**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**

Behavioral and Topological Heterogeneities in Network Versions of Schelling’s Segregation Model

Will Deter, Hiroki Sayama

Center for Collective Dynamics of Complex Systems, Department of Systems Science and Industrial Engineering, Binghamton University, State University of New York

\*Paste corresponding author name(s) here.

**Email:**  [wdeter1@binghamton.edu](mailto:wdeter1@binghamton.edu), sayama@binghamton.edu

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, agent heterogeneity, topological heterogeneity

**This PDF file includes:**

Main Text

Figures 1 to X

Tables 1 to X

**Abstract**

Agent-based network models of residential segregation have been of persistent interest to various research communities since their origin with Thomas Schelling. Frequently, these models have sought to elucidate the extent to which the collective dynamics of individual preferences may cause segregation to emerge. This open question has sustained relevance in U.S. jurisprudence. Previous investigation of heterogeneity of behaviors (preferences) by Xie & Zhou (2012) has shown reductions in segregation on networks. Previous investigation of heterogeneity of topologies by Gandica, Gargiulo, & Carletti (2016) has shown no significant impact to observed segregation levels. Recent work by Sayama and Yamanoi (2020) has shown the importance of representing realistic heterogeneities in dynamical social network models. In this work, the necessity of concurrent representation of both behavioral and topological heterogeneities in network segregation models is examined. Extending the previous works, additional network simulations were conducted using both Xie & Zhou’s and Schelling’s preference models on 2D lattices with varied levels of densification to create topological heterogeneities (i.e., clusters, hubs). Results show a richer variety of outcomes, including novel differences in resultant segregation levels, fragmentation, and hub composition. Notably, with concurrent increased representations of heterogeneous preferences and heterogenous topologies, reduced levels of segregation and fragmentation emerge. Implications and areas for future study are discussed

.

Gandica, Y., Gargiulo, F., & Carletti, T. (2016). Chaos, Solitons, and Fractals, 90, 46-54.

Sayama, H., & Yamanoi, J. (2020). NetSci-X 2020 Proceedings, pp. 171-181.

Xie, Y., & Zhou, X. (2012). PNAS, 109(29), 11646-11651.

**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 (1-9). Residential segregation appears to be a robust and resilient phenomenon with many possible contributing factors at multiple scales (2). This work focuses primarily on one of the two traditions of segregation theory identified by Fossett (4): “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 (10). 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 (7) 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 (7) and Gandica, Gargiulo, & Carletti (5) examine heterogeneous preferences and hetergeneous topology, respectively. This work attempts to examine the impact of the combination of these two dimensions of heterogeneity in network models of residential 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 (7), 15% of nodes are reserved as excess housing. The remaining 85% of nodes are randomly assigned either a red or blue occupant. As in Schelling (6), each node is assigned an identical tolerance threshold. , indicating a preference for at most ~39.62% opposite colored neighbors. To draw a more direct comparison, the -value is obtained by taking the population mean tolerance threshold from agents generated as described in the next section. 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.

**2.2 Heterogeneous Preferences**

Following Xie & Zhou (7), agents are provided with heterogeneous preferences (tolerance thresholds) aligned with the Guttman scale and rank-ordered logit model derived from Bruch & Mare’s (1) 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 (7) 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 .

**2.3 Heterogeneous Topology**

Departing from Gandica, Gargiulo, & Carletti (5), 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. 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.

**3. Results**

Ten unique simulation settings were used, half using uniform Schelling (6) tolerance threshold assignments and half using heterogenous Xie & Zhou (7) 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. Observed mean degrees were 3.875, 7.150, 10.030, 12.485, and 14.729, respectively.

**3.1 Assortativity**

For convenience, Newman’s (12) 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:

where is the total number of occupied nodes and and are the assortativity coefficient and size of the th component, respectively. For connected components with homogenous composition, a value of 1 was assigned for . Results are shown in Fig 3 and summarized in Tables 1 & 2. As anticipated, observed mean final assortativity was higher for all Schelling-tolerance simulation runs. As in Xie & Zhou (7), heterogeneity of tolerances did, on their own, result in reduced assortativity. Consistent with Gandica, Gargiulo, & Carletti’s (5) observations, increasing topological heterogeneity alone did not produce results with reduced assortativity, in fact, this increased assortativity to increase in our simulations. Finally, as expected, the combination of both dimensions of heterogeneity resulted in progressive reductions in assortativity. The lowest levels of assortativity observed occurred with Xie & Zhou (7) preferences and 128 densifications. Interestingly, low, non-zero levels of densification resulted in increased assortativity when Schelling (6) preferences were used.

