Environment and Planning B: Planning and Design 2009, volume 36, pages 149 – 169

doi:10.1068/b33042

Effects of city size, shape, and form, and neighborhood size and shape in agent-based models of residential segregation: are Schelling-style preference effects robust?

Mark Fossett
Department of Sociology, Texas A&M University, 311 Academic Building, College Station, TX 77843-4351, USA;
e-mail: m-fossett@tamu.edu

David R Dietrich
Department of Sociology, Duke University, Box 90088, Building 9, Durham, NC 27708, USA;
e-mail: david.dietrich@duke.edu
Received 7 April 2006; in revised form 6 August 2007; published online 15 December 2008

Abstract. This paper investigates the effects of city size, shape, and form, and neighborhood size and shape in agent-based models of residential segregation. We find that, in many key respects, model-generated segregation outcomes are not influenced in important ways by variation in these factors. For example, the expression of segregation based on agent preferences for coethnic contact in agent-based models does not vary with city size, city shape, city form, or the shape of neighborhoods involved in agent vision, or the use of distance-decay functions for evaluating neighbors. These findings indicate that results obtained from model-based segregation studies are likely to be robust relative to choices regarding these aspects of model specification. We do find important effects of the size or scale of neighborhoods involved in agent vision: model-generated segregation outcomes vary in complex ways with agent vision. Significantly, however, the effect of neighborhood size or scale does not appear to vary in important ways with neighborhood shape. Thus, what is important is the scale of agent vision—that is, the number of neighbors they 'see'—not the particular spatial arrangement of those neighbors. With the exception of the effect of the spatial scale of agent vision, our results suggest that researchers can generally presume that their findings regarding how model-generated segregation outcomes vary with substantive factors, such as agent preferences for coethnic contact or the ethnic demography of the city, are not contingent on choices regarding model implementation of city size, city shape, city form, and neighborhood shape. These findings are welcome because they suggest that simulation studies can devote less attention to technical specification choices and more attention to assessing substantive questions regarding the effects of social dynamics and sociodemographic distributions in the context of model systems.

Introduction

In 1971 Thomas Schelling outlined a model in which group segregation emerged out of the residential decisions of individual agents who were guided by relatively weak preferences for same-group contact, preferences that on first consideration appeared to be compatible with integration (Schelling, 1971). This agent-based model has proven to be an important tool in the study of segregation dynamics and the role of individual preferences in producing patterns of segregation. Rigorous formal studies (Young, 1998; Zhang, 2004a) have examined the underpinnings of important theoretical insights obtained from the model. An even larger number of empirical studies have explored the role of selected model parameters. Schelling initiated this line of research by examining several variations on his initial model. Subsequent work by others has explored different aspects of the model in more detail. For example, later studies have focused on varying implementations of agent vision (Laurie and Jaggi, 2003), refining decision rules that agents use when choosing their residential location on the basis of their neighbors (Pancs and Vriend, 2003; Sander et al, 2000), introducing competing residential preferences (Wasserman and Yohe, 2001), using unbounded torus landscapes

(Laurie and Jaggi, 2003), and considering different methods by which agents move (Meen and Meen, 2003; Young, 2001; Zhang, 2004b).

The key thing that these and related modeling efforts have in common is that they take Schelling's theoretical insights and the basic structure of his simulation model as their initial point of departure. Beyond that, however, model implementations can be highly variable. Indeed, there appears to be little consensus on appropriate choices for fundamental model elements such as city shape, city size, city form, neighborhood form, neighborhood size, neighborhood shape, and other basic elements of the modeling framework. Are these choices important? Proponents of agent modeling efforts often speculate that decisions regarding these aspects of model specification may be consequential. While their conjectures are usually offered in the spirit of refining the model, they may have an unwelcome impact: they may provide a basis for those who are skeptical of effects discovered using agent models to discount the findings of these studies. Critics of Schelling-style preference effects (eg Goering, 2006; Yinger, 1995) often speculate that choices regarding various aspects of model specification may be consequential. But they are not motivated by hopes of refining Schelling-style agent models. Instead, they are concerned that core substantive implications of model-generated results may depend in crucial ways on particular specification choices.

In fairness to critics, specification choices in agent models can, to those who are unfamiliar with such models, sometimes appear to be abstract, highly stylized, and of questionable relevance for understanding real-world social dynamics. The question arises, are Schelling-style preference effects in agent-based models contingent on particular choices of model specification? Or are they robust across a wide range of alternative choices? Proponents of the Schelling-style agent models would be interested in the answer because it would speak to the issue of which topics should be the focus of future research. Critics of the Schelling model would be interested in the answer because it would speak to their concern that Schelling-style preference effects are fragile and have limited relevance outside of simple, game-like models. To date, views on these matters are speculative, as there has been little systematic empirical analysis to assess the robustness of Schelling-style effects relative to basic model specification choices. This paper addresses this gap in the literature by examining how city size, shape, and form, and neighborhood size and shape affect Schelling-style preference effects in agent-based models of ethnic residential segregation.

Background

Models inspired by Schelling's study have implemented different specifications for the size, shape, and form of the city landscape, and the size and shape of the neighborhoods which agents 'see' when they evaluate residential satisfaction (Epstein and Axtell, 1996; Laurie and Jaggi, 2003; Pancs and Vriend, 2003; Zhang, 2004b). Often these modeling choices are justified on the basis of untested assumptions that city shape, size, and form, and neighborhood size and shape are likely to have important effects on model-generated segregation outcomes. Here we review some of the more common options considered in the literature.

Agent-based segregation models specify neighborhoods for at least two purposes. One purpose is to delimit agent 'vision', where vision refers to the neighborhood and neighbors that agents 'see' when they evaluate any particular residential location. The other purpose is to delimit neighborhoods for the purpose of measuring segregation outcomes (eg computing city-wide segregation indices). The specifications used for these two purposes need not be, and often are not, the same. This paper focuses attention as neighborhood specifications relating to agent vision. Fossett (2005a) gives attention to neighborhood specifications relating to the measurement of segregation outcomes.

We draw on the strategies for measuring segregation in computational models outlined in that paper and discuss them in more detail later in this paper.

Neighborhood specifications of agent vision can vary in at least three dimensions: shape, size or spatial scale, and boundedness. Bounded neighborhoods are nonoverlapping areas delimited by fixed spatial boundaries. Under this specification of vision, neighbors are defined on the basis of whether or not households are inside a particular bounded region. Households can be close together in space, even adjacent, but not be treated as neighbors, if they are on opposite sides of an area boundary. Bounded areas are usually implemented as square subsections of a lattice with their scale being determined by the width of the squares.[(1)] Not surprisingly, agent vision that follows bounded areas tends to produce segregation patterns that closely follow area boundaries. If bounded areas are considered in complete isolation (eg without concern for adjacent bounded areas), they can produce 'checkering', a segregation pattern in which individual bounded areas are ethnically homogeneous but the ethnic composition of adjacent bounded areas is uncorrelated.

Alternatively, neighborhoods may be specified as overlapping, site-centered areas. In this case neighborhoods are delimited in terms of local regions centered on a focal residential location with every residential location thus having its own unique neighborhood. Specifications of site-centered neighborhoods can vary considerably in shape and spatial scale. The most common shapes include squares, diamonds, and circles. Two special-case, small-scale neighborhoods have received particular attention: the Von Neumann neighborhood, a four-unit, diamond-shaped neighborhood also known as a 'Rook's neighborhood' (Laurie and Jaggi, 2003; Zhang, 2004b), and the Moore neighborhood, an eight-unit, square neighborhood also known as a 'Queen's' neighborhood. A simple way to compare site-centered areas on 'scale' or size is by the count of housing units in a single neighborhood. This is a nonlinear function of some aspect of the neighborhood form: for diamonds, the distance from center to vertex; for circles, the radius; for squares, the width; and so on. Neighborhoods can be specified as desired to manipulate agent 'vision' (eg Laurie and Jaggi, 2003) from smaller to larger spatial scales.

Schelling (1971) considered segregation dynamics on the basis both of bounded area specifications of agent vision and of site-centered specifications of agent vision, including small-scale areas (ie Queen's neighborhoods) and intermediate-scale areas (eg 5×5 square neighborhoods). In subsequent agent-based studies, site-centered specifications of agent vision have been used more widely than bounded area specifications. We focus on site-centered specifications in the present study because they afford greater opportunities to vary shape and scale.

