Evolutionary Game Selection Leads to Emergent Inequality

Mengesha, Isaak; Roy, Debraj

DOI
10.1007/978-3-031-97557-8_21

Publication date
2025

Document Version
Final published version

Published in
Computational Science – ICCS 2025 Workshops

Citation for published version (APA):
Mengesha, I., & Roy, D. (2025). Evolutionary Game Selection Leads to Emergent Inequality. In M. Paszynski, A. S. Barnard, & Y. J. Zhang (Eds.), *Computational Science – ICCS 2025 Workshops: 25th International Conference, Singapore, Singapore, July 7–9, 2025 : proceedings* (Vol. II, pp. 284-297). (Lecture Notes in Computer Science; Vol. 15908). Springer. https://doi.org/10.1007/978-3-031-97557-8_21

Evolutionary Game Selection Leads to Emergent Inequality

Isaak Mengesha and Debraj Roy(✉)

Computational Science Lab, University of Amsterdam, Amsterdam, The Netherlands
{i.mengesha,d.roy}@uva.nl

Abstract. The emergence of collective cooperation within competitive environments is well-documented in biology, economics, and social systems. Traditional evolutionary game models primarily investigate the evolution of strategies within fixed games, neglecting the simultaneous coevolution of strategies and the environment. Here, we introduce a game selection model where both the strategies employed by agents and the games themselves evolve dynamically through evolutionary processes. Our results demonstrate that these coevolutionary dynamics foster novel collective phenomena, including changed cooperative interactions. When applied to structured populations, the network's architecture, and agent properties such as risk-aversion and bounded rationality significantly influences outcomes. By exploring the interplay between these factors, our model provides novel insights into the persistent social dilemmas observable in real-world systems.

Keywords: Evolutionary Game Theory · Persistent Inequality · Social Networks · Multiple Equilibria · Agent-Based Modeling

1 Introduction

Wealth inequality has declined on a global scale over the past two centuries, but persistent and often worsening disparities remain a significant challenge. In 1800, around 80% of the global population was living below the international poverty line, whereas by 2022 that figure had fallen to approximately 8.5% [1]. This remarkable progress, documented by the World Bank, is tempered by persistent doubts about whether the international poverty line fully captures the reality of poverty and by the observed slowdown in poverty reduction over the last decade [2]. The COVID-19 pandemic, which adversely affected economies worldwide, has underscored the fragility of these gains [3].

Many researchers have sought to explain why low-income groups sometimes become trapped in cycles of poverty despite overall economic growth. They distinguish between "friction-driven" and "scarcity-driven" poverty traps. While the former requires multiple market failures (e.g., lack of credit and indivisibilities in production); the latter can exist without external frictions due deprivation leading to strong behavioural changes (e.g.low savings, underinvestment in human

M. Paszynski et al. (Eds.): ICCS 2025 Workshops, LNCS 15908, pp. 284–297, 2025.
https://doi.org/10.1007/978-3-031-97557-8_21

capital, and myopic decision-making) [4]. In the "poor but neoclassical" view, such adaptations are rational under the given constrains but as a consequence lead to a low-level equilibrium trap that individuals or communities struggle to escape from [5]. Contrary to this is the notion of classical macroeconomic theories like the Solow model of a single, universal equilibrium to which individuals and societies ultimately converge [6,7]. However, both perspectives face mixed evidence, with there being little evidence for the existence of strong traps [8], and simultaneously large and persistence difference in wealth accumulation across individuals and societies [9,10]. Identifying isolated causal mechanisms across heterogeneous populations is methodologically and ethically challenging. In response, modeling studies have become an attractive approach to probe how various micro-level factors can create and sustain wealth inequalities [11].

Risk aversion has been widely discussed as a catalyst for low-risk, low-reward behaviors, which may inhibit innovation and drive persistent wealth stratification [12]. Agents who are sensitive to risk in uncertain environments often prefer modest but reliable payoffs, a habit that can limit their capacity to accumulate resources in the long run. These mechanism are also studied from a modeling approach [13]. Bounded rationality introduces another layer to this dynamic [14]. Real-world decision-makers operate with finite computational abilities and incomplete information, often relying on heuristics rather than optimizing [15]. Furthermore we can see that the introduction of risk and uncertainty in the absence of comprehensive insurance can result in severe changes of the wealth distribution [16]. This condition can exacerbate poverty traps when agents fail to adopt better strategies even when they are locally observable.

