

# Understanding the social context of the Schelling segregation model

William A. V. Clark<sup>a,b</sup> and Mark Fossett<sup>c</sup>

<sup>a</sup>University of California, Los Angeles, CA 90095; and <sup>c</sup>Texas A&M University, College Station, TX 77843

Edited by Susan Hanson, Clark University, Worcester, MA, and approved January 22, 2008 (received for review August 30, 2007)

**A recent article [Vinkovic D, Kirman A (2006) *Proc Natl Acad Sci USA* 103:19261–19265] showing that the Schelling model has a physical analogue extends our understanding of the model. However, prior research has already outlined a mathematical basis for the Schelling model and simulations based on it have already enhanced our understanding of the social dynamics that underlie the model, something that the physical analogue does not address. Research in social science has provided a formal basis for the segregative outcomes resulting from the residential selection process and simulations have replicated relevant spatial outcomes under different specifications of the residential dynamics. New and increasingly detailed survey data on preferences demonstrates the embeddedness of the Schelling selection process in the social behaviors of choosing alternative residential compositions. It also demonstrates that, in the multicultural context, seemingly mild preferences for living with similar neighbors carry the potential to be strong determinants for own race selectivity and residential segregation.**

preferences | simulation | ethnicity | integration | neighborhoods

A recent article outlined a model that can explain the way in which separation or segregation (clustering) can arise in physical processes and is thus a parallel to the clustering outcomes of the Schelling segregation model (1). The physical analogue is interesting and it is intriguing to learn that there are physical parallels to social processes with specific commonalities in the physical processes of clustering and the social process of residential separation and segregation. That said, it is not completely clear that we have advanced our understanding of segregation and segregation dynamics by generating a physical analogue to the Schelling model. Although the physical analogue explains clustering and separating, the most important issue in the Schelling model, from a social perspective, is how choices play out in the social fabric and lead to segregated residential patterns. We show here that there are well articulated mathematical explanations for social segregation, that simulation studies with relatively simple utility structures can replicate complex and sometimes subtle segregation patterns seen in real urban environments, and that data from surveys of preferences reiterate the role of social distance in segregation outcomes.

The original Schelling agent model was disarmingly simple in its construction (2, 3). It posited that an agent, a model representation of a household that could be white or black, preferred to be on a square on a checkerboard in which half or more of the eight adjacent neighbors were of a similar color. In the economic context, this was seen as having utility one compared with having utility zero. Schelling used simple simulations based on such hypothetical preference schedules to show that the adjustments of individual households responding to changes in composition on the checkerboard invariably lead to complete segregation (3). New survey evidence for different groups makes it possible to base preference schedules for simulation studies on empirical findings instead of using hypothetical schedules as Schelling did (4). As we review below, preference schedules based on surveys indicate that

households do typically favor residing with households of similar color.

The Schelling model was of mostly theoretical interest and was rarely cited (5) until a significant debate about the extent and explanations of residential segregation in U.S. urban areas was engaged in the 1980s and 1990s (6, 7). To that point, most social scientists offered an explanation for segregation that invoked housing discrimination, principally by whites, as the major force in explaining why there was residential separation in the urban fabric (8). Reevaluation of the explanations for residential separation suggested that individual preferences alone, and in combination with economic differences among ethnic groups, could play an important role in explaining patterns of residential separation (9). The Schelling model was critical in providing a theoretical basis for viewing residential preferences as relevant to understanding the ethnic patterns observed in metropolitan areas. Furthermore, its implications have been buttressed by accumulated findings on preferences in multicultural settings which show that all major racial and ethnic groups hold preferences that are as strong as or stronger than the relatively mild preferences Schelling considered in his original two-group formulation.

