Schelling's Spatial Proximity Model of Segregation Revisited*

Romans Pancs,[§] Stanford University
Nicolaas J. Vriend,[§§] Queen Mary, University of London

September 2004

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. We also 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.

[§] R. Pancs, Stanford University, Department of Economics, Landau Economics Building, 579 Serra Mall, Stanford, CA 94305-6072.

[§§] Corresponding author: N.J. Vriend, Queen Mary, University of London, Department 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 seminar, workshop and conference participants of WEHIA (Trieste), CEF (Aix-en-Provence and Seattle), Complexity 2003 (Aix-en-Provence), Game Theory World Congress (Marseille), Queen Mary (London), GREQAM (Marseille), CeNDEF (Amsterdam), CERGE-EI (Prague), Warwick, Dundee, S. Anna (Pisa) and Stanford, and Antonio Cabrales, Herbert Dawid, Cees Diks, Steven Durlauf, Sanjeev Goyal, Joe Harrington, Yannis Ioannides, Alan Kirman, and Michael Thornton for helpful comments and discussions. 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. Lo and behold, an unraveling process starts from a more or less integrated city into a rather segregated one.

Schelling's neighborhood segregation model has become one of the most widely cited and acclaimed models in economics.[1] There may be three main reasons for this. First, the emergence of segregation seems intellectually 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, it 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. It is unusually simple. Combined with its intellectual appeal, this makes it a convenient means to illustrate 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 with a pen and paper. Finally, Schelling's model seems so widely cited because segregation is a serious social and political issue in the USA[2] and increasingly also in Europe.[3] [4]

Notwithstanding the general recognition given to Schelling's model nowadays, it leaves some question marks concerning the assumptions made 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

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

[2] As Mr. Trent 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.

[3] See, e.g., The Economist [2001].

[4] Notice that 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, between workers with different skills in different firms, or between different species occupying their own territory.

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 a formal as well as numerical analysis of the dynamic and equilibrium properties of Schelling's model of segregation.

Focusing on the driving force behind the dynamic behavior of Schelling's model, the individual preferences, we keep all other details as general and simple as possible. The main insight from our analysis can be summarized as follows. Schelling's model of segregation is very robust to a range of variations in its specification. Perhaps surprisingly, Schelling's mild assumptions concerning the individual preferences can be made considerably more extreme. Strict preferences for perfect integration by all individual agents will still lead to neighborhood segregation, and in the one-dimensional setup we prove formally that such preferences are a sufficient condition for complete segregation.

Apart from this main insight, our analysis leads to a better understanding of Schelling's model in the following three senses. First, the set of elements of the model that appear to be essential to explain segregation can be narrowed down considerably. Second, there is a significant difference between the version on a line and the variant on a lattice. 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, segregation occurs even if the individuals strictly prefer perfect integration with no bias whatsoever in favor of the agent's own type. In the one-dimensional case complete segregation is the unique long-term outcome. We also characterize (numerically on a lattice and formally on a line) the complete set of Myopic Nash Equilibria. This characterization makes clear that the dynamics in the model are not simply an equilibrium selection issue. On the one hand, in many cases the set of Myopic Nash Equilibria itself is biased towards segregation. 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, 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 original setup segregation was not a problem as, from a social welfare point of view, a completely segregated city may be just as well off as a perfectly integrated one, but with strict preferences for integration this is no longer true. This suggests that the effect on welfare of educating people to have preferences for integration might be adverse. Second, suppose a social planner wanted to maximize integration. 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), could a migration subsidy or tax system prevent segregation and implement integration? Although this question is beyond the scope of the current paper, the analysis suggests that a simple reward system for integrating moves or taxation of segregating moves might not work.

The remainder of this paper is structured as follows. In Section 2 we briefly recapitulate Schelling's [1969, 1971a, 1971b, 1978] spatial proximity model of segregation and outline the features of the model that we are going to analyze in detail. Section 3 presents some analytical tools and benchmark allocations. Schelling's two-dimensional model is analyzed in Section 4, while the one-dimensional, linear model is 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).[5] 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.[6] 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 is created by distributing equal numbers of agents of each type in random order. The dynamics, then, are an iterative process of agents choosing best-responses. At each stage all agents that are not satisfied are put in some arbitrary order. When an agent's turn comes, he

[5] 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).

[6] 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.

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.[7] 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 agents, who can each occupy one cell of the board. But now there are also some free cells left, as in a 5x5 example in Figure 1b. The neighborhood of an individual agent is 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 than relative positions characterize this two dimensional (2D) setup, and agents can only move to empty positions.[8]

The preferences considered in Schelling [1971a] are the same as for the one-dimensional model (each agent accepts up to 50% of unlike neighbors), whereas Schelling [1971b, 1978] also considers the possibility that agents accept up to 2/3 of unlike neighbors. The starting state is typically highly integrated.[9] The best-response dynamics, then, work as follows. All unsatisfied agents are put on a list in some arbitrary order. When an agent's turn comes, 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. Again, the typical outcome is a highly segregated state.

[7] Alternatively, one could interpret the 1D model as one with a continuous action space, and the agents taking up a neglegible amount of space themselves, such that there is always space between any two agents available.

[8] 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 (column) would see their neighborhood altered, as would all agents on the adjacent rows (columns).

[9] Schelling [1971a] starts with a random initial distribution of agents, while Schelling [1971b, 1978] creates the starting configuration by reshuffling a perfectly integrated board.

2.2 Schelling's model revisited

Many details of Schelling [1969, 1971a, 1971b, 1978] can be varied, but our focus is on the driving force behind the dynamic process, the individual preferences, while other details are kept as simple and general as possible. We consider 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. Our analysis is invariant to any positive monotonic transformation of these utility functions, which form a logical order.

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:

$$u(x) = \begin{cases} a + d\,(50 - |x - 50|) & \text{for } x<50 \text{ and } 50<x\leq c \\ a + d\,(50 - |x - 50|) + b & \text{for } x=50 \\ 0 & \text{for } x>c \text{ and if } x \text{ not defined (i.e., no neighbors),}[10] \end{cases}$$

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

Fig. 2a. flat utility

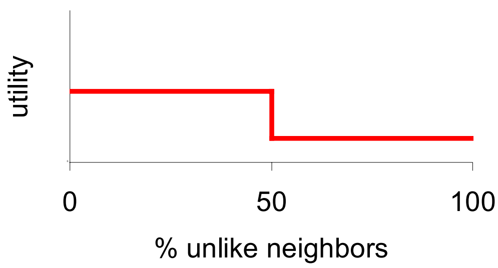


Fig. 2b. p50 utility

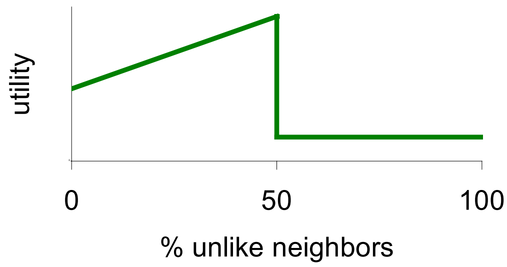


Fig. 2c. p100 utility

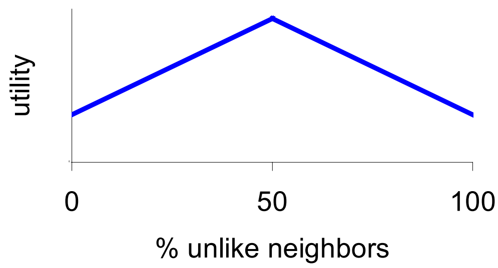
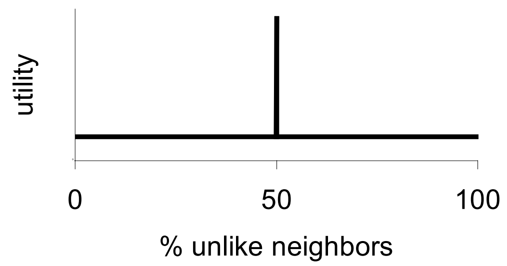


Fig. 2d. spiked utility



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

An essential change we introduce next concerns the rising part up to a peak at a 50-50 neighborhood (d>0), while retaining the cut-off point and the flat part beyond it, i.e., 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 peaked utility function (d>0) with the cut-off point removed (c=100). This is represented by the perfectly symmetric single-peaked utility function p100 in Figure 2c. 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.