Social capital, expressed through the structure of the network and the community's norms, represents a key dimension in explaining inequality [17,18]. Network connectivity allows for the exchange of knowledge, the adoption of profitable strategies, and the accumulation of social influence. Agents with more connections, or a higher network degree, can more easily access beneficial information and resources. In contrast, individuals who remain socially isolated have fewer opportunities to learn or innovate. For example, upward social mobility is greatly affected by the number of high SES connections [19]. These networks can co-evolve as agents establish or sever links based on shared characteristics or observed success, often reinforcing patterns of inequality. The notion that network degree functions as social capital suggests that the "rich-get-richer" effect may operate not only financially but also socially.

Computational models have further explored how simple local interaction dynamics can spontaneously generate persistent inequality, even absent explicit reinforcing mechanisms. For instance, the seminal Sugarscape model demonstrated that minimal behavioral rules alone can lead to substantial wealth disparities, despite agents sharing initially homogeneous conditions [20,21]. Network effects like preferential attachment and homophily further exacerbate these inequalities, allowing certain groups to accumulate disproportionate advantages in resources and information access [22]. Parallel to this, evolutionary game theory has investigated how cooperation can emerge and stabilize within competitive environments [23]. Axelrod's pioneering simulations of the iterated Prisoner's

Dilemma showed that repeated interactions enable reciprocity and foster persistent cooperative behaviors, even among self-interested agents [24]. Subsequent studies demonstrated that network structures and spatial clustering further promote cooperation by allowing cooperators to preferentially interact, thus protecting against exploitation by defectors [25–27]. These mechanisms highlight the significance of agent interaction structure in driving social outcomes, motivating our exploration of the role of social learning (game selection) on inequality.

In this paper, we develop a minimalistic agent-based model to examine how risk aversion, bounded rationality, and social capital can jointly lead to persistent inequality. By abstracting away many real-world complexities, the study aims to identify the core mechanisms of stratification that may persist in diverse contexts. We specifically avoided direct self-reinforcing mechanisms to define traps, or compounding returns to investment. The approach draws on concepts from evolutionary game theory and network science to capture how agents adapt both their strategies and type interactions they engage in (e.g. games played). The experimental design resembles Sadekar *et al.* work on the emergence of cooperative environments [28]. However, among other important distinctions, we focus on the resulting emergence of inequality and overall welfare. The results provide a theoretical account of how uneven wealth distributions may arise even when initial conditions are relatively homogeneous.

Our findings show that local decisions, strategy adaptation, and rewiring of social connections can compound even small initial differences into significant wealth disparities. Boundedly rational agents, who weigh risks conservatively, often settle on suboptimal behaviors, and those with slightly higher connectivity (social capital) gain preferential access to information (games), magnifying their wealth advantage over time. This outcome is consistent with real-world observations that social isolation hinders poverty alleviation and suggests that inequality can emerge and persist through minimal assumptions alone, without needing elaborate institutional or macroeconomic drivers. By illuminating how microlevel biases in learning and networking reinforce each other, this study underscores that small interventions or policy changes at the local level, such as reducing network fragmentation, could prove critical to preventing the entrenchment of persistent poverty.

2 Methods

To systematically investigate how minimal micro-level mechanisms generate and sustain inequality, we employ an agent-based evolutionary game-theoretic framework. This approach enables explicit modeling of individual heterogeneity, adaptive behavior, and dynamic interactions within structured populations. The model comprises four core components: heterogeneous agents characterized by distinct behavioral parameters, payoff matrices defining economic interactions, strategy selection driven by bounded rationality, and evolving network dynamics shaped by wealth-based homophily. Each component is selected to isolate fundamental drivers of inequality, abstracting complex real-world interactions into tractable, theoretically grounded processes. The following subsections formally detail each model component and its corresponding assumptions.

2.1 Individual Payoff Matrices

We use a general two-player, two-strategy symmetric game as our fundamental unit of analysis, due to its conceptual simplicity and analytical tractability. This formulation allows us to abstract essential strategic interactions into a generalizable framework, making it possible to clearly classify games based on payoff structures and systematically explore how different strategic conditions shape inequality dynamics. Formally, the payoff matrix for one agent (Alice) playing against another agent (Bob) is defined as follows:

$$\begin{pmatrix} R & S \\ T & P \end{pmatrix} \xrightarrow{(1)} \begin{pmatrix} R-P & S-P \\ T-P & 0 \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} 1 & \frac{S-P}{R-P} \\ \frac{T-P}{R-P} & 0 \end{pmatrix} \xrightarrow{(3)} \begin{pmatrix} 1 & U \\ V & 0 \end{pmatrix}. \qquad (1)$$

