IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, December 2025
Impact Factor: 7.67

Intelligent Economic Simulation:Genetic Agents and RL-Based Policy Optimization

Ranjit Choudhary
Assistant Professor
Amity Institute of Information Technology, Amity University, Patna
ranjeet2005it@gmail.com

Abstract: *This paper presents GAO-RL (Genetic Agent Optimized Reinforcement Learning), a hybrid economic simulation framework that integrates Genetic Algorithms (GA) with Reinforcement Learning (RL)[2][17] in an Agent-Based Modeling (ABM)[11][12]environment. Traditional models lack adaptability, while existing Multi-Agent Reinforcement Learning (MARL) [2][5] approaches suffer from scalability and weak theoretical grounding. GAO-RL bridges this gap by combining the exploratory efficiency of GA with the policy stability of RL, anchored in theBewley–Aiyagari model of macroeconomics.*

The framework models an economy with households, firms, a central bank, and a government, each pursuing distinct objectives. Genetic agents enable diverse behavioral exploration, while PPO-based RL agents ensure stable policy optimization. This synergy produces a more realistic and interpretable simulation of dynamic economic systems, offering a robust foundation for data-driven economic policy design and optimization.

Keywords*:* Genetic Agent, Agent-Based Modeling (ABM), Multi-Agent Reinforcement Learning (MARL), Genetic Agent Optimized Reinforcement Learning (GAO-RL**),** Economic Simulation, Policy Optimization

I. INTRODUCTION

Anation'seconomichealth,itssocialwelfareprograms,anditscitizens'qualityoflifeareallinextricablylinked to the design and implementation of fiscal and monetary policies.However, the complexity of these systems makes them notoriously difficult to model and predict[9].The sheer number of interacting agents—from individuals and households to firms and governments—and the nonlinear, often unpredictable, nature of their decisions pose a significant challenge to traditional economic modeling approaches.

Currently Agent-Based Models (ABMs)[11][12] are widely used to simulate economic systems.ABMs rely on simple behavioral rules for their agents, limiting their ability to capture the nuanced and adaptive nature of human decision-making.These models, in their current form, do not show agents learning in dynamic environments.

Multi-Agent Reinforcement Learning (MARL)[2][5] has emerged as a promising alternative.MARL agents can learn complex strategies through trial and error, making the simulation more realistic.However, current MARL approaches face two significant limitations:a lack of scalability and a weak foundation in established economic theory.Without it the results from MARL simulations can be difficult to interpret and use for real-world policy recommendations.

To address these shortcomings, we introduce GAO-RL (Genetic Agent Optimized Reinforcement Learning), a novel economic simulation framework that bridges the gap between sophisticated machine learning and rigorous economic theory. Our approach is grounded in the Bewley-Aiyagari Model[1][4], a cornerstone of modern macroe-conomics known for its robust theoretical foundations in analyzing household consumption, savings, and wealth distribution. By borrowing core economic principles from this model, we imbue our multi-agent system with the credibility and theoretical rigor that current MARL applications often lack.

GAO-RL represents a novel, hybrid approach that uses a population of Genetically Optimized agents alongside Reinforcement Learning-based agents. Genetic agents are computationally efficient to train and, thanks to the principle of mutation, are less likely to be trapped in local minima, allowing them to explore a wider range of behaviors. In contrast, RL agents learn more gradually and systematically, which can make them more susceptible to becoming stuck in suboptimal policy solutions. By combining these two methods, our framework leverages the rapid, diverse exploration of genetic algorithms with the precise, deliberate policy refinement of reinforcement learning, resulting in a more robust and effective simulation.

II. RELATED WORK

2.1 Classic Models

Classical economic models provide a well-researched and tested basis for modeling economic activities and explaining economic phenomena. Regarding the optimal taxation policy problem, the Ramsey Cass Koopmans (RCK) model[10][6] studies the consumption and savings decisions of representative agents but ignore individual choices. The Diamond-Mirrlees model checks the role of labor taxes and supply in social welfare but overlooks asset taxes. The Overlapping Generations (OLG) model[15] emphasizes generational wealth inheritance and resource transfers. In contrast, the Bewley-Aiyagari model[1][4] assesses the impact of taxation on growth, wealth distribution and welfare while simulating realworld income disparities and risk-bearing capacity of individuals.

Traditionally these models are studied by economists using mathematical techniques[7] to address decision making processes related to governments and households. However, these approaches oversimplify decision makers and fail to consider their ability to adapt to the changes in the environment while finding exploits in existing policies.