The underlying basis of residential preferences is a matter of ongoing debate and a full review of the issues is beyond the scope of this study. In brief, preferences have complex origins and emerge from the social psychology, group dynamics, and history of racial and ethnic relations. Many plausible views have been offered. For example, it has been suggested that preferences reflect attachments to group identity and group culture (e.g., language, religion, beliefs, norms, customs, music, art, etc.) formed in early socialization, the salience of ethnicity in personal identity and sense of social position, and ethnocentric evaluations wherein individuals favor their own group's culture over other group cultures. It also has been suggested that preferences are shaped by expectations and stereotypes, founded or unfounded, that coethnics will provide reciprocal acceptance and support; that other groups may be indifferent, unwelcoming, or overtly hostile; that residing in ethnic neighborhoods may be convenient and carry practical benefits; and that neighborhood ethnic mix is predictive of future neighborhood conditions and home values.

None of these possibilities is implausible and none are mutually exclusive of others, so all can be entertained simultaneously. Significantly, however, the basis of residential preferences is not crucial to assessing their implications for segregation. This is clear in mathematical formulations and computer simulation models. In such contexts, *particular preferences for coethnic contact produce identical residential choices and patterns of residential segregation regardless of the presumed basis of the preference*. Thus, from any point of view, groups' patterns of evaluating

Author contributions: W.A.V.C. and M.F. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>b</sup>To whom correspondence should be addressed. E-mail: wclark@geog.ucla.edu.

© 2008 by The National Academy of Sciences of the USA

the presence of households similar to and different from themselves may be relevant to understanding the outcomes of residential selection.

The more important distinction in the literature is the distinction between preferences and discrimination. Two key points are especially relevant. First, scholarly research and fair housing legislation maintain logical and legal distinctions between voluntary residential choices guided by preferences and involuntary constraints on housing choices resulting from housing discrimination, intimidation, and violence. Second, policy options for promoting integration differ dramatically for these two factors. Discrimination is prohibited by law and is subject to a variety of legal remedies. Preferences are outside the purview of fair housing law and remedies, if sought, will necessarily be fundamentally different. In view of this, it is crucial to gain a better understanding of the impact preferences may have on segregation.

### Mathematical Derivations and Agent-Based Models

Schelling's analyses of the implications of preference schedules suggested that mixed-race residential neighborhoods are not likely to be stable. Young (10) provides rigorous theoretical analyses supporting and extending Schelling's insights. Similarly, Zhang (11, 12) also has shown mathematically that segregation will result under the Schelling model, even when everyone desires integration. Specifically, he shows that segregation is a stochastically stable state "that tends to emerge and persist in the long run regardless of the initial state" (12, p 548). The finding that separation will occur even when most people prefer integrated neighborhoods is important because it may help us explain the persistence of separation even in a world that is increasingly tolerant and where society has made significant strides toward reducing discrimination in the housing market. "Without any discriminatory behavior in the housing market a slight preference for like-color neighbors . . . can give rise to a high level of residential segregation and cause it to persist" (11, p 164).

Formal analysis indicates that even asymmetric preferences for same-group contact can be sufficient to create segregated patterns under the Schelling model. Zhang established this for the two-color case with a preference for own color by one group and neutral preferences by the other group by drawing on the theory of stochastic dynamical systems and the method of agent modeling. Building on work in evolutionary game theory, he showed that even slight asymmetry in residential preferences between two groups will produce endogenous segregation. The basic model considers a set of  $n$  neighbors, either four in the rook's case, or eight in the queen's case, on a checkerboard lattice. Each of the locations may have a black agent or a white agent, or may be vacant. Price of housing is introduced, but not housing quality (all units are assumed to be of the same quality) and a market mechanism determines the price of housing. In the "housing market" (the locations of the lattice checkerboard), a "natural vacancy rate" facilitates housing market turnover. There is no restriction on the agent choices, but white agents prefer to live near whites whereas black agents are neutral with respect to their choices. In each period, a pair of locations is chosen randomly and a set of possible outcomes includes a black agent moving to a vacant location, a white agent moving to a vacant location, a black agent exchanging locations with a white agent, a black agent with a black agent, and a white agent with a white agent. The agents make choices based on their individual utilities but decisions are boundedly rational and agents may make utility-decreasing moves.

