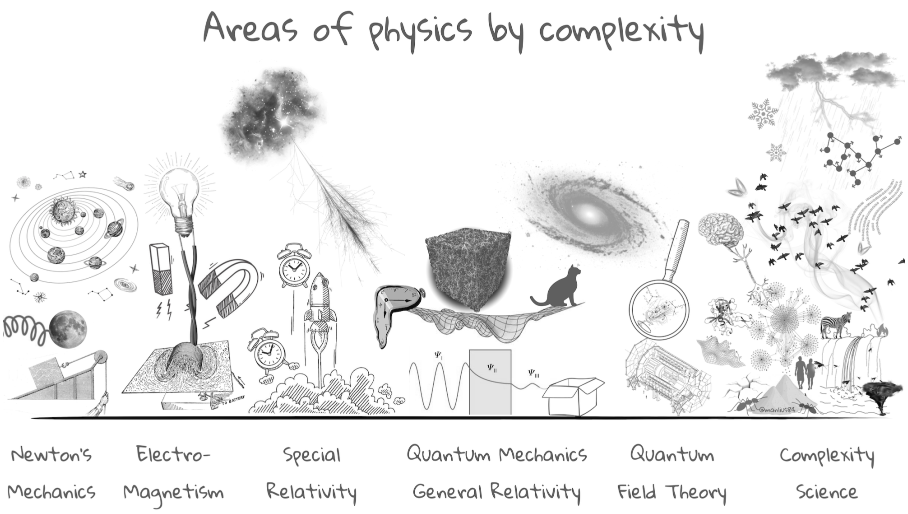
Final Report

ynamics

Università
degli Studi
di Padova



Project #30 and #44

Avilés Moreno, Miguel

Last update: September 15, 2025

Contents

1 | Axelrod model for dissemination of culture

Task leader(s): *Miguel Avilés Moreno*

1.1 Introduction

Why don't all differences in beliefs, attitudes, and behaviors disappear as people interact and become more alike? The Axelrod model, an agent-based model, simulates convergent social influence, where individuals are more likely to adopt traits from similar neighbors. Despite local convergence, global polarization can emerge, influenced by factors like trait count, interaction range, and territory size [2]. This report simulates the Axelrod model's key results [2], applies it to survey data from the American National Election Studies (ANES) [1], and examines the phase transition from homogeneous to polarized cultural states [3].

1.2 Axelrod Model Description

The Axelrod model involves N agents as nodes in an interaction network. Each agent i has a state vector of F cultural features: $\sigma_i = (\sigma_{i1}, \sigma_{i2}, \ldots, \sigma_{iF})$, where each σ_{if} is one of q integer traits, initially assigned with probability $1/q$. The dynamics proceed as follows:

1. Randomly select a connected pair of agents (i, j).
2. Compute their overlap: $l_{ij} = \sum_{f=1}^{F} \delta_{\sigma_{if}, \sigma_{jf}}$.
3. If $0 < l_{ij} < F$, agents interact with probability l_{ij}/F. During interaction, choose a feature g where $\sigma_{ig} \neq \sigma_{jg}$, and set $\sigma_{ig} = \sigma_{jg}$.

The model has q^F cultural options and reaches consensus if one option dominates the system [3].

1.3 Axelrod Application to ANES Survey

We apply the Axelrod model to data from the American National Election Studies (ANES) survey, conducted during recent U.S. presidential elections. Four cultural features are selected: gender, party registration, primary election participation, and voting intention, with up to six traits per feature (e.g., voting intention includes six candidate options).

The simulation uses a 19×19 grid ($N = 361$ agents), with $F = 4$ features and $q = 6$ traits, shown in Figure 1.1. Initial agent states reflect diverse survey responses. After 2000 iterations, agents form homogeneous groups, reaching a stationary state after 3693 iterations. However, the model's limitation is evident: simple interactions may not reduce diverse responses to just two dominant cultural states in reality.

1.4 | Original Axelrod model simulation

In this section, we evaluate how the number of features (F) and the number of traits per feature (q) affect the number of cultural regions in the stationary state. Following Axelrod's original approach, we simulated agents on a grid of size $L = 10$, assigning initial states uniformly at random. For each combination of F and q, the simulation was repeated five times to compute average results. The Table 1.1

The results show that increasing the number of features tends to reduce the number of stable regions and promotes convergence. This occurs because more features increase the probability that neighbors share at least one trait, facilitating interaction. Conversely, increasing the number of traits per feature has the opposite effect, making it more likely that agents will have no shared traits and leading to fragmentation into more distinct cultural regions.

