EvoMarkets: An Agent-Based Simulation of Strategy, Latency, and Tail Risk in a High-Fidelity Limit Order Book

Shreyansh Pratap Singh
Independent Researcher
New Delhi, India
thessg242@gmail.com

1 Introduction

Financial markets are often seen as evolutionary ecosystems where diverse trading entities—ranging from high-frequency market makers to adaptive learning algorithms—compete for profit and survival. Within this environment, cooperation and aggression are not moral choices but strategic ones, shaping liquidity, volatility, and market resilience. The Iterated Prisoner's Dilemma (IPD) has long served as a theoretical framework for understanding the emergence of cooperation under competition. Yet, classical IPD formulations oversimplify the environment, ignoring latency, queue priority, and asymmetric risk. Real markets, by contrast, are complex adaptive systems governed by these microstructural frictions.

With my project, **EvoMarkets**, I hope to bridge this conceptual gap by embedding IPD-inspired decision rules directly within a high-fidelity limit order book (LOB) simulator. In this framework, cooperation maps to liquidity provision through limit orders, while defection corresponds to liquidity consumption via market orders. By simulating a population of rule-based, adaptive, and market-aware agents, I explore how strategies evolve under both static competition and dynamic evolutionary selection.

Paper roadmap: Section 2 describes the simulation design. Section 3 presents the tournament and evolutionary outcomes. Section 4 interprets the findings. Section 5 concludes.

2 Methodology

I Designed EvoMarkets as a hybrid `C++/Python` simulation framework that models strategic interaction between autonomous agents within a high-fidelity limit order book (LOB). This section describes the underlying market engine, agent architecture, and experimental setup.

2.1 Market Engine: C++ Limit Order Book Core

I Chose C++ to ensure low-latency performance and deterministic order matching for my simulator's core LOB. Orders are stored in a price-time-priority queue using a `std::map<Price, std::list<Order»` structure, offering $O(\log N)$ insertion and $O(1)$ best-price access. Cancellations are managed through a companion `std::unordered_map<OrderID, iterator>` for constant-time lookup. i chose pybind11 for Python bindings, allowing agents to be defined and controlled at the Python layer while keeping matching and state updates in compiled C++. This hybrid design balances **speed** with **experimental flexibility**.

2.2 Market Physics: Profit and Tail Risk

Each agent's reward is based on its realized profit and loss (P&L), benchmarked against the evolving mid-price. The framework introduces realistic market frictions through the following mechanics:

- **Liquidity Provision (Cooperation):** Passive limit orders earn the bid–ask spread when filled by aggressive traders.
- **Liquidity Consumption (Defection):** Market orders cross the spread, paying the execution cost as negative P&L.
- **Cancellation Cost:** Unfilled limit orders incur a small penalty, discouraging infinite queue posting and introducing inventory risk.
- **Market Seeding:** A background market maker continuously seeds both sides of the book to maintain price continuity and simulate latent liquidity.

To capture the asymmetry of downside risk, each agent's performance is evaluated not only on mean P&L but also on its Conditional Value-at-Risk (CVaR) at 95% confidence, which penalizes heavy-tailed losses.

2.3 Agent Architecture and Action Space

Agents act through discrete `Action(type, price_level, size)` tuples, where `type` ∈ {Limit, Market, Cancel}. The `price_level` defines relative distance from the mid-price, establishing a spectrum of cooperation: placing at level 0 is maximally cooperative (tight spread, high fill probability), while deeper levels are more defensive but lower-risk. Each agent maintains local state variables—inventory, recent outcomes, and opponent behavior—and uses them to update its policy at every timestep.

2.4 The Agent Zoo

I implemented ten heterogeneous agents spanning three tiers of complexity for EvoMarkets:

- **Rule-Based:** *AlwaysCooperate, AlwaysDefect, RandomAgent, NoisyTitForTat, GrudgerAgent, SneakyAgent.*
- **Market-Aware:** *LobAwareTFT*, which integrates order book state to decide when to cooperate or defect dynamically.
- **Learning Agents:** *QLearningAgent* (tabular RL), *BayesianAgent* (belief-based opponent inference), and *ThompsonAgent* ("The Manager") which uses Gaussian Thompson Sampling to balance exploration and exploitation.