Finally, we consider a spiked utility function that emphasizes the strict preference for perfect integration by rendering an individual indifferent with respect to any other configuration, setting a=d=0 and b>0 (see Figure 2d). Here agents are driven exclusively by their obsession to live in a perfectly integrated neighborhood.

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. 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. That is, a player will choose a best-response with probability one, but in case of indifference between more

[10] 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.

than one optimal position (possibly including his current position), a player chooses (uniform) randomly among all positions offering the highest utility.

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.[11] We simply select at each stage one agent uniform randomly, and ask him to choose a best-response.[12] We will separately analyze the possibility that all agents simultaneously choose a best-response to the current state.[13]

Next, we specify the spatial setting. In all setups analyzed, we consider neighborhoods defined in terms of the eight nearest surrounding positions (for agents in the interior). 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 2D 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. 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 2D torus, is relatively natural (e.g., in the form of a ring road in a city, chairs around a table, the shoreline of an island or lake, or the 24 hours around the clock). 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 also the case for even k. Furthermore, with the peaked and spiked utility functions that we consider *all* (pure strategy) equilibria disappear for *any* k.[14]

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.

---

[11] 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.

[12] 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.

[13] In case of conflicting choices in the 2D 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 1D setup, because of the conceptual difficulties arising from conflicting choices.

[14] 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.

3.1 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. This equilibrium is myopic since the dynamic structure of the game is disregarded.[15]

Definition 1: Let $z_i \in Z_i$ be the location of player i, such that $z_i \neq z_j$ for $i \neq j$, and let $v_i(z)$ be the payoff of agent i from z. 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 all $z_i \in Z_i$ such that $z_i \neq z_j^*$ for $i \neq j$ for each i.

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: In a random allocation 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

We use the following measures of segregation.

Clusters: This measure counts the number of clusters that can be distinguished. Two agents 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.[16] The one-dimensional version of the cluster measure is a straightforward simplification, as there are no blanks. A cluster measure, which is equivalent to the average cluster size, was used in Schelling [1969, 1971a, 1971b, 1978], without being

[15] In other words, the equilibrium is in locations rather than strategies.

[16] 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.

formally defined. 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, and make one full turn to observe all his neighbors. Let m_i be the number of agent 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 neighborhoods are, as seen by the individual agents.[17]

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: $m_i = \left| 0.50 - g_i/(g_i + f_i) \right|$, 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: $(1/N) \sum_i (m_i)$. In the 2D setup, the mix deviation differs from the switch rate in two senses. First, for the mix deviation an agent surrounded by XXXXOOOO agents (in such order) would be in the same situation as an agent surrounded by a XOXOXOXO neighborhood, whereas the switch rate would pick up the difference. Second, for the switch rate a XXXOXXXX neighborhood would be the same as XXOOOOXX, whereas the mix deviation would spot the 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.

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

---

[17] 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.

based measure, because this gives us an exogenous criterion to assess the consequences of different individual utility functions.[18]

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.[19] Figure 3 shows the distribution of clusters for a million random allocations. The distribution is rather symmetric, with the 20 agents forming on average 7.8 clusters.

Fig. 3. Random allocations, 5x5 board

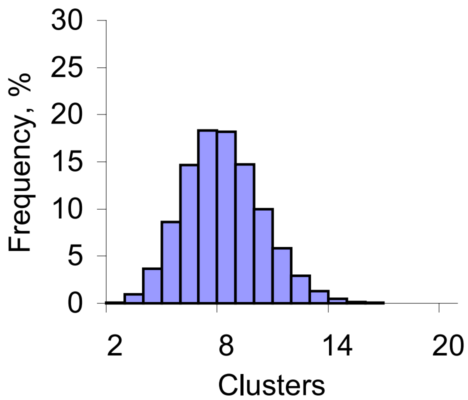


Table 1 shows the number of existing MNE for the flat and two peaked (p50 and p100) utility functions.[20] The number of MNE is of a similar order of magnitude for the flat and the

[18] 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. Typically, the literature on segregation uses measures designed on the basis of pre-defined and distinguishable neighborhoods (e.g., Census tracts). These measures can be classified along five dimensions: evenness (dissimilarity), exposure (isolation), concentration, centralization, and clustering (see, e.g., Cutler & Glaeser [1997], Cutler et al. [1999], Frankel & Volij [2004], Massey & Denton [1988], White [1983], or White [1986]). While there are no pre-defined neighborhoods in the Schelling model, our mix deviation and share measures can be seen as measures of evenness, our ghetto and switch rate measure exposure, and our cluster and distance measures concern clustering.

[19] The reason to start with a 5x5 board is its tractability. In the analysis of the 2D we always allocate 40% of locations to each type, leaving 20% empty.

[20] 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.

p100 utility functions, whereas it is much lower for the p50 function.[21] 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

	MNE		
	flat utility	***p50***	***p100***
non-strict MNE	430,110	2880	351,472
strict MNE	0	0	36,482
total MNE	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.[22] 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 4.0 clusters on average.

[21] When counting allocations, those obtained by swapping Os and Xs are not distinguished. Mirrored or rotated allocations are, however, distinguished because, for example, the number of distinct rotations depends on the degree of symmetry of a particular allocation.

[22] If it is not possible to find a better location for an agent with the p50 utility function, then it is also impossible 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, whereas an agent with the p50 utility function would deviate to that empty position.