City landscapes vary in size, shape, and form. The city shape used most often is a square (eg Epstein and Axtell, 1996; Laurie and Jaggi, 2003; Pancs and Vriend, 2003; Schelling, 1971) but a circular shape has also been used (eg Fossett, 2006; Fossett and Waren, 2005). Studies using the square shape have treated it as bounded in form (eg Schelling, 1971) and in the form of an unbounded torus, presented visually as a square city in which the edges are construed as 'wrapping' around and connecting to each other (Epstein and Axtell, 1996; Laurie and Jaggi, 2003; Pancs and Vriend, 2003). City size can be measured by the number of housing units in the city. In most agent-based studies, city size is small (eg a 100×100 square lattice), especially in comparison to a real metropolitan area. Modern computers make it feasible to implement a large city size if desired, but smaller sizes are the norm because they facilitate visual inspection

(1) Other geometric forms are rarely used with fixed boundaries. For an interesting exception, see Macy and van de Rijt's (2006) use of irregular-shaped bounded areas.

of segregation patterns based on graphical displays of the city landscape. Additionally, while it is feasible to implement large cities, computational burden increases with large size and this negative consideration must be weighed against whatever benefits are thought to accrue from using large cities. So far, no important substantive benefits have been identified for using large city size, but the matter has yet to be systematically explored.

Agent-based modelers and their critics have speculated about a variety of possible effects of city size, shape, and form, and neighborhood size and shape. Critics have raised concerns that findings from agent-based studies may be peculiar to the small size of the cities implemented in most models. On this basis they question the relevance of model-generated segregation behavior for understanding 'real world' segregation dynamics. Laurie and Jaggi (2003) argued for the use of torus specifications of city form to avoid complications of possible 'edge effects', where locations on the edge of a bounded city have fewer neighbors than other locations. Fossett and Waren (2005) used a nontorus, circular form arguing that, if 'edge effects' and city shape effects do exist, the bounded circular form is a better approximation of the form of real cities. Laurie and Jaggi (2003) identified important effects of agent vision—in particular, how the effects of preferences for coethnic contact are contingent on whether vision is restricted or expansive. Fossett and Waren (2005) documented how this interaction (of agent vision and preferences) is itself conditioned by the ethnic demography of the population. Finally, many have speculated that neighborhood shape (diamond, circle, square, etc) may have effects over and above neighborhood scale. For example, the studies by Laurie and Jaggi (2003) and Fossett and Waren (2005) document complex effects using a diamond-shaped specification of neighborhood vision. Are these effects likely to hold up when vision is specified using other shapes?

For the most part, notions about the effects of city size, shape, and form, and neighborhood size and shape remain at the level of conjecture. A few studies have explored the impact of varying one of these parameters within the context of a particular study (eg Laurie and Jaggi, 2003; Zhang, 2004b). Ours is the first study to undertake a systematic investigation of the effects of city size, shape, and form, and neighborhood size and form on segregation outcomes in the basic Schelling model.

Research design

In this study we examine the effect of city size, shape, and form, and neighborhood size and shape using SimSeg—a computational model that implements the essential features of the Schelling model along with numerous extensions.[2] Key model components for the present study include: housing units arranged on a lattice or city 'landscape'; a population of agents consisting of two groups arbitrarily designated as blacks and whites; a percentage of vacancies; agent preferences for coethnic contact; and a utility function that translates the percentage of like neighbors into a residential satisfaction score. Using this model we carried out experiments in which we varied city size and form and neighborhood size and form to assess their effects on segregation outcomes. We also varied a key parameter—city ethnic mix—which is known to impact model-generated segregation outcomes in order to determine whether Schelling-style segregation effects were in some way contingent on settings for city size, shape, and form, and neighborhood size and form.

The residential dynamics in the model were implemented as follows. In each simulation the SimSeg program created a city according to the settings for the model's parameters. All simulations begin with a randomly generated residential distribution.

[2] The full SimSeg model extends the basic Schelling model in many ways. Here we set the parameters of the SimSeg model to restrict it to the simpler form of 'pure' Schelling models.

Each simulation runs for thirty cycles, where cycles are periods of activity during which a random selection of 25% of households are given the opportunity to move. Each searching household assesses its satisfaction with its current residential location and compares this against its potential satisfaction with a dozen randomly selected prospective residential locations sampled from the available supply of vacant housing units.

In all cases the searching household moves if it is able to improve its residential satisfaction by doing so.(3) Otherwise, it remains in its original location subject to two conditions which serve to insure the full expression of preference effects within a given simulation experiment. The first is that all households are required to move on the first search. This insures that all households will reside in a location which they identify through search and choice. The second condition is that, in 20% of the searches, determined randomly, the searching household is required to move to the least objectionable residential alternative it identified through the search. In these cases the household must move even if it would prefer to stay in its current location. Substantively, this dynamic mimics fundamental demographic processes (eg household life-cycle transitions and regular population entries and exits) which produce continuous population turnover and random vacancies in all neighborhoods throughout the city. Technically, this dynamic protects against misleading equilibrium outcomes that can occur under some alternative search algorithms.(4)

We assessed model-generated segregation patterns using two measures of segregation. One is the index of dissimilarity (D), long a staple of research on residential segregation. The other is the pairwise contact difference, a measure developed by Fossett (2007) for the purpose of measuring segregation at small spatial scales.(5) We computed these measures using two different approaches. The first was a conventional implementation in which we calculated segregation scores using neighborhood data based on a neighborhood grid with fixed boundaries and equal-sized neighborhoods. The second approach was a site-centered measurement approach outlined by Fossett (2005b; 2007), which registers residential segregation using unique, local neighborhoods defined separately for each individual housing unit. As we note below, measures computed using these two approaches tracked each other very closely, so we present only the results for the more conventional approach of computing measures using data for bounded areas.(6)

(3) The evaluation function is simple. If a neighborhood meets or exceeds an agent's preference for coethnic contact, the agent is satisfied. Otherwise the agent is dissatisfied and the level of dissatisfaction increases as a simple linear function of the size of the discrepancy.

(4) Fossett and Waren (2005) review the importance of this dynamic for Schelling-style preference models. They point out that Schelling models must incorporate an appropriate mechanism that permits preference effects to be expressed, otherwise they can produce pathological equilibrium outcomes and misleading conclusions about the stability of residential patterns. Two possible mechanisms include the dynamic of 'compulsory' moves (which we use here and which might be more aptly labeled the population 'life-cycle' dynamic) and the use of very high vacancy rates (used by Schelling).

(5) Fossett (2005a) reviews measures of segregation used in agent-based models. Fossett (2007) provides a detailed introduction to the separation index, a measure that is similar to the variance ratio, but corrects for technical deficiencies of that index. The separation index has multiple interpretations. Here we emphasize its interpretation as an index of group difference in residential contact. This substantive interpretation is simple and appealing: the index registers the difference between the contact a focal group has with itself and the contact the comparison group has with the focal group.

(6) The computing formula for D for bounded areas is $100\sum|q_i - Q|/2PQ$, where i is an index for bounded areas, q_i is the proportion black in the area, Q is the black proportion in the city, and P is $1 - Q$. The computing formula for the contact difference is $P_{\mathrm{BB}} - P_{\mathrm{WB}}$, where P_{BB} is the average black contact with black neighbors given by $100\sum(b_i/B)(b_i - 1)/(t_i - 1)$ and P_{WB} is the average white contact with black neighbors given by $100\sum(w_i/W)b_i/(t_i - 1)$, where t, b, and w refer to counts of total, black, and white households, respectively, for areas and B and W refer to counts of black and white households, respectively, for the city.

