Nanosecond Microstructure: High-Frequency Traders Participation Stylized Facts

Gabin Taibi[a,b], Joerg Osterrieder[a], Stefan Schlamp[c]

[a] *Faculty of Behavioral Management and Social Sciences, University of Twente, Enschede, Netherlands*
[b] *Department of Applied Data Science and Finance, Bern University of Applied Sciences, Bern, Switzerland*
[c] *Market Data and Services, Deutsche Börse, Eschborn, Germany*

Abstract

This paper introduces a novel methodology for classifying market participants in electronic financial markets based on their reaction times and measuring their participation. As technological innovation reshapes trading, accurately distinguishing trader types is critical for understanding market dynamics and informing regulation. Using nanosecond-level timestamp data from Deutsche Börse, we analyze post-trade latencies to separate ultra-fast traders (UFTs) relying on field-programmable gate arrays (FPGAs), high-frequency traders (HFTs), and other conventional participants. Transparent latency thresholds enable this classification, after which we document their behavior in terms of (i) participation shares, (ii) price discovery via a 15-second mark-out signal-to-noise ratio, and (iii) average reaction latency. We pair these with market quality metrics including order-imbalance volatility, Amihud illiquidity, and high-frequency return diagnostics such as autocorrelation and variance-ratio tests. The methodology is applied to Euro STOXX 50 Index Futures (FESX), examining the most actively traded contract each day from January to August 2025.

Keywords: Market Microstructures, High-Frequency Trading, Price Discovery, Market Efficiency, Liquidity

Email addresses: `gabin.taibi@utwente.nl` (Gabin Taibi), `joerg.osterrieder@utwente.nl` (Joerg Osterrieder), `stefan.schlamp@deutsche-boerse.com` (Stefan Schlamp)

1. Introduction

1.1. Context of the Study

The rapid advancement of electronic markets has reshaped the financial landscape, driven by technological innovations such as ultra-low-latency trading systems, high-performance computing, and Field-Programmable Gate Array (FPGA) technology. These technological developments have enabled the rise of Ultra-Fast Traders (UFTs) and High-Frequency Traders (HFTs), who leverage ultra-fast reaction times to gain a competitive edge in financial markets. The emergence of these sophisticated market participants has fundamentally altered price formation mechanisms, liquidity provision, and the competitive dynamics of modern trading venues.

In modern financial markets, a variety of participants operate with different objectives and strategies, including market makers, speculators, hedgers, and institutional investors. Each group contributes to market liquidity and efficiency in distinct ways, yet the precise classification of these participants is critical to understanding their impact on price formation, market dynamics, and liquidity provision. Particularly, UFTs and HFTs, equipped with cutting-edge technological infrastructures, stand out due to their ability to react to market signals within microseconds, distinguishing them from conventional participants.

The technological arms race in financial markets has intensified over the past decade, with firms investing massively in infrastructure to shave microseconds off their reaction times. This competition has profound implications for market quality, fairness, and stability. While proponents argue that HFT improves liquidity and price efficiency, critics raise concerns about market manipulation, flash crashes, and the creation of a two-tiered market system.

1.2. Research Questions and Objectives

This paper addresses two fundamental research questions that guide our empirical investigation. First, we examine how to accurately classify market participants based on their technological capabilities, specifically their reaction times to market events. Then, we analyze how different categories of high-speed traders (UFTs versus HFTs) differ in their market impact and trading strategies.

To answer these questions, we develop a novel classification methodology using nanosecond-level timestamp data from Deutsche Börse's High-

Performance Timestamp (HPT) system and the A7 Analytics Platform. Our approach measures the reaction latency between trigger events (trades or order book updates) and subsequent order submissions, allowing us to categorize traders into three distinct groups. UFTs rely on Field-Programmable Gate Array (FPGA) technology and advanced network optimizations to achieve deterministic response times and access markets at record speeds, often targeting latencies in the sub-microsecond range. Unlike conventional processors, FPGAs are low-computation hardware devices that do not support complex algorithms but are optimized to execute simple conditional logic (e.g., price or size thresholds) with extremely low latency. High-Frequency Traders utilize optimized software systems, fast data-based decision algorithms and co-location facilities to accurately forecast short-term price movements, aiming for single digit microsecond latencies. Non-HFT participants, represent traditional market participants using standard electronic trading infrastructure. This classification reflects fundamental technological constraints, as each tier corresponds to distinct hardware and software architectures.

1.3. Main Contributions

This paper makes several key contributions to the literature on high-frequency trading and market microstructure.

First, we propose a novel, data-driven framework to classify market participants based directly on their measured reaction times, rather than relying on indirect proxies such as order-to-trade ratios or aggregate trading volumes. This approach provides a more precise and technologically grounded categorization that reflects the actual capabilities of different trading groups.

Second, we leverage a unique dataset from Deutsche Börse that contains nanosecond-level timestamps for both trades and all market events across Eurex, Xetra, and EEX instruments. This level of detail allows us to observe the technological arms race with exceptional granularity and to distinguish traders operating at different latency tiers. Covering the period from January to August 2025, the dataset captures market dynamics under varying participation intensities. The combination of high-precision timestamps and full order book reconstruction enables a direct analysis of participation rates and their effects on liquidity and price formation.

Third, we present new empirical evidence on the impact of high-speed trading across distinct trader categories. We examine the participation patterns of UFTs, HFTs, and conventional participants, and evaluate their respective contributions to liquidity provision, order book stability, and short-horizon price efficiency. The results reveal systematic differences in how each category interacts with the market.

Fourth, we inform the policy debate on high-frequency trading by disentangling the effects of UFTs and HFTs. Our evidence shows that these groups influence market quality in different ways, suggesting that regulation should distinguish across speed tiers rather than treating all high-speed traders as a single group. Such a differentiated perspective provides a foundation for more targeted interventions that preserve the efficiency benefits of electronic market making while addressing potential risks.

1.4. Paper Structure

The remainder of the paper is structured as follows. Section 2 reviews the existing literature on trader classification, high-frequency trading, and its effects on market microstructure. Section 3 details our dataset, classification methodology, and the metrics employed to evaluate market quality. Section 4 reports the empirical findings, focusing on how UFTs, HFTs, and conventional traders differ in their trading behavior and market impact. Section 5 interprets these results in the broader context of market design, technological competition, and regulatory implications. Finally, Section 6 summarizes the key insights, outlines the limitations of the study, and suggests avenues for future research.

2. Literature Review

The following section highlights the literature of key concepts and research relevant to understanding the mechanisms driving ultra-fast trading in modern financial markets. We begin by exploring the foundational role of Limit Order Books (LOBs) and the broader market microstructure in shaping price formation and liquidity dynamics. Next, we examine the rise of UFTs and HFTs, focusing on their strategies and the technologies that enable them to operate within these high-speed environments. Finally, we address the nanosecond races, where competition to minimize latency has intensified, giving certain participants a critical edge in trade execution and market impact.

2.1. Ultra-Fast and High-Frequency Traders Activity Monitoring

The classification and monitoring of high-frequency trading activity has been a central challenge in the market microstructure literature. Early work by Hasbrouck and Saar [24] established the importance of low-latency trading in modern markets, demonstrating that speed advantages translate into profitable trading opportunities. Menkveld [29] provided one of the first detailed analyses of a single HFT firm, revealing that HFTs act primarily as market makers, earning profits from the bid-ask spread while managing inventory risk.

The technological aspects of high-frequency trading have been extensively documented. Aldridge [3] describes the evolution from software-based to hardware-based trading systems, with FPGAs offering deterministic latency in the sub-microsecond range. Budish et al. [18] characterize the continuous limit order book as a flawed market design that incentivizes socially wasteful investment in speed, proposing discrete-time batch auctions as an alternative.

Several approaches have been developed to identify HFT activity in market data. Kirilenko et al. [27] use account-level data from the E-mini S&P 500 futures market to classify traders based on their daily trading patterns. They identify HFTs as those with high volume, low inventory, and short holding periods. Brogaard et al. [17] use a similar approach with NASDAQ data, finding that HFTs participate in 68% of traded volume. However, these classification methods rely on proprietary data not generally available to researchers.

Our approach differs fundamentally by using reaction times as the primary classification criterion. This method is inspired by Baron et al. [13], who show that speed is the defining characteristic of successful HFT strategies, and Aquilina et al. [8] who demonstrate that even microsecond advantages can be monetized in modern markets. By directly measuring reaction latencies at nanosecond precision, we can distinguish between different technological capabilities rather than relying on indirect behavioral proxies that may conflate different trader types.

