Laboratory of Economics and Management
Sant'Anna School of Advanced Studies

Piazza Martiri della Libertà, 33 - 56127 PISA (Italy)
Tel. +39-050-883-343 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://www.lem.sssup.it/

LEM
Working Paper Series

Segregation in Networks

Giorgio FAGIOLO*
Marco VALENTE°
Nicolaas J. VRIENDˆ

*University of Verona, and Sant'Anna School of Advanced Studies, Pisa, Italy
°University of L'Aquila, Italy
ˆQueen Mary, University of London, UK

2005/22 **November 2005**

ISSN (online) 2284-0400

Segregation in Networks*

Giorgio Fagiolo[†] Marco Valente[‡] Nicolaas J. Vriend[§]

November 22, 2005

Abstract

Schelling (1969, 1971a,b, 1978) considered a simple model with individual agents who only care about the types of people living in their own local neighborhood. The spatial structure was represented by a one- or two-dimensional lattice. Schelling showed that an integrated society will generally unravel into a rather segregated one even though no individual agent strictly prefers this. We make a first step to generalize the spatial proximity model to a proximity model of segregation. That is, we examine models with individual agents who interact 'locally' in a range of network structures with topological properties that are different from those of regular lattices. Assuming mild preferences about with whom they interact, we study best-response dynamics in random and regular non-directed graphs as well as in small-world and scale-free networks. Our main result is that the system attains levels of segregation that are in line with those reached in the lattice-based spatial proximity model. In other words, mild proximity preferences can explain segregation not just in regular spatial networks but also in more general social networks. Furthermore, segregation levels do not dramatically vary across different network structures. That is, Schelling's original results seem to be robust also to the structural properties of the network.

Keywords: Spatial proximity model, Social segregation, Schelling, Proximity preferences, Social networks, Undirected graphs, Best-response dynamics.

JEL Classification: C72, C73, D62.

*We thank Yannis Ioannides, Giulia Iori, Alan Kirman, Akira Namatame, workshop participants of iNeck (Strasbourg), WEHIA 2005 (Colchester) and SCE 2005 (Washington DC) as well as the theory reading group at QM (London) for their comments. Part of this work was done while G.F. was visiting the Bureau d' Economie Théorique et Appliquée (BETA), Université Louis Pasteur, Strasbourg, France. G.F. gratefully acknowledges financial support by the CNRS (Centre National de la Recherche Scientifique), France. All usual caveats apply.

[†]University of Verona, Verona (Italy) and Sant'Anna School of Advanced Studies, Pisa (Italy).

[‡]University of L'Aquila, L'Aquila (Italy).

[§]Corresponding Author. Department of Economics, Queen Mary, University of London, Mile End Road, London E1 4NS, U.K. Email: `n.vriend@qmul.ac.uk`.

1 Introduction

Segregation has been for some time one of the most important socio-political and public economic issues in the USA, and has increasingly become one also in many Western-European countries. As segregation has increasingly been recognized as one of the most important public policy issues in countries such as the UK, the Netherlands, France, and Germany, various countries have started evaluating and questioning the effectiveness of decades of integration policies (see, e.g., Baldwin and Rozenberg (2004), and Commissie Blok (2004)). The widely accepted view is that these policies have essentially been failures as integration simply did not happen. As Trevor Phillips, chairman of the Commission for Racial Equality in the UK, puts it: "we are sleepwalking our way to segregation" (Phillips, 2005). The main objective of this paper is to improve our understanding of this issue.

The prevalent form of integration policy in countries such as the UK and the Netherlands has been one of promoting multiculturalism by focusing on the individual citizens' preferences[1]. The idea was that by promoting openness and tolerance with respect to diversity one would allow integration to take place.

Individual preferences are exactly what the spatial proximity model of Schelling (1969, 1971a,b, 1978) focuses on. Schelling considered a simple model with individual agents who only care about the types of people living in their own local neighborhood. The spatial structure was represented by a one- or two-dimensional lattice. Schelling showed that an integrated society will generally unravel into a rather segregated one even though no individual agent strictly prefers this. This segregation is due to the spontaneous dynamics of the economic forces, with all individuals following their incentives to move to the most attractive locations. In doing so, they create externalities for other people, who will respond to their changed incentives, etc.

The preferences considered in the spatial proximity model are said to be mild, as everybody would be happy in a perfectly integrated society[2]. Pancs and Vriend (2005)

[1] This focus can be explained by the practical difficulties with other policy measures aimed at integration (see Pancs and Vriend (2005) for details).

[2] As this occurs without any of the individuals involved explicitly designing this outcome, the sleep-

examined the robustness of the spatial proximity model. They showed that the model can be further simplified (rendering the individual preferences even more salient as an explanatory variable of segregation), and that these proximity preferences may be even more extreme in favor of integration. This focus on mild individual preferences or preferences that even favor integration is not to say that institutional constraints or racism may not hinder integration. But what the model shows is that even without such obstacles one should perhaps expect segregation. It seems that any integration policy must be based on a good understanding of these spontaneous dynamics.

The idea that people care about their spatial proximity can be justified by the fact that this is where people mow their lawn, where their children play outside, where they do their shopping, and where they park their car. The social environment is, however, not limited to this spatial proximity. People also interact through networks of friends, relatives, and colleagues, and through virtual communities on the internet. And they are likely to have preferences with whom they do this, just as they have preferences about their spatial proximity. Similarly, segregation need not necessarily occur at the spatial (neighborhood) level. One might conceive people who are socially segregated despite being spatially integrated[3].

Therefore, in this paper we will make a first step to generalize the spatial proximity model to a proximity model of segregation. That is, we will examine models with individual agents who interact 'locally' in a range of network structures with topological properties that are different from those of regular lattices, while having mild preferences about with whom they interact. We stick to standard assumptions as far as types and preferences are concerned, and we study best-response dynamics. Apart from the sociopolitical interest in this, this seems also intellectually interesting as, after all, a lattice is just a special type of network, and we may want to know whether mild proximity preferences can explain segregation also in more general types of networks.

The paper is organized as follows. In Section 2 we present the model. Section 3

walking metaphor may seem appropriate.

[3]This appeared to be the case with some of the recent terror suspects in the Netherlands and the UK.

discusses in more detail the classes of networks that we consider in our analysis. The various indices used to measure segregation in social networks are introduced in Section 4. Section 5 contains our analysis of the model, including a sensitivity analysis of the parameter setup. Finally, Section 6 concludes.

2 The Model

Consider a society composed of N agents who can locate themselves in one of the $M \geq N \geq 3$ available locations. Each location can contain at most one agent. Locations can be connected or not. We model locations and connections through a non-directed graph (NDG) G composed of M nodes and a collection of non-directed edges linking any pair of nodes. Edges are described by the (symmetric) $M \times M$ sociomatrix $W = \{w_{kh}\}$, where $w_{kk} = 0$ $\forall k = 1, ..., M$ and $w_{kh} = w_{hk} = 1$ if and only if there is an edge connecting nodes k and h, and zero otherwise. We define the "neighborhood" V_k (or the "interaction group") of a node k as the set of nodes that node k is linked to:

$$V_k = \{h \in I_M : w_{kh} = w_{hk} = 1\}, \tag{1}$$

where $I_M = \{1, ..., M\}$.

