complexities

MDPI

Article

Beyond Shannon Entropy in Schelling Segregation: Shape-Weighted Diagnostics of Bimodality, Polarization, and Economic Freezing

George-Rafael Domenikos [1,*] and Lock Yue Chew [2,*]

[1] Early Mental Potential & Wellbeing Research (EMPOWER) Centre, Nanyang Technological University, Singapore 308232, Singapore
[2] School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore
* Correspondence: georgios.rd@ntu.edu.sg (G.-R.D.); lockyue@ntu.edu.sg (L.Y.C.)

Abstract

Generalized Schelling models extend the classical relocation framework by introducing additional agent attributes, interaction rules, and mobility constraints. We study four variants within a common lattice formulation: a modified movement-only baseline, a local affiliation-interaction model, a model with local and random long-range interactions, and a monetary model in which movement and communication consume resources. The simulations use a 64×64 lattice and are evaluated over 30 independently seeded runs per scenario, with means and 95% confidence intervals reported throughout. At each step, occupied agents are classified as bound, activated, or, in the monetary scenario, trapped. Their conditional affiliation distributions are evaluated using Shannon entropy and a shape-weighted diagnostic whose fitted components distinguish normal-like, heavy-tailed, and bimodal structures. The diagnostic is interpreted jointly with its fitted weights and the underlying histograms rather than as a unique scalar reconstruction of shape. Validation against the bimodality coefficient, Esteban–Ray polarization, Moran's I, and sign assortativity separates distributional polarization from spatial segregation. The interaction-driven scenarios develop strongly polarized two-lobed bound-state distributions, while the monetary model additionally produces resource-constrained dynamical arrest. The results are robust across independent initial conditions and random streams; separate extended-horizon checks to 500 steps showed no qualitative reversal. The same regimes persist on 96×96 and 128×128 lattices and under one-factor variations of satisfaction, vacancy, long-range-contact, and monetary-cost parameters. The framework therefore complements Shannon entropy by exposing distributional geometry that uncertainty alone does not encode.

Keywords: Schelling model; agent-based modeling; generalized Schelling dynamics; segregation; opinion dynamics; polarization; bimodality; Shannon entropy; monetary constraints

Check for updates

Academic Editor: Jarosław Kwapień

Received: 15 June 2026
Revised: 16 September 2026
Accepted: 17 September 2026
Published: 21 September 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Schelling's segregation model is a foundational example of how individually mild local preferences can generate macroscopically segregated patterns [1–3]. Related checkerboard work and later agent-based studies established lattice simulation as a general tool for linking individual rules to collective outcomes [4–6]. Subsequent analyses have examined

the role of thresholds, density, system size, implementation choices, physical analogues, networks, and heterogeneous populations [7–16].

In this study, we compare four generalized Schelling models, referred to as Scenarios 1–4. All four place agents on a partially occupied square lattice, assign each agent a signed affiliation, and allow dissatisfied agents to relocate to nearby vacant sites. Scenario 1 is a modified movement-only baseline in which affiliations remain fixed. Scenario 2 adds local interactions that update an agent's affiliation using the signs of its neighbors' affiliations. Scenario 3 retains these local interactions and adds transient contacts sampled randomly from the occupied population without a spatial-distance restriction. Scenario 4 retains both interaction mechanisms but assigns finite monetary resources to agents: long-range contacts and relocation incur costs and can occur only when affordable, whereas local interaction remains free. These nested scenarios separate the effects of relocation, local affiliation reinforcement, nonlocal exposure, and resource constraints. Their detailed rules are specified in Section 2.

The Schelling literature is not restricted to the immutable binary racial labels of the original illustration. Published extensions explicitly introduce monetary resources and economic mobility constraints [17–20]; continuous agent types that can represent attributes such as income or position on a political left–right axis [21]; switching agents [22]; Schelling–Voter hybrids in which an unhappy agent may change either location or state [23]. Accordingly, endogenous affiliation updates do not, by themselves, make the term "generalized Schelling dynamics" inappropriate. Scenarios 2–4 retain the defining spatial structure of a Schelling model—local satisfaction, vacancies, and relocation—while adding an opinion-like state variable. We therefore retain the generalized Schelling terminology and clarify that these scenarios also overlap with the broader opinion-dynamics literature [24–27].

Scenario 1 is a deliberately simple movement-only baseline. Its fixed same-sign count criterion is not identical to Schelling's original non-minority criterion when vacancies are present. For example, an agent with three same-sign neighbors, two opposite-sign neighbors, and three vacancies is not in the minority, but may fail the fixed-count criterion used here depending on the selected threshold. We therefore describe Scenario 1 as a modified Schelling baseline rather than an exact reproduction of the original rule.

The analytical focus is the evolving distribution of a signed affiliation variable. Shannon entropy is a foundational uncertainty measure [28,29], and Jaynes' maximum-entropy framework links entropy to inference under constraints [30,31]. A Shannon value does not, however, identify whether a finite-support histogram is normal-like, heavy-tailed, or two-lobed. Those structures may have comparable uncertainty while representing different dynamical regimes. We therefore use a shape-weighted diagnostic based on comparisons with three reference families. The scalar diagnostic is interpreted jointly with its fitted weights and the conditional histograms; it is not claimed that one number uniquely reconstructs a distribution.

This distributional question is related to, but distinct from, polarization and spatial segregation. The bimodality coefficient tests whether a one-dimensional distribution has a sufficiently low-kurtosis/two-peaked form, while the Esteban–Ray index measures identification within affiliation groups together with alienation between them [32,33]. Neither quantity uses the agents' coordinates. Conversely, Moran's I and categorical assortativity use spatial adjacency to quantify clustering of similar values or signs [34,35]. A globally bipolar affiliation histogram can therefore be spatially mixed, and a spatially segregated lattice can retain an unchanged marginal affiliation histogram. The proposed diagnostic is intended to describe the first object; the spatial indices are used below as independent validation rather than as alternative names for the same quantity.

The contribution is two-fold. First, the paper compares four generalized Schelling variants that progressively add affiliation interaction, nonlocal communication, and monetary constraints. Second, it introduces a computationally simple decomposition that identifies whether normal-like, heavy-tail, or bimodal structure carries the entropy response. Scenario 1 is retained because a controlled movement-only baseline is essential for determining which effects arise only when affiliation dynamics are activated. Its simplicity is intentional rather than a claim that every variable must affect every scenario.

The study is organized around four explicit research questions:

RQ1. Does Shannon entropy alone distinguish the distributional reorganization produced by affiliation interaction and monetary constraints?

RQ2. Does the bimodality component of the shape-weighted diagnostic agree with established, independently computed measures of bimodality and polarization?

RQ3. When affiliation interaction is active, does the distributional reorganization coevolve with actual spatial segregation measured by standard spatial indices, and where does that correspondence break down?

RQ4. Are the conclusions robust to the diagnostic thresholds, larger lattices, and variations of the principal satisfaction, vacancy, long-range-contact, and monetary-cost parameters?

2. Methodology

The model separates agent dynamics from the statistical diagnostics applied to the resulting configurations. This section defines the lattice and agent variables, the rules for satisfaction, interaction and relocation, and the classification used to construct conditional affiliation histograms. Section 3 then defines the shape-weighted diagnostic applied to those histograms, and Section 4 specifies the simulation ensemble and robustness protocol. A simulation step is one complete sweep of the interaction and movement rules; it is an algorithmic time unit rather than a calibrated physical duration. The initial configuration is indexed by $t = 0$, and subsequent integer times denote configurations saved after complete sweeps.

2.1. Lattice, Variables, and Initialization

The model is implemented in Python 3.10 on a finite $N \times N$ square lattice with coordinates $(i, j) \in \{1, \ldots, N\}^2$. Each site is either vacant or occupied by one agent. The site-indexed notation $C_{ij}(t)$ and $M_{ij}(t)$ is an Eulerian representation: it records the affiliation and monetary resource carried by the agent occupying site (i, j) at time t. When an agent moves, both quantities move with the agent. A vacant cell is represented computationally by `NaN`; $C = 0$ is a valid occupied neutral affiliation state.

The sign of C_{ij} identifies one of two opposing affiliations, while $|C_{ij}|$ records its magnitude; $C_{ij} = 0$ denotes neutrality. The variable M_{ij} is a non-negative resource balance measured in the same arbitrary units as movement and communication costs. It is initialized in all scenarios, although it affects the dynamics only in Scenario 4. The lattice has open boundaries: opposite edges are not connected, and neither interaction nor movement wraps around the lattice.

A fraction $v = 0.30$ of sites is selected uniformly without replacement as vacant. On occupied sites,

$$C_{ij}(0) \sim \mathcal{U}_{\mathrm{d}}\{-10, -9, \ldots, 0, \ldots, 9, 10\}, \tag{1}$$

and

$$M_{ij}(0) \sim \mathcal{U}[5, 15). \tag{2}$$

Here, $\mathcal{U}_{\mathrm{d}}$ denotes a discrete uniform distribution, assigning probability $1/21$ to each integer affiliation, and $\mathcal{U}[5, 15)$ denotes a continuous uniform distribution on the indicated resource interval. Vacancy positions, affiliations, and resources use independent random streams. The implementation selects $\mathrm{round}(vN^2)$ vacant sites, so $N_{\mathrm{occ}} = N^2 - \mathrm{round}(vN^2)$ agents are present. Relocation conserves this number throughout a run: agents neither enter nor leave the system.

The affiliation variable is deliberately integer-valued. Discrete-state lattice formulations are foundational to the original Schelling model and remain widely used in subsequent analytical and computational studies [2,7,12,13,15,16]. The 21-level state space used here generalizes the conventional binary group label while retaining bounded finite support, directly enumerable empirical probabilities, and an unambiguous discrete support for the shape-weighted entropy analysis. A continuous representation would require an additional binning or density-estimation scale and is not required for the distributional comparisons considered here.

