Toward Simulating Networked Societies with Formal Institutions Using AI Agents

Michael Richards[1], Danny Cowser[2], Daniel Nielson[2], Jacob W. Crandall[1]

[1]Computer Science Department, Brigham Young University
[2]Department of Government, University of Texas at Austin
michael.richards256@gmail.com, dannycowser@utexas.edu, dan.nielson@utexas.edu, crandall@cs.byu.edu

Abstract

Institutions are key to creating societies that are efficient, fair, and benevolent. Despite their importance, the complexities of human (networked) societies make it difficult to understand how formal institutions form and how they shape human communities. Artificial intelligence (AI) can potentially raise understanding in this regard. Thus, in this paper, we present a simulation model utilizing AI agents to simulate networked societies that contain formal institutions. We then observe the outputs of the resulting model under different societal conditions and formal institutions, and (where applicable) compare and contrast these outputs with political and economic theories. Our model outputs (a) address how inequality impacts societal prosperity, (b) illuminate how institutions can potentially impact poverty, and (c) give insights into the attributes of formal institutions that individuals are inclined to support. These and future simulation models can potentially inform how AI can support the design and development of institutions that facilitate healthier communities and nations.

Supplementary material (code and appendices) —
https://github.com/jakecrandall/AAAI2026-SM.git

Introduction

Effective formal institutions are a key to efficient, fair, and benevolent societies (Acemoglu and Robinson 2012; Ostrom 1990). In addition to spurring economic growth, institutions can be important for reducing poverty (Tebaldi and Mohan 2010), improving health outcomes (Hadipour, Delavari, and Bayati 2023; Antonelli and Marini 2025), and increasing stability (Chowdhury et al. 2024). However, the complexities of human (networked) societies make it difficult to understand how formal institutions form and how they shape human communities (Frey and Lau 1968).

When used appropriately (Poile and Safayeni 2016; Gilbert 2019), simulation models can shed insights into the impact of formal institutions and how to establish them. As such, agent-based models for studying institutions have been created (Gavrilets and Currie 2023; Smajgl, Izquierdo, and Huigen 2010). Many of these simulation models combine an economic game with a political game (Hurwicz 1996; Acemaglu and Robinson 2008; Currie et al. 2021; Tverskoi, Senthilnathan, and Gavrilets 2021). While insightful, we note two opportunities for improving upon these models. First, these models often consider economic interactions that do not capture critical aspects of economic interactions, which can limit their applicability. Second, the formal institutions that can be considered in these models are somewhat restrictive. More sophisticated and flexible simulation models can help us more effectively study formal institutions.

One factor driving the usage of simpler economic models to study institutions is that human behavior is hard to encode in more sophisticated environments. However, increasingly powerful AI tools make modeling such behavior more realistic (Argyle et al. 2023). In this way, AI agents can facilitate the development of higher-fidelity simulation models that better match the intricacies of human societies, including the impact and evolution of institutions within them.

Toward this end, in this paper, we propose and analyze a simulation model of human societies with two intended attributes. First, it implements a somewhat sophisticated, but abstract, economic model based on a strategically rich network game (Skaggs et al. 2024) in which players can acquire wealth and power through economic exchange (trade) and conquest. Second, the model facilitates the study of a wide range of formal institutions under different conditions.

Agents in this model come in two forms: institutional players and society members. Each institutional player follows scripted behavior according to codified rules related to taxation, behavior enforcement, and wealth redistribution. These players can be added to the model in any number and combination. The behavior of society members is intended to follow human behavior. We base this behavior on the CAB algorithm (Skaggs et al. 2024), which captures many aspects of human behavior in our economic game (Skaggs and Crandall 2025). We extend these agents to engage in our model's political process, wherein players affiliate with institutions. The resulting model produces a hierarchical power structure centered around institutions, in which institutional players gain power and help society prosper.

After presenting our model, we explore its outputs in a series of experiments. In so doing, we compare and contrast model outputs with existing theories related to institutions, economic growth, and equality. We consider three questions: (1) how does inequality impact economic growth? (2)

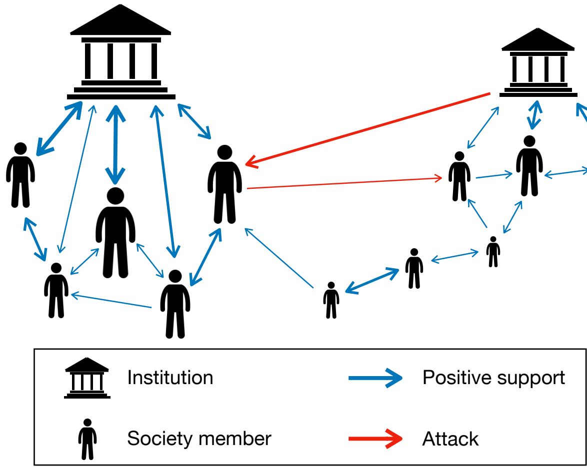


Figure 1: Illustration of our simulation model, which induces a weighted, signed, and directed graph defining the flow of capital among players. Thicker arrows indicate larger exchanges, and larger player icons indicate greater wealth. Any number of institutions can be added to the model.

how do institutions impact economic growth, equality, and poverty? and (3) what are the attributes of institutions that survive and prosper? By studying these questions with this and future models, we hope to better understand how to establish institutions that benefit societies and individuals.