2.2. Fair-Value Modeling and Price Discovery

The concept of fair value in high-frequency markets has evolved significantly with the rise of algorithmic trading. The traditional midpoint of the bid-ask spread, while simple to calculate, has been shown to be a biased estimator of the true price when order flow is imbalanced or when there is asymmetric information. Modern approaches incorporate order book depth, recent trade direction, and other microstructure signals to produce more accurate fair value estimates that better reflect instantaneous supply and demand dynamics.

Hasbrouck [23] introduces the concept of microprice, which weights the bid and ask prices by the inverse of their depths, providing a more accurate estimate of fair value in the presence of order book imbalance. The microprice is calculated as:

$$P_{micro} = \frac{P_{ask} \cdot V_{bid} + P_{bid} \cdot V_{ask}}{V_{bid} + V_{ask}} \tag{1}$$

where P_{bid} and P_{ask} are the best bid and ask prices, and V_{bid} and V_{ask} are the corresponding volumes.

The information content of trades and quotes has been extensively studied using the framework of Hasbrouck [22], who develops the information share metric to measure the contribution of different markets to price discovery. Gonzalo and Granger [21] propose an alternative measure based on the permanent-transitory decomposition of prices. Recent work by Brogaard et al. [16] shows that HFTs contribute disproportionately to price discovery, accounting for twice their share of volume in terms of information incorporation.

The role of latency in price discovery has gained attention recently. Hendershott et al. [26] demonstrate that algorithmic trading improves price discovery by incorporating information more quickly into prices. Conrad et al. [19] find that HFT activity is associated with improved price efficiency, particularly for small and mid-cap stocks. However, Biais et al. [14] warn that excessive speed competition can lead to adverse selection and reduced market quality.

2.3. Variance and Jump Modeling in High-Frequency Settings

The estimation of volatility from high-frequency data presents unique challenges due to market microstructure noise. The seminal work of Andersen et al. [7] introduces realized volatility as a consistent estimator of integrated variance when sampling at appropriate frequencies. For a day with N intraday returns r_i, realized variance is:

$$RV = \sum_{i=1}^{N} r_i^2 \tag{2}$$

However, at ultra-high frequencies, microstructure noise dominates the signal. Zhang et al. [31] proposes the two-scales realized volatility (TSRV) estimator that combines estimates at different sampling frequencies to mitigate noise. Barndorff-Nielsen et al. [9] develop the realized kernel estimator, which uses a kernel weighting function to optimally balance noise and discretization error:

$$RK = \gamma_0 + 2\sum_{h=1}^{H} k\left(\frac{h}{H+1}\right)\gamma_h \tag{3}$$

where γ_h is the h-th order autocovariance of returns and $k(\cdot)$ is a kernel function.

The multivariate case introduces additional complexity. Barndorff-Nielsen et al. [11] extend the kernel approach to estimate realized covariance matrices, crucial for understanding cross-asset dynamics in the presence of HFT. Hautsch and Huang [25] show that proper noise correction is essential for accurate correlation estimates, particularly when traders operate across multiple assets simultaneously.

Recent advances focus on separating continuous price variation from jumps. Aït-Sahalia and Yu [2] develop tests for common jumps across assets, finding that HFT activity can both trigger and dampen jump cascades depending on market conditions. Consequently, the detection of price jumps is critical for understanding extreme market events and the role of high-speed traders in their propagation or mitigation. Barndorff-Nielsen and Shephard [12] develop the bipower variation measure to separate continuous variation from jumps:

$$BV = \frac{\pi}{2}\sum_{i=2}^{N} |r_i||r_{i-1}| \tag{4}$$

The test statistic for jump detection is then:

$$J = \frac{RV - BV}{\sqrt{\theta \cdot TQ}} \tag{5}$$

where TQ is the tri-power quarticity for estimating the variance of the test statistic.

Lee and Mykland [28] propose an alternative approach using a threshold-based test that identifies individual jumps rather than testing for their presence over a day. Their method has been particularly useful for studying the intraday timing of jumps and their relationship to news arrivals. The ability to pinpoint jump timing with precision enables researchers to distinguish between news-driven price discontinuities and endogenous market dynamics generated by trading algorithms.

2.4. Market Quality and HFT Impact

The impact of HFT on market quality remains contentious. Hendershott et al. [26] find that algorithmic trading narrows spreads and improves liquidity, particularly for large-cap stocks. Hasbrouck and Saar [24] shows that HFTs enhance price efficiency by quickly incorporating information into prices. However, Zhang [30] argues that HFT can increase volatility and generate excess price movements unrelated to fundamentals.

The liquidity provision role of HFTs has been extensively studied. Menkveld [29] documents that a single HFT firm acts as a modern market maker, providing liquidity and earning profits from the spread. Brogaard et al. [15] find that HFTs supply liquidity when it is expensive and demand it when it is cheap, suggesting they improve market efficiency. However, Anand and Venkataraman [6] show that HFT liquidity provision is fragile and can evaporate during stress periods.

The welfare implications of HFT are analyzed by Biais et al. [14], who develop a model where speed competition can lead to excessive investment in technology and adverse selection for slower traders. Budish et al. [18] estimate the social waste from the arms race at $5 billion annually in global equity markets. Aquilina et al. [8] provide empirical evidence that speed advantages are valuable but create negative externalities for other market participants.

Regulatory responses to HFT have varied across jurisdictions. Friederich and Payne [20] evaluate the French financial transaction tax, finding it reduced HFT activity but also harmed market quality. Conrad et al. [19] examine the SEC's market access rule, showing it improved risk controls without significantly impacting HFT strategies. Our analysis contributes to this debate by providing granular evidence on how different speed tiers affect market outcomes, informing more targeted regulatory approaches.

3. Methodology

This section presents our comprehensive methodology for classifying market participants and analyzing their impact on market quality. We detail our unique data construction process, the novel classification algorithm based on reaction times, and the econometric approaches used to identify potential effects. Our methodological framework combines cutting-edge data infrastructure, relevant metrics computation, and natural experiments to provide robust evidence on the stratification of modern electronic markets.

3.1. Data Construction

3.1.1. Deutsche Börse Trading Infrastructure

Deutsche Börse offers a robust technological infrastructure to support a wide range of market participants, from traditional investors to cutting-edge HFT and UFT participants. The T7 trading platform is designed with an emphasis on ultra-low-latency performance, providing high-speed data feeds and connectivity options. This sophisticated infrastructure enables nanosecond-precision timestamp collection at multiple points throughout the order processing pipeline, creating unprecedented visibility into market dynamics. Additionally, the A7 Analytics Platform allow for rapid C++ algorithm development and deployment by providing a dedicated API enabling to obtain high-precision ready data.

The Xetra and Eurex platforms operate under Deutsche Börse's T7 electronic trading system, for equities and derivative products respectively, ensuring high-speed execution of trades. They are responsible for a large portion of stock market liquidity in Germany and across Europe. The unified T7 architecture across both platforms facilitates cross-asset strategies and enables consistent latency measurement across different instrument types.

3.1.2. High-Precision Timestamp Data

The High-Precision Timestamp (HPT, containing only trade events) and HPT All (HPTA, containing all market events: trades and orderbook updates) files are integral to facilitating ultra-fast trading activities, providing granular transaction data that allows market participants to refine their trading strategies. The data feed captures every event received in the system, from order entry to execution, following the strict price-time priority mechanism in the T7 matching engine. We implemented a Level 1 order book reconstruction on the A7 Analytic Platform. At every market event, the algorithm updates the best bid and best ask price and quantity, producing a time-aligned stream of top-of-book states keyed to HPT timestamps.

