Proceedings of the 5th International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/2026.39205

Artificial Intelligence in Financial Market Microstructure: Implications for Liquidity, Price Discovery, and Systemic Risk

Kaixi Xu[1], Hezi Yin[2,*]

[1]*Guangzhou Zhixin High School, Ersha Island Campus, Guangzhou, China*
[2]*The Potter's School, Springfield, USA*
**Corresponding Author. Email: yineudora@gmail.com*

Abstract. Market microstructure research is about the understanding of how trading rules, information asymmetries and the behaviour of the participants affect asset prices and market quality. The future of financial markets around the world is being transformed by artificial intelligence (AI). AI has reshaped market execution, transforming the dynamics of liquidity, price discovery, and market-making, from high-frequency trading (HFT) to deep learning and large language models (LLMs). This paper takes a holistic approach to synthesizing the role of AI in the markets, leveraging 15 fundamental papers and studies on classical market microstructure, algorithmic execution, deep learning order book forecasting, and decentralized finance (DeFi) game theory. This research considers AI as a new General-Purpose Technology (GPT) that augments the classical continuous auction models. This study then examines the empirical impact of algorithmic and HFT on price spreads and market liquidity. The paper then explores deep learning models and LLMs for forecasting order books and the sentiment analysis of unstructured data. Lastly, it examines the systemic risks associated with algorithmic authority and rational opacity and suggests future regulatory and economic modelling.

Keywords: Artificial Intelligence, Market Microstructure, Algorithmic Trading, Price Discovery, Systemic Risk

1. Introduction

In the past few decades, financial markets have changed from paper floor trading to electronic trading. MacKenzie analysed the Chicago Mercantile Exchange (CME) and documented what happened when open outcry was replaced by electronic matching systems [1]. This swap paved the way for mechanized trading executions, floor architecture, and order-matching systems, and left the technical and institutional groundwork for the high frequency trading and algorithmic trading. This structural change in addition to increasing the speed of execution modified strategic choices and human-machine interaction, and it was the beginning of modern algorithmic finance.

As computing power and data have increased exponentially, Artificial Intelligence (AI) has transformed from being a technology that automates only processes to a General Purpose Technology (GPT) in its entirety. The impact of AI in financial markets is not just about efficiency; it's about a paradigm shift in how market dynamics are understood and assessed through empirical

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/2026.39205

studies. AI applications are now used in execution, risk management, market surveillance and regulatory compliance, changing the financial market microstructure from the inside.

Before the emergence of machine learning, traditional market microstructure theory had given a solid analytical framework for the analysis of information flow and price discovery. Kyle's continuous auction model continues to be a basic paradigm [2]. The framework includes strategic interaction between three market participants: an informed trader, noise traders and competitive market makers. The informed trader adjusts the quantity of his/her order based on the information gathered and the market maker adjusts prices dynamically as a function of order flow. This interaction produces the standard quantitative measure of the market illiquidity and price impact, Kyle's lambda. The framework is at its essence a description of the process of private information gradually becoming part of prices via order flow, with noise trading acting as liquidity backup.

Today, the term informed trader has come to refer to anything beyond the classic case of fund managers buying fundamental tips. Today, automated systems enable by high-performance compute and adaptive learning models, to process high-dimensional, unstructured public data instantaneously, to feed faint predictive signals. As a result, the literature assessing the impact of AI on microstructure has grown like Topsy.

In response, this review aims to answer the research question: What are the effects of artificial intelligence on the financial market microstructure, specifically liquidity, price discovery, market efficiency and systemic risk. The rest of this review covers the basic theory, empirical execution results, modern deep learning architectures, systemic mechanisms, and suggestions for future research.