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

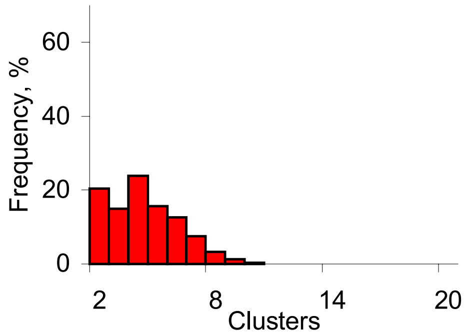


Fig. 4b. MNE, p50 utility, 5x5 board

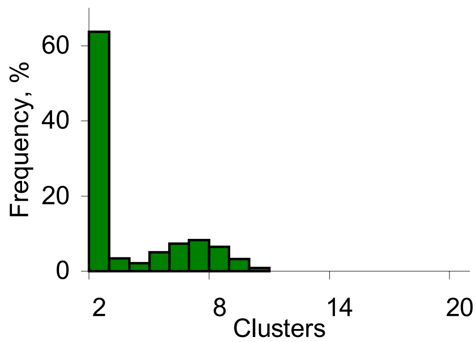


Fig. 4c. MNE, p100 utility, 5x5 board

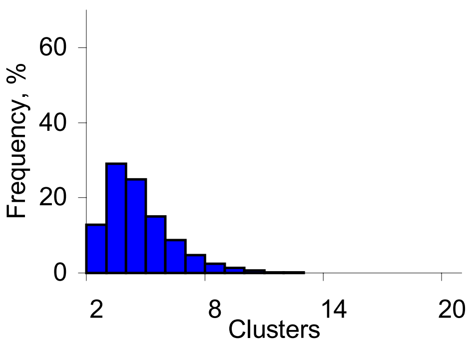


Fig. 5a. BR, 100,000 periods, flat utility, 5x5 board

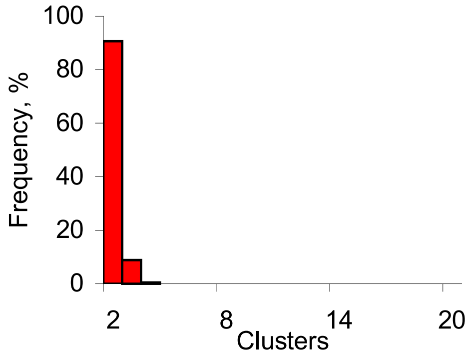


Fig. 5b. BR, 100,000 periods, p50 utility, 5x5 board

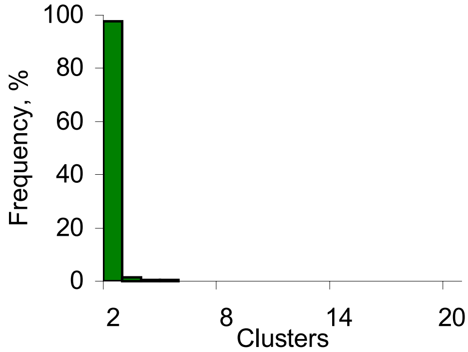
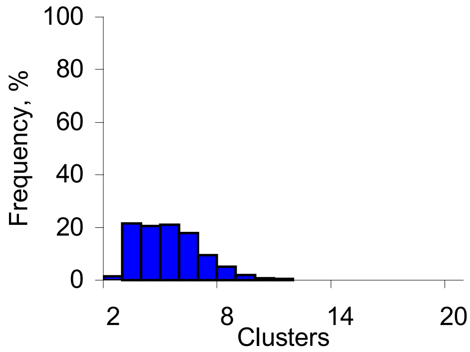


Fig. 5c. BR, 100,000 periods, p100 utility, 5x5 board



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 clusters for 1000 runs after

100,000 periods.[23] 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, 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. In Figure 5c, 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 diagrams 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.[24]

Table 2. Final distributions, 5x5 board

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

[23] 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 weakly preferred location.

[24] The reason to emphasize the clusters measure is that it seems to capture the notion of segregation best at the intuitive level.

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

	BR dynamics (100,000 periods)		
	flat	***p50***	***p100***
Observations	1000	1000	1000
non-strict MNE	404	370	145
strict MNE	n.a.	n.a	854
total MNE	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 random 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 100,000th period as evidence of the limiting behavior is reasonable, 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 already 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 p50 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

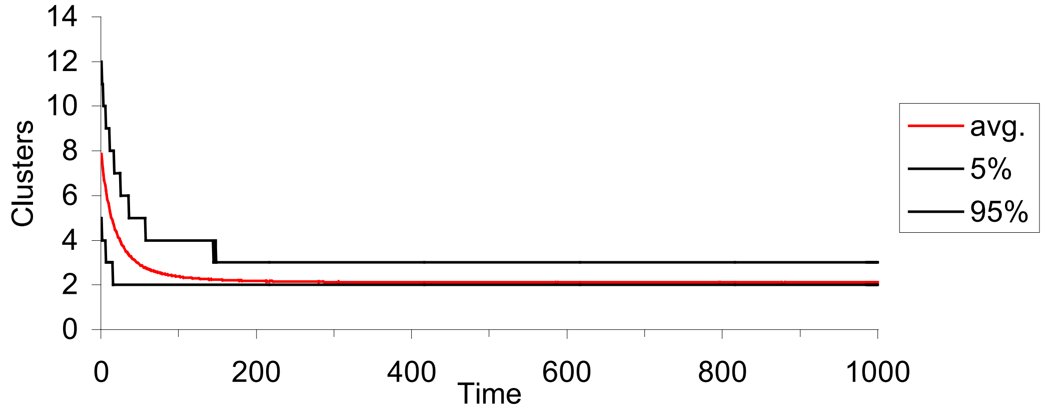


Fig. 6b. BR dynamics, p50 utility, 5x5 board, 10,000 runs

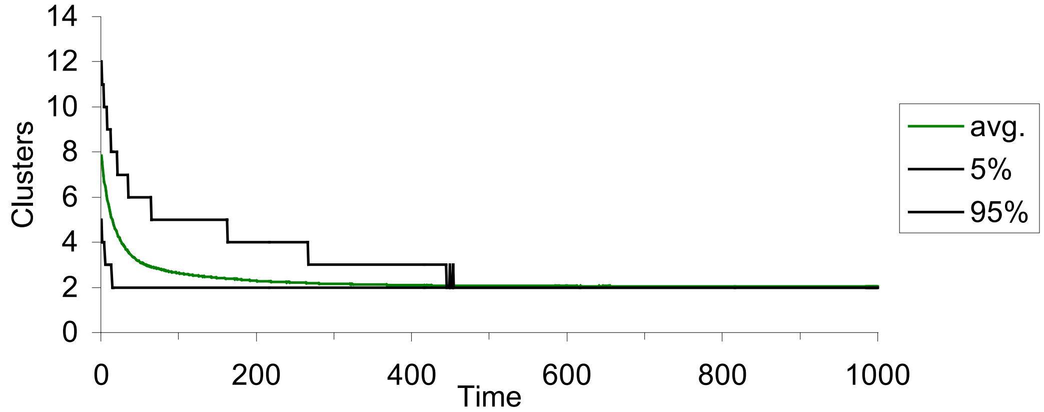
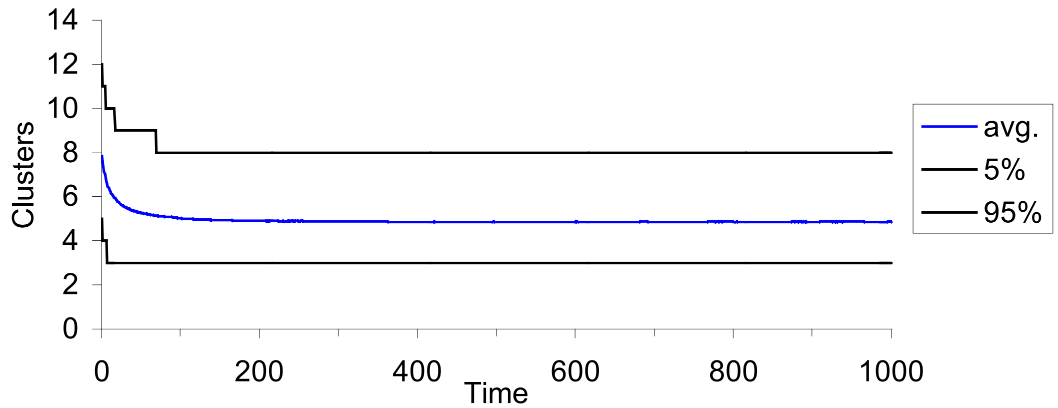