The data captures multiple critical timestamps (figure 1) with nanosecond precision throughout the order lifecycle, enabling thin granularity in latency

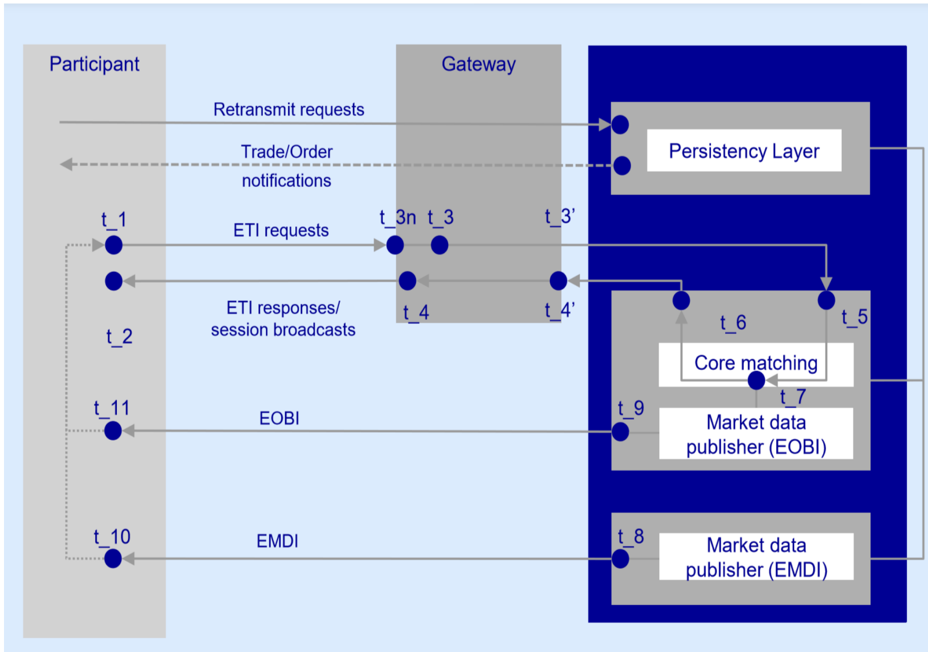


Figure 1: Timestamps collection workflow in Deutsche Börse T7

measurement. RequestTime (t_{3a}) marks when requests arrive at the Access Network layer, representing the entry point into the exchange infrastructure. RequestTime (t_{3n}) indicates transmission to the T7 matching engine, capturing internal network latency. AggressorTime (t_5) records the processing of aggressive orders that trigger executions, while ExecID (t_7) provides the exact execution timestamp when trades occur. TransactTime (t_9) documents the transaction recording in the exchange's systems, and EOBICaptTime (t_{9d}) marks public dissemination through market data feeds. This comprehensive timestamp cascade enables precise measurement of latencies at each processing stage, revealing how different participant types exploit specific segments of the trading infrastructure.

3.1.3. Sample Selection and Data Coverage

Our analysis covers the period from January 1, 2025, to August 31, 2025, focusing exclusively on the Euro STOXX 50 Index Futures (FESX) traded on Eurex. We select the most liquid FESX security for each trading day, typically the front-month contract until it approaches expiration, at which point liquidity shifts to the next maturity. This single-instrument focus allows us to examine market microstructure dynamics without confounding effects from cross-instrument heterogeneity.

For analyzing participant reactions to trades, we include all FESX futures contracts to capture the full scope of market responses across the maturity spectrum. This method ensures we identify all reactive trading behavior, whether participants respond in the same contract or arbitrage across different maturities.

To align the data sources, we proceeded day by day. First, ran a selection algorithm to obtain the most traded FESX instrument for that date. For this contract, we ingest the corresponding HPT All file (our reaction events) together with the Level-1 order-book reconstruction data. In parallel, we load the HPT files for all FESX contracts for the same date to capture trigger trades. Then, following established practice in the HFT literature, the final dataset is filtered to cover only the regular trading session from 9:30 to 17:00.

3.2. Classification of Market Events

3.2.1. Reaction Time Measurement

Our classification methodology centers on measuring reaction latency, measured as the time between a trigger trade happen on a FESX security and the participants responses (trades or order book updates on the most traded security). For each (reaction) event i submitted at time t_i, we identify the most recent (trigger) event j occurring at time $t_j \leq t_i - \epsilon$ (with $\epsilon = 10ns$). This approach captures the fundamental speed advantage that defines modern high-frequency trading, where success depends on reacting to market changes faster than competitors. The reaction time is:

$$\Delta t_{i,j} = t_i - t_j \tag{6}$$

where timestamps are measured in nanoseconds.

To accurately measure this latency, we match all events of the most liquid FESX security with a corresponding triggering event (hypothesizing in fact that they are all reacting to events), according to the following methodology. As shown on figure 2, the HPT files provide us with the t_{3a} (receiving order time) and t_{9d} (time at which the information update is dispatched to all market participants). Since all market participants use co-location and the cable length is the same for every participant, the minimum latency between t_{9d} and t_{3a} is fixed and known from Deutsche Bôrse. Our strategy is then to

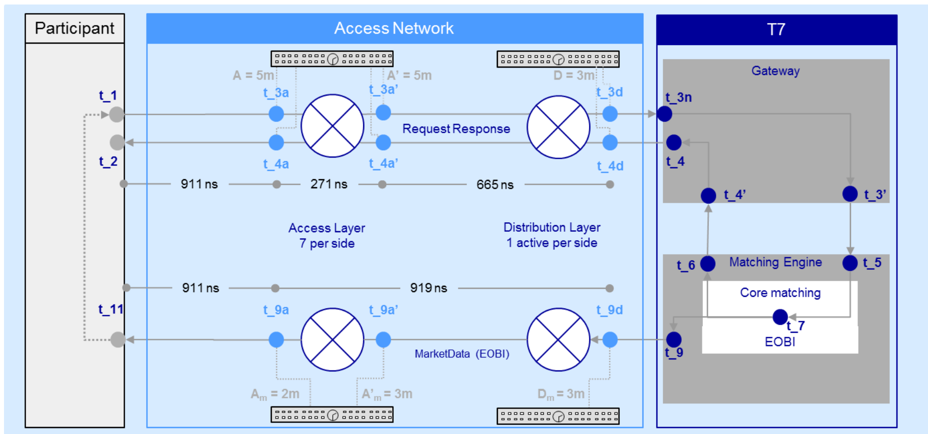


Figure 2: Deutsche Börse T7 matching engine

find, for all reaction event i the latest trigger event j with $t_{9d}^{(j)} <= t_{3a}^{(i)} - \delta - \epsilon$. The latency is then measured as:

$$\Delta t_{i,j} = t_{3a}^{(i)} - t_{9d}^{(j)} \tag{7}$$

3.2.2. Classification Thresholds

Based on Deutsche Börse empirical analysis and technological capabilities, we establish the following classification thresholds. These thresholds reflect natural discontinuities in the reaction time distribution corresponding to different hardware and software architectures. The classification boundaries align with known technological constraints, where sub-microsecond reactions require FPGA hardware, microsecond responses utilize optimized software, and slower reactions indicate human or less sophisticated algorithmic trading:

$$\text{Trader Type}_i = \begin{cases} \text{Noise} & \text{if } \Delta t_{i,j} < 0 \\ \text{UFT} & \text{if } 0 \leq \Delta t_{i,j} < 1,000 \text{ ns} \\ \text{HFT} & \text{if } 1,000 \leq \Delta t_{i,j} < 10,000 \text{ ns} \\ \text{Non-HFT} & \text{if } \Delta t_{i,j} \geq 10,000 \text{ ns} \end{cases} \tag{8}$$

With a trading day spanning 2.7×10^{10} microseconds (7.5 hours) and approximately 1 million daily events, the event probability per microsecond equals 3.7×10^{-5}. This translates to merely a 0.04% probability of random coincidence within a 10-microsecond window, confirming that observed sub-microsecond reactions represent genuine technological capabilities rather than statistical artifacts.

3.3. Market Regimes

To better contextualize the activity of ultra-fast and high-frequency traders, we compute several market regime indicators. Specifically, we incorporate multiple realized volatility estimators and a jump detection framework to characterize prevailing market conditions. This allows us to distinguish between normal and stressed environments, and to analyze how participation and informational quality vary across different volatility levels, the presence of jumps, and time-of-day segments.

3.3.1. Continuous Fair-Value Estimation

A continuous proxy for fair value must live inside the best quotes, respond to order-imbalance, and avoid spurious jumps from tick discretization. The naïve mid-price

$$P_t^{\mathrm{mid}} \;=\; \frac{P_t^{\mathrm{ask}} + P_t^{\mathrm{bid}}}{2} \tag{9}$$

is unbiased in expectation but is notoriously noisy at high frequency due to the discrete grid of tick sizes.

A common refinement is the *volume-weighted microprice* at the top of the book,

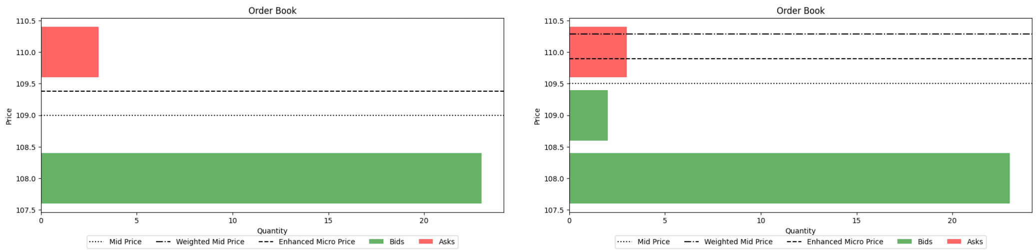
$$P_t^{\mathrm{vw}} \;=\; \frac{P_t^{\mathrm{ask}}\, V_t^{\mathrm{bid}} + P_t^{\mathrm{bid}}\, V_t^{\mathrm{ask}}}{V_t^{\mathrm{bid}} + V_t^{\mathrm{ask}}} \tag{10}$$

which tilts the mid toward the side with more resting depth. While intuitive, this estimator can exhibit undesirable sensitivity to microscopic changes at the best level: a small order improving the best price by one tick can wipe out large depth at the second level, causing P_t^{vw} to jump in a direction that is counterintuitive relative to the added same-side interest.

