allocation should be a MNE.  $\square$

**PROPOSITION C2.** On a circle with  $m$  agents of each type, where  $m > k$ , and a spiked utility function, there are no strict MNE.

**PROOF.** The proof proceeds by contradiction. Assume a strict MNE and pick any agent, let it be X. Take it out. Lemma C1 applies: on the remainder of the ring there will always be at least two positions with positive utility. One of these is the position the agent occupied (by Proposition C1), but there will be at least one more position elsewhere. Hence, there is always a choice to move elsewhere to a place with positive utility. This contradicts the definition of a strict MNE. Therefore, there are no strict MNE.  $\square$

**PROPOSITION C3.** On a circle with  $m$  agents of each type, where  $m > k$ , and a spiked utility function, in a MNE there can be no cluster larger than  $k+1$ , where  $k$  is the size of a neighborhood.

**PROOF.** A cluster bigger than  $k+1$  requires that all its agents but those on the two edges have zero utility. But by Proposition C1 this contradicts the assumption that the cluster is a part of a MNE allocation. Hence, no cluster can be larger than  $k+1$ .  $\square$

PROPOSITION C4. For an allocation on a circle to be a MNE with spiked preferences and neighborhood size  $k$ , where  $k$  is odd, it is necessary and sufficient that agents  $i$  and  $i+k+1$  for all  $i$  are of opposite types.

PROOF. The proof is constructive. Given an arbitrary segment of  $2k+1$  agents  $a_0 a_1 \dots a_k \dots a_{2k}$ , with the central agent  $a_k$  being satisfied, it is sometimes possible to continue the string adding agents  $a_{2k+1}$  onwards so that agents  $a_{k+1}$  onwards be also satisfied. Call this a complementary step procedure. The initial segments, for which it is possible to continue the process until the repetition starts, are compatible with positive utility for all agents. Since positive utility for every agent is a necessary and sufficient condition for an allocation to be a MNE (by Proposition C1), this procedure will allow to construct all patterns, which can serve to build a MNE, given the size of a neighborhood  $k$ .

The complementary step procedure will necessarily reproduce the initial segment. It is impossible to produce a string of an infinite length for which all segments consisting of  $2k+1$  agents would be unique, because the number of such segments is finite. Thus the string obtained by means of this procedure can be joined in a circle. This ensures that not only agents  $a_k$  onwards are satisfied, but also that agents from  $a_0$  through  $a_{k-1}$  are satisfied by construction.<sup>38</sup>

Let  $a_i = -1$  if location  $i$  is occupied by  $X$  and  $a_i = 1$  otherwise. Then the recursive formula used in the complementary step procedure to add agent  $i$  given the preceding segment is

$$a_i = a_{i-2k-1} - a_{i-k-1} + a_{i-k}. \quad (1)$$

The formula ensures that if in segment  $a_{i-2k-1} a_{i-2k} \dots a_{i-k-1} a_{i-k} \dots a_{i-1} a_i$  agent  $a_{i-k-1}$  was satisfied, then  $a_{i-k}$  will also be satisfied. It is undefined for  $a_{i-2k-1} = a_{i-k} = -a_{i-k-1}$ .

To prove necessity, assume the contrary:  $\exists r : a_r = a_{r+k+1} = \alpha_0$ . Without the loss of generality, set  $r=0$ . Applying (1) yields  $a_{2k+1} = a_0 - a_k + a_{k+1} = 2\alpha_0 - a_k$ . Hence,  $a_{2k+1} = a_k = \alpha_0$ . Applying (1) again gives  $a_{3k+1} = a_k - a_{2k} + a_{2k+1}$ , so that  $a_{3k+1} = a_{2k} = \alpha_0$ . Reiteratively using (1), the following obtains:

$$a_{zk} = a_{zk+1} = \alpha_0, \quad (2)$$

where  $z \in \mathbb{Z}$ . Apply (1) once again:  $a_{zk} = a_{(z-2)k-1} - a_{(z-1)k-1} + a_{(z-1)k}$ . Now by (2) it follows that

$$a_{(z-2)k-1} = a_{(z-1)k-1} \text{ or, equivalently, } a_{zk-1} = \alpha_1. \quad (3)$$

