GROUPE D'ANALYSE ET DE THÉORIE ÉCONOMIQUE LYON - ST ÉTIENNE

halshs-00502758, version 1 - 15 Jul 2010

W P 1017

**Dynamic models of residential ségrégation :
an analytical solution**

Sebastian Grauwin, Florence Goffette-Nagot, Pablo Jensen

Juillet 2010

Documents de travail | Working Papers

GATE Groupe d'Analyse et de Théorie Économique Lyon-St Étienne

93, chemin des Mouilles 69130 Ecully – France
Tel. +33 (0)4 72 86 60 60
Fax +33 (0)4 72 86 60 90

6, rue Basse des Rives 42023 Saint-Etienne cedex 02 – France
Tel. +33 (0)4 77 42 19 60
Fax. +33 (0)4 77 42 19 50

Messagerie électronique / Email : gate@gate.cnrs.fr
Téléchargement / Download : http://www.gate.cnrs.fr – Publications / Working Papers

DYNAMIC MODELS OF RESIDENTIAL SEGREGATION: AN ANALYTICAL SOLUTION

Sébastian GRAUWIN[a,b,c], Florence GOFFETTE-NAGOT[a,d,*], Pablo JENSEN[a,b,c,e]

[a]*Université de Lyon, Lyon, F-69007, France*
[b]*Institut rhônalpin des systèmes complexes, IXXI, Lyon, F-69007, France*
[c]*ENS-LYON, Laboratoire de Physique, UMR 5672, Lyon, F-69007, France*
[d]*CNRS, GATE Lyon-St Etienne, UMR 5824, Ecully, F-69130*
[e]*CNRS, Laboratoire d'Économie des Transports (LET), UMR 5593, Lyon, F-69363, France*

Abstract

We propose an analytical resolution of Schelling segregation model for a general class of utility functions. Using evolutionary game theory, we provide conditions under which a potential function, which characterizes the global configuration of the city and is maximized in the stationary state, exists. We use this potential function to analyze the outcome of the model for three utility functions corresponding to different degrees of preference for mixed neighborhoods. Schelling original utility function is shown to drive segregation at the expense of collective utility. If agents have a strict preference for mixed neighborhoods but still prefer being in the majority *versus* in the minority, the model converges to perfectly segregated configurations, which clearly diverge from the social optimum. Departing from earlier literature, these conclusions are based on analytical results. These results pave the way to the analysis of many structures of preferences, for instance those based on empirical findings concerning racial preferences. As a by-product, our analysis builds a bridge between Schelling model and the Duncan and Duncan segregation index.

Keywords: Residential segregation, Schelling, dynamic model, potential function, social preferences.

JEL Codes: C63, C72, C73, D62, J15.

*Corresponding author. GATE, 93 chemin des Mouilles, F-69130 Ecully. Tel 33+472866078. Fax 33+472866090
Email addresses: sebastian.grauwin@ens-lyon.fr (Sébastian GRAUWIN), goffette-nagot@gate.cnrs.fr (Florence GOFFETTE-NAGOT), Pablo.Jensen@ens-lyon.fr (Pablo JENSEN)

1. Introduction

Ethnic and immigrant residential segregation is a striking feature of most Western cities. Extensive views of segregation patterns in the U.S. have been provided recently by Cutler *et al.* (2008), Iceland and Scopilliti (2008) and Reardon *et al.* (2008). Cutler *et al.* (2008) examine a range of potential determinants of immigrant segregation, including cultural traits of immigrants and nativist sentiment among U.S. natives. Card *et al.* (2008) results on racial segregation for the 1970-2000 period show evidence of tipping-like behaviors: the rise of the minority share in a neighborhood above a certain threshold leads to a further decrease in the white population. This analysis is one of the first providing clear empirical evidence of non linear dynamic aggregate behaviors, as those predicted by social interaction models. According to these results, whites' utility in a neighborhood seems to exhibit a sharp decrease beyond a certain minority share. A direct link between white attitudes toward minority members and aggregate configurations, as measured by the location of tipping point, is also shown by the authors. The theoretical relationship between individual preferences and aggregate configurations has however not been fully explored to date.

An early contribution was provided by Schelling, who proposed a model aiming at formalizing the aggregate consequences of individual preferences regarding the social environment (Schelling, 1969, 1971, 1978). The two basic ingredients of Schelling 1971 model (Dynamic models of segregation, *Journal of Mathematical Sociology*) are an individual utility function that determines entirely the level of satisfaction enjoyed by an agent in a location and a dynamical rule that drives agents' location changes and therefore the evolution of the city configuration. Using an inductive approach, Schelling showed that if the preferences considered are such that an environment of more than 50% of own-group agents is highly preferred to a less than 50% of own-group environment, then the equilibrium configuration exhibits high levels of segregation, although there is no preference for segregation *per se*. Schelling 1971 paper is widely known thanks to this apparently paradoxical effect: mild individual preferences for own-group neighbors lead to a complete segregation at the global scale. However, a moment of reflexion suffices to understand that, given the highly asymmetrical utility function used in this model, it could hardly lead to an integrated environment. Yet, later research showed that even a peaked utility function, that is, a function achieving its maximum for a perfectly mixed environment, can lead to a fully segregated equilibrium as soon as this function is asymmetric (Zhang, 2004b; Pancs and Vriend, 2007; Barr and Tassier, 2008).

Criticizing the realism of Schelling model is straightforward: it ignores institutional causes of segregation, income effects, or cities' social structure. Anyway, the model has become a favorite example, in the modeling of social systems, of the unintended macro-level consequences of individual behavior, and Schelling 1971 paper is his most widely cited publication (more than 460 as of 2010, June 10th). After years of relatively

low citation records, this paper accrues since 2003 around 40 citations per year, showing the renewed interest in Schelling model. It is interesting to notice that citations arise from widely different fields: economics and sociology represent the two strongest contributors (40% of the total number of citations) but computer science, mathematics and physics gather 24% of the citations. This substantial scientific activity has lead to new insights: the interpretation of the emergence of segregation patterns as the result of a coordination problem (Zhang, 2004a,b); a physical analogue of Schelling's model (Vinkovic and Kirman, 2006); the robustness of Schelling's results with respect to different definitions of individual utilities and/or environment (Pancs and Vriend, 2007; Fagiolo *et al.*, 2007); the impact of heterogeneous agents and public policies (O'Sullivan, 2009); the exploration of tipping behaviors (Zhang, 2010).[1] However, most of this work relies on agent-based simulations.

Attempts to solve Schelling model analytically include Dokumaci and Sandholm (2007); Mobius and Rosenblat (2000); Pollicott and Weiss (2001); Zhang (2004a,b, 2010). Zhang (2004a,b, 2010)'s contributions represent to date the closest achievement in this direction. He proposes variations of Schelling model which he analyzes formally using the concept of potential function developed in evolutionary game theory. Zhang (2004a) considers a model with vacant cells and linear utility functions. Zhang (2004b, 2010) use an asymmetrically peaked utility function in a model with no vacant cells. The latter choice raises the issue of individual rationality, as it is assumed that, for individual moves to occur, two agents have to coordinate and agree on exchanging locations. As such, this analysis departs from Schelling original framework. Furthermore, these three contributions only cover two specific utility functions.

In the present paper, we build on Zhang (2004a,b, 2010) and provide a more general framework, allowing to solve analytically dynamic segregation models for a broad range of utility functions. In contrast to Zhang (2004b, 2010), our model considers general utility functions and remains in Schelling spirit by including vacant cells. We can predict the global pattern emerging from different utility functions, which, to the best of our knowledge, was never done before. Our analysis also builds a connection with a commonly used segregation measure and draws a parallel with general concepts of coalitional games.

Following Zhang (2004a,b, 2010), we place Schelling model in the context of evolutionary game theory. In this context, a potential function of the game is a function defined at the city level, of which variation with an individual's move is equal to the change of utility of this individual. We consider bounded neighborhoods, *ie* blocks where all the agents share the same neighbors, which permits us to formalize the externalities for general utility functions. We show that a potential function of the model exists if and only if the utility functions verify a non limiting condition, an interpretation of which is that the externalities generated by

---

[1]See Clark and Fossett (2008) for a literature review.

one type of agents are symmetric to the ones generated by the other type of agents. The case with no vacant cells is a limit case of this condition. Under this condition, a general form of the potential function is found. The main property of this potential function is that it reflects both the macro and micro scale. On one hand, this aggregate function only depends on the number of agents of each type in each block. On the other hand, it keeps tracks of the individual level since it corresponds to a sum of individual utility changes generated by individual moves.

We use this potential function to characterize the segregation level of the stationary configurations of the model for different utility functions, representing different degrees of preference for mixity. We examine successively (i) linear utility functions, with a continuous preference for segregated environments, (ii) Schelling's original utility function in which there is a mild preference for a mixed environment and (iii) asymmetrically peaked utility functions, according to which agents clearly exhibit a preference for a mixed environment. We show that there is no divergence between individual moves and social welfare with increasing linear utility functions, although segregation prevails in stationary configurations. We also show that even with the strongest preference for mixity - in the asymmetrically peaked utility function case - the model has segregated stationary configurations. This case is also the one in which the divergence between the stationary segregation level and the optimal segregation level is the highest. These results complement those obtained by Zhang (2004b, 2010) and by Pancs and Vriend (2007) based on simulations. In addition, a few simulation results are shown for illustrative purpose. In summary, our work provides a very general solution to Schelling's model, that encompasses previous work on this model and that paves the way to the analysis of many structures of preferences, for instance those based on empirical findings concerning racial preferences.

2. A general dynamic model of segregation

2.1. The city and the agents

Our artificial city is a two-dimensional NxN square lattice with periodic boundary conditions, *ie* a torus containing N^2 cells. Each cell corresponds to a dwelling unit, all of equal quality. We suppose that a certain characteristic divides the population of this city in two groups of households that we will refer to as red and green agents. Each location may thus be occupied by a red agent, a green agent, or may be vacant. We denote by N_V the number of vacant cells, and by N_R and N_G the number of respectively red and green agents. The parameter N thus controls the size of the city, the parameter $v = N_V/N^2$ its vacancy rate, and the fraction $n_R = N_R/(N_R + N_G)$ its composition.

We define a *state* x of the city as a N^2-vector, each element of this vector labeling a cell of the NxN lattice. Each state x thus represents a specific configuration of the city. We note X the set of all possible

4

configurations, the demographic parameters (N, v, n_R) being fixed.

2.2. Neighborhoods

Since Schelling (1969)'s work, two ways of conceiving the neighborhood of an agent have been developed and used in analytical and simulation models.

Bounded neighborhood models (Fig 1a) describe cities divided into geographical units within which all agents are connected. The neighborhood of an agent is thus composed entirely and exclusively of the locations present in the same geographical unit as his own. In the following, when we refer to a bounded neighborhood model, we will implicitly assume that the city is divided in a set $\mathcal{Q}$ of blocks, each of which contains $H + 1$ locations, where H is a fixed integer that corresponds to the number of locations in an agent's neighborhood (hence, the relation $|\mathcal{Q}|(H + 1) = N^2$ must hold). Note that since some locations remain empty, the size H of the neighborhood of an agent can also be interpreted as the maximum number of neighbors an agent can have. For a given configuration $x \in X$ of the city, we denote by $R_q(x)$ and $G_q(x)$ the number of red and green agents that live inside the block $q \in \mathcal{Q}$. Taking into account that some locations of each block may remain empty, the $\{R_q\}$ and the $\{G_q\}$ must thus verify :

$$\sum_q R_q = N_R \tag{2.1}$$

$$\sum_q G_q = N_G \tag{2.2}$$

$$(R_q, G_q) \in E_{H+1} \equiv \{(R, G), 0 \le R + G \le H + 1\} \tag{2.3}$$

Fig 1a displays an example of a city divided into square blocks, which corresponds to the kind of bounded neighborhoods we use in the simulations.

<div align="center">Insert Figure 1 about here.</div>

Continuous neighborhood models (Fig 1b) describe cities where the neighborhoods do not correspond to a zoning at the city level, but are centered on the local perception of each agent. In a continuous neighborhood description, one assumes that the neighborhood of an agent is composed of the H nearest locations surrounding him. The $H = 4$ "Von Neumann neighborhood" and the $H = 8$ "Moore neighborhood" that are displayed among other examples on Fig 1.b are the most commonly used in agent-based computational models.

In the following, we place ourselves in a bounded neighborhood description, unless otherwise mentioned. Bounded neighborhoods can be thought of as reproducing the effects of the administrative divisions of real cities such as census areas or school districts. Furthermore, there is no argument in favor of bounded or continuous neighborhoods from the viewpoint of assumption realism. We present in section 4.3 some

<div align="center">5</div>

simulation results illustrating the impact of the neighborhood description on the forms of segregation at the city scale. Those results show that the global characteristics of our model remain qualitatively independent of any specific definition of neighborhood, provided its size H is relatively small compared to the size N^2 of the city, in order to maintain the "local" property of neighborhood.