To reduce tick-induced volatility and limit sensitivity to thin best levels, we adopt a *reverse-weighted mid price*:

$$P_t^{\mathrm{micro}} \;=\; P_t^{\mathrm{mid}} \;+\; \frac{\Delta}{2} \cdot OIR_t, \qquad OIR_t \;=\; \frac{V_t^{\mathrm{bid}} - V_t^{\mathrm{ask}}}{V_t^{\mathrm{bid}} + V_t^{\mathrm{ask}}}, \tag{11}$$

where Δ is the tick size. This choice keeps P_t^{micro} within $[P_t^{\mathrm{bid}}, P_t^{\mathrm{ask}}]$, moves it monotonically with queue order imbalance ratio OIR_t, and caps any single update's adjustment at half a tick. In practice, (11) preserves the desirable directionality of the volume-weighted microprice but removes large, mechanically-driven swings when best-level depth flips or when a tiny improvement order appears inside the spread.



(a) Order book state at time t

(b) Order book state at time $t + 1$

Figure 3: Continuous fair-value estimation scenario.

Consider the two books in Fig. 3 (tick size $\Delta = 1$). Between t and $t+1$ a small buy improves the best bid by one tick while substantial depth remains just below. The classical volume-weighted microprice can jump by more than a tick and even lie outside the spread, as it scales imbalance by the prevailing spread. By contrast, our half-tick imbalance estimator moves smoothly, remains strictly within $[P^{\text{bid}}, P^{\text{ask}}]$, and increases monotonically with the buy-side imbalance. In this "thin inside quote" scenario, the enhanced microprice tracks the one-tick improvement without spurious overreaction, providing a stable, spread-bounded proxy for continuous fair value at nanosecond horizons.

3.3.2. Volatility Estimation

We take the enhanced microprice as the fair-value proxy and compute returns $r_i = \log \mu_{t_i} - \log \mu_{t_{i-1}}$ on raw high-frequency as well as on resampled grids (1 s, 5 s, 15 s). All volatility estimators below are produced at the daily level.

Realized Variance (RV). For a day d with N_d returns,

$$\mathrm{RV}_d = \sum_{i=1}^{N_d} r_i^2. \tag{12}$$

Median Realized Volatility [7] (MedRV, noise-robust). Let $N = N_d$. With the rolling median-of-three operator,

$$\mathrm{MedRV}_d = \frac{\pi}{6 - 4\sqrt{3} + \pi} \cdot \frac{N}{N-2} \sum_{i=3}^{N} \mathrm{median}\big(|r_i|, |r_{i-1}|, |r_{i-2}|\big)^2. \tag{13}$$

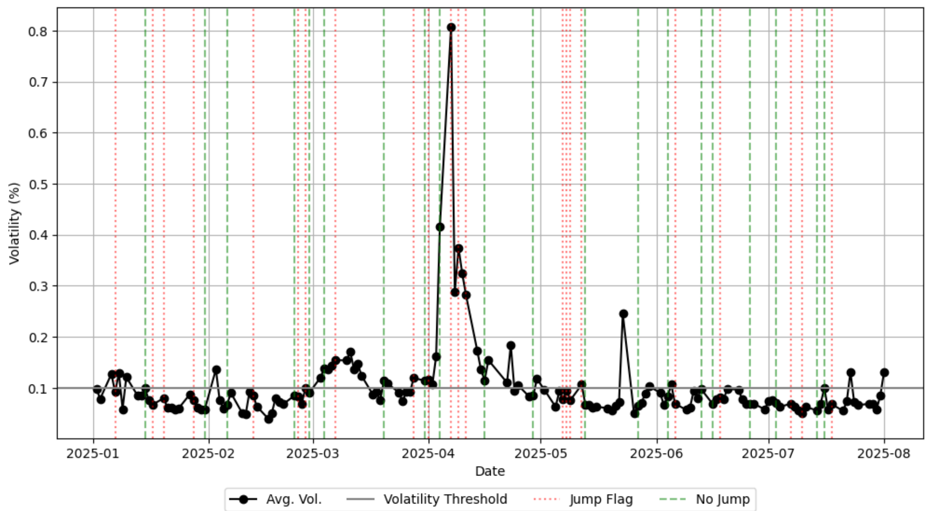


Figure 4: Volatility and Jump Market Regimes

Truncated Power Variation [1] (jump robust). For order $p > 0$, truncation level $\alpha > 0$, and $\omega \in (0, 1/2)$,

$$\mathrm{TPV}_d^{(p)} = \frac{\Delta^{1-\frac{p}{2}}}{m_p} \sum_{i=1}^{N} |r_i|^p \,\mathbf{1}\{|r_i| \le \alpha\,\Delta^{\omega}\}\,, \tag{14}$$

where Δ is the sampling interval, $m_p = \mathbb{E}(|Z|^p)$ for $Z \sim \mathcal{N}(0, 1)$. We use chose this estimator ($p = 2$) as our noise-robust estimate of the continuous variation when constructing jump measures.

Realized Kernel [10] (noise- and autocorrelation-robust). With bandwidth H and Parzen kernel $k(\cdot)$,

$$\mathrm{RK}_d = \sum_{h=-H}^{H} k\left(\frac{h}{H+1}\right) \hat{\gamma}_h, \qquad \hat{\gamma}_h = \sum_{i=|h|+1}^{N} r_i\, r_{i-|h|}. \tag{15}$$

We select the optimal bandwidth H based on the process described in their paper.

We averaged the 16 volatility estimators to obtain a single value (cf. figure 4), and arbitrarily selected a threshold of 10% volatility to flag high- and low-volatility settings.

3.3.3. Jump Detection

We identify jump days using the test from [1], based on the aforementioned truncated power variations technique. For $p > 3$ and integer $k \geq 2$, define

$$\widehat{S}_d(p,k) = \frac{\sum_{j=1}^{\lfloor N/k \rfloor} \left| \sum_{\ell=1}^{k} r_{(j-1)k+\ell} \right|^p}{\sum_{i=1}^{N} |r_i|^p}. \tag{16}$$

Under no jumps, $\widehat{S}_d(p,k) \xrightarrow{p} k^{\frac{p}{2}-1}$; when jumps are present, $\widehat{S}_d(p,k) \xrightarrow{p} 1$. We standardize $\widehat{S}_d(p,k)$ using the variance formula with either truncated power or multipower estimates of the continuous part and obtain an asymptotic $N(0,1)$ statistic; days are flagged as Jump when the null of continuity is rejected at the chosen level. We conducted the tests under both the jump and no-jump hypotheses, flagging days accordingly only when both tests yielded consistent results and discarding cases of disagreement, leading to the results shown on figure 4

3.3.4. Time of Day

Lastly, to account for strong intraday seasonality, we partition the trading session into three segments: Morning (09:30–12:00), Mid-day (12:00–15:00), and Close (15:00–17:00).

3.4. Market Conditions Measures

To analyze the interaction between UFTs/HFTs and market quality, we construct a set of liquidity and price efficiency measures. All metrics are computed at high frequency (raw or 1-second intervals), then aggregated into 15-minute non-overlapping buckets to provide consistent intraday panels of market conditions.

3.4.1. Liquidity Measures

Order Imbalance Ratio (OIR): Order imbalance captures the relative pressure on the bid and ask sides of the order book. For each order book snapshot, we compute OIR_t as described in subsection 3.3. To quantify intraday fluctuations, we calculate the standard deviation of OIR_t over each 1-second window and then average these values into 15-minute intervals:

$$\sigma_{OIR} = \sqrt{\frac{1}{N} \sum (OIR_t - \mu_{OIR})^2}. \tag{17}$$

HFT Illiquidity: We adapt the Amihud illiquidity measure [4] to the HFT context. For each 1-second return r_t and unsigned trade volume V_t, we define the 15-minutes measurement interval as:

$$ILLIQ = \frac{1}{N} \sum \frac{|r_t|}{V_t}. \tag{18}$$

3.4.2. Price Efficiency Measures

While closely related to liquidity, price efficiency measures provide insight into how quickly and accurately information is incorporated into prices.

Autocorrelation: We compute the first-order autocorrelation ($q = 1$) of log returns from the 1-second resampled reverse-weighted microprice series.

Variance Ratio (VR): To test for random walk behavior, we apply a variance-ratio test as per [5], using horizons $q = \{2, 5, 10, 20\}$. Both the VR value and its associated statistics are retained, based on the 1-second enhanced microprice log returns. This allows us to detect deviations from martingale pricing at multiple horizons.

