Policy Optimization in Stock Market: An Agent-Based Model Approach Integrating Evolutionary Computation

DU YUSEN
Laboratory of Evolutionary Computation, Master's Program in Computer Science
CS Seminar, University of Tsukuba, Tsukuba, Japan
s2620706@u.tsukuba.ac.jp

Abstract

Macroeconomic regulations frequently lead to unintended consequences due to the irrationality and heterogeneous interactions of market participants. An example is the 2016 Chinese stock market circuit breaker mechanism, where a policy designed to curb panic instead accelerated liquidity depletion and market crash. Traditional macroeconomic frameworks, such as Computable General Equilibrium (CGE) and Dynamic Stochastic General Equilibrium (DSGE) models, based on representative agents and rational expectation assumptions, causing them hard to capture micro-level irrational. In this paper, we propose a quantitative policy optimization framework that integrates an Agent-Based Model (ABM) of the stock market and plan to apply the Genetic Algorithm (GA) to it. Built upon a Central Limit Order Book (CLOB), our ABM contains Institutional, Normal Retail, and Noise Retail traders whose price beliefs originate from fundamental valuation, reputation driven Softmax social contagion, and trend extrapolation amplified by regulatory magnet effects. By injecting controlled shocks via event slot, the GA evolves a multi dimensional regulatory policy vector comprising price limits, circuit breaker thresholds, leverage caps, and settlement rules, to maximize market stability and liquidity while minimizing systemic bankruptcy. After we built the framework of simulation, We validated the effectiveness of the simulation using the heavy-tailed effect and volatility clustering.

1 Introduction

Financial markets are complex systems that market indicators emerge from interactions of diverse participants. When central banks or government introduce interventions such as interest rate adjustments, reserve requirement shifts, or trading curb rules, market reactions often deviate sharply from classical theoretical predictions. For example, from September 18, 2024, when the Federal Reserve announced a 50 basis point federal funds rate cut, to November 7, when the US announced another rate cut, the Dow Jones Industrial Average rose 4%. And in addition, experience has shown that the Chinese stock market frequently react counterintuitively to monetary policy announcements compared to developed markets. [1].

Such policy failures are even more pronounced in micro structural trading regulations. To mitigate severe turbulence following the 2015 crash, the Shanghai and Shenzhen Stock Exchanges implemented a circuit breaker rule on the first trading day of 2016. Contrary to stabilizing the market, the policy triggered a severe "magnet effect" [2], as the index approached the halt threshold, retail investors panicked over impending illiquidity and accelerated their sell offs, halting the market twice within four days and forcing regulators to abandon the policy [3].

These unexpected reactions happen not because the market is purely random, but because classical models ignore individual irrationality. Individual rationality does not imply overall rationality, and no formal justification supports the assumption that macro level market behavior mirrors a single maximizing representative individual [4]. To prevent regulatory backfires under individual irrationality, policymakers require a simulation system of evaluating how heterogeneous traders adapt to regulatory boundaries before policies are deployed in reality.

In this research, we develop an Agent-Based Model (ABM) stock market simulator integrated with Evolutionary Computation. Rather than reproducing historical market trajectories, our framework works backward from desired macroeconomic stability metrics to automatically search for optimal quantitative regulatory policies under extreme market shocks using a Genetic Algorithm (GA).

2 Preliminaries

2.1 Macroeconomic Policy Models: CGE, DSGE, and ABM

In traditional policy analysis, Computable General Equilibrium (CGE) models work top-down on macroeconomic that ignore micro-level behavioral heterogeneity. Conversely, Dynamic Stochastic General Equilibrium (DSGE) models adopt a micro-founded bottom-up perspective but are based on the *homo economicus* and rational expectations paradigm [4]. Consequently, both frameworks fail to account for endogenous panics and herding during financial crises.

The ABM overcomes these limitations by modeling limited rational, autonomous agents interacting directly without imposing a priori equilibrium constraints [4]. Pioneering work by Dosi et al. [5] established the "Schumpeter meeting Keynes" paradigm, demonstrating that ABMs can serve as rigorous policy laboratories for evaluating fiscal and monetary rules against business cycles.