2.3. Agent's utility function

Each agent has a utility level which depends only on his neighborhood composition. Let us consider an agent whose neighborhood is composed of R red agents, G green agents and V vacant cells. Since $R+G+V = H$, one needs two independent parameters to describe the composition of the agent's neighborhood. In all generality, we can thus write the utility of an agent for example as a function of R and G. Like most models of the literature, we assume for simplification that agents of a same group share the same utility function.[2]

Without any loss of generality, we write the utility of an agent as:

$$u = u_R(R, G) \quad \text{for a red agent with } R \text{ red and } G \text{ green neighbors,}$$

$$u = u_G(R, G) \quad \text{for a green agent with } R \text{ red and } G \text{ green neighbors.}$$

The utility of an agent is thus a function of $E_H \rightarrow \mathbb{R}$. More specifically in the case of a bounded neighborhood description, we will have :

$$u = u_R(R_q - 1, G_q) \quad \text{for a red agent living in block } q,$$

$$u = u_G(R_q, G_q - 1) \quad \text{for a green agent living in block } q.$$

In order to facilitate the comparison of different utility functions, utility in the examples presented below is such that a zero utility level denotes a complete dissatisfaction of the agent and a utility of one denotes complete satisfaction.

We also introduce a notation in order to characterize the level of utility on the global (city) scale:

$$U(x) = \sum_k u_k \tag{2.4}$$

where u_k is the utility of agent k and $U(x)$ denote the collective utility of a configuration x.

2.4. A behavioral rule: the logit dynamical rule

The core of dynamic segregation models is that agents are given opportunities to move to increase their individual utility. Once the static description of the model is specified, one must add a dynamic rule that governs these moves. In the following, the city configuration evolves according to an iterative process. At

---

[2]See O'Sullivan (2009) for a treatment with heterogeneous agents.

6

the first iteration, an initial configuration is randomly chosen. At each iteration, one agent and one vacant cell are picked at random.[3] The picked agent then chooses to move in that vacant cell with a probability $Pr\{move\}$ that depends on the utility gain Δu he would achieve if he was to move, as follows:

$$Pr\{move\} = \frac{1}{1 + e^{-\Delta u/T}} \tag{2.5}$$

where $T > 0$ is a fixed parameter.

Eq. 2.5 represents a logit choice function as developed e.g. in Anderson *et al.* (1992). From the microeconomic viewpoint, it can be given a justification in terms of a random-utility model where the random part of the utility function is a way to take into account criteria other than the observable choice characteristics or a way to model the agents' bounded rationality: it may happen that an agent takes a utility-decreasing move, either because he is making a mistake or because of a lack of information

The scalar T is used to determine the relative importance of the random part with respect to the deterministic part of the random utility function (see Appendix A). Clearly, the probability for an agent to take a utility-decreasing move drops down as $T \rightarrow 0$ and the described rule thus converges to the non-strict best response rule. For any finite $T > 0$, the agents choose non-best replies with a non-zero probability, but actions that yield smaller payoffs are chosen with smaller probability. This kind of perturbed best-response dynamics has been developed in Young (1998) in the context of evolutionary games. Here, neighborhood's composition is supposed to be the main determinant of agents' actions : we restrict our analysis to the case of low values of T.

Beside being more realistic from a behavioral point of view, the logit rule also provides a strong analytical framework to Schelling model. Obviously, it implies that the probability that the state at the tth iteration x^t is equal to a given state x only depends on the state at the previous iteration x^{t-1}:

$$\Pr(x^t = x|x^{t-1}, \ldots, x^1, x^0) = \Pr(x^t = x|x^{t-1}) \tag{2.6}$$

The dynamic rule thus yields a finite Markov process.

It is then easy to figure out that the Markov chain describing our system is irreducible (since $T > 0$ each imaginable move has a non-zero probability to happen and it is thus possible to get to any state from any state), aperiodic (given any state x and any integer k, there is a non-zero probability that we return to state x in a multiple of k iterations) and recurrent (given that we start in state x, the probability that we will never return to x is 0). These three properties ensure that the probability to observe any state x after t iterations

---

[3]Instead of assuming, as Schelling did, that the agents move to the nearest satisfactory position (the idea being that the cost of moving increases with distance), we suppose here that the distance between the current and envisaged locations of an agent does not intervene in his deciding whether to move or not. This could be justified by assuming that the cost of moving for the largest possible distance is smaller than any possible strictly positive difference in utility between two locations.

starting from a state y converges toward a fixed limit independent of the starting state y as $t \to \infty$.

In other words, for each set of parameters and dynamic rule, there exists a stationary distribution

$$\Pi : x \in X \to \Pi(x) \in [0,\, 1] \,,\, \sum_{x \in X} \Pi(x) = 1 \tag{2.7}$$

which gives the probability with which each state x will be observed in the long run.

Clearly, for $T \to \infty$, the randomness introduced in the dynamical rule prevails and the stationary distribution is just a constant. Similarly, for any finite $T > 0$, our dynamical system (the city) evolves toward an attractor composed of a subset A of X. It follows that any measure $\mathcal{M}$ - such as the global utility U - performed on the state space X will in the long run fluctuate around a mean value $\mathcal{M}_\infty = \sum_{x \in A} \Pi(x) \mathcal{M}(x)$. These mean values may depend on the intensity of the noise T, but the amplitude of the fluctuations decreases as $T \to 0$.

In the following, we refer to two states x and y as *immediately communicating states* (ICS) if we can switch from state x to state y by moving one single agent. We also note $\Delta_{xy} u$ the variation of utility of this agent induced by this particular move and P^T_{xy} the probability to be in state y at a given iteration if the system was in state x at the previous iteration. According to the dynamic rule presented in the previous section, one has:

$$P^T_{xy} \;=\; \gamma(1 + e^{-\Delta_{xy} u / T})^{-1} \quad \text{if } x \text{ and } y \text{ are ICS} \tag{2.8}$$

$$P^T_{xy} \qquad\quad = \; 0 \qquad\qquad \text{if } x \text{ and } y \text{ are not ICS} \tag{2.9}$$

where the parameter $\gamma = 1/(N_V(N_R + N_G)) = 1/(v(1-v)N^4)$ takes into account the probability to pick the right agent and the right vacant cell that allow to pass from x to y. P^T thus corresponds to the probability transition matrix for a fixed T and the stationary distribution Π is by definition the unique normalized function defined on X that verifies for all $x \in X$:

$$\sum_y P^T_{yx} \Pi(y) = \Pi(x) \tag{2.10}$$

3. Model solving with a potential function

3.1. Definitions and properties

Following Zhang (2004a,b), we place our model in the context of evolutionary game theory and use the concept of potential function to solve it. In game theory, the concept of potential function was proposed by Monderer and Shapley (1996). A game is said to be a potential game if the incentive of all players to choose their strategy can be expressed in one global function, which is called the potential function. In our context, the definition of a potential function takes the rather simple following form:

8

Definition 1. Let $\mathcal{F} : x \in X \to \mathcal{F}(x) \in \mathbb{R}$ be an aggregate function describing each of the potential configurations. By definition, $\mathcal{F}$ will be a (cardinal) potential function of our model if and only if each gain in utility Δu of a moving agent is equal to the variation $\Delta \mathcal{F}$ that is induced on the global level by the move of this agent.[4] A cardinal potential function will thus verify: $\mathcal{F}(y) - \mathcal{F}(x) = \Delta_{xy} u$ with $\Delta_{xy} u$ previously defined (section 2.4).

The main property of a potential function is to link the variation of a purely individual function (the utility of the moving agent) to the variation of a global function defined on the space X of all possible configurations. The ensuing lemma points out even more the value of the potential function as an analytical tool.

Lemma 1

If $\mathcal{F}$ is a potential function of the system, then the stationary distribution Π is such that for any configuration x:

$$\Pi(x) = \frac{e^{\mathcal{F}(x)/T}}{\sum_{z \in X} e^{\mathcal{F}(z)/T}} \tag{3.1}$$

It follows that for $T \to 0$, the stationary configurations are those that maximize $\mathcal{F}$.

Proof. The following proof follows the classical argument presented in Young (1998).

Let π be the function defined as $\pi : X \to [0, 1]$; $x \to \pi(x) = e^{\mathcal{F}(x)/T} / \sum_z e^{\mathcal{F}(z)/T}$. The first step of the proof consists in checking that π satisfies the *detailed balance condition*:

$$\pi(x) P_{xy}^T = \pi(y) P_{yx}^T \tag{3.2}$$

If x and y are two different and not communicating states, equality 3.2 is trivially satisfied since in this case $P_{xy}^T = P_{yx}^T = 0$. If $x = y$, the detailed balance condition is also trivially verified. In the case where $x \neq y$ and x and y are two communicating states, one has:

$$
\begin{aligned}
\pi(x) P_{xy}^T &= \pi(x) \gamma \frac{1}{1 + e^{-\Delta_{xy} u/T}} = \pi(x) \gamma \frac{1}{1 + e^{-(\mathcal{F}(y) - \mathcal{F}(x))/T}} = \pi(x) \gamma \frac{e^{\mathcal{F}(y)/T}}{e^{\mathcal{F}(x)/T} + e^{\mathcal{F}(y)/T}} \\
&= \pi(y) \gamma \frac{e^{\mathcal{F}(x)/T}}{e^{\mathcal{F}(x)/T} + e^{\mathcal{F}(y)/T}} = \pi(y) \gamma \frac{1}{1 + e^{-(\mathcal{F}(x) - \mathcal{F}(y))/T}} = \pi(y) \gamma \frac{1}{1 + e^{-\Delta_{yx} u/T}} \\
&= \pi(y) P_{yx}^T
\end{aligned}
$$

recalling that $\gamma = 1/(N_V (N_R + N_G)) = 1/(v(1 - v) N^4)$.

---

[4]Games can be either ordinal or cardinal potential games. In cardinal games, the difference in individual payoffs for each player from individually choosing one's strategy *ceteris paribus* has to have the same value as the corresponding difference in value for the potential function. In ordinal games, only the signs of the differences have to be the same.

Hence the detailed balance condition is always verified and

$$\sum_{x \in X} \pi(x) P_{xy}^T = \sum_{x \in X} \pi(y) P_{yx}^T = \pi(y) \sum_{x \in X} P_{yx}^T = \pi(y) \cdot 1 = \pi(y), \qquad (3.3)$$

which defines π as a stationary distribution of the process. Because the Markov chain is finite and irreducible, it has a unique stationary distribution. Hence, for each state x, $\Pi(x) = \pi(x) = e^{\mathcal{F}(x)/T} / \sum_z e^{\mathcal{F}(z)/T}$.

Define then X_F as the subset of X of the states that maximize the potential function $\mathcal{F}$:

$$X_F = \{y, \forall x \in X \ \mathcal{F}(y) \geq \mathcal{F}(x)\} \qquad (3.4)$$

The second part of the lemma can now be proved as follows: for two states x and y of X_F, we will have $\mathcal{F}(x) = \mathcal{F}(y)$ and therefore $\Pi(x)/\Pi(y) = e^{[\mathcal{F}(x)-\mathcal{F}(y)]/T} = 1$, which means that two states that strictly maximize $\mathcal{F}$ are observed in the long run with the same probability; for two states $x \in X \setminus X_F$ and $y \in X_F$, we will have $\mathcal{F}(x) - \mathcal{F}(y) \leq 0$ and therefore $\Pi(x)/\Pi(y) = e^{[\mathcal{F}(x)-\mathcal{F}(y)]/T} \to 0$ as $T \to 0$. This means that for $T \to 0$, the probability to observe a state that does not maximize the potential function $\mathcal{F}$ becomes in the long run infinitesimally small. $\square$

The potential function is hence a very powerful analytical tool. By establishing a relation between individual changes in utility and a global characteristic of the city configuration, and because stationary configurations can be defined as those maximizing the potential function for low noise levels, the existence of a potential function allows to qualify analytically stationary configurations. The fact that the knowledge of Δu is sufficient to say something on the global level is highly non-trivial since (for example) there is no way to determine the externalities produced by the move of an agent - *ie* the variation of the utility of his former and new neighbors - only by the knowledge of Δu. A low level of T is required for the maximum of the potential function to be achieved at stationary configurations.[5] Still, T has to remain strictly positive to avoid blocked states, that could arise if the noise was high compared to the utility level produced by the neighborhood composition.

3.2. Main result: existence of a potential function

It is possible, using the potential function, to examine the outcome of the model for different utility functions, representing different degrees of preference for mixity. To do so, two questions are to be answered first: given any pair of utility functions (u_R, u_G), does a potential function exist and can we compute it? Reciprocally, given a potential function, can we find a pair of utility functions (u_R, u_G) that can be translated

---

[5]In the case of finite values of the noise level ($T > 0$), it can be demonstrated using standard tools of statistical physics that the states which are the more probable to appear are those which maximize $F(x) + TS(x)$ where $S(x)$ is an entropy-like global function taking into account the number of ways of locating R_q red agents and G_q green agents in each block q of the city (Grauwin *et al.*, 2009a).

into this specific potential function?