Here, R is the payoff for Alice if both players choose action 1, S is the payoff for Alice if she chooses action 1 and Bob chooses action 2, and so on. Note that adding a constant to all payoffs does not change the strategic structure of the game. Similarly, subtracting P from each payoff does not alter the game's structure, which happens in step (1). By relabeling the strategies, we can assume that $R > P$. The payoffs lack a natural unit of measure, allowing us to rescale them by any positive number. In (2) we make a convenient choice is to rescale by $R - P$. Lastly we define $U = \frac{S-P}{R-P}$ and $V = \frac{T-P}{R-P}$. This representation allows us to analyze games by plotting them in two dimensions, as shown in the figure below (Fig. 1).

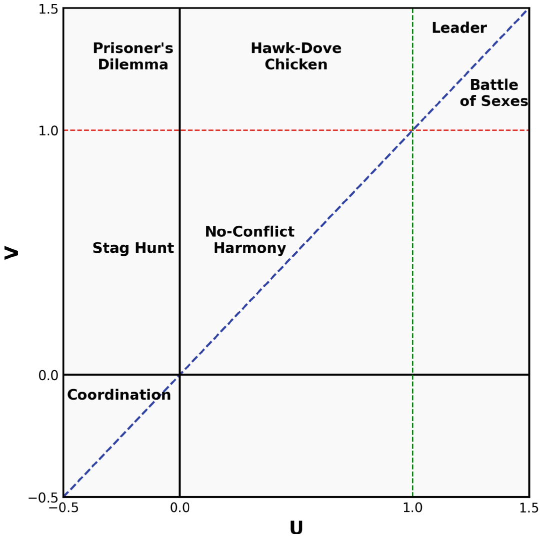

Fig. 1. Games classification - The U-V space of cooperation-defection games. Each agent is assigned one point in the U-V space. In this paper, we have constrained the space to $U > 0$ $V > 0$ space.

Note that there exist qualitative differences between the games depending on the possible orderings of U and V relative to 0 and 1. There are 12 possible orderings, corresponding to 12 different types of games, some of which are labeled in the figure. We limit ourself to the reduction to U and V as defined above.

2.2 Agent Attributes

The stochastic evolutionary game model is populated by heterogeneous agents, each with unique characteristics that influence their decision-making processes and economic outcomes. These agent-specific attributes form the foundation for the emergent inequality patterns observed in the simulation.

Each agent a_i in the model is characterized by a set of key attributes that evolve throughout the simulation. These attributes capture both accumulated resources and intrinsic behavioral tendencies:

- W_i **Wealth**: The cumulative payoff accumulated over all previous time steps. Represents economic resources and serves as the primary measure of inequality. Compounds over time and influences both strategic decisions and network formation.
- W_i^R **Recent Wealth**: The discounted sum of payoffs accumulated over the most recent five time steps. Reflects short-term performance and adaptability, providing a responsive measure of current economic trajectory.
- η_i **Risk Aversion**: A fixed parameter indicating willingness to engage in risky decisions. Higher values correspond to higher risk aversion, leading to selection of strategies with lower potential returns despite uncertainty. Remains constant throughout the simulation.
- λ_i **Bounded Rationality**: A fixed cognitive limitation parameter influencing decision-making precision. Higher values lead to more optimal strategy selection, while lower values result in more random choices. Captures heterogeneity in agents' ability to evaluate strategic opportunities.

The dynamic attributes—Wealth and Recent Wealth—are updated after each interaction based on the payoffs received. The Recent Wealth is calculated using a discounted sum of recent payoffs to reflect the greater relevance of more recent outcomes:

$$W_{i,t}^R = \sum_{j=0}^{4} (1-\delta_d)^j P_{i,t-j}, \tag{2}$$

where δ_d represents a constant discounting factor (set to 0.05 in our simulations), and $P_{i,t}$ is the payoff received by agent i at time t. For the initial time steps where $t < 5$, the recent wealth equals the total wealth: $W_{i,t}^R = W_{i,t}$. The total Wealth is simply the accumulated payoff over time. The risk aversion η changes the curvature in the isoelastic utility function:

$$U(\pi) = \begin{cases} \dfrac{\pi^{1-\eta}-1}{1-\eta}, & \eta \neq 1, \\ \ln(\pi), & \eta = 1, \end{cases}$$

where π represents consumption and η is the constant relative risk aversion parameter.