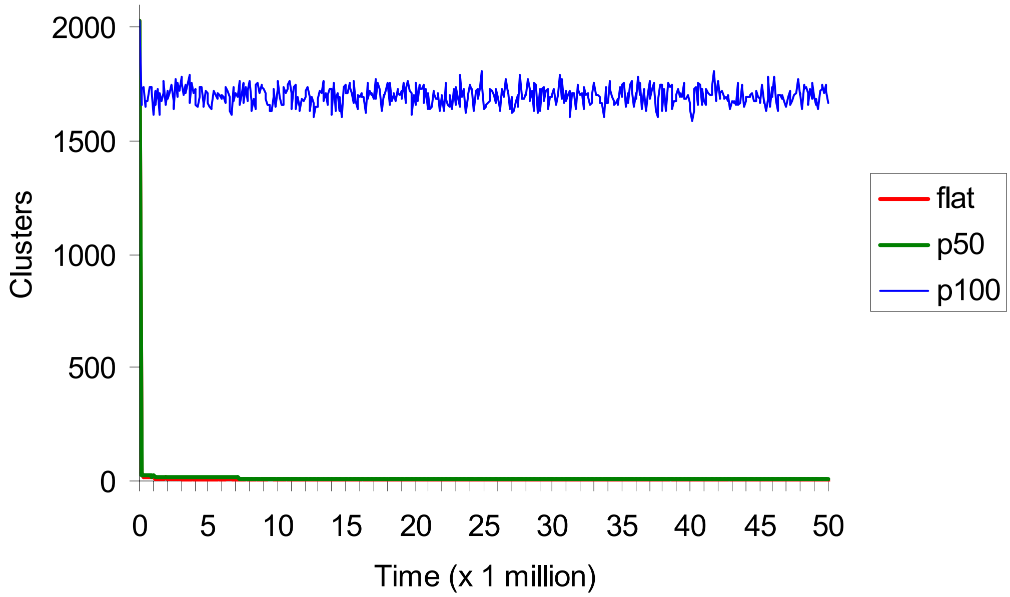
Fig. 6c. BR dynamics, p100 utility, 5x5 board, 10,000 runs



4.2 100x100 board

We now turn to a larger board. There are 4000 agents of each type on a 100x100 board. The distinction between a board and a torus is blurred as a board becomes bigger, because the share of agents at borders declines. Figure 7 plots the clusters segregation measure against time under BR moves for a 100x100 board. The three representative 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 p50 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 almost 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.[25]

Since the degree of segregation in the long run, according to the above diagram, is rather stable, a look at a particular final allocation could be informative. Figures 8a to 8c show a random board and two final allocations (after 50 million periods) for the flat and the p50 utility functions.[26] 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 more alien 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.

Fig. 8a. random allocation, 100x100 board

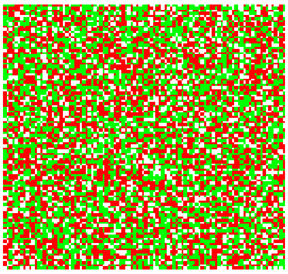


Fig. 8b. BR, 50m periods, flat utility, 100x100 board

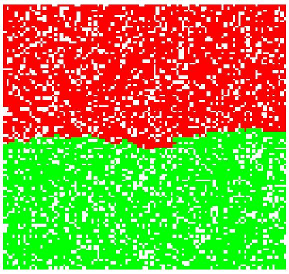
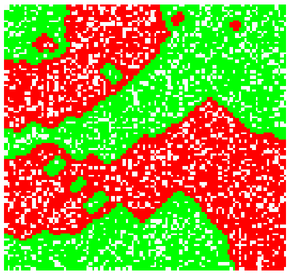


Fig. 8c. BR, 50m periods, p50 utility, 100x100 board



type X   type O   blank

4.3 Discussion of findings in 2D setup

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, and without inertia. In Pancs & Vriend [2003], we establish the robustness of the obtained results by introducing in turn inertia, simultaneous moves, and a torus. For each of these three cases, the following two questions are addressed. First, was the original assumption made by Schelling [1969, 1971a, 1971b, 1978] on inertia, a preference for nearby positions, and the specific way in which the order of moves was determined an essential element of his model of segregation? Second, does the introduction of this variation concerning 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 change our

[25] The switch measure is not sensitive to such small clusters, and in addition it shows the difference between the flat and p50 function immediately (see Pancs & Vriend [2003]).

[26] The board for the p100 function looks very similar to a random one, just as Figures 7 suggests, and is omitted.

findings concerning the relative effect of the utility functions with a strict preference for integration relative to Schelling's flat utility? In each of the three cases the answer to both questions is negative (see Pancs & Vriend [2003] for an extensive analysis).

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 and without a preference for nearby positions exhibits rapid segregation, even when the flat utility function is strengthened to imply strict preferences for perfect integration.

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.

This finding is consistent with Zhang [2004], who extends the analysis of Young [1998] for the 1D setup to two dimensions. Both Zhang [2004] and Young [1998] consider a modification of Schelling's one-dimensional model, with agents swapping locations, compensating payments between moving players, and in the presence of noise (mistakes). They argue that complete segregation is the only viable long-run outcome of the best-response dynamics if the agents' preferences are biased in favor of their own type.

The concept of a 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,[27] 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

[27] 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 Pancs & Vriend [2003] shows, this is not the case.

already biased towards segregation, especially for the p50 function, the BR dynamics favor the most segregated among them.

5. Analysis of one-dimensional setup

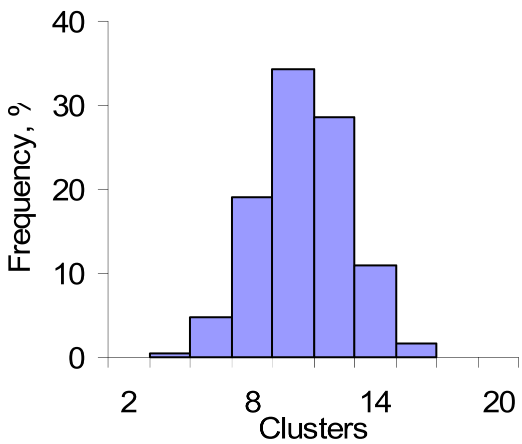
We now turn to an analysis of the 1D version of Schelling's spatial proximity model. In this section we present a number of sharp theoretical results on segregation in 1D. In addition, a numerical analysis serves to highlight implications of alternative specifications of the model, and to demonstrate that asymptotic theoretical results are attained in finite time.

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 most of 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 9 shows the distribution of the cluster measure for all 92,378 possible allocations.[28] On average the 20 agents are located in 10.5 clusters.

Fig. 9. 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, because there are fewer possible allocations (since there are no

[28] Again, we neglect equivalent allocations obtained by swapping Os and Xs.

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.

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

	MNE	
	flat utility	***peaked or spiked***
non-strict MNE	28	18
strict MNE	0	0
total MNE	28	18

Figures 10a and 10b 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 the peaked (p50) utility function 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.