Next,  $a_{zk-1} = a_{(z-2)k-2} - a_{(z-1)k-2} + a_{(z-1)k-1}$ , which gives  $a_{zk-2} = \alpha_2$ , and in general

$$a_{zk-l} = \alpha_l, \quad (4)$$

---

<sup>38</sup> It is also crucial that a complementary step procedure is uniquely defined, so it is possible to ‘scroll back’ to the initial allocation. For assume the sequence does eventually repeat, but not from the initial segment. Then it would be possible from any place on the formed circle, which excludes the initial segment, to return both to the initial segment and never return to it. This is a contradiction. Hence repetition starts from the initial segment.

where  $1 \leq l \leq k - 2$ . Combining (2), (3) and (4) ensures that a string consists of the following repeated segments:

$$\alpha_{k-2}\alpha_{k-3}\dots\alpha_2\alpha_1\alpha_0\alpha_0$$

Take a fragment of the string as given below

$$\alpha_0\alpha_0\alpha_{k-2}\alpha_{k-3}\dots\alpha_2\alpha_1[\alpha_0]\alpha_0\alpha_{k-2}\alpha_{k-3}\dots\alpha_2\alpha_1\alpha_0.$$

The condition for the bracketed agent to be satisfied, so that the fragment could be a part of a MNE, is (having divided through by 2)

$$2\alpha_0 + \alpha_1 + \alpha_2 + \dots + \alpha_{k-3} + \alpha_{k-2} = 0. \quad (5)$$

Recall that  $k$  is odd and  $|\alpha_i| = 1$ . Hence condition (5) can never be satisfied and

$$\forall i : a_i = -a_{i+k+1}. \quad (6)$$

To prove sufficiency consider any segment containing a full neighborhood (the index of the first element is normalized to 0)

$$a_0a_1\dots a_k\dots a_{2k-1}a_{2k}$$

The condition for the middle agent to be satisfied is

$$a_0 + a_1 + \dots + a_{k-1} + a_{k+1} + \dots + a_{2k-1} + a_{2k} = 0 \quad (7)$$

Substitution of (6) into (7) readily shows that (7) is an identity. Hence if (6) holds, then all agents have positive utility and an allocation is a MNE.  $\square$

**PROPOSITION C5.** For an allocation on a circle to be a MNE with spiked preferences and neighborhood size  $k$ , where  $k$  is even, it is necessary and sufficient that either (i) agents  $i$  and  $i+k+1$  are of opposite types or (ii) agents  $i$  and  $i+k$  are of the same type and  $a_i + a_{i+1} + \dots + a_{i+k-2} + a_{i+k-1} = 0$  for all  $i$ .

**PROOF.** The proof draws heavily on the proof of Proposition C4. In the necessary part of the proof of Proposition C4 it has been demonstrated that if an allocation is a MNE, then either agents  $i$  and  $i+k+1$  are of opposite types or agents  $i$  and  $i+k$  are of the same type. Moreover if agents  $i$  and  $i+k$  are of the same type, then condition (5) holds, and in any pattern there is at least one pair of agents of the same type (of type  $\alpha_0$ ). The only pattern with all adjacent agents of different types is a string of alternating X and O. This MNE falls under type (i). Condition (5) is equivalent to one in (ii), where two agents of the same type are gathered in one term. Condition (5) can now be satisfied, because  $k$  is even.

To prove sufficiency, consider any segment containing a full neighborhood (the index of the first element is normalized to 0)  $a_0a_1\dots a_k\dots a_{2k-1}a_{2k}$ . The condition for the middle agent to be satisfied is (7). If a pattern is of type (i), then substitution of (6) into (7) turns it into identity. Hence if (6) holds, then all agents have positive utility and an allocation is a MNE. If a pattern is of type (ii), then the fact that  $a_i = a_{i+k}$  and the condition in (ii) ensures that (7) and (5) hold, so that the allocation is a MNE.  $\square$

**PROPOSITION C6.** MNE is only possible, if Xs and Os are in equal numbers.

PROOF. Having mapped X into  $-1$  and O into  $1$ , it is a property of a neighborhood of each satisfied agent  $i$ , that

$$a_{i-k} + \dots + a_{i-1} + a_{i+1} + \dots + a_{i+k} = 0 \quad (1)$$

Add condition (1) for every agent on a circle:  $2k \sum a_i = 0$  or

$$\sum_{i \in \{i: a_i=1\}} a_i = \sum_{i \in \{i: a_i=-1\}} |a_i|. \quad (2)$$

Equation (2) says that agents should come in equal numbers.  $\square$

LEMMA C2. In a MNE with the flat utility function adjacent agents of different type enjoy perfect integration.