The fixed Risk Aversion (η) and (bounded) rationality (λ)—influences how agents select strategies and respond to opportunities. These parameters create heterogeneity in decision-making processes, allowing the model to explore how cognitive and behavioral differences contribute to economic stratification even in the absence of explicit institutional advantages or disadvantages. By incorporating these diverse agent attributes, the model captures essential aspects of economic inequality dynamics: the path-dependence of wealth accumulation, the role of risk preferences in economic outcomes, and the impact of decision-making limitations on long-term prosperity.

At $t = 0$, all agents are initialized with zero wealth ($W_i = 0$) and zero recent wealth ($W_i^R = 0$). The fixed personality traits are sampled from standard probability distributions:

$$\eta_i \sim \mathcal{N}(1, 0.5), \quad \lambda_i \sim \text{Log-}\mathcal{N}(0, 1). \tag{3}$$

The inverse risk aversion parameter η_i follows a normal distribution with mean 1 and standard deviation 0.5, creating a population where most agents have moderate risk preferences but with meaningful variation. Higher values of η_i indicate greater willingness to pursue high-risk, high-reward strategies. The bounded rationality parameter λ_i follows a lognormal distribution with location parameter 0 and scale parameter 1, resulting in a right-skewed distribution where most agents have relatively low cognitive precision while a small number have significantly higher decision-making capabilities. This distribution reflects empirical observations of skill heterogeneity in human populations [29,30]. These fixed traits generate persistent heterogeneity in the agent population, as they remain constant throughout the simulation, influencing strategic choices and consequently wealth accumulation patterns over time.

2.3 Payoff Matrix and Updates

Each agent participates in a two-strategy asymmetric game, where their payoff matrix is given by:

$$M_i = \begin{bmatrix} 1 & U_i \\ V_i & 0 \end{bmatrix}. \tag{4}$$

Agents interact by playing against a randomly selected opponent, adjusting their payoffs dynamically based on their dependence parameter $\delta \in [0, 1]$. The dependence parameter determines the degree to which an agent's payoffs are influenced by their opponent's payoffs:

- $\delta = 0$: The agent retains its original payoffs (fully independent play).
- $\delta = 1$: The agent fully adopts the average payoffs between itself and its opponent (fully dependent play).

For an agent i interacting with opponent j, the adjusted payoffs are computed as:

$$U_i' = (1-\delta)U_i + \delta\frac{(U_i + U_j)}{2}, \quad (5)$$

$$V_i' = (1-\delta)V_i + \delta\frac{(V_i + V_j)}{2}. \quad (6)$$

Symmetrically, the opponent updates their payoffs. After updating their payoff matrices, agents independently select strategies based on their respective payoffs. The resulting payoffs are determined by their own updated matrices, rather than a shared game structure, preserving individual strategic diversity. This formulation ensures that when δ is low, agents retain distinct payoffs, leading to heterogeneous strategic behaviors. As δ increases, their payoff structures converge, promoting more homogeneous interactions while still allowing for dynamic adaptation over repeated games.

2.4 Strategy and Game Selection

Strategies are selected using the Logit Quantal Response Equilibrium (LQRE), an equilibrium concept that accounts for bounded rationality and decision-making errors by agents [31]. Unlike the classical Nash equilibrium, which assumes perfectly rational agents, the LQRE allows for stochastic behavior where the probability of choosing a particular strategy increases with its expected payoff but remains sensitive to payoff differences. Specifically, agents follow a logistic choice rule to probabilistically select strategies, given their individual rationality parameter λ_i. The probability of agent i selecting a strategy s_l is given by:

$$P_i(s_l) = \frac{e^{\lambda_i \pi_i(s_l)}}{\sum_k e^{\lambda_i \pi_i(s_k)}}, \quad (7)$$

where $\pi_i(s_j)$ denotes the expected payoff to agent i from playing strategy s_j. A higher rationality parameter λ_i indicates more precise optimization and thus less randomness in strategy choice.

In the specific context of two-strategy interactions, the probability that agent 1 selects strategy S_1 ($P^1_{S_1}$) in interaction with another agent who selects strategy S_1 with probability $P^2_{S_1}$ is explicitly represented as:

$$P^1_{S_1} = \frac{e^{\lambda_1(P^2_{S_1}\pi^1_{11} + (1-P^2_{S_1})\pi^1_{12})}}{e^{\lambda_1(P^2_{S_1}\pi^1_{11} + (1-P^2_{S_1})\pi^1_{12})} + e^{\lambda_1(P^2_{S_1}\pi^1_{21} + (1-P^2_{S_1})\pi^1_{22})}}, \quad (8)$$

where π^1_{11}, π^1_{12}, π^1_{21}, and π^1_{22} represent the payoffs to agent 1 corresponding to the strategy pairs (S_1, S_1), (S_1, S_2), (S_2, S_1), and (S_2, S_2) respectively. Together with the other agent this defines a system of two coupled equations with two unknowns that we solve numerically for each agent up to a certain error tolerance.