2.2. Neighborhood, Satisfaction, and Movement

Let $\mathcal{N}_{ij}$ be the valid Moore neighborhood of site (i, j): the surrounding sites whose horizontal and vertical offsets are each at most one, excluding the focal site itself. There are eight such sites in the interior, five on an edge, and three at a corner. Vacant sites contribute neither agreement nor signed influence. A neutral focal agent has no same-sign neighbors by definition, and neutral neighbors do not count as agreeing with either sign. The number of agreeing neighbors is

$$a_{ij}(t) = \sum_{(k,l)\in\mathcal{N}_{ij}} \mathbf{1}\big[C_{ij}(t)C_{kl}(t) > 0\big]. \tag{3}$$

The indicator $\mathbf{1}[\cdot]$ equals one when its condition holds and zero otherwise. The product condition counts two occupied agents as agreeing precisely when both affiliations have the same nonzero sign. Agreement therefore depends on sign, not magnitude. An agent is satisfied when $a_{ij}(t) \geq h$, with $h = 4$. This fixed same-sign threshold differs from Schelling's original non-minority criterion and is therefore described as a modified Schelling rule.

The same absolute threshold is used at boundaries; it is not divided by the number of available neighbors. Consequently, at the baseline $h = 4$, a corner agent cannot be satisfied because it has only three neighboring sites.

After the interaction stage described below, occupied sites are reshuffled, and dissatisfied agents are processed sequentially. Each dissatisfied agent identifies vacant sites within Manhattan distance $R_{\mathrm{move}} = 4$ and selects one uniformly at random. If no candidate exists, the agent remains in place. In Scenario 4, an actual move is executed only when $M_{ij} \geq X_{\mathrm{move}} = 1$; the cost is deducted, and C_{ij} and M_{ij} move together.

The Manhattan distance from (i, j) to (k, l) is $|k - i| + |l - j|$. A candidate destination must lie inside the lattice, be vacant at the time of the move, and satisfy this distance bound. Destinations are not screened for improved satisfaction. Satisfaction and the vacancy set are evaluated on the current configuration as movement proceeds so that an earlier relocation can affect a later decision. The movement schedule is a shuffled list of occupied coordinates recorded at the start of this stage; coordinates that have become empty are skipped. The origin becomes vacant after a successful move. In Scenario 4, affordability is checked using the balance remaining after communication costs, and a movement cost is charged only for an executed relocation.

2.3. Affiliation Interactions and Exact Update Order

Scenarios 2–4 include synchronous affiliation interactions. From copies C^{prev} and M^{prev} made at the start of the sweep, the local signed influence is

$$I_{ij}^{\text{local}}(t) = \sum_{(k,l)\in\mathcal{N}_{ij}} \text{sign}\Big(C_{kl}^{\text{prev}}(t)\Big), \tag{4}$$

where vacancies are omitted and $\text{sign}(c)$ is -1, 0, or $+1$ according to whether c is negative, zero, or positive. The superscript prev denotes the configuration before any affiliation changes in the current sweep. All occupied agents participate in local interaction, including satisfied agents and those unable to afford movement. Thus, the update is the full signed sum of occupied neighbors, not merely a one-unit step.

Scenarios 3 and 4 additionally include $n_{\text{call}} = 2$ transient long-range contacts per occupied agent and sweep. Each contact is an independent uniform draw from the occupied population, with replacement; the contact set is redrawn at every sweep. The distant influence is

$$I_{ij}^{\text{distant}}(t) = \sum_{r=1}^{2} \text{sign}\Big(C_{K_r L_r}^{\text{prev}}(t)\Big). \tag{5}$$

Here, (K_r, L_r) is the coordinate selected for contact r. Sampling from the full occupied population permits repeated partners, local neighbors, and the focal agent itself. The term "long-range" therefore describes the absence of a spatial-distance restriction rather than a requirement that each sampled partner be distant. A contact contributes the sampled partner's sign to the focal agent's update; it does not impose a reciprocal update on that partner.

Networked Schelling formulations are well established and can introduce additional effects associated with degree heterogeneity, clustering, link persistence, and network topology [14]. The present study instead uses annealed random mixing as a controlled nonlocal-interaction mechanism: distant contacts are sampled independently at each sweep, so the comparison between Scenarios 2 and 3 isolates the effect of transient long-range exposure without adding persistent network topology or memory as further model parameters. The conclusions concerning distant interaction are therefore limited to this specified mixing rule. In Scenario 4, the pair of contacts is executed only when the agent can pay the combined cost $2X_{\text{call}} = 0.50$, where $X_{\text{call}} = 0.25$ per contact; the combined amount is deducted once before the two contacts.

Affiliation is then updated synchronously as

$$C_{ij}(t^+) = \text{clip}_{[-10,10]}\Big[C_{ij}(t) + I_{ij}^{\text{local}}(t) + I_{ij}^{\text{distant}}(t)\Big], \tag{6}$$

where t^+ denotes the intermediate configuration after interaction but before movement. The clipping operator replaces values above 10 by 10 and values below -10 by -10, leaving values inside the interval unchanged. In Scenario 2, the distant term is zero; Scenario 1 skips the affiliation update entirely. In Scenario 4, an agent unable to afford the complete contact set makes no paid contacts in that sweep, but still receives local influence. Resources are not replenished, and local interaction carries no cost. Movement is performed only after this interaction update. The exact order within each sweep is therefore synchronous interaction, affiliation clipping, reshuffling, and sequential movement.

2.4. Four Scenarios

Table 1 summarizes the four model configurations defined by the preceding rules. In Scenario 1, affiliation values remain fixed and dissatisfied agents can relocate to available vacant sites without resource costs. Scenario 2 retains this relocation rule and adds syn-

chronous affiliation updates from local neighbors. Scenario 3 retains both mechanisms and additionally samples two contacts per agent and sweeps from the occupied population without a spatial restriction. Scenario 4 uses the same local and nonlocal interaction rules, but long-range contacts and relocation require sufficient resources and deduct their respective costs; local interaction remains free. Monetary balances therefore constrain the dynamics only in Scenario 4. All four use the same initialization distributions, boundaries, satisfaction criterion, and movement radius. Thus, the successive comparisons distinguish the effects of local interaction, nonlocal exposure, and resource depletion within a common model framework.

Table 1. Simulation scenarios.

Scenario	Name	Mechanism
1	Modified movement-only baseline	Fixed affiliation values; dissatisfied agents relocate.
2	Local affiliation interaction	Scenario 1 plus synchronous local signed affiliation influence.
3	Local and long-range interaction	Scenario 2 plus two transient uniformly sampled long-range contacts per occupied agent and sweep.
4	Monetary generalized Schelling model	Scenario 3 with movement and communication costs and resource-constrained trapping.

2.5. Microstates and Conditional Affiliation Distributions

At each saved configuration, occupied sites are assigned one of three mutually exclusive labels, $X \in \{B, A, T\}$. Here, "microstate" denotes an agent's current satisfaction and mobility category, rather than the complete microscopic configuration of the lattice. An occupied agent is bound when it is satisfied. A dissatisfied agent is activated when it can pay the movement cost and trapped when it cannot:

$$X_{ij}(t) = \begin{cases} B, & a_{ij}(t) \geq 4, \\ A, & a_{ij}(t) < 4 \text{ and } M_{ij}(t) \geq 1, \\ T, & a_{ij}(t) < 4 \text{ and } M_{ij}(t) < 1. \end{cases} \tag{7}$$

The displayed thresholds are the baseline values $h = 4$ and $X_{\text{move}} = 1$; classification uses the corresponding configured values in parameter sweeps. Satisfaction takes precedence over affordability: a satisfied agent is classified as bound even when its balance is low. An activated agent is financially eligible to move but need not actually relocate, for example, if no nearby vacancy exists. A trapped agent is dissatisfied and financially unable to move; the label does not imply that its affiliation cannot subsequently change through local interaction. Labels are recomputed after each complete sweep and are not permanent agent identities.

Money is not depleted in Scenarios 1–3, so $q_T(t) = 0$ identically in those scenarios. If J_t is a site sampled uniformly from the occupied sites, then

$$q_X(t) = \Pr[X_{J_t}(t) = X] = \frac{N_X(t)}{N_{\text{occ}}(t)}. \tag{8}$$

Here, $N_X(t)$ counts agents in category X, and $N_{\text{occ}}(t)$ counts all occupied sites. The subscript J_t denotes the uniformly sampled occupied location. The fractions satisfy $q_B + q_A + q_T = 1$ and exclude vacancies from their denominator. They describe the composition of the occupied population, whereas the distributions defined next describe affiliation within each category.

For every nonempty microstate,

$$p_X(c;t) = \frac{n_X(c;t)}{N_X(t)}, \qquad c \in \{-10, -9, \ldots, 0, \ldots, 9, 10\}. \tag{9}$$

The count $n_X(c;t)$ is the number of agents that simultaneously have category X and affiliation c, so $N_X(t) = \sum_c n_X(c;t)$. Consequently, $p_X(c;t)$ is the empirical conditional probability of affiliation c given category X, and sums to one for each nonempty category. The common support $\mathcal{C} = \{-10, \ldots, 10\}$ includes all possible affiliation levels, including those with zero observed count.

The label n in a histogram subtitle is $N_X(t)$, the number of occupied agents contributing to that conditional histogram. The smoothed distribution is

$$\widetilde{p}_X(c;t) = \frac{n_X(c;t) + \alpha}{N_X(t) + 21\alpha}, \qquad \alpha = 0.5. \tag{10}$$

The constant α is a pseudo-count added to each of the 21 affiliation bins. The denominator normalizes the augmented counts to one. This symmetric Dirichlet smoothing makes every bin positive in a nonempty category, permitting finite logarithms in the subsequent entropy and divergence calculations. It modifies the conditional histogram used for the diagnostic, but does not modify the simulated affiliations, the category counts, or the fractions q_X. Its influence is greater when N_X is small, which is why conditional sample sizes are reported alongside the histograms.

An empty microstate is neither smoothed nor fitted and contributes zero through $q_X = 0$. Dirichlet smoothing, therefore, cannot create a trapped distribution when the trapped class is absent.

3. Shape-Weighted Entropy Diagnostic

3.1. Motivation

Shannon entropy measures the uncertainty or spread of a probability distribution,

$$H_{\mathrm{Sh}}(p) = -\sum_c p(c) \log p(c). \tag{11}$$

This makes it a natural baseline for quantifying disorder in agent-based systems. However, Shannon entropy is insensitive to the specific geometric form of the distribution beyond its probability weights. In particular, two distributions may have comparable spread while representing very different structural states of the system.

