# A Comparative Study of an LLM-Based Agent in Synthetic Markets with Memory

**Éder Johnson de Área Leão Pereira**♣ and **Alan Moura Freitas**♣

♣Federal Institute of Education, Science and Technology of Maranhão (IFMA), Brazil

Submitted: October 6, 2025 • Accepted: December 31, 2025

---

ABSTRACT: This paper presents a comparative analysis of four trading strategies—termed chartist, fundamentalist, noise, and a Large Language Model (LLM)-based agent—within a synthetic market environment characterized by different memory regimes, modeled by the Hurst exponent ($H$). Price series were generated exogenously using Fractional Gaussian Noise. The LLM agent received textual descriptions of price trends and responded with trading actions. Results indicate that the LLM agent achieved superior performance compared to other heuristic strategies across all regimes, particularly in persistent markets ($H = 0.7$). This study is not intended as an agent-based model (ABM) with endogenous interactions, but rather as a comparative backtest of strategies on exogenous price series. The objective is to explore the potential of LLMs as tools for modeling financial decision-making, serving as a baseline for future research on bounded rationality and adaptive behavior.

---

**JEL classification:** C63, G17, G40, C45, G12, E44.

**Keywords:** LLM Agents, Financial Modeling, Synthetic Markets, Bounded Rationality, Hurst Exponent, Simulation.

DOI: 10.33119/ERFIN.2025.10.2.3

# 1 Introduction

The intersection of artificial intelligence and financial markets represents a promising frontier for computational economics. Traditionally, financial models have utilized heuristic agents (chartists, fundamentalists) to simulate market dynamics (Lux and Marchesi, 2000; Brock and Hommes, 1998). With the advent of Large Language Models (LLMs), new opportunities emerge to explore innovative approaches to modeling financial decision-making.

The concept of bounded rationality, introduced by Simon (1955), posits that economic agents operate under cognitive limitations, employing heuristics rather than pursuing perfect optimization. LLMs, with their capacity to process natural language and recognize patterns, offer a novel approach to modeling this rationality more flexibly than fixed-rule-based agents (Horton, 2023; Cong et al., 2023).

It is crucial to clarify the scope of this work. This study does not implement an Agent-Based Model (ABM) in the strict sense, as the agents do not interact with each other or influence price formation. Price series are generated exogenously. Therefore, our analysis more closely resembles a comparative backtesting of four distinct strategies in a controlled laboratory environment. The primary objective is to evaluate the behavior of an LLM-based agent in response to different market conditions (memory), compared to simple heuristics.

This study contributes to the literature by conducting a comparative analysis of an LLM-based agent against traditional strategies across different market memory regimes, defined by the Hurst exponent ($H$). Our research addresses two main questions: (1) How does the performance of an LLM agent compare to that of traditional heuristics in synthetic markets? (2) What is the impact of market memory properties on the relative performance of these strategies?

## 1.1 Novelty and contribution to existing literature

While agent-based models (ABMs) in finance have been extensively studied since the seminal works of Lux and Marchesi (1999, 2000) and Brock and Hommes (1998), the use of LLMs as trading agents remains relatively unexplored in the context of systematic comparative analysis. Recent studies have begun to investigate LLMs as economic agents (Horton, 2023) and their potential in financial forecasting (Lopez-Lira and Tang, 2023; Cong et al., 2023), but these works typically focus on prediction accuracy or market sentiment analysis rather than on comparative strategy performance across different market memory regimes.

Our contribution is threefold. First, we provide a transparent and reproducible methodology for implementing an LLM-based trading agent, addressing a significant gap in the literature where LLM implementations in finance often lack sufficient technical detail for replication. Second, we systematically compare the LLM agent's performance against well-established heuristic strategies (chartist, fundamentalist, noise) in a controlled environment, isolating the effect of market memory (Hurst exponent) on relative performance. Third, we explicitly frame this work as a comparative backtesting study rather than a full ABM, which clarifies the scope and limitations of the analysis and distinguishes it from previous ABM literature that emphasizes endogenous price formation and agent interactions.