Additionally, agents need to make choices w.r.t. what games to play in the future. Taking into account the wealth difference between players δW_{ij} as well

λ_i and η_i. The utility function $U(\eta_i, \delta W_{ij})$ is the isoelastic Utility function. Specifically, the probability that agent i selects to interact with agent j in a future game is modeled using a softmax choice rule, closely resembling the QRE:

$$P_{i \to j} = \frac{\exp\left(\lambda_i \cdot U(\eta_i, \delta W_{ij})\right)}{\sum_k \exp\left(\lambda_i \cdot U(\eta_i, \delta W_{ik})\right)}, \tag{9}$$

This selection rule enables agents to strategically evaluate and probabilistically select opponents based on wealth-driven incentives, risk attitudes, and cognitive limitations, contributing to the dynamic evolution of the system's economic interactions.

2.5 Network Topology and Evolution

Homophily—the tendency for agents to associate with others who are similar to themselves—is a fundamental mechanism in social and economic networks. In wealth-driven settings, agents with similar wealth are more likely to interact, forming clusters that reflect their economic similarity. A formalization of such behavior must capture two essential aspects: first, the increased likelihood of forming connections when agents' wealth levels are similar; and second, the possibility of breaking connections when wealth disparities become significant. Agents interact on a network that updates adaptively based on wealth-driven homophily. The probability of forming a new link follows:

$$P_{\text{con}} = \frac{1}{1 + (1 - \beta \cdot \max(|W_i - W_k|, \epsilon))^{-\alpha}}, \tag{10}$$

where α and β modulate the strength of homophily. Here, $|W_i - W_k|$ represents the wealth difference between agents i and k, and ϵ is a small constant that ensures the expression is well-defined even when the wealth difference is very small. The parameter β scales the impact of the wealth difference, while α controls how sharply the connection probability decreases as this difference increases. Connections may be severed based on the inverse of $P_c on$.

2.6 Simulation Process

At the start of each simulation, agents are placed onto an initial network topology generated using either a Random Regular, Watts-Strogatz, or Holme-Kim algorithm. In our simulations, results did not differ meaningfully across these topologies, suggesting robustness of the observed inequality patterns to the choice of initial network structure, as we demonstrate convergence to the same equilibrium network topology. Agents are initialized with zero wealth, and their fixed behavioral parameters (bounded rationality λ_i and inverse risk aversion η_i) are randomly drawn from the distributions described above. Optionally, entries of the payoff matrices may be normalized by the sum of all payoffs to facilitate comparison across games. Each simulation then proceeds for 100 discrete time steps. Within each time step, every agent sequentially executes the actions summarized in Table 1. Payoffs obtained during interactions are exclusively allocated to the focal agent, ensuring that the wealth effects are not dominated by network dynamics.

Table 1. Detailed simulation steps in the evolutionary game model.

1.	**Opponent Selection**	An agent selects an opponent randomly from its current network neighborhood.
2.	**Payoff Adjustment**	The agent adjusts its payoffs towards the opponent's payoffs based on the dependence parameter δ.
3.	**Strategy Selection**	The agent probabilistically selects a strategy using the Logit Quantal Response Equilibrium (LQRE), influenced by bounded rationality λ_i.
4.	**Play Game**	Agents play the selected game, receive payoffs, and update their Wealth W_i and Recent Wealth W_i^R.
5.	**Choose Game**	Agents choose future games based on past performance, wealth differences, and individual risk aversion η_i.
6.	**Network Update**	Network connections form or dissolve according to wealth-driven homophily rules.

Throughout the simulation, we specifically track two main outcomes: economic inequality—measured by agents' wealth distributions—and the learning rate, capturing how rapidly agents adapt their strategy and game choices in response to past experiences.