We suppose that each node is empty (i.e., it does not contain an agent) with probability $\theta \in (0, 1)$, while it is occupied with probability $1 - \theta$. Therefore, on average, there are $N = (1 - \theta)M$ agents in the society. Each agent can be one of two types, say -1 and $+1$. Time is discrete, and time ticks are labeled by $t = 0, 1, 2, ...$

Agents have standard, binary, Schelling-type preferences: they are happy if and only if the relative frequency of agents of their own type is greater or equal than 0.50 in their neighborhood. More formally, if node i is occupied by an agent of type $s \in \{-1, +1\}$ at time t:

$$u_{it} = u_{it}(s) = \begin{cases} 1 & if \quad x_{it}(s) \geq 0.5 \\ 0 & \quad \text{otherwise} \end{cases}, \tag{2}$$

where $u_{it} = u_{it}(s)$ is the utility of agent i (of type s) at time t and $x_{it}(s)$ is the current relative frequency of agents (i.e., filled nodes) of type s in V_i[4].

The initial state of the system is characterized by: (i) an instance of the network structure, i.e., a graph $G_0 = \{I_M, W_0\}$ (more on that below); (ii) an allocation of agents and types across the M available nodes. The initial allocation of agents and types across the M nodes is drawn uniform randomly. Thus, at $t = 0$, each node $i \in I_M$ will be either empty or occupied. If it is occupied, this will be either a -1 or a $+1$ agent, each with probability 0.5. Thus, in the society there will be, on average, $N/2$ agents of type -1 and $N/2$ agents of type $+1$.

The dynamics is as follows. At each $t > 0$, an agent is drawn at random (and independently) from $I_N = \{1, ..., N\}$. This agent checks every available node in the network G_0, i.e., his current node plus all empty nodes, and computes the utility that he could earn at each of these nodes. The agent chooses the node that provides the highest achievable utility level. Agents resolve ties by randomizing among all nodes providing the same maximal utility level.

Notice that we assume no inertia in the agents' choices. That is, the agents' current locations do not bias their choices (e.g., because of moving costs). We also assume that agents can move to any empty node in the network, i.e., there are no information or moving constraints or costs. In Section 5.3 we will study the effect of removing these assumptions.

3 Network Structures

To investigate the scope of proximity preferences explaining segregation, we do not constrain the graph G_0 to be a lattice-type of network. Instead, we explore a number of classes of NDGs characterized by very different structural and topological properties. This allows us to investigate how segregation levels, emerging out of the best-response dynamics described above, may depend on the type of network. We study six classes

[4] In line with Pancs and Vriend (2005), we assume that the utility associated to an empty neighborhood is zero.

of NDGs: two-dimensional lattices with Von-Neumann neighborhoods (*2D-VN*), two-dimensional lattices with Moore neighborhoods (*2D-M*), regular NDGs (*REG*), and random (*RAND*), small-world *(SW)* and scale-free (*SF*) NDGs. We now discuss each of these network structures in more detail.

1. **Two-dimensional boundary-less lattices with Von-Neumann neighborhoods (2D-VN)**. The two lattice-type of networks are considered to benchmark our analysis against the standard Schelling model. The difference between these two classes of lattices lies in the metrics employed to compute the distance among any two nodes. Neighborhoods are accordingly defined as containing all nodes that lie within a certain integer interaction radius $r \geq 1$. In the 2D-VN lattice, the "Manhattan" metrics is used and neighborhoods of radius $r \geq 1$ have a "diamond" shape. That is, if any node h has coordinates (x_h, y_h), the distance is defined as:

$$\delta(k, k') = |x_k - x_{k'}| + |y_k - y_{k'}|, \tag{3}$$

where k and k' are any two nodes. We avoid singularities in the lattice by placing nodes on a torus. A node's neighborhood is thus defined as:

$$V_k(r) = \{h = 1, ..., M : \delta(k, h) \leq r\}, \tag{4}$$

where r is the neighborhood radius. Notice that the degree d_k of any node k, i.e., the number of inward (and outward) links to (and from) k, is:

$$d_k(r) = |V_k(r)| = 2r(r + 1). \tag{5}$$

See Figure 1, panel (a), for an example of the shape of $V_k(r)$ for $r = 1$ in the case of 2D-VN. Notice that the links actually in place in a 2D-VN lattice with $r > 1$ differ from those present in the lattice (with $r = 1$). In fact, any agent is connected through additional direct links with all agents placed $2, ..., r$ steps away. For example, if $r = 2$, the agent placed in the node (x, y) is not only linked

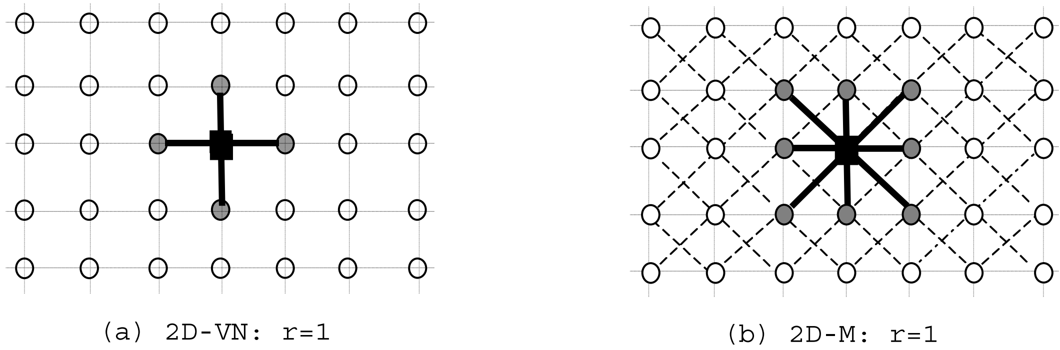

Figure 1: An Example of Neighborhood Shapes with 2-Dimensional Von-Neumann (2D-VN) and Moore (2D-M) lattices for $r = 1$

with the agents placed in $(x-1, y), (x+1, y), (x, y-1), (x, y+1)$ (whose links with (x, y) are present in the lattice), but also with agents placed two steps away in the underlying lattice.

2. **Two-dimensional boundary-less lattices with Moore neighborhoods (2D-M)**. In this case, the lattice is endowed with the "spherical" metrics, which entails "box-shaped" neighborhoods:

$$\delta(k, k') = \max\{|x_k - x_{k'}|, |y_k - y_{k'}|\}, \tag{6}$$

where k and k' are any two nodes. Again, we avoid singularities in the lattice by placing nodes on a torus. Here, the degree d_k of any node k is:

$$d_k(r) = |V_k(r)| = 4r(r+1), \tag{7}$$

where $V_k(r)$ is defined as in (4). Notice that given any interaction radius $r \geq 1$, 2D-M lattices have neighborhoods twice as large as those of 2D-VN lattices. An alternative way to put this (for the case $r = 1$) is to say that the 2D-VN lattices have only lateral (horizontal and vertical) links, whereas the 2D-M lattices have all diagonal links in place as well. An example of the shape of $V_k(r)$ for $r = 1$ in the case of 2D-M graphs is reported in Figure 1, panel (b), where such additional links for the agent concerned are shown with dashed lines.