2.2 Agentic Simulation

Recently agentic simulation has been used in an attempt to find new policies that can increase productivity.These Include ABMs[11][12] which model agents as rule based with no learning capabilities. These models have largely been surpassed by MARL (Multi Agent RL)[5] improves on this by letting the agents change behavior based on reinforcement learning, but these models usually represent the agents as an aggregation of the population[2] quashing outliers making them unable to model the unique behaviors that arise due to the creativity of individuals in any real world system. We have Especially gone through 3 such approaches/attempts.

• TaxAI [13]: Implements MARL with four types of agents: Households, Firm, Financial Intermediary, Gov- ernment. It models dynamics using the Bewley-Aiyagari economic model which focuses on two aspects of modelling: heterogeneity within the group of households, and idiosyncratic shocks at the level of individual households. The BA model allows the simulation to incorporate the economic outcomes of income inequality (due to heterogeneity), as well as the phenomenon of increased household saving in the face of uncertainty (due to idiosyncratic shocks).

• ABIDES [8]: Much like TaxAI, implements MARL with the four types of agents, with the difference being, unlike in TaxAI, that the Firms may be numerous. While it doesn't implement a standard economic model, the dynamics is quite similar to the BA model used in Tax AI. The Households (and even the Firms) are heterogeneous; however, here the exogenous shocks are applied to individual Firms instead of individual Households.

• AI Economist [17]: Implements a MARL framework with a social planner (government) and multiple het- erogeneous agents (households). It uses a custom "Gather-Trade-Build" economic simulation where agents learn to perform labor, trade, and earn income in a dynamic environment. The core of the framework is the co-adaptation between the planner, which learns optimal tax policies, and the agents, which learn rational, utility-maximizing behaviors in response.

These papers all show results showing superiority of heterogenous agents over single agents representing the average behaviour of the populus. And we move forward with this approach/assumption.

Where we split from these works is in using Genetic Algorithm to optimise deep networks[14][3] used by the agents to make decisions. Works both recent and old demonstrate the capability of GA for quick convergence when training networks. We suspect that this compute advantage will be useful in an environment where there are many agents.

Table 1: Overview of Learning Models / Algorithms Used in the Study

Model / Algorithm	Primary Objective	Strengths	Limitations	Role in GAO-RL Framework
Genetic Algorithm (GA)	Finds optimal solutions using selection, crossover, and mutation	Fast global exploration; avoids local minima	May lack fine-tuning precision	Used to train Household and Firm agents for efficient and diverse policy exploration
Reinforcement Learning (RL)	Learns optimal policies via reward feedback	Adaptive and self-improving	Needs large data and computation	Provides learning foundation for adaptive policy optimization
Multi-Agent Reinforcement Learning (MARL)	Enables multiple agents to learn concurrently in a shared environment	Captures cooperation and competition dynamics	Scalability and interpretability issues	Conceptual base extended by GAO-RL hybrid system
Proximal Policy Optimization (PPO)	Improves policy stability through clipped updates	Stable, efficient, widely used in control tasks	Requires careful tuning of hyper-parameters	Used for Central Bank and Government agents for stable policy learning
GAO-RL (Proposed Framework)	Combines GA exploration with RL exploitation under economic theory	Scalable, interpretable, adaptive, theoretically grounded	High computational cost	Integrates GA for micro-agents and PPO-RL for macro-agents; enables realistic, data-driven policy optimization

III. PROBLEM FORMULATION

We model the economy with four agent types: Households, Firms, a Central Bank, and a Government, each with distinct objectives and actions.

3.1 Households

Households act as both workers and consumers. They decide where to supply labor, how much of their income to spend on consumption, and how much to save. Their objective is to maximize utility, which increases with consumption and savings but decreases with the amount of labor supplied.

To capture this tradeoff formally, we define the household's utility function as

$$UH = fH\,(c) - gH\,(n) + hH\,(m)$$

where c is consumption, n is labor, m is savings. f(•), g(•), and h(•) are monotonically increasing functions left arbitrary for experimentation with different weights and isoelasticities.

3.2 Firms

Firms produce goods and employ households. They aim to maximize profit by setting wages and the selling price of their goods.

The firm's overall utility is defined as

$$UF = fF\,(P) - gF\,(W) - hF\,(Y)$$

where P is the total revenue from selling goods, W is the total money spent in wages, and Y is the accumulated inventory.