3 Results

In the standard experimental runs, we recover basic stylized facts about broader income distributions. We first investigate the complementary cumulative distribution of income to compare with empirical data [16]. The results qualitatively depend on whether the payoff matrices are normalized by the sum of their entries. If not normalized, the income distribution follows a classical power-law that slightly flattens over time and eventually saturates due to finite size effects. Similarly, in the normalized case, however, the income distribution approximates a heavy-tailed form, with visual resemblance to a power-law, although the absolute range is limited (recent wealth spans only a few discrete levels), preventing definitive tail characterization. This indicates the existence of multiple equilibria with varying productivity. Our inequality measures using the Gini coefficient confirm this, landing at 0.12 and 0.41 (normalized) respectively. We interpret the non-normalized case as reflecting technological innovation that increases baseline productivity. With normalization, there are no overall payoff advantages between games; instead, games differ primarily in their ease of coordination toward beneficial outcomes (Fig. 2).

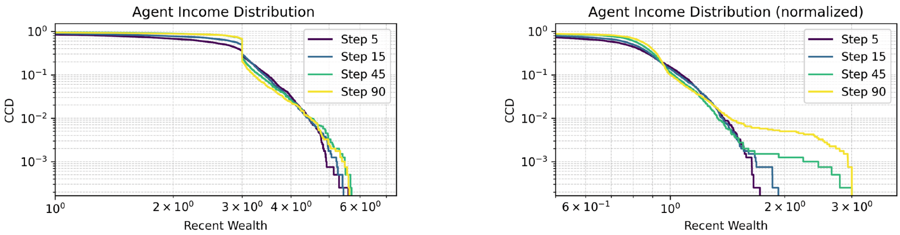

Fig. 2. The complementary cumulative distribution of income over a fixed time interval mimics a classical power-law for standard games and a log-normal shape for normalized games.

This brings us to learning, as illustrated in Fig. 3, where we observe a single equilibrium in the absence of normalization and multiple equilibria with normalization. Notice that in the normalized case, the real difference lies in the concentration level of payoffs, making it easier for players to coordinate toward them. These preferred states generally correspond to the corners of the UV state space. Also note that, due to compromises between payoff matrices, interacting with individuals playing a symmetric but different game effectively reduces one's payoff. This contributes to the overall lower wealth observed under normalization.

Lastly, we run experiments on system-level learning by varying the population-level rationality and risk-aversion parameters. We achieve this by adding a fixed constant to empirically justified distributions of these parameters. We approximate the learning speed of the system by measuring the rate at

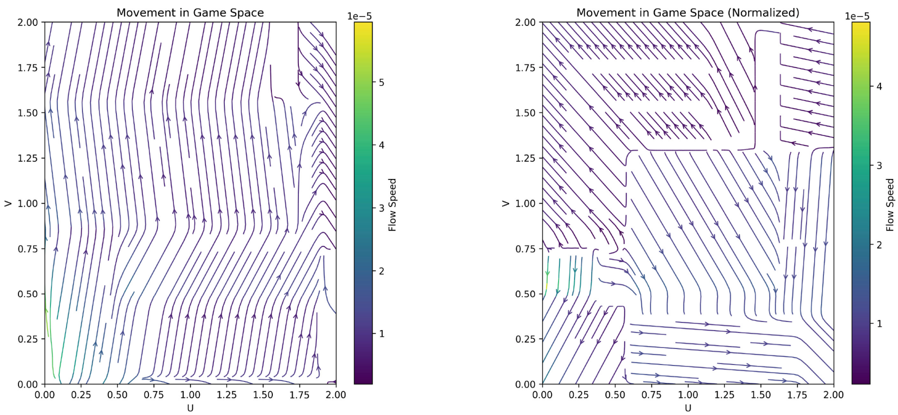

Fig. 3. Movement in the UV state space. Generally, attractors correspond to higher payoffs. In the non-normalized case, the upper right corner is the attractor state, preferred due to high payoffs for both players. In the normalized case, high singular values for both U and V are preferred as payoff concentrates value in one cell.