2.2 Heterogeneous Agent Models and Market Microstructure

This subsection introduces the theoretical tools we adopt to build our simulator and the limitations of recent studies that we aim to overcome. First, for model validation, Cont [6] defined the standard statistical properties of real asset returns, such as heavy tails and volatility clustering, and Lux and Marchesi [7] proved in *Nature* that these properties emerge from interactions between fundamentalists and chartists. We use these properties as our validation benchmark. Second, for agent design, we adopt the discrete choice framework of Brock and Hommes [8, 9] and Chiarella and He [10], where traders select strategies based on past profits, as the mathematical tool for our Softmax social network. We also build our three-factor valuation model upon the value, trend, and noise trader

structures of Farmer and Joshi [11] and Hommes [12].

Recently, Albers et al. [13] modeled belief-driven role emergence under market shocks, while Evans and Ganesh [14] calibrated boundedly rational market behaviors using reinforcement learning. However, their limitation is that they only focus on forward simulation to reproduce market phenomena, without optimizing regulatory policies.

2.3 Regulatory Assessment and Evolutionary Optimization

Regarding trading regulations, Ackert et al. [15] and Hu [16] analyzed the behavioral and statistical properties of circuit breakers and price limits. Yeh and Yang [17] examined price limits in an artificial stock market; however, existing studies based on manual comparison of a few static policy presets. Systematically optimizing a multi dimensional regulatory policy vector via evolutionary computation remains an unexplored frontier that our work directly addresses.

3 Multi-Agent Stock Market Model

We construct a multi-sector, multi-stock ABM simulation platform featuring a continuous Central Limit Order Book (CLOB) and an out-of-market social savings capital pool, as illustrated in Fig. 1.

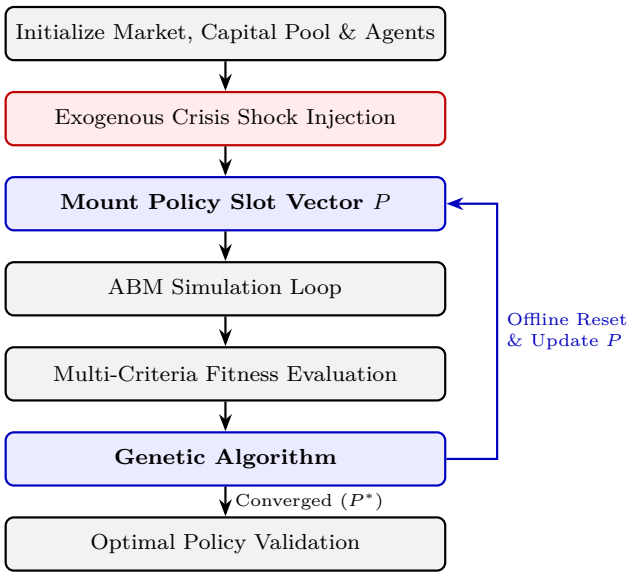


Figure 1: Offline workflow of the ABM stock market simulator and GA policy optimizer.

3.1 Market Environment, CLOB, and Capital Pool

The simulated market comprises multiple industrial sectors, each containing stocks characterized by intrinsic fundamental attributes: Total Assets, Total Liabilities, Shareholders' Equity (SE), Earnings (E), Number of Shares (NOS), Book Value Per Share (BPS = SE/NOS), and Earnings Per Share (EPS), alongside real time market attributes (Price P_t, Volume, Liquidity).

Order matching is governed by a continuous CLOB enforcing strict Price-Time Priority: incoming bid orders with price $\geq$ lowest ask execute immediately against the earliest ask order; otherwise, they enter the order book as limit orders.

To model macro liquidity inflows and outflows, total system capital is partitioned into a Market Active Capital Pool (1/3 initially) and an Out-of-Market Social Savings Reserve Pool (2/3 initially). When traders go bankrupt, their remaining assets are liquidated and returned to the reserve pool. Conversely, during market booms, each 1% sustained increase in the market index raises the probability of generating new retail agents from the reserve pool by 5% above a 1% baseline.