This issue is central in segregation dynamics. The transition of interest is not simply a loss of disorder. Instead, the system often evolves from a broad, approximately unimodal, or normal-like affiliation distribution toward a polarized two-lobed distribution. In such a state, probability mass accumulates near the two extremes of the affiliation axis, while the center becomes depleted. This is not merely a reduction in uncertainty; it is the formation of two competing affiliation basins.

Figure 1 illustrates this distinction using two microstate-conditioned affiliation histograms from the simulation. The activated population retains a broader, more normal-like shape, whereas the bound population exhibits stronger edge accumulation and a clearer bimodal tendency. Both distributions are defined on the same affiliation support, but their structural interpretation is different. This motivates the use of an entropy diagnostic that does not only measure spread, but also identifies whether the observed distribution is better characterized by normality, heavy tails, or bimodality.

We therefore introduce a shape-weighted entropy that combines three structural contributions:

1. a normality or Shannon-like bulk contribution;
2. a heavy-tail contribution, capturing probability mass near extreme affiliation values;
3. a bimodality contribution, capturing the emergence of two polarized lobes separated by a depleted center.

The contribution of each term is determined by data-driven weights computed at each simulation step. The scalar value alone does not identify the histogram shape. It is interpreted jointly with the fitted weight vector (w_N, w_T, w_B), the microstate population N_X, and the conditional histogram. The term "entropy" is retained because the diagnostic contains concave Shannon and Tsallis cores, but the full data-adaptive quantity is used as a descriptive score rather than as a thermodynamic state function.

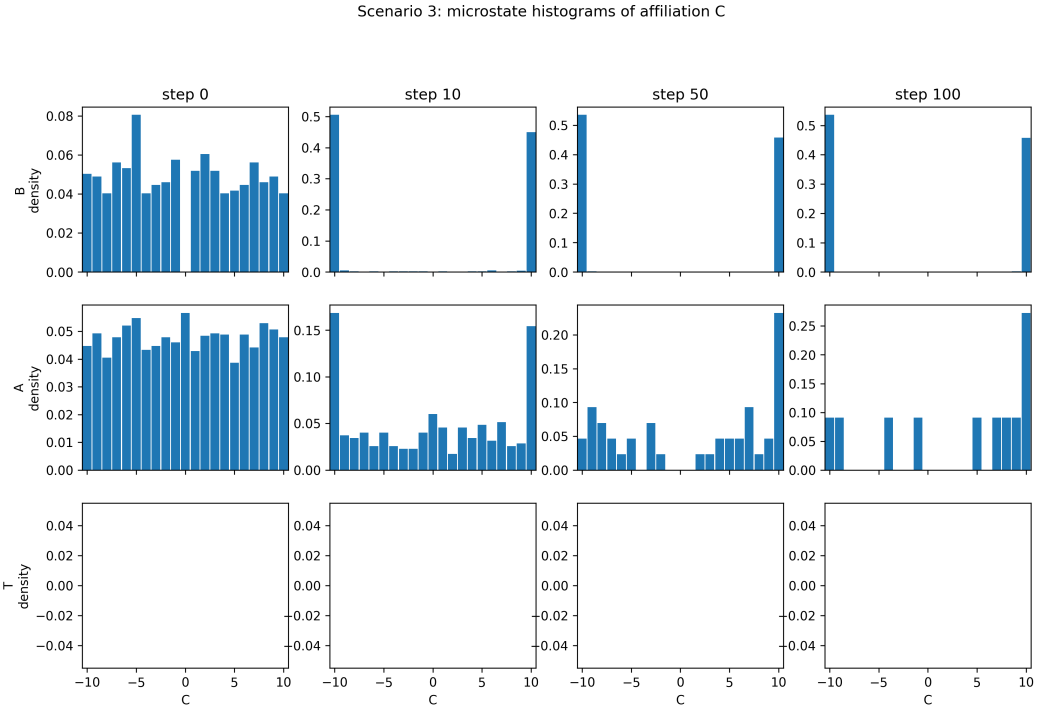


Figure 1. Representative Scenario 3 microstate-conditioned affiliation histograms. The bound population becomes strongly two-lobed, whereas the late activated population is small and its normalized shape is correspondingly sample-sensitive. The trapped panels are empty because $q_T = 0$ in Scenario 3. The title value n is the number of agents in the displayed conditional histogram.

The weights of these contributions are learned directly from the empirical microstate-conditioned histograms at each simulation step. The resulting entropy therefore functions both as a scalar diagnostic and as an interpretable decomposition of the distributional structure of the system.

3.2. Reference Distributions

For each microstate histogram $\tilde{p}_X(c;t)$, we compare the empirical distribution with three reference shapes defined on the same discrete support $\mathcal{C}$.

The first reference is a discrete Gaussian distribution,

$$r_N(c;\sigma) = \frac{\exp\left(-c^2/(2\sigma^2)\right)}{Z_N(\sigma)}, \tag{12}$$

where

$$Z_N(\sigma) = \sum_{c \in \mathcal{C}} \exp\left(-c^2/(2\sigma^2)\right). \tag{13}$$

Here, $\sigma > 0$ is the width, or standard-deviation, parameter of the discrete Gaussian reference.

The second reference is a discrete heavy-tailed distribution, represented by a Student-type shape,

$$r_T(c;\nu) = \frac{\left(1 + c^2/\nu\right)^{-(\nu+1)/2}}{Z_T(\nu)}, \tag{14}$$

where

$$Z_T(\nu) = \sum_{c \in \mathcal{C}} \left(1 + c^2/\nu\right)^{-(\nu+1)/2}. \tag{15}$$

Here, $\nu > 0$ is the degrees-of-freedom parameter controlling the tail thickness of the Student-type reference; smaller ν yields heavier tails.

The third reference is a symmetric bimodal distribution:

$$r_B(c;\mu,\sigma) = \frac{\frac{1}{2}\exp\left[-\frac{(c-\mu)^2}{2\sigma^2}\right] + \frac{1}{2}\exp\left[-\frac{(c+\mu)^2}{2\sigma^2}\right]}{Z_B(\mu,\sigma)}, \tag{16}$$

with normalization

$$Z_B(\mu,\sigma) = \sum_{c \in \mathcal{C}} \left[\frac{1}{2}\exp\left(-\frac{(c-\mu)^2}{2\sigma^2}\right) + \frac{1}{2}\exp\left(-\frac{(c+\mu)^2}{2\sigma^2}\right)\right]. \tag{17}$$

Here, $\mu > 0$ is the location/mean parameter of each symmetric component, so the two lobes are centered near $\pm\,\mu$, and $\sigma > 0$ is the standard-deviation/width parameter of each component.

The Gaussian reference captures early unimodal structure, the heavy-tailed reference captures mass accumulation near extreme affiliations, and the bimodal reference captures the emergence of polarized two-basin structure.

For each histogram, the Gaussian width is $\sigma = \max(\sigma_{\text{emp}}, 0.5)$. The Student reference is selected from $\nu \in \{2, 3, 5, 10\}$ by minimum KL divergence. The equal-weight symmetric bimodal reference uses $\mu = \max(1, |c_{\text{mode}}|)$ and component width $\max(\sigma, 1)$. All references are normalized on the same discrete support.

These three families were selected to form a minimal, interpretable comparison rather than an exhaustive density-model search. The centered discrete Gaussian supplies a light-tailed unimodal reference; the Student family changes tail thickness without adding asymmetry; the equal-weight symmetric Gaussian mixture is the lowest-parameter two-lobe reference compatible with the sign symmetry of the initialization and update rules. The empirical width and mode set their scales, the same candidate set is used for every scenario and time step, and no reference parameter is optimized against the reported outcome. Agreement with external bimodality and polarization indices is evaluated in Section 5.3 so that the interpretation does not rest solely on these selected families.

3.3. KL-Derived Shape Weights

For each empirical distribution $\tilde{p}_X(c;t)$, we compute the Kullback–Leibler divergence to each reference shape:

$$d_k^X(t) = D_{\text{KL}}(\tilde{p}_X(\cdot;t) \| r_k), \qquad k \in \{N, T, B\}. \tag{18}$$

We also compute the divergence to the uniform distribution,

$$u(c) = \frac{1}{|\mathcal{C}|}, \tag{19}$$

given by

$$d_U^X(t) = D_{\mathrm{KL}}(\tilde{p}_X(\cdot;t)\|u). \tag{20}$$

The relative fit of each reference shape is then defined as

$$F_k^X(t) = \max\left[0, 1 - \frac{d_k^X(t)}{d_U^X(t) + \epsilon}\right], \tag{21}$$

where $\epsilon = 10^{-12}$ is used only to prevent division by zero. These fit scores are normalized to obtain the shape weights,

$$w_k^X(t) = \frac{F_k^X(t)}{\sum_{\ell\in\{N,T,B\}} F_\ell^X(t)}, \qquad k \in \{N, T, B\}. \tag{22}$$

If all fit scores vanish, the weights are set to $1/3$.

The weights satisfy

$$w_N^X(t) + w_T^X(t) + w_B^X(t) = 1, \tag{23}$$

and provide a step-resolved description of which distributional structure best explains each microstate histogram. Thus, w_N measures the normal or Shannon-like contribution, w_T measures the heavy-tail contribution, and w_B measures the bimodal contribution.

3.4. Tail and Bimodality Masks

To explicitly quantify the location of probability mass, we introduce two fixed masks on the affiliation support. The tail mask is

$$m_{\mathrm{tail}}(c) = \mathbf{1}[|c| \geq c_{\mathrm{tail}}], \tag{24}$$

where $c_{\mathrm{tail}} = 7$, so only the two outer affiliation levels on either side are treated as extreme.

The bimodality mask is defined as

$$m_{\mathrm{bi}}(c) = \begin{cases} +1, & |c| \geq c_0, \\ -\kappa, & |c| < c_0, \end{cases} \tag{25}$$

where $c_0 = 4$ and $\kappa = 1$. For the reported support $C_{\max} = 10$, the mask contrasts the outer affiliation levels with the central region. The thresholds were fixed before comparing scenarios and were not tuned separately by scenario.

The thresholds are expressed as fractions of the affiliation range: $c_{\mathrm{tail}}/C_{\max} = 0.70$ isolates the most extreme 30% of magnitudes, while $c_0/C_{\max} = 0.40$ separates the central band from the two outer lobes. The unit central penalty makes the bimodality score an outer-minus-central probability contrast on the smoothed conditional histogram. To test that these conventions do not select the result, the robustness analysis evaluates all nine combinations $c_{\mathrm{tail}} \in \{6, 7, 8\}$ and $c_0 \in \{3, 4, 5\}$ on the same 120 seeded histories.