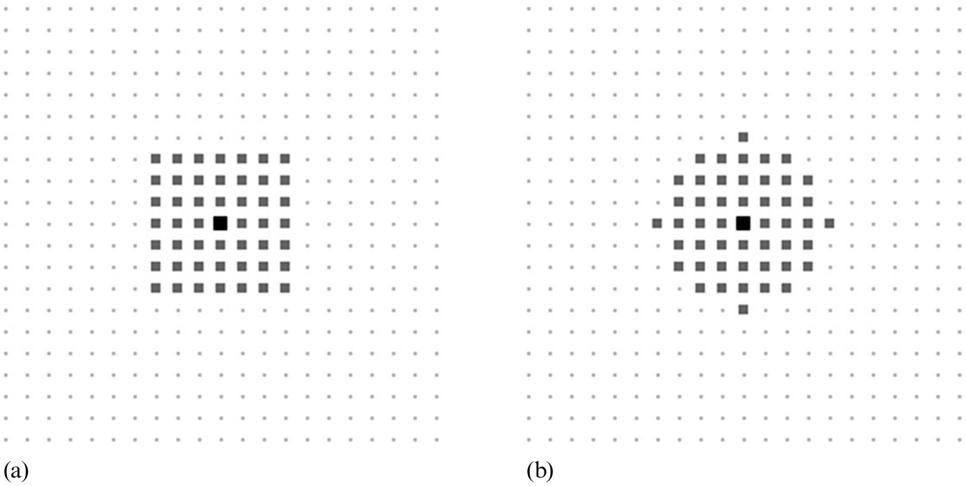


Figure 1. Spatial domains used in computing conventional and site-centered scores for segregation measures: (a) forty-eight-unit nonoverlapping grid section; (b) forty-eight-unit overlapping circular area.

All else being equal, it is easier to detect segregation, and obtain higher values of segregation scores, when it is measured using smaller areal units. Consequently, we held the spatial scale of the neighborhoods used in computing segregation scores constant across all simulations. Fixed boundary neighborhoods were held constant at forty-nine total units, based on nonoverlapping 7×7 sections of the city landscape's housing lattice. Site-centered neighborhoods were specified in terms of a circular region also consisting of forty-nine total units, but centered on a particular housing unit. The form of these neighborhoods is depicted in figure 1. Taking account of the focal unit, the scale of the neighborhoods is forty-eight units.

Note that the size and shape of areal units used in computing segregation scores are specified separately from those of neighborhoods involved in agent vision. The latter vary in our simulations. The former are fixed across all simulations.

Variations in city size, city form, neighborhood size, and neighborhood form

We implemented three variations in city form: a circular form treated as having fixed outer boundaries, a square form treated as having fixed outer boundaries, and a square form treated as an unbounded torus. Examples of these three forms are depicted in figure 2. We implemented two specifications for city size—'small' and 'large'. The two sizes are depicted in figure 3 for the circular city form.

When city size is set to 'small' the city landscape is comprised of nonoverlapping, 7×7 square sections of housing units organized within an 11×11 square area grid. In the circular city form (shown in figure 3), the city has 97 bounded areas and 4753 housing units. In the square and torus city forms, the cities have 121 bounded areas and 5929 housing units. When city size is set to 'large', the city landscape is comprised of nonoverlapping, 7×7 square sections of housing units organized within a 21×21 square area grid. In the circular form, the city has 349 bounded areas and 17 101 housing units. In the square and torus city forms, the cities have 441 areas and 21 609 housing units.

In our study the smaller city is in the intermediate range, compared with city size specifications used in other simulation studies. The larger city is more than three times the size of the smaller city and is substantially larger than cities used in other simulation studies; it thus provides a reasonable basis for assessing the effects of city size.

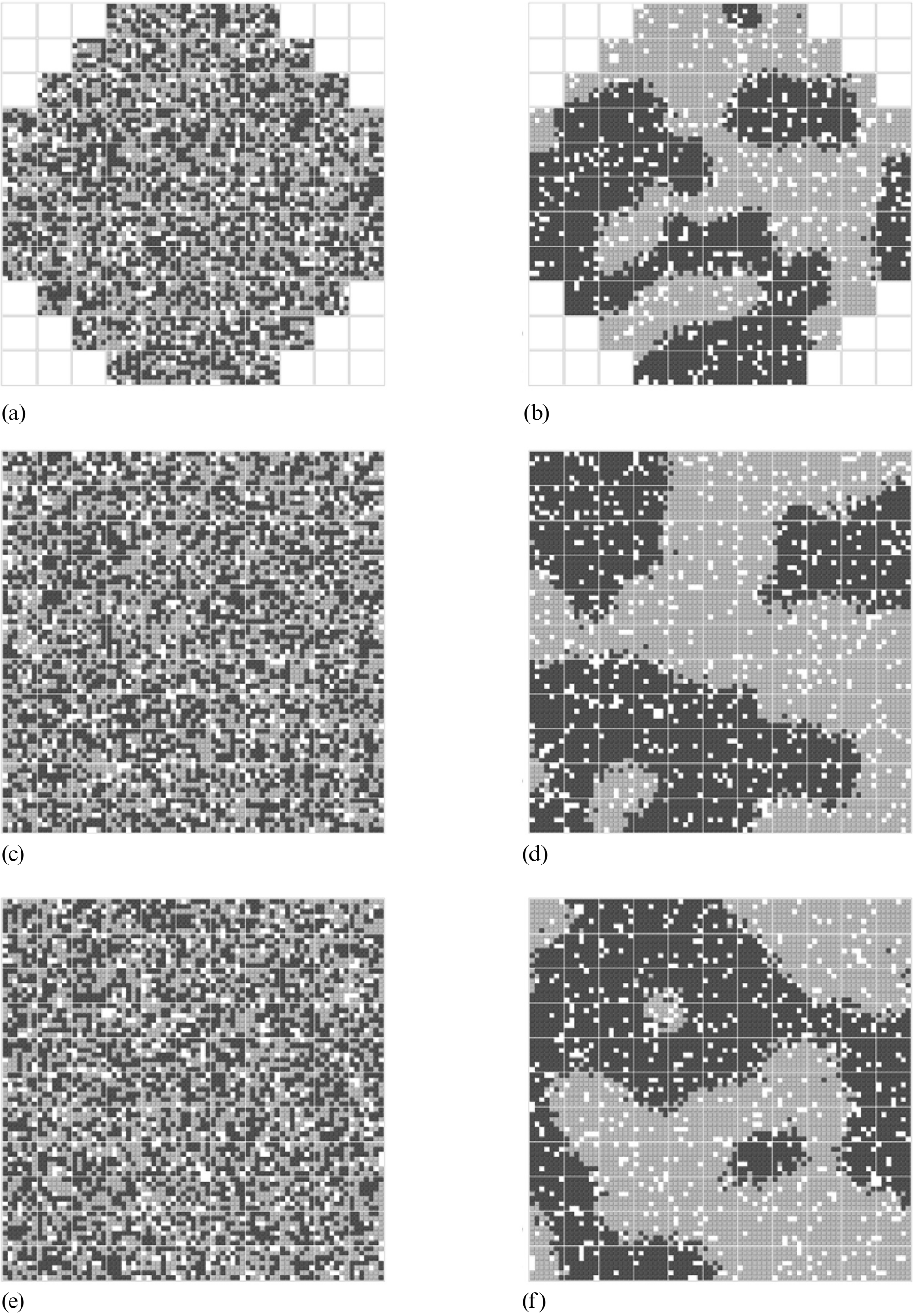
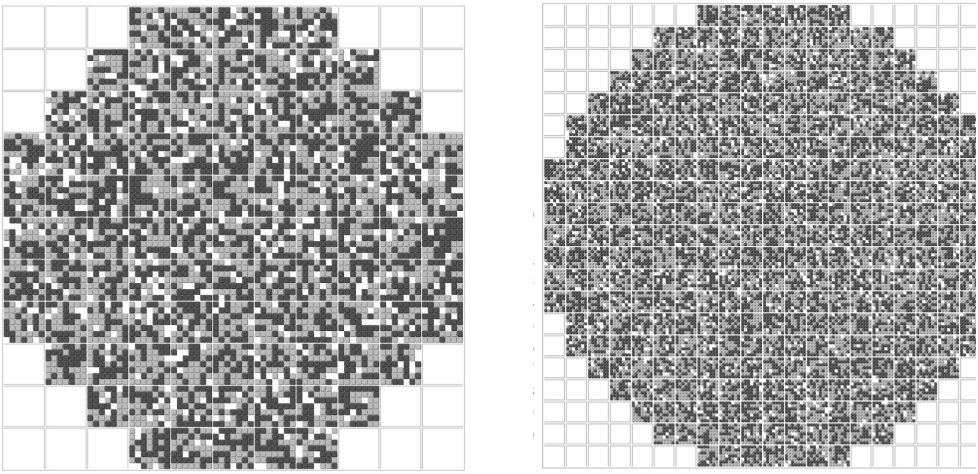


Figure 2. Examples of city form with random initial residential distribution and residential distribution at completion of a simulation experiment (a) circular city at initialization; (b) circular city at completion; (c) square city at initialization; (d) square city at completion; (e) torus city at initialization; (f) torus city at completion.