**3.2 Fragmentation**

The number of connected components (CCs) was observed throughout each simulation run as a measure of emergent fragmentation due to spatial organization of vacancies. For lower levels of densification (), greater average fragmentation was observed when Xie & Zhou (7) preferences were used. For higher levels of densification (), greater average fragmentation was observed when Schelling (6) preferences were used. These results are summarized in Fig 8 and Tables 1 & 2.

**3.3 Hub Composition**

Mean degree centrality for vacancies was calculated for each final graph and the batch average was obtained. On average, for all simulations, observed mean vacant node degree centrality was slightly higher than mean degree centrality for the graph. However, the substantial vacancy accumulations in highly dense housing areas noted by Gandica, Gargiulo, & Carletti (5) were not observed. No significant difference was observed between the accumulations of vacancies in Xie & Zhou preference simulations and Schelling preference simulations. This difference was most pronounced with 32 and 64 densifications. These results are shown in Fig 2 and summarized in Tables 1 & 2.

For simulations with Xie & Zhou preferences on densified graphs, on average, individuals in the group with the highest tolerance thresholds () had substantially higher degree centralities than those in the group with the lowest tolerance thresholds (). These results are shown in Figs 6 & 7 and summarized in Tables 1 & 2.

**4. Discussion**

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 to residential segregation as much as previously thought when sufficient heterogeneity is represented. 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.

The introduction of heterogeneity of topology had the effect of increasing assortativity in some Schelling-preference simulation runs. These increases disappeared when heterogeneity of tolerances was introduced. Slight accumulation of vacant nodes in densified neighborhoods was observed. Substantial differences, such as those in Gandica, Gargiulo, & Carletti (5), were not observed. This is likely a result of differences in network wiring. When Xie & Zhou (7) preferences were used, substantial accumulations of highly tolerant nodes appeared to fill highly dense neighborhoods. The observed self-organization of tolerant nodes in dense areas and intolerant nodes in sparse areas may also explain the difference in levels of fragmentation observed between densified models with Xie & Zhou (7) and Schelling (6) preferences.

The probability that an individual will exceed its tolerance threshold, can be given as a function of the probability, , that its nth neighbor will be dissimilar:

where is the individual’s tolerance threshold and is the neighborhood size. Since our models begin with a randomly mixed population where and increases monotonically, we need only consider cases where , as shown in Fig 9. Thus, when tolerance thresholds are equal, individuals in in smaller neighborhoods should be more likely to exceed their thresholds. Thus, vacancies in highly dense areas are more likely to be filled earlier in the simulation while vacancies in rural areas are more likely to be created concurrently. In this way, highly dense areas can act as “stores” of tolerance. In Schelling-preference models, tolerance is stored until a tipping point is reached at which the hub will begin to shed non-dominant color nodes. When combined with heterogeneity of preferences, these hubs can become “stores” of diversity as well as tolerance as highly tolerant nodes become highly stable in their positions.

The results collected thus far have been limited to simulations on static networks. Further work should include adaptive network simulations. 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. The topology diversification process used in this work was designed to reflect real-world spatial properties while maintaining 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 (7), 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**