Perspective on Simulation Models

Before presenting and analyzing our model, we highlight foundational considerations regarding simulation models. Specifically, Poile and Safayeni (2016) note that (1) *equifinality* (vastly different models can produce the same outputs) and (2) the many assumptions that must be made in constructing simulation models make it difficult to prove any theory using simulation models. As such, Poile and Safayeni (2016) argue that the primary value of simulation models is the iterative design process itself. In this process, researchers identify theories that are both supported and unsupported by the model. They then jointly consider how the model could be changed to conform with theories and how the model may potentially provide insights a theory may have glanced over.

We adopt this perspective. While we have attempted to create an effective model, we do not consider it to be a final product. Thus, rather than seeking to prove that the model predicts real-world outcomes, we compare and contrast our model outputs with various political and economic theories. When our model contrasts with a theory, we consider that interesting. We thus consider why the model fails to explain the theory and how the model or theory could potentially be changed to achieve conformity. In this way, our model can offer valuable insights despite these foundational issues.

The Simulation Model

Our model is illustrated by Figure 1. As in prior work (Hurwicz 1996; Acemaglu and Robinson 2008; Currie et al. 2021; Tverskoi, Senthilnathan, and Gavrilets 2021), this model combines an economic game with a political game. In this section, we describe these games along with the algorithms used to model institutions and society members.

The Economic Game

The economic game defines the flow of capital among players due to trade and conquest. We model this flow of capital using a strategic network game called the Junior High Game (JHG) (Skaggs et al. 2024), which induces a weighted, signed, and directed graph defining the flow of capital.

In the JHG, players seek to become wealthy. Initially, all players are assigned a wealth. In each round, each player has N tokens (representing capital), each of which it can (1) *keep* (to maintain its own wealth and defend itself), (2) *give* to another player (to increase that player's wealth), or (3) use to *attack* a player (which steals wealth from that player). After all players allocate their tokens, the wealth of each player is updated based on these token allocations, and a new round then begins. Importantly, allocations from more wealthy players have greater impact on wealth, thus inducing power asymmetries into the society.

JHG scenarios can be configured via trade and attack coefficients. For simplicity, we assume these coefficients are the same for all players. The parameter settings for the JHG that we use are given in the supplementary material (SM 1).

Using the JHG as the economic game offers several benefits. First, it provides strategic flexibility in an abstract setting, modeling both positive and negative exchanges between pairs of players as well as power asymmetry. These properties of the real world are not all simultaneously modeled in public goods games and other commonly used economic games. Second, the JHG allows institutions to be added as players. No special rules are needed to encode institutions into the simulation. This gives substantial flexibility with regards to the number and type of institutions it can model. Third, once institutional players are added to the game, a political game naturally follows.

Institutional Behavior

In contrast to existing work (Pryor 1973; Acemaglu and Robinson 2008; Currie et al. 2021; Tverskoi, Senthilnathan, and Gavrilets 2021) wherein the institution is specified in game rules, we take an agent-based approach in which institutions are simply players in the game. In our model, institutional players potentially levy taxes or membership fees, punish violations of their codified rules, and redistribute capital. In the JHG, these functions are carried out in each round via token allocations made by the institutional player.

Our institutional players follow scripted behavior defined by several parameters (Table 1). These parameters can be set to model rules of many kinds of institutions, thus allowing us to probe theories about inclusive verses extractive institutions, collective action, state formation, etc. In each round, these players first determine and announce how much taxes or membership fees (number of tokens) that each society member should pay to the institution. The institutional player then allocates tokens to defend itself against attacks from other players. Next, the institutional player allocates tokens to punish other players for violations they have

Parameter	Possible values
Fee schedule	{Progressive, Flat, Regressive}
% Power	$[0, 100]$
Redistribution	{Progressive, Regressive, Proportional, Merit}
Remuneration	{Yes, No}
Punish nonpayment	{Yes, No}
Punish attacks	{Yes, No}
Theft threshold	≥ 0
Punish strength	≥ 0
Extractive	{Yes, No}

Table 1: Parameters of institutional players.

committed, such as tax evasion or attacks on its members. Third, if the institutional player still has remaining tokens, it gives tokens to society members according to its distribution schedule. Algorithmic details are given in SM 1.