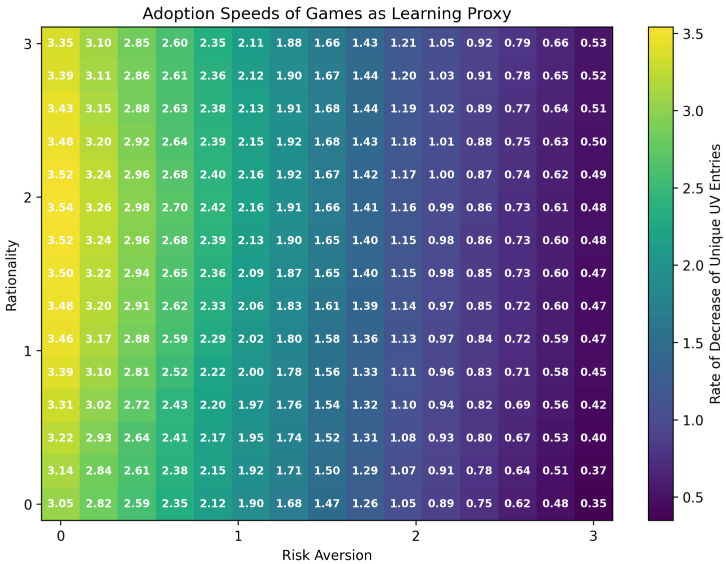

Fig. 4. While keeping the overall shapes of rationality and risk-aversion distributions fixed, we shift the distributions by adding a constant. This results in different system-level learning behaviors, approximated by the speed at which the system converges to the desired games.

which the number of unique games decreases within the population. As individuals converge toward equilibria, more individuals share similar games. We confirm that increased risk aversion decreases the learning speed, an effect stronger than the sensitivity observed with rationality. Interestingly, as rationality increases, adoption speed peaks before reaching maximum rationality values. We attribute this to the fact that when all players closely approach the Nash equilibrium, the normalized games become less differentiated. This mitigates coordination difficulty in a more rational population, thereby reducing payoff differences and slowing game adoption speed (Fig. 4).

4 Discussion

Our results indicate that the system can settle into multiple stable equilibria, leading to enduring wealth stratification. Even with identical initial conditions and rules, some agent communities gravitate toward high-payoff cooperative regimes, while others remain in low-payoff conventions. This finding reinforces the concept of poverty traps: stable low-level equilibria from which escape is difficult [4]. In line with the "poor but rational" perspective, we observe that purely adaptive behaviors under constraints (without external market failures) can lock populations into low-level wealth equilibria [5]. This suggests that persistent inequality arises endogenously through path-dependent dynamics, supporting theories that allow multiple long-run equilibria rather than a single universal outcome [7]. Thus, small differences can yield substantially divergent trajectories, explaining persistent wealth disparities amidst overall economic growth [8,9].

Normalization of payoffs, removing inherent technological advantages, underscores the importance of coordination among agents. Without exogenous innovation, agents prosper only through effective collective action, whereas miscoordination traps groups in suboptimal outcomes. Such dynamics mirror institutional traps described in growth theory, where societies fail to organize collective action necessary for development [4]. Practically, our findings emphasize social cohesion and norm alignment as critical to avoiding coordination failures, especially in innovation-poor contexts.

The interplay between rationality and risk aversion strongly influences equilibrium convergence. Highly rational, risk-averse agents swiftly settle into safe yet suboptimal equilibria, while moderate bounded rationality or reduced risk aversion facilitates exploration and potential discovery of superior outcomes. This outcome aligns with bounded rationality and risk-sensitive decision-making theories, where heuristic-driven choices and aversion to novel strategies prevent optimal outcomes [14–16]. Hence, neither extreme rationality nor excessive caution guarantees optimal collective outcomes—strategic exploration significantly improves long-term performance.

4.1 Limitations and Future Research

Our minimalist approach isolates essential inequality mechanisms but excludes many real-world complexities, potentially understating certain inequality drivers (e.g., inherited advantages or market institutions). Explicitly excluding compounding mechanisms such as capital investment returns further limits the realism [12]. Future work could also examine whether the emergent network structures exhibit real-world properties such as degree heterogeneity, clustering, or small-world topology. Additional extensions might include exploring network-based interventions, incorporating realistic agent behaviors (e.g., memory-driven learning), and validating models against empirical data. Such efforts would bridge theoretical findings with real-world implications, enhancing both the robustness and applicability of these insights.

Acknowledgment. We thank Karin Brinksma for her valuable contribution in developing the simulation code used in this study. We also acknowledge support and resources provided by University of Amsterdam.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Roser M.: The history of global economic inequality. Our World in Data (2017). https://ourworldindata.org/the-history-of-global-economic-inequality