PROOF. Consider two neighboring agents:  $X'$  and  $O'$ , each of which prefers at least  $k$  neighbors of his own type. The neighborhood of  $X'$  is different from the neighborhood of  $O'$  in that it contains two more agents, one of which is  $O'$ , and omits two agents, one of which is  $X'$ . Both  $X'$  and  $O'$  can be satisfied only if the neighborhood of  $O'$  contains exactly  $k$  Os, then it is possible to ensure  $k$  Xs for  $X'$  by stipulating that an extra agent is of X type and an omitted agent is of O type. Thus, adjacent agents of different type can only have perfectly integrated neighborhoods.

Using a mapping as in Proposition C4:

$$\Omega_{X'} = \Omega_{O'} + 2 + a' - a'' \leq 0 \text{ and } \Omega_{O'} \geq 0, \quad (*)$$

where  $\Omega_{X'}$  and  $\Omega_{O'}$  is the 'sum' of agents in the neighborhood of  $X'$  and  $O'$  respectively,  $a'$  and  $a''$  are additional and excluded agents. Both inequalities are satisfied only if  $\Omega_{X'} = \Omega_{O'}$ .

So far it has been implicitly assumed that we are either operating on a circle or considering agents on a line who are sufficiently remote from edges. The difference the possibility of edges introduces is that either  $a'$  or  $a''$  can turn into zero. If  $X'$  is closer to the border than  $O'$  then  $a'$  might turn into zero, otherwise  $a''$  might be zero. If this is the case, then it is never possible to satisfy (\*): both  $X'$  and  $O'$  cannot be satisfied simultaneously. Hence, the allocation is not a MNE.  $\square$

PROPOSITION C7. The set of MNE for single peaked utility functions, with the peak at perfect integration, coincides with the set of MNE for the spiked utility function.

PROOF. For a circle, Lemma C1 guarantees (provided  $m > k$ ) an option for any agent to migrate to a perfectly integrated location. Since the considered utility functions have peaks at perfect integration, it is impossible for any agent in a MNE not to enjoy a perfectly integrated neighborhood. But, by Proposition C1, this is a necessary and sufficient condition for an allocation to be a MNE with a spiked utility function. Hence, all MNE are common for single peaked and spiked utility functions (with the peak at perfect integration).  $\square$

PROPOSITION C8. The set of all MNE on a circle with a flat utility function and the neighborhood size  $k$  ( $m > k$ ) contains the set of all MNE on a circle with a peaked utility function for corresponding  $k$ , plus all possible combinations of clusters, each populated by at

least  $k+1$  agents with at least one cluster exceeding  $k+1$  agents. For  $k \leq 4$  there are no other MNE than these.

PROOF. In any MNE with the spiked utility function all agents enjoy maximum possible utility. With the flat utility function, all agents would continue to enjoy the highest possible utility for each of these allocations. Hence all MNE with the spiked function are also MNE with the flat function. Further, if all clusters are  $k+1$  or larger, than agents on edges enjoy perfectly integrated neighborhoods, while the rest are surrounded by a majority of their own type. Hence, no one has an incentive to move and such an allocation is a MNE. It remains to be shown that there are no other MNE.

In MNE, clusters of size  $k+1$  cannot co-exist with smaller clusters. Assume the opposite. Consider any cluster of size at least  $k+1$ , let it be X-type, which neighbors a cluster of  $k$  agents or less (O-type):

...X{O}<sub><k</sub> · O{X}<sub>≥k+1</sub>...

Then the agent marked with a dot will be in a minority, and hence willing to go to an allocation with higher utility, which is known to exist. It needs to be demonstrated, that there is no MNE with all cluster sizes less or equal to  $k$  (for  $k \leq 4$ ) where some agents would not enjoy a perfectly integrated neighborhood.