2. AI applications and market microstructure outcomes

2.1. Algorithmic trading strategies and market making mechanism

Addy et al. highlighted that modern algorithmic trading has moved beyond static statistical arbitrage and mean reversion strategies in the integration of AI in quantitative execution strategies [3]. Modern execution is rather based on adaptive market-making and dynamic order routing, with deep learning and reinforcement learning algorithms. These architectures regularly analyse the market in real time and dynamically re-stagger orders across the fragmented market. In the case of dynamic liquidity provision, Hendershott and Riordan examined the order book data at high frequencies and showed that algorithmic traders proactively monitor and provide liquidity [4]. When the bid-ask spread increases, algorithmic market participants systematically provide liquidity to the market through limit orders, while they demand liquidity through market orders during times of contraction of the bid-ask spread. This dynamic offering helps to stabilize liquidity and improve price discovery. Moreover, while human traders typically react to market shocks within a few hours or days, automated algorithms can respond in milliseconds.

2.2. Market quality, bid-ask spread and high-frequency trading games

HFT is a specific form of algorithmic trading and has mixed, intricate impacts on market quality. Research integrated the structural effect of HFT on market stability, bid-ask spreads and social welfare [5]. In general, HFT reduces the cost on the explicit side of trading by the retail and institutional investor by narrowing the inside spread by sending and cancelling orders in high frequencies. High-frequency market makers utilise automated queue prioritisation and arbitrage

DOI: 10.54254/2754-1169/2026.39205

across multiple venues to narrow the spread between the bids and asks of correlated assets to near marginal operational cost, thus providing pricing efficiency.

But, in times of extreme stress these high-frequency algorithms quickly pull their limit orders to reduce inventory exposure, which leads to the reduction of venue liquidity and can produce "arms races" based on speed, harming market quality. On the basis of these, Nahar et al. pointed out that the adverse selection risk faced by the market participants is increasing in the HFT era due to latency asymmetries [6]. HF traders are able to take advantage of sub-millisecond execution to make pre-emptive trades or cross-venue latency arbitrage. Upon arrival of news on the public feeds, the ultra-low-latency algorithms are able to remove limit orders that are "stale" - posted by slower participants in the market than themselves - before they can be cancelled.

This mechanism deters traditional, slower speed liquidity providers and increases microstructural fragility during more volatile trading periods. Non-HFT liquidity providers compensate for the increased likelihood of being picked off by a latency-arbitrageur by increasing their quote spread or decreasing the size of the quote. This strategic interaction can be considered as a non-cooperative game between the low-latency firms and the traditional market makers. As a result, there is a constant trade-off to be struck between the execution effectiveness and structural integrity when regulators are considering a new matching engine, whether in the form of speed bumps or regular batch auctions, to ensure fair treatments.

2.3. Zero-commission environment, retail investor behaviour and HFT micro-interaction

Fintech developments and payment for order flow (PFOF) revenue models have dramatically reshaped the way retail trades and engender new types of microstructural relationships. Eaton et al. took advantage of exogenous service disruptions on commission-free trading platforms (such as Robinhood) to study the relationship between retail order flow on these platforms and high-frequency market makers [7]. They found that the retail order flow has a significant proportion of noise and correlation, which increases inventory holding costs and market makers' directional exposure.

When retail order flow dries up suddenly as happens during platform outages, the high-frequency market makers reduce the bid-ask spreads in lit exchanges. This structural response implies that algorithmic spread pricing is directly related to retail noise trading. Off exchange internalizes often internalize retail order flow that they know is often being done by un-informed traders. But when retail trading surges tend to become very correlated (either from social media momentum or gamification of platform interfaces) retail order flow acts as a sum of directional shocks and not as an uncorrelated source of "noise. With very little fundamental return predictability in retail flows at long time scales, strong retail order flow bursts raise the market maker's inventory risk in the short term, and reduce the baseline market liquidity in secondary markets.

3. The modelling approach

3.1. High-frequency limit order book (LOB) and deep learning modelling

Although limit order books (LOBs) contain rich information about the high-frequency supply and demand dynamics at various price levels, they are extremely non-linear, very high dimensional and noisy. Raw tick data is detailed information on the placing or cancelling of limit orders, and market order executions. To gain predictive signals from raw LOB state vectors, it needs computational

DOI: 10.54254/2754-1169/2026.39205

architectures that can process spatial relationships at different price levels, and temporal relationships between different order arrival sequences.