2. Alkire, S., Nogales, R., Quinn, N.N., Suppa, N.: On track or not? Projecting the global multidimensional poverty index. J. Dev. Econ. **165**, 103150 (2023)
3. Mahler, D.G., Yonzan, N., Lakner, C.: The impact of Covid-19 on global inequality and poverty (2022
4. Ghatak M.: Theories of poverty traps and anti-poverty policies. World Bank Econ. Rev. **29**(suppl_1), S77–S105 (2015)
5. Duflo, E.: Poor but rational. Underst. Poverty **24**, 367–379 (2006)
6. Acemoglu, D.: Introduction to Modern Economic Growth. Princeton university press (2008)
7. Banerjee, A.V., Duflo, E.: Growth theory through the lens of development economics. Handb. Econ. Growth **1**, 473–552 (2005)
8. Kraay, A., McKenzie, D.: Do poverty traps exist? Assessing the evidence. J. Econ. Perspect. **28**(3), 127–148 (2014)
9. Barrett, C.B., Carter, M.R.: The economics of poverty traps and persistent poverty: empirical and policy implications. J. Dev. Stud. **49**(7), 976–990 (2013)
10. Aiyar, S., Duval, R., Puy, D., Yiqun, W., Zhang, L.: Growth slowdowns and the middle-income trap. Jpn. World Econ. **48**, 22–37 (2018)
11. De Nardi, M.: Quantitative models of wealth inequality: a survey. National Bureau of Economic Research (2015). https://api.semanticscholar.org/CorpusID:35849844
12. Haushofer, J., Fehr, E.: On the psychology of poverty. Science **344**(6186), 862–867 (2014)
13. Iglesias, J.-R., Gonçalves, S., Abramson, G., Vega, J.-L.: Correlation between risk aversion and wealth distribution. XXPhys. A **342**(1–2), 186–192 (2004)
14. De Bruijn, E.-J., Antonides, G.: Poverty and economic decision making: a review of scarcity theory. Theor. Decis. **92**(1), 5–37 (2022)
15. Morton, J.M.: Reasoning under scarcity. Australas. J. Philos. **95**(3), 543–559 (2017)
16. Banerjee, A.V., Newman, A.F.: Risk-bearing and the theory of income distribution. Rev. Econ. Stud. **58**(2), 211–235 (1991)
17. Puttnam, R.: The role of social capital in development : an empirical assessment (2002). https://api.semanticscholar.org/CorpusID:197648850
18. Wooldridge, M., Jennings, N.R.: Pitfalls of agent-oriented development. In: Proceedings of the Second International Conference on Autonomous Agents, pp. 385–391 (1998)
19. Chetty, R., et al.: NBER working paper series social capital I: measurement and associations with economic mobility (2022). https://api.semanticscholar.org/CorpusID:251255471
20. Epstein, J.M., Axtell, R.: Growing Artificial Societies: Social Science from the Bottom Up. Brookings Institution Press, Washington, D.C. (1996). 9780262550253
21. Föllmer, H.: Random economies with many interacting agents. J. Math. Econ. **1**(1), 51–62 (1974)
22. DiMaggio, P., Garip, F.: Network effects and social inequality. Ann. Rev. Sociol. **38**, 93–118 (2012). https://doi.org/10.1146/annurev.soc.012809.102545
23. Sanchez Carrera, E.J.: Evolutionary dynamics of poverty traps. J. Evol. Econ. **29**(2), 611–630 (2019)
24. Axelrod, R., Hamilton, W.D.: The evolution of cooperation. Science **211**(4489), 1390–1396 (1981). https://doi.org/10.1126/science.7466396
25. Nowak, M.A.: Five rules for the evolution of cooperation. Science **314**(5805), 1560–1563 (2006). https://doi.org/10.1126/science.1133755
26. Santos, F.C., Pacheco, J.M.: Scale-free networks provide a unifying framework for the emergence of cooperation. Phys. Rev. Lett. **95**(9), 098104 (2005)

27. Kasthurirathna, D., Piraveenan, M.: Emergence of scale-free characteristics in socio-ecological systems with bounded rationality. Sci. Rep. **5**(1), 10448 (2015)
28. Sadekar, O., Civilini, A., Gómez-Gardeñes, J., Latora, V., Battiston, F.: Evolutionary game selection creates cooperative environments. Phys. Rev. E **110**(1), 014306 (2024)
29. Mincer, J.: Progress in human capital analyses of the distribution of earnings. In: The Personal Distribution of Incomes (Routledge Revivals), pp. 136–192. Routledge (2013)
30. Gualandi, S., Toscani, G.: Human behavior and lognormal distribution. A kinetic description. Math. Models Methods Appl. Sci. **29**(04), 717–753 (2019)
31. McKelvey, R.D., Palfrey, T.R.: Quantal response equilibria for normal form games. Games Econ. Behav. **10**(1), 6–38 (1995)