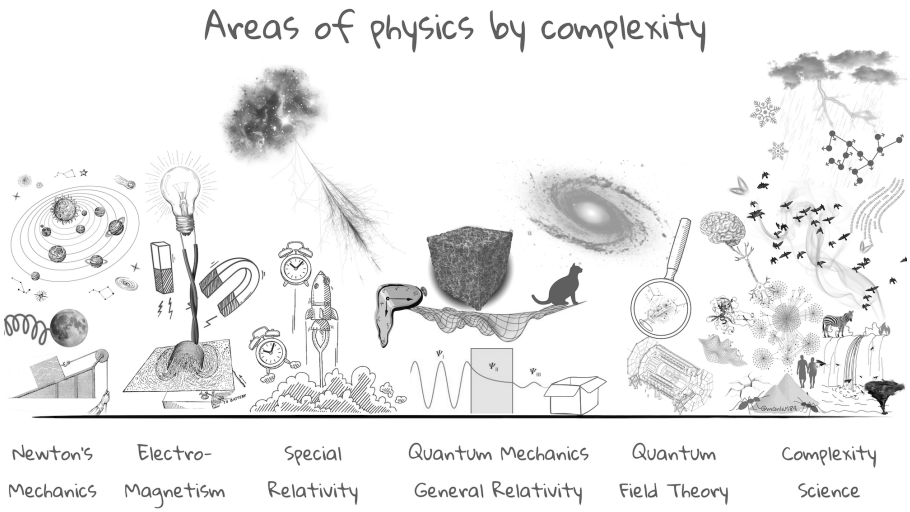
Final Report

Physics of Complex Networks: Structure and Dynamics

Università degli Studi di Padova



PoCN Projects

Vigorelli, Lorenzo

Last update: September 15, 2025

Contents

1 | Sociophysics: Game Theory on networks

Task leader(s): *Lorenzo Vigorelli*

1.1 | Ultimate game

The Ultimatum Game captures [7] how fairness and rejection of unfair proposals spread in a population, and is often linked to the emergence of cooperation. It is a two-player game where the *proposer* offers a division p of a fixed reward, and the *responder* accepts if $p \geq q$ or rejects otherwise. If accepted, both share the reward; if rejected, both receive nothing. The static solution (Nash equilibrium) predicts the proposer offering the minimum amount and the responder accepting it.

On a network, each node represents a player interacting with its neighbors. There are three strategy types:

- **Fair**: $q = p$
- **Pragmatic**: $q = 1 - p$ (same payoff regardless of role)
- **Random**: random (p, q)

At each time step, every player plays the game with all its neighbors in both roles. Strategies then evolve according to one of two update rules:

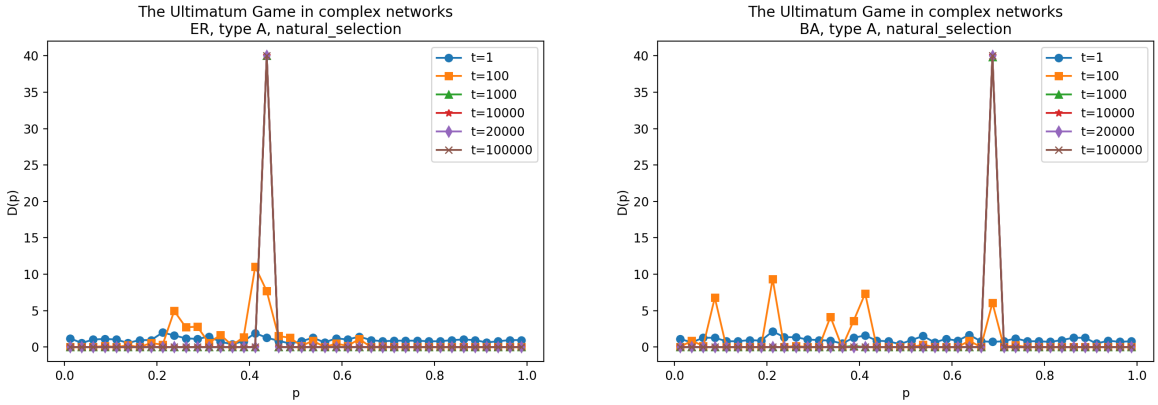
- **Natural selection:** player i compares its payoff Π_i with a random neighbor j and adopts (p_j, q_j) with probability

$$P_{ij} = \frac{\Pi_j - \Pi_i}{2 \max\{k_i, k_j\}} \quad \text{if } \Pi_j > \Pi_i.$$

- **Social penalty:** the lowest-payoff player and all its neighbors are replaced by new players with random strategies, keeping the same links.

1.2 | Simulations

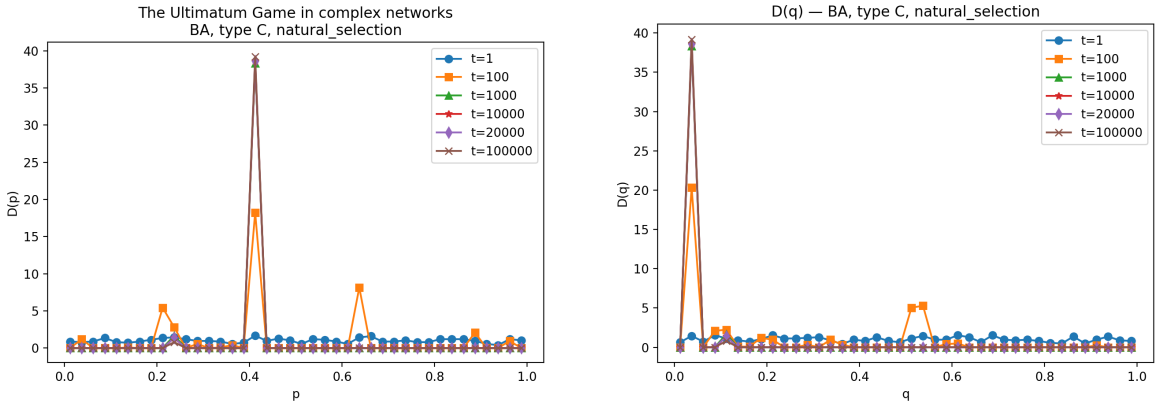
Each network has been tested with both update rules (natural selection, social penalty) and different player types (A, B, C, mixed). Three network topologies were considered: Erdős-Rényi (ER), Barabási-Albert (BA), and an empirical network (Karate Club). The following figures highlight the main effects of network structure, player type, and update rule. The remaining cases are reported in the appendix.