1. Bruch, E. E., & Mare, R. D. (2006). Neighborhood choice and neighborhood change. The American journal of sociology, 112(3), 667-709. Retrieved from <https://doi.org/10.1086/507856>
2. Clark, W. A. (1991). Residential Preferences and Neighborhood Racial Segregation: A Test of the Schelling Segregation Model. Demography, 28(1), 1-19. Retrieved from <https://doi.org/10.2307/2061333>
3. Clark, W. A., & Fossett, M. (2008). Understanding the social context of the Schelling segregation model. Proceedings of the National Academy of Sciences - PNAS, 105(11), 4109-4114. Retrieved from <https://www.pnas.org/doi/full/10.1073/pnas.0708155105>.
4. Fossett, M. (2006). Ethnic Preferences, Social Distance Dynamics, and Residential Segregation: Theoretical Explorations Using Simulation Analysis. The Journal of Mathematical Sociology, 30(3-4), 185-273. Retrieved from <https://doi.org/10.1080/00222500500544052>
5. Gandica, Y., Gargiulo, F., & Carletti, T. (2016). Can topology reshape segregation patterns? Chaos, Solitons, and Fractals, 90, 46-54. Retrieved from <https://doi.org/10.1016/j.chaos.2016.02.024>
6. Schelling, T. (1971). Dynamic models of segregation. The Journal of mathematical sociology, 1(2), 143-186. Retrieved from <https://doi.org/10.1080/0022250X.1971.9989794>
7. Xie, Y., & Zhou, X. (2012). Modeling individual-level heterogeneity in racial residential segregation. Proceedings of the National Academy of Sciences - PNAS, 109(29), 11646-11651. Retrieved from <https://doi.org/10.1073/pnas.1202218109>
8. Zhang, J. (2004). A Dynamic Model of Segregation. The Journal of Mathematical Sociology, 28(3), 147-170. Retrieved from <https://doi.org/10.1080/00222500490480202>.
9. Zhang, J. (2004). Residential segregation in an all-integrationist world. Journal of Economic Behavior & Organization, 54(4), 533-550. Retrieved from <https://doi.org/10.1016/j.jebo.2003.03.005>.
10. Frankenberg, E., & Taylor, K. (2018). De Facto Segregation: Tracing a legal basis for contemporary inequality. Journal of Law and Education, 47(2), 189-233. Retrieved from <https://www.proquest.com/scholarly-journals/de-facto-segregation-tracing-legal-basis/docview/2041728504/se-2>.
11. Sayama, H., & Yamanoi, J. (2019). Beyond Social Fragmentation: Coexistence of Cultural Diversity and Structural Connectivity Is Possible with Social Constituent Diversity. Retrieved from <https://doi.org/10.48550/arxiv.1911.04381>
12. Newman, M. E. (2003). Mixing patterns in networks. Physical review. E, Statistical, nonlinear, and soft matter physics, 67(2 pt 2), 026126–026126. Retrieved from <https://doi.org/10.1103/PhysRevE.67.026126>

**Figures and Tables**

**A picture containing diagram

Description automatically generated**

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

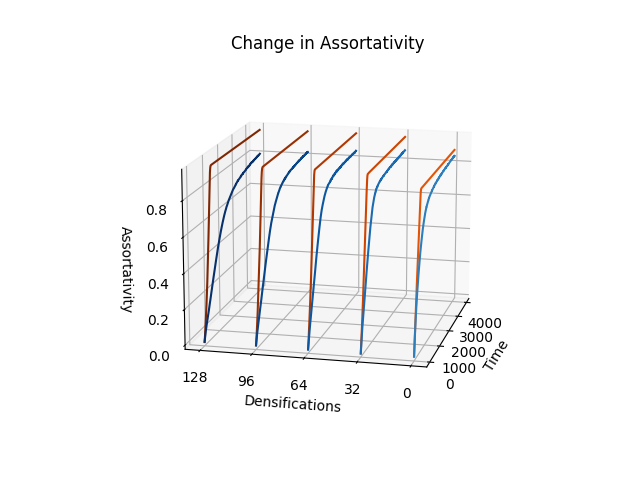
<insert page break here>

A picture containing background pattern

Description automatically generated

**Figure 2.** Sample final graph images for 0, 32, 64, 96, and 128 densifications using Xie & Zhou preferences. Redder nodes are less tolerant, greener nodes are more tolerant.