To overcome these problems, Zhang et al. proposed the DeepLOB architecture which is the integration of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks [8]. DeepLOB is able to automatically learn spatial features from the multi-level price-volume matrices and also learn the temporal order flow dependencies. Spatial convolution layers are used to discover microstructural patterns, including order book imbalances, support/resistance walls, and queue exhaustion, by looking at the price levels and the size of orders. The temporal sequence of these spatial feature vectors is then fed into the LSTM layers to capture long-memory dynamics and persistence of the order flow. Empirically, DeepLOB performs much better than traditional linear econometric models like Vector Autoregressions (VARs) in predicting short-horizon mid-price movements in benchmark datasets, such as FI-2010, and LSE tick data.

Popa and Muresan examined multi-scale financial forecasting via classic regularized linear models and gradient-boosted decision trees, and deep neural architectures [9]. They showed how financial modelling is evolving from univariate time-series forecasting to multi-modal models, which combine high-frequency LOB dynamics with macroeconomic indicators. Tree-based ensembles are still a good baseline for tabular problems, because they are less prone to overfitting and are more interpretable; however, the advantage of deep learning models in high-frequency classification problems is that they offer greater performance. Unique to ultra-high-frequency environments are the deep neural networks which are able to model non-linear interactions between features without the need for manual feature engineering.

3.2. Large language model (LLM) and unstructured emotional processing

With the advent of the LLMs, unstructured text processing has undergone a revolution in financial markets. In the past financial sentiment analysis has been based on the static lexicon-based dictionary (i.e. Loughran-McDonald) or bag-of-words NLP classifiers. The challenge of these traditional approaches lies in their inability to cope with context, negation, financial jargon and complicated syntactic structures in corporate reports and earnings call transcripts.

Lopez-Lira and Tang tested ChatGPT's ability to read sentiment from news headlines, showing that sentiment scores obtained from LLMs predict the direction of short-term stock returns [10]. The authors were able to statistically prove the predictive power of LLM sentiment scores for next-day and intraday returns by asking the LLM to classify a news headline as positive, negative or neutral for a particular firm's stock price. These types of models greatly surpass the classic dictionary-based methods and previous NLP models. Importantly, the return predictability of sentiment is present in more recent models (GPT-3.5 onwards), but not in earlier models (GPT-1, GPT-2, BERT), suggesting that understanding economic sentiment is an emergent property of model scale.

The unstructured passing of sentiment is especially strong in emerging markets. Taking Chinese equity markets as examples, research analysed the text information from social media (e.g., Eastmoney, Guba) and concluded that the investor sentiment has a non-linear predictability on the excess returns and intraday volatility of the markets [11]. The study developed a Naïve Bayes text classification model and integrated it with LSTM to show that social sentiment indices show direct correlation with the volatility indices of the markets. In emerging markets, retail investors are very active on social media, and there are feedback effects between social sentiment, retail order flow and stock price volatility. Therefore, the use of AI to enable text mining in the retail sector validates that sentiment is not orthogonal noise but rather a clear microstructural price signal.

DOI: 10.54254/2754-1169/2026.39205

3.3. Financial AI architecture, knowledge enhancement and autonomous agents

Recent literature pays attention to knowledge enhanced AI architectures to overcome the black-box nature and hallucination issues of purely data-driven deep learning. Liu summarized cutting-edge financial AI pipelines, notably the use of Graph Neural Networks (GNNs) to capture intricate corporate ownership, supply-chain, customer-supplier, and institutional co-ownership relationships [12]. Traditional time-series models assume the assets of a particular firm are independent processes. By contrast, GNNs can model cross-asset spillovers by conveying the message vectors from one node to another over the graph edges that represent the economic relationships. Relational knowledge can improve the interpretability and economic validity of high-frequency forecast model in time-series models and help algorithms predict the cross-asset return momentum.

In parallel, financial AI is going beyond passive predictive instruments to become independent trading agents. LLM cores can be integrated into multi-agent systems to handle parsing real-time financial filings, generate trading strategies, make automated risk management decisions, and execute orders through a matching engine. These independent agents are called specialized synthetic market participants. But, when it starts using autonomous LLM agents in production, significant issues emerge concerning the reliability of the system, transparency in decisions, the time it takes for the machine to execute, and regulatory compliance. LLMs can also be probabilistic, meaning that their execution results may be unpredictable in case of an out-of-distribution market, which can result in operational risks for quantitative funds.