Under the conditions outlined earlier, Zhang shows that  $S$ , the set of all states that maximizes the sum of all agents' utilities, is stochastically stable. Significantly, if blacks are color neutral, and whites have a slight preference for like-colored neighbors (say at least a 60–70% preference), the results are intriguing: residential segregation emerges, vacancy rates are higher in areas with black

concentrations, and whites pay more for housing. The results Zhang obtains by using a game-theoretic framework are powerful; they move the Schelling model from a set of general statements and an inductive approach to neighborhood selection to a "precisely defined mathematical concept" (11, p 164) where neighborhood transitions are determined by a set of utility functions. Young (10) similarly provides rigorous formal foundations for Schelling's most basic insights.

### Preferences, Tolerance, and the Basis of Simulation Models

There has been a long history of simulating the Schelling model. Early models were spare and simple, but advances in computing capabilities make it more feasible than ever to use simulation approaches to explore how residential outcomes vary over a wide range of increasingly complex and realistic model inputs (13, 14). For example, a recent model incorporates the role of multiple types of preferences, multiple ethnic groups, urban and demographic conditions, and intergroup inequality in socioeconomic status (15). Models of this sort expand our ability to investigate the residential patterns that emerge when agent choices are guided by different preference distributions, including both hypothetical distributions and ones fashioned after results from survey studies.

A core feature of agent models is the use of game-theoretic contexts with feedback loops. In the standard two-group simulation, agents may be open to living in mixed neighborhoods and seek only to avoid being in the minority (i.e., less than half the population). However, their moves to satisfy this objective—that is, moving from an area where their group is in the minority to one where it is at least half—change the ethnic mix in both locations in a way that precipitates further movement. The neighborhood they leave becomes less attractive to members of their group and the neighborhood they enter becomes less attractive to members of the other group. Over successive iterations, homogeneous neighborhoods emerge.

The limitations of nonexperimental research make it difficult to evaluate the effects of different factors on segregation in real cities. Consequently, there has been a contentious debate between those who focus on the role of social distance and preferences as potential explanatory forces in creating separation, and those who invoke the effects of housing discrimination. Because neither group can establish empirically grounded evaluations of the effects of either preferences or discrimination, recent work has focused on the logical standing of explanations based on social distance asking, "Is it reasonable to entertain the hypothesis that social distance and preference dynamics could generate and sustain significant levels of segregation in the absence of discrimination" (15). The mathematical formulations we reviewed earlier suggest that segregation can arise from simple preferences for own-race neighbors. Below we show that simulations based on observed ethnic preferences also provide quantitative and qualitative (visual) evidence for this possible residential dynamic.

An extensive literature exists on preferences and their implications for residential separation. An early study of Detroit used survey flash cards depicting different racial residential neighborhood combinations to establish that blacks generally preferred neighborhoods that were half black and half white, but whites generally preferred neighborhoods that were nearly all white (16). Subsequent studies of other metropolitan areas replicated this finding (17, 18) and also documented similar patterns in the preferences of other racial and ethnic groups (9). Focusing on white–black differences, the finding that most white households want majority white neighborhoods, whereas most African American households prefer neighborhoods that are at least half black—with a 50/50 black–white composition most preferred—is central to arguments that preferences may contribute to group residential separation in cities. Two lines of reasoning support this conclusion.