The corresponding mask scores are linear functionals:

$$\langle m_{\mathrm{tail}}, \tilde{p}_X \rangle = \sum_{c\in\mathcal{C}} m_{\mathrm{tail}}(c)\tilde{p}_X(c;t), \tag{26}$$

and

$$\langle m_{\text{bi}}, \tilde{p}_X \rangle = \sum_{c \in \mathcal{C}} m_{\text{bi}}(c) \tilde{p}_X(c;t). \tag{27}$$

3.5. Shape-Weighted Entropy per Microstate

For each microstate X, we first compute the Shannon entropy,

$$H_N(\tilde{p}_X) = -\sum_{c \in \mathcal{C}} \tilde{p}_X(c;t) \log \tilde{p}_X(c;t), \tag{28}$$

and a generalized heavy-tail entropy core,

$$H_q(\tilde{p}_X) = \sum_{c \in \mathcal{C}} \frac{\tilde{p}_X(c;t) - \tilde{p}_X(c;t)^q}{q-1}, \qquad q \neq 1. \tag{29}$$

In the limit $q \to 1$, H_q reduces to the Shannon form.

The shape-weighted entropy of microstate X is then defined as

$$\begin{aligned} S_X(t) = {} & w_N^X(t) H_N(\tilde{p}_X) + w_T^X(t) H_q(\tilde{p}_X) \\ & + w_T^X(t) \beta_{\text{tail}} \langle m_{\text{tail}}, \tilde{p}_X \rangle + w_B^X(t) \beta_{\text{bi}} \langle m_{\text{bi}}, \tilde{p}_X \rangle. \end{aligned} \tag{30}$$

All reported calculations use ($q = 1.5$), ($\beta_{\text{tail}} = 1$), and ($\beta_{\text{bi}} = 1$). The value ($q = 1.5$) retains concavity while modestly increasing sensitivity to low-probability mass, while the unit mask scales keep the additive terms on their natural probability-mass scale. For the simulations evaluated in this study, the reported lattice-level diagnostic $S_{\text{lattice}}(t)$ defined in Equation (31) is positive for every scenario, independent run, and time step. Its positivity is also preserved over the complete local sensitivity grid described in the subsection "Diagnostic and Model Robustness". The diagnostic is interpreted jointly with its fitted component weights and the corresponding conditional histograms.

Equation (30) is designed to retain the uncertainty-sensitive component of Shannon entropy while adding explicit sensitivity to the structural features that characterize segregation dynamics. The corresponding weights provide an interpretable decomposition of the entropy at each step.

3.6. Lattice-Level Entropy and Shannon Comparison

The microstate entropies are aggregated into a lattice-level diagnostic by weighting each microstate contribution by its population fraction:

$$S_{\text{lattice}}(t) = \sum_{X \in \{B,A,T\}} q_X(t) S_X(t). \tag{31}$$

Here, $q_X(t)$ is the fraction of occupied agents belonging to microstate X at time t. This aggregation preserves the distinction between bound, activated, and trapped agents while producing a single scalar quantity that summarizes the full lattice state.

For comparison, we also compute a lattice-level Shannon entropy using the same microstate-conditioned histograms:

$$H_{\text{lattice}}(t) = \sum_{X \in \{B,A,T\}} q_X(t) H_N(\tilde{p}_X). \tag{32}$$

This ensures that the comparison between $S_{\text{lattice}}(t)$ and $H_{\text{lattice}}(t)$ is performed on the same microstate decomposition. The difference is that $H_{\text{lattice}}(t)$ measures only the Shannon uncertainty of the affiliation histograms, whereas $S_{\text{lattice}}(t)$ additionally incorporates the tail and bimodality terms through the data-driven weights.

The purpose of this comparison is not to replace Shannon entropy as a general measure of uncertainty. Rather, it is to test whether a shape-sensitive entropy can better describe simulations in which the relevant transition is not simply a change in spread, but the emergence of a specific histogram structure. In the present Schelling-type models, this structure is often a two-lobed affiliation distribution with mass accumulating near the positive and negative extremes. The weights $w_N^X(t)$, $w_T^X(t)$, and $w_B^X(t)$ therefore provide an interpretable decomposition of the entropy response, indicating whether each microstate is being described primarily by a normal-like, heavy-tailed, or bimodal contribution at each step.

All weights are computed directly from the empirical microstate histograms before the entropy is evaluated. They are therefore used as diagnostic coefficients rather than as variables optimized by an entropy maximization procedure. This makes the method computationally simple and directly tied to the observed distributional shape of the simulated system.

Accordingly, two equal values of S_{lattice} need not imply equal shapes. Shape is inferred only from the joint object $\{S_{\text{lattice}}, w_N, w_T, w_B, p_X, N_X\}$. The scalar score summarizes the magnitude of the weighted response; the weights identify the reference geometry responsible for it; and the histogram verifies whether the fitted interpretation is visually and statistically plausible.

3.7. External Distributional and Spatial Benchmarks

Four independent indices are added to clarify what the proposed diagnostic does and does not measure. First, the bias-corrected bimodality coefficient is [32]

$$\text{BC} = \frac{g^2 + 1}{k + 3(n-1)^2/[(n-2)(n-3)]}, \tag{33}$$

where g and k are the bias-corrected sample skewness and excess kurtosis of all occupied affiliation values and n is the occupied population. The conventional uniform-distribution benchmark is $5/9$; values above it are compatible with bimodality, although BC is not by itself a significance test.

Second, distributional polarization is measured using the discrete Esteban–Ray identification–alienation index [33],

$$P_{\text{ER}} = \frac{1}{P_{\max}} \sum_i \sum_j p_i^{1+\alpha} p_j |y_i - y_j|, \qquad \alpha = 1.6, \tag{34}$$

where the occupied affiliation positions y_i are rescaled to $[0, 1]$, and $P_{\max} = 2(1/2)^{2+\alpha}$ is the value of the equally populated two-endpoint reference. Thus, BC measures moment-based two-peakedness, whereas P_{ER} additionally weights within-group identification and between-group distance. Both remain independent of the lattice coordinates.

Actual spatial organization is measured in two established ways. With binary Moore-contiguity weights a_{ij} between occupied sites, global Moran's statistic is [34]

$$I = \frac{n}{W} \frac{\sum_i \sum_j a_{ij}(C_i - \bar{C})(C_j - \bar{C})}{\sum_i (C_i - \bar{C})^2}, \qquad W = \sum_i \sum_j a_{ij}. \tag{35}$$

For the affiliation sign, Newman's nominal assortativity is computed on the same occupied Moore-edge graph [35],

$$r_{\text{sign}} = \frac{\text{Tr}(e) - \sum_s a_s b_s}{1 - \sum_s a_s b_s}, \tag{36}$$

where e is the directed sign-mixing matrix induced by the undirected spatial edges, $a_s = \sum_t e_{st}$, and $b_s = \sum_t e_{ts}$. Vacancies and neutral $C = 0$ agents are omitted from the sign measure. We also report the directly interpretable fraction of non-neutral occupied neighbor pairs with the same sign.

For a like-for-like comparison with the external indices, the explicit lattice-level bimodality component of the proposed diagnostic is recorded separately as

$$B(t) = \sum_{X \in \{B,A,T\}} q_X(t) \langle m_{\mathrm{bi}}, \tilde{p}_X(\cdot; t) \rangle. \tag{37}$$

This component lies on an outer-minus-center probability scale and is part of Equation (30); it is not presented as a replacement for S_{lattice}. The scalar entropy, its shape components, the external distributional indices, and the spatial indices answer different questions and are compared explicitly rather than conflated.

3.8. Concavity for Fixed Weights

For fixed non-negative weights w_N^X, w_T^X, w_B^X, the entropy functional is concave in the empirical distribution $\tilde{p}_X$. The Shannon core is strictly concave on the interior of the probability simplex, since

$$\nabla^2 H_N(p) = -\mathrm{diag}\left(\frac{1}{p_i}\right) \prec 0. \tag{38}$$

Here, $\mathrm{diag}(1/p_i)$ is the diagonal matrix whose entries are $1/p_i$, and $\prec 0$ means that the Hessian is negative definite. This is the usual second-derivative condition for concavity on the interior of the probability simplex. The generalized entropy core H_q is concave for the parameter range used here, since

$$\frac{d^2}{du^2}\left(\frac{u - u^q}{q - 1}\right) = -q u^{q-2} \leq 0. \tag{39}$$

The tail and bimodality masks are linear functionals of p, and are therefore both convex and concave. Consequently, the weighted sum in Equation (30) remains concave for fixed weights.

This concavity statement applies only after the weights are held fixed. Because the weights are recomputed from each empirical histogram, it does not establish global concavity of the full data-adaptive map $p \mapsto S[p, w(p)]$. The weights are diagnostic shape coefficients, not variational degrees of freedom.

4. Simulation Protocol and Reproducibility

The reported calculations use the fixed configuration in Table 2. No parameter is left implicit.

The reported batch comprises 120 simulations: 30 independent runs for each of the four scenarios. The base seed is 20260519. For scenario s and zero-based run number r, the offset is

$$b_{s,r} = 20260519 + 100000s + 1000r, \tag{40}$$

and the independent seed streams are $b_{s,r} + 11$ for vacancies, $+23$ for affiliation, $+37$ for money, and $+51$ for dynamics. Thus, repetitions vary the initial vacancy pattern, affiliation values, monetary values, processing order, movement choices, and long-range contacts.

For every numerical variable and step, the pipeline records the sample mean, sample standard deviation, run count, and 95% confidence interval

$$\bar{x} \pm 1.96\, s_x / \sqrt{30}. \tag{41}$$

Complete per-run trajectories, representative lattice histories, exact run parameters, and the plotting code are included in the public repository. Appendix A describes the repository contents and uncertainty summaries. The main figures show the 30-run mean and 95% confidence band where appropriate. The 100-step horizon is the common comparison window; separate runs extended to 500 steps retained the same qualitative regimes and showed no reversal of the principal conclusions. This is an empirical horizon check, not a claim of a unique asymptotic equilibrium.