Table 1: Heterogeneous Agent Composition and Attributes

Attribute	Institutional	Normal Retail	Noise Retail
Population Ratio	5%	57%	38%
Capital Ratio	70%	18%	12%
Invest / Savings Split	80%/20%	20%/80%	20%/80%
Stock Holding Limit	20–100	5–50	5–50
Trading Interval	3–5 Days	1–3 Days	1–3 Days
Base Fund. Weight $w_{\text{fund}}^{\text{base}}$	80%	50%	10%
Bankruptcy Threshold	10% Init.	20% Init.	20% Init.

3.2 Heterogeneous Trader Composition

Following empirical market demographics [13, 11], agents are divided into three distinct classes: Institutional Traders, Normal Retail Traders, and Noise Retail Traders. Their structural and behavioral parameters are summarized in Table 1. Each agent i maintains an active cash balance, a reserve savings account, a portfolio holdings vector Θ_i, and a price belief vector ϕ_i.

3.3 Three Factor Valuation and Social Network

At each trading step t, agent i forms a price expectation $\phi_i(m,t)$ for stock m via a dynamic weighted combination of fundamental analysis, social contagion, and price trend extrapolation [12]:

$$\begin{aligned}\phi_i(m,t) = w'_{\text{fund}} V_{\text{fund}}(i,m,t) + w'_{\text{social}} V_{\text{social}}(i,m,t) \\ + w'_{\text{trend}} V_{\text{trend}}(i,m,t) + \epsilon\end{aligned} \tag{1}$$

where $w'_{\text{fund}} + w'_{\text{social}} + w'_{\text{trend}} = 1$, and ϵ denotes zero-mean Gaussian noise.

1) Fundamental Valuation (V_{fund}): Anchored by sector average Price-to-Book (PB_S) and Price-to-Earnings (PE_S) ratios, the intrinsic value of stock m is evaluated as:

$$V_{\text{fund}}(m,t) = \begin{cases} \dfrac{(\text{BPS}_m \cdot \text{PB}_S) + (\text{EPS}_m \cdot \text{PE}_S)}{2}, & E_m > 0 \\ \text{BPS}_m \cdot \text{PB}_S, & E_m \leq 0 \end{cases} \tag{2}$$

2) Reputation Driven Social Contagion (V_{social}): To prevent portfolio homogeneity and capture realistic herding, we adapt the evolutionary discrete choice model of Brock and Hommes [8, 9] and Chiarella and He [10] to a directed social network. Each agent j accumulates a decayed historical performance reputation score $U_j(t)$ based on realized return rates. Agent i allocates attention weights across its neighbor set $\mathcal{N}_i$ via a temperature scaled Softmax function:

$$W_{ij}(t) = \frac{\exp(U_j(t)/\tau_s)}{\sum_{k\in\mathcal{N}_i} \exp(U_k(t)/\tau_s)} \tag{3}$$

$$V_{\text{social}}(i,m,t) = \sum_{j\in\mathcal{N}_i} W_{ij}(t)\phi_j(m,t) \tag{4}$$

where τ_s is the social temperature governing choice intensity.

3) Trend Extrapolation (V_{trend}): Capturing momentum chasing and panic selling [11], agents extrapolate the short term return $R_t(\tau) = (P_t - P_{t-\tau})/P_{t-\tau}$ over look-back window τ using an individual extrapolation factor $\lambda_i \sim \mathcal{N}(0.5, 0.2^2)$:

$$V_{\text{trend}}(i,m,t) = P_t \cdot (1 + \lambda_i \cdot R_t(\tau)) \tag{5}$$

4) Regulatory Magnet Effect on Weights: Following Subrahmanyam [2] and Wong et al. [3], when market

volatility $\sigma_t = \sqrt{\frac{1}{T}\sum_{k=1}^{T}(R_{t-k} - \bar{R})^2}$ approaches the circuit breaker threshold L_{breaker}, or when stock price P_t nears the price limit boundary L_{price}, irrational herding and trend chasing intensify relative to fundamental belief $w_{\text{fund}}^{\text{base}}$:

$$\tilde{w}_{\text{social}} = w_{\text{social}}^{\text{base}} \cdot \left(1 + \beta_i \cdot \frac{\sigma_t}{L_{\text{breaker}}}\right) \tag{6}$$

$$\tilde{w}_{\text{trend}} = w_{\text{trend}}^{\text{base}} \cdot \left(1 + \beta_i \cdot \frac{|R_t(\tau)|}{|L_{\text{price}} - P_t| + \delta}\right) \tag{7}$$

where β_i is agent i's social sensitivity and $\delta > 0$ prevents division by zero. Normalizing $(w_{\text{fund}}^{\text{base}}, \tilde{w}_{\text{social}}, \tilde{w}_{\text{trend}})$ yields the active weights $(w'_{\text{fund}}, w'_{\text{social}}, w'_{\text{trend}})$ in Eq. (1).

5) Trading Actions: Based on $\phi_i(m,t)$ and current price P_t, each active agent chooses from three actions: Buy, Sell, or Hold. If $\phi_i(m,t) > P_t$ and the agent has available cash or leverage capacity (L_{leverage}), it submits a limit buy order within the price limit L_{price}. If $\phi_i(m,t) < P_t$ and the agent holds shares unlocked under the $T + L_{\text{settle}}$ rule, it submits a limit sell order. Otherwise, or when the market halts under L_{breaker}, the agent holds.

3.4 Shock Injection and Policy Slot Mechanism

Following the macroeconomic ABM policy-laboratory approach [5], our simulator implements a modular `ShockManager` and `PolicySlot` architecture. Evaluating crisis regulations in a calm, mean-reverting market is ineffective because regulatory boundaries are rarely triggered. Therefore, after the market reaches a steady state, `ShockManager` injects exogenous fundamental or liquidity shocks to push the market into extreme contraction or overheating regimes.

Simultaneously, the `PolicySlot` dynamically mounts a regulatory policy vector selected from the broader macroeconomic policy matrix C_{policy}. Focusing on direct triggers of market volatility, we parameterize the four-dimensional regulatory vector:

$$P = [L_{\text{price}},\ L_{\text{breaker}},\ L_{\text{leverage}},\ L_{\text{settle}}] \tag{8}$$

representing daily stock price limit ratio $L_{\text{price}} \in (0, 1]$, index circuit breaker threshold $L_{\text{breaker}} \in (0, 1]$, maximum margin financing leverage $L_{\text{leverage}} \geq 1.0$, and settlement lock-in period $L_{\text{settle}} \in \{0, 1, 2, \dots\}$ ($T + n$ trading rule).

4 Validation and Preliminary Results

To validate the simulator before future policy GA, we implement an `StylizedFactsService` based on Cont's empirical benchmarks [6] with three market properties, and run simulation in closed baseline market scenario for 3,000 days:

1. **Mean Reversion:** In closed baseline runs without exogenous shocks, the market composite price index starts at 3,000 and converges to an equilibrium level of approximately 2,686 around $t \approx 100$ (Fig. 2). Following Day 100, the index fluctuates stably around this equilibrium mean, demonstrating clear convergence and mean reverting dynamics.

2. **Heavy Tails:** Over 3,000 simulated trading days, excess kurtosis reaches 24.3565, with skewness of -2.9197 and a Jarque-Bera statistic of 78,390.67 ($p < 0.0001$). These results indicate strongly non-Gaussian, negatively skewed returns with pronounced heavy tails (Fig. 3) [7].