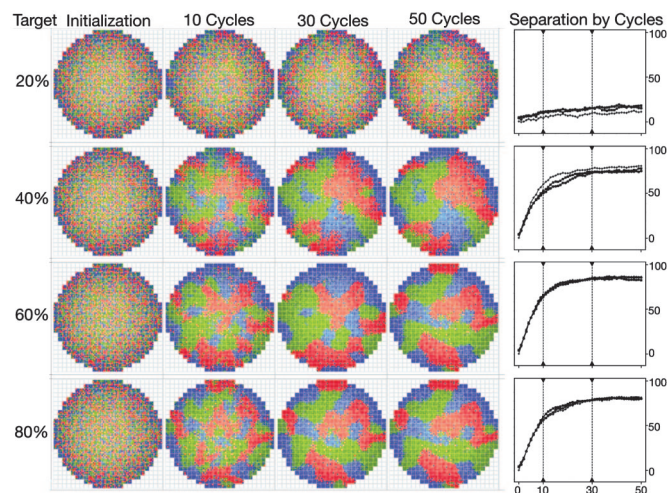




“virtual households with the ability to search in virtual housing markets” and make residential choices. Households have preferences for coethnic contact specified as the percentage presence of coethnic households in the neighborhood the household lives in or is considering moving to. The SimSeg model we use runs in a fashion similar to all agent models, but with greater capabilities than most previous simulations of residential dynamics. The model’s key features and algorithms are described in detail elsewhere (15). In brief, it creates housing units of varying quality and places them at residential locations across the city. It then populates this virtual city with households of different ethnic and socioeconomic status (income) and distributes them randomly to housing units. Ethnic segregation in the city thus initially reflects only random departure from even distribution. During the simulations households move according to the rules of the model by making residential choices guided by three separate and independent preference concerns: goals for housing quality, goals for neighborhood income, and goals for neighborhood ethnic composition. The program calculates segregation scores over the course of the simulation so trends in segregation can be assessed quantitatively and qualitatively based on graphical representations of the city’s residential distributions.

We present results from simulation experiments crafted to explore the implications of ethnic preferences in multigroup situations similar to those discussed previously. For simplicity, we consider a situation with three groups—blue, red, and green—and even demographic proportions (34%, 33%, and 33%, respectively).<sup>d,e</sup> Households are assigned ethnic preferences; that is, targets for same-group presence in neighborhoods.<sup>f</sup> We vary these in magnitude in different simulations to bracket the range of preferences reported in surveys. In addition, all households in all groups are assigned preferences for high-quality housing (relative to what they can afford) and preferences to reside in areas with high-average income. Households give identical weight to all three preferences, so goals regarding area ethnic composition are just one among several important goals shaping the location decisions households make.

Households “see” neighborhoods based on diamond-shaped local areas consisting of the 40 housing units within four cardinal moves of a focal housing unit.<sup>g</sup> Households evaluate neighborhood ethnic composition and average income based on the population residing in the local area. As in real urban areas, the city is given a significant city–suburb gradient on housing quality with low-income housing being concentrated in the center. In addition, the city has a high-income group (blues) and two low-income groups (reds and greens) whose average income is 30% below the average for the high-income group. The city–suburb housing gradient creates pronounced segregation along socioeconomic lines, but its impact on ethnic segregation is modest. At initialization, the two low-income ethnic groups are



**Fig. 3.** Animation sequences and segregation trends by targets for coethnic contact.

more centralized, but their overall segregation from the high-income ethnic group is not great (15).

We present representative examples of simulation results in Fig. 3.<sup>h</sup> Each row corresponds to a single simulation experiment and contains a sequence of images depicting the city’s residential distribution at four points in the experiment. Ethnic segregation is seen visually when households of different colors (i.e., blue, red, and green) cluster together on the city landscape. In addition, a quantitative segregation index—the variance ratio ( $V$ )—is computed for each of the three possible two-group comparisons.<sup>i</sup> The scores are plotted in the graph on the right-hand side of the figure. Socioeconomic segregation is seen visually based on high-income households being depicted in darker shades and low-income households in lighter shades. Accordingly, the central areas are lighter and the outlying areas are darker. Socioeconomic segregation is not our main focus, so we have omitted quantitative results to save space.

The key feature of the simulations is that we vary in-group preference targets assigned to households from levels substantially below those documented in surveys (e.g., 20%) to levels similar to those documented in surveys (e.g., 40–60%) and beyond.<sup>j</sup> Several important patterns are evident in the results. One is that ethnic segregation is low at the beginning of every simulation; not surprising because it primarily reflects random departure from even distribution. Another is that ethnic segregation emerges quickly and stabilizes at high levels when preferences for same-group contact are similar to those documented in surveys. The crucial transition occurs when preference targets reach or exceed a group’s population representation (22). Thus, high levels of segregation emerge when targets for same-group contact shift from 20% to 40%. Segregation is low when households seek 20% same-group contact because all households can be satisfied in neighborhoods that match the city’s ethnic mix. This is not the case when all households seek 40% same-group contact; accordingly, high levels of segregation emerge. Perhaps surprising to some, raising targets for same-

<sup>d</sup>SimSeg can implement two- and three-group simulations. The shift from two to three groups is sufficient to reveal many differences in the implications of preferences in the multigroup context.