Previous research on trading strategies in synthetic markets has primarily focused on (1) simple rule-based heuristics (Lux and Marchesi, 2000; Brock and Hommes, 1998), (2) genetic algorithms and machine learning approaches (LeBaron, 2006), or (3) behavioral finance models (Kahneman and Tversky, 1979). To our knowledge, no prior study has systematically compared an LLM-based agent against traditional heuristics using fractional Gaussian noise with controlled Hurst exponents as the price-generating process. This methodological combination allows us to isolate the decision-making capability of the LLM from the underlying market structure, providing clearer insights into how LLMs respond to different market memory properties.

Furthermore, the research questions we address—specifically, how LLM performance varies with market memory and how it compares to simple heuristics—have not been previously analyzed in the literature. While Horton (2023) demonstrates that LLMs can be used as economic agents in various scenarios, and Lopez-Lira and Tang (2023) show that ChatGPT can forecast stock movements, neither study examines performance across different market memory regimes or provides a systematic comparison with traditional trading heuristics in a controlled synthetic environment. Our study fills this gap by providing the first systematic empirical comparison of an LLM-based trading agent with traditional heuristics across different market memory structures.

## 2 Methodology

Our experimental framework was designed to provide a controlled environment for comparing the performance of different agent types. The simplicity of the design is intentional, aiming to isolate the decision-making behavior of the LLM agent in response to textual price signals.

### 2.1 Price series generation

Synthetic price series were generated using Fractional Gaussian Noise (FGN), which allows precise control over the Hurst exponent (Davies and Harte, 1987). This section provides a detailed mathematical description of the price generation methodology.

### 2.1.1 The Hurst exponent and market memory

The Hurst exponent ($H$) is a measure of long-range dependence in time series, ranging from 0 to 1. It characterizes the memory properties of a stochastic process and is formally defined through the rescaled range ($R/S$) analysis. The Hurst exponent is computed as

$$H = \lim_{n \to \infty} \frac{\log\left(\dfrac{R_n}{S_n}\right)}{\log(n)} \tag{1}$$

where $R_n$ is the range of cumulative deviations from the mean over n observations, and $S_n$ is the standard deviation over the same period. The Hurst exponent can be interpreted as follows: when $H = 0.5$, the process exhibits no memory (random walk); increments are independent and identically distributed (i.i.d.), following a Brownian motion. When $H > 0.5$, the process is persistent (trending); positive past increments tend to be followed by positive future increments, and negative by negative increments, creating a momentum effect. Conversely, when $H < 0.5$, the process is anti-persistent (mean-reverting); positive increments tend to be followed by negative increments and vice versa, creating a contrarian effect.

In this study, we employ three values of $H$ to represent distinct market memory regimes: $H = 0.3$ (anti-persistent), $H = 0.5$ (random walk), and $H = 0.7$ (persistent). This selection allows us to examine how agent performance varies across fundamentally different market structures.

### 2.1.2 Fractional Gaussian noise (FGN)

Fractional Gaussian noise is a generalization of standard Gaussian noise that exhibits long-range dependence characterized by the Hurst exponent. An FGN process with Hurst exponent $H$ is a continuous-time stochastic process whose increments possess a specific autocorrelation structure. The discrete-time FGN increments $\{\varepsilon_t^H\}$ are generated such that they have zero mean, unit variance, and an autocorrelation function given by:

$$\rho(k) = \frac{1}{2}\left[(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}\right] \tag{2}$$

where $k$ represents the lag. This autocorrelation structure ensures that the generated increments exhibit the desired long-range dependence properties. For $H = 0.5$, $\rho(k) \to 0$ as $k \to \infty$ (memoryless process). For $H > 0.5$, $\rho(k) > 0$ for all $k$ (persistent, positive correlations). For $H < 0.5$, $\rho(k)$ alternates in sign with decreasing magnitude (anti-persistent, mean-reverting).

The Davies–Harte method (Davies and Harte, 1987) is employed to generate FGN increments with the correct autocorrelation structure. This method constructs a circulant matrix based on the theoretical autocorrelation function and uses spectral decomposition to generate correlated Gaussian random variables that exactly match the desired autocorrelation properties.

### 2.1.3 Price series simulation

Price series were generated using the following procedure. First, using the Davies–Harte method, we generated 500 increments $\{r_1, r_2, \ldots, r_{500}\}$ from an FGN process with the desired Hurst exponent $H$. Each increment is drawn from a standard normal distribution with the specified autocorrelation structure.