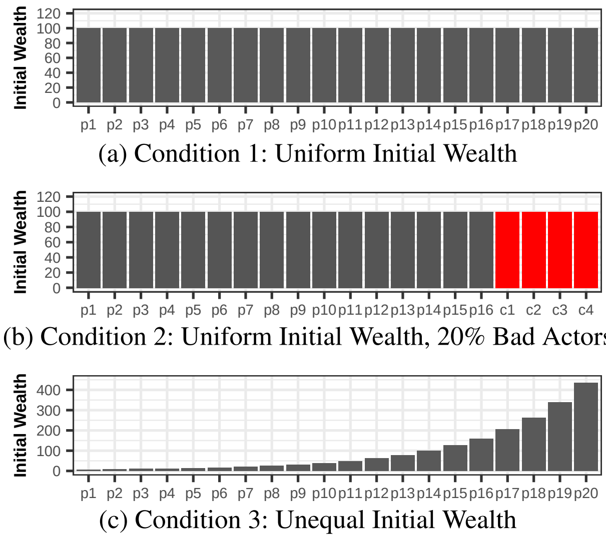
The Behavior of Society Members

While LLMs are powerful tools for modeling human behavior in some tasks (e.g., (Argyle et al. 2023; Hashemi and Macy 2025), they are currently not well-suited for many numerical reasoning tasks such as optimizing economic outcomes (Rahman 2025; Morge 2025). Thus, we instead use the CAB algorithm (Skaggs et al. 2024), developed for playing the JHG, to encode behavior into society members. A CAB agent follows a three-step process for determining token allocations. First, it identifies communities based on prior token allocations. Second, it determines which community it would like to belong to. Third, it allocates tokens to form and strengthen the desired community. These processes are modulated by a set of parameters Θ. Skaggs and Crandall (2025) computed a set of 100 parameterizations of the CAB algorithm that together approximate human behavior in the JHG. Thus, we utilize these 100 parameterized CAB agents, which we refer to as hCABs, to model society members. Let Φ denote this set of parameterizations.

We modify hCABs in three ways to adapt them to scenarios with institutions. First, to maintain transactional relationships between formal institutions and non-institutional players, we constrain hCABs to not give more than specified fees to institutional players. We do not, however, prohibit them from attacking institutional players.

Second, we add an additional parameter θ_{aff} to the agents to dictate institutional affiliation. For simplicity, we assume an agent only affiliates with (i.e., pays fees to) a single institution (if any). Thus, given two institutional players denoted 1 and 2, then $\theta_{\text{aff}} \in \{1, 2, -1\}$, where -1 denotes no affiliation. The set of parameterizations is, then, $\Phi' = \Phi \times \theta_{\text{aff}}$.

Finally, to tune agent behavior in a specific *situation*, we evaluate each $\phi \in \Phi'$ by running a series of training simulations. We define a situation to be the configuration of institutional players in society and the agent's initial economic standing: lower-, middle-, or upper-class. Within simulations, parameterizations are selected (to dictate each agent's behavior) proportional to the economic growth they are estimated to provide agents, such that more successful parameterizations are more likely to be selected (see SM 1).



(a) Condition 1: Uniform Initial Wealth

(b) Condition 2: Uniform Initial Wealth, 20% Bad Actors

(c) Condition 3: Unequal Initial Wealth

Figure 2: Three societal conditions considered in our simulations. Dark bars indicate players that follow hCAB strategies. Red bars indicate *bad actors* (CATs).

In some of our experiments, we also find it useful to observe societal outcomes in the presence of *bad actors*. We model these bad actors as Conspiring Autonomous Thieves (CATs) *Skaggs et al.* (Skaggs et al. 2024). CATs coordinate to iteratively attack the weakest, most vulnerable members of society, which decimates individuals and drains wealth from the society. Overcoming them requires some form of collective action from members of society.

The Political Game

The political game considered in our simulation model follows directly from the prior three model components. That is, we insert multiple institutional players, which potentially differ in the rules they embody, into the JHG. At the beginning of each scenario, each of these institutional players is given no wealth, and as such has no power. Only when society members choose to pay fees do institutional players become empowered. The more (and stronger) society members that pay fees to the institutional player, the more power it gains due to JHG dynamics. Thus, affiliation choices made by hCABs constitute choices made in the political game.

Results: Already-Established Governments

To begin to explore our model, we consider its outputs in the absence of a political process. That is, we conducted simulations in which a single institutional player, designed to optimize some criteria, is empowered as the government. We then observe model outputs regarding mean wealth and inequality under that government.

Experiments

We conducted simulations under the three different conditions illustrated in Figure 2. In addition to the government

Name	Description of objective function
None	No government
Prioritize equality	Minimize the Gini Index, with only minor consideration for citizen wealth
Balanced priorities	Equal consideration for maximizing citizen wealth and minimizing the Gini Index
Prioritize ave. wealth	Maximize the mean wealth of citizens, with only minor consideration for Gini Index
Prioritize gov. wealth	Maximize the wealth of the government

Table 2: Objective functions used to derive governments.