(a) (b)

Figure 3. Examples of small and large city sizes. (a) Small city at initialization; circular form with 4753 housing units. (b) Large city at initialization; circular form with 17 101 housing units.

We performed exploratory analyses using even larger city size specifications, including specifications of up to 250 000 households (which, assuming four persons per household, would imply a city of 1 000 000 inhabitants). These exploratory analyses gave no indication that further increases in city size added value to the analysis. In view of this, and in view of the fact that computational demands increase dramatically with increments in city size, we limited our analysis to just two settings for city size.

We specified agent vision using site-centered neighborhoods and varied neighborhoods through the dimensions of shape and size.(7) We used three shapes: circular, diamond shaped, and square. Examples are provided in figure 4. Diamond-shaped neighborhoods are based on cardinal moves of V where V is the range of vision from a focal housing unit. At $V = 1$ this yields a 'Von Neumann' or 'Rook's' neighborhood with four units. Circular neighborhoods are defined on the basis of a radius of V extending out from the focal unit. At $V = 1$ this also yields a 'Rook's' neighborhood with four units. Square-shaped neighborhoods are based on horizontal or vertical distance V from the focal unit. At $V = 1$, this yields a 'Moore' or 'Queen's' neighborhood with eight units.

In the simulations we varied neighborhood size by generating random values for V from 1 to 7. At any value of V (with the exception of $V = 1$), the diamond-shaped neighborhoods have half the units of square-shaped neighborhoods and the circular neighborhoods have an intermediate number of units. We term the total number of units in the neighborhood 'scale'. We use scale, not V, as our measure of neighborhood size.

Settings for model parameters known to affect segregation outcomes

Two parameters—preferences for coethnic contact and city ethnic mix—have especially important consequences for segregation outcomes in agent-based models. Schelling (1971) and Fossett and Waren (2005) have documented how preferences for coethnic contact have effects on segregation that are not simple, but instead are profoundly conditioned by the ethnic demography of the city. Generally speaking, preferences for coethnic contact promote segregation via the dynamic of congregation whenever a group's preferred level of coethnic presence exceeds the group's supply in the population.

(7) See Fossett (2005b) for a review of options in SimSeg for specifying city size, shape, and form, neighborhood shape and size, and distance decay in agent vision.

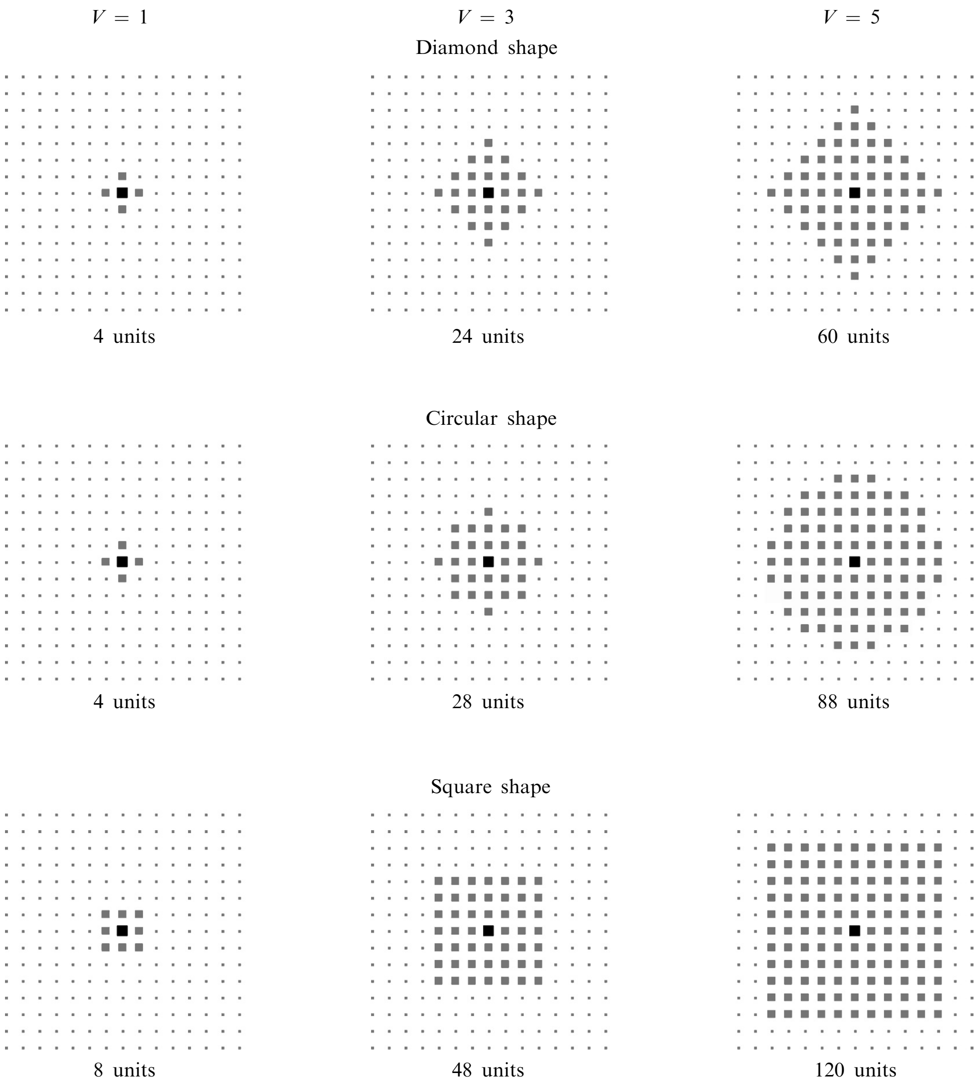


Figure 4. Examples of variations in agent vision by neighborhood shape and scale [as determined by range of vision (V)].

For the purposes of the present analysis, we set agent preference for coethnic contact to 30%. This setting is relatively low; at least it is low compared with levels of preference for coethnic contact measured in surveys (eg Clark, 1991; 1992; 2002). We vary the ethnic mix of the city from 10% black to 90% black. As will be seen below, this variation specification highlights a well-known Schelling effect—modest preferences for coethnic contact can produce surprisingly high levels of segregation under certain demographic conditions.

The general expectation that congregation dynamics will produce segregation when demand for coethnic neighbors exceeds supply leads to the following prediction: segregation will be lower when the city's ethnic mix is in the range 30–70% black and will be higher when it is outside this range. Analyses reported below confirm this prediction, but only under certain conditions. More specifically, the prediction is confirmed when agent vision is 'expansive' (ie involves larger neighborhoods); it is not confirmed when agent vision is 'narrow' (ie involves very small neighborhoods).

Laurie and Jaggi (2003) were the first to explore in detail how vision conditions the effects of preferences on segregation. Focusing on the special case in which the city has a 50 – 50 ethnic mix, they found that expanded vision is permissive of integration when preferences for coethnic contact are moderate (eg agents seek 30% coethnic contact). Fossett and Waren (2005) extended their findings, pointing out that the combination of moderate ethnic preferences and balanced ethnic mix they used was optimal for achieving integration in agent-based models and that the same preferences would produce high levels of segregation when the city ethnic mix was imbalanced. In addition, Fossett and Waren documented that preferences, ethnic mix, and vision interact in complex ways. For example, they showed that expanded vision produces higher, not lower, levels of segregation when congregation dynamics are strongly predicted (eg when group preferences for coethnic contact exceeded group percentages in the city population).