3.5. Impact Analysis Framework

To link HFT activity with market quality, we compute measures of trading intensity and predictive power at the participant level. All metrics are also aggregated into 15-minute windows to align with market condition measures.

3.5.1. HFT Activity Measures

Participation: For each participant i, we compute the relative participation rate as the share of events (orders or trades) classified under the latency thresholds defined earlier in this paper. Participation shares are standardized by window length:

$$Participation_{i,t} = \frac{share_i}{900},$$

where 900 is the number of seconds in a 15-minute interval.

Signal-to-Noise Ratio (SNR): Beyond simple participation, we assess the *quality* of trading activity by measuring participants' ability to predict short-term price moves. Following the logic of mark-out returns, we define the Signal-to-Noise Ratio (SNR) for participant i in bucket t as:

$$SNR_{i,t} = \frac{\sum_{j=1}^{N} \text{Signal}_{i,j}}{\sum_{j=1}^{N} \text{Noise}_{i,j}}, \tag{19}$$

where each event j is classified as *signal* or *noise* depending on whether the reaction price moves toward or away from the 15-second mark-out price. Specifically:

$$\text{Signal}_{i,j} = \begin{cases} \frac{|R_j - T_j|}{R_j}, & \text{if } T_j \leq R_j \leq M_j \text{ or } T_j \geq R_j \geq M_j, \\ \frac{|M_j - T_j|}{R_j}, & \text{if } T_j \leq M_j \leq R_j \text{ or } R_j \leq M_j \leq T_j, \\ 0, & \text{otherwise,} \end{cases} \tag{20}$$

$$\text{Noise}_{i,j} = \begin{cases} \frac{|R_j - T_j|}{R_j}, & \text{if } R_j \leq T_j \leq M_j \text{ or } R_j \geq T_j \geq M_j, \\ \frac{|R_j - M_j|}{R_j}, & \text{if } T_j \leq M_j \leq R_j \text{ or } R_j \leq M_j \leq T_j, \\ 0, & \text{otherwise.} \end{cases} \tag{21}$$

with T_j the trigger price (the microprice when at the trigger trade time), R_j the reaction price (the microprice at the reaction time t), and M_j the mark-out price at t+15 seconds. An $SNR_{i,t} > 1$ indicates predictive trading (signal dominates), while $SNR_{i,t} < 1$ reflects adverse or noise-driven trading. Unlike traditional information share metric or aggregate price impact measures, our novel mark-out signal-to-noise ratio directly quantifies each participant's short-term predictive ability at the event level, enabling a granular assessment of informational efficiency in nanosecond trading environments.

3.5.2. Quantile Analysis

To assess heterogeneity across activity levels, we partition the distribution of participation into quantiles. In particular, we compare the top (Q5) and bottom (Q1) quantiles of participation rates to evaluate how market quality and predictive power differ between low-activity and high-activity participants. This framework highlights whether outsized participation at high speed is associated with stabilizing or destabilizing effects on market dynamics.

4. Empirical Results

This section presents our empirical findings on the classification of market participants and their impact on market quality. We present first the analysis of trading patterns, and finish with potential evidences of HFT on financial markets.

4.1. Participant Classification Evaluation

4.1.1. Distribution of Reaction Times

We classify events as triggered by UFTs, HFTs, or other participants based on measured reaction times. Figure 5 shows the distribution of reaction latencies across all 15-minute windows over the full sample period, revealing a clear separation between technological classes.

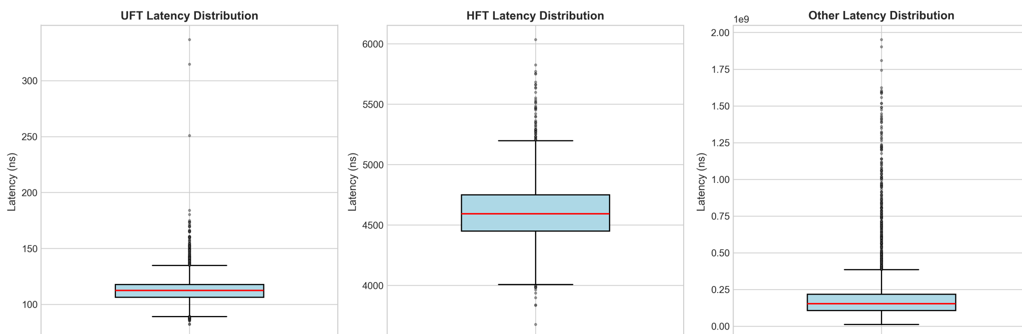


Figure 5: Distribution of Reaction Times by Trader Type

Ultra-fast traders (UFTs) exhibit extreme latency precision, with values ranging from $0.08\,\mu s$ to $0.33\,\mu s$, and a dense concentration between $0.11\,\mu s$ and $0.12\,\mu s$. This sharp clustering suggests intense speed competition at the nanosecond level, where even a $0.01\,\mu s$ advantage can determine queue priority.

High-frequency traders (HFTs), though slower in absolute terms, display a similarly competitive band, with reaction times ranging from $3.7\,\mu s$ to $6.0\,\mu s$ and a concentration between $4.4\,\mu s$ and $4.8\,\mu s$. This implies that shaving off even $1\,\mu s$ of internal latency may shift a participant from the bottom to the top of the HFT speed curve, with potentially significant implications for execution performance and profitability.

Other participants operate on a much slower scale, with latencies typically spanning from a few milliseconds to several hundred. The first and third quartiles lie at 107 ms and 218 ms, respectively. The multimodal nature of the latency distribution confirms that speed tiers correspond to structural differences in technology and execution infrastructure, rather than arbitrary thresholds.

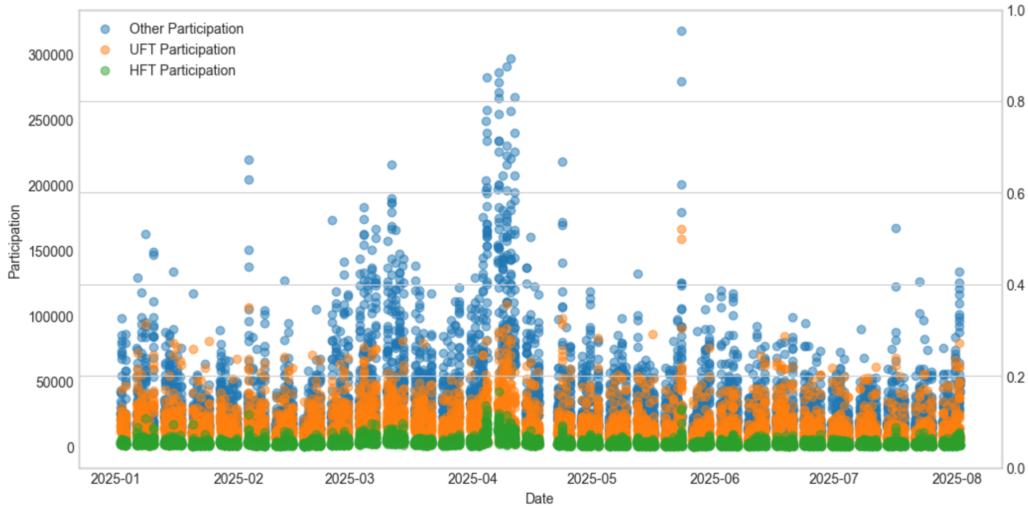


Figure 6: Participation Share per Participant for the Whole Period

4.1.2. Trading Activity by Participant Type

Participation dynamics, like latency distributions, reveal distinct behavioral patterns across trader classes. Figure 6 plots the daily share of total participation for HFTs, UFTs, and other participants from January to August 2025. High-frequency traders maintain a remarkably stable share over time, with only modest fluctuations, except for a temporary spike in April 2025. This period coincides with a surge in overall trading activity and volatility triggered by renewed uncertainty surrounding the U.S. trade policy, which likely caused increased responsiveness from algorithmic traders.

UFTs, while generally accounting for a larger share than HFTs (roughly six times greater on average) showcase greater variability. Their participation closely co-moves with that of conventional participants, suggesting that UFTs are primarily reactive: they capitalize on flow initiated by non-HFT agents. This pattern underscores the latency-sensitive nature of UFT strategies-mainly market making-which appear to monitor and respond to slower market participants rather than proactively driving liquidity.

Notably, the participation share of conventional traders exhibits the widest range and highest volatility across the period. Peaks in their share often coincide with spikes in absolute participation levels, reinforcing the idea that market regimes with increased retail or institutional flow. Overall, these dynamics support a layered structure of market interaction, where UFTs respond to conventional flow, HFTs arbitrage across liquidity fragments, and other traders inject exogenous demand into the system.