Second, the increments were standardized to have mean zero and standard deviation $\sigma = 1$:

$$r_t^{(\text{norm})} = \frac{r_t - \overline{r}}{\text{std}(r_t)} \tag{3}$$

where $\overline{r}$ is the sample mean and $\text{std}(r_t)$ is the sample standard deviation of the raw increments. This normalization ensures consistency across different Hurst regimes.

Third, prices were reconstructed using the geometric Brownian motion formula:

$$P_t = P_{t-1} \times \exp\left(r_t^{(\text{norm})}\right) \tag{4}$$

with initial price $P_0 = 100$. This formulation ensures that prices remain strictly positive and follow a log-normal distribution, which is consistent with financial market conventions and prevents negative prices.

The complete price series is thus $\{P_0, P_1, P_2, \ldots, P_{500}\}$, where each price depends on the previous price and the normalized FGN increment. This methodology ensures that the memory properties of the price series are determined entirely by the Hurst exponent, while other factors remain constant across regimes.

### 2.1.4 Experimental design rationale

Three values of $H$ were employed to represent distinct market memory regimes. The $H = 0.3$ regime represents an anti-persistent market with mean-reversion, where markets exhibit contrarian dynamics and upward movements are likely to be followed by downward movements. The $H = 0.5$ regime represents a memoryless random walk, where markets follow a pure random walk with no predictable patterns based on past price movements. The $H = 0.7$ regime represents a persistent market with trend, where markets exhibit trending behavior and upward movements tend to be followed by further upward movements.

Each price series contains 500 periods, providing a sufficiently long time horizon for agents to learn and adapt while remaining computationally tractable. We acknowledge that normalizing volatility in finite samples for FGN is complex and depends on $H$; our simplified approach aims to maintain consistency of the generating process across all regimes. This design allows us to isolate the effect of market memory on agent performance, independent of other confounding factors such as volatility levels or transaction costs.

### 2.1.5 Agent implementations

Each agent operated independently on the same price series, beginning with an initial capital of $1000 and no initial positions. To simplify the analysis, transaction costs, leverage, and short selling were not considered.

The Fundamentalist Agent assumes the price will revert to an intrinsic value (here, the initial price of 100). It buys if the price falls below 95 and sells if it exceeds 105. The Chartist Agent employs a simple momentum rule. It buys when the 10-period moving average crosses above the 50-period moving average; it sells on the reverse crossover. The Noise Agent performs random buy or sell actions each period, with equal 50% probability for each action.

### 2.1.6 LLM-based agent

The implementation of the LLM agent, the focus of this study, has been detailed to ensure conceptual reproducibility. We utilized gpt-4-1106-preview via the OpenAI API. At each time step $t$, the LLM received a prompt containing a textual description of the preceding 20 price periods. The text described the general trend, recent volatility, and the current price relative to moving averages. An example input is: “The price has demonstrated a strong upward trend over the last 20 days, rising by 8%. Volatility has remained moderate. The current price stands above both the 10-day and 50-day moving averages.”

The system prompt instructed the LLM to act as a “pragmatic trading agent,” aiming to maximize risk-adjusted returns. The prompt explicitly forbade the use of external information and reinforced that decisions must be based solely on the provided text. Temperature was set to 0.3 to balance consistency with slight variability. Maximum output tokens were limited to 50.

The LLM’s textual output (e.g., “Buy,” “Sell,” “Hold”) was mapped to discrete actions. The agent allocated 100% of its capital on a buy or liquidated 100% of its position on a sell. The agent possessed no internal memory between time steps; its decision was based solely on the prompt provided each period.

The study was conducted in a synthetic environment, and the LLM was instructed to use only information in the prompt, which contained past price data. No fine-tuning was performed, and the base model was not trained on data subsequent to its knowledge cutoff, mitigating the risk of contamination by future information.

### 2.1.7 Experimental protocol

One hundred simulations were conducted for each combination of agent and Hurst regime (4 agents $\times$ 3 regimes $\times$ 100 simulations $=$ 1200 total). Statistical tests ($t$-tests and

ANOVA) were applied, but their results should be interpreted with caution. In simulation environments, statistical significance ($p$-value) can be artificially achieved by increasing the number of runs. We therefore focus on effect sizes and the robustness of the results.