In the simulations we report here we vary the city ethnic mix and vision to generate a complex array of segregation outcomes. This provides a useful background against which to assess possible effects of city size, city form, and neighborhood shape. It allows us to speak to the issue of whether complex and substantively important Schelling-style effects are robust with regard to specifications of city size, city shape, city form, and neighborhood shape.

Results

Three variables in the simulations are specified as categorical: city shape and form (three categories), city size (two categories), and the shape of neighborhood vision (three categories). Two variables are specified as being interval in nature: the scale of neighborhood vision (ranging from 4 to 225) and the city ethnic mix (ranging from 10% to 90% black). To generate results for analysis, we ran 500 simulation experiments for each of the eighteen unique combinations of the three categorical variables. In each set of 500 experiments we varied the scale of agent vision (based on V) and city ethnic mix randomly within the ranges noted above. We ran each experiment for thirty cycles and recorded segregation scores at initialization and again at the conclusion of the experiment Residential distributions at initialization are determined by random assignment; residential distributions at the conclusion reflect model-generated segregation patterns.

We used four measures to assess model segregation behavior: the index of dissimilarity computed using data for bounded areas (ie fixed-size areas with nonoverlapping fixed boundaries); the index of dissimilarity computed using data for site-centered areas (ie overlapping areas of fixed size and shape centered on a focal housing unit); the contact difference index computed using data for bounded areas; and the contact difference index computed using data for site-centered areas. Formulas for computing the segregation scores were taken from Fossett (2005a; 2007). Scores for the index of dissimilarity were normed against expected values under random assignment using methods outlined in Winship (1977) and Fossett (2005a).[8] Substantive findings are not affected by using normed values of the index of dissimilarity, with one exception: normed scores are lower on average.[9]

[8] Winship (1977) and Fossett (2005a) document that expected values of the index of dissimilarity under random assignment vary systematically with the ethnic mix of the city. For example, under random assignment, the expected value of D is 11.9 when the ethnic mix is 50 – 50 and 20.0 when the ethnic mix is 10 – 90 or 90 – 10. The norming transformation $Z = 100(D - E)/(100 - E)$ removes this component of variation in scores for the measure. The expected value of Z is 0.0 under random assignment. Thus, nonzero values for Z can be interpreted as segregation that systematically exceeds chance deviations from an exactly even distribution.

[9] The correlation of normed and original dissimilarity scores exceeds 0.99.

In general, the substantive findings obtained were fundamentally similar using all four measures. There is one notable exception—segregation scores were higher on average when computed using overlapping, site-centered areas. Figure 2, introduced earlier, gives a visual sense of why this is the case. The figure shows that segregation patterns in three examples of final city landscapes do not closely follow the boundaries of the nonoverlapping, fixed boundary areas. The reason for this is that the simulations conducted for this study implemented agent vision in terms of overlapping, site-centered areas. Analyses not reported here show that, if agent vision is implemented in terms of bounded areas (eg simulating sensitivity to school district boundaries), segregation scores are higher on average using measures based on bounded areas rather than on site-centered areas.(10) Since substantive findings are similar when using the different measurement strategies, we generally present just the results of dissimilarity scores for bounded areas, because they are more familiar.

The key findings of our analyses can be easily summarized. Segregation outcomes vary in substantively important ways over different settings for the city ethnic mix and agent vision. Segregation outcomes do not vary in substantively important ways over different settings for city size, city form, and neighborhood shape.

The importance of city ethnic mix and agent vision is documented in figure 5. Separate curves are plotted for each of the separate settings for the scale of agent vision as measured by neighborhood size in housing units. The settings for agent vision assessed on this dimension range from 4 units, for the extremely small-scale Rook's

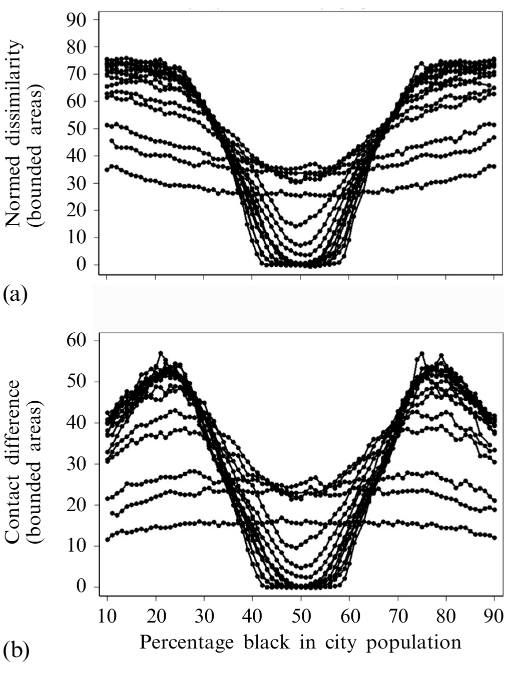


Figure 5. Segregation under modest preferences for coethnic contact according to city ethnic mix and varying agent vision: (a) normed dissimilarity; (b) contact difference. See text for discussion of the individual curves shown.

(10) More exactly, the average scores are higher using nonoverlapping, fixed-boundary areas if the boundaries are aligned with boundaries relevant for agent vision. If segregation measures are computed using arbitrary nonoverlapping fixed-boundary areas, the resulting scores are not necessarily higher than those computed using overlapping, site-centered areas.

neighborhood, to 224 units for the 15×15 square neighborhood (based on $V = 7$). The figure reveals a complex interaction between preferences, city ethnic demography, and vision previously examined in Laurie and Jaggi (2003) and Fossett and Waren (2005).

When vision is extremely narrow and involves neighborhoods of very small spatial scale, segregation outcomes are moderately high at all settings for city ethnic mix. This is seen in the three relatively flat curves observed in both graphs. For the dissimilarity scores these three curves open slightly upwards. For the contact difference scores they open slightly downwards. The settings for neighborhood size for these curves are four, eight, and twelve units, respectively.(11)

Laurie and Jaggi (2003) and Fossett and Waren (2005) account for this pattern as follows. Segregation is surprisingly high under the 50 – 50 ethnic mix for these settings because random entries and exits can easily 'tip' the ethnic mix of small-scale local neighborhoods and make them unattractive to one or the other of the two groups, even though both groups have only modest preferences for coethnic contact. This precipates congregation dynamics at small spatial scales which produce 'snaking' or dendritic patterns of segregation documented in example landscapes presented in Laurie and Jaggi (2003) and Fossett and Waren (2005). Segregation does not climb to high levels, at least not in the time frames used in our simulations, because the spatial scale of the dendritic pattern of segregation that emerges is smaller than the spatial scale of the neighborhood size we use in computing segregation measures.(12)

Laurie and Jaggi (2003) and Fossett and Waren (2005) document that the pattern of segregation changes when agent vision becomes more expansive. The critical transition range is roughly between 8 – 12 units on the low side and 24 – 40 units on the high side. This is highlighted in figure 6, which plots curves for agent vision based on diamond-shaped neighborhoods with 12, 24, 40, and 112 units, respectively. At 12 units the curve is relatively flat for both graphs. At 24 units the tails of the curve are clearly moving upward, signaling higher levels of segregation for cities with imbalanced ethnic demographies. At 40 units and beyond the tails continue to climb and, more importantly, the center region of the curve begins to dip, signaling lower levels of segregation for cities with balanced ethnic demographies.

The dramatic change in the shape of the curves as agent vision becomes more expansive is important on at least two levels. Technically, it highlights how preference-driven segregation behavior in agent models is strongly conditioned by the scale of agent vision. More substantively, it may establish a finding with relevance for understanding segregation in the real world. Specifically, since real-world residential dynamics almost certainly involve expanded vision, it suggests that weak to moderate preferences for coethnic contact may have a surprising capacity to support ethnic residential segregation.

This now brings us to the central question of this study: Are these theoretically interesting and complex Schelling-style preference effects robust? Or do they depend in critical ways on model specification choices regarding factors such as city size, city shape, city form, and neighborhood shape? The answer appears to be clear and simple; the effects just reviewed are highly robust across the model specification choices we consider. Our results suggest that, to the extent that the factors of city size, city shape, city form, and neighborhood shape have effects on segregation at all, they are slight and substantively uninteresting.

(11) These correspond to the four-unit Rook's neighborhood, the eight-unit Queen's neighborhood, and twelve-unit diamond-shaped neighborhood with vertices of two units.