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

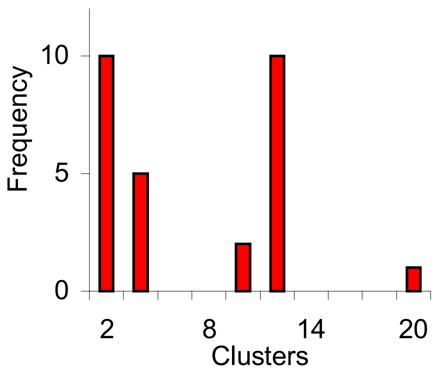


Fig. 10b. MNE, peaked or spiked utility, 10+10 ring

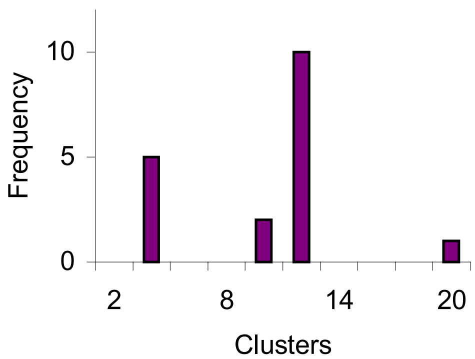


Fig. 11a. BR, 100,000 periods, flat utility, 10+10 ring

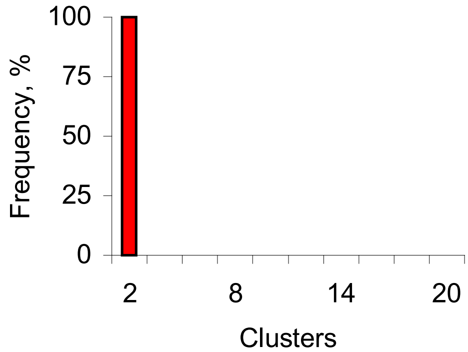
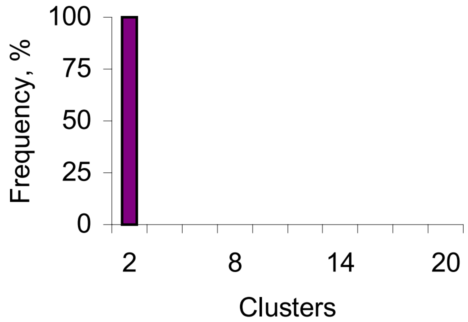


Fig. 11b. BR, 100,000 periods, peaked or spiked utility, 10+10 ring



Figures 11a and 11b 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 always selecting a particular MNE. However, with any of the peaked or spiked utility functions there is no MNE corresponding to this outcome.[29] 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.

Since complete segregation is not a MNE with the peaked or spiked utility functions, it cannot be a steady state. 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 summarizes the characterization of all possible allocations as well as those of the MNE and the BR outcomes using various segregation measures. MNE for the peaked and spiked utility functions are perfectly integrated according to the mix deviation, share and ghetto measures,[30] 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.

[29] 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.

[30] Although this cannot be read from the table, this applies in fact to each MNE for the peaked and spiked utility functions.

Table 12. 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
Clusters	*Average*	10.53	7.14	10.00	2.00	2.00
	5%	6	2	4	2	2
	95%	14	12	12	2	2
Distance	*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
Mix dev.	*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
Share	*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
Ghetto	*Average*	0.00	1.43	0.00	4.00	4.00
	5%	0	0	0	4	4
	95%	0	4	0	4	4

Fig. 12a. BR dynamics, flat utility, 10,000 runs, 10+10 ring

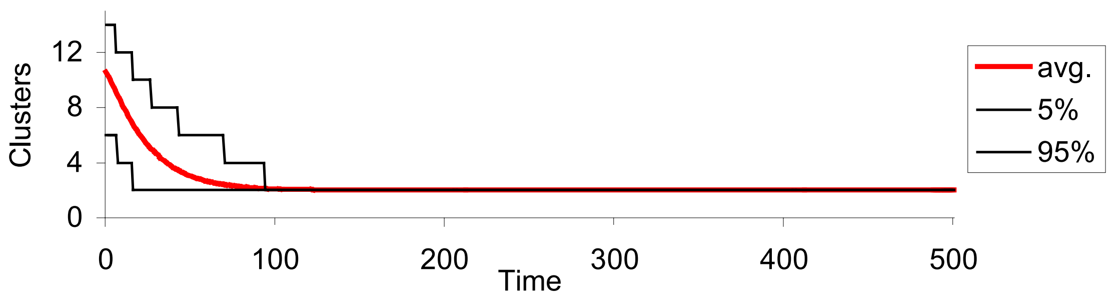
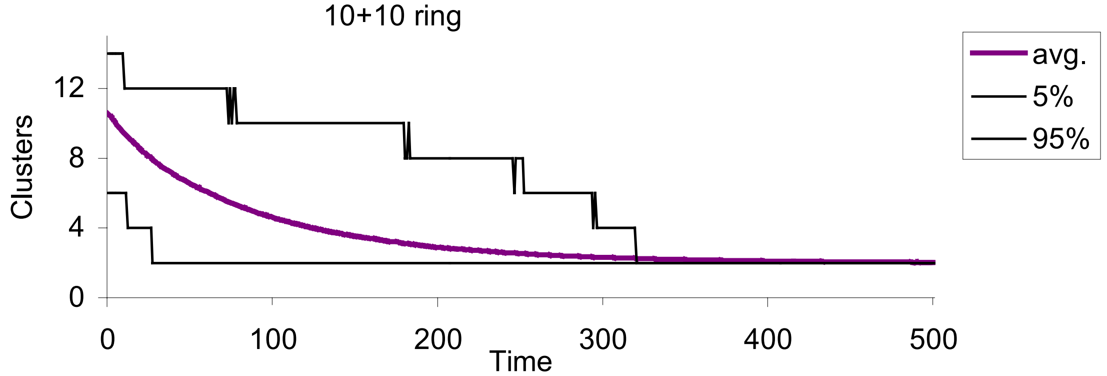


Fig. 12b. BR dynamics, peaked or spiked utility, 10,000 runs, 10+10 ring



Figures 12a and 12b 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 A1 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 p100) utility functions.

PROPOSITION A1. 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 class consisting of all completely segregated states.

DISCUSSION. The complete proof can be found in Appendix A. Here is a sketch. We need to show that it is possible to reach complete segregation from each initial configuration with positive probability, while all completely segregated states constitute a single recurrent class. On a completely segregated ring, there are two borders between ghettos. These borders are the only locations that offer positive utility. Hence, any agent can only move to such a border, which does not affect the 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, we show that it will eventually emerge whatever the initial allocation. Hence, the limit outcome of the BR dynamics is complete segregation.

COROLLARY 1. Proposition A1 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 A1 (see Appendix A). □

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 Appendix B.

On a ring there exist at least two positions where an agent can insert itself and enjoy perfect integration (Lemmas A2 and B1). This implies that to be a MNE it is necessary and sufficient that all agents enjoy the highest possible utility (Proposition B1), 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 (including the flat utility function), 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 B2), 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 B3). 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 implies 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.

The sets of MNE for the spiked and two peaked utility functions coincide (Proposition B6). This set is a subset of MNE with the flat utility function (Proposition B7). 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 B6 shows that BR moves for all utility specifications with a strict preference for integration are identical.