	Prioritize Equality	Balanced Priorities	Prioritize Mean Wealth	Prioritize Gov Wealth
	Cond. 1: Uniform Initial Wealth			
Fee Schedule	Progressive	Flat	Progressive	Flat
Redistribution	Regressive	Regressive	Regressive	Merit
% Power	0.0	3.0	1.5	49.5
	Cond. 2: Uniform Initial Wealth, 20% CATs			
Fee Schedule	Progressive	Progressive	Flat	Regressive
Redistribution	Progressive	Progressive	Merit	Merit
% Power	29.5	30.5	27.0	47.0
	Cond. 3: Unequal Initial Wealth			
Fee Schedule	Flat	Flat	Progressive	Flat
Redistribution	Progressive	Merit	Regressive	Merit
% Power	39.0	24.5	5.5	47.5

Table 3: Selected government parameter settings computed for each condition and objective function. Full parameter settings and descriptions are given in SM 1 and 2.

player, each society consists of 20 non-institutional players, each of which can be viewed as an individual in a small society or a cluster of individuals in a larger society. In the first condition, all non-institutional players begin the scenario with the same wealth and follow hCAB strategies. The second condition is identical except that four players are *bad actors* (CATs). In the third condition, non-institutional players begin the scenario having different initial wealth. The wealth distribution roughly models the distribution of total wealth found in the Great Britain in 2025 (Francis-Devine 2025), which has a wealth Gini Index (Gini 1921) of 0.59.

For each of the three societal conditions, we simulate societies with five distinct government conditions, the first of which has no government player. The other four contain government players whose parameters are selected, using genetic evolution, to maximize the objective functions described in Table 2. Details about these objective functions and the optimization algorithm are given in SM 2.

In these experiments, the government player begins the simulation already empowered. To do this, we transfer initial wealth from each non-institutional player to the government player. That is, if the government's *power* parameter is 30% (i.e., the government is designed to own 30% of society's wealth), 30% of each non-institutional player's initial wealth is transferred to the government prior to the first round.

Results are based on 100 30-round simulations conducted for each condition and government.

Model Outputs

Table 3 overviews the governments computed for each objective function and condition. Scatter plots showing mean wealth and Gini Index after 30 rounds for each simulation are given in Figure 3. Additional data, including the actions of governments and society members, are provided in SM 2.

We make several observations. First, Table 3 shows that the ideal government depends on both societal conditions and government objectives. No one set of government parameters is ideal across all scenarios. This outcome is consistent with some political theory (e.g., (Andrews 2010)).

Second, under uniform initial wealth (Condition 1), the governments selected to maximize each objective function (Table 3) have little or no power, except when government wealth is prioritized. This suggests that society members are able to effectively regulate themselves when they have relatively equal wealth and power. In this condition, strong governments that redistribute wealth and punish tax evasion and stealing lead to less economic growth and higher Gini Index.

Third, in Condition 2 (when 20% of players are bad actors), society suffers with respect to both overall wealth and equality. Thus, in these societies, all government objective functions produce strong governments that hold a high percentage of society's wealth. This gives the government power to combat the bad actors, and in so doing quickly renders them powerless before too much damage is done. Once the government reduces the power of the bad actors, these governments then give back to members of society in good standing, leading to some prosperity. However, even with these strong governments, these societies end up with less total wealth than those in Condition 1.

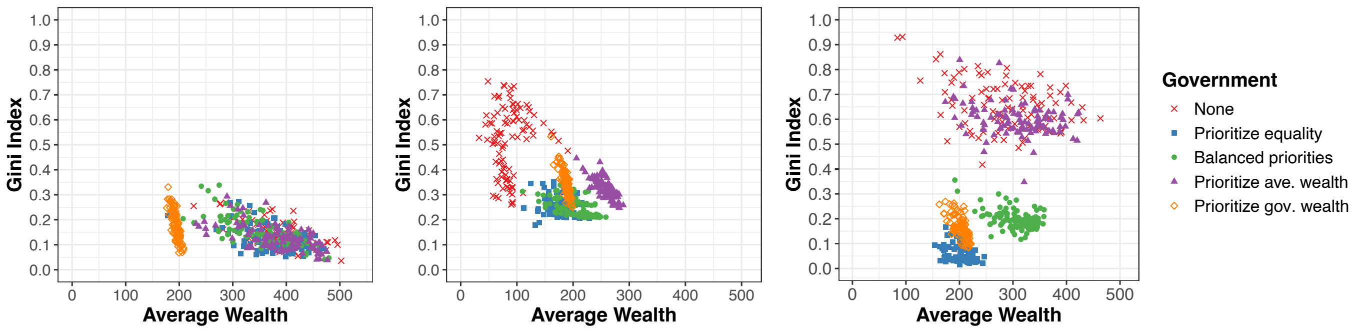
Fourth, under unequal initial wealth (Condition 3), maximizing most objective functions (except prioritizing individual wealth) produces strong governments. For example, with the government selected for balanced priorities, society initially has a high wealth Gini Index (0.59). However, the government uses a flat tax coupled with merit-based redistribution to substantially lower the Gini Index (to about 0.20 on average) while also substantially increasing total societal wealth (by 38% over the society with no government).