If  $k=1$  then a single agent forming a cluster of size 1 is always dissatisfied. If  $k=2$ , then any agent from a cluster sized 1 or 2, by Lemma C2, in MNE would enjoy a perfectly integrated neighborhood, if such a MNE exists. If  $k=3$ , again, by Lemma C2, we should consider only clusters of size 3. Such a cluster needs to be enveloped by two clusters of the opposite type comprising at least 4 agents each for bordering agents of the opposite type to be satisfied. Hence, no agent in cluster of size 3 will be satisfied, and this allocation will not be a MNE.

If  $k=4$ , then, by Lemma C2, we can consider only central agents in clusters of size 3 and 4. For the cluster of size 3, it is impossible that the central agent (e.g., O) is in a majority while the bordering agents of the opposite type (Xs) are satisfied:

...XXXOXOOOXOOX...

The cluster of size 4, trivially, will be surrounded by clusters of the opposite type of 5 agents or more for the border agents of the opposite type to be satisfied. This renders central agents dissatisfied:

...XXXXXOOOOXXXXX...

Thus it is impossible to construct a MNE with cluster sizes not exceeding  $k$  for ( $k \leq 4$ ), such that there exists an agent who is in a majority. □

#### Appendix D. On MNE on a line

LEMMA D1. On any line of size  $2m$ , where  $m$  is the number of each type,  $m > (k^2 - 1)/2$ , where  $k$  is the size of a  $k+k$  neighborhood, there exists at least one perfectly integrated position between two neighboring agents.

PROOF. Define  $\sigma_i$  as the number of Xs that a segment of length  $2k$  contains, where  $2k$  is the size of a neighborhood and  $i$  is the index of a position from which  $k$  neighbors in each

direction are counted to form a  $k+k$  neighborhood. If boundaries of the line do not allow to count all  $k$  agents in each direction, then as many as the boundaries allow are included. It follows that  $\sigma_i \in \{0, 1, \dots, 2k\}$  and  $i \in \{1, 2, \dots, 2m+1\}$ . Position  $i$  provides positive utility if and only if  $\sigma_i = k$ . If we go through all such possible segments once, then each agent will be covered at most  $2k$  times. Therefore

$$2km - k^2 + k \leq \sum_{i=1}^{2m+1} \sigma_i \leq 2km. \quad (*)$$

The lower boundary on the sum assumes that  $k-1$  positions at either end of the line are occupied by Xs, which ensures that they are counted least often. The upper limit assumes that all Xs are concentrated away from the edges and thus each is counted  $2k$  times. Expression (\*) is valid for  $m \geq 2(k-1)$ , which is ensured by the premise  $m > (k^2 - 1)/2$ .

The mapping  $\sigma$  has a property for all  $i$  that

$$|\Delta\sigma_i| = |\sigma_i - \sigma_{i+1}| \leq 1, \quad (**)$$

If segment  $i$  is shifted by one position, it is possible that

- one X enters the segment, one X leaves, then  $|\Delta\sigma_i| = 0$
- one X enters, none leave, then  $|\Delta\sigma_i| = 1$
- none enter, none leave, then  $|\Delta\sigma_i| = 0$
- none enter, one leaves, then  $|\Delta\sigma_i| = 1$ .

The proof proceeds by contradiction. Assume  $\neg\exists i: \sigma_i = k$ . It follows that  $\exists i: \sigma_i < k$ , for otherwise the sum in (\*) exceeds the upper boundary. From (\*\*) and the assumption that a position with positive utility does not exist, it follows that  $\neg\exists i: \sigma_i > k$ , hence  $\forall i: \sigma_i < k$ . Consequently,

$$\sum_{i=1}^{2m+1} \sigma_i \leq (2m+1)(k-1).$$

For the condition (\*) to hold, it is necessary that

$$(2m+1)(k-1) \geq 2km - k^2 + k \text{ or } m \leq (k^2 - 1)/2.$$

But this contradicts the premise, which requires  $m > (k^2 - 1)/2$ . Therefore  $\exists i: \sigma_i = k$ .  $\square$

**PROPOSITION D1.** For a line of size  $2m$ , where  $m$  is the number of each type,  $m > (k^2 - 1)/2$ , where  $k$  is the size of  $k+k$  neighborhood, and  $m > 1$  if  $k = 1$  and  $m > 3$  if  $k = 2$ , with a spiked utility function, no MNE exist.<sup>39</sup>