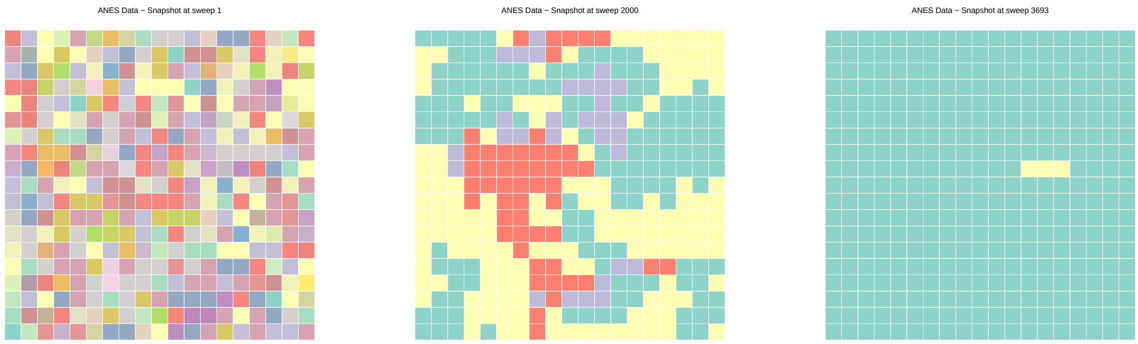
1.5 | Phase Transition in the Size of the Largest Connected Component

In this analysis, we explore how the number of traits per feature (q) influences the emergence of large homogeneous cultural regions, a phenomenon that can be interpreted as a nonequilibrium order–disorder transition. To quantify this behavior, we define an order parameter as the mean relative size of the largest uniform cultural domain, $\langle S_{\max}\rangle/N$. When q is smaller than a critical threshold value q_c, the system tends to reach consensus, resulting in a monocultural state where almost all agents share the same cultural configuration ($\langle S_{\max}\rangle/N \approx 1$). Conversely, for $q > q_c$, the dynamics freeze into a polarized or multicultural state with fragmented domains ($\langle S_{\max}\rangle/N \ll 1$). This transition becomes sharper as the system size increases. In two-dimensional lattices, such as the one used here, the transition is of first order, while in one dimension it becomes continuous. These results are obtained by initializing the system in random, uncorrelated configurations.

The Figure 1.2 shows the evolution of the order parameter as a function of q, illustrating the transition.

For this simulation, we used the following parameters: grid size $L = 20$, number of features $F = 10$, a maximum of 50,000 sweeps to ensure convergence, and 10 independent runs to average the results. The values of q were varied from 1 to 120 in increments of 5 to cover the region around the critical threshold q_c.

Supplementary Materials



Iteration 1 Iteration 2000 Iteration 3693

Figure 1.1: Snapshots of Axelrod model at different iterations (ANES survey).

	$q = 5$	$q = 10$	$q = 15$
$F = 5$	1.2	2.4	5.0
$F = 10$	1.2	1.4	1.6
$F = 15$	1.2	1.6	1.8

Table 1.1: Average number of cultural regions.

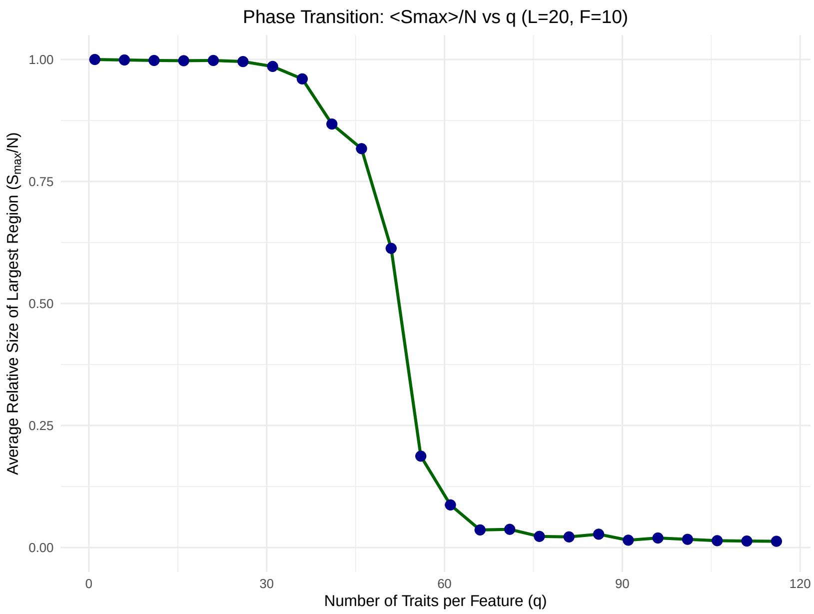


Figure 1.2: Relative size of the largest homogeneous region vs. q.