Finally, a comparison of Figures 3a and c shows that, in our model, society appears to be better off with respect to both societal wealth and wealth equality when wealth inequality was initially low. This raises the following question: How does wealth inequality impact economic growth?

Results: Inequality's Impact on Prosperity

Our model's answer to this question is shown in Figure 4, which shows the average economic growth of our simulated societies in the absence of institutional players for a range of initial wealth distributions. The figure shows that economic growth substantially decreases as inequality increases.

Whether or not inequality impacts (or even is correlated with) economic growth in human societies has been hotly contested. Some analysis of data suggests that there is indeed a negative relationship between Gini Index and eco-



(a) Uniform initial wealth (b) Uniform wealth, bad actors (c) Unequal initial wealth

Figure 3: Average individual wealth and Gini Index after 30 rounds.

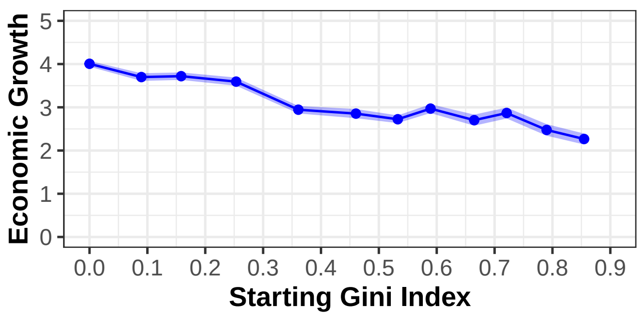


Figure 4: The impact of inequality on economic growth without a government. Economic growth is measured as the number of times wealth increases over 30 rounds.

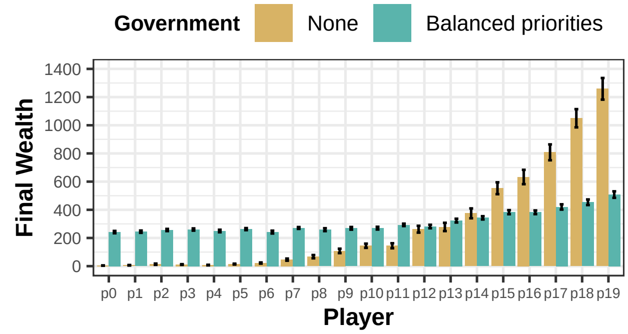


Figure 5: Impact of government on player wealth in societies with initially unequal wealth after 30 rounds.

nomic growth (Ostry, Berg, and Tsangarides 2014; Lee 2012). One proposed reason for this trend is that higher inequality may lead to lower demand, which slows economic growth (Bivens 2017). Our model backs this theory. Players tend to keep more (less trade/lower demand) and attack more under unequal initial wealth than under equal initial wealth.

However, others argue that there is a more nuanced relationship between inequality and economic growth. Kuznets (1955) promoted a U-shaped relationship wherein developing nations experience a surge in both inequality and economic growth as parts of the economy begin to thrive. This increase is followed by a reduction in inequality as the nation further develops. Lederman and Brückner (2015) argue that economic growth is increased by inequality in poor countries, but is decreased in high- and middle-income countries.

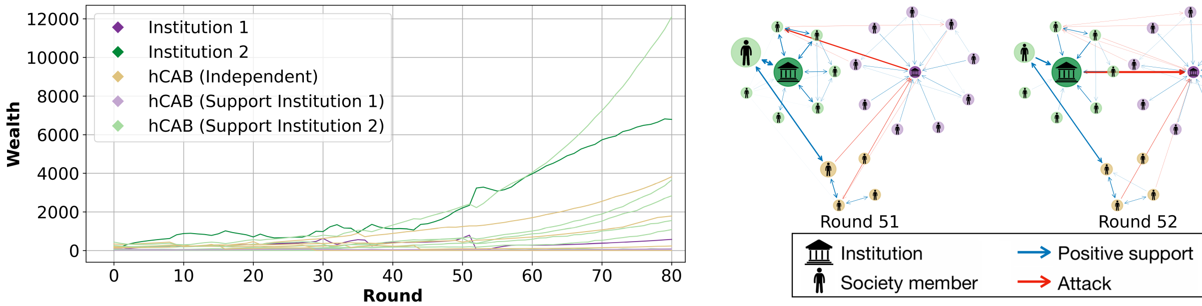
The difference between these latter theories and our model causes us to reflect: are these theories incorrect or is our model flawed? Could our model be altered in a reasonable way so that it can support these theories? As we consider these questions, we can observe that the simplifying assumption we made in our model regarding equal trade and attack coefficients for all players could prohibit our model from conforming with these theories. However, the model could potentially be altered. Lower trade coefficients could be used to model less developed societies. As some players within these societies innovate, these trade coefficients could be increased for those players, thus possibly causing them to become more wealthy than other players. Future work could consider how to vary these coefficients in a realistic way to simulate these trends as well as other potential alterations (e.g., agent behaviors), after which the revised model could again be compared against these theories.

