**PNAS Template for Main Manuscript**

This PNAS template for the Main Manuscript may be used to organize your main text source file. The template is intended to provide a clearly organized PDF to facilitate the review process. Further information is available in our [Author Center](https://www.pnas.org/authors/submitting-your-manuscript#article-types).

**Using the template**

Paste the appropriate text from your manuscript into the relevant section of the template. You may maintain the template formatting, or reapply styles after pasting your text into the template.

Figures should be placed on separate pages with legends set immediately below each figure. Table titles should be set immediately above each table. Please do not use field codes for figures and tables to ensure the numbering is correct after the PDF conversion.

References cited in the main text should be included in a separate reference list at the end of this file. Examples of the PNAS citation style are included below.

**Notes about submission**

The following items are required on the title page: **Title, Author Line, Author Affiliations, Corresponding Author information**. Each of these items **must** be provided on the title page for us to proceed with processing your paper.

You are not required to adhere to the section order outlined below. For example, you may combine your Results and Discussion or use alternate section headings. Materials and Methods should be included after the Results and Discussion in most cases. If your paper does not include the standard section headings, please provide a brief explanation in the **Comments to Editorial Staff** field of the submission form.

You may include subheadings within the standard headings listed below. You may also include line numbers.

**Submitting your main manuscript**

Delete this first page, and then save your completed main text file as a PDF for submission, following instructions available [here](http://www.pnas.org/site/authors/procedures.xhtml#preparation).

*Updated February 2022*



**Main Manuscript for**

Necessary Heterogeneity in Network Models of Residential Segregation

W.H. Deter

Paste the full affiliation list here.

\*Paste corresponding author name(s) here.

**Email:**  [xxxxx@xxxx.xxx](mailto:xxxxx@xxxx.xxx)

PNAS strongly encourages authors to supply an [ORCID identifier](https://orcid.org/) for each author. Do not include ORCIDs in the manuscript file; individual authors must link their ORCID account to their PNAS account at [www.pnascentral.org](http://www.pnascentral.org). For proper authentication, authors must provide their ORCID at submission and are not permitted to add ORCIDs on proofs.

**Author Contributions:** Paste the author contributions here.

**Competing Interest Statement:** Disclose any competing interests here.

**Classification:** Paste the major and minor classification here. Dual classifications are permitted, but cannot be within the same major classification.

**Keywords:** segregation models, population heterogeneity, topological heterogeneity

**This PDF file includes:**

Main Text

Figures 1 to X

Tables 1 to X

**Abstract**

The necessity of topological and behavioral heterogeneity in network models of segregation is examined. Previous investigation of heterogeneity of behaviors (preferences) have shown reductions in segregation on networks. Previous investigation of heterogeneity of topologies has shown no significant change in segregation levels. This work examines the impact of the combination of these heterogeneities to identify interaction between them.

**Significance Statement**

Paste your significance statement here. Please note that it should not exceed 120 words, but should be at least 50 words in length. It should not include any references.

**Main Text**

**1. Introduction**

Residential segregation is a persistent topic of discussion in the social sciences whose investigation is often enriched by computational models, particularly agent-based models (ABMs) with network structure (Bruch & Mare, 2006; Clark W. A., 1991; Clark & Fossett, 2008; Fossett, 2006; Gandica, Gargiulo, & Carletti, 2016; Schelling, 1971; Xie & Zhou, 2012; Zhang, 2004a; Zhang, 2004b). Residential segregation itself appears to be a robust and resilient phenomenon with many possible contributing factors at multiple scales (Clark W. A., 1991). This work focuses primarily on one of the two traditions of segregation theory identified by Fossett (2006): “individual preferences,” which asserts that segregation is an emergent property arising from the collective dynamics of individuals and their choices in mostly-free housing markets. Identifying and understanding factors and behaviors which contribute to residential segregation has substantial sociopolitical significance. In the decades following Brown v. Board of Education, a legal debate surrounding de jure vs. de facto segregation developed which included substantial discussion of the role of individual preferences in the genesis of de facto segregation (Frankenberg & Taylor, 2018). It is clearly important to understand the extent to which de facto segregation might emerge as the result of the collective dynamics of individual preferences.

Recent work by Sayama & Yamanoi (2019) has highlighted the importance of heterogeneity in network models of social problems, showing clearly that certain emergent, system-level phenomena may only arise with sufficient heterogeneity of agent behaviors and/or characteristics. A review of the literature identified relatively few investigations of component heterogeneity in network models of segregation. Notably, Xie and Zhou (2012) and Gandica, Gargiulo, & Carletti (2016) examine heterogeneity of preferences and heterogeneity of topology, respectively. This work attempts to examine the impact of the combination of these two dimensions of heterogeneity in network models of segregation.