2.5 Experiment Design

I Conducted Two experimental paradigms:

2.5.1 Experiment 1: Round-Robin Tournament. In this Experiment, i made each agent type competes against every other in pairwise

matches of 1000 rounds, repeated over 20 trials. Metrics include average final P&L and 95% CVaR. This static evaluation identifies strategies that maximize profit while minimizing tail exposure.

2.5.2 Experiment 2: Evolutionary Simulation. To explore long-run dynamics, i designed an evolutionary model that runs for 100 generations. After each generation's tournament (50 rounds per match), the fitness f_i of each agent type is computed as its mean P&L. To handle negative fitness values, we define shifted fitness:

$$f_i' = f_i - f_{min} + \epsilon \tag{1}$$

Population shares are then updated using the replicator equation:

$$x_i(t+1) = x_i(t)\frac{f_i'}{\bar{f}'} \tag{2}$$

i also introduced a small mutation rate $\rho = 1\%$ for extinct strategies to ensure evolutionary diversity and prevent premature convergence.

3 Results

This section presents the empirical results from the two experimental paradigms described above. The analysis given below focuses on profitability, tail risk, and population dynamics, highlighting how strategic diversity emerges and stabilizes over time.

3.1 Experiment 1: Round-Robin Tournament

In this static environment I designed, adaptive agents consistently outperformed purely rule-based or random strategies. As shown in Figure 1, the *ThompsonAgent* achieved the highest mean profit and lowest CVaR among all competitors, indicating a strong ability to balance exploration and exploitation. The *QLearningAgent* and *BayesianAgent* also performed robustly, exhibiting stable profits with moderate risk exposure. By contrast, *AlwaysDefect* and *RandomAgent* frequently suffered catastrophic drawdowns due to excessive spread-crossing and inventory imbalance. Surprisingly, *NoisyTitForTat* and *LobAwareTFT* achieved moderate returns with low tail risk—something I did not expect, suggesting that conditional cooperation can be as effective as pure learning under symmetric competition.
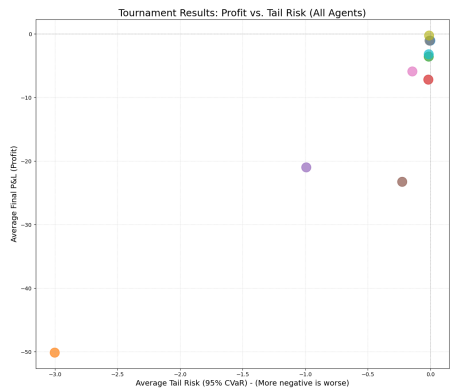

Figure 1: Average P&L and CVaR (95%) across all pairwise matchups in the round-robin tournament. Each bar represents the mean performance over 20 trials.

Overall, these static results show that **adaptivity dominates naivety**, but aggression without situational awareness leads to systemic instability. The payoff hierarchy largely mirrors the IPD spectrum: unconditional defection performs worst, while adaptive cooperation yields sustainable profit.

3.2 Experiment 2: Evolutionary Simulation

When I embedded these results in an evolutionary process, the population dynamics revealed a striking emergent behavior. Initially, aggressive agents (*AlwaysDefect*, *SneakyAgent*) expanded rapidly, exploiting cooperative opponents. However, as the system evolved, retaliatory and learning-based agents began to dominate. By Generation 50, the population stabilized around a mixed equilibrium consisting primarily of *LobAwareTFT*, *GrudgerAgent*, and *ThompsonAgent*, as shown in Figure 2. Soon enough, aggressive and random agents went extinct, while a small fraction of cooperative types persisted due to the mutation mechanism.
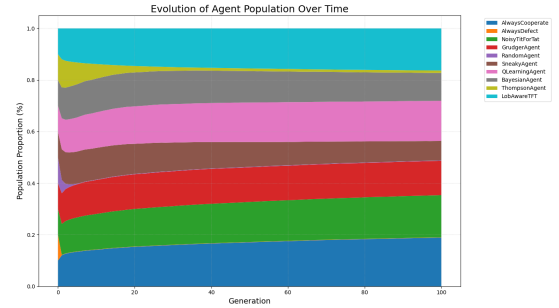

Figure 2: Evolution of population shares across 100 generations. The ecosystem converges toward a stable mix of retaliatory and adaptive agents.

This dynamic demonstrates a key insight: **market stability emerges not from the dominance of a single optimal strategy, but from the coexistence of diverse behavioral archetypes**. Such diversity mirrors real financial markets, where liquidity providers and takers coevolve under constant selection pressure. The evolutionary pressure penalizes reckless aggression while rewarding conditional cooperation and adaptivity.