We show in the following that in the context of bounded neighborhoods, one can achieve an analytical resolution of the model under a rather mild condition. Let us begin with some definitions.

Definition 2. Let $\mathbb{U}$ be the set of pairs of utility functions (u_R, u_G) that verify, for all $(R, G) \in E_H$, the following condition:

$$u_R(R, G) - u_R(R, G + 1) = u_G(R, G) - u_G(R + 1, G) \tag{3.5}$$

Condition 3.5 only imposes that if a block contains $R + 1$ red agents and $G + 1$ green agents, the utility gain a red agent would achieve if a green agent left must be the same than the utility gain a green agent would achieve if a red agent left. The results in the following apply to pairs of utility functions verifying this condition. As we show below, this condition is not strongly restrictive from a theoretical viewpoint, which means that our approach can be applied to virtually all the usual utility functions.

Definition 3. Let $\mathbb{F}$ be the set of aggregate functions of the form $\mathcal{F}(x) = \sum_{q \in \mathcal{Q}} F(R_q, G_q)$, where F is an intermediate function defined on the set E_{H+1} of all possible numbers of red and green agents that can be present in a block.

The main result of this paper consists in the following proposition:

Proposition 1

Each aggregate function $\mathcal{F} \in \mathbb{F} : x \to \mathcal{F}(x) = \sum_{q \in \mathcal{Q}} F(R_q, G_q)$ is a potential function to which corresponds at least[a] one pair (u_R, u_G) of utility functions of $\mathbb{U}$ that can be expressed as:

$$\begin{cases} u_R(R, G) = F(R + 1, G) - F(R, G) \\ u_G(R, G) = F(R, G + 1) - F(R, G) \end{cases} \tag{3.6}$$

Reciprocally, for each pair of utility functions (u_R, u_G) of $\mathbb{U}$, there exists one corresponding potential function $\mathcal{F}_{[u_R, u_G]} \in \mathbb{F}$. This function can be expressed through the functional $F_{u_R, u_G} : E_{H+1} \to \mathbb{R}$ - such that $\mathcal{F}_{u_R, u_G}(x) = \sum_{q \in \mathcal{Q}} F_{[u_R, u_G]}(R_q, G_q)$ - which is defined for all $(R, G) \in E_{H+1}$ by:[b]

$$F_{[u_R, u_G]}(R, G) = \sum_{r=1}^{R} u_R(r - 1, 0) + \sum_{g=1}^{G} u_G(R, g - 1) \tag{3.7}$$

$$= \sum_{r=1}^{R} u_R(r - 1, G) + \sum_{g=1}^{G} u_G(0, g - 1) \tag{3.8}$$

---

[a]Since the definition of the potential function makes only intervene its variation, $\mathcal{F}$ can be defined up to an additive constant. Similarly, a utility function can also be defined up to a constant. All the formula in this insert are written with the convention $u(0, 0) = F(0, 0) = F(0, 1) = F(1, 0) = 0$. For more details, see the proof in appendix A.2.
[b]Our notation assumes that a sum is null whenever its upper bound of summation is inferior to its lower bound of summation.

Proof. See Appendix B. □

Proposition 1 states that it is always possible to define a function F corresponding to the variation of utility of a moving agent, but only a pair of utility functions verifying condition 3.5 allows this function to be path-independent and therefore uniquely defined for any given configuration. Reciprocally, for any pair of utility functions belonging to $\mathbb{U}$, the game has a potential function that is maximized at stationary configurations, and this function is the sum of neighborhood-level intermediate components. As Eq 3.7 shows, the intermediate component of the potential function corresponds to the sum of utilities of the agents arriving in succession in the block. This sum is calculated starting from an empty block, agents being introduced one by one, first the red ones and then the green ones. As Eq 3.8 shows, the same sum is obtained if green agents are introduced first and red agents after. Before giving a more general interpretation in section 3.4, it is useful to give the following corollary, aimed at showing that utility functions belonging to $\mathbb{U}$ can be given a convenient formulation in which interactions between the two groups are expressed through the same function in the two utility functions and at giving a more general formulation for the potential function.

Corollary 1

Any pairs of utility functions (u_R, u_G) belonging to $\mathbb{U}$ can be written:

$$u_R(R, G) = \xi_R(R) + \sum_{g=0}^{G-1} \xi(R, g) \tag{3.9}$$

$$u_G(R, G) = \xi_G(G) + \sum_{r=0}^{R-1} \xi(r, G) \tag{3.10}$$

where ξ_R and ξ_G are arbitrary functions of $\{0, 1, .., H\} \to \mathbb{R}$ and ξ an arbitrary function of $E_H \to \mathbb{R}$.

Hence, for each pair of utility functions (u_R, u_G) verifying 3.9 and 3.10, thanks to Eq. 3.7, one can rewrite the general form of the potential function $\mathcal{F}_{[u_R, u_G]}$ as:

$$\mathcal{F}(x) = const + \sum_{q} \left(\sum_{r=0}^{R_q-1} \xi_R(r) + \sum_{g=0}^{G_q-1} \xi_G(g) + \sum_{r=0}^{R_q-1} \sum_{g=0}^{G_q-1} \xi(r, g) \right) \tag{3.11}$$

Proof. For any pairs of utility functions (u_R, u_G), one can define ξ_R and ξ_G, two functions of $\{0, 1, .., H\} \to \mathbb{R}$ and ξ_{RG} and ξ_{GR}, two functions of $E_H \to \mathbb{R}$ by

$$\begin{cases} \xi_R(r) &= u_R(r, 0) \\ \xi_G(g) &= u_G(0, g) \end{cases} \qquad \begin{cases} \xi_{RG}(r, g) &= u_R(r, g+1) - u_R(r, g) \\ \xi_{GR}(r, g) &= u_G(r+1, g) - u_G(r, g) \end{cases}$$

for all $0 \leq r \leq H$ and $0 \leq g \leq H$. By definition, one can then write the utility functions as

$$
\begin{aligned}
u_R(R, G) &= \xi_R(R) + \sum_{g=0}^{G-1} \xi_{RG}(R, g) \\
u_G(R, G) &= \xi_G(G) + \sum_{r=0}^{R-1} \xi_{GR}(r, G)
\end{aligned}
$$

for all $(R, G) \in E_H$.

The condition given by Eq. 3.5 is obviously equivalent to $\xi_{RG} = \xi_{RG} \equiv \xi$, which proves Corollary 1.

□

3.3. Interpretation

We propose here an interpretation of condition 3.5 and of the form of the potential function. Proposition 1 ensures that there exists a potential function for any pair of utility functions verifying condition 3.5. In its original form, this condition says that there is a symmetry in the externalities generated by green agents on red agents and by red agents on green agents: starting from a given neighborhood composition, the variation in utility produced by the departure of an agent of the other type must be the same for the two categories. This can be seen as rather limiting, as some real world situations do not conform to this condition. For instance, well-known surveys on the appreciation by white and black individuals of their preferred residential environment show that blacks are in favor of integrated neighborhoods, whereas whites favor all-white neighborhoods (Farley *et al.* (1978); see Farley *et al.* (1997) for recent figures).

However, condition 3.5 covers more general types of preferences when the rate of vacant cells is low. Namely, in the limit of a very low vacancy rate, there is no vacant cells in most of the blocks, *ie* in these blocks the relation $R_q + G_q = H + 1$ holds. Hence, one only needs one parameter among (R_q, G_q, V_q) to define a utility function and considering for instance that an agent's utility only depends on his number of similar neighbors is sufficient to describe all possible cases. This can simply be done by taking $\xi \equiv 0$ in Eq. 3.9 and Eq. 3.10, while keeping the functions ξ_R and ξ_G independent and free. In other words, it is clear that in the limit of no vacant cells, each agent arriving in a neighborhood receives a utility that is fully determined by the number of like-neighbors. Therefore, the order in which the agents settle in the neighborhood does not matter and the condition for having a potential function holds. The set $\mathbb{U}$ hence describes all possible pairs of utility functions in the limit $v \to 0$. It follows also that condition 3.5 holds for all pairs of utility functions in situations where vacant cells are considered in the same way as unlike-color neighbors.

Note finally that in Zhang (2004a), the symmetric effect of unlike-color neighbors on each type of agent emerges as the result of the determination of housing prices by densities, in a model where preferences over neighborhoods are determined by the number of like-agents only.

In its original form, the potential function $\mathcal{F}$ can be interpreted as the sum of the incentives the agents had (when they settled) to move into the neighborhood where they are located. Indeed, if $x(t)$ denotes the state of the city at iteration t, then the potential can be rewritten as

$$\mathcal{F}(x(t)) - \mathcal{F}(x(0)) = \sum_{t'=1}^{t} \Delta_{x(t'-1)x(t')} u \tag{3.12}$$

where $\Delta_{x(t'-1)x(t')} u = 0$ by definition if no move happens at iteration t' and where we can take $\mathcal{F}(x(0)) = 0$ since the potential is defined up to a constant.

Conversely, the potential function $\mathcal{F}$ can also be viewed as the minimum utility level each agent would require to accept quitting his neighborhood. As such, it represents, in the case $T \to 0$, the stability of the configuration x: the higher the potential function, the smaller the incentives for agents to move.

To interpret further the potential function, it is worth noting that condition 3.5 can also be written as follows:

$$u_R(R, G) + u_G(R + 1, G) = u_G(R, G) + u_R(R, G + 1) \tag{3.13}$$

which means that starting from any initial composition of a block, the sum of utilities of a red agent and a green agent entering successively in this block is the same whatever the order in which they enter. This expression stresses that, under condition 3.5, the value of function F in a given neighborhood q does not depend on the particular path of events that lead to the composition of this neighborhood. This is also particularly clear in the form of condition 3.5 given in corollary 1. It hence follows that the potential function $\mathcal{F}$, which is the sum of the F intermediate functions, is independent of the particular order in which the agents arrived in the neighborhoods.

Hence it is also possible to define $\mathcal{F}$ as the average over all the possible ways of ordering the agents, which is what states the corollary given in the following subsection.

3.4. A coalitional game formulation

The following corollary gives a general form of the potential function and shows that it can be interpreted in terms of the Shapley value of a non cooperative coalitional game (see Appendix C for a short presentation of coalitional games and Shapley value; Shapley, 1953).

Corollary 2

The potential function $\mathcal{F}$ can be written as the sum of the following intermediate functions F:

$$F_{[u_R, u_G]}(R, G) \;=\; \sum_{\substack{0 \leq r \leq R,\, 0 \leq g \leq G \\ (r,g) \neq (0,0)}} \binom{R}{r}\binom{G}{g} \frac{(r+g-1)!\,(R+G-r-g)!}{(R+G)!} \; \nu(r,g) \qquad (3.14)$$

where $\quad \nu(r,g) \;=\; r\, u_R(r-1, g) \;+\; g\, u_G(r, g-1) \qquad\qquad\qquad (3.15)$

is the collective utility in a block having r red and g green agents.

Let us consider a coalitional game defined by the set of the $R_q + G_q$ agents present in block q along with a coalition worth equal to the sum of their utilities $\nu_q(R_q, G_q) = R_q\, u_R(R_q - 1, G_q) + G_q\, u_G(R_q, G_q - 1)$, that is the collective utility at the neighborhood level. The sum of the potential functions of the $|Q|$ coalitional games defined on the $|Q|$ blocks is equal to the potential function $\mathcal{F}$.

Proof. The form of F given in expression 3.14 is derived in Appendix B. Then, acknowledging that in a neighborhood with $R_q + G_q$ agents, there are $\binom{R}{r}\binom{G}{g}$ possible coalitions having g green and r red agents and applying the formula of the potential function of the Shapley value derived in Hart and Mas-Colell (1989) (see Appendix C), one obtains exactly the formula of the potential function given in 3.14. Hence the potential function $\mathcal{F}$ can be written as the sum of these potential functions of the $|Q|$ coalitional games defined on the $|Q|$ blocks. $\square$

The potential function $\mathcal{F}$ of our non cooperative game can thus be written as the sum of the potential functions of $|Q|$ coalitional games defined on the $|Q|$ blocks.[6] Notice furthermore that the corresponding Shapley value can be straightforwardly identified, thanks to Eq. 3.6 and the relationship between the Shapley value and its potential, as the vector of $\mathbb{R}^{R_q + G_q}$ whose components are the utilities enjoyed by the agents inside block q. This also means that condition 3.5, that has to be verified by the utility functions to have a potential function, corresponds to the Balance Contribution property of the Shapley value taken in the particular case of two agents of different colors (see Appendix C for more details).