As happens with 2D-VN lattices, also here for $r > 1$ there are additional links in place. For example, for $r = 2$, also agents that would be two steps away from each other in the 2D- M lattice with $r = 1$ are directly linked.

3. **Regular NDG (REG)**. Two-dimensional lattices are regular NDGs, i.e., NDGs where all nodes hold the same number of edges (i.e., have the same *degree*). Lattices, however, possess further spatial homogeneity and symmetry properties, such as invariance to roto-translation, i.e., all neighborhoods are invariant up to a translation in space and/or a rotation around their center. Therefore, the third class of NDG we explore is the one of *regular* NDGs, which are simply defined as NDGs where all nodes have the same degree d, but do not necessarily satisfy the additional spatial homogeneity and symmetry properties that lattices do. At time $t = 0$, we choose at random a regular NDG of degree d using the algorithm proposed by Steger and Wormald (1999). Under this routine, regular graphs of degree d are (approximately) generated uniformly at random.

4. **Random NDG (RAND)**. We, then, discard the hypothesis of regularity, by considering the class of *random* NDGs. Given an average degree equal to $(M - 1)p$, we generate the graph by allowing each edge to be in place, independently of all other edges, with a probability p. Therefore, unlike in regular NDGs, nodes will generally have different degrees. In fact, the degree distribution is symmetric, appears to be quite dispersed over the support $\{0, ..., M - 1\}$, and has average degree $\overline{d} = (M - 1)p$.

5. **Small-World NDG (SW)**. Next, we consider two additional NDG classes that, due to their close relationships with empirically observed social and economic networks (for an introduction, see, e.g., Barabási (2003)), have received an increasing attention over the last decade. First, we study *small-world* NDGs *(SW)*. The main features of SW networks (Watts, 1999, 2003) are that they tend to have a small "path- length" (i.e., average distance between any two nodes) and a large "clustering coefficient" (i.e., likelihood that any two neighbors of an agents are also neighbors of each other). To generate a small-world NDG, we start from a two- dimensional

boundary-less lattice with Von-Neumann neighborhoods as in 1. above, for a certain value of r. Then, each edge (h, k) is independently rewired to a randomly chosen node, say k', outside $V_h(r)$ with some probability $\beta \in (0, 1)$. In case of rewiring, the edge (h, k) is deleted and replaced by the new edge (h, k'). If β is close to one and M is very large, this procedure yields a random NDG. Otherwise, the degree distribution will be symmetric, centered around $2r(r + 1)$ but less dispersed than the one associated to a random NDG. In the results presented, we employ $\beta = 0.5$[5].

6. **Scale-Free NDG (SF)**. Finally, we consider *scale-free* NDGs (*SF*). A scale-free NDG has a skewed, power-law degree distribution, with few nodes holding a large number of edges (i.e., the hubs of the network) and many nodes with few edges. To build a SF graph, we employ a standard "preferential attachment" procedure (Barabási and Albert, 1999), starting with M_0 nodes linked through a 2D-VN lattice with $r = 1$ (and thus an initial degree $d = 4$). One node at a time is added until a size M is reached. In any step, the additional node is allowed to form 4 links. Each new link is formed by choosing one of the existing nodes with a probability proportional to its current degree. The larger the initial number of nodes, the smaller is the degree heterogeneity (the less skewed the degree distribution). The underlying assumption of this setup is that any node can hold at no cost any arbitrarily large number of nodes (as M increases). When a SF network is in place, there is a small set of agents that hold a large number of links, while all the others hold a small number of links. Thus, the degree distribution is skewed to the right, with a relatively small average and a fat right tail. As expected, the average degree of a SF network generated from an initial node size M_0 tends to have always the same average degree. Moreover, this average degree *is not* monotonically increasing in M_0. In fact, our analysis shows that the following (approximate) relation holds:

$$d \simeq 0.00003 \cdot M_0^3 - 0.0062 \cdot M_0^2 + 0.3485 \cdot M_0 + 3.1916. \tag{8}$$

[5]See Watts and Strogatz (1998) All our results are not altered if one tunes β in the range of small-world graphs.

Hence, D grows for $M_0 \leq 39$ and decreases for $M_0 \geq 40$.

Given a certain class within this range of networks, we need to specify the system- and network-specific parameters. System parameters are M (number of nodes) and θ (average percentage of empty nodes). Network specific parameters characterize – given the class of networks to be implemented – the set of possible networks from which the one actually in place will be drawn (see Table 1).

	Parameters		
Network	System	Network-specific	Node degree
2D-VN		r	$d = 2r(r+1)$
2D-M		r	$d = 4r(r+1)$
REG	(M, θ)	d	d
RAND		p	$\overline{d} = (M-1)p$
SW		r	$\overline{d} = 2r(r+1)$
SF		M_0	See (8)

Table 1: Network Classes and System Parameters: M: Number of Nodes; θ average percentage of empty nodes; d: Graph Average Degree; r: Interaction Radius; p: Link Probability; M_0: Initial Number of Nodes in a preferential-attachment graph formation mechanism.

Table 1 also shows how the node degrees depend on the chosen network-specific parameter values. When we compare segregation levels in networks from different classes, we will keep the system parameters and the average node degree constant across the networks considered. We achieve this by tuning the network-specific parameters. Thus, the values for the radius r in the 2D-VN and 2D-M lattices, the probability p in the random graphs, the radius r in the starting lattice for the small world networks, and M_0 in the scale-free networks are all chosen such that the ensuing (average) node degrees d match those considered for the other network classes. In other words, implicitly the only network-specific parameter to be considered is the (average) node degree d.

In the lattice-case (i.e., 2D-VN and 2D-M), the initial graph is automatically defined once one specifies the interaction radius r (and consequently the degree d). In all other cases, given a choice for the network class and for the network-specific parameters of that class (e.g., the degree d in a regular graph), each time we draw G_0 (uniform) randomly from the set of all possible graphs belonging to that class and with the given network-specific parameters (e.g., all regular graphs with degree d).

4 Measuring Segregation in Social Networks

A number of indices have been suggested in the literature to measure segregation when the agents are located on generic NDGs (see, e.g., Freeman, 1972; Mitchell, 1978; Freeman, 1978; Fershtman, 1997; Echenique and Fryer, 2005, and references therein). We mainly employ two indices. The first one is *Freeman's segregation index* (FSI) (Freeman, 1972, 1978), see the Appendix for details. The rationale underlying the computation of the FSI is that if a given agent-attribute (in our case the type $+1$ or -1) does not matter for social relationships (i.e., for the link structure as described by G_0), then the links among the agents should be distributed randomly with respect to that attribute. Therefore, suppose we observe a given allocation of agent types across the M nodes, connected through the NDG G_0. Let us, then, split the agents in two groups according to their type and, *for each type*, let us count the number of cross-group links (i.e., the number of links connecting any pair of agents of different types), as well the number of within-group links (i.e., the number of links connecting any pair of agents of the same type). This gives us a 2×2 contingency table whose generic entry l_{xy} gives us the number of links between type-x and type-y agents in G_0. Similarly, one can compute the expected contingency table for a random allocation of agent types on G_0. The difference between the number of cross-group ties expected by chance and the number of observed ties (divided by expected ones) gives us the FSI. The index ranges between -1 and 1, with the highest segregation level obtained when there are no cross-group links in place.