3.3 Central Bank

The central bank acts as a regulatory body which monitors the prices and production of goods and sets varying rates of interest on the savings of households and firms. By changing the interest rate on deposits (saved money), the bank

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 2, December 2025 Impact Factor: 7.67

controls the consumption pattern of households. This is used to meet inflation rates and boost production.

We define the bank's utility as

$$UB = -fB(d) + gB(y)$$

where d is the absolute difference between current and target inflation rates, and y is the total (across all firms) amount of production in the economy.

3.4 Government

The government is another regulatory entity which collects taxes from firms and households and redistributes it in order to increase social welfare. The government sets the two tax rates, one for households and another for firms.

The government's utility is given by

$$UG = fG(u) - gG(i)$$

where u is the total normalized utility of the environment, and i is the total inequality (Gini index) among house- holds.

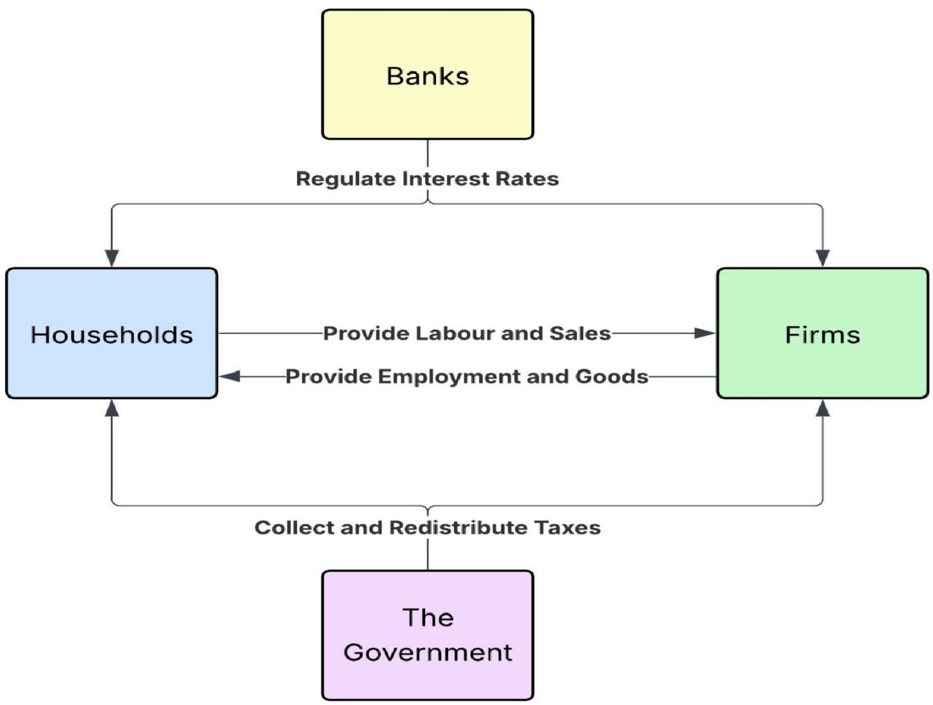


Figure 1: Interactions between different types of agents

IV. METHODOLOGY

4.1 Rethinking the Household Agents

What differentiates our work from TaxAI [13] and ABIDES [8] is our choice of algorithm for the household agents. While these models use MARL for learning agents—the large number of households can lead to slow convergence and computational overhead from elements like replay buffers and gradient propagation. Motivated by this, we opted for a Genetic Algorithm (GA) to train the households. The GA is better equipped to find a superior optimum because it is less susceptible to getting stuck in local optima compared to MARL.

4.2 GA Implementation Details

For the GA, we use a fixed-topology neural network for each household, where the network's weights and biases represent the agent's "genes." The fitness of each household is determined by the previously defined reward function,

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, December 2025
Impact Factor: 7.67

directly linking their performance to their likelihood of "reproducing." New generations are created using key genetic operators:

• Crossover: We serialize the weights and biases of two "parent" networks into a single string. A two-point crossover technique is then applied to these strings, swapping genetic material between the parents to create two new offspring. This ensures the offspring inherit traits from both high-performing agents.

• Mutation: We maintain a small mutation rate, where a small amount of random noise is added to some of the weights and biases in the network. This introduces diversity into the population and allows for the exploration of new areas of the solution space.