Regardless, while equalizing wealth across society does increase economic growth in our current model, it would not be beneficial to wealthier society members. To see this, consider Figure 5, which compares player wealth in a society with no government to player wealth in a society with the balanced-priorities government from Condition 3. The balanced government produces vastly higher wealth for the initially poorer individuals in society. However, the originally wealthier players are worse off with this government.

Would society select such a government? 14 out of 20 society members could potentially be better off with the balanced government. However, over 76% of the wealth initially held in the society is held by the six most wealthy players. As such, these individuals are likely to possess much of the *de facto* (if not also the *de jure*) political power in society, thus allowing them to avoid becoming subject to such a government. We investigate this question in the next section.

Results: Who Gets Their Way?

To address the question with our model of whether the majority or elites get their way in this scenario, we conduct an



(a) Change in wealth over time

(b) Exchange network for selected rounds

Figure 6: A scenario in which two institutional players come into conflict.

Class	Inst. 1: Merit	Inst. 2: Proportional	None: Independent
Lower	69.8%	22.1%	8.1%
Middle	2.0%	42.2%	55.9%
Upper	1.5%	55.7%	42.9%
All Players	39.2%	34.5%	26.3%

Table 4: Percent of players from each class that supported each institution.

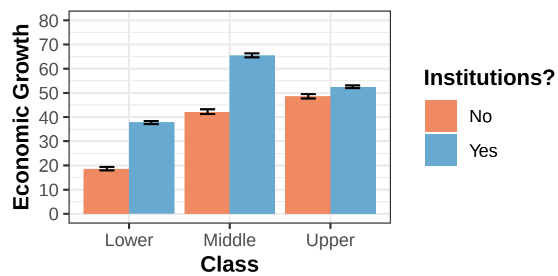


Figure 7: The economic growth of players in each class in (1) societies with no institutions (No) and (2) societies in which players are given a choice between institutions (Yes).

experiment using our full simulation model, wherein we add in the model's political process.

Experiment

For this experiment, we add two institutional players into Condition 3. These two institutions differ in the rules they encode. The first institution implements policies similar to the balanced government. Its policies are designed to lower wealth inequality. The second institution is identical except that it redistributes wealth proportionally to the amount of fees paid. As such, it is designed to not substantially impact the distribution of wealth. Full parameter settings for these two institutional players are provided in SM 3.

To account for potential randomness in training, we tuned ten different sets of hCABs for this scenario. We evaluated each set across 100 simulations. Results presented are an average of the resulting 1000 simulations. To allow the institutions to establish themselves and for society to evolve thereafter, we conducted 80-round games rather than just 30.

Model Outputs

Table 4 shows the average percentage of players, across all simulations, that selected each institution in each class. Overall, about 39% of players chose to support Institution 1, 35% supported Institution 2, and 26% chose to not support either institution. There was wide variation across classes, however. Players from lower classes tended to support Institution 1, while middle- and upper-class players tended to either support Institution 2 or to remain independent.

While more players chose to support Institution 1, it held, on average, far less wealth than Institution 2. Across all rounds of the simulations, Institution 1 held just 3.8% of societal wealth, while Institution 2 held 13.9%. This is because Institution 2 was supported by more wealthy individuals.

In many scenarios, the two institutional players are able to co-exist, each gaining enough power to help their members. However, in other scenarios, the two institutional players come in conflict leading one (typically Institution 2) to dominate the other. Figure 6 illustrates one such simulation. In this simulation, both institutions began to rise. However, in round 35 and again in round 50, a player affiliated with Institution 1 was attacked by a player affiliated with Institution 2. This caused Institution 1 to retaliate to protect its attacked member (e.g., Round 51), which in turn caused Institution 2 to attack Institution 1 (e.g., Round 52).

Despite these conflicts, the presence of these two institutions vastly improves the overall wealth of the society on average, particularly for lower- and middle-class players (Figure 7). That said, wealth inequality still increases over time with both institutions present (Gini Index grows from 0.59 to 0.64), but not as quickly as it does without institutions (where it grows to 0.72 over the same time). The presence of these two institutions greatly improves society with respect to both wealth and wealth equality.

We note two primary reasons that institutions help lower- and middle-class players so substantially in our model. First, with no institutions, lower-class players have difficulty finding trading partners, and, as a result, end up keeping 35% of their tokens. With the two institutions present, the amount of keeping among lower-class players falls to just 14%, as players can trade with the institutions and also often find collaborations with players affiliated with the same institution (Figure 6). Second, given institutions, upper-class players attack lower- and middle-class players less frequently.

These results are consistent with theories that state that institutions are a fundamental cause of long-term economic growth (Acemoglu, Johnson, and Robinson 2005), and that these institutions should be inclusive (Acemoglu and Robinson 2012). In this case, all players had access to institutions that were beneficial to them. Future work could verify whether this simulation model continues to match these economic and political theories in more scenarios, such as those in which broad access to institutions is not provided.