Drawing the parallel a bit further allows also to highlight that the same difference exists between the potential function of the coalitional game and the grand coalition worth as between the potential function $\mathcal{F}$ and the collective utility U. Whereas $\mathcal{F}$ represents the sum of the agents' utilities at the time when they have moved into their current location starting with a totally empty city (or are considered to have done so), U represents the sum of the agents' utilities once they are all settled. Hence, while stationary configurations

---

[6]This result illustrates a theorem presented in Ui (2000), which extends the notion of Shapley value to non cooperative games such as ours.

maximize $\mathcal{F}$ they do not necessarily (and the following examples show that they generally don't) maximize the collective utility.

4. Segregation for different levels of preference for mixity

The main property of the potential function obtained in the previous section is that it reflects both the macro and micro scale. On one hand, $\mathcal{F}$ is an aggregate function which only depends on the number R_q and G_q of red and green agents in each block. On the other hand, $\mathcal{F}$ also keeps tracks of the individual level since it corresponds to a sum of the utility differences generated by individual moves. When the stationary states are reached in the case $T \to 0$, $\mathcal{F}$ is maximized, which means that no agent can strictly improve her utility by moving. The potential function can now be used to assess the outcomes of our location model for different utility functions, representing different degrees of preference for mixity. We examine successively (i) linear utility functions, with a continuous preference for segregated environments, (ii) Schelling original utility function in which there is a mild preference for a mixed environment and (iii) an asymmetrically peaked utility function, according to which agents exhibit a strict preference for a mixed environment.

4.1. Linear utility functions

We consider here utility functions that exhibit a monotone effect of the number of same-color neighbors on utility, through linear utility functions. Zhang (2004a) proposes an analytical solution of a dynamic model of segregation with a linear utility function and shows that the halved sum of individual utilities is a potential function of the game thus defined. In this section, we show that proposition 2 allows to find similar results for all linear utility functions verifying condition 3.5 in the context of bounded neighborhoods.

Suppose that u_R and u_G are expressed as:

$$
\begin{aligned}
u_R(R, G) &= aR + bG \\
u_G(R, G) &= bR + dG
\end{aligned}
\tag{4.1}
$$

where a, b, d are constant parameters.[7]

One can easily verify that this particular pair of utility functions verifies condition 3.5 and compute the corresponding potential function:

$$
\mathcal{F}(x) = \frac{1}{2} \sum_q (aR_q(R_q - 1) + dG_q(G_q - 1) + 2bR_qG_q)
\tag{4.2}
$$

---

[7]Zhang (2004a)'s utility function corresponds to $b = d = -1$ and $a \geq -1$, the utility also including a fixed income term which makes it positive. In Zhang (2004a)'s framework, the impact of unlike neighbors is not due to preferences, but to the impact of density on housing prices.

16

One can rewrite this potential function as:

$$\mathcal{F}(x) = \frac{b-a}{2}\big[\rho_{RR}(x) + \rho_{RG}(x)\big] + \frac{b-d}{2}\big[\rho_{GV}(x) + \rho_{RG}(x)\big] - \frac{b}{2}\big[\rho_{RV}(x) + \rho_{GV}(x)\big] \quad (4.3)$$

with:

$$\rho_{RG} = \sum_q R_q G_q \qquad\qquad\text{the number of red-green pairs of neighbors,}$$

$$\rho_{RV} = \sum_q R_q(H + 1 - R_q - G_q) \quad\text{the number of red-vacant pairs of neighbors and}$$

$$\rho_{GV} = \sum_q G_q(H + 1 - R_q - G_q) \quad\text{the number of green-vacant pairs of neighbors.}$$

This last form allows a convenient interpretation of the potential function. Putting aside at this point the last two terms, $\mathcal{F}(x)$ is proportional to ρ_{RG}, that gives a measure of the relative contact between the two groups. Hence, the sign of the prefactor $b - (a + d)/2$ indicates whether mixed states (when positive) or segregated states (when negative) are obtained at the global level. Notice that the two groups do not need to both have strong preferences for like neighbors for segregation to emerge. It is the average preference over the two groups that determines the level of segregation.

The terms proportional to ρ_{RV} and ρ_{GV} show that agents avoid the proximity of vacant cells when $a > 0$ and $d > 0$. All these insights gained from the study of the potential function can be checked by means of simulations (Fig 2).

Insert Figure 2 about here.

Turning now to the link between segregation of the stationary configurations and collective utility, it is useful, using proposition 1, to write the potential function as follows:[8]

$$
\begin{aligned}
\mathcal{F}(x) &= \frac{1}{2}\sum_q (aR_q(R_q - 1) + dG_q(G_q - 1) + 2bR_q G_q) \\
&= \frac{1}{2}\sum_q (R_q u_R(R_q - 1, G_q) + G_q u_G(R_q, G_q - 1)) \\
&= \frac{1}{2}U(x)
\end{aligned}
$$

With this choice of utility functions, the potential function is thus proportional to collective utility and therefore lemma 1 ensures that, for low values of T, the stationary configuration maximizes collective utility. Reciprocally, one can verify (see proof in Appendix E) that if we want the potential function to be proportional to the collective utility, so that states that maximize the potential function also maximize collective utility, then the constant of proportionality is necessarily 0.5 and the pair of utility functions must take the form

---

[8]Note that the result is similar to the one obtained in the continuous neighborhood case. See Grauwin *et al.* (2009b)

displayed in Eq. 4.1 (up to a constant).[9]

To sum up, the linear utility functions as defined in 4.1 lead to segregated or mixed states depending on the values of the parameters. However, in all cases, there is no divergence between stationary configurations and the optimum: these utility functions are such that utility-improving moves also improve collective utility.

4.2. Schelling utility function

Suppose that the agents compute their utility with Schelling utility function (which is equal to 1 if their fraction of similar neighbors is superior or equal to 0.5, and equal to 0 otherwise). This utility function can be expressed in terms of the number of red and green neighbors as follows:

$$
\begin{aligned}
u_R(R,G) &= \Theta(R-G) = \frac{1}{2}(1 + |R+1-G| - |R-G|) \\
u_G(R,G) &= \Theta(G-R) = \frac{1}{2}(1 + |R-1-G| - |R-G|)
\end{aligned}
\tag{4.4}
$$

where Θ is the Heaviside function defined by: $\Theta(x) = 0$ if $x < 0$ and $\Theta(x) = 1$ if $x \geq 0$.

It is easy to figure out that this particular pair of utility functions respects condition 3.5, and is therefore in the set $\mathbb{U}$. Indeed, the form of the symmetric externality produced by a new unlike-color neighbor is the following:

$$
\begin{aligned}
u_R(R, G+1) - u_R(R, G) &= |R-G| - \frac{1}{2}(|R-G-1| + |R-G+1|) \tag{4.5} \\
&= u_G(R+1, G) - u_G(R, G) \tag{4.6}
\end{aligned}
$$

It is possible to compute the potential function rather directly thanks to its interpretation, as the sum of the utility of the agents being introduced one by one in the city, this sum being independent of the precise order of introduction of the agents. To do so for a given configuration $x \equiv \{R_q, G_q\}$, let us consider that we introduce in each block first the agents in majority (*ie* the red ones if $R_q > G_q$, the green ones if $G_q > R_q$, either the red or the green ones if $R_q = G_q$) and second the agents in minority. Each of the first agents has a utility of 1 as he settles in the city while each of the other minority agents has a zero utility when he settles.[10] Hence it is straightforward to write the potential as:[11]

---

[9]Grauwin *et al.* (2009b) show that in the context of continuous neighborhoods, linear utility functions are the only cases where a potential function exists.

[10]This example shows that to compute the potential function corresponding to a given pair (u_R, u_G) of utility functions, it may be worth to think ahead of a practical order of introduction of the agents. The computation of $\mathcal{F}$ is indeed easier and bears more meanings with an appropriate order.

[11]Notice that in this particular example, we do not use the convention $u(0,0) = 0$. See appendix Appendix D for more details.

$$\begin{aligned}
\mathcal{F}(x) &= const + \sum_{q \in \mathcal{Q}} \max(R_q, G_q) \\
&= const + \sum_{q \in \mathcal{Q}} \frac{1}{2}\Big(R_q + G_q + |R_q - G_q|\Big) \\
&= const' + \frac{1}{2} \sum_{q \in \mathcal{Q}} |R_q - G_q|
\end{aligned}$$

One can verify that the same expression can be found using relation 3.7 (see Appendix D), the computation being in this case more formal than what we present here.

The reader can recognize an expression well-known to scientists working on segregation. This potential function is indeed a linear form of the Duncan and Duncan dissimilarity index, which in the case where the total number of red and green agents in the city are equal ($N_R = N_G = N$), is written as $D(x) = \frac{1}{2}\sum_q |R_q/N_R - G_q/N_G| = \frac{1}{2N}\sum_q |R_q - G_q|$ (Duncan and Duncan, 1955). To the best of our knowledge, an analytical connection between the two "historical" works of Schelling and Duncan and Duncan on segregation has never been found before.

It is worth here investigating the link between the potential function and collective utility. The collective utility in a neighborhood q is:

$$\begin{cases}
U_q &= \frac{1}{2}\left(R_q + G_q + |R_q - G_q|\right) & \text{if } R_q \neq G_q \\
U_q &= R_q + G_q & \text{if } R_q = G_q
\end{cases}$$

The potential function can therefore be written:

$$\mathcal{F}(x) = const + U(x) - \sum_{q|R_q = G_q} \frac{1}{2}\Big(R_q + G_q\Big) \tag{4.7}$$

$$\tag{4.8}$$

This expression shows that the configuration that maximizes $\mathcal{F}$ does correspond to the maximum collective utility. The divergence is due to the existence of perfectly mixed neighborhoods. To be more specific, let us compare two configurations differing by the existence, in configuration x_1, of two perfectly mixed neighborhoods with $K < (H+1)/2$ agents of each color, that are changed to segregated ones in configuration x_2 due to the exchange of two agents of different color. The difference in the potential function between the two configurations is only due to the change affecting these two neighborhoods. It is written:

$$\Delta\mathcal{F} = \mathcal{F}(x_2) - \mathcal{F}(x_1) = 2 \tag{4.9}$$

because each of the two neighborhoods gained one agent of one color and lost an agent of the other color.

The difference in collective utility consists of the loss of utility of the agents who are now in minority in their neighborhood, that is:

$$\Delta U = U(x_2) - U(x_1) = -2(K-1) = 2 - 2K \tag{4.10}$$

Comparing the difference in the potential function and in collective utility between these two configurations, one observes that decreasing the number of perfectly mixed neighborhoods decreases collective utility (due to the loss of those who are in the minority in the new configuration) while increasing the value of the potential function. This is because the two moving agents have still a utility of 1 in their new neighborhood, while they clearly exert negative externalities on the agents of the group which is now in minority in this neighborhood. Stationary configurations will tend therefore to exhibit few perfectly mixed neighborhoods, at the expense of collective utility.

Our analysis based on Schelling original utility function shows thus two points. First, it provides an analytical demonstration of Schelling result, that this pair of utility functions leads to segregated stationary configurations. Second, it sheds light on the source of the discrepancy between the collective utility of stationary configurations and the maximum utility that could be attained with a perfectly mixed environment.

4.3. Asymmetrically peaked utility functions

In this section, we apply our analytical framework to asymmetrically peaked utility functions displayed on Fig. 3, that have been studied in Pancs and Vriend (2007).[12] We will see that our potential function provides a criteria for global segregation or for integration. These utility functions are particularly appealing for demonstrating Schelling's intuition, that the aggregate outcome of the game can run against individual preferences. Indeed, these functions consider a case where agents strictly prefer perfectly mixed neighborhoods against any level of segregation.

Insert Figure 3 about here.

In the following, we place ourselves in the limit $v \to 0$ and thus take $\xi \equiv 0$: the utility of each type of agents can be described entirely by the number of like-color agents. For simplicity, we suppose that the number H of possible neighbors of an agent is even. The asymmetrically peaked utility functions can then be written:

$$\begin{cases} u_R(R,G) = \xi_{ap}(R) \\ u_G(R,G) = \xi_{ap}(G) \end{cases}$$

---

[12]Refer to Grauwin *et al.* (2009b) for the study of other utility functions.

with:

$$
\begin{cases}
\xi_{ap}(s) & = 2s/H & \text{if } s \le H/2 \\
\xi_{ap}(s) & = 2 - m - 2(1-m)s/H & \text{if } s > H/2
\end{cases}
$$

with m a fixed parameter (see Fig. 3).