(a) ER: $D(p)$ with central plateau (b) BA: $D(p)$ with shifted plateau (hubs)

Figure 1.1: Comparison of $D(p)$ in ER and BA networks under natural selection.

ER vs BA under natural selection. Figure 1.1 shows the contrast between homogeneous and scale-free networks. In ER, strategies converge towards a stable plateau around intermediate offers p, reflecting uniform connectivity. In BA, the presence of hubs shifts and distorts this plateau, producing sharper peaks and greater heterogeneity. Overall, BA amplifies local instabilities, while ER smooths strategies into collective equilibrium.

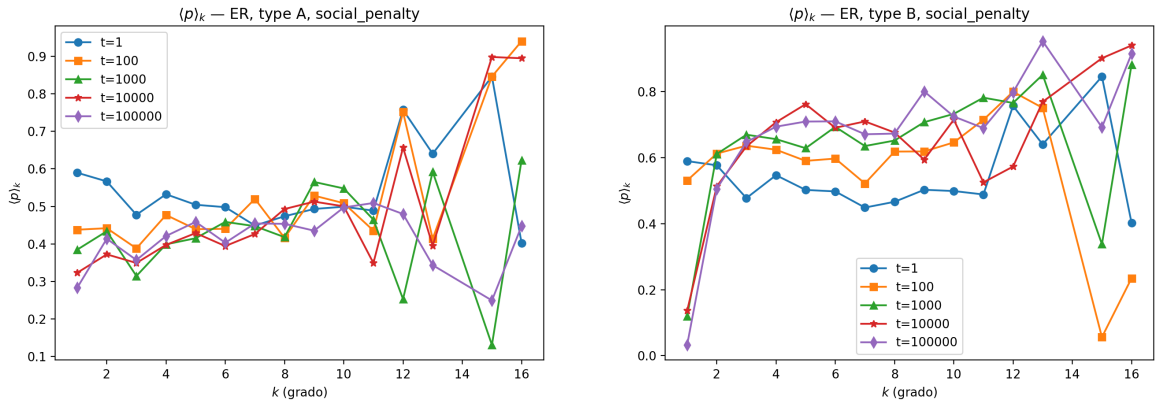


(a) BA: $D(p)$. (b) BA: $D(q)$.

Figure 1.2: Distributions of proposals (p) and acceptance thresholds (q) in BA network with random players under natural selection.

Asymmetry between proposals and acceptance (BA, C players). Figure 1.2 shows that proposals p remain variable and multimodal, stabilizing around several peaks (Fig. 1.2a), while acceptance thresholds q collapse to very low values (Fig. 1.2b). This asymmetry reflects hub influence: they spread opportunistic acceptance quickly, so even small offers are usually accepted, while offers themselves retain heterogeneity.

Social penalty and node degree. Figure 1.3 illustrates the degree dependence under social penalty. We show results by degree class because they highlight more clearly the contrast between low- and high-connectivity players: global distribu-



(a) Average proposal for A players

(b) Average proposal for B players

Figure 1.3: Average proposal $\langle p \rangle$ by node degree in ER network under social penalty.

tions tend to hide these effects, while averages by degree reveal the structural role of connectivity. For type A players (Fig. 1.3a), low-degree nodes converge to similar offers with only gradual adjustments. For type B players (Fig. 1.3b), convergence is slower and more polarized: low-degree nodes end up offering very little, while high-degree nodes move towards fairer offers. This strengthens heterogeneity and emphasizes that central nodes are pressured into fairness, while peripheral ones remain selfish. Compared to natural selection, convergence is slower and degree-based differences persist longer.

1.3 | Conclusions

The simulations show that cooperation and fairness in the Ultimatum Game are shaped by the interplay of network topology, player strategies, and update rules.

Network topology. ER networks smooth fluctuations and favor intermediate offers, while BA networks amplify hub influence, producing heterogeneity and opportunistic acceptance. Empirical networks display bimodal outcomes tied to community structure and degree.

Player type. Fair players (A) converge to intermediate offers in homogeneous settings but become asymmetric in heterogeneous ones. Pragmatic players (B) show strong degree effects: low-degree nodes offer very little, while hubs adopt fairer values. Random players (C) generate multimodal proposals, with acceptance thresholds collapsing to low values.

Update rule. Natural selection drives faster convergence and more uniform equilibria. Social penalty converges more slowly but enforces degree-based differentiation, with high-degree nodes pressured into fairness and peripheral ones remaining selfish.

Overall, fairness emerges not from strategies alone but from how they interact with network structure and update rules, with hubs acting as decisive drivers of global outcomes.

2 | Percolation: k-core & k-clique percolation

Task leader(s): *Lorenzo Vigorelli*

2.1 | Percolation

Percolation describes the effect of progressively removing nodes or edges from a network according to specific rules and studying whether a macroscopic connected component survives. This process reveals how network structure responds to failures or perturbations, and whether the transition from disconnected to connected phases is continuous (second order), discontinuous (first order), or hybrid. Here we focus on local structures, in particular on k-clique percolation [2]. Regarding the simpler case of percolation in ER networks, consider the procedure explained in the appendix.

2.2 | k-Clique Percolation

A *k-clique* is a fully connected subgraph of k vertices. A *k-clique percolation cluster* is built from adjacent cliques that share $k-1$ vertices, thus generalizing classical edge percolation. In Erdős–Rényi (ER) random graphs, a giant k-clique cluster emerges when the link probability p exceeds a critical threshold $p_c^{(k)}$, which scales as

$$p_c^{(k)} \simeq \left[(k-1)N^{\frac{1}{k-1}}\right]^{-1}. \tag{2.1}$$

For $k=2$, this reduces to the standard ER threshold $p_c = 1/N$.
Two natural order parameters can be used:

$$\phi_k = \frac{N^*}{N}, \quad \text{fraction of vertices in the largest cluster,} \tag{2.2}$$

$$\psi_k = \frac{N_k^*}{N_k}, \quad \text{fraction of cliques in the largest cluster.} \tag{2.3}$$

Here N^* is the number of vertices in the largest k-clique cluster, N_k^* the number of k-cliques it contains, and