Second, we compute the more sophisticated (but less intuitive) *"spectral" segregation index* (SSI) for social networks recently proposed by Echenique and Fryer (2005). The index has two remarkable properties: (i) it disaggregates at the level of individuals and types (that is, one can compute the extent to which each single individual – or a given type – is segregated in the society); (ii) the level of segregation of any individual increases linearly with the level of segregation of his neighbors. Notice that the FSI does not possess this "linearity" feature. In fact, the FSI counts indiscriminately any within- and cross-group link among any two agents belonging to a cluster of connected agents of the same type. The SSI, on the other hand, takes into account the fact that agents located close to the fringe of such a cluster are less segregated that those near its center. More formally, the SSI associated to a given type $s \in \{-1, +1\}$ is defined as the largest eigenvalue of the sub-matrix obtained from the sociomatrix W_0 by considering only the rows and the columns associated to the nodes whose current type is s. In our analysis, we report the average SSI, computed over the two types $s \in \{-1, +1\}$ and suitably rescaled to have an index ranging in the unit interval (see the Appendix for details)[6].

5 Analysis of the Model

In this section, we present an analysis of our model for a society of $M = 100$ nodes. Initially, we set the average percentage of empty nodes $\theta = 0.3$, and for any given network class we choose the network-specific parameters. Section 5.2 presents a sensitivity analysis across the parameter space as described in Table 1. Our analysis will take the form of a Monte Carlo (MC) analysis. The procedure is as follows.

For each choice of network class and network-specific parameters we generate a number of independent runs. For each run, where necessary, we randomly select a specific instance of the network class, and we generate an initial allocation of agents and types across the network uniformly at random. We, then, let the best-response dynamics run, and collect system statistics when either segregation measures or the configuration of types across the M nodes have reached a steady-state[7]. This typically happens well before $T = 50000$ time-steps with probability one. We independently repeat this exercise 1000 times, computing the Monte Carlo (MC) average and standard deviation for the relevant measures. Since across-run variability turns out to be very small (across-run standard deviations are of an order of magnitude of 10^{-5}) and MC distributions appear to be

[6]We also check our results against a number of alternative segregation indices, such as those proposed in Fershtman (1997) and Freeman (1978), and some of those originally developed in the lattice-case (see Pancs and Vriend, 2005). In particular, we compute the *mix deviation* index (MD), defined as the absolute deviation between a 50-50 neighborhood and the current frequency of like agents in the neighborhood, averaged over all agents. Our results are not qualitatively altered if one considers these additional segregation measures.

[7]Notice that whenever the system does not seem to converge to a stable configuration of types across the M nodes, a cyclical behavior is typically observed with agents who keep switching among a small sets of locations. In these cases, segregation levels do converge to stable levels well before the system begins cycling.

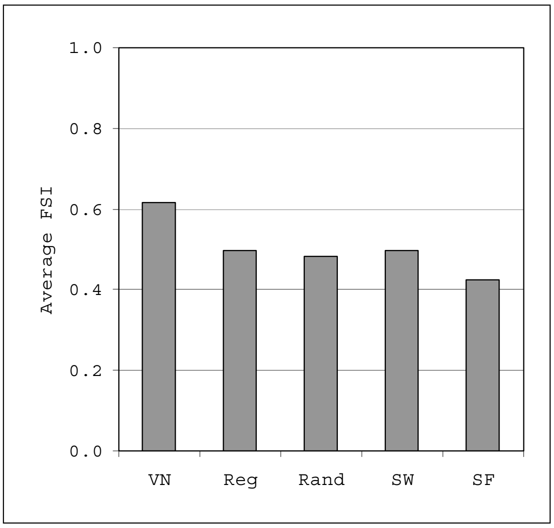

Figure 2: Average Freeman's Segregation Index (FSI) v. Network Classes. Average Degree $d = 4$. Parameters: $M = 100$, $\theta = 0.3$. MC Sample Size $= 1000$.

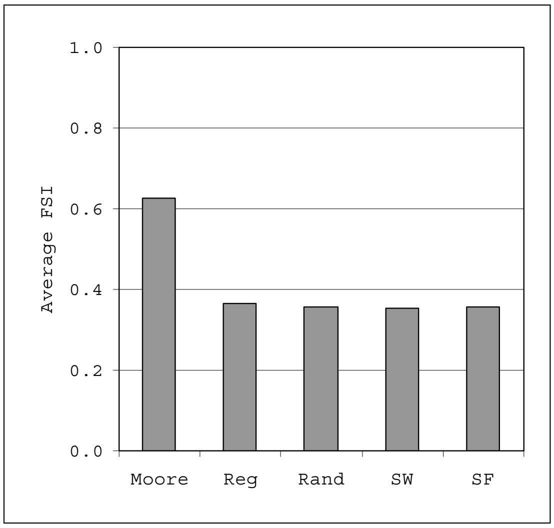

Figure 3: Average Freeman's Segregation Index (FSI) v. Network Classes. Average Degree $d = 8$. Parameters: $M = 100$, $\theta = 0.3$. MC Sample Size $= 1000$.

symmetric, we report below MC averages of segregation measures only.

5.1 Some Benchmark Results

The main question we are interested in here is whether mild proximity preferences can explain segregation also in more general network structures, e.g., when the underlying network is not necessarily a proxy of geographic space but it rather has structural properties that make it resemble more a social network. Secondly, we also want to know whether segregation levels emerging in non-lattice NDG vary across different families of networks, and if so, how.

To begin answering these questions, we compare MC averages of the FSI and SSI in the benchmark case $\theta = 0.3$ where agents are placed in one of our six different classes of NDGs (see Table 1). Initially, we restrict our attention to network-specific parameters implying NDGs with (average) degree $\overline{d} \in \{4, 8\}$. This allows us to compare non-lattice NDGs directly with either a 2D-VN lattice or a 2D-M lattice with $r = 1$ (and thus $d = 4$ or $d = 8$).

As Figures 2 - 5 show, segregation levels are rather similar in all network classes considered. Lattice-type networks seem to display a higher average FSI, but this effect is weaker for the SSI. This means that once one takes into account "linearity" in segregation

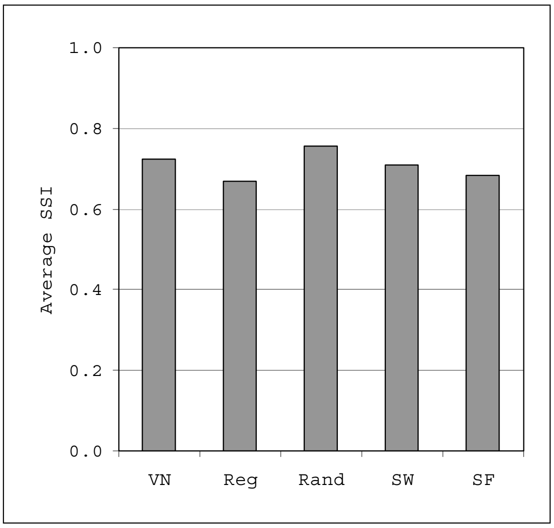

Figure 4: Average Spectral Segregation Index (SSI) v. Network Classes. Average Degree $d = 4$. Parameters: $M = 100$, $\theta = 0.3$. MC Sample Size $= 1000$.