(12) In unreported analyses we find that, given sufficient time, small-scale segregation tends to evolve into larger-scale segregation. But the necessary length of time can be quite long.

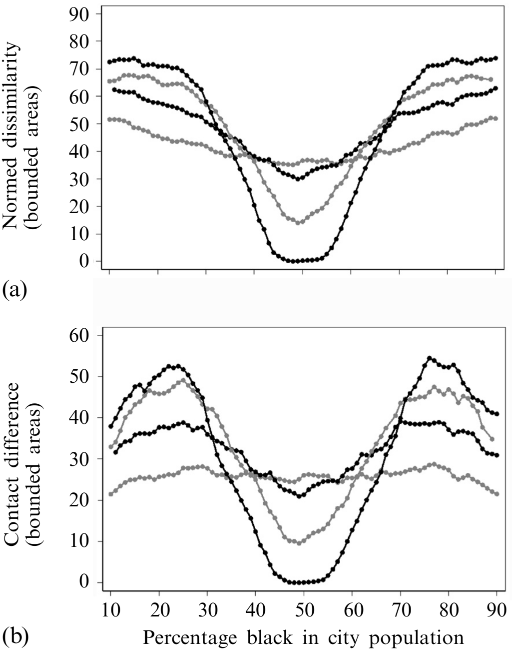
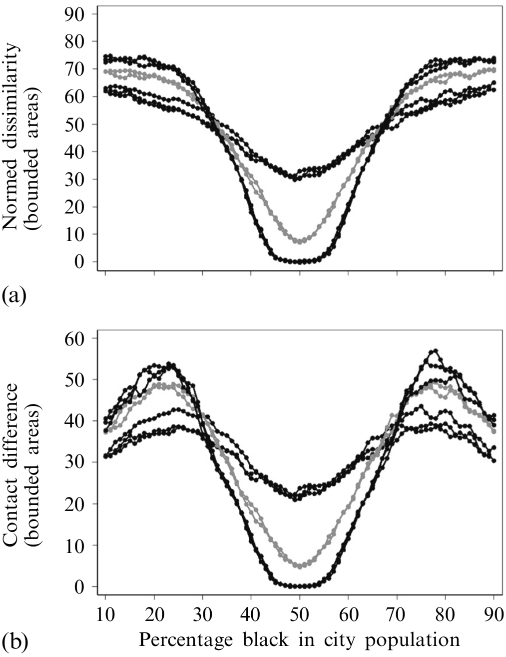


Figure 6. Segregation under modest preferences for coethnic contact according to city ethnic mix and four settings of neighborhood size: (a) normed dissimilarity; (b) contact difference. Curves in gray show agent vision values of 12 and 40; curves in black show agent vision values of 24 and 112.

Figure 7 presents graphical results consistent with this conclusion. The graphs display separate curves for agent vision associated with different neighborhood shapes that yield roughly similar neighborhood size. Consider, for example, the curves for diamond-shaped areas with 24 units, circular areas with 28 units, and square areas with 24 units. These three curves are plotted in black in figure 7 and yield the highest segregation scores when the city percentage black is 50%. Visually, the graphs suggest that differences between these three curves are not particularly interesting. The same is true at larger spatial scales. Curves for a square neighborhood with 48 units and a circular neighborhood with 48 units are plotted in gray in figure 7. They yield intermediate scores for segregation when the percentage black in the city is 50%. The differences between these two curves also appear trivial. Finally, curves for a diamond-shaped neighborhood with 112 units, a circular neighborhood with 120 units, and a square neighborhood with 120 units are plotted in black. They yield the lowest scores for segregation when the percentage black is 50%. The differences between these three curves also are slight and unimpressive.

Overall, the graphs in figure 7 provide a strong visual impression that variation in agent vision associated with neighborhood scale matters in a very important way. But, holding scale constant, variation in neighborhood shape does not appear to matter, at least it does not matter over the range of shapes considered here.[13] This visual impression is confirmed quantitatively by regression analysis of the effect of neighborhood shape and other factors on the model-generated segregation outcomes reported in tables 1 and 2. These tables report the results of ordinary least squares regressions of

[13] We should note that this conclusion applies to a range of neighborhood shapes that are relatively compact. We do not consider, for example, hourglass, star, or other irregular shapes, as such shapes have never been entertained in the literature.



(a)

(b)

Figure 7. Segregation under modest preferences for coethnic contact according to city ethnic mix for selected comparisons of neighborhood shape: (a) normed dissimilarity; (b) contact difference. Curves in gray show neighborhoods of 48 units. See text for further discussion of individual curves.

Table 1. Regressions of normed dissimilarity scores on selected model parameters.

Model parameter	Imbalanced ethnic mix		Balanced ethnic mix	
Agent vision based on ln of neighborhood size	11.83[a]	11.83[a]	−8.61[a]	−8.66[a]
City size				
small (reference)				
large		−0.03		−0.00
City shape and form				
circle (reference)				
square		−0.36		0.18
square torus		−1.04[a]		−0.27
Neighborhood shape for vision				
diamond (reference)				
circle		−0.61[a]		−0.31
square		−0.42[a]		0.71
Adjusted R^2	0.87	0.87	0.49	0.49
N	3941	3941	2062	2062

[a] Probability of chance occurrence is less than 0.01.

segregation scores from the end of each simulation on dummy variables used to assess the impact of city size, city shape, city form, and neighborhood shape.

The regressions include a control for agent vision measured by the natural logarithm of the number of units in the neighborhood seen by agents. The ln specification is used to capture nonlinearity in the effect of the scale of agent vision—its impact diminishes as the scale of vision increases. We take account of the complex interaction

Table 2. Regressions of contact difference scores on selected model parameters.

Model parameter	Imbalanced ethnic mix		Balanced ethnic mix	
Agent vision based on ln of neighborhood size	10.47[a]	10.48[a]	−5.74[a]	−5.77[a]
City size				
small (reference)				
large		−0.01		−0.01
City shape and form				
circle (reference)				
square		−0.18		0.11
square torus		−1.46[a]		−0.31
Neighborhood shape for vision				
diamond (reference)				
circle		−0.47[a]		−0.15
square		−0.53[a]		0.63
Adjusted R^2	0.81	0.82	0.44	0.44
N	3 941	3 941	2 062	2 062

[a] Probability of chance occurrence is less than 0.01.

between agent vision and city ethnic mix by performing regressions separately for two subsets of simulations: a subset of simulations in which the city ethnic mix is imbalanced, specified as simulations where the black population is 10–30% or 70–90%, and a subset of simulations in which the city ethnic mix is balanced, specified as simulations where the black population is 40–60%.

Table 1 presents the regression results for normed dissimilarity scores. The first column reports the results from the regression of normed dissimilarity on the scale of agent vision when the city ethnic mix is imbalanced. As suggested visually in figure 4, segregation increases with the scale of agent vision. The second column reports regression results assessing whether city size, city form, and neighborhood shape affect segregation outcomes. Owing to a large sample size some effects are statistically significant, but the effects are small and substantively unimportant. This is reflected in their trivial impact on model fit (as indicated by the adjusted R^2 statistic).

The third and fourth columns report parallel results for the regression of normed dissimilarity scores when the city ethnic mix is balanced. In this case the scale of agent vision has important *negative* effects on segregation—again confirming the visual impression imparted by figure 5. Here, the effects of city size, city form, and neighborhood shape are again small. None of the effects is statistically significant, possibly because the sample size is smaller. Again the impact of the effects on the model fit is trivial.

Table 2 reports the results of comparable regressions performed using the contact difference index to measure segregation outcomes. The substantive patterns are the same as before. Agent vision has important effects on segregation that are complex; segregation increases with the scale of vision when the ethnic demography of the city is imbalanced and segregation decreases with the scale of vision when the ethnic demography of the city is balanced. In contrast, city size, city form, and neighborhood shape do not have important effects on segregation outcomes.

We also replicated the results presented in tables 1 and 2 using site-centered versions of the normed dissimilarity index and the contact difference index. In the interests of space we do not present these results here because, in every important respect, they closely parallel the results just reviewed.