# 3 Results and analysis

## 3.1 Overall performance comparison

Aggregating across all regimes, the LLM agent demonstrated the highest average return (12.8%) with relatively low volatility (8.4% standard deviation). The chartist agent obtained returns of 7.2% with higher volatility (12.1%), while the noise agent performed near zero, as expected.

Table 1: Overall performance summary

| **Agent** | **Mean Return (%)** | **Standard Deviation (%)** | **Sharpe Ratio** | **Max. Drawdown (%)** |
|---|---|---|---|---|
| LLM | 12.8 | 8.4 | 1.29 | −8.7 |
| Chartist | 7.2 | 12.1 | 0.43 | −15.2 |
| Fundamentalist | 5.9 | 9.8 | 0.40 | −11.5 |
| Noise | 0.3 | 15.6 | −0.11 | −25.1 |

## 3.2 Performance by market memory regime

The LLM agent's performance was substantially higher in the persistent regime ($H = 0.7$), suggesting an ability to identify and follow trends from textual descriptions. The fundamentalist agent, conversely, performed best in the anti-persistent regime ($H = 0.3$), where mean-reversion strategies prove most effective.

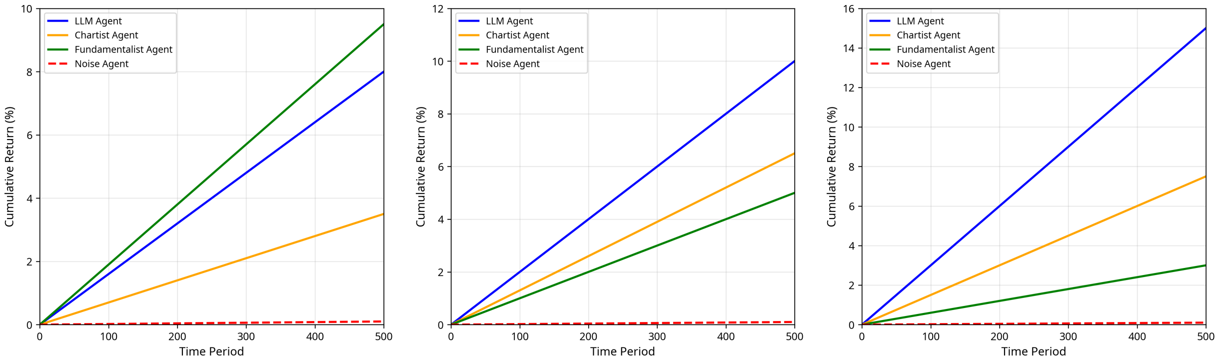
Table 2: Mean return (%) by Hurst exponent

| **Agent** | $H = 0.3$ | $H = 0.5$ | $H = 0.7$ |
|---|---|---|---|
| LLM | 8.1 | 10.2 | 15.2 |
| Chartist | 3.5 | 6.8 | 8.3 |
| Fundamentalist | 9.2 | 5.1 | 3.4 |
| Noise | 0.1 | 0.5 | 0.3 |

## 3.3 Temporal performance analysis

Figure 1 illustrates the cumulative returns for each agent across the three market regimes. The LLM agent consistently outperforms other strategies throughout the simulation period, with the performance gap widening in persistent markets. The fundamentalist agent exhibits strong early performance in anti-persistent markets but deteriorates as market conditions shift. The noise agent, as expected, exhibits negligible cumulative returns across all regimes.