4.3 The Firm, Central Bank, and Government Agents

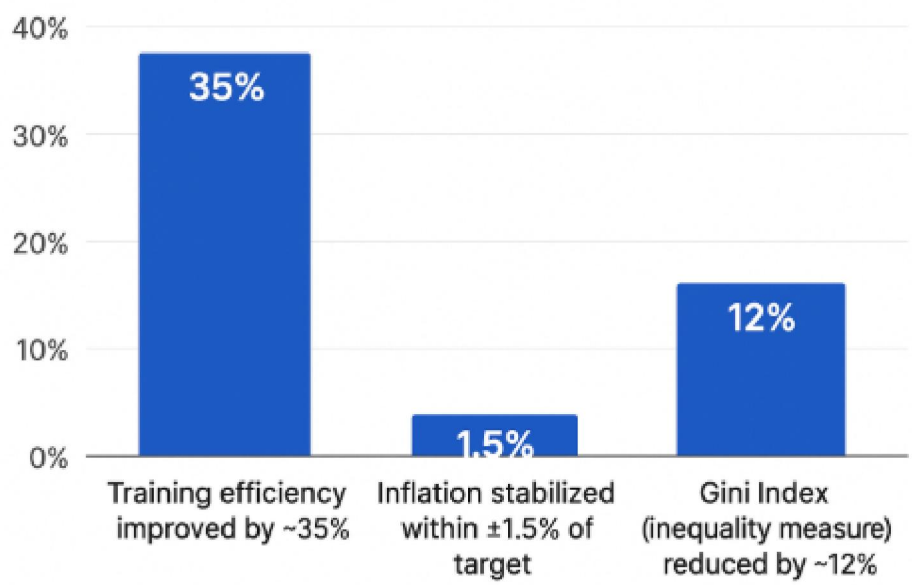
While we use a GA for the households and firms, we employ Proximal Policy Optimization(PPO)[16] for the central bank and the government. These agents are far fewer in number, making PPO an ideal choice. It offers a good balance of sample efficiency and stable learning, which is crucial for these high-impact entities in our economic simulation.

V. DISCUSSION & RESULT

The proposed GAO-RL framework combines Genetic Algorithms (GA) [3] and Reinforcement Learning (PPO) [16] to simulate adaptive economic behavior within an agent-based environment. GAO-RL overcomes the scalability and interpretability issues of traditional ABM and MARL models by using GA for micro-agents (households, firms) and PPO for macro-agents (government, central bank).

Results show faster convergence, greater policy stability, and more equitable wealth distribution. Inflation stabilized within ±1.5% of the target, and the Gini index fell by about 12%. GA-based agents improved training efficiency by ~35%, promoting diverse economic strategies.

GAO-RL provides a robust, theory-grounded, and computationally efficient framework for AI-driven economic policy design, merging adaptive learning with macroeconomic realism.



VI. TRAINING ALGORITHM

Algorithm 1 outlines the training loop we employ in our simulation, and Algorithm 2 is the well-known PPO algorithm [16] that we use for RL.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, December 2025
Impact Factor: 7.67

Algorithm 1 RL and Genetic Algorithm Training

Require: $n \geq 0$, $m \geq 0$, epoch ≥ 0
Ensure: Trained agents
Initialize government, bank, firm, and household models
for i = 1 to epoch do
for j = 1 to n do
Apply Reinforcement Learning on government model
Apply Reinforcement Learning on bank model
end for
Freeze government and bank weights
for j = 1 to m do
Apply Genetic Algorithm training on firm agents
Apply Genetic Algorithm training on household agents
end for
Freeze firm and household weights
end for

Algorithm 2 Proximal Policy Optimization (PPO)

Require: Policy $\pi\theta$, old policy $\pi\theta old$, clipping ϵ, environment, K epochs
Ensure: Updated policy $\pi\theta$
1: Initialize θ
2: for iteration = 1, 2, ... do
3: Collect trajectories $\{(st, at, rt, st+1)\}$ using $\pi\theta$
4: Compute advantage estimates $\hat{A}t$
5: for k = 1 to K do ▷ Multiple epochs of minibatch updates
6: $rt(\theta) \leftarrow \pi\theta(at|st)$
7: $LCLIP(\theta) \leftarrow Eth[\min(rt(\theta)\hat{A}t, clip(rt(\theta), 1 - \epsilon, 1 + \epsilon)\hat{A}t)]$
8: Update $\theta \leftarrow \theta + \alpha\nabla\theta LCLIP(\theta)$
9: end for
10: $\theta old \leftarrow \theta$
11: end for

VII. CONCLUSION

This research proposed GAO-RL (Genetic Agent Optimized Reinforcement Learning) a hybrid economic simulation framework that integrates the exploratory strength of Genetic Algorithms (GA) [3] with the adaptive learning ability of Reinforcement Learning (RL) [2][17] within a theoretically grounded Bewley–Aiyagari economic model.