4. Analysis

4.1. Algorithm authority and rational opacity

AI algorithms take over functions in pricing, order execution, and the characteristics of financial markets change. Frimpong and Mamuti created the term algorithmic price discovery loop and rational opacity, which assume that algorithms are now given the cognitive capacity to execute pricing [13]. Today, price discovery is done on an ongoing basis via an algorithm, not a person. While deep neural models are statistically very efficient in forecasting, they do not reveal causal mechanisms, resulting in the "rational opacity".

Market actors know that the price of a product on the market is reflecting the information regarding that product as of when it happens, but they do not have the ability to see what caused the algorithm to price that product and what weights the algorithm applied to the factors. This opacity makes the traditional regulatory frameworks more difficult to maintain, detection of insider trading is harder, and standards of market transparency are complicated. Algorithms can respond to non-linear signals at the same time, resulting in price changes without an apparent fundamental news flow.

To bridge the gap with this view, Rahmani et al. performed a survey on AI uses in trading, analytics and risk control [14]. In the future, markets will be increasingly dominated by algorithms acting on their own and replacing the role of classic market makers and human intermediaries, and the provision of market liquidity and the dynamics of market execution will increasingly be driven by deep learning and deep reinforcement learning algorithms. But in the case of several rival market participants learning algorithms with the same historical data and objective function, algorithmic strategies trend to become structurally synchronized. This strategy homogenization makes the system vulnerable, especially during times of market stress, when algorithms begin to draw on the liquidity pool simultaneously, or when a one-sided liquidation order is placed.

DOI: 10.54254/2754-1169/2026.39205

4.2. Decentralized finance (DeFi) and centralized liquidity market making (CLMM)

In addition to central exchanges, blockchain has also brought new decentralized microstructure systems. Tang et al. used game theory to assess strategic interactions between liquidity providers (LPs) in CLMMs (e.g., Uniswap V3) [15]. They set up a water-filling" liquidity provisioning strategy, which is based on a unique Nash equilibrium.

Automated AI agents are more and more acting as algorithmic LPs, dynamically rebalancing their liquidity ranges in line with asset volatility in decentralized venue structures. Empirical evidence shows that, while budget constrained LPs use 100% of their capital, unconstrained LPs maintain reserve capital to maximize yield. But for high-risk decentralized pools, these empirical LP behaviours differ from the theoretical Nash equilibria, indicating structural inefficiencies that can be leveraged by automated AI agents.

5. Research outlook

Cross-modal and Agent Game Modelling: It can construct a multi-agent reinforcement learning model to simulate algorithmic competition and emergent behaviours in a realistic high-frequency micro-environment by combining high-frequency limit order book data with large language model agents. Meanwhile, this configuration challenges the researchers to closely monitor the reliability and comprehensibility of LLM agents after practitioners deploy them. Algorithmic manipulation and measuring price discovery: Quantitative measures assessing the amount of "rational opacity" that stifles causal price discovery should be developed, and there is a need to investigate regulatory tools based on Explainable AI. New cross-chain and decentralized market-making with microstructure: It can integrate AI market-making strategies seamlessly into decentralized auction and liquidity protocols such as DeFi CLMM. In doing so, researchers can analyse the quality of markets and risk transfer in permissionless settings, and analysts can test their theories using real data from the chain. Moreover, the comparison of different markets and regulatory frameworks, as well as macro-micro linkages of AI on financial stability in the long-term, will provide fertile topics for future research.

6. Conclusion

In the end, AI is reshaping the financial micro-market landscape. In principle, the technology is an update to the classic Kyle model on "informed trading," and represents a new paradigm of computing and data-enabled trading. Real-life examples such as algorithmic market-making, DeepLOB, and high-frequency forecasting with the LLM have significantly increased market liquidity and accelerated information processing in the markets. However, at the level of mechanics and systemic risk, algorithmic dominance and rational opacity, coordinated flash crashes from copying dynamic strategies, and decentralized game dynamics in CLMM raise questions for the existing microstructural theories.