2 | Social connectedness index II from Facebook

Task leader(s): *Miguel Avilés Moreno*

2.1 | Introduction

This project focuses on the analysis of the Facebook Social Connectedness Index (SCI) of October 2021 to study the global structure of social ties. The dataset provides a scaled measure of social connections between pairs of geographic areas, including NUTS3 regions in Europe and GADM regions elsewhere. Our objective is to build a network of the top 100 countries (excluding the USA) where nodes represent GADM NUTS3 areas and edges encode friendship intensity between regions. The project involves generating two output files: one containing node information and another containing the edge list. Finally, we will perform a simplified network analysis to compare structural properties across countries.

The Social Connectedness Index (SCI) is a measure derived from anonymized Facebook data that quantifies the intensity of social ties between pairs of geographic areas.

2.2 | Data Description

For this analysis was used the dataset from [5], which provides the SCI as of October 2021. The main file, `gadm1_nuts3_counties-gadm1_nuts3_counties - FB Social Connectedness Index - October 2021.tsv`, contains pairwise connections between geographic regions, including each area's connections to itself. Each row records a pair of locations and the scaled SCI value. The second file provides the administrative level and type of each location code, facilitating interpretation of the regional identifiers, `gadm1_nuts3_counties_levels.csv`.

2.3 | Nodes File

To prepare the list of nodes, we excluded all connections involving the United States. From the rest of countries, we identified the top 100 with the highest number of unique regional nodes.

After selecting this subset, we retained only the locations appearing in pairs where both ends belonged to the same country, ensuring a consistent national context for the connections.

To locate the node, we retrieved geographical coordinates (longitude and latitude) corresponding to each region. For European NUTS3 regions, we used centroid coordinates derived from the Eurostat shapefiles dataset [4]. For GADM regions, we relied on version 2.5 of GADM boundaries included in the original data repository [5]. Because GADM codes in Facebook's dataset do not always match GADM's standard notation, we obtained consistent centroids defining a conversion function and the assistance of Gemini [6].

2.4 | Edges File

The edges represent the social connections between pairs of regional nodes. To avoid trivial loops, we removed all self-links (where a region connects to itself) and ensured that each pair of nodes was counted only once, eliminating duplicate entries arising from undirected network records.

Since nearly all region pairs have some level of social connectedness (with a strictly positive SCI), including every link would result in a fully connected graph with very high density, offering limited interpretive value. Therefore, we applied a threshold: for each country, we computed the mean SCI across all its internal pairs and retained only edges whose SCI exceeded this mean, improving the relevance of subsequent analyses. The resulting edge list includes the identifiers of the source and target nodes, the country to which the link belongs, and the scaled SCI value quantifying the strength of the connection.

2.5 | Network Analysis of Top Countries

In this section, the aim is to assess the structure and connectivity of the different regions within a country. The table below show a slice of the countries results for the metrics computed (Number of Nodes, Number of Edges, Network Density, Average Path Length, Clustering Coefficient and Number of Components).

We see that countries such as India and Canada have large networks with a high number of nodes and edges. Hong Kong exhibits a very dense and well-connected network with a high clustering coefficient, despite its relatively small number of nodes. Countries like Argentina show sparse networks, with low density and few edges.

As illustrated in Figure 2.1, we observe a clear positive relationship between the number of nodes and the number of edges in national networks.

2.6 | Degree Distribution and Network Topology

We performed an analysis of the degree distribution and network topology for two countries: Serbia (RS) and India (IND). The results show that India's network exhibits a higher average node degree, indicating dense interconnections across regions. In contrast, Serbia's network is relatively sparse, with fewer regional links and lower node connectivity overall. These structural differences are visually represented in Figure 2.2.

Supplementary Materials

Country	#Nodes	#Edges	Density	Avg. Path	Clustering	#Components
ARG	24	3	0.0109	1.00	0.000	5
CAN	293	2188	0.0511	4.24	0.185	1
IND	644	8717	0.0421	3.67	0.211	2
HKG	18	69	0.4510	1.57	0.579	1

Table 2.1: Sample of network metrics for selected top countries.