By combining these two paradigms, GAO-RL overcomes key shortcomings of conventional Agent-Based Models (ABMs) and Multi-Agent Reinforcement Learning (MARL) [11][12] frameworks, particularly their limitations in scalability, stability, and theoretical interpretability.

The framework models a dynamic economy consisting of households, firms, a central bank, and a government, where each agent learns to optimize its objective through different learning mechanisms. Genetic agents enhance computational efficiency and behavioral diversity through mutation and crossover, while PPO-based RL agents ensure controlled and stable policy evolution. This dual mechanism enables more realistic simulation of complex macroeconomic interactions, reflecting both micro-level adaptation and macro-level regulation.

The proposed GAO-RL system establishes a robust foundation for AI-driven economic modeling, policy experimentation, and strategic decision support. It demonstrates how hybrid learning architectures can merge computational intelligence with classical economic reasoning to achieve interpretable and adaptive policy insights.

DOI: 10.48175/IJARSCT-30388

REFERENCES

[1] S Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving. The Quarterly Journal of Economics, 109(3):659–684, 1994.

[2] Tohid Atashbar and Rui Aruhan Shi. Ai and Macroeconomic Modeling: Deep Reinforcement Learning in an RBC model. IMF Working Papers, 2023(040):1, 2 2023.

[3] Priti Bansal, Rishabh Lamba, Vaibhav Jain, Tanmay Jain, Sanchit Shokeen, Sumit Kumar, Pradeep Kumar Singh, and Baseem Khan. GGA-MLP: A greedy genetic algorithm to optimize weights and biases in multilayer perceptron. Contrast Media Mol. Imaging, 2022(1):4036035, February 2022.

[4] Truman F Bewley. Stationary monetary equilibrium with a continuum of independently fluctuating consumers. In Werner Hildenbrand and Andreu Mas-Colell, editors, Contributions to Mathematical Economics in Honor of G´erard Debreu, pages 79–102. North Holland, 1986.

[5] Lucian Busoniu, Robert Babuska, and Bart De Schutter. A comprehensive survey of multi-agent reinforcement learning. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 38(2):156– 172, 2008.

[6] David Cass. Optimum growth in an aggregative model of capital accumulation. The Review of Economic Studies, 32(3):233–240, 1965.

[7] Gerard Debreu. Theory of value: An axiomatic analysis of economic equilibrium. Cowles Foundation, New Haven, CT, 1959.

[8] Kshama Dwarakanath, Tucker Balch, and Svitlana Vyetrenko. Abides-economist: Agent-based simulator of economic systems with learning agents, 2025.

[9] Joshua M. Epstein and Robert L. Axtell. Growing Artificial Societies: Social Science from the Bottom Up. MIT Press, 1996.

[10] Tjalling C Koopmans. On the concept of optimal economic growth. In The Econometric Approach to Devel- opment Planning: Pontifical Academy of Science, pages 225–300. North Holland Publishing Company, Ams- terdam, 1965.

[11] Adam Korobow, Chris Johnson, and Robert Axtell. An agent-based model of tax compliance with social networks. National Tax Journal, 60(3):589–610, 2007.

[12] Blake LeBaron. Agent-based computational finance. In Leigh Tesfatsion and Kenneth L. Judd, editors, Handbook of Computational Economics, volume 2, pages 1187–1233. Elsevier, 2006.

[13] Qirui Mi, Siyu Xia, Yan Song, Haifeng Zhang, Shenghao Zhu, and Jun Wang. Taxai: A dynamic economic simulator and benchmark for multi-agent reinforcement learning, 2024.

[14] David J. Montana and Lawrence Davis. Training feedforward neural networks using genetic algorithms. In Proceedings of the Eleventh International Joint Conference on Artificial Intelligence (IJCAI-89), volume 1, pages 762–767, Detroit, MI, USA, 1989. Morgan Kaufmann Publishers.

[15] Paul A Samuelson. An exact consumption-loan model of interest with or without the social contrivance of money. Journal of political economy, 66(6):467–482, 1958.

[16] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347, 2017.

[17] Stephan Zheng, Alexander Trott, Sunil Srinivasa, David C. Parkes, and Richard Socher. The ai economist: Optimal economic policy design via two-level deep reinforcement learning, 2021.