<insert page break here>

****

**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>

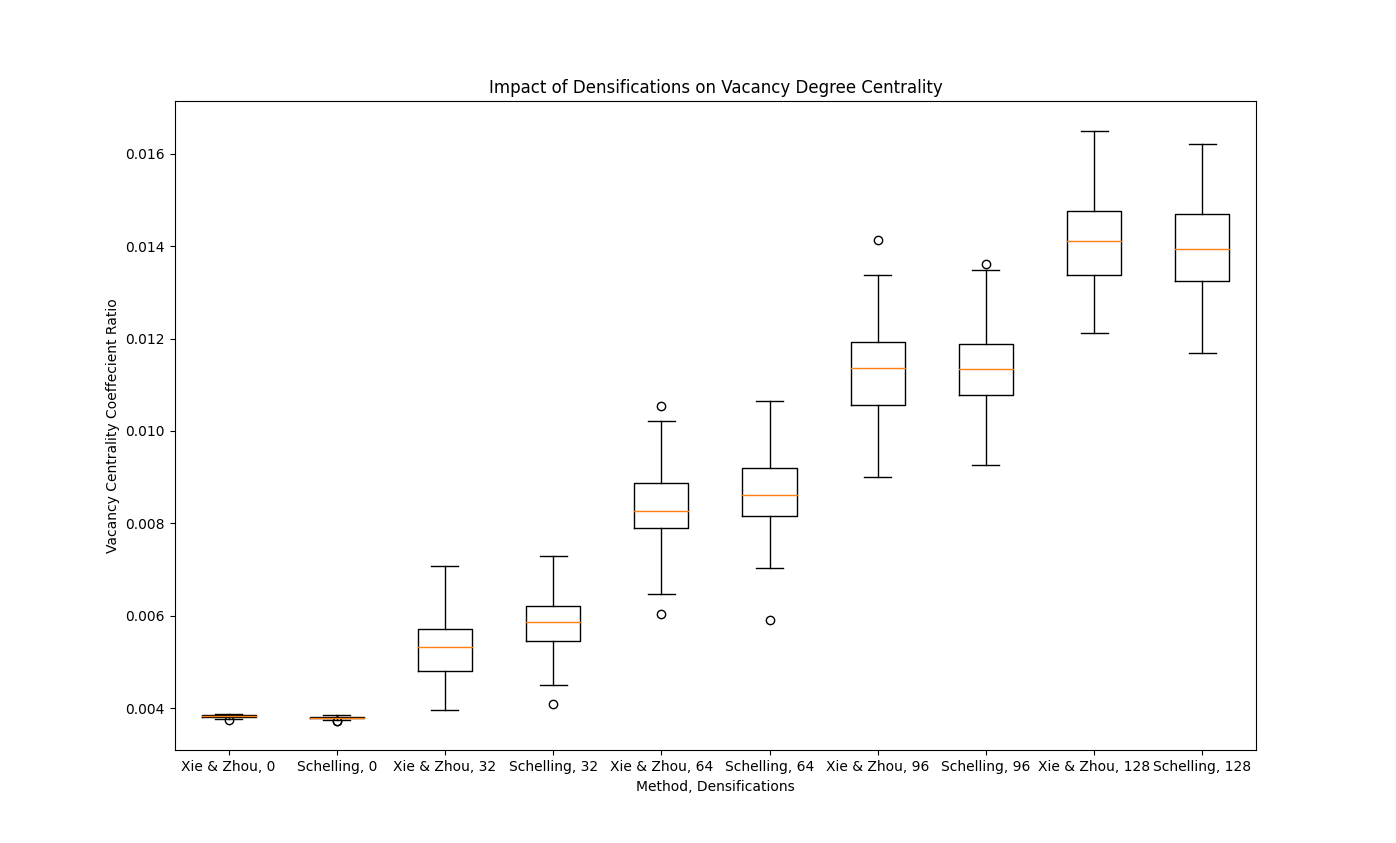
**Chart, radar chart

Description automatically generatedChart, scatter chart

Description automatically generated**

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

<insert page break here>

****

**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>

Chart, line chart

Description automatically generated

**Figure 9.** Vertical axis: , probability of exceeding tolerance threshold. Horizonal axis: probability of nth dissimilar neighbor. Moore (blue) neighborhood where and von Neumann (red) neighborhood where . .

<insert page break here>

**Table 1.** Results, Xie & Zhou preferences.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| D |  |  |  |  |  |  |  |  |  |  |
| 0 | 0.8452 | 0.0260 | 2.7400 | 1.4807 | 0.0038 | 0.0000 | 0.0038 | 0.0000 | 0.0037 | 0.0000 |
| 32 | 0.8645 | 0.0278 | 2.8600 | 1.5364 | 0.0053 | 0.0006 | 0.0069 | 0.0004 | 0.0050 | 0.0004 |
| 64 | 0.8465 | 0.0297 | 2.4900 | 1.2609 | 0.0084 | 0.0008 | 0.0101 | 0.0005 | 0.0063 | 0.0006 |
| 96 | 0.8258 | 0.0436 | 2.0700 | 1.1769 | 0.0113 | 0.0010 | 0.0127 | 0.0006 | 0.0078 | 0.0008 |
| 128 | 0.8009 | 0.0528 | 1.6100 | 0.8111 | 0.0141 | 0.0009 | 0.0151 | 0.0005 | 0.0094 | 0.0013 |

<insert page break here>

**Table 2.** Results, Schelling preferences..

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| D |  |  |  |  |  |  |
| 0 | 0.8804 | 0.0162 | 2.9400 | 1.4548 | 0.0038 | 0.0000 |
| 32 | 0.9427 | 0.0103 | 2.6700 | 1.2414 | 0.0058 | 0.0006 |
| 64 | 0.9503 | 0.0109 | 2.3900 | 1.0188 | 0.0087 | 0.0008 |
| 96 | 0.9488 | 0.0126 | 2.1600 | 1.1110 | 0.0114 | 0.0009 |
| 128 | 0.9438 | 0.0144 | 1.8600 | 0.9276 | 0.0140 | 0.0010 |