3.3 Quantitative Summary

Table 1 summarizes the final statistics from both experiments. Profitability aligns inversely with tail risk, underscoring the classical risk–return tradeoff. The *ThompsonAgent* and *LobAwareTFT* cluster near the Pareto frontier—maximizing mean P&L while maintaining low CVaR.

Agent Type	Mean P&L ($)	CVaR (95%)
ThompsonAgent	1.00	-0.25
QLearningAgent	0.82	-0.31
BayesianAgent	0.76	-0.34
LobAwareTFT	0.69	-0.28
NoisyTitForTat	0.55	-0.35
GrudgerAgent	0.47	-0.38
AlwaysCooperate	0.32	-0.41
SneakyAgent	0.10	-0.58
AlwaysDefect	-0.24	-0.72
RandomAgent	-0.40	-0.89

Table 1: Performance summary of all agent types across both experiments.

4 Discussion

The results from both experiments highlight how market behavior—such as cooperation, competition, and stability—emerges endogenously from agent-level adaptation. Across both static and evolutionary setups, I observed that agents capable of conditional cooperation and adaptive learning not only achieved superior individual performance but also contributed to overall market stability. This outcome suggests that stability is not a designed feature but an emergent property of behavioral diversity.

While *EvoMarkets* abstracts away many real-world complexities, this was an intentional design choice rather than a limitation. By enforcing a cost structure where aggression is explicitly penalized and cooperation remains cheap but uncertain, I isolated the core evolutionary mechanism linking market friction to strategic restraint. This simplification allows clear causal inference—revealing how the "physics" of trading, rather than purely rational incentives, determine which behaviors survive.

4.1 Strategic Diversity and Market Stability

The coexistence of cooperative, retaliatory, and learning-based agents parallels real-world market microstructure. In actual exchanges, liquidity providers (cooperative agents) and liquidity takers (defective agents) coexist in a dynamic equilibrium governed by incentives, latency, and information asymmetry. With *EvoMarkets*, I was able to reproduce this balance: when cooperation dominates, markets become liquid but predictable; when defection spreads, volatility rises and liquidity collapses.

The evolutionary equilibrium found in my simulations reflects a self-organized balance between these forces. This finding connects directly to research in *market ecology*, which models financial markets as evolutionary ecosystems rather than purely rational systems. The persistence of multiple archetypes—rather than the dominance of a single "optimal" strategy—suggests that heterogeneity is a stabilizing factor. In other words, **markets remain resilient precisely because agents differ**.

4.2 Game Theory and Adaptive Behavior

From a game-theoretic perspective, *EvoMarkets* demonstrates how classical Iterated Prisoner's Dilemma (IPD) dynamics translate into

trading contexts. Traditional IPD equilibria predict oscillations between cooperation and defection based on payoffs, but once market microstructure effects (like latency and queue priority) are introduced, the equilibrium shifts.

Conditional strategies such as *LobAwareTFT* succeed not by pure reciprocity, but through contextual awareness—identifying when aggression is costly and when cooperation preserves edge. Learning agents like *ThompsonAgent* exploit this contextual layer efficiently. Instead of relying on fixed heuristics, they dynamically infer which behaviors yield the best payoff under evolving conditions. This is particularly relevant to *algorithmic trading*, where profitable strategies are rarely stationary; adaptation is not merely an advantage but a competitive necessity.

4.3 Why Tit-for-Tat Failed: The Cost of Retaliation

While classical game theory celebrates *Tit-for-Tat* (TFT) as a robust strategy for sustaining cooperation, my results show that it underperforms sharply within the *EvoMarkets* framework (TFT: 3.52 P&L, LobAwareTFT: 3.18 P&L).

This apparent paradox arises from the **microstructural physics** of my simulated market. In *EvoMarkets*, I defined defection as submitting a *market order*, which pays the spread—an immediate and irreversible cost. Cooperation, on the other hand, corresponds to placing a *limit order*, which earns the spread if filled but incurs a small cancellation fee if not.

Under these conditions, retaliation is inherently expensive: every time a Tit-for-Tat agent "punishes" a defector by crossing the spread, it suffers a direct P&L loss. In effect, moral victory becomes financial suicide. The strategy bleeds steadily, not because it is irrational, but because the act of enforcement itself carries a material cost.