**PROOF.** If  $k$  is odd, then neighborhoods of the first and the last agents contain an odd number of neighbors. If  $k$  is even, then neighborhoods of the second and the next to last agents contain

---

<sup>39</sup> The conditions stated in Proposition D1 are not very restrictive. For instance, one should work with a string of 4 or bigger for  $k=1$ , a string of 8 for  $k=2$ , and with a string of at least 10 for  $k=4$ .

an odd number of neighbors. There is no way such neighborhoods can consist of equal numbers of both types. Therefore these agents will always be dissatisfied. The proof proceeds by contradiction. Assume there exists a MNE. Take the leftmost of the above mentioned dissatisfied agents. By Lemma D1, a position that offers positive utility necessarily exists. Let  $i^*$  be an index of such a position with positive utility. Two cases are possible.

- a) If  $k$  is odd and  $i^* \leq k + 1$ , then the rightmost dissatisfied agent is willing to migrate to  $i^*$ , hence the allocation cannot be a MNE. If  $i^* > k + 1$ , then the leftmost dissatisfied agent is willing to move to  $i^*$ , which again contradicts the assumption of MNE.
- b) If  $k$  is even and  $i^* \leq k + 2$ , then the rightmost dissatisfied agent will be willing to migrate to  $i^*$ , otherwise the leftmost dissatisfied agent will be willing to migrate to  $i^*$ , which is incompatible with a MNE.

The first condition on  $m$  ensures applicability of Lemma D1, while all three conditions make sure that  $i^*$  cannot belong to neighborhoods of the dissatisfied agents at the extremes of a line. Thus there is always an incentive for an agent to move, and hence no MNE exist.  $\square$

COROLLARY D1. Under the conditions stipulated in Proposition D1, MNE with a single peaked utility function do not exist either.

PROOF. Depending on whether  $k$  is odd or even, the first and the last or the second and the next to last agents will never enjoy a perfectly integrated neighborhood. Hence, these will never have the highest possible utility. They will be willing to migrate to a perfectly integrated neighborhood, which, according to Lemma D1 always exists.  $\square$

PROPOSITION D2. The set of all MNE on a line with the flat utility function and the neighborhood size  $k$  ( $k \leq 4$ ) contains all possible combinations of clusters populated by at least  $k+1$  agents each. Conditions stipulated in Proposition D1 should hold.<sup>40</sup>

PROOF. The proof follows that of Proposition C8. If all clusters are  $k+1$  or larger, then agents on edges enjoy perfectly integrated neighborhoods, while the rest are surrounded by a majority of their own type. Next, in MNE, clusters of size  $k+1$  cannot co-exist with smaller clusters. For otherwise an agent, who belongs to the smaller cluster and is situated at the border of the smaller and the bigger clusters would necessarily be dissatisfied.

It needs to be demonstrated that there is no MNE with all cluster sizes less or equal to  $k$  ( $k \leq 4$ ). Since there are no MNE in which all agents enjoy a perfectly integrated neighborhood, we shall look for a MNE in which some agent (say, O) has an own type majority neighborhood, with a cluster size not exceeding  $k$ . If  $k=1$  then a single agent forming a cluster of size 1 is always dissatisfied. If  $k=2$ , then any agent from a cluster sized 1 or 2, by Lemma C2, in MNE would enjoy a perfectly integrated neighborhood, if such a MNE exists.

If  $k=3$ , again, by Lemma C2, we should consider only clusters of size 3 and, as explained there, such a cluster can be assumed to be sufficiently far from the end of a line. It needs to be enveloped by two clusters of the opposite type. These two can never be satisfied, if O is to enjoy a majority of his own type:

...OXOOOXO...

---

<sup>40</sup> It is not required that at least one cluster exceeds  $k+1$ . This is because there are no MNE inherited from the spiked utility function, which alternating clusters of size  $k+1$  is an example of on a circle. The distinction was needed for a circle not to account for the ' $k+1$ ' equilibrium twice.