**2. Models**

**2.1 Baseline Model**

A baseline model was constructed using the NetworkX package in Python. The base model topology is a 32 by 32 regular lattice with edges connecting von Neumann neighborhoods and a closed boundary condition. As in Xie & Zhou (2012), 15% of nodes are reserved as excess housing. The remaining 85% of nodes are randomly assigned either a red or blue occupant. Following Schelling (1971), each node is assigned an identical tolerance threshold, , of 0.50 indicating a preference for at most 50% of neighbors of opposite color. The following procedure runs for 4000 timesteps:

At time, , the current neighborhood proportion of dissimilar neighbors for the th node in the th neighborhood, , is given as

where is the number of dissimilar neighbors and is the total number of neighbors in the th neighborhood. At each time step, , is calculated for all and the set of candidate transfers at time, , , is constructed. A node, , is randomly selected. For all vacant nodes, , neighborhood composition is calculated to create a list of candidate vacancies, , from which a destination node, is randomly selected. The objects at nodes and are then swapped: the occupant at the th node moves to the th node and leaves a vacancy in its place. If at any time, , , all nodes are satisfied, and no additional trades will be found. If at any time, , , the occupant at the selected node,, is unable to locate a satisfactory destination and remains in place.

**4.2 Heterogeneous Preferences**

Following Xie & Zhou (2012), agents are provided with heterogeneous preferences (tolerance thresholds) aligned with the Guttman scale and rank-ordered logit model derived from Bruch & Mare’s (2006) Detroit data. Tolerance thresholds were assigned by drawing values from a uniform distribution over a given interval: For 10.47% of individuals, fell within [0.0,0.07); for 18.10% of individuals, fell within [0.07,0.21); for 26.73% of individuals, fell within [0.21, 0.36); for 13.86% of individualsfell within [0.36,0.57); for 26.59% of individuals, fell within [0.057,1.00]. For these individuals, the simulation procedure described in the previous section was implemented with replacing . For the 4.25% of individuals for whom no tolerance threshold was set, Xie & Zhou’s (2012) rank-ordered logit model (eq. 4) was implemented to determine probability of transition to each candidate neighborhood. Transition destination, , is then randomly selected with probability, , given by the model:

where .

**4.3 Heterogeneous Topology**

Departing from Gandica, Gargiulo, & Carletti (2016), who implement a metapopulation model, primarily to make topological differences more explicit, we select random neighborhoods to densify. To do so, for each randomly selected node, a set of neighbors, , is constructed. For each , a set of neighbors, is constructed such that . A set of edges, is then drawn. A single iteration of this procedure is considered a single densification as shown in Fig. 1.

This method ensures that neighborhood densities follow a consistent gradient across G, so nodes and their neighbors cannot have unrealistic differences in degree. In other words, two individuals living in the same building ought to have the same number of neighbors. The result of multiple densifications is substantially greater variation in node degree across the network as well as a marked increase in the mean node degree. The result is a variety of neighborhood sizes. While in the base model, each neighborhood is a von Neumann neighborhood bordering another von Neumann neighborhood (except at the boundary), randomly densified lattices have a variety of neighborhood boundary relationships, e.g., a von Neumann neighborhood might be adjacent to a Moore neighborhood. This enables a richer diversity of neighbor relationships: Imagine two dissimilar neighbors, each with a 50% tolerance threshold. One neighbor (red) lives in a von Neumann neighborhood and the other (blue) lives in a Moore neighborhood. Red will have three additional neighbors and blue will have 7 additional neighbors. The function plotted below describes the probability that each will exceed their Chart, line chart

Description automatically generatedtolerance threshold given the probability of dissimilar neighbors in their neighborhood:

Figure . Probability of exceeding tolerance threshold.

As the network approaches near-perfect assortative mixing, the probability that individuals located in larger neighborhoods will exceed their tolerance threshold diminishes more rapidly. However, when the network approaches near-total disassortative mixing, the probability that individuals located in larger neighborhoods will exceed their tolerance threshold increases more rapidly. This behavior is maintained as long as tolerance thresholds are equal. Interestingly, this indicates that as segregation increases, it is more likely that the individual in the von Neumann neighborhood will vacate its place. This reinforces the position of the individual in the Moore neighborhood. As a result, tolerant individuals located in

**2. Results**

Ten unique simulation settings were used, half using uniform Schelling (1971) tolerance threshold assignments and half using heterogenous Xie & Zhou (2012) tolerance threshold assignments. For each of these groups, five separate batches of 100 simulations were conducted. Each batch had increasing numbers of densifications: 0, 32, 64, 96, 128, as shown in Fig. 2. As the number of densifications increased, substantial differences in degree distributions and mean node degree were observed, as shown in Fig. 5 and Table X.

For convenience, Newman’s (2003) assortativity coefficient, , is used as a measure of segregation levels during the simulations. Values of close to 1 indicate high levels of segregation while values of close to 0 indicate approximately random mixing. To account for vacant nodes on the network, the assortativity coefficient for the subgraph containing only occupied nodes is used. For all simulations where the derived subgraph consisted of multiple connected components, was calculated for each connected component and a weighted average was constructed. As anticipated, observed mean increase in assortativity was higher for all Schelling-tolerance simulation runs. As in Xie & Zhou (2012), heterogeneity of tolerances did, on their own, result in reduced segregation levels. Contrary to Gandica, Gargiulo, & Carletti’s (2016) observation, topological heterogeneity alone did produce some results with reduced segregation levels. Finally, as expected, the combination of both dimensions of heterogeneity resulted in progressive reductions in segregation.