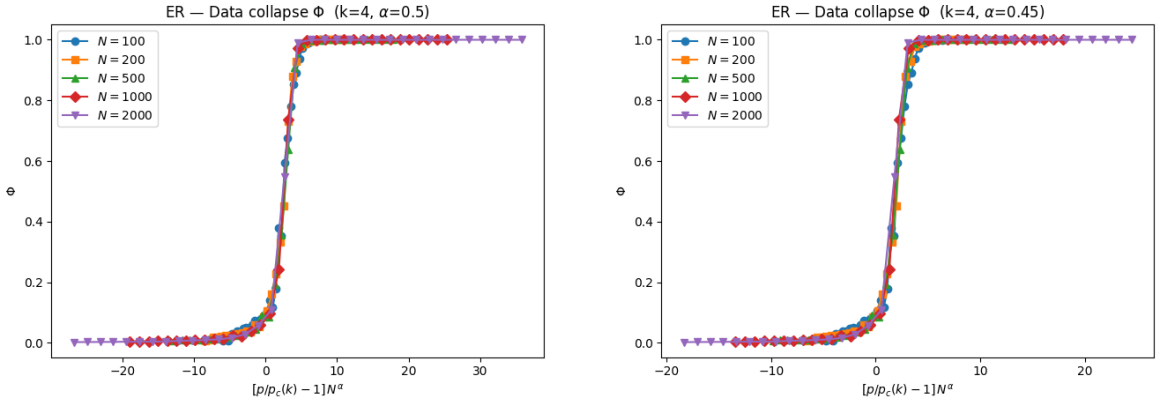
$$N_k \simeq \binom{N}{k} p^{k(k-1)/2} \tag{2.4}$$

is the expected total number of k-cliques in the network.

At the percolation threshold, the scaling of the largest k-clique cluster with system size N follows [2]:

$$N_c \sim \begin{cases} N^{-k/6}, & k \leq 3, \\ N^{1-\frac{k}{2}}, & k \geq 3. \end{cases}$$

2.3 | Simulations



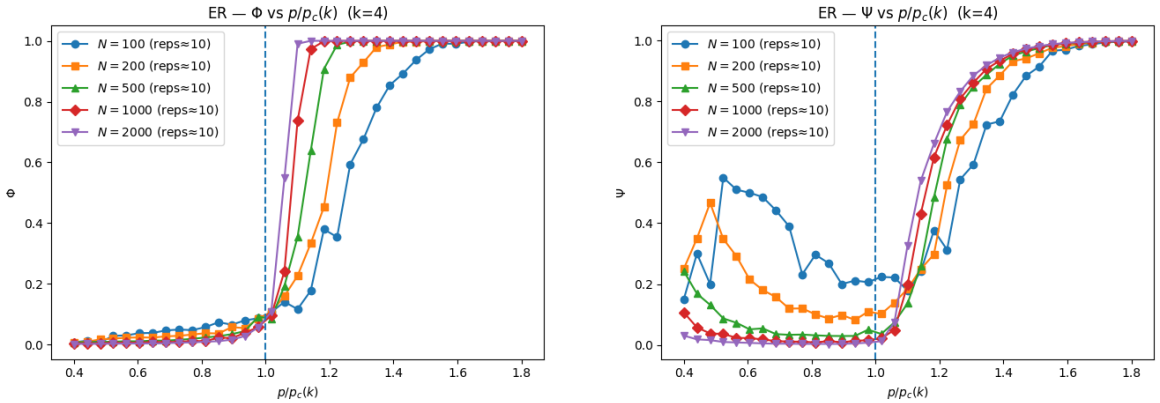
(a) Scaling collapse with $\alpha = 0.45$.

(b) Scaling collapse with $\alpha = 0.50$.

Figure 2.1: Comparison of scaling collapses for $k = 4$ using different critical exponents α.

- Collapse when rescaled accordingly with $\alpha = 0.45$: The scaling collapse in Fig. 2.1a shows an almost universal curve across system sizes N. However, finite-size effects are still visible, suggesting that $\alpha = 0.45$ is not the optimal critical exponent.
- Collapse where rescaled accordingly with $\alpha = 0.50$: When using $\alpha = 0.50$ (Fig. 2.1b), the collapse is significantly better, with nearly perfect overlap of curves from different N. This indicates that $\alpha = 0.50$ captures the correct finite-size scaling of the order parameter.

The contrast between the two panels in Fig. 2.1 highlights the importance of correctly identifying the critical exponent. Notice that the correct α depends on the clique size k.



(a) Order parameter ϕ_k for $k = 4$.

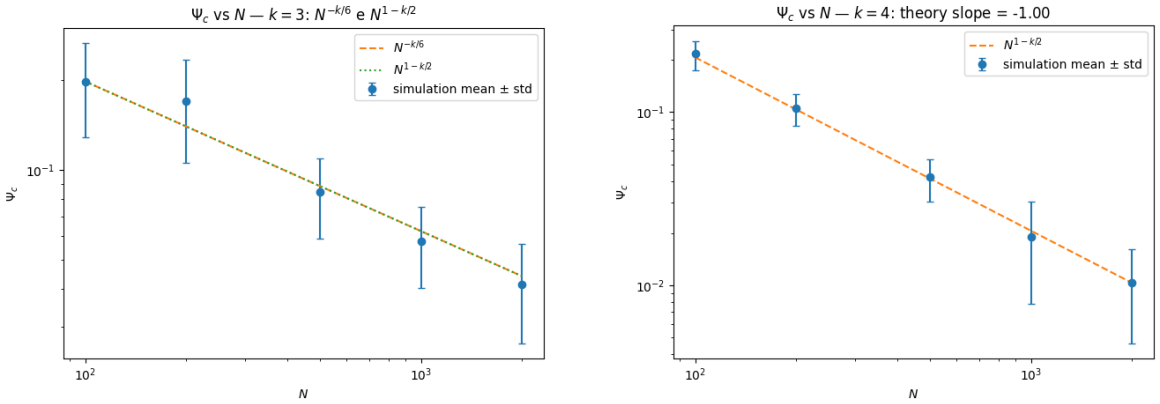
(b) Order parameter ψ for $k = 4$.

Figure 2.2: Order parameters for $k = 4$: relative number of cliques (left) and fraction of vertices (right).

Considering the case $k = 4$ (Fig. 2.2):

- ψ_k: The relative number of k-cliques in the largest cluster increases smoothly with p (Fig. 2.2b). This continuous trend suggests a second-order–like transition in clique space, converging to a step function in the thermodynamic limit. Simulations confirm that curves for different N collapse when rescaled, consistent with theoretical expectations.
- ϕ: In contrast, the fraction of vertices in the largest k-clique cluster (Fig. 2.2a) exhibits a much sharper jump near criticality. This nearly discontinuous behavior resembles a first-order transition, even though in the limit $N \to \infty$ it should converge to a function, which is 0 for $p \leq p_c^{(k)}$ and grows continuously to 1 for $p > p_c^{(k)}$.