The following properties allow 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 all locations i and i+k+1 are of opposite types (Proposition B4). 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 $a_i + a_{i+1} + ... + a_{i+k-2} + a_{i+k-1} = 0$, where a_j=1 if an agent in the jth position is X and a_j= −1 otherwise (Proposition B5). 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 alternating MNE are robust to the length of a ring.[31]

If k=4, as it is throughout this Section, then all equilibria described in Proposition B5 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.[32]

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. Proposition A1 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 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 B3, 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 13a and 13b 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

[31] 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.

[32] 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}%.

Figure 10b once the horizontal axis is scaled up by a factor 10, and is therefore not shown here.

Fig. 13a. All possible allocations, 100+100 ring

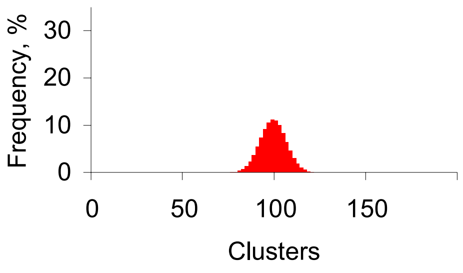
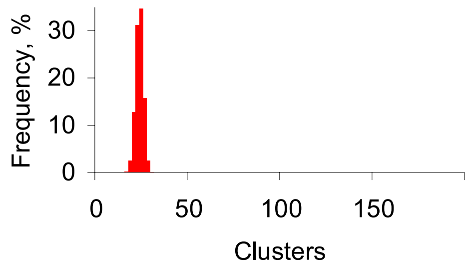


Fig. 13b. MNE, flat utility, 100+100 ring



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 14a shows the first 250,000 periods of a single run of 50 million periods, while Figure 14b depicts the same for the peaked and spiked utility functions.[33] 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. Thus, with the peaked or spiked utility functions, 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 or spiked utility functions, whereas they never seem to occur with the flat utility function.

[33] As argued in the previous section, in 1D the BR dynamics are identical for each of the peaked or spiked functions considered in this paper.

Fig. 14a. BR dynamics, flat utility, 1 run, 100x100 board

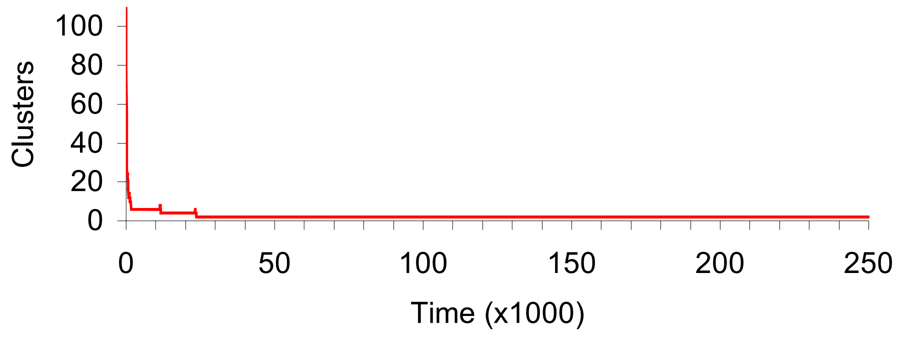
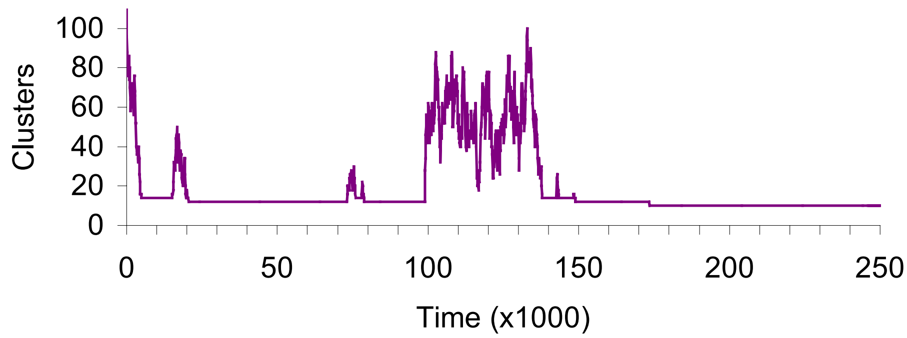


Fig. 14b. BR dynamics, peaked or spiked utility, 1 run, 100x100 board



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. 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 15a 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 a new opportunity to migrate as well. Figure 15b offers a close-up of this process. This increases the number of optimal locations within the growing 'mushroom', draining other clusters. Eventually this integration disappears.

Fig. 15a. BR dynamics, peaked or spiked utility function, 1 run, 100x100 ring

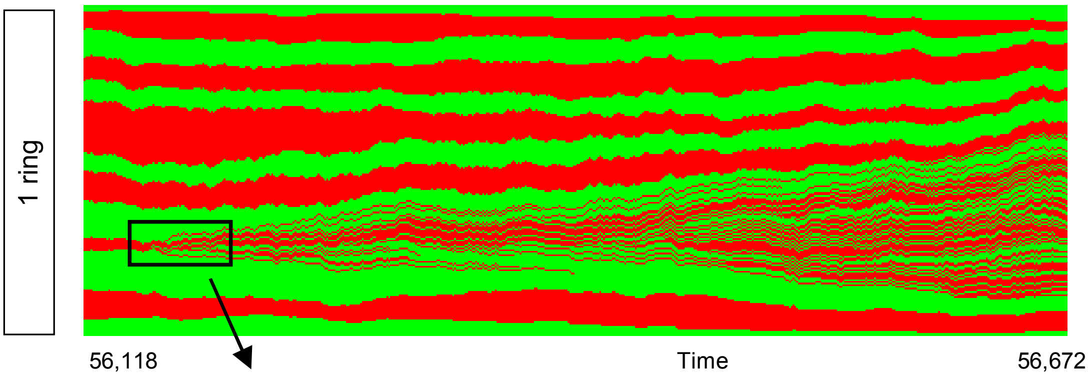
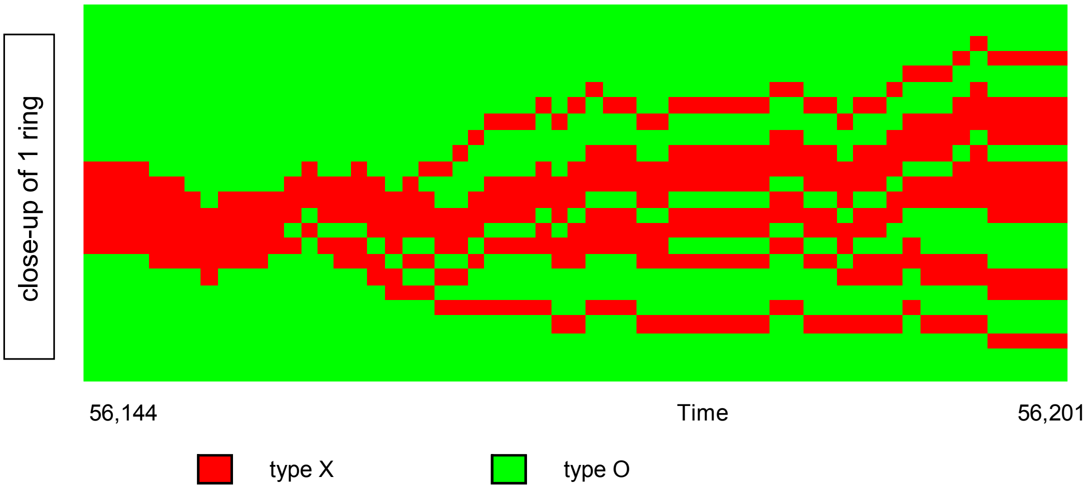
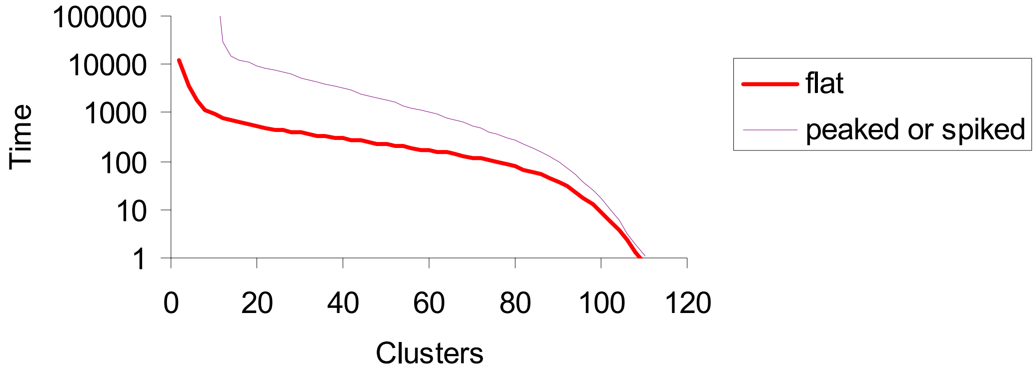