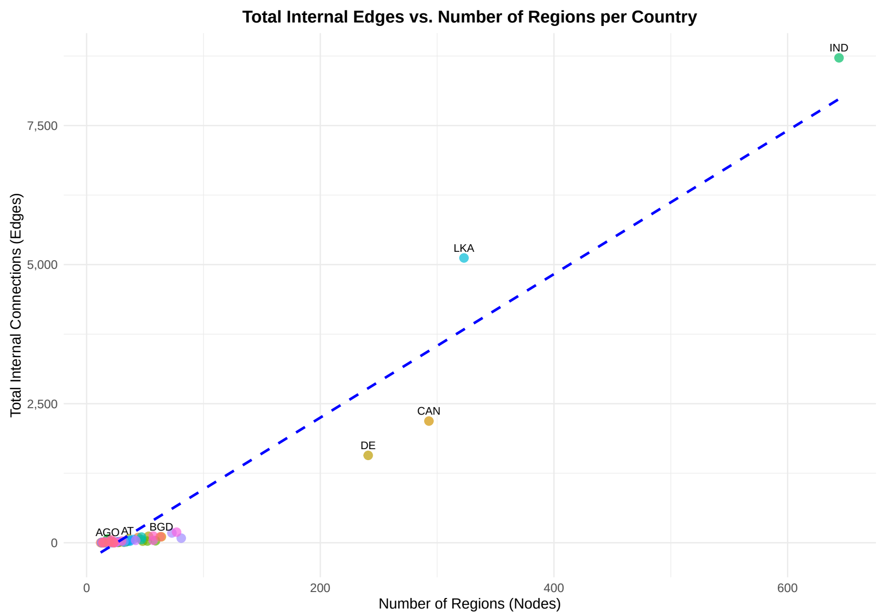
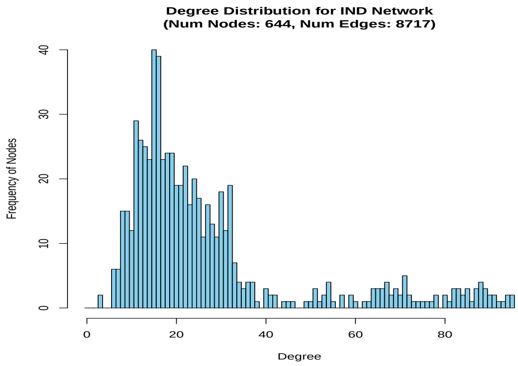
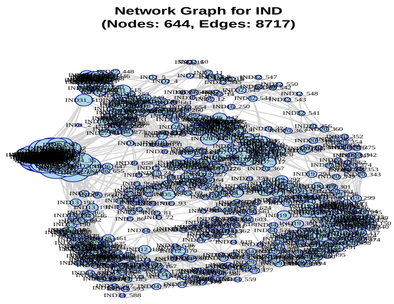


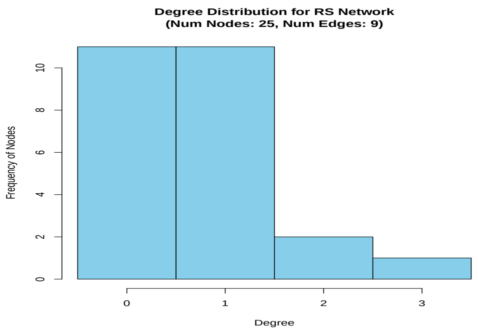
Figure 2.1: Number of nodes vs. number of edges across countries.



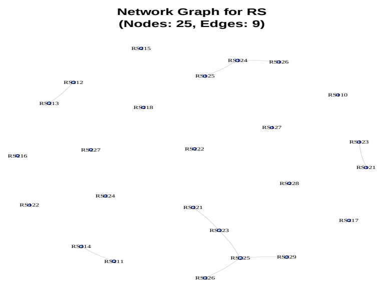
((a)) India: Degree Distribution



((b)) India: Network Topology



((c)) Serbia: Degree Distribution



((d)) Serbia: Network Topology

Figure 2.2: (a) India's degree distribution, (b) India's network, (c) Serbia's degree distribution, (d) Serbia's network.

3 | Bibliography

[1] American National Election Studies. ANES 2024 Time Series Study Preliminary Release: Combined Pre-Election and Post-Election Data. https://www.electionstudies.org, 2025. Dataset and documentation, April 30, 2025 version. Accessed 6 July 2025.

[2] Robert Axelrod. The dissemination of culture: A model with local convergence and global polarization. *The Journal of Conflict Resolution*, 41(2):203–226, 1997. URL http://www.jstor.org/stable/174371. Accessed 6 July 2025.

[3] Claudio Castellano, Matteo Marsili, and Alessandro Vespignani. Nonequilibrium phase transition in a model for social influence. *Phys. Rev. Lett.*, 85:3536–3539, Oct 2000. doi: 10.1103/PhysRevLett.85.3536. URL https://link.aps.org/doi/10.1103/PhysRevLett.85.3536.

[4] Eurostat. Nomenclature of Territorial Units for Statistics (NUTS). https://ec.europa.eu/eurostat/web/nuts, 2025. Webpage and documentation, latest update accessed 12 July 2025.

[5] Facebook Data for Good. Facebook Social Connectedness Index. https://data.humdata.org/dataset/social-connectedness-index, 2025. Dataset and documentation, October 2021 version. Accessed 12 July 2025.

[6] Google DeepMind. Gemini AI. https://deepmind.google/models/gemini/, 2025. Webpage and documentation, accessed 12 July 2025.