<sup>e</sup>This is a conservative setting because specifying groups as equal in size is optimal for maintaining integration in Schelling-style segregation models (22). Higher levels of segregation result when we use more realistic demographic distributions (e.g., 65, 25, and 10).

<sup>f</sup>Targets for same-group presence are the same for all groups. We ran simulations in which preference targets were dispersed around the mean (as seen in the distributions in Fig. 2) and simulations in which targets were homogeneous. The results were similar either way, indicating that, in the main, preference effects revolve around the central tendency of the preference distribution.

<sup>g</sup>This specification is used in recent studies (22, 23) and is similar to the  $5 \times 5$  square neighborhood Schelling used in some simulations (2). Significantly, preference effects are robust over varying specifications of neighborhood size, shape, and type (e.g., bounded areas versus site-centered areas).

<sup>h</sup>The results shown here are from representative simulation experiments. We performed 500 replications for each design to establish the statistical norm for quantitative results.

<sup>i</sup> $V$  is a well known measure with many attractive technical properties. We also computed scores for the more widely used dissimilarity index ( $D$ ). Reporting  $V$  is a conservative choice; scores for  $D$  are always higher and suggest even more dramatic segregation results.

<sup>j</sup>We omit results for simulations with preferences set to zero. Not surprisingly, segregation did not emerge in these simulations.



outcomes may arise from simple microbehavioral dynamics. The social sciences have been slow to embrace these tools, but their potential value is becoming increasingly apparent. As the literature advances, agent models have great potential to help us to better understand the real-world dynamics of the city. Already, they have provided a basis for arguing that preferences, social distance dynamics, and residential choices cannot be dismissed as possible explanations of segregation outcomes. As the models continue to be refined, they may inform the debate about alternate explanations for segregation in the urban landscape even further.

## Observations and Conclusions

Our article shows that there is now a rigorous mathematical basis for the Schelling model and increasingly refined methods for simulating the impact of preferences and social distance dynamics. The results establish clearly that to ignore the role of choice behavior based on own-race preferences is to ignore a potentially important influence of racial and ethnic dynamics in the residential fabric. Of course, much work remains to be done to assess how segregation dynamics may change in an increasingly multiethnic society. For example, as the number of mixed race/ethnic individuals continues to increase, segregation patterns will be further complicated by their residential selections. Preliminary evidence suggests that mixed-race households are more likely to live in integrated neighborhoods than in homogeneous concentrations of either of their parental races or ethnicities.

Intra- and intergroup inequality in socioeconomic status both continue at high levels, racial and ethnic status continue to have high salience in social relations, and survey evidence shows that people generally strive to live with others who are similar to them with respect to ethnic and socioeconomic status. All of this suggests that we will continue to see significant levels of residential separation in the residential fabric (26). Although it is a question of some debate, there is at present no basis for anticipating evenness to emerge in racial and ethnic spatial distributions in our cities. Practices of legal, institutional race exclusion have been substantially dismantled, but stereotypes and prejudice remain and can promote residential separation in the absence of housing discrimination. Equally importantly,

people with similar social characteristics often congregate in the same neighborhoods based on mutual attraction, common interests, and shared sensibilities (27).

Past patterns of coercion and institutional practices such as race-restricted covenants produced involuntary separation by severely constraining housing options for minorities. But the emergence of affluent African American suburban communities in contemporary metropolitan areas is not easily explained in these terms. To the contrary, it is occurring in conjunction with expanding housing opportunities for middle class minorities and the rapid demise of exclusively white residential areas. It is suggestive of the power of positive ethnic association and the tendency for people to gravitate to the similar, the familiar, and the comfortable, especially when residing with other groups that may be indifferent or unwelcoming is not crucial for realizing other important residential outcomes (e.g., good schools, low crime, etc.). It further suggests that separation in the residential mosaic may diminish only slowly even as overt discrimination in housing declines and expressions of "racial tolerance" increase.