Fig. 15b. Close-up of BR dynamics, peaked or spiked utility function, 1 run, 100x100 ring



The more segregrated 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. 16. BR dynamics, average 'first passage times', 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 no more than *x* clusters in that run.[34] An analysis of first passage times abstracts from the outbursts shown above. Figure 16 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.

5.4 Discussion of findings in one-dimensional setup

In Pancs & Vriend [2003] we also 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. 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? Again, we find that the answer to each of these questions is negative.[35]

---

[34] 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 $x=102$ and $x=103$ is 2.

[35] The main change that comes with a line is the non-existence of MNE for the peaked and spiked utility functions (see Pancs & Vriend [2003]), whereas the only MNE with the flat utility function consists of complete segregation.

This robustness allowed us to abstract from superfluous details on the order of moves, assume away inertia and a preference for nearby positions, 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. Here 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 to get complete *segregation* is that it implies a strict preference for perfect *integration*. Given that, the utility function may have multiple peaks, and it may even describe a preference for living in any minority neighborhood rather than in any majority neighborhood.[36]

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. Hence, 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.[37] 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.

Appendix A. 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 form a unique recurrent class while all other states be transient.

On a completely segregated ring, there are two borders between ghettos. These borders are the only locations that offer positive utility. Hence, any agent can only move to such a border, which does

[36] An example would be the horizontal mirror image of the peaked p50 utility function. Such features seem relevant from a biological perspective, where species might want to avoid living with too many like competitors.

[37] 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 nature of the moves as such that constitutes the essential difference.

not affect the 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. PPMs 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 assured. 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.

· $\{\mathrm{X}\}_n$ is a segment consisting of n Xs.

· Square brackets […] are used to highlight a seed.

· $\{\mathrm{X},\mathrm{O}\}_{l,n}$ is any segment consisting of l Xs and n Os in any arbitrary order.[38]

LEMMA A1. Any segment of the type $\{\mathrm{X},\mathrm{O}\}_{k-l,l}\mathrm{X}\{\mathrm{X},\mathrm{O}\}_{l,k-l}$, where $2k$ is the size of a neighborhood and $0 < l < k$, can be transformed into $\{\mathrm{O}\}_l\{\mathrm{X}\}_{k+1}\{\mathrm{O}\}_{k-l}$ 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,

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

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. □

LEMMA A2. 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 σ_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 $\sigma_i \in \{0,1,...,2k\}$ and $i \in \{1,2,...,2m\}$. 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

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

The mapping σ has a property for all i that

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

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 $|\Delta\sigma_i| = 0$

[38] For instance, $\{\mathrm{X},\mathrm{O}\}_{2,3}$ could be XOOOX, as well as OOXOX or OXXOO.

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

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. □

LEMMA A3. Any segment of the type $\{\mathrm{O}\}_l\{\mathrm{X}\}_{k+1}\{\mathrm{O}\}_{k-l}$, 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 $[\{\mathrm{O}\}_k\{\mathrm{X}\}_k]$ by means of positive probability moves.

PROOF. The proof is done by induction.
Inductive base: segment $\{\mathrm{O}\}_{k-1}\{\mathrm{X}\}_{k+1}\mathrm{O}$ can be preceded by either O or X. The former case yields a seed $[\{\mathrm{O}\}_k\{\mathrm{X}\}_k]\mathrm{XO}$.
Consider the latter case of $\mathrm{X}\{\mathrm{O}\}_{k-1}\{\mathrm{X}\}_{k+1}\mathrm{O}$. By Lemma A2 there exist at least two positions with positive utility. At the same time between any of the two Xs, which are inside the fragment $\{\mathrm{X}\}_{k+1}$ and immediately to the left of $\{\mathrm{X}\}_{k+1}$: $\sigma > k$. Consequently, $\sigma = k$ in at least two places outside this fragment.
a) If $\sigma = k$ immediately to the right of $\{\mathrm{X}\}_{k+1}$, then we have a seed.
b) If $\sigma = k$ anywhere enclosed between horizontal bars $\mathrm{X}|\{\mathrm{O}\}_{k-1}|\{\mathrm{X}\}_{k+1}\mathrm{O}$ (bars included), then by PPM we insert there O and obtain a seed $\mathrm{X}[\{\mathrm{O}\}_k\{\mathrm{X}\}_k]\mathrm{XO}$.
c) If $\sigma = k$ is outside the segment $\mathrm{X}\{\mathrm{O}\}_{k-1}\{\mathrm{X}\}_{k+1}\mathrm{O}$, then the following action will be a positive probability move: $\mathrm{X}\{\mathrm{O}\}_{k-1}\mathrm{X}\{\mathrm{X}\}_k\mathrm{O} \rightarrow \mathrm{X}\{\mathrm{O}\}_{k-1}\{\mathrm{X}\}_k\mathrm{O}$. Since the destination is outside the segment, the moved X could not have constituted a part of its neighborhood. Finally, $\mathrm{X}\{\mathrm{O}\}_{k-1}\{\mathrm{X}\}_k\mathrm{O} \rightarrow \mathrm{X}\{\mathrm{O}\}_{k-1}\mathrm{XO}\{\mathrm{X}\}_{k-1}\mathrm{O} \rightarrow \mathrm{X}[\{\mathrm{O}\}_k\{\mathrm{X}\}_k]\mathrm{O}$.