Using Θ the Heaviside function defined by: $\Theta(x) = 0$ if $x < 0$ and $\Theta(x) = 1$ if $x \ge 0$, this utility function can also be written:

$$
\xi_{ap}(s) = 2\frac{s}{H} - (2-m)\frac{2}{H}\left(s - \frac{H}{2}\right)\Theta\left(s - \frac{H}{2}\right)
$$

For a given state x of the city, simple calculations show that the collective utility can be written as $U(x) = \sum_q \left(\tilde{U}(R_q) + \tilde{U}(G_q)\right)$, with

$$
\tilde{U}(S) = 2\frac{S(S-1)}{H} - (2-m)\frac{2S}{H}\left(S - 1 - \frac{H}{2}\right)\Theta\left(S - 1 - \frac{H}{2}\right), \qquad \forall 0 \le S \le H \qquad (4.11)
$$

Likewise, the corresponding potential function is $\mathcal{F}(x) = const + \sum_q F(R_q, G_q) = const + \sum_q \left(\tilde{F}(R_q) + \tilde{F}(G_q)\right)$, where

$$
\begin{aligned}
\tilde{F}(S) &= \sum_{s=0}^{S-1} \xi_{ap}(s) \\
&= \frac{(S-1)S}{H} - \frac{2-m}{H}\left(S - 1 - \frac{H}{2}\right)\left(S - \frac{H}{2}\right)\Theta\left(S - \frac{H}{2} - 1\right) \\
&= \frac{1}{2}\left[\tilde{U}(S) + (2-m)\left(S - 1 - \frac{H}{2}\right)\Theta\left(S - 1 - \frac{H}{2}\right)\right]
\end{aligned} \qquad (4.12)
$$

Proof. See Appendix F. $\square$

This expression implies once again that the potential $\mathcal{F}$ and the collective utility U are linearly related when the individual utility is linear (case $m = 2$). The lower m, the less linear the individual utility and the greater the divergence from the $\mathcal{F} = const + U/2$ relation. Thus, relation 4.12 puts forward the crucial role of the asymmetric parameter m which is the driver of the moves that produce externalities.

To be more specific, let us compare two configurations x_1 and x_2, which differ only in the repartition of $H + 1$ red and $H + 1$ green agents in two neighborhoods. In configuration x_1, the repartition is rather homogeneous, with $H/2 + 1$ red and $H/2$ green agents in the first neighborhood and $H/2$ red and $H/2 + 1$ green agents in the second one. In configuration x_2, the repartition is more segregated, with $H/2 + 1 + K$ red and $H/2 - K$ green agents in the first neighborhood and $H/2 - K$ red and $H/2 + 1 + K$ green agents in the second one, $K \in \{0, 1, ..., H/2\}$ being an integer determining the level of segregation. The difference in the potential function between the two configurations is only due to the change affecting these two neighborhoods.

It is written:

$$
\begin{aligned}
\Delta \mathcal{F} &= 2\tilde{F}(H/2+1+K) + 2\tilde{F}(H/2-K) - 2\tilde{F}(H/2+1) - 2\tilde{F}(H/2) \\
&= \frac{2m}{H}K(K+1)
\end{aligned}
\tag{4.13}
$$

Proof. See Appendix F. □

As could be expected, $\Delta \mathcal{F}$ increases with m, which means that a given segregated configuration is more probable and stable than a mixed one as the asymmetry toward like-agents is stronger. It also increases with K, which means that for a given m a highly segregated block is more probable than a slightly segregated one. More importantly, a perfectly segregated block will be more probable than a perfectly mixed one if and only if relation 4.13 is positive, *ie* if and only if $m > 0$.

The corresponding difference in collective utility consists of the loss of all the agents. It can be written:

$$
\begin{aligned}
\Delta U &= 2\tilde{U}(H/2+1+K) + 2\tilde{U}(H/2-K) - 2\tilde{U}(H/2+1) - 2\tilde{U}(H/2) \\
&= 4K\left(m\left(\frac{1}{2} + \frac{K+1}{H}\right) - 1\right)
\end{aligned}
\tag{4.14}
$$

Proof. See Appendix F. □

It is obvious that increasing segregation also increases the collective utility as soon as $m \geq 1$. It is straightforward to verify based on equation 4.14 that $\Delta U \geq 0 \Leftrightarrow m \geq m^* = H(H+1)^{-1}$.[13]

Our analysis provides a microscopic criteria allowing to predict a global outcome. For $0 < m < m^* \simeq 1$, complete segregated configurations will be obtained at the expense of the collective utility and for $m < 0$, perfectly mixed configuration will be obtained. These results hold of course in the limit of a low noise level $(T \to 0)$.

The same is observed with simulations. The snapshots presented on the left panel of Fig. 4 are typical stationary configurations obtained by simulating an artificial city where the agents' preferences are given by the asymmetrically peaked utility function with bounded neighborhoods. These snapshots allow us to compare the analytical results obtained for $v \to 0$ and $T \to 0$ with a more realistic $v = 5\%$ and $T = 0.1$. We can see that for values of the asymmetry parameter m close to 0, the system converges toward randomly-organized mixed configurations which also maximize the utility of most agents. On the contrary, for higher values of m, completely segregated configurations are obtained. The transition between these two extreme

---

[13]The limit value is not strictly equal to 1 because of the precise definition of our model: the argument of the utility function is the number of neighbors, which does not include the agent himself. In the limit $H \gg 1$, m^* converges toward 1.

outcomes occurs for values of m included between 0.05 and 0.2. The relative smoothness of this transition is due to the non-zero values of the vacancy rate and of the level of noise.

<center>Insert Figure 4 about here.</center>

The outcome for values of m higher than 0.2 illustrates the paradox of Schelling model: large segregative patterns appear although they absolutely do not maximize the utility of most agents, as most of them are stuck inside an homogeneous area with a utility of 0.5. In this case, one of the key element driving segregation is the asymmetry of this utility function, ie, the fact that even if the agents have a strict preference for mixity, they still favor a large-majority status over a small-minority status. In particular, with the asymmetrically peaked utility function, a red (green) agent may move for example from a 49% red (green) neighborhood to a 51% red (green) neighborhood because it slightly increases her utility. Meanwhile, this move is likely to decrease the utility of the previous and new neighbors and therefore decrease the collective utility level. Both of these factors imply that a highly-segregated configuration is necessarily very stable. Indeed, once the city is divided into homogeneous areas, a red agent will have no incentive to go from the red area to the green one, because his utility would drop from 0.5 to 0.[14]

Finally, the snapshots displayed on the right panel of Fig. 4 present stationary configurations obtained using a continuous neighborhood description. All the other parameters of the simulations are otherwise identical to the ones used in the simulations with bounded neighborhood, which allows us to measure the impact of the choice of a neighborhood description. It is noticeable that mixed random configurations are obtained in both bounded and continuous neighborhood descriptions for low values of m, and segregative patterns also appear in both description for high values of m. The two descriptions produce a different "transition range" (roughly 0.1 ± 0.05 in the bounded neighborhood case *versus* 0.25 ± 0.05 in the continuous neighborhood case). However, these few observations seem to indicate that the choice of a neighborhood description is qualitatively not determinant of the model outcome in terms of segregation.

5. Further discussion and extensions

To this point, we have focused our attention on a formal version of a classical Schelling-type segregation model. But our approach provides an analytical framework that allows to investigate many extensions and to consider a broader range of issues than other classical approaches. Some preliminary suggestions, yet to be fully developed, are presented in this section.

---

[14] And even though a red agent goes from time to time into the green area by mistake, he will have a strong incentive to return to the red area because of the asymmetry in the utility function, and he will do so very likely before a second red agent joins him in the green area.

<center>23</center>

5.1. Segregation by ethnic origin, income, and preferences for public amenities

Up to now, we have always implicitly supposed that the sole characteristic that the agents use to evaluate a location is the composition of its neighborhood. Other determinants of residential location choice however exist that are not necessarily correlated to neighbors' characteristics, such as local public goods. The red and green labeling of our two groups thus may correspond to two different ethnic origins or two groups with different preferences for local public goods. Tiebout (1956)'s analysis of the sorting induced by local public goods is perhaps the main competitor of Schelling (1971) in terms of its influence on later work on neighborhood choice.

It is very easy to write versions of our model which take into account the agents' preferences for public goods while keeping the existence and properties of a potential function. Noting for example A the set of all the public facilities and $d_{i,a}$ the distance between an agent i and an amenity $a \in A$, the utility of an agent i could be rewritten in a general fashion as

$$u_i(R, G) \longrightarrow u_i(R, G) + \tilde{u}_i(\{d_{i,a}\}_{a \in A}) \tag{5.1}$$

and one could then easily derive the more general form of the potential function

$$\mathcal{F}(x) \longrightarrow \mathcal{F}(x) + \sum_i \tilde{u}_i(\{d_{i,a}\}_{a \in A}) \tag{5.2}$$

This generalized approach could provide a means to correct one of the bias of our analytical model, namely the lack of heterogeneity in locations. However, the extraction of the properties of the stationary states from this condensate global function would become quite challenging, as the dimension of the state variable of the system increases with the number of added amenities.

5.2. Taxation to sustain collective welfare

The basic concept at the center of a model *à la* Schelling is that of an agent deciding where to move according solely to the benefit Δu she would achieve if she was to move. Her move affecting her past and new neighbors, an implicit consequence is that she could generate externalities that amount to $\Delta U - \Delta u$ while moving.

Suppose now the existence of a benevolent planner who rewards positive externalities and taxes negative externalities. A way to model the action of that hypothetic benevolent ruler is to write the probability that a move happens as:

$$Pr\{move\} = \frac{1}{1 + e^{-\left(\Delta u + \alpha(\Delta U - \Delta u)\right)/T}} \tag{5.3}$$

where $0 \leq \alpha \leq 1$ is a parameter controlling the tax level. The limit case $\alpha = 0$ corresponds to a standard Schelling model and the limit case $\alpha = 1$ corresponds to a case where only the interest of the collectivity as

24

a whole is taken into account.

Following the path of the proofs developed in section 2.2, one can infer the stationary distribution in the bounded neighborhood framework:

$$\Pi(x) = \frac{e^{\left((1-\alpha)\mathcal{F}(x)+\alpha U(x)\right)/T}}{\sum_z e^{\left((1-\alpha)\mathcal{F}(z)+\alpha U(z)\right)/T}} \tag{5.4}$$

The potential function can thus in this context be generalized to $(1-\alpha)\mathcal{F}(x) + \alpha U(x)$. We already noted that the configurations maximizing $\mathcal{F}$ are not in general maximizing U and could even in certain cases (asymmetrically peaked utility function) be very unfavorable to U. In this context, the question of interest is to determine what level of tax α is necessary or sufficient to break undesired stationary configurations obtained in the classical Schelling model. We address such questions analytically in another paper (Grauwin *et al.*, 2009a).

6. Conclusion

In this paper, we used recent tools from evolutionary game theory to develop an analytical resolution of Schelling segregation model for bounded neighborhoods and two homogeneous groups of agents with general utility functions. This represents a major step forward compared to previous work, mostly limited to computer simulations or providing analytical results for specific models. We showed that the stationary configurations reached following the selfish individual moves of the agents maximize a potential function under mild conditions on the agents' utility functions. This potential function can be interpreted as the sum of the agents' utilities as they move into their neighborhood, starting from a totally empty city. In other words, the potential function cumulates the incentives the agents had to move into the neighborhood where they are located. Thanks to this potential function, we are able to solve Schelling model with general utility functions.

This step forward was enabled by a partial reduction in the heterogeneity of agents' neighborhoods through the use of bounded neighborhoods. This allows to keep track of how each individual move affects the global configuration. Instead, when continuous neighborhoods are used, this information is lost because the way a moving agent affects his past and new neighbors depends on factors (the type of their neighbors' neighbors) that are not fully determined by the agent's decision. Therefore, it is generally impossible to know how an individual move affects a function of the global configuration unless the utility functions are linear.

We used the potential function to assess the outcomes of our location model for different utility functions, representing different degrees of preference for mixity. We examined successively linear utility functions, Schelling original utility function and asymmetrically peaked utility functions. The first two utility functions lead to segregated stationary configurations. In the linear utility case (and for meaningful values of the

parameters), the segregated configurations, that maximize the potential function, also maximize the collective utility. With Schelling original utility function, a divergence between collective utility and the potential function appears; the potential function happens to be a linear form of the Duncan and Duncan segregation index. Asymmetrically peaked utility functions lead to segregated configurations even for a slight asymmetry, because this asymmetry provokes moves to slightly segregated neighborhoods that will never be compensated by reverse moves. Note finally that when the vacancy rate approaches 0, any pair of utility functions gives a potential function that allows to characterize the stationary configurations.

Our analytical approach helps understanding the ingredients that contribute to the paradoxical result that has generated interest for Schelling model. Even if the dynamics is governed by agents moving to improve their own utility, the evolution leads to city configurations in which most of the agents are far from being satisfied. Our paper shows rigorously what the two main ingredients of segregation are. First, the most important element driving segregation is the *asymmetry* of the utility function. Symmetric functions do not lead to segregation. Once utility functions favor a majority status over a minority status, segregation is found, even if agents have a strict preference for mixity, as in the asymmetrically peaked function. The second important element is the existence of externalities. As already noted by Zhang (2004b) and Pancs and Vriend (2007), the existence of externalities explains why individual preferences for integrated environments may lead to segregated configurations. Indeed, location choice by an agent is only based on her own utility level, even if it also affects her neighbors' utility levels. This makes mixed neighborhoods unstable and segregated configurations very stable. The unstability of mixed neighborhoods is particularly clear in the block configuration for the asymmetrically peaked function. Starting with the Nash equilibrium where $R_q = G_q = (H+1)/2$ and $T > 0$, the logit rule implies that there is a positive probability that an agent accepts a slight decrease of his utility, and leaves a block with composition $R_q = G_q = (H+1)/2$. The agents of the same colour remaining in his former block now have a lower utility and are even more likely to leave. This creates an avalanche which empties the block of agents of the same color, as each move away further decreases the utility of the remaining agents. Conversely, highly-segregated configurations are very stable. Indeed, once the city is divided into homogeneous areas, a red agent will have no incentive to go from the red area to the green one, his utility dropping from m to 0.