Table 2. Exact simulation and entropy parameters.

Quantity	Symbol	Value
Lattice side	N	64
Simulation horizon	$t_{\max}$	100 sweeps
Independent runs	R	30 per scenario
Vacancy fraction	v	0.30
Movement radius	R_{move}	4 (Manhattan)
Satisfaction threshold	h	4 same-sign neighbors
Affiliation support	$\mathcal{C}$	$\{-10, \ldots, 0, \ldots, 10\}$
Initial money	$M(0)$	$\mathcal{U}[5, 15)$
Movement cost	X_{move}	1.0
Long-range contacts	n_{call}	2 per agent per sweep
Cost per contact	X_{call}	0.25
Dirichlet smoothing	α	0.5
Tsallis parameter	q	1.5
Tail threshold	c_{tail}	7
Center/outer threshold	c_0	4
Tail, bimodality scales	$\beta_{\text{tail}}, \beta_{\text{bi}}$	1, 1
Central penalty	κ	1
KL/Shannon guard	ϵ	10^{-12}
Student candidates	ν	$\{2, 3, 5, 10\}$
Esteban–Ray exponent	α_{ER}	1.6
Spatial weights	a_{ij}	Binary occupied Moore adjacency
Larger-lattice checks	N	96, 128; 10 runs per scenario

Diagnostic and Model Robustness

Local entropy-parameter sensitivity was assessed using the full factorial grid ($q \in \{1.4, 1.5, 1.6\}$, $\beta_{\text{tail}} \in \{0.9, 1.0, 1.1\}$, and $\beta_{\text{bi}} \in \{0.9, 1.0, 1.1\}$). The 27 combinations were evaluated on the same 30 seeded runs per scenario. Across the four scenarios, the minimum Pearson correlation between a perturbed ensemble-mean trajectory and the reported trajectory was 0.9891, every start-to-final direction was preserved, and the maximum absolute change in a final-step mean was (6.91%). Thus, small simultaneous perturbations change the numerical scale modestly but do not alter the qualitative behavior or interpretation. The complete run-level results and sensitivity figure are provided in the public supplement hosted on GitHub (https://github.com/RDomenikos/Shape-Weighted_Schelling_Entropy/blob/main/SENSITIVITY_ANALYSIS.md (accessed on 16 September 2026)).

The mask-threshold check uses the complete 3×3 grid $c_{\text{tail}} \in \{6, 7, 8\}$ and $c_0 \in \{3, 4, 5\}$ on the same 120 histories. At the ensemble level, the minimum mean within-run correlation with the reported trajectory is 0.9451; the largest mean absolute final-step change is 12.25%. Every scenario–threshold combination preserves the start-to-final direction, and $S_{\text{lattice}}(t)$

remains positive for every run and step. The lower individual-run minimum ($r = 0.7753$) occurs in a Scenario 3 residual-class trajectory, but does not reverse its direction or the ensemble interpretation.

Finite-size robustness is evaluated without retuning at $N \in \{64, 96, 128\}$. The reported $N = 64$ ensemble contains 30 runs per scenario; each enlarged size contains 10 independent runs per scenario. Model robustness is assessed one factor at a time, also with 10 runs per scenario and 100 sweeps: $h \in \{3, 5\}$ around the reported value 4, $v \in \{0.20, 0.40\}$ around 0.30, $n_{\text{call}} \in \{1, 3\}$ around 2, and a common movement/call cost scale in $\{0.75, 1.25\}$ around 1. The cost scaling preserves $X_{\text{call}}/X_{\text{move}} = 0.25$. These tests vary the mechanisms already in the model; no network topology or additional behavioral rule is introduced.

5. Results

5.1. Spatial and Affiliation Dynamics

Figure 2 shows representative lattice evolution in the four scenarios. The independent-run ensemble confirms that the reported patterns are not tied to one initial state. By step 100, the mean bound fraction is 0.9159 ± 0.0080 in Scenario 1, 0.9960 ± 0.0017 in Scenario 2, 0.9956 ± 0.0023 in Scenario 3, and 0.8934 ± 0.0063 in Scenario 4 (mean $\pm$ SD, $n = 30$). Scenarios 2 and 3 therefore approach almost completely bound populations under active affiliation reinforcement. Scenario 4 instead retains a trapped fraction of 0.1066 ± 0.0063 as resources are depleted. In Scenarios 1–3, $q_T = 0$ exactly.

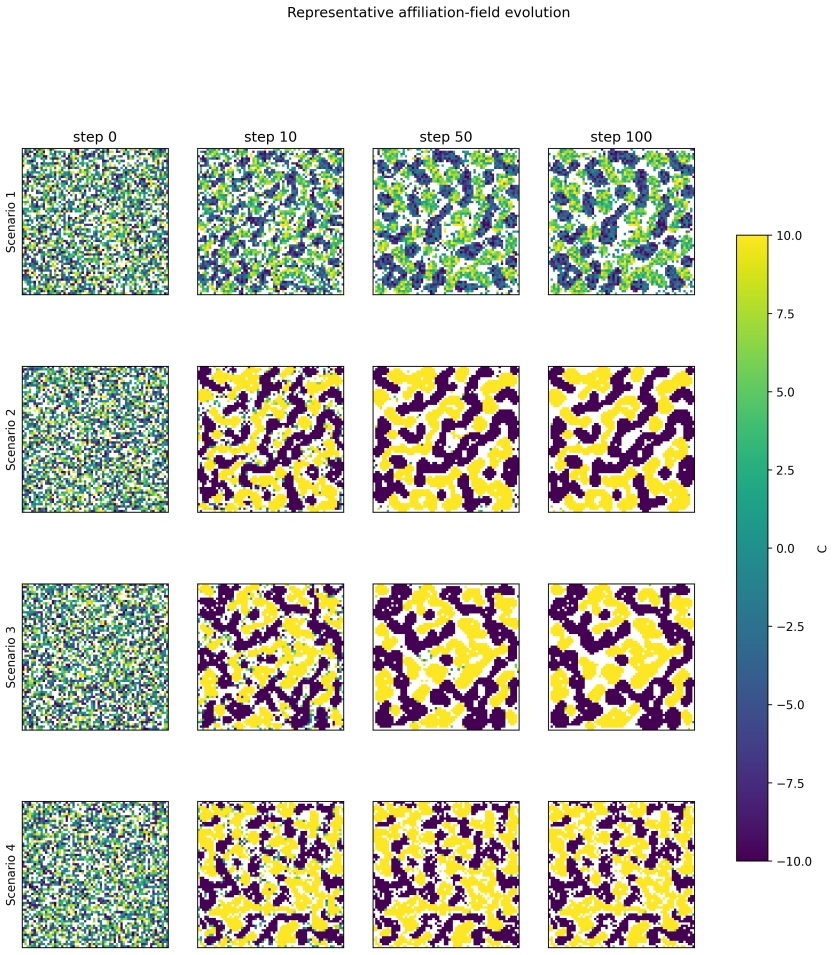


Figure 2. Representative lattice evolution in the four scenarios. Vacancies are distinct from occupied neutral agents. Scenario 1 isolates relocation with fixed affiliation, Scenarios 2 and 3 add affiliation reinforcement, and Scenario 4 adds resource-constrained communication and movement.

Scenario 1 remains a meaningful fixed-affiliation control. Affiliation magnitude is not causal in its movement rule, but applying the same conditional diagnostic shows which distributional features can arise from relocation and microstate selection alone. Any stronger edge accumulation in Scenarios 2–4 can then be attributed to the added affiliation updates rather than to a change in analysis.

5.2. Shannon and Shape-Weighted Diagnostics

Shannon entropy and the shape-weighted diagnostic respond differently because they encode different information. Shannon entropy measures uncertainty on the conditional support; it does not identify whether probability mass is unimodal, heavy-tailed, or concentrated in two opposing lobes. The fitted weight vector supplies this missing geometric decomposition. The scalar diagnostic is therefore never interpreted alone: the trajectory in Figure 3 is read together with Figure 4 and the conditional histograms in Figure 1.