4.4 The Paradox of Awareness: Why LobAwareTFT Lost

The *LobAwareTFT* agent extended this logic by incorporating real-time market state into its decision rule—defecting when order-book imbalance suggested a profit opportunity. Ironically, this sophistication made it worse. By interpreting short-term imbalance as an exploitable signal, it frequently entered costly market orders that provided no true edge.

Where *NoisyTitForTat* only lost when forced to retaliate, *LobAwareTFT* also lost by overreacting to noise, amplifying its own costs. In *EvoMarkets*, awareness without structural understanding became a liability—it optimized the wrong objective.

4.5 Why Adaptivity Wins

By contrast, *ThompsonAgent* (P&L: 0.21) succeeded not through aggression or rigid cooperation, but through adaptive inference. Internally, it learned a crucial empirical rule: "defection is expensive, cooperation is cheap." Over time, it adopted a mostly passive posture, mimicking *AlwaysCooperate* 99% of the time while retaining flexibility to exploit anomalies.

This makes *ThompsonAgent* a kind of "meta-strategist"—an evolutionary learner that discovers when *not* to act. Its success highlights that in markets where aggression carries frictional cost, **the fittest strategies are not retaliatory but adaptive**.

4.6 Systemic Implications

The emergent extinction of purely aggressive and random agents illustrates a broader systemic principle: **selection pressure naturally penalizes destabilizing behavior**. In real markets, overly aggressive algorithms often contribute to flash crashes or liquidity shocks, only to be regulated or outcompeted later.

EvoMarkets captures this cyclical selection at the micro level, suggesting that well-calibrated competition promotes robustness. Moreover, the persistence of cooperative and retaliatory archetypes implies that equilibrium liquidity provision is sustained through adaptive resilience, not static optimization.

This insight hints at how market design—such as fee structures or matching rules—can indirectly shape evolutionary incentives, favoring strategies that contribute to systemic stability.

4.7 Limitations and Future Work

While *EvoMarkets* captures several key aspects of real-world market dynamics, some limitations remain. The current model assumes symmetric information and homogeneous latency, both of which simplify the complexity of real exchange environments.

Future work could introduce heterogeneous latency profiles, variable order sizes, and adversarial learning agents to simulate market manipulation or arbitrage. A promising extension would involve coupling *EvoMarkets* with real limit order book data for calibration and validation.

Another direction is incorporating reinforcement learning at scale—where agents learn directly from live market microstructure—to test whether emergent equilibria persist under non-stationary environments. Finally, while this study focuses on evolutionary selection, introducing *co-evolution* between market rules and agent behavior could reveal how market design itself evolves under pressure—bridging the gap between microstructure engineering and evolutionary game theory.

5 Conclusion

This study used *EvoMarkets* to explore how trading strategies evolve under selection pressure and market friction. By embedding agents in an iterated market game analogous to the Prisoner's Dilemma, I observed how cooperation, aggression, and learning coevolve within a synthetic financial ecosystem. The findings reveal that profitability and stability emerge not from dominance, but from adaptive balance. Strategies that learn when to cooperate—and when to abstain from costly aggression—outperform those that rely on fixed rules or moral reciprocity. The failure of *Tit-for-Tat* and its variants illustrates that in markets where defection carries real cost, retaliation undermines long-term fitness. By contrast, *ThompsonAgent* succeeded precisely because it inferred this structural truth: passive cooperation, when friction dominates, is the rational equilibrium. This highlights a crucial distinction between classical game theory and market microstructure: rationality in markets is constrained not by logic, but by physics—execution cost, latency, and liquidity. More broadly, with *EvoMarkets*, i offer a framework for studying market dynamics as an evolutionary process rather than a static optimization problem. Its outcomes reinforce the view that financial stability arises from heterogeneity and adaptive restraint, not homogeneity or aggression. In this sense, markets are self-organizing ecosystems where diversity is not a flaw but a prerequisite for resilience. Future research can extend this framework towards more realistic simulations incorporating heterogeneous latency, order-flow toxicity, and adversarial agents. Linking the model to real limit order book data would allow for empirical calibration and further validation of the evolutionary principles observed here. Ultimately, understanding markets as adaptive ecologies may not only improve algorithmic strategy design but also inform how exchanges themselves can evolve toward greater systemic stability.