In summary, for $k = 4$ the choice of order parameter is crucial: ψ_k suggests continuous growth, while ϕ reveals a sharp structural jump. This duality illustrates the coexistence of second-order-like behavior in vertex space and first-order-like behavior in clique space.



(a) Scaling of ψ_c with N for $k = 3$.

(b) Scaling of ψ_c with N for $k = 4$.

Figure 2.3: Scaling of the critical order parameter ψ_c with system size N for $k = 3$ and $k = 4$.

- Critical ψ_c vs N, $k = 3$: The scaling in Fig. 2.3a follows the theoretical prediction $\psi_c \sim N^{-2/3}$, showing that the relative size of the giant 3-clique cluster decreases with N. This implies that, although a giant 3-clique cluster exists for finite networks, its relative weight vanishes in the thermodynamic limit.
- Critical ψ_c vs N, $k = 4$: For $k = 4$, Fig. 2.3b shows an even steeper decrease with N, consistent with $\psi_c \sim N^{-1}$. This indicates that the giant 4-clique cluster becomes negligible even faster, vanishing completely as $N \to \infty$.

The comparison in Fig. 2.3 highlights the change in scaling: for $k = 3$, the giant clique cluster shrinks as $N^{-2/3}$, while for $k = 4$ it decays more sharply as N^{-1}. In both cases the relative size tends to zero in the thermodynamic limit, but the suppression is stronger for higher-order cliques.

It would be particularly interesting to extend this analysis to scale-free networks, exploring how different percolation modes (node/edge, random failures or targeted attacks) affect the nature of the transition.

3 | Social connectedness index II from Facebook

Task leader(s): *Lorenzo Vigorelli*

3.1 | Social Connectedness Index

The Social Connectedness Index (SCI) measures the intensity of social ties between two locations i and j based on Facebook friendship links [5]. It is defined as

$$SCI_{i,j} = \frac{FB\_Connections_{i,j}}{FB\_Users_i \cdot FB\_Users_j}, \tag{3.1}$$

where $FB\_Connections_{i,j}$ is the number of links between users in i and j, and FB_Users_i, FB_Users_j are the number of users in each location. The index expresses the relative probability of a friendship between i and j, and is rescaled between 1 and 1,000,000,000 for comparability.

To ensure data quality and privacy, locations with very few users are excluded, Gaussian noise is added to connection counts, values are averaged over multiple runs (covering approximately 99% of users), and some regions are not reported. The resulting structure is an undirected, weighted graph where nodes correspond to geographic areas and edge weights to $SCI_{i,j}$.

3.2 | Dataset

The dataset employed is the `GADM1NUTS3` snapshot of 15 December 2021, available through the Humanitarian Data Exchange portal [5]. Geographic units are defined as follows:

- **Europe**: subdivided into NUTS3 regions according to Eurostat boundaries [3].
- **Canada and South Asia** (Bangladesh, India, Nepal, Pakistan, Sri Lanka): subdivided into GADM2 units; in the United States, counties are used following U.S. Census Bureau shapefiles [8].
- **Other countries**: subdivided into GADM1 units, except those with a population below 1 million, which remain unsplit.

Each row of the SCI file represents a pair of these regions. Geographic boundaries are sourced from the Database of Global Administrative Areas (GADM, versions 2.8 and 4.1) [4], with additional reference mapping files provided by Meta [6].

Since both the SCI collection methods and geographic codings have evolved over time, harmonization was required to align administrative levels. The ultimate objective is the construction, for each country, of an undirected weighted graph consisting of:

- a **node list**, including latitude and longitude (representative points of the regions);
- an **edge list**, where edge weights correspond to the SCI values between regions.

3.3 | Node list and edge list

Node list. The set of nodes is obtained by intersecting the location codes in the SCI file (`user_loc` and `fr_loc`) with the reference table of admissible units [6]. The target granularities are `NUTS3`, `GADM2`, `GADM1`, and `COUNTY`. For each type, the corresponding geometries are loaded from GADM [4], Eurostat [3], or U.S. Census Bureau shapefiles [8], converted to WGS84, and assigned a representative point to determine latitude and longitude. As the coding systems differ, keys are normalized before spatial matching (e.g. GADM2 → `ISO3.ADM1.ADM2_1`, GADM1 → `ISO3.ADM1_1`, U.S. counties → 5-digit `GEOID`).

To avoid duplicates, the following priority is adopted: all `NUTS3` and `COUNTY` units are retained, `GADM2` units are kept, and any `GADM1` unit already covered at GADM2 level is discarded. The final node list contains: `nodeID`, `nodeLabel`, `nodeName` (if available), `latitude`, `longitude`, and the dataset of origin.

Edge list. From the SCI file, only pairs with both endpoints successfully mapped are retained. Self-loops are removed and the network is treated as undirected: duplicate pairs (i, j) and (j, i) are merged by averaging the `scaled_sci` values. Edge weights correspond to the scaled SCI, and no thresholding is applied at this stage. Intra-country subnetworks are obtained by restricting edges to nodes within the same country.

Limitations. The intended GADM v2.8 could not be used; instead, v3.x and later versions were adopted [4], which employ different identifiers and coverage. Consequently, some GADM2 units could not be matched, reducing coverage in countries where ADM2 is the main resolution. Additional challenges include heterogeneous code formats, representative points for multipolygons, and the SCI's built-in privacy noise and minimum-user thresholds, which produce many very small edge weights.

3.4 | Network analysis

As the SCI subnetworks are complete and undirected, weighted versions of classical metrics are considered.

Betweenness centrality. As shown in Fig. 3.1, average betweenness remains low and independent of size, indicating that flows are not dominated by specific nodes. Maximum betweenness, however, varies considerably: some networks rely on a few bridge nodes, while others distribute flows more evenly.