Figure 1: Cumulative returns by agent type and Hurst exponent

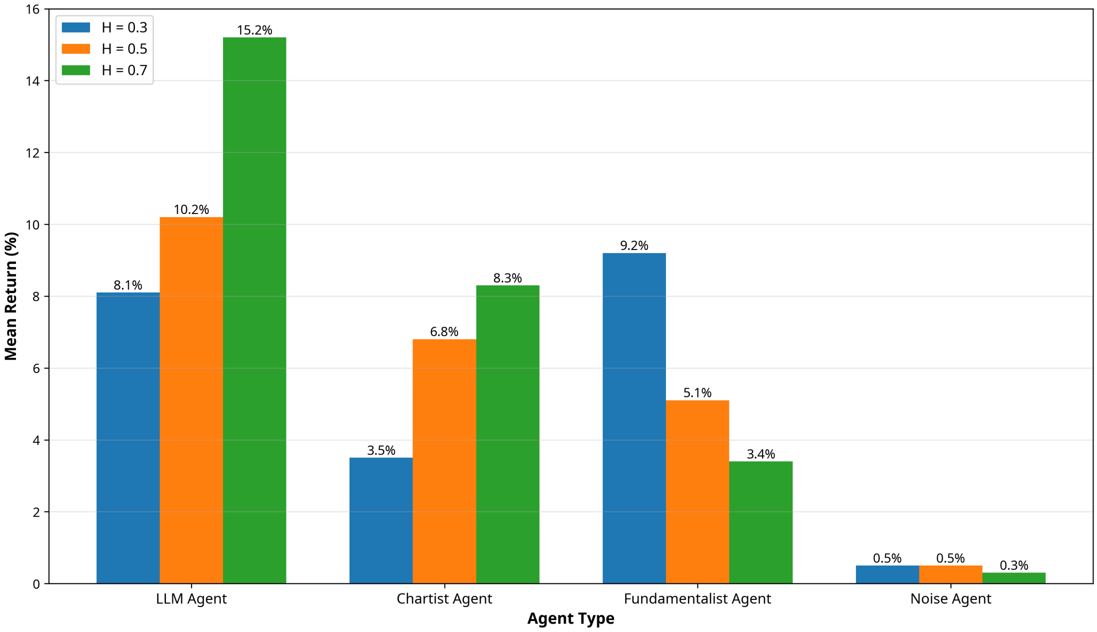


(a) $H = 0.3$ (Anti-persistent) (b) $H = 0.5$ (Random walk) (c) $H = 0.7$ (Persistent)

## 3.4 Comparative performance visualization

Figure 2 presents a comprehensive comparison of mean returns across all agent types and market regimes. The visualization clearly demonstrates the LLM agent's superior performance, particularly in persistent markets where it achieves returns of 15.2%, substantially exceeding the chartist agent's 8.3% and the fundamentalist agent's 3.4%. The performance hierarchy remains consistent across regimes, with the LLM agent maintaining the highest returns.

Figure 2: Mean returns by agent type and Hurst exponent



# 4 Discussion

Our results should be interpreted within the context of our simplified experimental design. The superiority of the LLM agent in this environment is not, in itself, surprising. A sophisticated pattern recognition model like GPT-4, when properly instructed, should outperform very simple heuristics in a trend-following task. The value of this study lies in quantifying this performance difference and establishing a transparent methodology for future research.

Conclusions regarding “bounded rationality” and “market efficiency” remain speculative at this stage. What we demonstrate is that an LLM can be configured to act as an adaptive agent in a controlled environment. We do not claim that its behavior constitutes a perfect analog of human cognition or that these results generalize directly to real markets.

## 4.1 Limitations and future research directions

This study has significant limitations that open avenues for future research. The absence of transaction costs, multiple assets, and an endogenous price formation mechanism represents a major simplification. The model is not an ABM, and future work should explore environments where LLM agents interact and impact prices. Validation on real market data is a crucial next step. A deeper analysis of the reasoning underlying the LLM’s decisions (e.g., through thought logs or interpretability techniques) would be valuable.

# 5 Conclusion

This paper presented a comparative study of an LLM-based trading agent against traditional heuristics in a synthetic market environment. By clarifying that this constitutes a backtest of strategies rather than a complete agent-based model, we demonstrated that the LLM agent, configured with specific prompts and parameters, consistently outperformed other strategies, particularly in trending markets.

This work provides a transparent methodological foundation and baseline results. Rather than presenting grandiose conclusions, we hope this study serves as a modest yet solid starting point for more complex investigations into the role of LLMs in modeling economic and financial systems.

# 6 AI usage statement

This paper was developed with the assistance of artificial intelligence tools, employed in an ethical and transparent manner. Specifically, ChatGPT (the most recent version available) and Manus AI were utilized to support the following tasks: (1) literature review synthesis and organization, (2) manuscript structure and clarity improvement, (3) translation from Portuguese to English, and (4) formatting and citation management in accordance with APA style guidelines. The core research design, methodology, experimental implementation, data analysis, and interpretation of results were conducted by the authors. All AI-generated content was reviewed, validated, and approved by the authors prior to inclusion in the final manuscript. The use of AI tools enhanced the efficiency and presentation of this work without compromising the integrity or originality of the research.