The analytical tool given here will permit to consider the outcomes of other types of utility functions, in particular those that emerge from empirical findings on social preferences. It is now conceivable to analyze the theoretical outcomes of these preferences and possibly to test the effect of introducing public policy instruments aimed at decreasing segregation.

The kind of solution that is developed here has been used in equilibrium statistical mechanics. Equilibrium statistical mechanics has developed powerful tools to link the microscopic and macroscopic levels. These tools

26

are usually limited to physical systems, where dynamics is governed not by selfish criteria but by a global quantity such as the total energy. By using the potential function, which is analogous to state functions in thermodynamics, we have extended the analytical framework of statistical mechanics to Schelling model. By doing so, our work paves the way to analytical treatments of a much wider class of social systems, where dynamics is governed by individual strategies.

Appendix A. Justification of the logit rule

This appendix aims at showing the role of T in the logit rule. It can easily be shown that equation 2.5 derives from a random utility function $\tilde{u}$ composed of two terms: a deterministic term u which depends on the agent's neighborhood composition and the product of a random term ϵ and parameter T:

$$\tilde{u} = u + T\epsilon$$

The random term ϵ in the payoff function can be interpreted as a way to account of criteria other than the observable choice characteristics or as a way to model the agents' bounded rationality: it may happen that an agent takes a utility-decreasing move, either because he is making a mistake or because of a lack of information. The parameter T is a positive constant which determines the relative importance of the random term with respect to the utility level, which is comprised between 0 and 1. If T is close to 0, the random term is not important and can be neglected. If it is close to infinity, the random term is very important and the neighborhood's compositions do not play any role.

Appendix B. Proof of Proposition 1

In this section, we place ourselves in a bounded neighborhood description. proposition Let us first prove the first part of 1, that is that any aggregate function $\mathcal{F} = \sum_{q \in \mathcal{Q}} F(R_q, G_q)$ is a potential function that corresponds to (at least) one pair of utility functions (u_R, u_G) of $\mathbb{U}$.

Suppose that $\mathcal{F} = \sum_{q \in \mathcal{Q}} F(R_q, G_q)$ is a potential function of the game, where the intermediate function F is known. Let us assume that an agent is moving from a block 1, characterized by the numbers $(R_1, G_1) \in E_{H+1}$ of red and green agents who live in it, to a block 2 characterized similarly by the numbers $(R_2, G_2) \in E_H$ of red and green agents living in it (since there must be at least one vacant location in block 2 for an agent to move in it, we necessarily have $R_2 + G_2 < H + 1$). By definition, the utility variation of a moving agent must be equal to the variation of $\mathcal{F}$ it induces. Hence :

- to cover the cases when the moving agent is a red one: for all $(R_1, G_1) \in E_{H+1}$ with $R_1 \geq 1$,

$$u_R(R_2, G_2) \quad - \quad u_R(R_1 - 1, G_1) =$$
$$F(R_2 + 1, G_2) + F(R_1 - 1, G_1) - F(R_2, G_2) - F(R_1, G_1) \tag{B.1}$$

- to cover the cases when the moving agent is a green one: for all $(R_1, G_1) \in E_{H+1}$ with $G_1 \geq 1$,

$$u_G(R_2, G_2) \quad - \quad u_G(R_1, G_1 - 1) =$$
$$F(R_2, G_2 + 1) + F(R_1, G_1 - 1) - F(R_2, G_2) - F(R_1, G_1) \tag{B.2}$$

Taking $R_2 = G_2 = 0$ in equations B.1 and B.2, one finds that the utility functions u_R and u_G verify for all $(R, G) \in E_H$:

$$u_R(R, G) - u_R(0, 0) \quad = \quad F(R + 1, G) - F(R, G) - F(1, 0) + F(0, 0) \tag{B.3}$$

$$u_G(R, G) - u_G(0, 0) \quad = \quad F(R, G + 1) - F(R, G) - F(0, 1) + F(0, 0) \tag{B.4}$$

These relations define (up to a constant $u(0, 0)$) the utility functions the agents necessarily have if $\mathcal{F} = \sum_{q \in \mathcal{Q}} F(R_q, G_q)$ is a potential function of the game. It still remains to prove that this pair of utility functions belongs to the set $\mathbb{U}$. According to relations B.3 and B.4, one has for all $(R, G) \in E_H$:

$$u_R(R, G) - u_R(R, G + 1) \quad = \quad \big(F(R + 1, G) - F(R, G)\big) - \big(F(R + 1, G + 1) - F(R, G + 1)\big)$$
$$= \quad \big(F(R, G + 1) - F(R, G)\big) - \big(F(R + 1, G + 1) - F(R + 1, G)\big)$$
$$= \quad u_G(R, G) - u_G(R + 1, G)$$

Hence relation 3.5 holds, which means by definition that the pairs of utility functions (u_R, u_G) defined by relations B.3 and B.4 belongs to $\mathbb{U}$. Notice that in our demonstration no particular constraint has to be assumed on the form of function F. As a consequence, any aggregate function $\mathcal{F} = \sum_{q \in \mathcal{Q}} F(R_q, G_q) \in \mathbb{F}$ is a potential function of the game as soon as the pair of agents' utility functions is chosen so that relations B.3 and B.4 hold.

Let us now prove the second part of proposition 1, which is that to any pair of utility functions (u_R, u_G) of $\mathbb{U}$ corresponds a potential function of the form $\mathcal{F} = \sum_{q \in \mathcal{Q}} F(R_q, G_q)$.

Let $(u_R, u_G) \in \mathbb{U}$ be a pair of utility functions verifying condition 3.5. Suppose that $F(0, 0)$, $F(0, 1)$ and $F(1, 0)$ are given and let us define recursively the function F on E_{H+1} by the following equations, verified for all $(R, G) \in E_H$:

$$F(R + 1, G) - F(R, G) \quad = \quad F(1, 0) - F(0, 0) + u_R(R, G) - u_R(0, 0) \tag{B.5}$$

$$F(R, G + 1) - F(R, G) \quad = \quad F(0, 1) + F(0, 0) + u_G(R, G) - u_G(0, 0) \tag{B.6}$$

The most important thing to notice is that these two relations are consistent with each other thanks to

condition 3.5 that links the two utility functions u_R and u_G. By summing Eq. B.5 on R and then Eq.B.6 on G, one finds the following expression for function F:

$$F(R,G) - F(0,0) = R\Big(F(1,0) - F(0,0)\Big) + \sum_{r=1}^{R}\Big(u_R(r-1,0) - u_R(0,0)\Big)$$
$$+ G\Big(F(0,1) - F(0,0)\Big) + \sum_{g=1}^{G}\Big(u_G(R,g-1) - u_G(0,0)\Big) \qquad \text{(B.7)}$$

or conversely by summing Eq. B.6 on G then Eq.B.5 on R,

$$F(R,G) - F(0,0) = R\Big(F(1,0) - F(0,0)\Big) + \sum_{r=1}^{R}\Big(u_R(r-1,G) - u_R(0,0)\Big)$$
$$+ G\Big(F(0,1) - F(0,0)\Big) + \sum_{g=1}^{G}\Big(u_G(0,g-1) - u_G(0,0)\Big) \qquad \text{(B.8)}$$

Hence, since $\mathcal{F} = \sum_{q \in \mathcal{Q}} F(R_q, G_q)$ one can obtain, after rearranging the different terms, a symmetric expression of the potential:

$$\mathcal{F} = |\mathcal{Q}|F(0,0) + N_R\big(F(1,0) - F(0,0) - u_R(0,0)\big) + N_G\big(F(0,1) - F(0,0) - u_G(0,0)\big)$$
$$+ \frac{1}{2}\Big[\sum_{r=1}^{R}\Big(u_R(r-1,0) + u_R(r-1,G)\Big) + \sum_{g=1}^{G}\Big(u_G(0,g-1) + u_G(R,g-1)\Big)\Big] \qquad \text{(B.9)}$$

Since the potential can be chosen up to a constant, it is clear from the previous expression that the choice of $F(0,0)$, $F(0,1)$, $F(1,0)$, $u_R(0,0)$ and $u_G(0,0)$ does not really matter. This justifies our choice to put them to zero to simplify the generic expressions of the potential given in Eq. 3.7 and 3.8.

According to Eq B.7, $F(R,G)$ can be interpreted as the sum of the settling utility of R red and G green agents, these agent settling one by one in an initially empty block, the red agents first and then the green ones. The same goes for Eq. B.8 while the green agents settle first, the red coming next. In fact, relation 3.5 ensures that the sum does not depend on the exact order with which the agents settle (see section 3.4 for more precisions). Hence $F(R_q, G_q)$ can be written as the mean of the settling utilities of the agents over all possible orders of settlement:

$$F(R_q, G_q) = \sum_{\substack{0 \le r \le R_q, \\ 0 \le g \le G_q \\ (r,g) \ne (0,0)}} \Big(\frac{\alpha_R(r,g)}{\alpha(R_q, G_q)}u_R(r-1,g) + \frac{\alpha_G(r,g)}{\alpha(R_q, G_q)}u_G(r,g-1)\Big)$$

where $\alpha(R_q, G_q) = (R_q + G_q)!$ is the number of settling orders of the $R_q + G_q$ agent living in block q, $\alpha_R(r,g)$ is the number of such orders in which a red agent settle after $r-1$ red and g green, in which case his utility is $u_R(r-1,g)$, and similarly $\alpha_G(r,g)$ is the number of such orders in which a green agent settles

30

after r red and $g - 1$ green, in which case the utility of this agent is $u_R(r - 1, g)$.

$\alpha_R(r, g)$ is the product of:

- R_q, the number of ways of choosing the rth red settling agent,

- $\binom{R_q - 1}{r - 1}\binom{G_q}{g}$, the number of ways of sharing out the $R_q - 1 + G_q$ other agents - either before or after this agent,

- $(r + g - 1)!$, the number of ways of ordering the agents who settle before and

- $(R_q + G_q - r - g)!$, the number of ways of ordering the agents who settle after.

$\alpha_G(r, g)$ can be similarly computed. We thus end up with a new formula for $F(R_q, G_q)$:

$$F(R_q, G_q) = \sum_{\substack{0 \leq r \leq R_q,\ 0 \leq g \leq G_q \\ (r,g) \neq (0,0)}} \binom{R_q}{r}\binom{G_q}{g} \frac{(r + g - 1)!\,(R_q + G_q - r - g)!}{(R_q + G_q)!} \big(r\, u_R(r - 1, g) \ + \ g\, u_G(r, g - 1) \big)$$

$$\text{(B.10)}$$

Appendix C. Some notions of coalitional games

A coalitional game is a pair $(\mathcal{N}, \nu)$ where $\mathcal{N} = \{1, ..., N\}$ is a set of $N = |\mathcal{N}|$ players and ν is a map that assigns to each subset or coalition $S \in \mathcal{P}(\mathcal{N})$ a real number or *worth* $\nu(S)$ such that $\nu(\varnothing) = 0$. Depending on the specific assumption of the given model, this worth can in certain cases be interpreted as a coalition utility. The main assumption in cooperative game theory is that the grand coalition $\mathcal{N}$ will form. The challenge is then to allocate the payoff $\nu(\mathcal{N})$ among the players in some fair way. A *solution concept* is a vector $\phi(\nu) \in \mathbb{R}^N$ whose components are the individual compensations. Researchers have proposed different solution concepts based on different notions of fairness.

The Shapley value Sh is one particular solution of which individual components are defined by:

$$Sh_i(v) = \sum_{S \subseteq \mathcal{N} \setminus \{i\}} \frac{|S|!\,(N - |S| - 1)!}{N!} (\nu(S \cup \{i\}) - \nu(S)) \tag{C.1}$$

This formula can be justified if one imagines the coalition being formed one actor at a time, with each actor demanding their contribution $\nu(S \setminus \{i\}) - \nu(S)$ to the coalition worth as a fair compensation, and then averaging over the possible different permutations in which this coalition can form. Properties verified by the Shapley value include:

- Efficiency: The payoff vector exactly splits the total value:

$$\sum_{i \in \mathcal{N}} Sh_i(v) = \nu(\mathcal{N})$$

- Symmetry: The solution concept ϕ allocates equal payments $\phi_i = \phi_j$ to symmetric players i and j where two players i and j are symmetric if $\nu(S \cup \{i\}) = \nu(S \cup \{j\}), \forall\, S \subseteq \mathcal{N} \setminus \{i, j\}$; that is, we can exchange one player for the other in any coalition that contains only one of the players and not change the payoff.