Strength and clustering. Figure 3.2 shows that average strength decreases with network size, as weights spread across more nodes. Clustering is low and declines

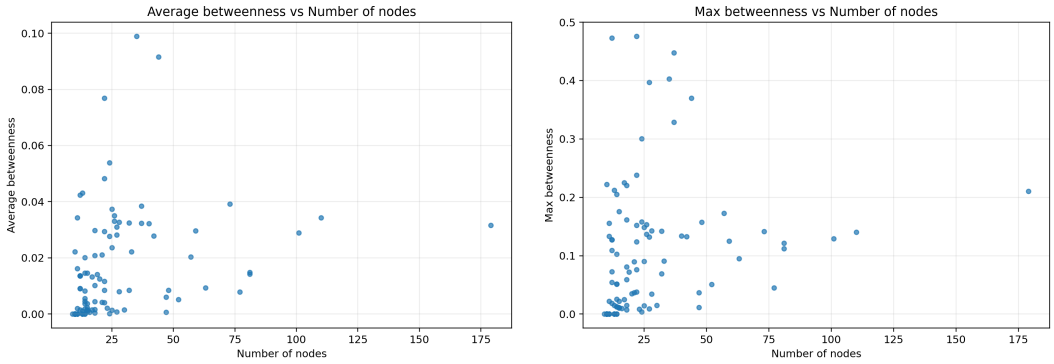


Figure 3.1: Average (left) and maximum (right) betweenness centrality vs. network size.

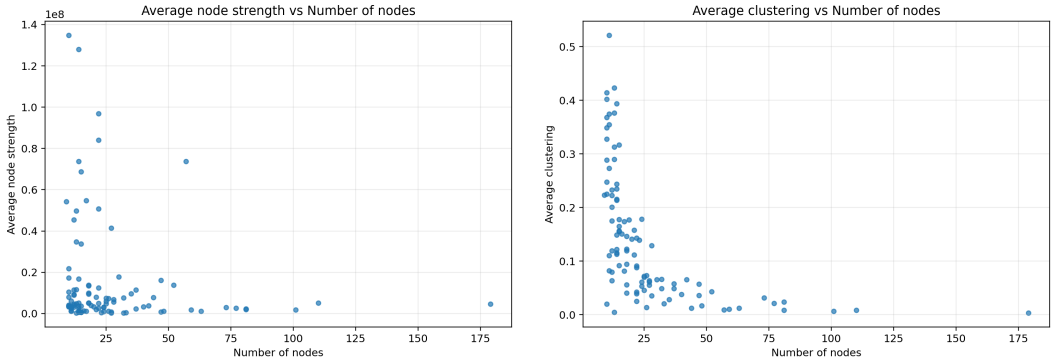


Figure 3.2: Average node strength (left) and clustering coefficient (right) vs. network size.

as networks grow, indicating weaker local cohesion. Together, this suggests that larger SCI subnetworks are more diffuse both globally and locally.

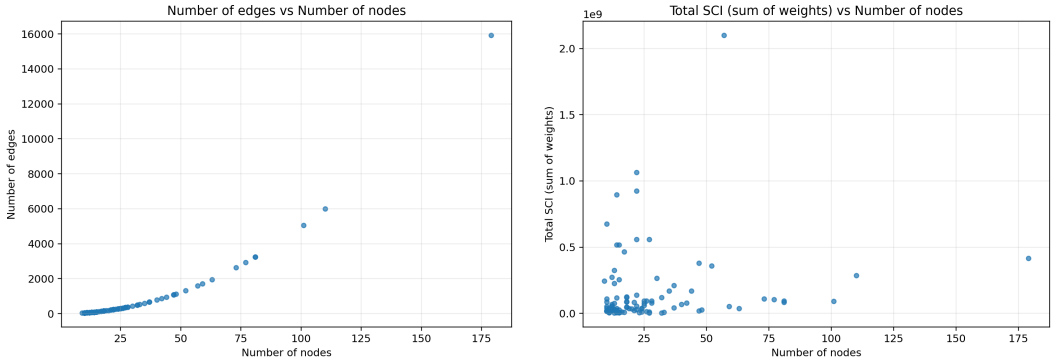


Figure 3.3: Number of edges (left) and total SCI weight (right) vs. network size.

Edges and total weight. As seen in Fig. 3.3, the number of edges grows superlinearly with nodes, confirming the density of SCI graphs. Total SCI weight also scales with size but varies widely across countries, reflecting demographic and usage differences.

4 | Bibliography

[1] Alessio Cardillo, Jesús Gómez-Gardeñes, Daniele Vilone, and Angel Sánchez. Co-evolution of strategies and update rules in the prisoner's dilemma game on complex networks. *New Journal of Physics*, 12:103034, 2010. doi: 10.1088/1367-2630/12/10/103034.

[2] Imre Derényi, Gergely Palla, and Tamás Vicsek. Clique percolation in random networks. *Physical Review Letters*, 94(16):160202, 2005. doi: 10.1103/PhysRevLett.94.160202.

[3] Eurostat. Nuts boundaries (nuts 2016, version 2.8). https://ec.europa.eu/eurostat/web/gisco/geodata/statistical-units/territorial-units-statistics, 2016. Geometries and centroids of NUTS regions used in SCI.

[4] GADM Project. Database of global administrative areas (gadm). https://biogeo.ucdavis.edu/data/gadm3.1/gadm41.shp.zip, 2022. Geospatial data for administrative areas, versions 2.8 and 4.1.

[5] Meta Platforms, Inc. Social connectedness index (sci). https://data.humdata.org/dataset/social-connectedness-index, 2021. Dataset on Facebook friendship ties between geographic regions.

[6] Meta Platforms, Inc. gadm1_nuts2_levels.csv. https://data.amerigeoss.org/dataset/social-connectedness-index/resource/dd9753a9-188d-4c65-a9a1-b25d6bca0621, 2024. Reference table of administrative regions used in SCI (modified 15 March 2024).

[7] Roberta Sinatra, Jaime Iranzo, Jesús Gómez-Gardeñes, Luis M. Floria, Vito Latora, and Yamir Moreno. The ultimatum game in complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, (09):P09012, 2009. doi: 10.1088/1742-5468/2009/09/P09012.

[8] U.S. Census Bureau. Cartographic boundary shapefiles – counties. https://public.opendatasoft.com/explore/dataset/us-county-boundaries/export/?disjunctive.statefp&disjunctive.countyfp&disjunctive.name&disjunctive.namelsad&disjunctive.stusab&disjunctive.state_name, 2021. Geospatial boundaries for U.S. counties used in SCI.