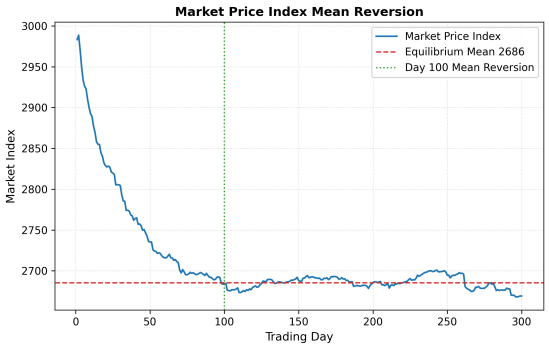


Figure 2: Market Price Index Mean Reversion

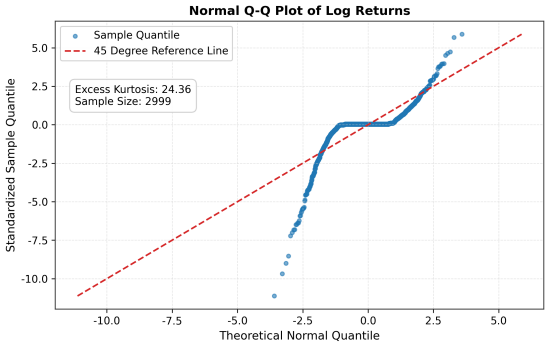


Figure 3: Normal Q-Q Plot of Log Returns

3. **Volatility Clustering:** The absolute return ACF at lag 1 is 0.4178, while the five lag ARCH-LM statistic is 393.27 ($p < 0.0001$). These results provide strong evidence of volatility clustering (Fig. 4).

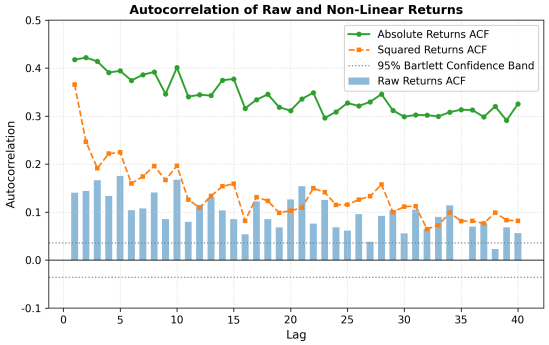


Figure 4: Autocorrelation of Raw and Non-Linear Returns

5 Next Step: GA

5.1 Chromosome and Genetic

Each candidate policy is encoded as a mixed variable chromosome $P = [L_{\text{price}}, L_{\text{breaker}}, L_{\text{leverage}}, L_{\text{settle}}]$, where the first three genes are continuous real variables and L_{settle} is a discrete integer gene. Across generations, the population evolves via:

- **Selection:** Tournament Selection, preserving selection pressure while avoiding premature convergence under noisy ABM evaluations.
- **Crossover:** Simulated Binary Crossover (SBX) applied to continuous regulatory thresholds, with integer rounding enforced on L_{settle}.
- **Mutation:** Gaussian Mutation within feasible regulatory bounds, enforcing the structural constraint $L_{\text{breaker}} \leq L_{\text{price}}$ so circuit breakers are not masked by narrower stock price limits.

5.2 Fitness Function

To capture the true regulatory trade-off between stability and market functionality, we formulate the fitness score $\mathcal{F}(P)$ over the post-shock evaluation window $[t_0, t_1]$ as:

$$\begin{aligned}\mathcal{F}(P) = \omega_1 S_{\text{growth}}(P) + \omega_2 S_{\text{liquidity}}(P) \\ - \omega_3 S_{\text{bankruptcy}}(P)\end{aligned} \tag{9}$$

1. **Stable Growth (S_{growth}):** Measured by the risk-adjusted index recovery ratio $\frac{\bar{R}_{\text{index}}}{\sigma_{\text{index}}+\delta}$ penalized by Maximum Drawdown (MDD), rewarding policies that arrest crashes without inducing secondary oscillations.

2. **Liquidity Preservation ($S_{\text{liquidity}}$):** Defined as the ratio of average daily trading volume during the post-shock regulation period to baseline pre-shock volume, preventing artificial market freezes.

3. **Bankruptcy Penalty ($S_{\text{bankruptcy}}$):** The cumulative proportion of agents whose active capital falls below their bankruptcy threshold, penalizing systemic wealth destruction.