- Null Players: A null player should receive zero. A player i is null if $v(S \cup \{i\}) = v(S)$ for all S not containing i.

- Balanced contributions (BC): For all $i, j \in \mathcal{N}$,

$$\phi_i(\nu) - \phi_i(\nu^{\mathcal{N} \setminus \{j\}}) = \phi_j(\nu) - \phi_j(\nu^{\mathcal{N} \setminus \{i\}}). \tag{C.2}$$

BC means that for any two players, the benefit of each of them from the participation of the other is the same.

The Shapley value solution is the only efficient solution for which the BC rule is verified. Hart and Mas-Colell (1989) also show that the Shapley value is the only efficient solution to admit a *potential*. Let us denote by Γ the set of all coalitional games $(\mathcal{N}, \nu)$. A solution $\phi : \Gamma \to \mathbb{R}^N$ admits a potential if there exists a map $P : \Gamma \to \mathbb{R}$, called then the potential of ϕ, such that

$$\phi_i(\nu) = P(\nu) - P(\nu^{\mathcal{N} \setminus \{i\}}) \tag{C.3}$$

They show that the Shapley value's potential is given by:

$$P(\nu) = \sum_{S \subseteq \mathcal{N}} \frac{(|S| - 1)!\,(N - |S|)!}{N!} \nu(S) \tag{C.4}$$

Appendix D. Calculation of a potential function in Schelling case

Suppose that the agents compute their utility with Schelling utility function (which is equal to 1 if their fraction of similar neighbors is superior or equal to 0.5, and equal to 0 otherwise). This utility function can be expressed in terms of the number of red and green neighbors as follows:

$$
\begin{aligned}
u_R(R, G) &= \Theta(R - G) = \frac{1}{2}(1 + |R + 1 - G| - |R - G|) \\
u_G(R, G) &= \Theta(G - R) = \frac{1}{2}(1 + |R - 1 - G| - |R - G|)
\end{aligned} \tag{D.1}
$$

where Θ is the Heaviside function defined by: $\Theta(x) = 0$ if $x < 0$ and $\Theta(x) = 1$ if $x \geq 0$. Notice that in this example (and in this example only) the convention $u(0, 0) + 0$ used in proposition 1 is not respected. The

32

form we choose to write Schelling utility function imposes $u_R(0,0) = u_G(0,0) = 1$. It is easy to figure out that this particular pair of utility functions respect condition 3.5, and is therefore in the set $\mathbb{U}$. Indeed,

$$u_R(R,G) - u_R(R,G+1) = \Theta(R-G) - \Theta(R-G-1) = \begin{cases} 0-0 & = 0 \quad \text{if } R \leq G-1 \\ 1-0 & = 1 \quad \text{if } R = G \\ 1-1 & = 0 \quad \text{if } R \geq G+1 \end{cases}$$

and

$$u_G(R,G) - u_G(R+1,G) = \Theta(G-R) - \Theta(G-R-1) = \begin{cases} 1-1 & = 0 \quad \text{if } R \leq G-1 \\ 1-0 & = 1 \quad \text{if } R = G \\ 0-0 & = 0 \quad \text{if } R \geq G+1 \end{cases}$$

Hence relation $u_R(R,G) - u_R(R,G+1) = u_G(R,G) - u_G(R+1,G)$ is always verified.

To compute a corresponding potential function, one can refer to the general form of Eq. B.9 (since we do not use the convention $u(0,0) = 0$ in this particular example) which can be written here as:

$$
\begin{aligned}
\mathcal{F} &= const + \frac{1}{2} \sum_{q \in \mathcal{Q}} \Big[\sum_{r=0}^{R_q-1} \Big(u_R(r,0) + u_R(r,G_q) \Big) + \sum_{g=0}^{G_q-1} \Big(u_G(0,g) + u_G(R_q,g) \Big) \Big] \\
&= const + \frac{1}{4} \sum_{q \in \mathcal{Q}} \Big[\sum_{r=0}^{R_q-1} \Big(3 + |r+1-G_q| - |r-G_q| \Big) + \sum_{g=0}^{G_q-1} \Big(3 + |R_q-1-g| - |R_q-g| \Big) \Big] \\
&= const + \frac{1}{4} \sum_{q \in \mathcal{Q}} \Big[\Big(3R_q + |R_q-G_q| - G_q \Big) + \Big(3G_q + |R_q-G_q| - R_q \Big) \Big] \\
&= const + \frac{1}{2} \sum_{q \in \mathcal{Q}} \Big(R_q + G_q + |R_q-G_q| \Big) \\
&= const + \frac{1}{2}(N_R + N_G) + \frac{1}{2} \sum_{q \in \mathcal{Q}} |R_q - G_q| \\
&= const' + \frac{1}{2} \sum_{q \in \mathcal{Q}} |R_q - G_q|
\end{aligned}
$$

Appendix E. Relation between the potential function $\mathcal{F}$ and the collective utility U

Let us suppose that $(u_R, u_G) \in \mathbb{U}$, and that the potential function of the system can be expressed as a linear function of the collective utility, ie $\mathcal{F}(\{R_q, G_q\}) = \lambda U(\{R_q, G_q\}) + \mu$. Since the potential function can

be defined up to constant, we can take $\mu = 0$. Writing the utility functions under the form

$$u_R(R,G) \;=\; \xi_R(R) + \sum_{g=0}^{G-1} \xi(R,g)$$

$$u_G(R,G) \;=\; \xi_G(G) + \sum_{r=0}^{R-1} \xi(r,G)$$

introduced in Eq. 3.9 and 3.10, the relation of proportionality between the potential and the collective utility can be written as

$$\sum_q \Big(\sum_{r=0}^{R_q-1} \xi_R(r) + \sum_{g=0}^{G_q-1} \xi_G(g) + \sum_{r=0}^{R_q-1}\sum_{g=0}^{G_q-1} \xi(r,g) \Big)$$

$$= \; \lambda \sum_q \Big(R_q \xi_R(R_q-1) + R_q \sum_{g=0}^{G_q-1} \xi(R_q-1,g) + G_q \xi_G(G_q-1) + G_q \sum_{r=0}^{R_q-1} \xi(r,G_q-1) \Big)$$

Since this relation must hold for all $\{R_q, G_q\}$, it follows that that for all $(R,G) \in E_H$, the following holds:

$$\sum_{r=0}^{R-1} \xi_R(r) + \sum_{g=0}^{G-1} \xi_G(g) + \sum_{r=0}^{R-1}\sum_{g=0}^{G-1} \xi(r,g)$$

$$= \; \lambda \Big(R\xi_R(R-1) + R \sum_{g=0}^{G-1} \xi(R-1,g) + G\xi_G(G-1) + G \sum_{r=0}^{R-1} \xi(r,G-1) \Big) \quad\quad\quad \text{(E.1)}$$

Taking successively $G = 0$ and $R = 0$ in that last equation provides three independent relations dissociating the three functions ξ_R, ξ_G and ξ:

$$\forall R > 0, \quad\quad\quad\quad \textstyle\sum_{r=0}^{R-1} \xi_R(r) = \lambda R \xi_R(R-1) \quad\quad\quad\quad\quad\quad \text{(E.2)}$$

$$\forall G > 0, \quad\quad\quad\quad \textstyle\sum_{g=0}^{G-1} \xi_G(g) = \lambda G \xi_G(G-1) \quad\quad\quad\quad\quad\quad \text{(E.3)}$$

$$\forall (R,G), \in E_H \quad \textstyle\sum_{r=0}^{R-1}\sum_{g=0}^{G-1} \Big(\lambda\xi(R-1,g) + \lambda\xi(r,G-1) - \xi(r,g) \Big) = 0 \quad\quad\quad \text{(E.4)}$$

Notice moreover that the convention $u(0,0) = 0$ implies $\xi_R(0) = \xi_G(0) = 0$. Let us also define $a = \xi_R(1)$, $d = \xi_G(1)$ and $b = \xi(0,0)$. Starting from equations E.2 to E.4, it is straightforward to prove recursively that

$$\lambda = 1/2$$

$$\forall R > 0, \quad \xi_R(R) = aR$$

$$\forall G > 0, \quad \xi_G(G) = dG$$

$$\forall (R,G) \in E_H \quad \xi(R,G) = b$$

34

Hence the agents' utility functions corresponds exactly to those introduced in Eq. 4.1:

$$u_R(R, G) = aR + bG$$

$$u_G(R, G) = bR + dG$$

The individual utilities are thus necessarily linear in the numbers of similar and dissimilar neighbors in case the potential function $\mathcal{F}$ is proportional to the collective utility U □

Appendix F. Potential and collective utility in the case of the asymmetrically peaked utility

Proof of equation 4.12

We have to derive the expression $\tilde{F}(S) = \sum_{s=0}^{S-1} \xi_{ap}(s)$, where

$$\begin{cases} \xi_{ap}(s) & = 2s/H & \text{if } s \le H/2 \\ \xi_{ap}(s) & = 2 - m - 2(1-m)s/H & \text{if } s > H/2 \end{cases}$$

For $S - 1 \le H/2$, it is straightforward to write:

$$\tilde{F}(S) = \frac{2}{H} \sum_{s=0}^{S-1} s = \frac{(S-1)S}{H} \tag{F.1}$$

For $S - 1 > H/2$, one has

$$
\begin{aligned}
\tilde{F}(S) & = \frac{2}{H} \sum_{s=0}^{H/2} s + (2-m)(S-1-H/2) - (1-m)\frac{2}{H} \sum_{s=H/2+1}^{S-1} s \\
& = \frac{2}{H}(2-m) \sum_{s=0}^{H/2} s + (2-m)(S-1-H/2) - (1-m)\frac{2}{H} \sum_{s=0}^{S-1} s \\
& = (2-m)(H/4 + 1/2 + S - 1 - H/2) - (1-m)\frac{(S-1)S}{H} \\
& = (2-m)\left[S - H/4 - 1/2 - \frac{(S-1)S}{H} \right] + \frac{(S-1)S}{H} \\
& = \frac{(S-1)S}{H} - (2-m)\frac{1}{H}\left[-\left(S - \frac{H}{2}\right)\left(\frac{H}{2}\right) - (S-1)\left(\frac{H}{2}\right) + (S-1)S \right] \\
& = \frac{(S-1)S}{H} - \frac{2-m}{H}\left(S - \frac{H}{2} - 1\right)\left(S - \frac{H}{2}\right) \tag{F.2}
\end{aligned}
$$

Thanks to the Heaviside function both results can then be written under the general form:

$$\tilde{F}(S) = \frac{(S-1)S}{H} - \frac{2-m}{H}\left(S - \frac{H}{2} - 1\right)\left(S - \frac{H}{2}\right)\Theta\left(S - \frac{H}{2} - 1\right) \tag{F.3}$$

Proof of equations 4.13 and 4.14

35

The computation of $\Delta\mathcal{F}$ in relation 4.13 is based on the expression of $\tilde{F}(S)$ (equation F.3), which gives, with $K \in \{0, 1, ..., H/2\}$:

$$
\begin{aligned}
\Delta\mathcal{F} &= 2\tilde{F}(H/2 + 1 + K) + 2\tilde{F}(H/2 - K) - 2\tilde{F}(H/2 + 1) - 2\tilde{F}(H/2) \\
&= \frac{2}{H}\left[\left(\frac{H}{2} + K\right)\left(\frac{H}{2} + K + 1\right) + \left(\frac{H}{2} - K - 1\right)\left(\frac{H}{2} - K\right) - \left(\frac{H}{2}\right)\left(\frac{H}{2} + 1\right) - \left(\frac{H}{2} - 1\right)\left(\frac{H}{2}\right)\right] \\
&\quad - \frac{2(2 - m)}{H}K(K + 1) \\
&= \frac{2m}{H}K(K + 1)
\end{aligned}
$$

In the same way,

$$
\begin{aligned}
\Delta U &= 2\tilde{U}(H/2 + 1 + K) + 2\tilde{U}(H/2 - K) - 2\tilde{U}(H/2 + 1) - 2\tilde{U}(H/2) \\
&= \frac{4}{H}\left[\left(\frac{H}{2} + K\right)\left(\frac{H}{2} + K + 1\right) + \left(\frac{H}{2} - K - 1\right)\left(\frac{H}{2} - K\right) - \left(\frac{H}{2}\right)\left(\frac{H}{2} + 1\right) - \left(\frac{H}{2} - 1\right)\left(\frac{H}{2}\right)\right] \\
&\quad - \frac{4(2 - m)}{H}K(H/2 + K + 1) \\
&= \frac{8}{H}K(K + 1) - \frac{4(2 - m)}{H}K(H/2 + K + 1) \\
&= \frac{4K}{H}\big(m(H/2 + K + 1) - H\big) \\
&= 2\Delta\mathcal{F} - 2(2 - m)K \tag{F.4}
\end{aligned}
$$

36

References

Anderson, S.P., De Palma, A., Thisse, J.-F., 1992. *Discrete Choice Theory of Product Differentiation*. MIT Press, Cambridge (MA).