To evaluate how different types of market participants impact financial market quality, we compare market metrics across two extreme participation levels—quintile 1 (Q1, representing the lowest 20% of activity) and quintile 5 (Q5, representing the highest 20%)—per trader type. For each group, we compute average values of key market quality indicators, namely the standard deviation of the order imbalance ratio (OIR Std.), Amihud illiquidity, autocorrelation, and variance ratio statistics (VR_q) for $q \in \{2, 5, 10, 20\}$. These metrics, presented in Table 1, allow us to assess the informational efficiency and liquidity of the market under varying participation intensities.

Table 1: Market quality metrics by participation quintiles (Q1: low, Q5: high) for each participant type.

Metric	**UFT**		**HFT**		**Other**	
	Q1	Q5	Q1	Q5	Q1	Q5
Participation	5,562	40,060	948	7,408	14,423	96,875
OIR Std.	0.1108	0.2367	0.1009	0.2658	0.0887	0.2862
Amihud Illiq. ($\cdot 10^{-9}$)	1.09	1.43	1.02	1.63	0.093	1.76
Autocorrelation	0.0116	0.0210	0.0138	0.0208	0.0253	0.0159
VR_2	0.5078	0.5090	0.5089	0.5095	0.5143	0.5072
VR_5	0.2038	0.2052	0.2040	0.2053	0.2064	0.2040
VR_{10}	0.1023	0.1024	0.1026	0.1024	0.1038	0.1016
VR_{20}	0.0518	0.0518	0.0520	0.0517	0.0526	0.0514

The results reveal several important patterns. First, increasing non-HFT participation is associated with a sharper rise in order imbalance volatility and Amihud illiquidity, indicating that higher traditional algorithmic trader activity may reduce market depth and increase sensitivity to trades. However, a contrasting trend appears in the autocorrelation of returns: while increased UFT and HFT activity is associated with higher autocorrelation, the opposite is true for other participants, whose increased presence reduces autocorrelation. This suggests that high-speed participants may contribute to short-term predictability in prices—potentially due to latency arbitrage or better price forecasting—whereas slower participants add more noise and long-memory effects to price dynamics.

Regarding efficiency, variance ratio metrics provide further insight. Across all VR horizons ($q = 2, 5, 10, 20$), we observe negligible changes for UFTs and HFTs, indicating that their activity does not significantly affect longer-horizon return variance scaling. In contrast, higher activity by non-HFT participants leads to a consistent drop in VR values—especially VR_2—suggesting that these participants may degrade price efficiency through noisier or less informed trading.

In sum, while UFT and HFT activity appears to preserve long-horizon efficiency and increase short-term return autocorrelation—possibly reflecting more informed or anticipatory trading—greater non-HFT participation is linked to deteriorating market conditions. Specifically, higher non-HFT activity coincides with increased illiquidity, more volatile order imbalances, reduced short-term autocorrelation, and lower variance ratios, indicating noisier price dynamics and impaired informational efficiency.

4.2. Short-Horizon Price Informativeness

Price discovery, the process by which markets incorporate information into asset prices, represents a fundamental function of financial markets. Understanding which participants contribute most to this process has important implications for market design and regulation. We measure contribution to price discovery using the Signal-to-Noise Ratio described earlier in section 3.5, indicating the short-term informativeness of trading actions. Unlike traditional information share metrics that rely on cointegration and price series decomposition, the SNR focuses on whether a participant's activity pushes the price toward or away from its short-term future value, thereby offering a microstructural view of predictive trading.

Figure 7 shows the distribution of SNR values across the three trader categories. The results reveal that HFTs exhibit the highest median and a notably wide dispersion in SNR, suggesting that they contribute most consistently and dynamically to the price discovery process. This aligns with the idea that HFTs—while not operating at the ultra-fast latencies of UFTs—deploy fast, information-rich strategies based on predictive modeling and machine learning. Their actions appear to encode forward-looking information over short horizons (e.g., 1 to 15 seconds), leading to a stronger informational footprint.

UFTs, on the other hand, display lower median SNRs with a moderately wide distribution. Although they react the fastest, their behavior seems

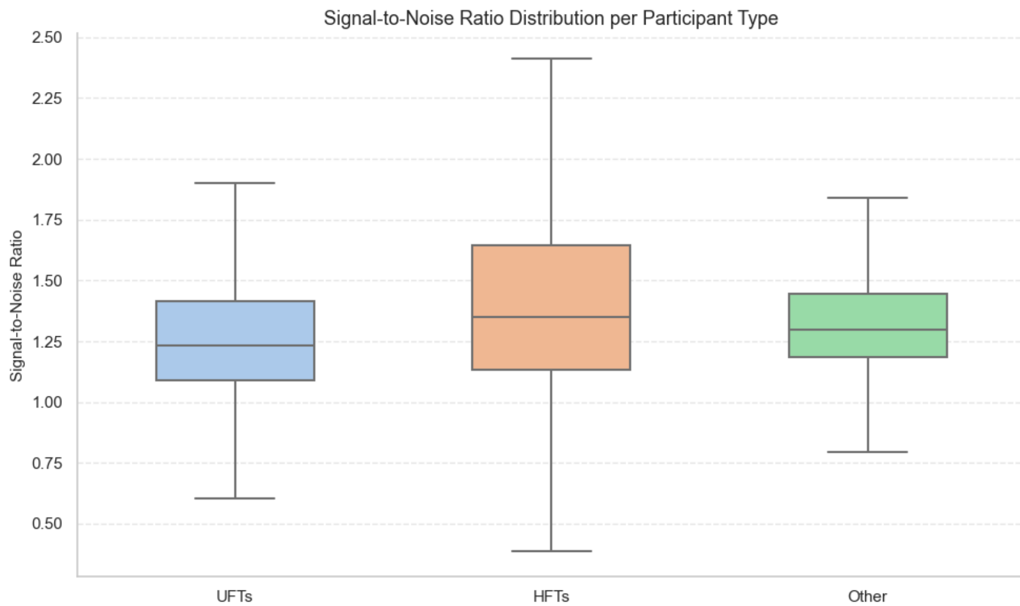


Figure 7: Signal-to-Noise Ratio distribution per Participant

largely mechanical or anticipatory, exploiting latency advantages rather than embedding new information into prices. This results in weaker alignment with future price direction, confirming their limited role in information incorporation.

Interestingly, non-HFT participants show a median SNR similar to UFTs but with a tighter interquartile range and lower upper tail. This pattern suggests more homogeneous trading behavior, potentially driven by execution algorithms, institutional order splitting, or slower, non-reactive strategies. While these participants may reflect long-term fundamental views or macroeconomic positioning, their short-term informativeness appears weaker and less variable compared to HFTs.

These findings highlight a clear stratification of informational roles: HFTs appear to drive short-term informational efficiency, UFTs optimize for latency without significant price informativeness, and other participants contribute heterogeneously, potentially incorporating slower-moving or exogenous signals.

5. Discussion

This section interprets our empirical findings, discusses their implications for market design and regulation, and acknowledges the limitations of our analysis. We explore the mechanisms driving our results and their broader significance for understanding modern financial markets.

5.1. Summary of Key Findings

Our study provides new insights into the role of latency-sensitive traders in shaping short-term market dynamics. We classify traders into Ultra-Fast Traders (UFTs), High-Frequency Traders (HFTs), and other participants based on nanosecond-level reaction times, and evaluate their participation rates, signal-to-noise ratio (SNR), and impact on key market quality metrics. The analysis reveals that HFTs exhibit the highest SNR and contribute most consistently to short-term price discovery. UFTs, while extremely fast, display lower SNR values, suggesting their actions are largely reactive. Other participants, despite slower reactions, show heterogeneous contributions, occasionally incorporating valuable exogenous information.

5.2. Information and Price Discovery

The fact that HFTs contribute disproportionately to price discovery raises fundamental questions about the nature of information in high-frequency markets. The concentration of price discovery among the fastest and most technologically sophisticated traders suggests that speed and predictive modeling have become major sources of informational advantage. The landscape of competition in market microstructure has shifted from traditional fundamental analysis to dominance by those with superior data processing, model sophistication, innovative data sources, and infrastructure.

Several mechanisms may explain the superior price discovery contribution of these traders. First, speed effectively converts public information into private information. HFTs, by processing market data within microseconds, can act on imbalances, order flow patterns, and latent signals before these are reflected in prices. Second, their cross-market presence enables them to arbitrage inefficiencies across venues and instruments, thereby enforcing price consistency. Finally, sophisticated statistical and machine learning models allow HFTs to forecast short-term price movements with high precision, making their activity particularly informative on horizons of 1 to 15 seconds.