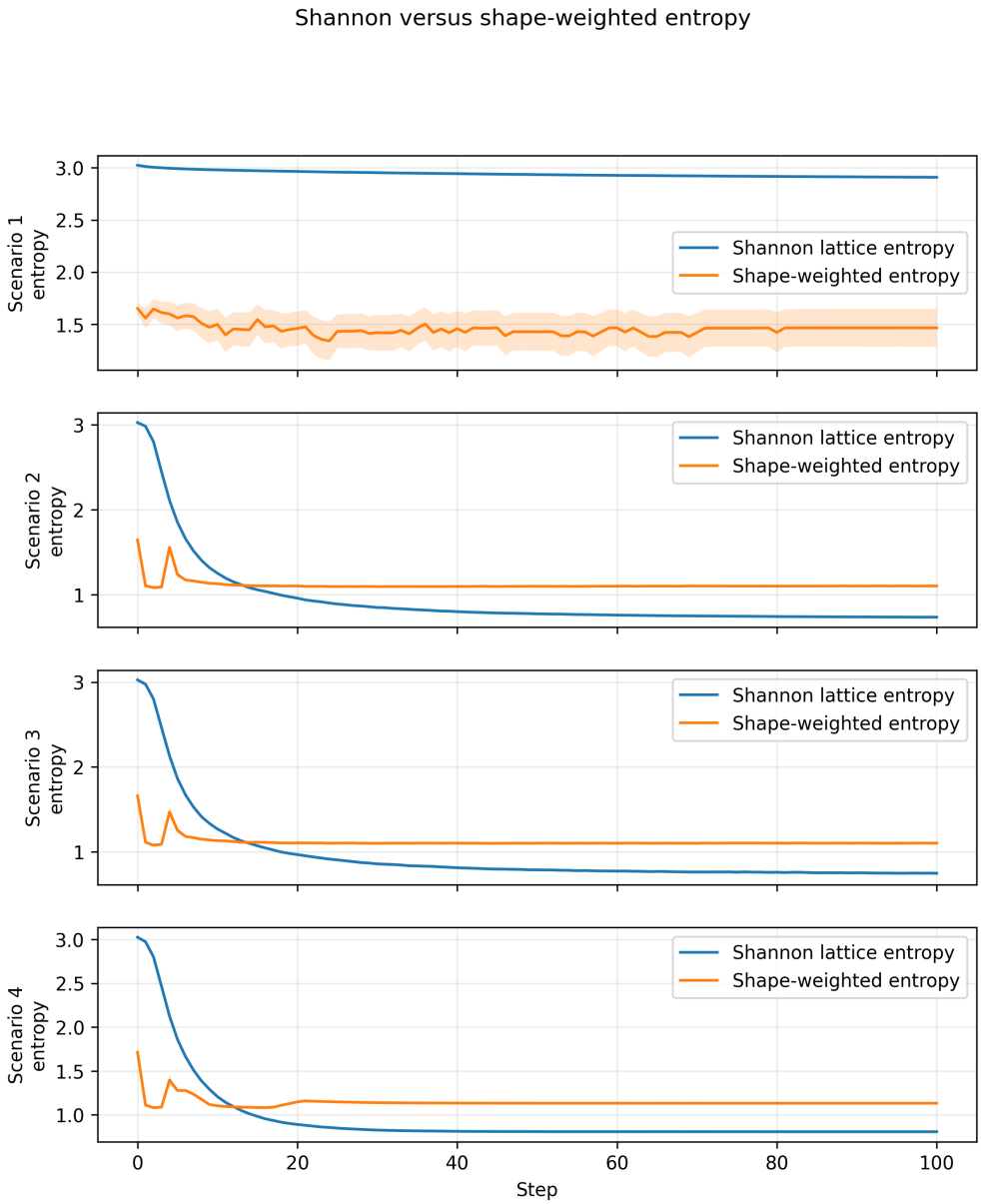


Figure 3. Lattice-level Shannon entropy and shape-weighted entropy across 30 independent runs per scenario. Lines show ensemble means and shaded bands show 95% confidence intervals. The horizontal axis is simulation step.

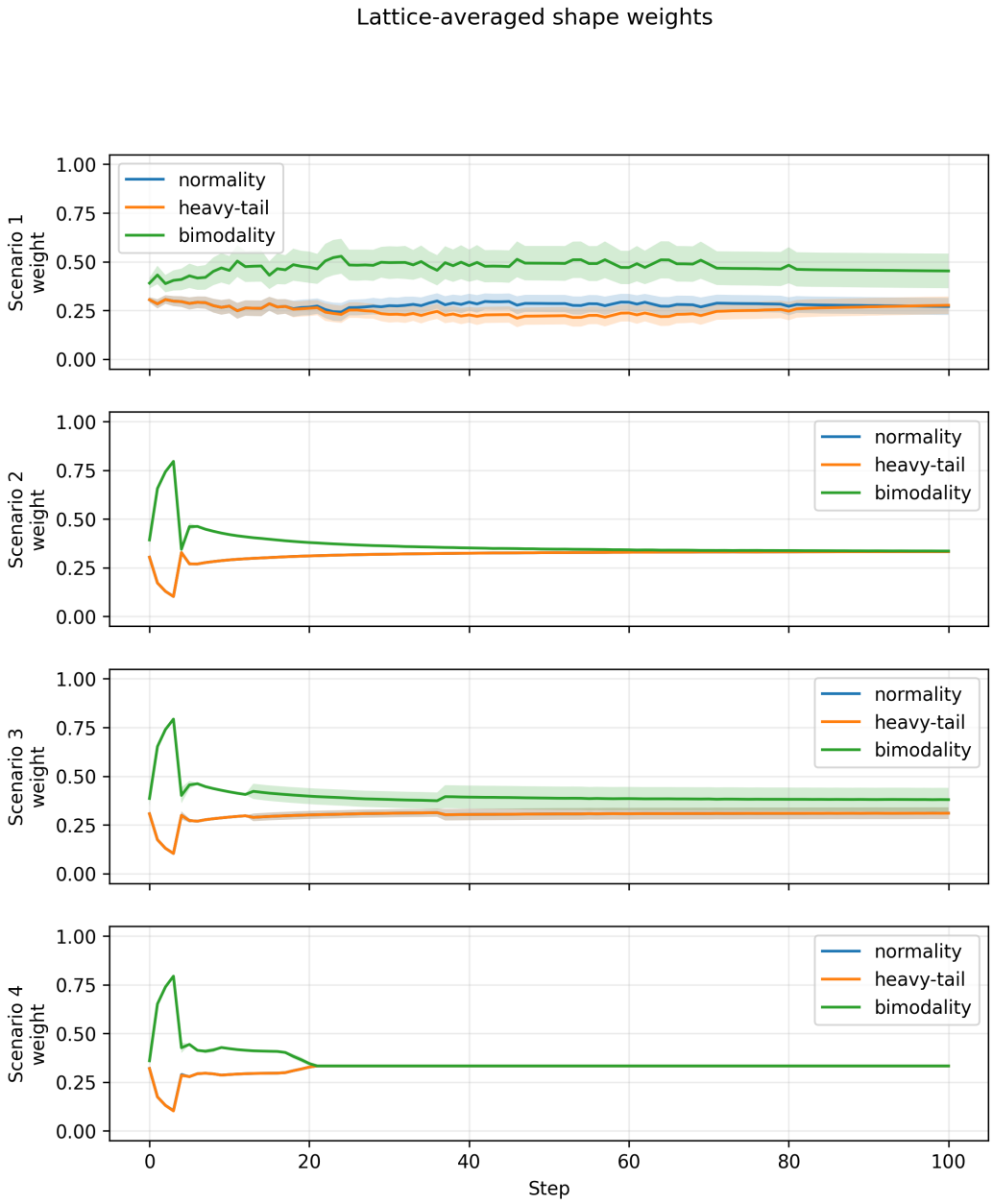


Figure 4. Lattice-averaged fitted weights for normal-like (w_N), heavy-tailed (w_T), and bimodal (w_B) reference structures. Lines show means and shaded bands show 95% confidence intervals over 30 runs. Some trajectories overlap and are not separately visible.

Independent structural indices support the same interpretation. At step 100, polarization is 0.5246 ± 0.0062 in the fixed-affiliation baseline, rising to 0.9983 ± 0.0008, 0.9972 ± 0.0010, and 0.9920 ± 0.0011 in Scenarios 2–4. Central depletion similarly rises from 0.6689 ± 0.0089 to 0.9989 ± 0.0007, 0.9988 ± 0.0008, and 0.9946 ± 0.0012. These independently computed quantities confirm that the two-lobed histograms are not an artifact of the entropy fitting procedure.

The apparently abrupt late change in the Scenario 3 activated histogram is a conditional small-sample effect, not an unexplained system-wide transition. In the representative run, the activated class contains only 11 agents at steps 93 and 100, while the bound class contains 2856. Across the 30-run ensemble, the final activated fraction is 0.00438 ± 0.00232. A normalized conditional histogram based on this small residual class is necessarily sen-

sitive to a few observations; the population-level result is the stable, strongly polarized bound distribution.

5.3. External Validation Against Distributional and Spatial Indices

Figure 5 answers RQ2 and RQ3 without treating the metrics as interchangeable. In the movement-only control, affiliation values never change: $B(t)$, BC, and Esteban–Ray polarization therefore remain constant, while Moran's I rises from approximately zero to 0.6774 and sign assortativity rises to 0.8507. This is the expected counterexample showing that spatial sorting need not change the marginal affiliation distribution.

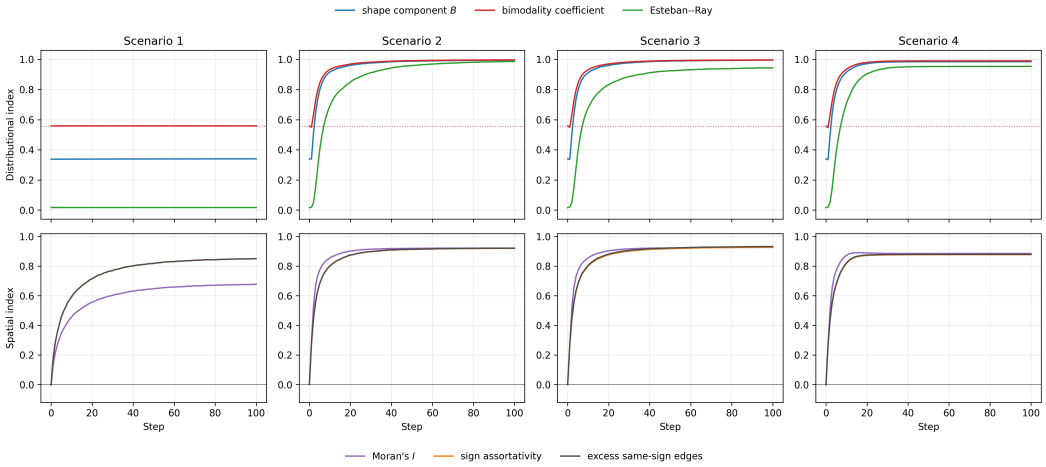


Figure 5. External validation across 30 independent runs per scenario. The upper row compares the explicit lattice bimodality component $B(t)$ with the bias-corrected bimodality coefficient and normalized Esteban–Ray polarization; the dotted line is the conventional BC benchmark 5/9. The lower row shows Moran's I, sign assortativity, and the excess same-sign edge fraction $2f_{\text{same}} - 1$. Curves are ensemble means. Scenario 1 separates unchanged distributional shape from increasing spatial segregation, whereas Scenarios 2–4 show their joint evolution.

When affiliation interaction is active, the distributional and spatial changes converge. Averaging the within-run time correlations over the 30 seeds, $B(t)$ correlates with BC at $r = 0.9974$–0.9979 and with Esteban–Ray polarization at $r = 0.9471$–0.9537 across Scenarios 2–4. Its correlations with Moran's I are $r = 0.9523$–0.9680, and with sign assortativity $r = 0.9812$–0.9827. At step 100, BC is 0.9966 ± 0.0008, 0.9960 ± 0.0011, and 0.9901 ± 0.0012 in Scenarios 2–4, respectively, while Moran's I is 0.9218 ± 0.0106, 0.9265 ± 0.0103, and 0.8853 ± 0.0102 (mean $\pm$ SD). Thus, the two-lobed distributional transition identified by the diagnostic is accompanied by strong spatial clustering in the interaction-driven models, but the Scenario 1 control establishes that this correspondence is empirical rather than definitional.