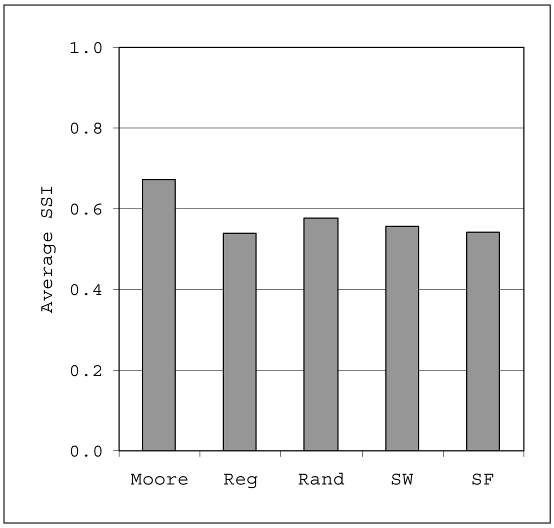

Figure 5: Average Spectral Segregation Index (SSI) v. Network Classes. Average Degree $d = 8$. Parameters: $M = 100$, $\theta = 0.3$. MC Sample Size $= 1000$.

(that is, an individual's segregation level increases with that of his peers), little difference in segregation can be detected across different network structures[8]. The structural properties of the network do not seem to engender a strong impact on segregation levels attained in the long- run by the system.

To put the values of the segregation indices found into perspective, we make the following two observations. First, for each network class and combination of network-specific parameters, we compare the values of the FSI and SSI obtained through best-response dynamics with the distribution of values for these indices over the set of all possible allocations of agents and types across $M = 100$ nodes (keeping $\theta = 0.3$ as above). We numerically generated proxies for these "theoretical" distributions by computing our segregation indices over 100,000 random allocations of agents and types for each given network class and network-specific parameters. Consider, for example, the FSI distributions (similar results hold also for SSI). The resulting simulated "theoretical" distribution of the FSI appears to be symmetric around 0. The corresponding MC distributions obtained by running our model lie clearly to the right of the simulated "theoretical" distributions (see

[8]According to standard statistical tests for the difference between two means, all reported values are not statistically different. Notice, however, that one can make any pair of average segregation levels statistically different by sufficiently increasing the MC sample size. Therefore, we do not report statistical tests for the difference between any two average indices.

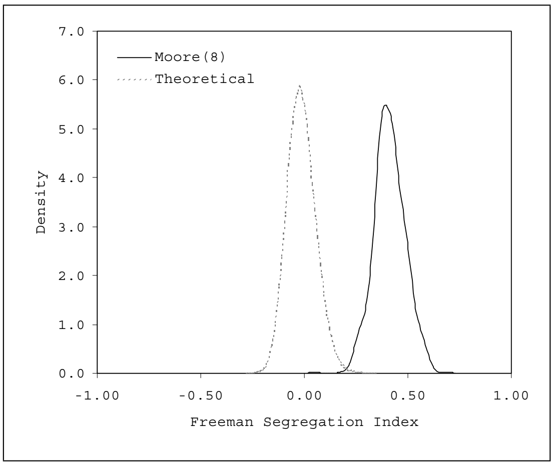

Figure 6: FSI levels generated by the model (solid line) v. FSI Simulated Theoretical Distribution (dashed line). Network: 2D-M lattice with $r = 1$ ($d = 8$). Parameters: $M = 100$, $\theta = 0.3$. Sample Sizes: Model = 1000; Simulated Index Distribution = 100,000.

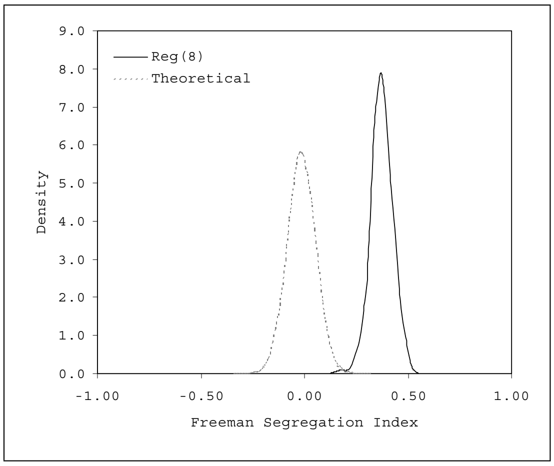

Figure 7: FSI levels generated by the model (solid line) v. FSI Simulated Theoretical Distribution (dashed line). Network: Regular lattice with $d = 8$. Parameters: $M = 100$, $\theta = 0.3$. Sample Sizes: Model = 1000; Simulated Index Distribution = 100,000.

Figures 6 and 7 for the case of the FSI in 2D-M and REG networks with degree $d = 8$)[9].

Second, the average "mix deviation" index reaches – in both lattice and non-lattice networks – levels that are similar to those obtained for lattices in Pancs and Vriend (2005). For example, Pancs and Vriend (2005) find an average mix deviation level of 0.19 for a 2D-M lattice (without boundaries) for $r = 1$ and $\theta = 0.2$. In our 2D-M setup, we find a similar mix-deviation level, i.e., 0.23. Furthermore, we find that for corresponding parameter values (i.e., $\theta = 0.2$ and $d = 8$), the mix deviation reaches 0.21 in REG and in RAND graphs, and drops to 0.18 in SF graphs. Finally, if we employ 2D-VN lattices or SW graphs (with $r = 1$ and thus $d = 4$) the MD respectively attains values of 0.27 and 0.22 respectively.

5.2 Sensitivity Analysis: Empty Spaces and Connectivity

The findings of the last section show that proximity preferences can explain segregation in a wide range of networks. In this section we turn to a sensitivity analysis of the model. We are interested in assessing how, for each given class of networks, segregation levels depend on the particular parametrization as far as the average percentage of the empty nodes (θ) and the connectivity of the network (measured by its average degree $\overline{d}$) are

[9] All densities have been estimated using a normal kernel with a 0.20 bandwidth.

concerned. Therefore, we consider a range of different values of θ and $\overline{d}$, and compute MC averages of the FSI and SSI over 1000 independent runs for each combination of parameter values. Figure 8 shows the SSI against average degree $\overline{d}$ for various fractions of empty nodes θ, with each panel concerning one of the network classes considered[10] [11].

— Figure 8 about here —

We see that segregation levels are decreasing with the average degree in every network class, and for any value of θ. Very high segregation levels are attained by the system when the society is poorly connected and there is a small percentage of empty nodes. As the connectivity increases, segregation becomes somewhat less pronounced, but even in very connected societies, segregation levels remain significantly higher than the expected level for random allocations.

— Figure 9 about here —

Figure 8 also shows that the SSI decreases with the percentage of empty nodes. This is, however, not true for the FSI. The FSI, unlike the SSI, declines for small values of θ and then remains fairly stable in each given network class. Figure 9 presents a comparison of FSI and SSI against θ for the case of small-world networks. This difference seems due to the linearity taken into account by the SSI. That is, only if one measures segregation by taking into account the structure of emerging clusters, does a higher percentage of empty nodes negatively affect the measured segregation levels. This means that the larger the degrees of freedom the agents have in their moving choices, the less the society ends up in "thick" segregated clusters, where many agents are "far" from the fringe. When θ is very large, the society tends to be segregated in less structured clusters, with many agents located on the "border" between clusters of agents of the same type and an empty space. That is why the FSI does not detect large drops of segregation levels even for large values of θ.