The role of UFTs in this context appears more nuanced. While they are the fastest actors in the market, their strategies are primarily reactive and rule-based. Our results show that UFTs exhibit the lowest SNR values with narrow dispersion, indicating that although they succeed in securing queue priority and capturing microstructure-driven profits, their actions often lack informational content. Rather than generating or incorporating new information, UFTs primarily optimize latency to exploit fleeting arbitrage opportunities, particularly around auctions and large incoming orders. This limits their contribution to true price discovery despite their dominant speed.

Non-HFT participants display heterogeneous SNR profiles. Despite longer reaction times, they may incorporate broader information sets, including macroeconomic news or portfolio-level rebalancing, especially during low-volatility or off-peak hours. Their influence on short-term price efficiency appears limited, but they may contribute to longer-horizon informational dynamics not captured by our SNR measure.

The welfare implications of this stratified price discovery process are complex. On the positive side, the dominance of informed HFT activity accelerates the incorporation of available information into prices, improving short-term market efficiency and reducing adverse selection risks for institutional investors. The rapid updating of quotes and predictive trading behavior enhances allocative efficiency and supports more accurate valuation mechanisms across asset classes. However, these benefits come with costs. The intense competition for speed has driven an arms race that channels vast resources into infrastructure, colocation, and custom hardware development, with questionable marginal social returns. The dominance of speed-based strategies may crowd out fundamental research, contributing to a trading ecosystem where technological prowess outweighs economic insight. Moreover, the extreme speed advantages held by a few firms may distort fairness and accessibility, creating a bifurcated market structure.

Importantly, the marginal contribution to price informativeness decreases as latency shrinks further. This finding supports regulatory perspectives that seek to mitigate the negative externalities of extreme latency competition while preserving the benefits of efficient information processing. A nuanced policy approach should thus focus on calibrating rather than suppressing latency-sensitive activity.

5.3. Strategic Motivations of UFTs and HFTs

5.3.1. Strategic Differences Between UFTs and HFTs

Our empirical findings highlight clear strategic divergences between UFTs and HFTs, rooted in their differing latency capabilities and trading objectives. UFTs specialize in ultra-low latency arbitrage strategies, including cross-venue arbitrage and index arbitrage. Their approach is highly opportunistic and concentrated around events that create momentary price discrepancies. These strategies are capital-light, typically involve minimal inventory risk, and rely on speed to capture value before it dissipates. The high success rate of UFTs in executing aggressive orders underlines their ability to seize fleeting liquidity advantages, especially during auctions and news-driven volatility spikes.

In contrast, HFTs balance speed with sophistication. Operating within a slightly slower latency band, HFTs implement strategies that combine passive liquidity provision with active inventory management. Their presence throughout the trading session suggests a focus on market making and directional positioning over short horizons. They continuously update quotes, manage order queues, and hedge exposures, often relying on advanced machine learning algorithms to forecast price paths and adapt dynamically. Their ability to hold and manage inventory underscores a higher level of engagement with market fundamentals and risk.

This functional differentiation points to a form of implicit division of labor in modern electronic markets. UFTs dominate in latency-critical environments, enforcing price uniformity and capturing time-sensitive profits, while HFTs sustain market liquidity and informational efficiency over slightly longer horizons. Rather than competing head-to-head, these actors often complement each other in shaping intraday price dynamics.

5.3.2. Economic Rents from Speed

The economic value extracted from speed differentials underscores the competitive intensity and capital requirements of high-speed trading. We propose now to discuss the structural conditions under which such rents arise.

The competition among UFTs and HFTs is fierce. Even microsecond-level disadvantages can render a strategy unprofitable. Maintaining competitiveness in this domain requires significant investment in infrastructure, proprietary hardware, low-latency networks, as well as human capital—including traders, researchers, and engineers. Firms operating in this space must invest continuously to refine execution logic, risk models, and adaptive strategies. The barrier to entry is not merely technical; it is financial and organizational.

Our analysis of tick size events confirms that ultra-fast participants can systematically extract value from transient inefficiencies. Their ability to reposition quickly in order queues, capture liquidity, and avoid adverse selection translates into measurable profitability. However, these advantages are fragile, transient, and contingent on continuous technological superiority.

5.4. Regulatory Considerations

Regulators across jurisdictions have adopted different approaches to addressing high-speed trading. In Europe, MiFID II has imposed a wide range of behavioral requirements, including order-to-trade ratio limits, market-making obligations, algorithm testing protocols, and harmonized tick sizes. In contrast, the U.S. framework under Reg NMS emphasizes market structure. By enforcing best execution across venues, regulating access fees, and restricting sub-penny pricing, the U.S. system seeks to maintain competitive and transparent markets. Our European data suggests that structural interventions—particularly tick size harmonization—may be more effective than behavioral restrictions in preserving market quality while avoiding the unintended suppression of beneficial competition.

Our evidence thus supports a more balanced regulatory philosophy. Rather than targeting specific participant behaviors, regulators should shape incentives through carefully designed structural rules. Tick sizes, resting times, and access fees are more likely to affect market-wide outcomes than prescriptive rules about participant conduct.

Building on these insights, we propose a tiered regulatory framework tailored to the latency capabilities and market roles of different participants. For ultra-fast traders operating below the 1-microsecond threshold, the primary focus should be on systemic risk and market integrity. These firms should be subject to enhanced transparency requirements, including real-time reporting and participation in risk mutualization mechanisms.

HFTs operating between 1 and 10 microseconds should be regulated in line with their role as liquidity providers. Market-making obligations, minimum resting times, and capital adequacy requirements should reflect their continuous exposure to inventory risk and their importance for maintaining orderly markets. During stress periods, incentives and penalties tied to liquidity provision would encourage stability.

5.5. Limitations

Despite the high granularity of our dataset, several limitations warrant caution. Our analysis is confined to Eurex, which, while one of the most liquid derivatives markets globally, does not capture the full scope of cross-venue strategies employed by high-speed traders. The inability to track participants across days due to anonymized IDs limits our capacity to observe strategic learning or long-term positioning. We also focus on the most liquid futures contracts, raising questions about the generalizability of our findings to smaller or less liquid markets.

Classification uncertainty presents another limitation. Although our latency-based methodology is grounded in microstructural logic, the precise mapping of reaction times to firm identities is not available. This introduces potential noise in our categorizations, particularly for participants operating near classification thresholds.

A broader concern relates to the lack of a natural control group. As noted in preliminary discussions, comparing markets with and without HFT presence is problematic because such differences are endogenous. Market features like liquidity, regulation, and investor composition all co-evolve with HFT activity, making causal inference difficult. Our use of tick size changes as quasi-natural experiments addresses this to some extent but does not fully resolve the equilibrium selection problem.

5.5.1. Welfare Analysis

Our empirical metrics speak primarily to short-term market quality and informational efficiency. However, a complete welfare analysis must also consider unmeasured costs. The infrastructure arms race entails substantial fixed costs with limited spillovers outside finance. The dominance of speed-centric strategies may disincentivize longer-term research or investment approaches. Systemic risks may arise if correlated latency strategies amplify shocks. And most critically, perceptions of unfairness could deter participation by retail or institutional actors less equipped to compete in the latency domain.

Furthermore, the distributional consequences of high-speed trading are non-trivial. Rents from speed accrue to a small number of geographically concentrated firms, potentially exacerbating inequality. Barriers to entry are high, not just technologically but also financially. These considerations lie beyond the direct scope of our study but warrant close attention in future work.

5.6. Future Research Directions

We plan to validate our classification framework using proprietary Deutsche Börse participant identifiers, enabling a confusion matrix of misclassifications and more accurate strategy tagging. We also plan to explore more avenues, expanding our scope to include more Eurex and Xetra products—particularly less liquid instruments—to test the robustness of our findings in different market environments.

Auctions represent another important area. These are periods of concentrated liquidity and strategic interaction, where UFT activity intensifies. Understanding how auctions affect trading behavior and market quality across participant types would enhance the realism of our conclusions.

We also propose a regime-specific extension of our analysis. While we computed multiple market regimes using volatility clustering, we did not yet assess how trader behavior and market quality metrics shift across these regimes. A regression-based extension of our current quintile approach would allow us to isolate the marginal effect of each participant type within different volatility, jump presence and time-of-day conditions.

Finally, we outline two important robustness checks. First, the out-of-trading hours analysis offers a quasi-natural experiment in trader presence, as lower volumes may reduce the profitability of high-speed strategies. Second, tick size change events, which we already documented, serve as strong exogenous shocks. These changes directly affect the profitability and viability of latency-sensitive strategies and provide compelling evidence for causal interpretation.

Together, these extensions would strengthen both the internal and external validity of our conclusions, laying the groundwork for more informed market design and regulatory debates.