A | Ultimatum Game

A.1 | Related works: Evolutionary Game Theory on Graphs

Evolutionary game theory on graphs is relevant because populations interact only with a fixed subset of neighbors, which has been proposed as a mechanism for the emergence of cooperation among unrelated individuals [1]. In this framework, each agent is characterized by a strategy and an update rule. For example, in the weak Prisoner's Dilemma, strategies are Cooperation (C) or Defection (D), while update rules include replicator (REP), Moran-like (MOR), or unconditional imitation (UI). Cardillo et al. [1] studied the co-evolution of strategies and update rules on adaptive networks. A different example is the Ultimatum Game [7], where strategies evolve depending on past payoffs and the network structure influences the spread of fair or egoistic behavior. This setting highlights the role of altruistic acts (costly to the individual but beneficial to others) as a key mechanism for cooperation.

A.2 | Empirical Network (Karate Club)

Basic features. Nodes: $n = 34$, edges: $m = 78$, undirected and unweighted. Degrees range from 1 (peripheral members) to 17 (instructor hub). Average degree $\bar{k} \approx 4.6$, clustering coefficient ≈ 0.57, and average path length ≈ 2.4 (diameter $= 5$). The network naturally splits into two groups (instructor vs president), frequently used as a benchmark for community detection.

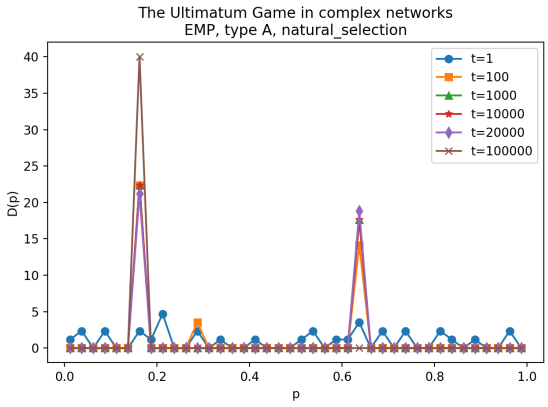


Figure A.1: Proposal distribution $D(p)$ in EMP network with A players under natural selection.

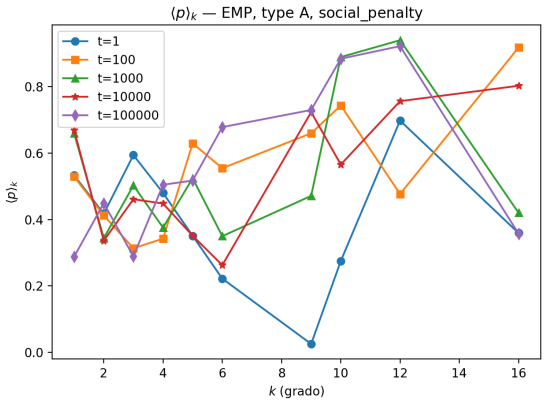
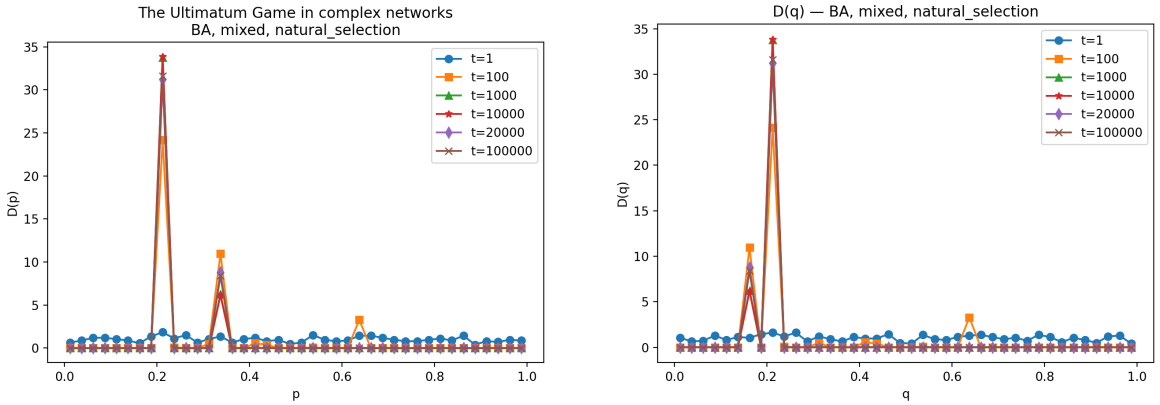


Figure A.2: Average proposal $\langle p \rangle$ by degree class in EMP network under social penalty.

In Figure A.1, proposals show two clear peaks (intermediate and high p), reflecting the heterogeneity of the empirical network. When strategies are analyzed by degree (Figure A.2), high-degree nodes tend to propose more. This suggests that under social penalty they are pressured into fairer behavior, since their removal would have a strong impact on the system.

A.3 | Other simulations



(a) Proposals $D(p)$ with mixed players (A,B,C) in BA network.

(b) Acceptances $D(q)$ with mixed players (A,B,C) in BA network.

Figure A.3: BA network with a mixed population of A, B, and C players under natural selection.

In the BA network with mixed player types (Figure A.3), proposals $D(p)$ become multi-modal (Fig. A.3a) and acceptance thresholds $D(q)$ show several persistent peaks (Fig. A.3b). This indicates that heterogeneous strategies can coexist: fair, pragmatic, and random behaviors interact, generating diversity both in offers and in acceptances. Compared to homogeneous populations, the presence of multiple player types prevents convergence to a single equilibrium and sustains residual variability.

The heatmap in Figure A.4 shows the joint distribution of offers p and thresholds q in the ER network with C players under social penalty at $t = 100000$. The highest density lies in the lower-right region (high p, low q), meaning that most players converge to offering generous proposals while at the same time accepting almost any amount. This reflects the effect of the social penalty: nodes are pressured to avoid the lowest payoff, which drives proposers towards fairness and responders towards opportunism. Additional scattered clusters at intermediate values of p and q reveal alternative equilibria, but the dominant attractor corresponds to this "asymmetric cooperative" regime. The joint representation is therefore crucial, as it shows that fairness (high p) and tolerance (low q) emerge together, a pattern hidden in separate distributions of p and q.