[10] For the scale-free networks, we employed values for $M_0 \in \{9, 16, 25, 36\}$, which implies approximate average node degrees $d \in \{5.82, 7.36, 8.50, 9.11\}$.

[11] Similar results are obtained also for the FSI.

5.3 Sensitivity Analysis: Inertia and Local Moves

In the basic model presented in Section 2, agents choose pure best- responses. That is, if an agent is drawn at time t to revise his current state, he constructs his choice set, which includes his current location plus all vacant locations, and choose the best option in this set, randomizing in case the best-response is not unique. Although our analysis shows that mild proximity preferences alone are sufficient to explain segregation in a wide range of general network structures, it may be interesting to analyze whether additional constraints on the individual behavior lead to different segregation levels.

Therefore, we first examine the effect of introducing *inertia* in individual decisions. Suppose that at time $t > 0$ the s-type agent located in node i is drawn at random from $I_N = \{1, ..., N\}$. With inertia, this agent stays put if there is no vacant location that he would strictly prefer to his current location. The idea of inertia is based on the implicit modelling assumption of some small costs of moving (smaller than the smallest possible difference in satisfaction between any two locations, but otherwise arbitrarily small). Notice that under the *inertia* rule, satisfied agents will never move.

Second, we restrict the choices of the individual agents to *local-moves-only.* With local-moves-only the agents' moves are restricted to the agents' direct neighborhoods only. Suppose the agent drawn, say the one located in node k, considers to move. He checks every node that is currently empty *in his interaction group* V_k *only*, and then behaves as in the basic model as far as his decision is concerned. This can be based either on the additional assumption of moving costs that increase with the distance travelled in the network such that they are greater than the greatest possible improvement in utility derived from the neighborhood for any move beyond the current neighborhood, or on some information costs preventing agents to observe anything that is outside their current neighborhood. Notice that in both cases, the agents will tend to explore a smaller number of options and the ensuing dynamics will be more 'sticky'.

— Figure 10 about here —

Figure 10 presents the average degree in four network classes (with $\theta = 0.3$) for three cases: (i) the basic model without inertia and allowing global moves, (ii) the model with *inertia* and global moves, and (iii) the model with *inertia* and *local-moves-only*[12]. As expected, segregation levels decrease when one introduces subsequently the assumptions of *inertia* and *local-moves-only*, especially in lattice-type of networks. In all other cases, the effect is relatively small.

6 Conclusions

Considering six different classes of networks, we generalized the spatial proximity model due to Schelling (1969, 1971a,b, 1978). For each network structure, we considered two types of agents who occupy nodes in the network, with some nodes being empty. The agents have preferences about the composition of their own neighborhood in the network, and move to available vacant locations following myopic best-responses. We analyzed the ensuing dynamics, performing also a sensitivity analysis of a range of parameter values and setups. The main result of our analysis is an affirmative answer to the question whether mild proximity preferences as such may suffice to explain segregation in a wide range of network structures. In other words, our analysis confirms mild proximity preferences as an important possible explanation of segregation not only in regular spatial networks, but also in more general social networks.

An example of social interactions fitting this kind of model is a network of professional relationships, i.e., interactions with colleagues, suppliers, customers, etc. Typically an agent searching a job needs to find a vacancy (empty node), and when starting a new job he 'inherits' the links that come with his new position, while severing ties related to his previous job. A similar pattern applies to positions in sports clubs or, say, orchestras.

Obviously, not all social interactions are adequately described by a *given* network structure. Therefore, the next step of our research will be to study models of segregation based on proximity preferences in which the network structure may evolve endogenously, with individual agents forming new ties or severing existing ones according to their proximity preferences in response to their environment.

[12]These results do not qualitatively vary if one considers the SSI.

References

Baldwin, T. and G. Rozenberg (2004), "Britain 'must scrap multiculturalism"', The Times, April 2003.

Barabási, A.-L. (2003), *Linked.* Cambridge: Plume Books.

Barabási, A.-L. and R. Albert (1999), "Emergence of Scaling in Random Networks", *Science*, 286: 509–512.

Commissie_Blok (2004), "28 689 Eindrapport Onderzoek Integratiebeleid", Den Haag: Sdu.

Echenique, F. and R. Fryer (2005), "On the Measurement of Segregation", Working Paper No. 11258, NBER.

Fershtman, M. (1997), "Cohesive Group Detection in a Social Network by the Segregation Matrix Index", *Social Networks*, 19: 193–207.

Freeman, L. (1972), "Segregation in social networks", *Sociological Methods and Research*, 6: 411–427.

Freeman, L. (1978), "On Measuring Systematic Integration", *Connections*, 2: 9–12.

Mitchell, J. (1978), "On Freeman's Segregation Index: An Alternative", *Connections*, 2: 9–12.

Pancs, R. and N. Vriend (2005), "Schelling's Spatial Proximity Model of Segregation Revisited", mimeo.

Phillips, T. (2005), "After 7/7: Sleepwalking to Segregation", Speech given at the Manchester Council for Community Relations, 22-09-2005.

Schelling, T. (1969), "Models of Segregation", *American Economic Review*, 59: 488–493.

Schelling, T. (1971a), "Dynamic Models of Segregation", *Journal of Mathematical Sociology*, 1: 143–186.

Schelling, T. (1971b), "On the Ecology of Micromotives", *The Public Interest*, 25: 61–98.

Schelling, T. (1978), *Micromotives and Macrobehavior.* New York: W.W. Norton and Company.

Steger, A. and N. Wormald (1999), "Generating random regular graphs quickly", *Combinatorics, Probability and Computing*, 8: 377–396.

Watts, D. (1999), *Small Worlds.* Princeton: Princeton University Press.

Watts, D. (2003), *Six Degrees: The Science of a Connected Age.* New York: W.W. Norton and Company.

Watts, D. and S. Strogatz (1998), "Collective Dynamics of 'Small-World' Networks", *Nature*, 393: 440–442.

Appendix

A Segregation Indices

In this appendix we formally define the segregation indices that we employ in our analysis.