Results: Which Institutions Are Chosen?

In the previous section, we considered how society would choose between institutions with merit-based verses proportional redistribution policies. In this section, we consider other institutional choices with the goal of understanding the characteristics of institutions that tend to be selected.

Experiments

These experiments are identical to the experiments reported in the last section, except that we consider alternative institutional choices. These choices are: (1) an institution that remunerates victims verses one that does not; (2) institutions that differ in the amount of fees they require, one requiring more fees, the other less; (3) institutions that differ with respect to how strongly they punish attacks on their members, one being more vengeful than the other; and (4) institutions that are identical except that one institution is *extractive*, meaning it attacks society members that are not affiliated with it without cause. Parameter values for each institutional player are given in SM 3.

Model Outputs

Results are summarized in Table 5. We make two observations. First, with respect to both economic growth and Gini Index, all combinations of institutions are better for society than having no institutions. In each configuration, the majority of society affiliates with an institution. Second, institutional players designed to become powerful, using higher fees, more vengeance, and extraction, tend to receive larger support than those that do not (see Configurations 4-6). In our model, these institutions offer their members greater protection and transfer wealth more readily from non-members to members.

This latter result is particularly salient in the case of an extractive institution (Configuration 6). Since the extractive institution attacks non-members without cause, it is dangerous to not be affiliated with it once it is strong. As such, the agents learn to almost exclusively support it. However, once everyone joins this institution, it stops being able to extract wealth (because everyone is a member). As such, this institution eventually creates a strong and prosperous society.

This latter result is reminiscent of theories developed in the study of institutional formation in Congo (Sánchez de la Sierra 2020) and Iraq (Allen, Bertazzini, and Heldring 2023). In both cases, government institutions arose from extractive firms which became sufficiently strong that they eventually morphed into de facto governments.

	Societal Choice	**(%)**	**Economic Growth**	**Gini Index**
Config 1	No institutions		43.9±0.7	0.723±0.003
Config 2	Proportional:	34	62.8±0.5	0.638±0.002
	Merit:	39		
	Independent:	26		
Config 3	Remuneration:	32	67.3±0.6	0.637±0.002
	No Remuneration:	32		
	Independent:	36		
Config 4	Lower fees:	17	67.7±0.6	0.613±0.002
	Higher fees:	59		
	Independent:	24		
Config 5	More Vengeful:	47	68.1±0.5	0.639±0.002
	Less Vengeful:	18		
	Independent:	35		
Config 6	Extractive:	87	71.6±0.4	0.653±0.002
	Not Extractive:	8		
	Independent:	5		

Table 5: Summary results for societies with choices between two institutions. ± indicates standard error of the mean.

That said, we note that our model predicts that extractive institutions thrive only when when they are allowed to become strong. In our model, the extractive institution does not become strong when another institution is already empowered, as this empowered institution tends to mitigate the extractive institutional player once it attacks its members. Future work could further explore the conditions under which extractive institutions gain and maintain power.

Conclusion

We presented a simulation model of human (networked) societies containing formal institutions. The model is based on a strategically rich network game that allows institutional players to be flexibly added to the simulation in any number and type. Agent behaviors are encoded using AI algorithms designed to model human behaviors. Our analysis shows that many of the model outputs appear to match a variety of political and economic theories, but not others. In cases of mismatch, we have hypothesized ways in which the model could change to potentially match those theories.

Future work should address limiting assumptions made in our current model. For example, our AI agents are trained under the assumption that they will select institutions based on their own self interest. However, other approaches could potentially be more accurate. Additionally, our current model made the simplifying assumption that all society members have the same trade coefficients. While determining how to effectively relax this assumption is challenging, doing so could potentially allow the model to conform with more nuanced political theories regarding the impact of inequality on economic growth. As the AI community continues to build increasingly effective models of human societies, we can better understand how to establish formal institutions that are fair, efficient, and benevolent.

References

Acemaglu, D.; and Robinson, J. A. 2008. Persistence of Power, Elites, and Institutions. *American Economic Review*, 98: 267–293.

Acemoglu, D.; Johnson, S.; and Robinson, J. A. 2005. Chapter 6 Institutions as a Fundamental Cause of Long-Run Growth. In Aghion, P.; and Durlauf, S. N., eds., *Handbook of Economic Growth*, volume 1, 385–472. Elsevier.

Acemoglu, D.; and Robinson, J. A. 2012. *Why nations fail: The origins of power, prosperity, and poverty*. Currency.

Allen, R. C.; Bertazzini, M. C.; and Heldring, L. 2023. The Economic Origins of Government. *American Economic Review*, 113(10): 2507–2545.

Andrews, M. 2010. Good Government Means Different Things in Different Countries. *Governance*, 23(1): 7–35.

Antonelli, M. A.; and Marini, G. 2025. Do institutions matter for citizens' health status? Empirical evidence from Italy. *The European Journal of Health Economics*, 26: 95–115.