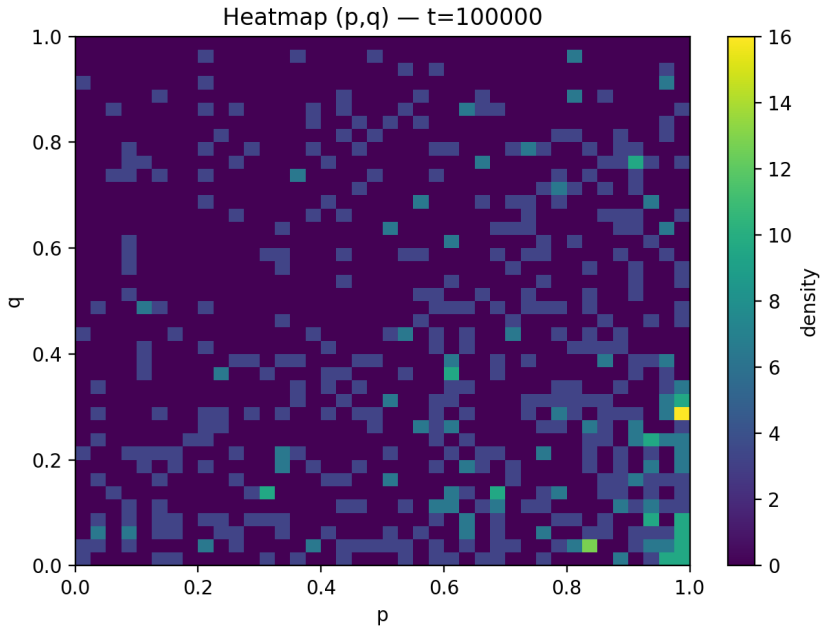


Figure A.4: Heatmap of joint distribution $D(p, q)$ in ER network with C players under social penalty ($t = 100000$).

Summary of additional results

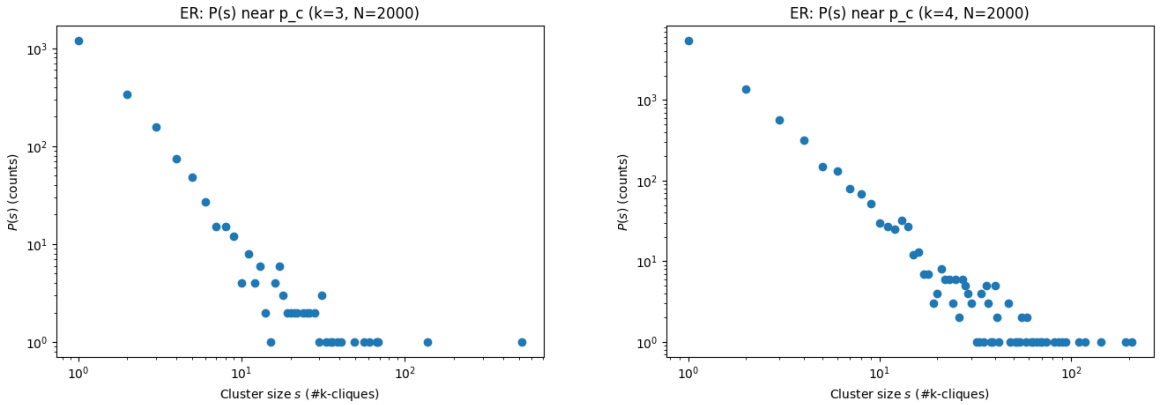
The additional simulations confirm the trends observed in the main text:

- Empirical networks show bimodality and degree effects, with hubs driven to fairer strategies under social penalty.
- BA networks with mixed player types (A, B, C together) preserve strong heterogeneity: fair, pragmatic, and random strategies coexist, leading to multimodal distributions in both p and q.
- Heatmaps of (p, q) reveal that the dominant equilibrium corresponds to high offers (p large) and low acceptance thresholds (q small). This asymmetric cooperative regime is sustained by the social penalty mechanism and is not visible from marginal distributions alone.

Overall, these supplementary cases reinforce the idea that network structure, player diversity, and update dynamics jointly determine how cooperation and fairness spread in populations.

B | k-clique percolation

B.1 | Behaviour at criticality



(a) $P(s)$ distribution for $k = 3$, $N = 2000$.

(b) $P(s)$ distribution for $k = 4$, $N = 2000$.

Figure B.1: Cluster size distributions $P(s)$ in ER graphs at criticality. For $k = 3$ the distribution decays rapidly, while for $k = 4$ the tail is more extended, with rare but larger clusters.

Considering the cluster size distribution at the critical point, it shows that it behaves differently depending on k:

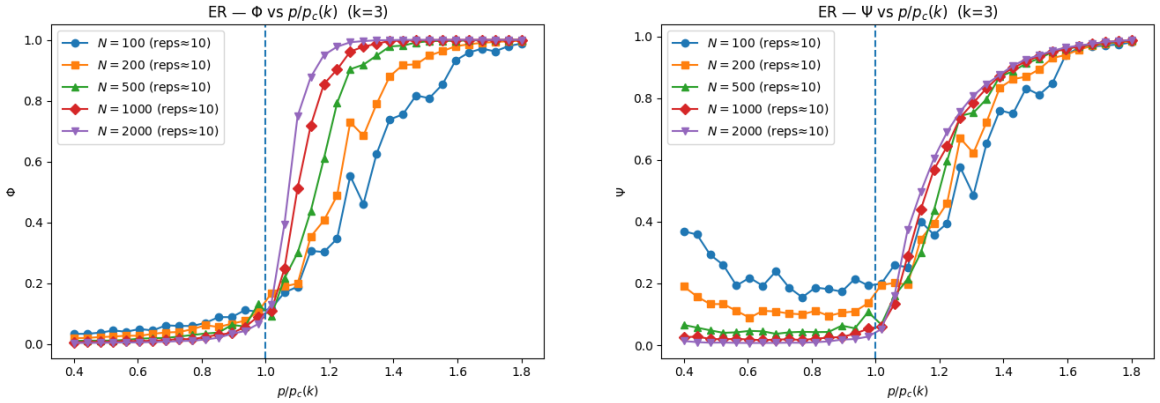
- For $k = 3$ (Fig. B.1a), $P(s)$ decays steeply and large clusters are essentially absent beyond a certain value. This indicates critical fluctuations typical of a continuous transition, but the range of accessible cluster sizes is relatively narrow.

- For $k = 4$ (Fig. B.1b), the distribution extends further: although large clusters are rare, they appear up to $s > 100$. This heavier tail reflects that, in finite systems, clique percolation with $k \geq 4$ can still generate sizeable clusters near criticality, even though theory predicts that the giant component vanishes in the thermodynamic limit.

Thus (Fig. B.1), $k = 3$ shows a continuous transition with limited cluster range, while $k = 4$ displays rarer but larger clusters, consistent with finite-size effects and the theoretical expectation that the giant component disappears as $N \to \infty$.