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

Consider the latter case when X precedes (**): $\mathrm{X}\{\mathrm{O}\}_{l-1}\{\mathrm{X}\}_{k+1}\{\mathrm{O}\}_{k-l+1}$. As with the inductive step, we know that inside $\{\mathrm{X}\}_{k+1}$ and immediately to the left of $\{\mathrm{X}\}_{k+1}$ we have $\sigma > k$.
a) If $\sigma = k$ immediately to the right of $\{\mathrm{X}\}_{k+1}$, then we already have a seed.
b) If $\sigma = k$ anywhere enclosed between horizontal bars $\mathrm{X}|\{\mathrm{O}\}_{l-1}|\{\mathrm{X}\}_{k+1}\{\mathrm{O}\}_{k-l+1}$ (bars included), then by PPM move there O from the right hand side so that the resulting segment $\mathrm{X}\{\mathrm{O}\}_l\{\mathrm{X}\}_{k+1}\{\mathrm{O}\}_{k-l}$ contains (*).

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

i) A seed has formed: $\mathrm{X}\{\mathrm{O}\}_{l-1}\mathrm{X}\left[\{\mathrm{X}\}_k\{\mathrm{O}\}_k\right]$.

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

iii) If there is no position in the rightmost $\{\mathrm{O}\}$ with $\sigma = k$, then the following is possible: $\mathrm{X}\{\mathrm{O}\}_{l-1}\{\mathrm{X}\}_{k-l}\mathrm{X}\{\mathrm{X}\}_l\{\mathrm{O}\}_{k-l+1} \to \mathrm{X}\{\mathrm{O}\}_{l-1}\{\mathrm{X}\}_{k-l}\{\mathrm{X}\}_l\{\mathrm{O}\}_{k-l+1} \to \mathrm{X}\{\mathrm{O}\}_{l-1}\{\mathrm{X}\}_{k-l+1}\mathrm{O}\{\mathrm{X}\}_{l-1}\{\mathrm{O}\}_{k-l+1} \to \{\mathrm{O}\}_{l-1}\{\mathrm{X}\}_{k-l+1}\mathrm{O}\{\mathrm{X}\}_l\{\mathrm{O}\}_{k-l+1} \to \{\mathrm{O}\}_{l-1}\{\mathrm{X}\}_{k-l}\mathrm{O}\{\mathrm{X}\}_{l+1}\{\mathrm{O}\}_{k-l+1} \to \{\mathrm{O}\}_{l-1}\{\mathrm{X}\}_{k-l-1}\mathrm{O}\{\mathrm{X}\}_{l+2}\{\mathrm{O}\}_{k-l+1} \to \ldots \to \{\mathrm{O}\}_{l-1}\mathrm{XO}\{\mathrm{X}\}_k\{\mathrm{O}\}_{k-l+1} \to \{\mathrm{O}\}_l\{\mathrm{X}\}_{k+1}\{\mathrm{O}\}_{k-l}\mathrm{O}$.

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

PROPOSITION A1. 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 class 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, some agent will go to a location that ensures one (it exists by Lemma B1). 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 $\{\mathrm{X},\mathrm{O}\}_{k-l,l}\mathrm{X}\{\mathrm{X},\mathrm{O}\}_{l,k-l}$. By Lemma A1 this can be transformed into $\{\mathrm{O}\}_l\{\mathrm{X}\}_{k+1}\{\mathrm{O}\}_{k-l}$ and then by Lemma A3 into $\left[\{\mathrm{O}\}_k\{\mathrm{X}\}_k\right]$ with positive probability. Thus a seed can always be constructed and complete segregation then built. On a completely segregated ring, there are two borders between ghettos. These borders are the only locations that offer positive utility. Hence, any agent can only move to such a border, which does not affect the integrity of the ghettos. □

Appendix B. On MNE on a ring

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

LEMMA B1. 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 σ_i and its properties from Lemma A2. Without loss of generality assume that there are m Os and m–1 Xs. The following equality will hold:

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

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 σ_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

$(2m-2)(k-1) + k \geq 2k(m-1)$ 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 $\sigma_i = k$. □

PROPOSITION B1. 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 B1 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. □

PROPOSITION B2. 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 B1 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 B1), 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. □

PROPOSITION B3. 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 B1 this contradicts the assumption that the cluster is a part of a MNE allocation. Hence, no cluster can be larger than *k*+1. □

PROPOSITION B4. 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 2*k*+1 agents $a_0 a_1 ... a_k ... a_{2k}$, 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 B1), 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 2*k*+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.[39]

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 $a_{i-2k-1}a_{i-2k}...a_{i-k-1}a_{i-k}...a_{i-1}a_i$ 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 $z \in Z$. 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

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

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

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

Take a fragment of the string as given below

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

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)

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

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)

$$a_0a_1...a_k...a_{2k-1}a_{2k}$$

[39] 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.

The condition for the middle agent to be satisfied is

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

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. □

PROPOSITION B5. 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 $a_i + a_{i+1} + ... + a_{i+k-2} + a_{i+k-1} = 0$ for all i.

PROOF. The proof draws heavily on the proof of Proposition B4. In the necessary part of the proof of Proposition B4 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 α_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) $a_0 a_1 ... a_k ... a_{2k-1} a_{2k}$. 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. □

LEMMA B2. 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.

Let $a_i = -1$ if location i is occupied by X and $a_i = 1$ otherwise, then

$$\Omega_{X'} = \Omega_{O'} + 2 + a' - a'' \le 0 \text{ and } \Omega_{O'} \ge 0, \qquad (*)$$

where $\Omega_{X'}$ and $\Omega_{O'}$ are 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. □

PROPOSITION B6. 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 B1 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 B1, 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). □

PROPOSITION B7. 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 an 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):

$$...\mathrm{X}\{\mathrm{O}\}_{<k}\dot{\mathrm{O}}\{\mathrm{X}\}_{\geq k+1}...$$

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 B2, in MNE would enjoy a perfectly integrated neighborhood, if such a MNE exists. If *k*=3, again, by Lemma B2, 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 B2, 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. □

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.
Clark, W.A.V. (1991). Residential Preferences and Neighborhood Racial Segregation: A Test of the Schelling Segregation Model. *Demography*, *28*, 1-19.
Cutler, D.M., & Glaeser, E.L. (1997). Are Ghettos Good or Bad? *Quarterly Journal of Economics*, *112*, 827-872.
Cutler, D.M., Glaeser, E.L., & Vigdor, J.L. (1999). The Rise and Decline of the American Ghetto. *Journal of Political Economy*, *107 (3)*, 455-506.
Dixit, A.K., & Nalebuff, B.J. (1991). *Thinking Strategically. The Competitive Edge in Business, Politics, and Everyday Life*. New York: Norton.
The Economist (2001). Britain: Living apart, June 30th, 52-53.
Frankel, D., & Volij, O. (2004). *Measuring Segregation* (mimeo).
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.
Massey, D.S., & Denton, N.A. (1988). The Dimensions of Residential Segregation. *Social Forces*, *67 (2)*, 281-315.
Pancs, R., & Vriend, N.J. (2003). *Schelling's Spatial Proximity Model of Segregation Revisited* (Dept. of Economics Working Paper No. 487) Queen Mary, University of London.
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.
White, M.J. (1983). The Measurement of Spatial Segregation. *American Journal of Sociology*, *88 (5)*, 1008-1018.
White, M.J. (1986). Segregation and Diversity Measures in Population Distribution. *Population Index*, *52*, 198-221.
Young, H.P. (1998). *Individual and Social Structure: An Evolutionary Theory of Institutions*. Princeton, NJ: Princeton University Press.
Zhang, J. (2004). Residential Segregation in an All-integrationist World. *Journal of Economic Behavior & Organization*, *54*, 533-550.