Authorship

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] MacKenzie, D. (2015). Mechanizing the Merc: The Chicago Mercantile Exchange and the rise of high-frequency trading. *Technology and Culture, 56*(3), 646–675. https://doi.org/10.1353/tech.2015.0102

DOI: 10.54254/2754-1169/2026.39205

[2] Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica, 53*(6), 1315–1335. https://doi.org/10.2307/1913210
[3] Addy, W. A., Ajayi-Nifise, A. O., Bello, B. G., Tula, S. T., Odeyemi, O., & Falaiye, T. (2024). Algorithmic trading and AI: A review of strategies and market impact. *World Journal of Advanced Engineering Technology and Sciences, 11*(1), 258–267. https://doi.org/10.30574/wjaets.2024.11.1.0054
[4] Hendershott, T., & Riordan, R. (2013). Algorithmic trading and the market for liquidity. *Journal of Financial and Quantitative Analysis, 48*(4), 1001–1024. https://doi.org/10.1017/S0022109013000471
[5] Menkveld, A. J. (2016). The economics of high-frequency trading: Taking stock. *Annual Review of Financial Economics, 8*, 1–24. https://doi.org/10.1146/annurev-financial-121415-033010
[6] Nahar, J., Nishat, N., Shoaib, A. S. M., & Hossain, Q. (2024). Market efficiency and stability in the era of high-frequency trading: A comprehensive review. *International Journal of Business and Economics, 1*(3), 1–13. https://doi.org/10.62304/ijbm.v1i3.166
[7] Eaton, G. W., Green, T. C., Roseman, B. S., & Wu, Y. (2022). Retail trader sophistication and stock market quality: Evidence from brokerage outages. *Journal of Financial Economics, 146*(2), 502–528. https://doi.org/10.1016/j.jfineco.2022.08.002
[8] Zhang, Z., Zohren, S., & Roberts, S. (2019). DeepLOB: Deep convolutional neural networks for limit order books. *IEEE Transactions on Signal Processing, 67*(11), 3001–3012. https://doi.org/10.1109/TSP.2019.2907260
[9] Popa, F. G., & Muresan, V. (2025). Artificial intelligence in finance: From market prediction to macroeconomic and firm-level forecasting. *AI, 6*(11), Article 295. https://doi.org/10.3390/ai6110295
[10] Lopez-Lira, A., & Tang, Y. (2026). Can ChatGPT forecast stock price movements? Return predictability and large language models. *Journal of Financial Economics, 184*, Article 104335. https://doi.org/10.1016/j.jfineco.2026.104335
[11] Li, Y., Bu, H., Li, J., & Wu, J. (2020). The role of text-extracted investor sentiment in Chinese stock price prediction with the enhancement of deep learning. *International Journal of Forecasting, 36*(4), 1541–1562. https://doi.org/10.1016/j.ijforecast.2020.05.001
[12] Liu, J. (2024). A survey of financial AI: Architectures, advances and open challenges [arXiv preprint, arXiv:2411.12747]. https://doi.org/10.48550/arXiv.2411.12747
[13] Frimpong, V., & Mamuti, A. (2026). Opaque price control and algorithmic authority in financial markets. *Encyclopedia, 6*(1), Article 19. https://doi.org/10.3390/encyclopedia6010019
[14] Rahmani, A. M., Rezazadeh, B., Haghparast, M., Chang, W.-C., & Ting, S. G. (2023). Applications of artificial intelligence in the economy, including applications in stock trading, market analysis, and risk management. *IEEE Access, 11*, 80769–80793. https://doi.org/10.1109/ACCESS.2023.3300036
[15] Tang, W., El-Azouzi, R., Lee, C. H., Chan, E., & Fanti, G. (2025). Game theoretic liquidity provisioning in concentrated liquidity market makers. *Proceedings of the ACM on Measurement and Analysis of Computing Systems, 9*(1), Article 7, 1–45. https://doi.org/10.1145/3711700