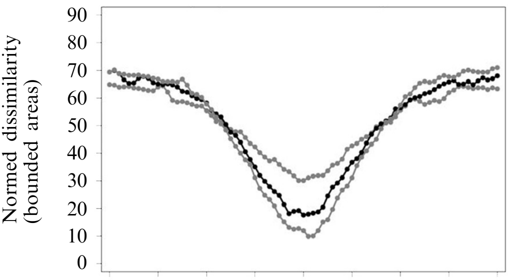
Agent vision with distance decay

Finally, we also report results from analyses assessing the impact of specifying agent vision using a distance-decay function in which nearest neighbors are given more weight than distant neighbors. The key findings are highlighted in figure 8, which plots segregation according to city ethnic mix for three different sets of 2000 simulation experiments, each using different specifications of distance decay for agent vision.

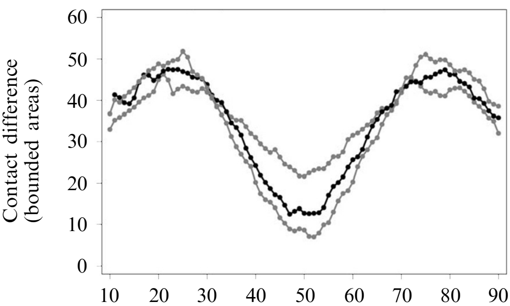
The basic design elements in these simulations are the same as before: preferences for coethnic contact are set at low levels (ie agents seek 30% coethnic presence) and the ethnic mix in the city is varied randomly between 10% and 90% black. Agent vision in these simulations is specified using a circle-shaped, site-centered neighborhood.

The key factor that varies across the three designs is the combination of the 'nominal' scale of agent vision and the presence or absence of distance decay in agent vision. The two curves plotted in light gray in figure 8 are for designs in which distance decay is not active so all neighbors are weighted equally. One uses a smaller neighborhood ($V = 3$) with twenty-eight neighbors and the other uses a larger neighborhood ($V = 4$) with forty-eight neighbors. As seen earlier, the results for the larger neighborhood differ from the results for the smaller neighborhood; the curve which plots segregation according to city ethnic mix is lower in the middle and higher at the tails when vision is specified using the larger scale neighborhood.

The curve plotted in black in figure 8 is for results from simulations which also use the larger neighborhood ($V = 4$) with forty-eight total neighbors. The difference in this case, however, is that agent vision is specified using the distance-decay function, in which nearer neighbors are given greater weight and distant neighbors are given lesser weight. Specifically, the nearest neighbor is given a weight of 1.0 and the most



(a)



(b)

Figure 8. Segregation under modest preferences for coethnic contact according to city ethnic mix for selected comparisons with and without distance decay: (a) normed dissimilarity; (b) contact difference. Curves in gray show designs for which distance decay is not active with $V = 3$ and $V = 4$; the curve in black shows a design with $V = 4$ for which a distance-decay function was used.

distant neighbor is given a weight of 0.1. Neighbors at intermediate distances are given weights between 0.1 and 1.0 on the basis of a simple linear function of distance. The total or unweighted count of neighbors is 48.0; the weighted count of neighbors is 27.7.

The number of total neighbors invites a comparison with the results for the larger (forty-eight-neighbor) neighborhood with the distance decay inactive. The number of weighted neighbors invites a comparison with the results for the smaller (twenty-eight-neighbor) neighborhood with the distance decay inactive. Interestingly, the curve reflecting the results obtained for the simulations in which distance decay was active falls between these two points of comparison. This suggests that implementing agent vision with distance decay has an effect similar to reducing the scale of agent vision—that is, the curve shifts back in the direction of the curve for the smaller neighborhood. At the same time, however, the results also indicate that the impact of specifying agent vision using a larger neighborhood is not completely attenuated as the curve shifts to a position that is intermediate between the two curves for the larger and smaller neighborhoods.

We conducted additional explorations of the impact of implementing distance decay in agent vision. We replicated the reports just reported using other neighborhood shapes (eg diamonds and squares as well as circles) and using other ranges of vision (V). In the interests of space, we do not report these results at length because they were similar in all important respects to the results already reported—that is, the implementation of distance decay has an effect very similar to that of specifying reduced agent vision. We also explored the impact of using convex and concave distance decay functions in addition to the linear distance-decay function used in the results reported in figure 8. Selected results from these analyses are presented in figure 9.

The experiments reported in figure 9 are all based on a circular neighborhood form with $V = 3$, which yields twenty-eight neighbors. A set of 2000 experiments was conducted without distance decay. Five additional sets of 2000 experiments were conducted on the basis of five different distance-decay functions: strongly convex, moderately convex, linear, moderately concave, and strongly concave.[14] The five distance-decay functions range from building mild—with decay being slight with increasing distance initially and accelerating with increasing distance—to very aggressive—with decay being rapid with increasing distance initially and diminishing with increasing distance. As noted above, the unweighted count of neighbors is 28.0. As distance decay progresses from mild to aggressive, the weighted count of neighbors declines as follows: 20.8, 18.4, 16.4, 14.5, and 12.5.

In figure 9, the curves for the experiments in which distance decay was not active are plotted in black. They take the lowest values of any of the plotted curves when ethnic mix is balanced (ie 50–50) and they take the highest values of any of the plotted curves when ethnic mix is unbalanced (eg 20–80 or 80–20). The curves based on the most aggressive distance-decay function are plotted in light gray. They take the highest values of any of the plotted curves when ethnic mix is balanced and they take the lowest values of any of the plotted curves when ethnic mix is imbalanced. The other curves using less aggressive distance-decay functions are plotted in a darker gray. They fall between the two previous curves in a straightforward progression: as distance decay becomes milder, the curves shift toward the curves for which the distance decay was not active.

Figure 9 highlights that the basic impact of implementing distance decay is similar to that of reducing the scale of agent vision. Compared with the curve for which distance decay is not active, all of the curves for results using distance decay have

[14] See Fossett (2005b) for more detailed discussion of the distance-decay functions.

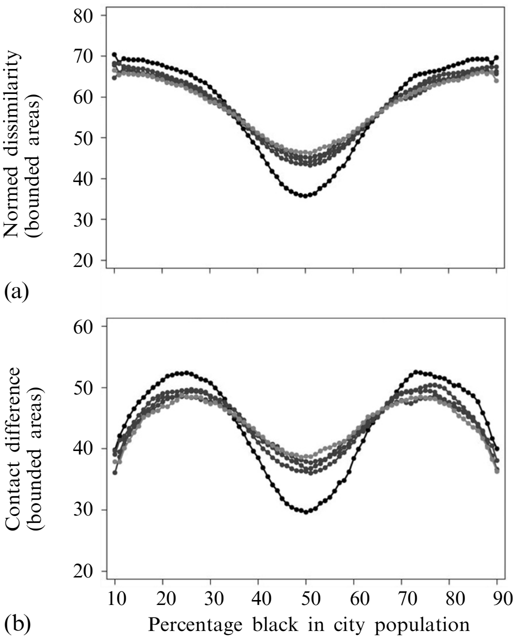


Figure 9. Segregation under modest preferences for coethnic contact according to city ethnic mix for selected comparisons both without distance decay, and with five different distance-decay functions: (a) normed dissimilarity; (b) contact difference. Curves in black, light gray, and dark gray show designs for which distance decay is not active, most aggressive, and less aggressive, respectively. See text for further discussion of curves.

higher segregation when ethnic mix is balanced and lower scores when ethnic mix is unbalanced. This is the type of change observed when the scale of agent vision is reduced and distance decay is not active. The differences among the curves for the five different distance-decay functions are not dramatic, indicating that the precise form of distance decay is not especially important in this context.

Beyond the effects just reported, we could find no other substantively interesting effects of distance decay. The practical implications for simulation studies appear to be straightforward. Neighborhood scale is the crucial dimension of agent vision. Its consequences can be explored with or without implementing distance decay in agent vision. Distance-decay approaches to specifying agent vision may be preferred on theoretical grounds, but the practical consequence appears to be limited to altering the expression of neighborhood scale in ways that are not surprising and not especially important. Thus, we conclude that investigators can implement distance decay as a part of agent vision at their discretion. Once the consequence of effectively attenuating neighborhood scale is taken into account, the choice is not likely to be important.