The full S_{lattice} is not re-labelled as a segregation coefficient. It retains Shannon and heavy-tail contributions and decreases as the interaction-driven systems become more ordered; the positive $B(t)$ component identifies which geometric reorganization accompanies that decline. This explains why the scalar entropy, BC/Esteban–Ray polarization, and Moran/assortativity indices should be read together rather than ranked as substitutes.

5.4. Lattice-Size and Model-Parameter Robustness

The finite-size results in Figure 6 show no qualitative size dependence. Across $N = 64, 96, 128$, final B remains in $[0.9949, 0.9975]$ for Scenario 2, $[0.9947, 0.9975]$ for Scenario 3, and $[0.9847, 0.9890]$ for Scenario 4. The corresponding Moran ranges are $[0.9208, 0.9218]$, $[0.9265, 0.9346]$, and $[0.8853, 0.8890]$. The final bound fraction varies by

less than 0.008 within every scenario across these three sizes, and the Scenario 4 trapped fraction remains between 0.0992 and 0.1066.

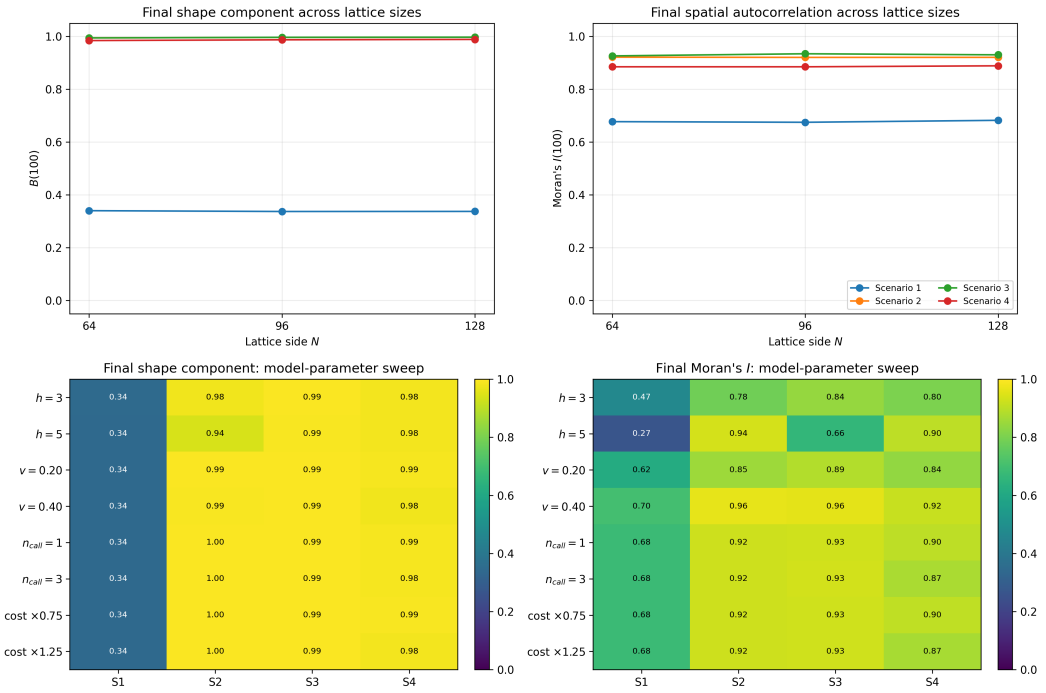


Figure 6. Robustness to lattice size and principal model parameters. Top: final $B(100)$ and Moran's $I(100)$ at $N = 64, 96, 128$; the $N = 64$ points use 30 runs per scenario and the enlarged sizes use 10. Bottom: final ensemble means for one-factor variations of satisfaction threshold, vacancy fraction, long-range contacts, and the common monetary-cost scale, with 10 runs per scenario and configuration.

The parameter sweeps change the magnitude of satisfaction and trapping in physically expected ways without removing the interaction-driven regime. Across all eight one-factor variations and Scenarios 2–4, final B remains between 0.9402 and 0.9951, Moran's I between 0.6606 and 0.9643, and sign assortativity between 0.7535 and 0.9614. The start-to-final direction of B, Moran's I, sign assortativity, q_B, and q_T is preserved in every evaluated configuration. The broadest quantitative changes occur for $h = 5$, which directly raises the difficulty of becoming bound, and in Scenario 4 where the trapped fraction ranges from 0.0129 to 0.2811 as satisfaction, vacancy, calls, and costs are varied. These are parameter effects on the strength of the mechanisms, not reversals of the distributional or spatial conclusions.

5.5. Economic Freezing in Scenario 4

The monetary scenario separates dissatisfaction from mobility. Agents that remain dissatisfied after their resources fall below the movement cost enter the trapped microstate, producing a final mean trapped fraction of 0.1066 ± 0.0063. The simultaneous evolution of movement activity, remaining resources, and q_T in Figure 7 supports the interpretation as resource-constrained dynamical arrest rather than a plotting category or vacancy count.

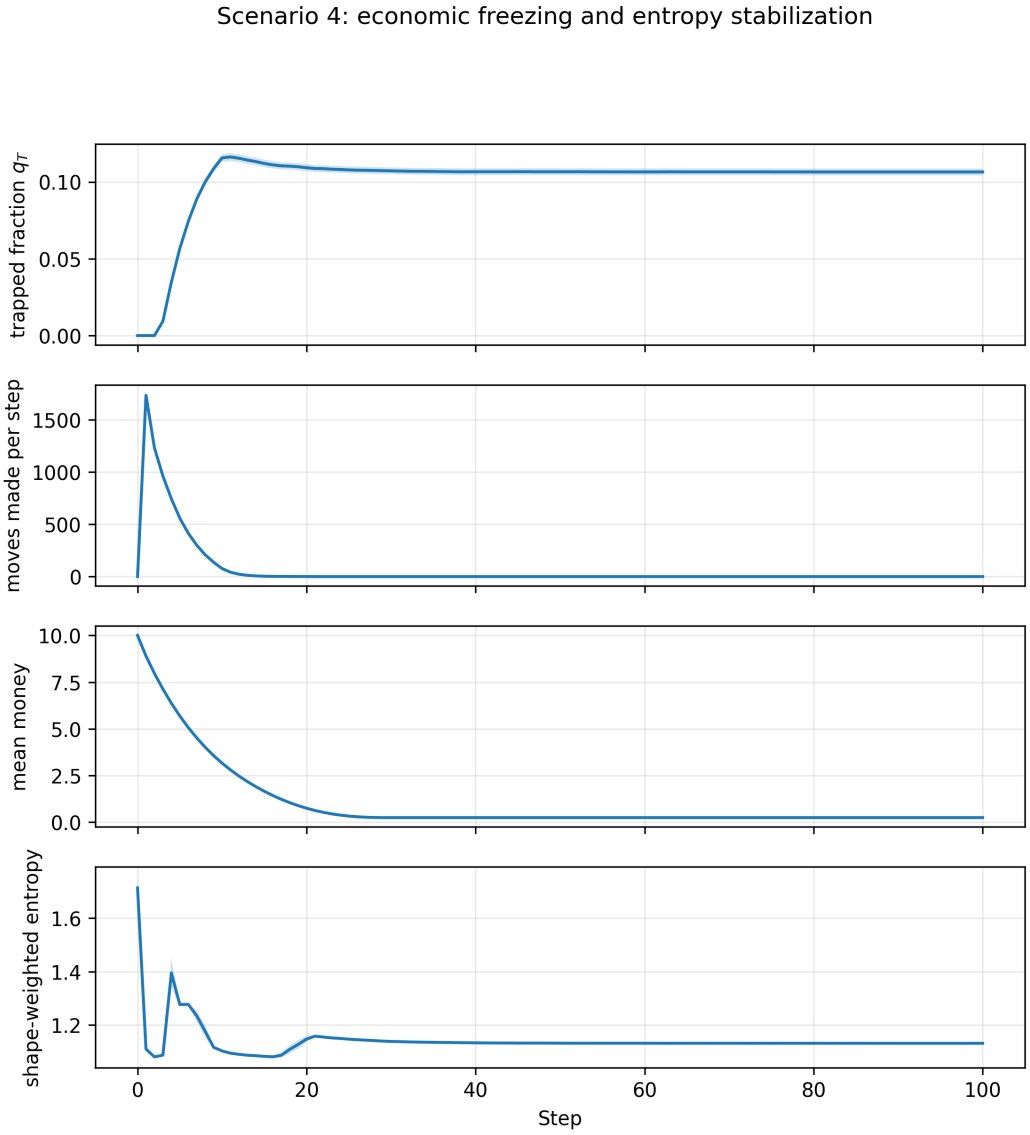


Figure 7. Scenario 4 economic freezing over 30 independent runs. Curves show the trapped fraction, movement activity, remaining money, and entropy diagnostics; shaded bands show 95% confidence intervals.

6. Discussion

The four models are appropriately described as generalized Schelling dynamics. Their common spatial structure is lattice occupancy, vacancies, local satisfaction, and relocation. Scenarios 2–4 add an endogenous affiliation state and consequently overlap with opinion dynamics, as do multiple published extensions that combine Schelling relocation with economic attributes, evolving types, or state switching. The overlap is stated explicitly, but it does not require abandoning the established generalized-Schelling classification.

The long-range mechanism is deliberately minimal. Independent random contacts provide an annealed nonlocal-mixing benchmark with no persistent topology. Replacing it with a realistic social network would answer a different question by adding degree heterogeneity, clustering, repeated ties, and homophily. The present manuscript reports the model actually studied and gives the exact partner-selection rule required to reproduce it.

The integer affiliation support is likewise a modelling decision rather than a defect. The update remains graded because an agent receives the full summed signed influence of its local and long-range contacts. A continuous, complex, or otherwise generalized state

space may be studied separately, but is not needed to make the present finite-state model mathematically or computationally coherent.

The shape-weighted diagnostic addresses a limitation of Shannon entropy without claiming to replace it. Shannon entropy is not a Maxwell–Boltzmann fit, but it is shape-blind: distributions with materially different geometries can have similar entropy. The present method adds explicit reference-shape weights and outer-versus-center scores. The histogram shape cannot be inferred uniquely from the scalar value alone; it is inferred from the scalar, the fitted weights, the conditional sample size, and the histogram together.