Argyle, L. P.; Busby, E. C.; Fulda, N.; Gubler, J.; Rytting, C.; and Wingate, D. 2023. Out of One, Many: Using Language Models to Simulate Human Samples. *Political Analysis*, 31(3): 337–351.

Bivens, J. 2017. Inequality is slowing U.S. economic growth. Technical report, Economic Policy Institute.

Chowdhury, M. A. F.; Prince, E. R.; Shoyeb, M.; and Abdullah, M. 2024. The threshold effect of institutional quality on sovereign debt and economic stability. *Journal of Policy Modeling*, 46(1): 39–59.

Currie, T. E.; Campenni, M.; Flitton, A.; Njagi, T.; Ontiri, E.; Perret, C.; and Walker, L. 2021. The Cultural Evolution and Ecology of Institutions. *Philosophical Transactions of the Royal Society London B*, 376: 20200047.

Francis-Devine, B. 2025. Wealth in Great Britain. Technical report, House of Commons Library.

Frey, B.; and Lau, L. J. 1968. Towards a mathematical model of government behaviour. *Journal of Economics*, 28: 355–380.

Gavrilets, S.; and Currie, T. E. 2023. Mathematical models of the evolution of institutions. In Richardson, P. J.; Bednar, J.; Currie, T. E.; Gavrilets, S.; and j. J. Wallis, eds., *Institutional Dynamics and Organizational Complexity: How Social Rules Have Shaped the Evolution of Human Societies Throughout Human History*, 101–112. Open Access Book, Cultural Evolution Society, 3rd edition.

Gilbert, N. 2019. *Agent-based models*. Sage Publications.

Gini, C. 1921. Measurement of Inequality of Incomes. *The Economic Journal*, 31(121): 124–126.

Hadipour, A.; Delavari, S.; and Bayati, M. 2023. What is the role of institutional quality in health outcomes? A panel data analysis on 158 countries from 2001-2020. *Heliyon*, 9(9): e20251.

Hashemi, F.; and Macy, M. 2025. Collective Social Behaviors in LLMs: An Analysis of LLMs Social Networks. In *Proceedings of the SciSoc LLM Workshop: Large Language Models for Scientific and Societal Advances (SciSocLLM @KDD 2025)*. ACM, New York, NY, USA, 16 pages.

Hurwicz, L. 1996. Institutions as Families of Game Forms. *Japanese Economic Review*, 47(2): 113–132.

Kuznets, S. 1955. Economic Growth and Income Inequality. *The American Economic Review*, 45(1): 1–28.

Lederman, D.; and Brückner, M. 2015. Effects of Income Inequality on Aggregate Output. Technical report, World Bank Policy Discussion Paper 7317.

Lee, D. B. 2012. *Occupy This: The Effect of Income Inequality on GDP Per Capita Growth Using Panel Data in the United States from 1963 to 2009*. CMC Senior Thesis. 311, Claremont College.

Morge, M. 2025. Is Generative Artificial Intelligence Ready for Computational Social Science? In *LNAI*. Mathieu, Philippe and Corchado, Juan Manuel and De La Prieta, Fernando, Lille, France: Springer.

Ostrom, E. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press.

Ostry, J. D.; Berg, A.; and Tsangarides, G. D. 2014. Redistribution, Inequality, and Growth. Technical report, IMF Staff Discussion Note No. SDN/14/02.

Poile, C.; and Safayeni, F. 2016. Using computational modeling for building theory: A double edged sword. *Journal of Artificial Societies and Social Simulation*, 19(3).

Pryor, F. L. 1973. Simulation of the impact of social and economic institutions on the size distribution of income and wealth. *The American Economic Review*, 63(1): 50–72.

Rahman, R. 2025. Large Language Models in Numberland: A Quick Test of Their Numerical Reasoning Abilities. *arXiv:2504.00226*.

Sánchez de la Sierra, R. 2020. On the Origins of the State: Stationary Bandits and Taxation in Eastern Congo. *Journal of Political Economy*, 128(1): 32–74.

Skaggs, J.; Richards, M.; Morris, M.; Goodrich, M. A.; and Crandall, J. W. 2024. Fostering Collective Action in Complex Societies using Community-Based Agents. In *Proceedings of the International Joint Conference on Artificial Intelligence*. Jeju, Korea.

Skaggs, J. B.; and Crandall, J. W. 2025. Modeling Human Behavior in a Strategic Network Game with Complex Group Dynamics. *arXiv:2505.03795*.

Smajgl, A.; Izquierdo, L. R.; and Huigen, M. 2010. Rules, knowledge and complexity: How agents shape their institutional environment. *Journal of Modelling & Simulation of Systems*, 1(2).

Tebaldi, E.; and Mohan, R. 2010. Institutions and poverty. *The Journal of Development Studies*, 46(6): 1047–1066.

Tverskoi, D. A.; Senthilnathan, A.; and Gavrilets, S. 2021. The Dynamics of Cooperation, Power, and Inequality in a Group-structured Society. *Scientific Reports*, 18: 18670.