6. Conclusion

This paper provides a detailed empirical analysis of ultra-fast and high-frequency trading activity on the Eurex exchange using nanosecond-level data. By classifying participants based on reaction time thresholds and analyzing their behavior across a wide range of metrics—including latency distributions, participation profiles, market quality impacts, and a novel measure of signal-to-noise ratio—we offer new insights into the stratified structure of modern electronic markets.

Our findings underscore a clear segmentation of strategic behavior and market function across latency tiers. Ultra-Fast Traders (UFTs), operating below the 1-microsecond threshold, dominate in terms of aggressive execution and event-driven latency arbitrage, yet exhibit limited contribution to short-term price informativeness. High-Frequency Traders (HFTs), active in the 1–10 microsecond window, demonstrate greater heterogeneity in strategy, balancing passive and aggressive participation while showing the strongest contribution to price discovery via our signal-to-noise ratio framework. Other participants—those outside these speed thresholds—display more irregular trading patterns and less consistent informational contribution.

Importantly, our novel signal-to-noise ratio metric allows us to move beyond traditional participation-based proxies to quantify the quality of trading activity. We show that HFTs consistently trade in ways that anticipate short-term price evolution, suggesting a predictive component to their strategies that goes beyond mere speed. In contrast, UFT behavior appears to rely more on reaction and positioning than forecasting, often capitalizing on momentary imbalances rather than contributing to efficient price formation.

These empirical patterns carry significant implications for market design and regulation. While speed clearly enhances certain aspects of market efficiency—particularly in reducing latency arbitrage and narrowing spreads—our analysis suggests that the marginal benefits of further latency reductions diminish sharply beyond the microsecond level. This supports a more nuanced regulatory approach that distinguishes between beneficial speed competition and potentially wasteful or destabilizing latency races. Tiered regulation, accounting for both participant speed and function, may provide a more effective framework than blunt constraints or one-size-fits-all policies.

Our work also highlights methodological innovations that may be useful in future research. The classification of participant types based on reaction times offers a practical alternative to proprietary IDs, enabling granular behavioral analysis without relying on firm-level identifiers. Moreover, our combination of participation metrics with a mark-out-based measure of informativeness provides a richer perspective on the link between trading activity and market quality.

Several limitations remain. Our analysis focuses on a single exchange and a limited set of instruments, and our anonymized participant classification cannot capture the full strategic complexity or cross-venue behavior of modern trading firms. Nevertheless, our results open the door to further investigation into the dynamic interplay between speed, strategy, and informational contribution in electronic markets.

Looking ahead, future research could extend our framework to multi-venue data, incorporate more robust classification validation using proprietary tags, and analyze regime-specific dynamics during stress events or structural changes. In doing so, we may move closer to a comprehensive understanding of how modern financial markets function—and how they might be designed to serve the broader goals of efficiency, fairness, and stability.

Acknowledgements

The authors gratefully acknowledge:

- The Deutsche Börse for hosting the PhD internship, particularly the Quantitative Analytics teams and Dr. Stefan Schlamp for supervision during the internship.

- The Swiss National Science Foundation (SNSF) for funding the PhD research project "Narrative Digital Finance: a tale of structural breaks, bubbles & market narratives" (grant number IZCOZ0-213370).

- the collaboration between ING Group and the University of Twente, which has contributed to research in artificial intelligence applications in finance.

- the International Advanced Fellowship-UBB program, funded by Babeș-Bolyai University (contract nr. 21PFE/30.12.2021, ID: PFE-550-UBB),

- Bern University of Applied Sciences, the author's employer.

- University of Twente, the host university awarding the PhD degree, and PhD promotor and supervisor Prof. Dr. Jörg Osterrieder and co-promotor and supervisor Dr. Xiaohong Huang.

- The COST Action 19130 Fintech and Artificial Intelligence for providing exposure to an important research network of researchers and industry members. COST Actions support collaboration and knowledge exchange among researchers across Europe.

- This project has received funding from the Horizon Europe research and innovation programme under the Marie Skłodowska-Curie Grant

Agreement, acronym: DIGITAL, No. 101119635. We are grateful to the MSCA network for offering access to leading industry partners and a vast network of researchers, professionals, and students.

References

[1] Aït-Sahalia, Y., Jacod, J., 2009. Testing for jumps in a discretely observed process .

[2] Aït-Sahalia, Y., Yu, J., 2009. High frequency market microstructure noise estimates and liquidity measures. The Annals of Applied Statistics 3, 422–457. doi:10.1214/08-AOAS200.

[3] Aldridge, I., 2013. High-frequency trading: a practical guide to algorithmic strategies and trading systems. John Wiley & Sons.

[4] Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5, 31–56.

[5] Ammar, I.B., Hellara, S., 2021. Intraday interactions between high-frequency trading and price efficiency. Finance Research Letters 41, 101862.

[6] Anand, A., Venkataraman, K., 2016. Market conditions, fragility, and the economics of market making. Journal of Financial Economics 121, 327–349.

[7] Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2001. The distribution of realized exchange rate volatility. Journal of the American Statistical Association 96, 42–55.

[8] Aquilina, M., Budish, E., O'Neill, P., 2022. Quantifying the high-frequency trading "arms race". The Quarterly Journal of Economics 137, 493–564.

[9] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2008. Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. Econometrica 76, 1481–1536.

[10] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2009. Realized kernels in practice: Trades and quotes.

[11] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2011. Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. Journal of Econometrics 162, 149–169.

[12] Barndorff-Nielsen, O.E., Shephard, N., 2006. Econometrics of testing for jumps in financial economics using bipower variation. Journal of Financial Econometrics 4, 1–30.

[13] Baron, M., Brogaard, J., Hagströmer, B., Kirilenko, A., 2019. Risk and return in high-frequency trading. Journal of Financial and Quantitative Analysis 54, 993–1024.

[14] Biais, B., Foucault, T., Moinas, S., 2015. Equilibrium fast trading. Journal of Financial Economics 116, 292–313.

[15] Brogaard, J., Hagströmer, B., Norden, L., Riordan, R., 2015. Trading fast and slow: Colocation and liquidity. The Review of Financial Studies 28, 3407–3443.

[16] Brogaard, J., Hendershott, T., Hunt, S., Ysusi, C., 2019. Price discovery without trading: Evidence from limit orders. The Journal of Finance 74, 1621–1658.

[17] Brogaard, J., Hendershott, T., Riordan, R., 2014. High-frequency trading and price discovery. The Review of Financial Studies 27, 2267–2306.

[18] Budish, E., Cramton, P., Shim, J., 2015. The high-frequency trading arms race: Frequent batch auctions as a market design response. The Quarterly Journal of Economics 130, 1547–1621.

[19] Conrad, J., Wahal, S., Xiang, J., 2015. High-frequency quoting, trading, and the efficiency of prices. Journal of Financial Economics 116, 271–291.

[20] Friederich, S., Payne, R., 2015. Order-to-trade ratios and market liquidity. Journal of Banking & Finance 50, 214–223. doi:`https://doi.org/10.1016/j.jbankfin.2014.10.005`.

[21] Gonzalo, J., Granger, C., 1995. Estimation of common long-memory components in cointegrated systems. Journal of Business & Economic Statistics 13, 27–35.

[22] Hasbrouck, J., 1995. One security, many markets: Determining the contributions to price discovery. The Journal of Finance 50, 1175–1199.

[23] Hasbrouck, J., 2019. Price discovery in high resolution. Journal of Financial Econometrics 17, 1–30.

[24] Hasbrouck, J., Saar, G., 2013. Low-latency trading. Journal of Financial Markets 16, 646–679.

[25] Hautsch, N., Huang, R., 2012. The market impact of a limit order. Journal of Economic Dynamics and Control 36, 501–522. doi:`https://doi.org/10.1016/j.jedc.2011.09.012`.

[26] Hendershott, T., Jones, C.M., Menkveld, A.J., 2011. Does algorithmic trading improve liquidity? The Journal of Finance 66, 1–33.

[27] Kirilenko, A., Kyle, A.S., Samadi, M., Tuzun, T., 2017. The flash crash: High-frequency trading in an electronic market. The Journal of Finance 72, 967–998.

[28] Lee, S.S., Mykland, P.A., 2008. Jumps in financial markets: A new nonparametric test and jump dynamics. The Review of Financial Studies 21, 2535–2563.

[29] Menkveld, A.J., 2013. High frequency trading and the new market makers. Journal of Financial Markets 16, 712–740.

[30] Zhang, F., 2010. High-frequency trading, stock volatility, and price discovery. Working paper .

[31] Zhang, L., Mykland, P.A., Aït-Sahalia, Y., 2005. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. Journal of the American Statistical Association 100, 1394–1411.