B.2 | Other simulations

Considering the case $k = 3$ (Fig. B.2):



(a) Order parameter ϕ_k for $k = 3$.

(b) Order parameter ψ for $k = 3$.

Figure B.2: Order parameters for $k = 3$. Both ϕ_k (left) and ψ (right) increase smoothly with p, showing consistent finite-size scaling and confirming a continuous transition.

- ϕ_k For $p/p_c(k) < 1$ we have $\phi_k \approx 0$, indicating the absence of a giant clique component. Around $p/p_c(k) \simeq 1$ the parameter increases rapidly, with a sharper transition as N grows. For small N values, smoothing and finite-size fluctuations appear. This behavior is consistent with a continuous (second-order) transition.
- ψ Below the threshold ψ remains very small, especially for large N, indicating fragmented cliques. Beyond the critical point it increases regularly, showing the emergence of a large-scale interconnected structure. For large N, the curves collapse onto each other, confirming finite-size scaling.

As for $k = 4$ they reveal a continuous transition at $p/p_c(k) = 1$ in the clique space and discontinuous transition in the vertex. The simulations exhibit finite-size scaling effects that vanish as N increases, in agreement with the theoretical predictions of clique percolation.

B.3 | Erdős–Rényi Baseline

The Erdős–Rényi random graph $G(N, p)$ is defined on N nodes, where each possible edge is included independently with probability p. Varying p directly controls the average degree of the network,

$$\langle k \rangle = p(N-1) \simeq pN, \tag{B.1}$$

making ER graphs particularly convenient for percolation studies. By simply tuning p, one generates ensembles of graphs with different expected connectivities, a feature more direct than in scale-free or small-world models.

A giant connected component emerges once the average degree exceeds unity, yielding the percolation threshold

$$p_c = \frac{1}{N}. \tag{B.2}$$

At criticality, the fraction S of nodes in the largest connected component satisfies

$$S = 1 - e^{-\langle k \rangle S}. \tag{B.3}$$

This provides the natural baseline: $k = 2$ clique percolation is exactly equivalent to classical ER percolation, while $k \geq 3$ introduces genuinely higher–order phenomena.

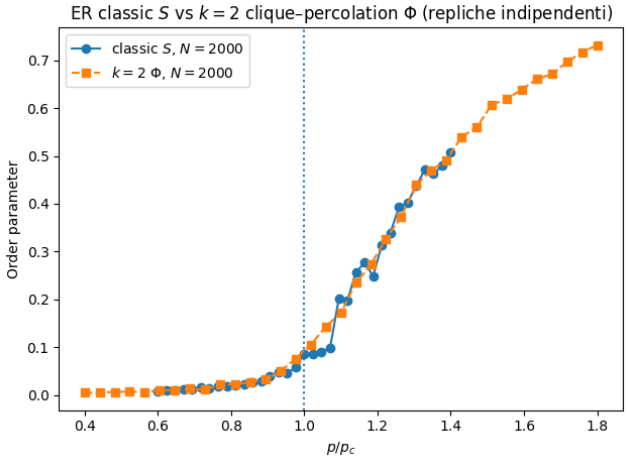


Figure B.3: Comparison between classical percolation (S) and $k = 2$ clique percolation (Φ) in an ER graph with $N = 2000$. The two curves overlap, confirming their equivalence.

As shown in Fig. B.3, the order parameter S of classical ER percolation and Φ for $k = 2$ clique percolation coincide perfectly, validating that clique percolation reduces to the classical case when $k = 2$.
The missing points of classical ER percolation were not simulated outside of the range $[0.6, 1.4]$

C | SCI

C.1 | Global Network anakysis

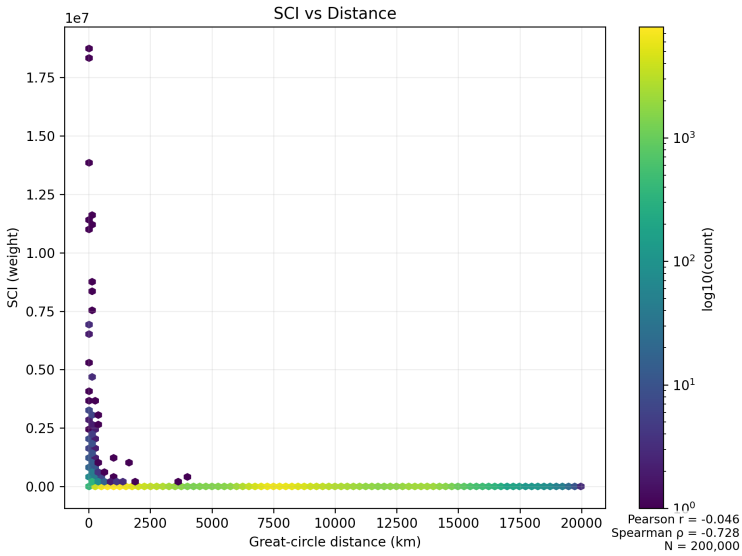


Figure C.1: Relationship between SCI weights and great-circle distance. The color scale indicates the $\log_{10}$ of the number of pairs.

SCI vs. Distance. Figure C.1 illustrates how social connectedness decreases with geographic distance. This considers not the network of each country but a random sampling (to have affordable computations, even though could generate a bias and invalidate the results) of the global network. Most pairs of regions lie in the short-range regime (below ∼2000 km), where SCI values are highest and connections most frequent. Beyond this range, SCI values drop sharply, confirming a strong distance–decay effect: geographically closer regions are far more likely to be socially connected.

Despite this general trend, a small number of long-distance pairs retain relatively high SCI values. These outliers often reflect historical ties, migration flows, or cultural and linguistic affinities that sustain connections despite large geographic separation. The color scale highlights that the majority of pairs correspond to weak ties at short distances, while strong long-distance links are rare but significant.

The correlation measures reported in Figure C.1 further quantify this relationship. The Pearson correlation coefficient is close to zero ($r \approx -0.046$), indicating that the raw linear association between distance and SCI is weak. However, the Spearman

rank correlation is strongly negative ($\rho \approx -0.728$), confirming a pronounced monotonic decreasing trend: as distance increases, SCI values systematically decline. This discrepancy highlights that the distance–decay effect is not linear but strongly monotonic, consistent with the heavy concentration of high SCI values at short distances and their rapid drop beyond $\sim$2000 km.