New Results

**3. Discussion**

Results clearly show that network models of segregation which combine heterogeneity of tolerances with heterogeneity of topologies show substantially different behavior than those which employ only one dimension of heterogeneity. This result is important as it may have bearing on the important question raised previously: to what extent can the collective dynamics of individual preferences lead to residential segregation? These results indicate that such preferences may not contribute as much as previously thought to residential segregation when sufficient heterogeneity is considered. At minimum, these results indicate the necessity of representing both heterogeneity of tolerances and heterogeneity of topologies in network models of residential segregation, as the omission of one or the other will result in substantial loss.

Curiously, the introduction of heterogeneity of topology had the effect of increasing segregation levels in some simulation runs. These increases disappeared when heterogeneity of tolerances was introduced. As in Gandica, Gargiulo, & Carletti (2016), significant accumulation of vacant nodes in high-degree areas was observed when tolerances were assigned via the Schelling method. This clustering did not appear to take place when tolerances were assigned via the Xie & Zhou method. This may indicate an important difference in model behaviors, especially if such clustering is the source of aforementioned observed greater segregation levels. Further work is required to investigate these phenomena and establish any linkages.

The results collected thus far have been limited to simulations on static networks. Further work should include adaptive network simulation as well. Addition and removal of housing stock in response to individuals’ behaviors should be represented. Such changes could reveal important adaptive dynamics which could lead to novel mixing behaviors or patterns, such as self-organized migration of racial groups.

One important limitation of this work is that increases in mean degree were not independently evaluated. It is possible that diversity of degree distributions is less important than asserted. Further work is required to establish the extent of importance of degree distribution versus the importance of higher mean degree. While the topology diversification process used in this work was designed to reflect real-world spatial properties while maintaing the traditional lattice frame employed in previous network models of segregation, it would be worthwhile to continue experiments with empirical topology data. Such study could illuminate similarities or differences in real-world versus expected levels of segregation.

Finally, further work is required to understand the relationship between heterogeneity and mixing patterns. We might reasonably conclude that heterogeneity of tolerances in segregation models enables a latent property of heterogeneity of neighborhood sizes to be expressed. Then, as the latter is expressed, the effect of the former is enhanced. Revisiting Sayama & Yamanoi (2019), it is natural to wonder whether neighborhoods might be expressed as individuals with their own cultural tolerance parameters, and, further, whether such an adaptation might be useful in constructing network models of segregation.

**Acknowledgments**

Paste your acknowledgments here.

**References**

Paste your main manuscript references here. They should be listed in order of citation.

Example References:

1. J.-M. Neuhaus, L. Sticher, F. Meins, Jr., T. Boller, A short C-terminal sequence is necessary and sufficient for the targeting of chitinases to the plant vacuole. *Proc. Natl. Acad. Sci. U.S.A.* 88, 10362–10366 (1991).
2. E. van Sebille, M. Doblin, Data from “Drift in ocean currents impacts intergenerational microbial exposure to temperature.” Figshare. Available at [https://dx.doi.org/10.6084/m9.figshare.3178534.v2. Deposited 15 April 2016](https://dx.doi.org/10.6084/m9.figshare.3178534.v2.%20Deposited%2015%20April%202016).
3. A. V. S. Hill, “HLA associations with malaria in Africa: Some implications for MHC evolution” in Molecular Evolution of the Major Histocompatibility Complex, J. Klein, D. Klein, Eds. (Springer, 1991), pp. 403–420.

**Figures and Tables**

**A picture containing diagram

Description automatically generated**

**Figure 1.** A densified portion of the lattice.

<insert page break here>

A picture containing background pattern

Description automatically generated

**Figure 2.** Sample final graph images for 0, 32, 64, 96, and 128 densifications. Left, Xie & Zhou method; right, Schelling method. Blue and red nodes are individuals, grey nodes are vacancies.

<insert page break here>

**Chart

Description automatically generated**

**Figure 3.** Average change in assortativity over time. Oranges, Schelling method; blues, Xie & Zhou method.

<insert page break here>

**Figure 4.** Average change in assortativity over time, Schelling method.

<insert page break here>

**Chart, radar chart

Description automatically generated**

**Figure 5.** Impact of densifications on graph degree distribution.

<insert page break here>

**Chart, box and whisker chart

Description automatically generated**

**Figure 6.** Impact of densifications and method on vacancy degree centrality.

<insert page break here>

**Chart, box and whisker chart

Description automatically generated**

**Figure 7.** Impact of densifications on degree centrality amongst tolerant vs. intolerant individuals, Xie & Zhou method. Tolerant individuals are those with thresholds at or above 0.75 while intolerant individuals are those with thresholds at or below 0.25.

<insert page break here>

**Chart, box and whisker chart

Description automatically generated**

**Figure 8.** Impact of densifications and method on final number of connected components.

<insert page break here>

**Table 1.** Type or paste table title here. Paste table below the title.

<insert page break here>

**Table 2.** Type or paste table title here. Paste table below the title.