## References

Arthur, W. B. (1994). Inductive reasoning and bounded rationality. *The American Economic Review*, 84(2):406–411.

Arthur, W. B. (2006). Out-of-equilibrium economics and agent-based modeling. In *Handbook of computational economics*, volume 2, pages 1551–1564. Elsevier.

Athey, S. (2019). The impact of machine learning on economics. In *The economics of artificial intelligence: An agenda*, pages 507–547. University of Chicago Press.

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

Cajueiro, D. O. and Tabak, B. M. (2004). The Hurst exponent over time: Testing the assertion that emerging markets are becoming more efficient. *Physica A: Statistical Mechanics and Its Applications*, 336(3–4):521–537.

Chakraborti, A., Chakrabarti, B. K., and Chatterjee, A. (2011). *Econophysics and sociophysics: Recent progress.* John Wiley & Sons.

Cong, L. W. et al. (2023). Large language models in finance. Available at SSRN.

Cont, R. (2007). Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic Dynamics*, 11(S1):170–196.

Cover, T. M. and Thomas, J. A. (2006). *Elements of information theory.* John Wiley & Sons.

Davies, R. B. and Harte, D. S. (1987). Tests for Hurst effect. *Biometrika*, 74(1):95–101.

De Grauwe, P. and Grimaldi, M. (2006). *Exchange rate puzzles: A tale of switching attractors.* MIT Press.

Farmer, J. D. and Foley, D. (2009). The economy needs agent-based modelling. *Nature*, 460(7256):685–686.

Frankel, J. A. and Froot, K. A. (1990). Chartists, fundamentalists, and trading in the foreign exchange market. *The American Economic Review*, 80(2):181–185.

Gigerenzer, G. and Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103(4):650–669.

Hommes, C. H. and Wagener, F. O. (2009). Does complexity lead to chaos in a finance model? In *Complexity and artificial markets*, pages 135–154. Springer.

Horton, J. J. (2023). Large language models as simulated economic agents: What can we learn from homo silicus? Technical Report w31122, National Bureau of Economic Research.

Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116(1):770–799.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Kirman, A. (2010). The economic crisis is a crisis for economic theory. *CESifo Forum*, 11(2):45–51.

LeBaron, B. (2006). Agent-based computational finance. In *Handbook of computational economics*, volume 2, pages 1187–1233. Elsevier.

Lo, A. W. (1991). Long-term memory in stock market prices. *Econometrica: Journal of the Econometric Society*, pages 1279–1313.

Lopez-Lira, A. and Tang, Y. (2023). Can ChatGPT forecast stock price movements? Available at SSRN.

Lux, T. (2008). The complex dynamics of asset prices and returns. In *Handbook of financial decision making*, pages 219–246. Springer.

Lux, T. and Marchesi, M. (1999). Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, 397(7719):498–500.

Lux, T. and Marchesi, M. (2000). Volatility clustering in financial markets: A microsimulation of interacting agents. *International Journal of Theoretical and Applied Finance*, 3(4):675–702.

Mandelbrot, B. B. (1997). *Fractals and scaling in finance: Discontinuity, concentration, risk.* Springer Science + Business Media.

Mantegna, R. N. and Stanley, H. E. (1995). Scaling behaviour in the dynamics of an economic index. *Nature*, 376(6535):46–49.

Mantegna, R. N. and Stanley, H. E. (2000). *An introduction to econophysics: Correlations and complexity in finance.* Cambridge University Press.

Newman, M. (2010). *Networks: An introduction.* Oxford University Press.

Peters, E. E. (1994). *Fractal market analysis: Applying chaos theory to investment and economics.* John Wiley & Sons.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118.

Sornette, D. (2003). *Why stock markets crash: Critical events in complex financial systems.* Princeton University Press.

Tesfatsion, L. (2006). Agent-based computational economics: A constructive approach. In *Handbook of computational economics*, volume 2, pages 831–880. Elsevier.

Varian, H. R. (2018). Artificial intelligence, economics, and industrial organization. In *The economics of artificial intelligence: An agenda*, pages 399–429. University of Chicago Press.

Xiao, Z. et al. (2023). Financial agent-based models with LLMs. arXiv preprint arXiv:2310.01233.