- **Freeman Segregation Index (FSI)**

  Consider a NDG G over M nodes described by the sociomatrix W. Nodes can be of three types: $+1$, -1 or 0 (with 0 meaning that the node is empty). Let N_A and N_B the number of agents of types $A = +1$ and $B = -1$. Let us define the 2x2 *mixing* matrix P where entry p_{ab} with $a, b \in \{A, B\}$ counts the number of links connecting an a-type node with a b-type node. It is easy to see that:

  $$P = E'AE,$$

  where A is the $N \times N$ socio-matrix obtained from W by deleting rows/columns associated to empty nodes, and E is the $N \cdot 2$ matrix (*state indicator* matrix) where columns refer to the type $\{A, B\}$ and rows to the node. Any row i can either be $(1, 0)$ if node i is of type A or $(0, 1)$ if node i is of type B. Given the *mixing* matrix of cross-group ties P, define the expected matrix of cross-group ties by $E(P)$ using the standard expected contingency matrix. Operationally, let

  $$P = \begin{bmatrix} p_{AA} & p_{AB} \\ p_{BA} & p_{BB} \end{bmatrix}$$

  be the *mixing* matrix. Freeman (1972) asked how we could identify segregation in a social network. Theoretically, he argues, if a given attribute (group label) does not matter for social relations, then relations should be distributed randomly with respect to the attribute. Thus, the difference between the number of cross-group ties expected by chance and the number of observed ties (divided by expected ones) measures segregation:

  $$FSI = \frac{[E(p_{AB}) + E(p_{BA})] - [p_{AB} + p_{BA}]}{E(p_{AB}) + E(p_{BA})}.$$

To compute expected values, consider the row and column sums:

$$\begin{aligned}
p_{A\cdot} &= p_{AA} + p_{AB} \\
p_{B\cdot} &= p_{BA} + p_{BB} \\
p_{\cdot A} &= p_{AA} + p_{BA} \\
p_{\cdot B} &= p_{AB} + p_{BB} \\
p_{\cdot\cdot} &= p_{AA} + p_{AB} + p_{BA} + p_{BB}
\end{aligned}$$

and let:

$$\begin{aligned}
E(p_{AB}) &= \frac{p_{A\cdot} \times p_{\cdot B}}{p_{\cdot\cdot}} \\
E(p_{BA}) &= \frac{p_{B\cdot} \times p_{\cdot A}}{p_{\cdot\cdot}}
\end{aligned}$$

Notice that in principle Freeman's index should be computed as:

$$FSI^* = \max\{FSI, 0\}$$

because it measures whether there are less cross-group links than expected. As Mitchell (1978) argues, there is no way to compute systematic integration. Freeman (1978) then suggest to compute two different indices, one for segregation and one for integration, namely:

$$FSI^{**} = \frac{[E(p_{AB}) + E(p_{BA})] - [p_{AB} + p_{BA}]}{E(p_{AB}) + E(p_{BA}) - \min\{E(p_{AB}) + E(p_{BA})\}}$$

Since $\min\{E(p_{AB}) + E(p_{BA})\} = 0$, this means that when measuring systematic segregation we have $FSI^{**} = FSI$. Similarly, the index to measure systematic integration is:

$$FII = \frac{[E(p_{AB}) + E(p_{BA})] - [p_{AB} + p_{BA}]}{E(p_{AB}) + E(p_{BA}) - \max\{E(p_{AB}) + E(p_{BA})\}},$$

where $\max\{E(p_{AB}) + E(p_{BA})\} = N_A \cdot N_B$. We compute FSI and FII for agents of types $A = -1$ and $B = +1$. A larger value for any of the two indices means "more" of what it measures.

- **Spectral Segregation Index (SSI)**

  The Spectral segregation index is computed as follows (see Echenique and Fryer, 2005, for additional details and applications). Consider the $M \times M$ socio-matrix W and the current configuration of types across the M nodes. Define as A_x the symmetric, square sub-matrix obtained from W by considering only the rows/columns associated to agents of type $x \in \{-1,+1\}$. If there are $K = N/2$ agents of each type (assuming N even), both A_{-1} and A_{+1} are $K \times K$ symmetric matrices. For each type $x \in \{-1,+1\}$, define the Spectral segregation index with SSI(x) as the largest eigenvalue of A_x. The entries of the eigenvector associated to SSI(x), suitably scaled so that the average of the entries are equal to the eigenvalue SSI(x), give the segregation levels of the individuals belonging to group x. Notice that:

$$d_{\min}(x) \leq \overline{d}(x) \leq SSI(x) \leq d_{\max}(x),$$

where $d_{\min}(x)$, $\overline{d}(x)$ and $d_{\max}(x)$ are, respectively, the minimum, average and maximum degree of the nodes associated to agents of type x in the sub-graph W composed of nodes hosting an agent of type x. Therefore, to have an index ranging in the unit interval, one can employ:

$$SSI^*(x) = \frac{SSI(x) - d_{\min}(x)}{d_{\max}(x) - d_{\min}(x)}.$$

In our results, we typically get that $SSI^*(-1)$ and $SSI^*(+1)$ are very close in each run and parametrization. This is because of symmetry between types in the initial random allocation, with equal numbers of agents of each type. Therefore, we report the average SSI index:

$$SSI = \frac{SSI^*(-1) + SSI^*(+1)}{2}$$

as our spectral measure of segregation for the society as a whole.

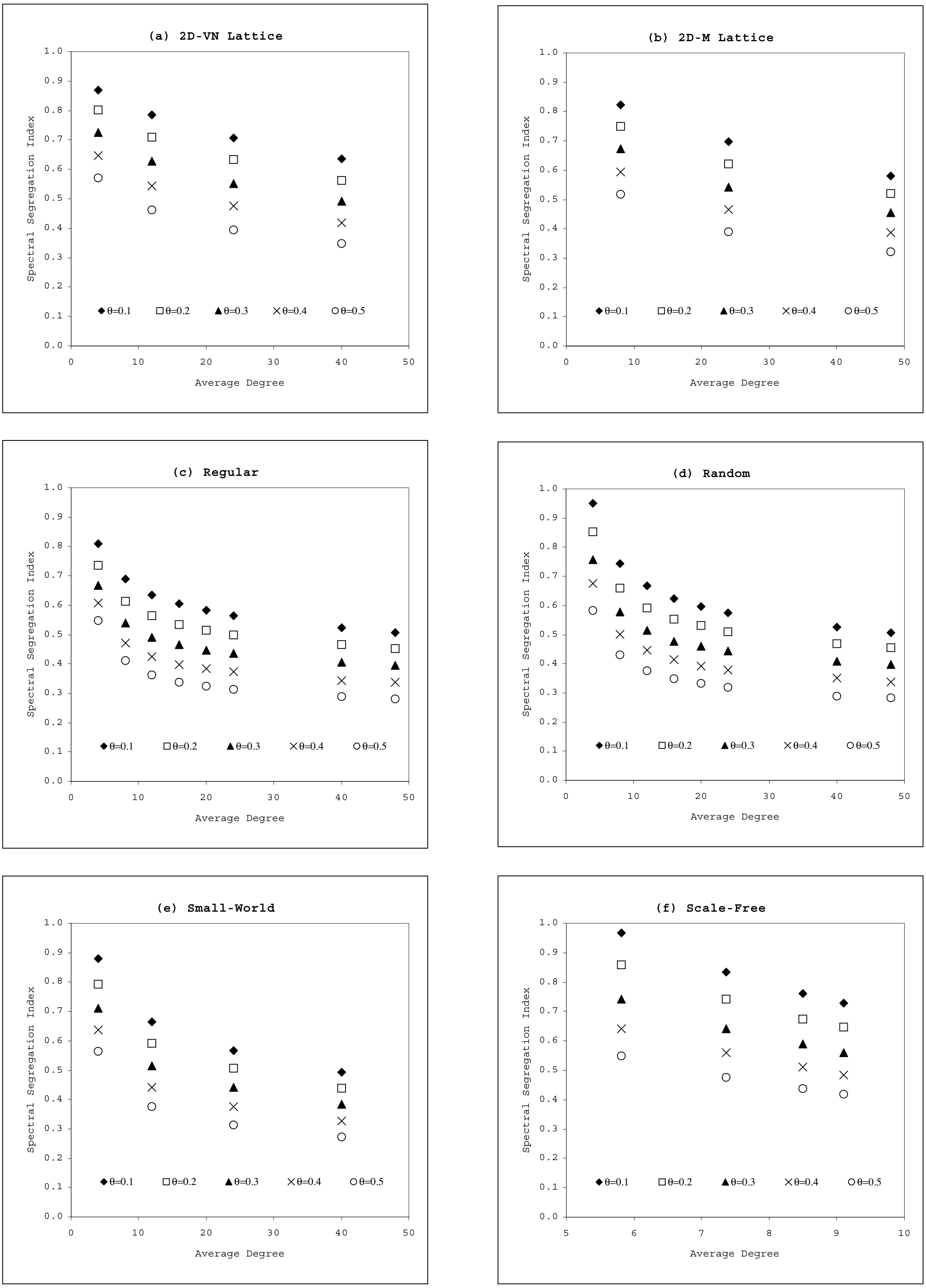

Figure 8: SSI v. Average Degree for Different Levels of θ. Parameters: $M = 100$, MC Sample Size: 1000