The implication of much commentary on urban areas is that the urban fabric would quickly become integrated if discrimination and other constraints on residential opportunities were eliminated (28). The preference and social distance work both in its mathematical and agent-based forms, and the survey results raise serious questions about this view (29). They suggest that mere tolerance and the absence of virulent housing discrimination will not produce integration under prevailing patterns of ethnic preference, at least, not in the short run. That the Schelling model has a physical analogue is interesting and theoretically illuminating is without question, but it is the social content of the Schelling model that gives it its power and importance. The implications of the mathematical formulations, simulation results, and survey evidence on preferences converge to suggest that we cannot presently expect rapid movement toward even spatial distributions of ethnic and racial groups and socioeconomic status groups. In view of this, further attention should be given to understanding how segregation in our metropolitan areas is shaped by the combination of substantial socioeconomic inequality within and between groups and social distance dynamics that reflect the salience of race-ethnicity and socioeconomic status in residential decisions.

- Vinkovic D, Kirman A (2006) *Proc Natl Acad Sci USA* 103:19261–19265.
- Schelling TC (1971) *J Math Sociol* 1:143–186.
- Schelling TC (1978) *Micromotives and Macrobehavior* (Norton, New York).
- Clark WAV (1991) *Demography* 28:1–19.
- Skvoretz J (2006) *J Math Sociol* 30:181–184.
- Galster G (1988) *Popul Res Policy Rev* 7:93–112.
- Clark WAV (1988) *Popul Res Policy Rev* 7:113–121.
- Massey D, Denton N (1993) *American Apartheid* (Harvard Univ Press, Cambridge, MA).
- Clark WAV (1992) *Demography* 30:451–466.
- Young HP (1998) *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions* (Princeton Univ Press, Princeton).
- Zhang J (2004) *J Math Sociol* 28:147–170.
- Zhang J (2004) *J Econ Behav Organ* 54:533–550.
- Epstein JM, Axtell R (1996) *Growing Artificial Societies: Social Science from the Bottom Up* (MIT Press, Cambridge MA).
- Gaylord RJ, D'Andria LJ (1998) *Simulating Society: A Mathematica Toolkit for Modeling Socioeconomic Behavior* (Springer, Amsterdam).
- Fossett M (2006a) *J Math Sociol* 30:185–274.
- Farley R, Schuman H, Bianchi S, Colasanto D, Hatchett S (1978) *Soc Sci Res* 7:319–344.
- Armor D, Clark WAV (1995) Housing segregation and school desegregation. *Forced Justice*, ed Armor D (Oxford Univ Press, Oxford), pp 117–153.
- Farley R, Fielding E, Krysan M (1997) *Housing Policy Debate* 8:763–800.
- Galster G (1982) *J Am Real Estate Urban Econ Assoc* 10:39–60.
- Sermos M (2000) *Geogr Anal* 32:225–246.
- Ellen IG (2000) *Sharing America's Neighborhoods: The Prospects for Stable Racial Integration* (Harvard Univ Press, Cambridge, MA).
- Fossett M, Warren W (2005) *Urban Stud* 42:1893–1917.
- Laurie AJ, Jaggi NK (2003) *Urban Stud* 40:2687–2704.
- Fossett M, Dietrich D (2008) *Environ Plan B*, in press.
- Brown LA, Chung SY (2008) *Ann Assoc Am Geogr* 98:180–212.
- Clark WAV (2002) *Urban Geogr* 23:237–256.
- St. John C (2006) *J Math Sociol* 30:327–332.
- Zubinsky C, Bobo L (1996) *Soc Sci Res* 25:335–374.
- Fossett M (2006b) *J Math Sociol* 30:289–298.