Barr, J., Tassier, T., 2008. Segregation and strategic neighborhood interaction. *Eastern Economic Journal*, 34, 480–503.

Card, D., Mas, A., Rothstein J., 2008. Tipping and the dynamics of segregation. *The Quarterly Journal of Economics*, 123, 177–218.

Clark, W.A.V., Fossett, M., 2008. Understanding the social context of the Schelling segregation model. *Proceedings of the National Academy of Sciences*, 105, 4109–4114.

Cutler, D.M., Glaeser, E.L., Vigdor, J.L., 2008. Is the melting pot still hot? Explaining the resurgence of immigrant segregation. *Review of economics and statistics*, 90, 478–497.

Dokumaci, E., Sandholm, W.H., 2007. *Schelling Redux: An Evolutionary Dynamic Model of Residential Segregation*. Unpublished manuscript, University of Wisconsin-Madison.

Duncan, O.D., Duncan, B., 1955. A methodological analysis of segregation indexes. *American Sociological Review*, 20, 210–217.

Fagiolo, G., Valente, M., Vriend, N.J., 2007. Segregation in networks. *Journal of Economic Behavior and Organization*, 64, 316–336.

Farley, R., Schuman, H., Bianchi, S., Colasanto D., Hatchett, S., 1978. Chocolate city, vanilla suburb: will the trend toward racially separate communities continue? *Social Science Research*, 7, 319–344.

Farley, R., Fielding, E.L., Krysan, M., 1997. The residential preferences of balcks and whites: a four-metropolis analysis. *Housing Policy Debate*, 8, 763–800.

Grauwin, S., Bertin, E., Lemoy R., Jensen, P., 2009a. Competition between collective and individual dynamics. *Proceeding of the National Academy of Sciences USA*, 106, 20622-20626.

Grauwin, S., Goffette-Nagot, F., Jensen, P., 2009b. Dynamic models of residential segregation: brief review, analytical resolution and study of the introduction of coordination. GATE Working Paper, 2009 (available at ftp://ftp.gate.cnrs.fr/RePEc/2009/0914.pdf).

Hart, S., Mas-Colell, A., 1989. Potential, value and consistency. *Econometrica*, 57, 589–614.

Iceland, J., Scopilliti, M., 2008. Immigrant residential segregation in US metropolitan areas. 1990–2000. *Demography*, 45, 79–94.

Mobius, M.M., Rosenblat, T.S., 2000. *The Formation of Ghettos as a Local Interaction Phenomenon*. Unpublished manuscript, Harvard University.

Monderer, D., Shapley, L.S., 1996. Potential games. *Games and Economic Behavior*, 14, 124–143.

O'Sullivan, A., 2009. Schelling's model revisited: Residential sorting with competitive bidding for land. *Regional Science and Urban Economics*, 39, 397–408.

Pancs, R., Vriend, N.J., 2007. Schelling's spatial proximity model of segregation revisited. *Journal of Public Economics*, 91, 1–24.

Pollicott, M., Weiss, H., 2001. The Dynamics of Schelling-type segregation models and a nonlinear graph laplacian variational problem. *Advances in Applied Mathematics*, 27, 17–40.

Reardon, S.F., Matthews, S.A., O'Sullivan, D., Lee, B.A., Firebaugh, G., Farrell, C.R., Bischoff, K., 2008. The geographic scale of metropolitan racial segregation. *Demography*, 45, 489–514.

Schelling, T.C., 1969. Models of Segregation. *American Economic Review, Papers and Proceedings*, 59, 488–493.

Schelling, T.C., 1971. Dynamic models of segregation. *Journal of Mathematical Sociology*, 1, 143–186.

Schelling, T.C., 1978. *Micromotives and Macrobehavior*. Norton, New York.

Shapley, L.S., 1953. A Value for n-Person Games. in: *Contributions to the Theory of Games II*, ed. by Kuhn H. W., TuckerA. W., 1953. Princeton University Press, Princeton. 307–317.

Tiebout, C.M., 1958. A pure theory of local expenditures. *The Journal of Political Economy*, 64, 416.

Ui, T., 2000. A Shapley value representation of potential games. *Games and Economic Behavior*, 31, 121–135.

Vinkovic, D., Kirman, A., 2006. A physical analogue of the Schelling model. *Proceedings of the National Academy of Sciences*, 103, 19261–19265.

Young, H.P., 1998. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton University Press, Princeton.

Zhang, J., 2004a. A dynamic model of residential segregation. *The Journal of Mathematical Sociology*, 28, 147–170.

Zhang, J., 2004b. Residential segregation in an all-integrationist world. *Journal of Economic Behavior and Organization*, 54, 533–550.

Zhang, J., 2010. Tipping and residential segregation: a unified Schelling model. *Journal of Regional Science*, forthcoming.
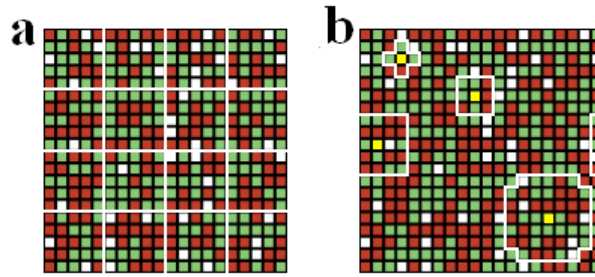
Figure 1: **Different forms of neighborhood.** Red, green and white squares denote respectively red agents, green agents and vacant cells. **a.** Example of a bounded neighborhood in which the city is divided in square blocks containing $H + 1 = 25$ cells/locations; **b.** In the case of a continuous neighborhood description, the neighborhood of an agent corresponds to his H nearest cells/locations. Around the agents marked in yellow, we enlightened by the white frontiers a $H = 4$, a $H = 8$, a $H = 24$ and a $H = 44$ continuous neighborhood. [If you printed this document in black and white, the red and green squares should appear respectively in dark grey and soft grey.]
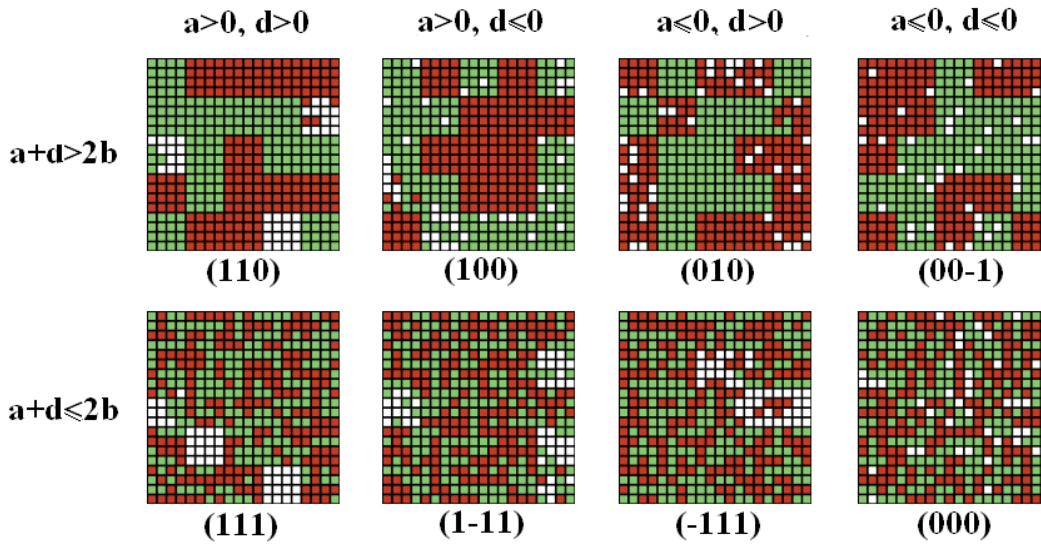

Figure 2: **Typical stationary configurations obtained by simulations for different values of (adb). Top panel:** for $2b - (a + d) < 0$, the system evolves towards segregated configurations where red and green agents tends to live in different blocks. **Bottom panel:** for $2b - (a + d) > 0$, the system evolves towards mixed configurations where the number of red-green pairs of neighbors is maximized. **From left to right:** the sign of a and d controls the tendency of red and green agent to prefer to live in dense or uncrowded areas. The demographic parameters are ($N = 20, v = 10\%, n_R = 0.5$). Neighborhood size is fixed to $H + 1 = 16$ and the level of noise is $T = 0.1$
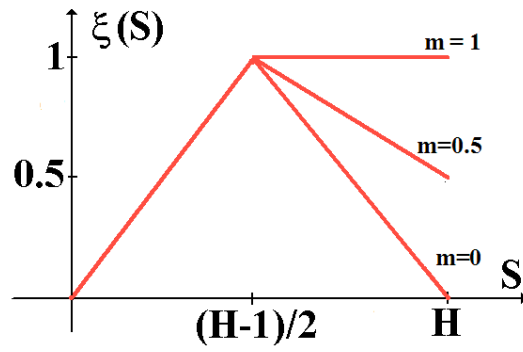
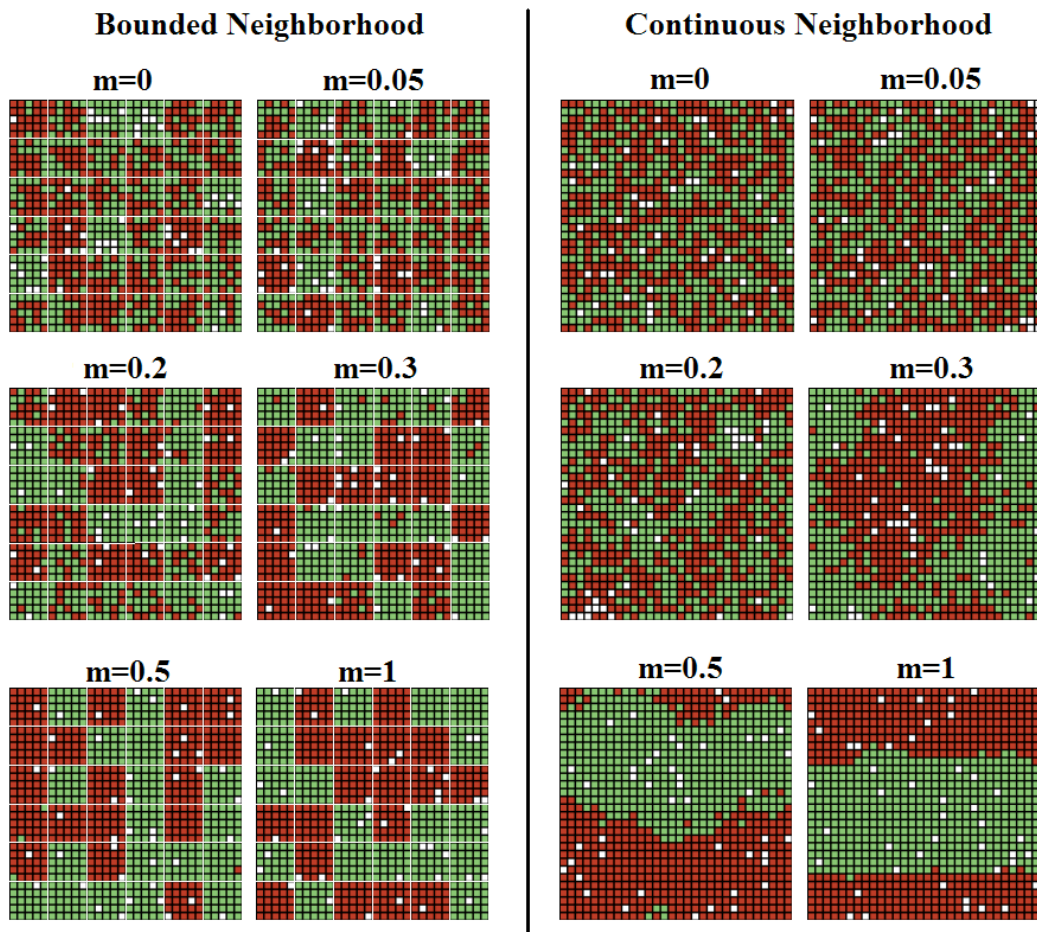Figure 3: **Asymmetrically peaked function for some values of** m.

Figure 4: **Typical stationary configurations obtained by simulations with the asymmetrically peaked utility function.** The demographic parameters are $(N = 30, v = 5\%, n_R = 0.5)$. Neighborhood sizes are fixed to $H = 24$, and the level of noise is fixed to $T = 0.1$. Left: with a bounded neighborhood description. Right: with a continuous neighborhood description.