If  $k=4$ , then, again, by Lemma C2, we can consider only central agents in clusters of size 3 and 4 that are remote from both edges. For the cluster of size 3, it is impossible that the central agent (e.g. O) is in a majority while the bordering agents of the opposite type (Xs) are satisfied:

...XXXOXOOOXOOX...

The cluster of size 4, trivially, will be surrounded by clusters of the opposite type of 5 agents or more for the border agents of the opposite type to be satisfied. This renders central agents dissatisfied:

...XXXXXOOOOXXXXX...

Thus it is impossible to construct a MNE with cluster sizes not exceeding  $k$  (for  $k \leq 4$ ) such that there exists an agent who is in majority of his own type.  $\square$



## References

- Akerlof, G.A. (1997). Social Distance and Social Decisions. *Econometrica*, 65 (5), 1005-1027.
- Arrow, K.J. (1998). What Has Economics to Say about Racial Discrimination? *Journal of Economic Perspectives*, 12 (2), 91-100.
- Barkley Rosser Jr., J. (1999). On the Complexities of Complex Economic Dynamics. *Journal of Economic Perspectives*, 13 (4), 169-192.
- Binmore, K.G. (1992). *Fun and Games. A Text on Game Theory*. Lexington, MA: Heath.
- Blume, L.E. (1997). Population Games. In W.B. Arthur, S.N. Durlauf & D.A. Lane (Eds.), *The Economy as an Evolving Complex System II* (Vol. XXVII. Santa Fe Institute Studies in the Sciences of Complexity, pp. 425-460). Reading, MA: Addison-Wesley.
- Brock, W.A., & Durlauf, S.N. (2001). Discrete Choice with Social Interactions. *Review of Economic Studies*, 68 (2), 235-260.
- Dixit, A.K., & Nalebuff, B.J. (1991). *Thinking Strategically. The Competitive Edge in Business, Politics, and Everyday Life*. New York: Norton.
- Glaeser, E., & Scheinkman, J.A. (2002). *Non-Market Interactions*. (Mimeo - Invited lecture Econometric Society World Congress in Seattle, 2000).
- Ioannides, Y.M., & Seslen, T.N. (2002). Neighborhood Wealth Distributions. *Economics Letters*, 76 (3), 357-367.
- Krugman, P. (1996). *The Self-Organizing Economy*. Cambridge, MA: Blackwell.
- Lindbeck, A., Nyberg, S., & Weibull, J.W. (1999). Social Norms and Economic Incentives in the Welfare State. *Quarterly Journal of Economics*, 114 (1), 1-35.
- Manski, C.F. (2000). Economic Analysis of Social Interactions. *Journal of Economic Perspectives*, 14 (3), 115-136.
- Schelling, T.C. (1969). Models of Segregation. *American Economic Review, Papers and Proceedings*, 59, 488-493.
- Schelling, T.C. (1971a). Dynamic Models of Segregation. *Journal of Mathematical Sociology*, 1 (2), 143-186.
- Schelling, T.C. (1971b). On the Ecology of Micromotives. *The Public Interest*, 25, 61-98.
- Schelling, T.C. (1972). A Process of Residential Segregation: Neighborhood Tipping. In A.H. Pascal (Ed.), *Racial Discrimination in Economic Life*. Lexington, MA: Lexington Books.
- Schelling, T.C. (1978). *Micromotives and Macrobehavior*. New York: Norton.
- Skyrms, B., & Pemantle, R. (2000). A Dynamic Model of Social Network Formation. *Proceedings of the National Academy of Sciences of the USA*, 97 (16), 9340-9346.
- Young, H.P. (1998). *Individual and Social Structure: An Evolutionary Theory of Institutions*. Princeton, NJ: Princeton University Press.

**This working paper has been produced by  
the Department of Economics at  
Queen Mary, University of London**

**Copyright © 2003 Romans Pancs and Nicolaas J. Vriend  
All rights reserved.**

**Department of Economics  
Queen Mary, University of London  
Mile End Road  
London E1 4NS  
Tel: +44 (0)20 7882 5096 or Fax: +44 (0)20 8983 3580  
Email: [j.conner@qmul.ac.uk](mailto:j.conner@qmul.ac.uk)  
Website: [www.econ.qmul.ac.uk/papers/wp.htm](http://www.econ.qmul.ac.uk/papers/wp.htm)**