Discussion and implications

The analyses reported in this study have several important implications for computational studies of residential segregation. First, holding the scale of vision constant, the neighborhood shape dimension of agent vision appears to be largely irrelevant to model-generated segregation outcomes; at least, the shapes considered here—diamonds, circles, and squares—can apparently be used interchangeably without raising concerns that the choice matters in any important way. In particular, there is no evidence that using one shape or another will either obscure or exaggerate model-generated segregation outcomes produced by more substantively interesting segregation dynamics associated with ethnic preferences, ethnic demography, and scale of agent vision. The same appears to be true for the distance-decay dimension of agent vision; once consequences for neighborhood scale are considered, distance decay does not appear to have important consequences for segregation dynamics.

In a like manner, city form also appears to exert only slight and substantively unimportant effects on model-generated segregation outcomes. It would appear that researchers can use circular forms, square forms, and torus forms for the city shape as they wish. We can find no evidence to suggest that model-generated segregation outcomes produced by more substantively interesting segregation dynamics will be obscured or exaggerated by the choice concerning this aspect of model specification.

Finally, city size does not appear to be an important factor in agent-based segregation studies. Model-generated segregation outcomes produced by substantively interesting segregation dynamics such as ethnic preferences, ethnic demography, and scale of agent vision are statistically identical in smaller and larger cities. We offer one qualification to this conclusion. While the larger city size considered here is more than three times the size of the smaller city, we must acknowledge it is still small compared to the size of large metropolitan areas. Perhaps differences in city size only matter when the range of variation in size is enormous.

To consider this, we conducted exploratory analyses using cities that were much larger in size—ten times larger than the larger city size used here. We could find no visual or quantitative evidence of size effects in these analyses. In view of this, and given the dramatic increase in computational burden associated with performing analyses using very large cities, we did not pursue this systematically.(15) We feel safe in predicting that our overall findings would have been the same if we had pursued the matter further. We base this prediction both on empirical and on theoretical grounds. The exploratory analyses we conducted provide preliminary empirical grounding. The theoretical grounding involves both positive and negative elements. On the negative side, those who speculate that Schelling-style effects might be peculiar to small cities offer no compelling theoretical rationale for this view. On the positive side, we have compelling evidence that microlevel agent decisions based on the relative attractiveness of alternative residential locations will produce segregation in model systems. The key theoretical ingredients of this process—agent preferences for coethnic contact and variation in the ethnic mix of local regions within the city—will be present in large as well as small cities. Accordingly, we have good reason to expect Schelling-style segregation effects to be expressed in model cities of any size.

Taken together the findings of this study are welcome. They suggest that substantive findings about segregation dynamics gleaned from agent-based computational models do not depend in crucial ways on choices regarding several aspects of model specification. The finding that these choices are mundane, rather than critical, may help to overcome the skepticism some direct toward simulation studies of segregation dynamics. We suspect that many who question the results of these studies are wary in part because some of the stylized conventions of computational models such as torus landscapes and diamond-shaped neighborhoods are unfamiliar and highly abstract. These design choices, coupled with the small size of simulated cities and the game-like quality of computational models, feed concerns that insights from these model residential systems may have limited relevance for real-world segregation patterns. This paper obviously does not address every concern that skeptics may raise about agent models of segregation, but it contributes to addressing these concerns by systematically

(15) The SimSeg simulation can generate cities with over 250 000 households. However, it can take hours or even days to run an individual simulation of the type conducted in this study using this specification of city size.

investigating whether factors such as city size, city shape, city form, and neighborhood shape matter for model-generated segregation outcomes. Our results suggest that they do not, at least not in any important and obvious way. This shrinks the boundaries of legitimate concerns and directs future research on this issue in new directions.

In closing, we entertain the following question, "if factors such as city size, shape, and form, and neighborhood shape are not important in Schelling-style models of residential segregation dynamics, what factors do matter?" A preliminary answer seems clear. Preferences for coethnic contact matter; the ethnic demography of the city matters; and the scale of agent vision matters. All three are crucial in their own right and, as the results reviewed above demonstrate, they come together in complex ways to shape model-generated segregation outcomes. These basic insights were first established in Schelling's (1971) landmark study. They have been elaborated in more detail and with greater quantitative rigor in subsequent studies. The present study contributes to this tradition by highlighting some of the more theoretically compelling Schelling-style effects and demonstrating that they are robust over a wide range of model design choices. We invite others to explore and to test the boundaries of these insights in new ways.

Acknowledgements. The research reported here and the development of the SimSeg simulation model has been supported in part by NIH Grants R43HD038199 (Simulating Residential Segregation Dynamics: Phase I) and R44HD038199 (Simulating Residential Segregation Dynamics: Phase II). An earlier version of this paper was presented at the Annual Meeting of the Southwestern Sociological Association, New Orleans, Louisiana, 23 March 2005.

References

Clark W A V, 1991, "Residential preferences and neighborhood racial segregation: a test of the Schelling segregation model" *Demography* **28** 1 – 19

Clark W A V, 1992, "Residential preferences and residential choices in a multi-ethnic context" *Demography* **30** 451 – 466

Clark W A V, 2002, "Ethnic preferences and ethnic perceptions in multi-ethnic settings" *Urban Geography* **23** 237 – 256

Epstein J M, Axtell R, 1996 *Growing Artificial Societies* (MIT Press, Cambridge, MA)

Fossett M, 2005a, "Measuring uneven distribution: conceptual and practical considerations for segregation studies", SimSeg technical paper; available from author

Fossett M, 2005b, "The urban geography of SimSeg: the city landscape and selected spatial domains in the SimSeg simulation model", SimSeg technical paper; available from author

Fossett M, 2006, "Ethnic preferences, social distance dynamics, and residential segregation: theoretical explorations using simulation analysis" *Journal of Mathematical Sociology* **30** 185 – 274

Fossett M, 2007, "The separation index: a new alternative for measuring uneven distribution", SimSeg technical paper; available from author

Fossett M, Waren W, 2005, "Overlooked implications of ethnic preferences for residential segregation in agent-based models" *Urban Studies* **25** 1893 – 1917

Goering J, 2006, "Shelling redux: how sociology fails to make progress in building and empirically testing complex causal models regarding race and residence" *Journal of Mathematical Sociology* **30** 299 – 317

Laurie A J, Jaggi N K, 2003, "Role of 'vision' in neighbourhood racial segregation: a variant of the Schelling segregation model" *Urban Studies* **40** 2687 – 2704

Macy M W, van de Rijt A, 2006, "Ethnic preferences and residential segregation: theoretical explorations beyond Detroit" *Journal of Mathematical Sociology* **30** 275 – 288

Meen D, Meen G, 2003, "Social behaviour as a basis for modeling the urban housing market: a review" *Urban Studies* **40** 917 – 935

Pancs R, Vriend N J, 2003, "Schelling's spatial proximity model of segregation revisited", WP 487, Department of Economics, Queen Mary College, University of London

Sander R H, Schreiber D, Doherty J, 2000, "Empirically testing a computational model: the example of housing segregation", in *Proceedings of the Agent 2000 Workshop on Simulation of Social Agents: Architectures and Institutions* Eds D Sallach, T Wolsko, Argonne National Laboratory, Argonne, IL, pp 108 – 115, http://agent2007.anl.gov/proceedings/2000.pdf

Schelling T C, 1971, "Dynamic models of segregation" *Journal of Mathematical Sociology* **1** 143 – 186
Wasserman H, Yohe G, 2001, "Segregation and the provision of spatially defined local public goods" *The American Economist* **45** 13 – 24
Winship C, 1977, "A reevaluation of indexes of residential segregation" *Social Forces* **55** 1058 – 1066
Yinger J, 1995 *Closed Doors, Opportunities Lost: The Continuing Costs of Housing Discrimination* (Russell Sage Foundation, New York)
Young H P, 1998 *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions* (Princeton University Press, Princeton, NJ)
Young H P, 2001, "The dynamics of conformity", in *Social Dynamics* Eds S N Durlauf, H P Young (MIT Press, Cambridge, MA) pp 133 – 154
Zhang J, 2004a, "A dynamic model of residential segregation" *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