The external benchmarks sharpen this scope. The near-unity correlations of $B(t)$ with BC and its strong agreement with Esteban–Ray polarization validate the intended distributional interpretation without relying on the three fitted reference families alone. Moran's I and sign assortativity then show that the same transition is spatially realized in Scenarios 2–4. Scenario 1 is equally important: it produces substantial spatial segregation while its marginal affiliation shape remains fixed. The diagnostic therefore complements rather than replaces established segregation measures, and claims about spatial organization are based on the spatial indices themselves.

For every simulation evaluated in this study, the reported lattice-level diagnostic $S_{\text{lattice}}(t)$ remains positive at every time step. This positivity also persists across the complete local sensitivity grid in the subsection "Diagnostic and Model Robustness". The diagnostic is interpreted over the domain evaluated here and jointly with its fitted weights, conditional sample sizes, and histograms. The complete present functional, constants, code, seeds, per-run outputs, and uncertainty summaries are now public, allowing the reported construction to be reproduced exactly.

Finally, the ensemble design directly addresses stochastic robustness. The conclusions are supported by 30 independently seeded simulations per scenario and narrow uncertainty bands, not by a single favorable realization. The separate 500-step checks address the finite-horizon concern without making an unsupported claim of formal convergence.

The robustness analyses extend that conclusion beyond random seeds. Perturbing both mask thresholds preserves trajectory direction and positivity; enlarging the lattice to 128×128 leaves the final distributional and spatial regimes effectively unchanged; varying the model's principal control parameters changes the degree of binding or trapping without eliminating polarization and segregation in the interaction-driven scenarios. These tests support the methodological conclusions over the evaluated neighborhood of model and diagnostic parameters, while not asserting universality outside it.

Author Contributions: Conceptualization, G.-R.D. and L.Y.C.; methodology, software, formal analysis, investigation, visualization, and writing—original draft, G.-R.D.; supervision and writing—review and editing, L.Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No external empirical data were used. The complete Python implementation, exact parameter configuration, run-level seed values, 120 per-run trajectories, aggregate means, standard deviations, 95% confidence intervals, representative lattice histories, and figure-generation scripts are publicly available in the GitHub repository: https://github.com/RDomenikos/Shape-Weighted_Schelling_Entropy (accessed on 16 September 2026). The reproducibility package includes the external-index calculations, paired mask-threshold analysis, enlarged-lattice runs, one-factor model sweeps, and the CSV summaries underlying Figures 5 and 6.

Acknowledgments: The authors thank the editor and reviewers for their assessment of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Repository Contents and Uncertainty Summaries

The public repository contains the exact source files used for the reported model and analysis: `simulation_model.py`, `entropy_metrics.py`, `run_paper_batch.py`, and `plot_paper_figures.py`. The shell script `run_full_paper.sh` records the complete command line. The file `run_parameters.json` stores every run's parameters and seed values. The directory `per_run/` contains all 120 trajectories, and `summary_timeseries.csv` contains the mean, sample standard deviation, run count, and 95% confidence bounds at every step.

The robustness and validation analyses are reproduced by `run_robustness_validation.py`. It writes `mask_sensitivity_per_run.csv`, `mask_sensitivity_summary.csv`, `sweep_checkpoints.csv`, `sweep_summary.csv`, and `sweep_comparisons.csv`. The main batch output records BC, Esteban–Ray polarization, Moran's I, sign assortativity, the same-sign edge fraction, and the explicit lattice-level tail and bimodality components at every step.

The corrected plotting pipeline treats vacancies as `NaN` and retains $C = 0$ as a valid occupied neutral state. It also treats the trapped class as empty in Scenarios 1–3. Consequently, $q_T = 0$ exactly for those scenarios, and no vacancy count is displayed as a trapped population.

References

1. Schelling, T.C. Models of Segregation. *Am. Econ. Rev.* **1969**, *59*, 488–493.
2. Schelling, T.C. Dynamic Models of Segregation. *J. Math. Sociol.* **1971**, *1*, 143–186. [CrossRef]
3. Schelling, T.C. *Micromotives and Macrobehavior*; W. W. Norton: New York, NY, USA, 1978.
4. Sakoda, J.M. The Checkerboard Model of Social Interaction. *J. Math. Sociol.* **1971**, *1*, 119–132. [CrossRef]
5. Epstein, J.M. Agent-Based Computational Models and Generative Social Science. *Complexity* **1999**, *4*, 41–60. [CrossRef]
6. Bonabeau, E. Agent-Based Modeling: Methods and Techniques for Simulating Human Systems. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7280–7287. [CrossRef] [PubMed]
7. Dall'Asta, L.; Castellano, C.; Marsili, M. Statistical Physics of the Schelling Model of Segregation. *J. Stat. Mech. Theory Exp.* **2008**, *2008*, L07002. [CrossRef]
8. Singh, A.; Vainchtein, D.; Weiss, H. Schelling's Segregation Model: Parameters, Scaling, and Aggregation. *Demogr. Res.* **2009**, *21*, 341–366. [CrossRef]
9. Daudé, E.; Langlois, P. Comparison of Three Implementations of Schelling's Spatial Segregation Model. In *Agent-Based Modelling and Simulation in the Social and Human Sciences*; The Bardwell Press: Cumnor, UK, 2007; pp. 295–326.
10. Pancs, R.; Vriend, N.J. Schelling's Spatial Proximity Model of Segregation Revisited. *J. Public Econ.* **2007**, *91*, 1–24. [CrossRef]
11. Vinković, D.; Kirman, A. A Physical Analogue of the Schelling Model. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 19261–19265. [CrossRef] [PubMed]
12. Gauvin, L.; Vannimenus, J.; Nadal, J.P. Phase Diagram of a Schelling Segregation Model. *Eur. Phys. J. B* **2009**, *70*, 293–304. [CrossRef]
13. Rogers, T.; McKane, A.J. A Unified Framework for Schelling's Model of Segregation. *J. Stat. Mech. Theory Exp.* **2011**, *2011*, P07006. [CrossRef]
14. Collard, P.; Mesmoudi, S.; Ghetiu, T.; Polack, F.A.C. Emergence of Frontiers in Networked Schelling Segregationist Models. *Complex Syst.* **2013**, *22*, 35–59. [CrossRef]
15. Hatna, E.; Benenson, I. The Schelling Model of Ethnic Residential Dynamics: Beyond the Integrated–Segregated Dichotomy of Patterns. *J. Artif. Soc. Soc. Simul.* **2012**, *15*, 6. [CrossRef]
16. Abella, D.; San Miguel, M.; Ramasco, J.J. Aging Effects in Schelling Segregation Model. *Sci. Rep.* **2022**, *12*, 19376. [CrossRef] [PubMed]
17. Mantzaris, A.V. Incorporating a Monetary Variable into the Schelling Model Addresses the Issue of a Decreasing Entropy Trace. *Sci. Rep.* **2020**, *10*, 17005. [CrossRef] [PubMed]
18. Panagiotakopoulos, T.; Domenikos, G.R.; Mantzaris, A.V. Exploring Simulated Residential Spending Dynamics in Relation to Income Equality with the Entropy Trace of the Schelling Model. *Mathematics* **2022**, *10*, 3323. [CrossRef]

19. Zhao, Z.; Randall, D. A Heterogeneous Schelling Model for Wealth Disparity and its Effect on Segregation. In Proceedings of the 2022 ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization, Arlington, VA, USA, 6–9 October 2022; pp. 1–10. [CrossRef]
20. Domenikos, G.R.; Mantzaris, A.V.; Choudur, L. A Thermodynamic State-Space Description of a Monetary Schelling Segregation Model. *Phys. A Stat. Mech. Its Appl.* **2026**, *697*, 131738. [CrossRef]
21. Bilò, D.; Bilò, V.; Döring, M.; Lenzner, P.; Molitor, L.; Schmidt, J. Schelling Games with Continuous Types. In Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, Macao, China, 19–25 August 2023; pp. 2520–2527. [CrossRef] [PubMed]
22. Hazan, A.; Randon-Furling, J. A Schelling Model with Switching Agents: Decreasing Segregation via Random Allocation and Social Mobility. *arXiv* **2012**. [CrossRef]
23. Caridi, I.D.; Pinasco, J.P.; Saintier, N.; Schiaffino, P. Characterizing Segregation in the Schelling–Voter Model. *Phys. A Stat. Mech. Its Appl.* **2017**, *487*, 125–142. [CrossRef]
24. Deffuant, G.; Neau, D.; Amblard, F.; Weisbuch, G. Mixing Beliefs among Interacting Agents. *Adv. Complex Syst.* **2000**, *3*, 87–98. [CrossRef]
25. Hegselmann, R.; Krause, U. Opinion Dynamics and Bounded Confidence: Models, Analysis and Simulation. *J. Artif. Soc. Soc. Simul.* **2002**, *5*, 2.
26. Lorenz, J. Continuous Opinion Dynamics under Bounded Confidence: A Survey. *Int. J. Mod. Phys. C* **2007**, *18*, 1819–1838. [CrossRef]
27. Castellano, C.; Fortunato, S.; Loreto, V. Statistical Physics of Social Dynamics. *Rev. Mod. Phys.* **2009**, *81*, 591–646. [CrossRef]
28. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
29. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 623–656. [CrossRef]
30. Jaynes, E.T. Information Theory and Statistical Mechanics, Parts I. *Phys. Rev.* **1957**, *106*, 620–630 [CrossRef]
31. Jaynes, E.T. Information Theory and Statistical Mechanics, Parts II. *Phys. Rev.* **1957**, *108*, 171–190. [CrossRef]
32. Pfister, R.; Schwarz, K.A.; Janczyk, M.; Dale, R.; Freeman, J.B. Good Things Peak in Pairs: A Note on the Bimodality Coefficient. *Front. Psychol.* **2013**, *4*, 700. [CrossRef] [PubMed]
33. Esteban, J.M.; Ray, D. On the Measurement of Polarization. *Econometrica* **1994**, *62*, 819–851. [CrossRef]
34. Moran, P.A.P. Notes on Continuous Stochastic Phenomena. *Biometrika* **1950**, *37*, 17–23. [CrossRef]
35. Newman, M.E.J. Mixing Patterns in Networks. *Phys. Rev. E* **2003**, *67*, 026126. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.