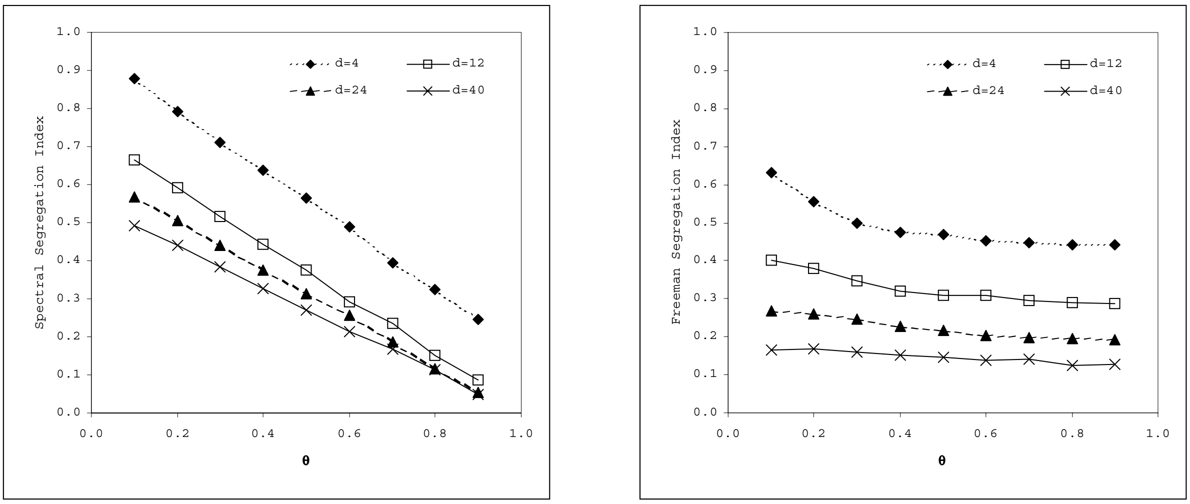

Figure 9: Small-World Networks. SSI (left) and FSI (right) v. θ for Different Average Degrees. Parameters: $M = 100$, MC Sample Size: 1000

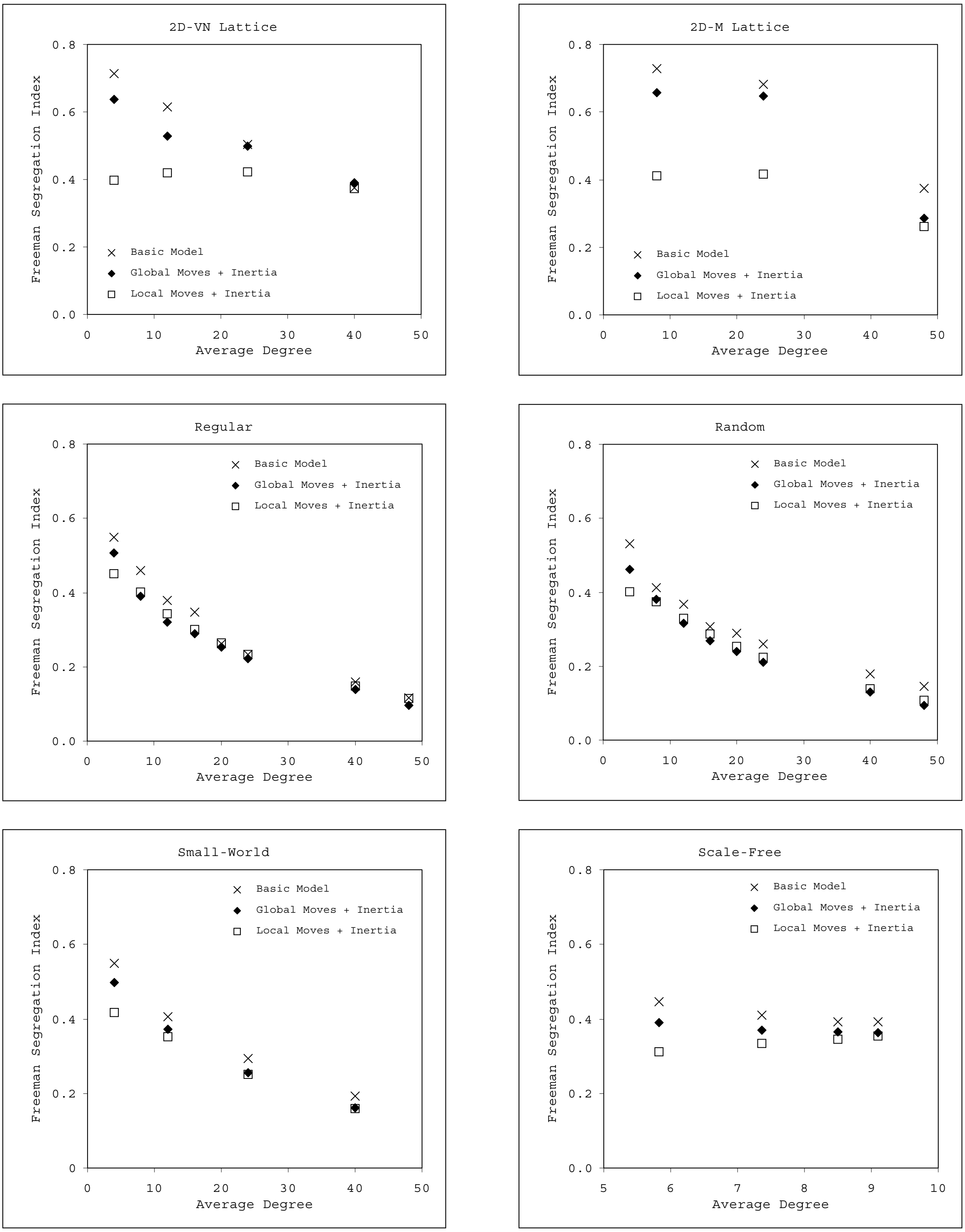

Figure 10: FSI v. Average Degree in three alternative setups: (i) Basic Model; (ii) Inertia in individual decision rules; (iii) Inertia in individual decision rules and local moves. See Section 5.3 for details. Parameters: $M = 100$, $\theta = 0.3$, MC Sample Size: 1000