5.3 Optimization Objective and Evaluation

The target scenario is a predefined crisis shock inspired by the January 2016 Chinese circuit breaker episode. The target is to identify policies that improve market stability under stress. We will compare the GA selected policy against a no intervention baseline, a fixed reference policy, and randomly sampled policies. All policies will be evaluated under the same shock settings and a common set of random seeds. The GA will be considered effective if it achieves higher mean fitness than the baselines across repeated runs.

6 Conclusion

This study presents an integrated ABM-GA framework to address macroeconomic and microstructural policy failures caused by boundedly rational investor behavior. By grounding heterogeneous agent valuation, Softmax reputation contagion, and regulatory magnet effects in established computational finance literature, our simulator provides a realistic testbed for crisis intervention. Coupling the completed simulation engine with evolutionary optimization will offer a quantitative decision support tool for designing resilient financial market regulations. As of CS Seminar, the construction and verification of the simulation system have been completed. Our next step will be to integrate the GA into our system and evaluate it against the historical crisis benchmarks.

References

[1] Jun Hu, George J. Jiang, and Guanzhong Pan. Market reactions to central bank interest rate changes: Evidence from the chinese stock market*. *Asia-Pacific Journal of Financial Studies*, 49(5):803–831, 2020.

[2] AVANIDHAR SUBRAHMANYAM. Circuit breakers and market volatility: A theoretical perspective. *The Journal of Finance*, 49(1):237–254, 1994.

[3] Kin Ming Wong, Xiao Wei Kong, and Min Li. The magnet effect of circuit breakers and its interactions with price limits. *Pacific-Basin Finance Journal*, 61:101325, 2020.

[4] Giorgio Fagiolo and Andrea Roventini. Macroeconomic policy in dsge and agent-based models. *Revue de l'OFCE*, N° 124(5):67–116, November 2012.

[5] Giovanni Dosi, Giorgio Fagiolo, and Andrea Roventini. Schumpeter meeting keynes: A policy-friendly model of endogenous growth and business cycles. *Journal of Economic Dynamics and Control*, 34(9):1748–1767, 2010. Computational perspectives in economics and finance: Methods,dynamic analysis and policy modeling.

[6] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236, 2001.

[7] Thomas Lux and Michele Marchesi. Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, 397(6719):498–500, Feb 1999.

[8] William A. Brock and Cars H. Hommes. A rational route to randomness. *Econometrica*, 65(5):1059–1095, 1997.

[9] William A. Brock and Cars H. Hommes. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 22(8):1235–1274, 1998.

[10] Carl Chiarella and Xue-Zhong He. Heterogeneous beliefs, risk and learning in a simple asset pricing model. *Computational Economics*, 19(1):95–132, Feb 2002.

[11] J.Doyne Farmer and Shareen Joshi. The price dynamics of common trading strategies. *Journal of Economic Behavior & Organization*, 49(2):149–171, 2002.

[12] Cars H. Hommes. Chapter 23 heterogeneous agent models in economics and finance. volume 2 of *Handbook of Computational Economics*, pages 1109–1186. Elsevier, 2006.

[13] Evan Albers, Mohammad T Irfan, and Matthew J. Botsch. Beliefs, shocks, and the emergence of roles in asset markets: An agent-based modeling approach. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '24, page 40–48, Richland, SC, 2024. International Foundation for Autonomous Agents and Multiagent Systems.

[14] Benjamin Patrick Evans and Sumitra Ganesh. Learning and calibrating heterogeneous bounded rational market behaviour with multi-agent reinforcement learning. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '24, page 534–543, Richland, SC, 2024. International Foundation for Autonomous Agents and Multiagent Systems.

[15] Lucy F. Ackert, Bryan Church, and Narayanan Jayaraman. An experimental study of circuit breakers: The effects of mandated market closures and temporary halts on market behavior. *Journal of Financial Markets*, 4(2):185–208, 2001.

[16] Yourou Hu. Statistical properties of price limit hits in the chinese stock markets. *Advances in Applied Mathematics*, 12(9):4112–4121, 2023.

[17] Chia-Hsuan Yeh and Chun-Yi Yang. Examining the effectiveness of price limits in an artificial stock market. *Journal of Economic Dynamics and Control*, 34(10):2089–2108, 2010.