LUISS

DEPARTMENT OF ECONOMICS AND FINANCE

CHAIR OF CAPITAL MARKETS

"The Evolution of Program Trading: A Comprehensive Study on High-frequency Trading and its Impact on Capital Markets"

Prof. Paolo Vitale
SUPERVISOR

Prof. Giovanni Rillo
CO-SUPERVISOR

Leonardo Roversi
ID: 272521
CANDIDATE

ACADEMIC YEAR 2023/2024

ACKNOWLEDGMENTS

Before proceeding with this thesis, I would like to thank Professor Paolo Vitale and Professor Giovanni Rillo for their invaluable guidance, and insightful feedback, which have been essential to the completion of this work. This thesis was an opportunity for me to challenge myself, and it would not have been possible without their support, for which I am grateful,

Furthermore, I would like to thank my parents for always supporting me in my aspirations and decisions, and for always believing in me.

Finally, I want to thank myself, for always holding myself to a higher standard and always pushing beyond my limits.

INDEX

Abstract **4**

Introduction **5**

Digitalization of Financial Markets **5**

1.1 From the Ticker Tape to the Computer 5

1.2 The Role of Program Trading in the October 1987 Black Monday 7

1.3 The rise of Automated Trading Practices 10

1.3.1 Algorithmic Trading 11

1.3.1.1 Efficiency-driven Algorithms 12

1.3.1.2 Prediction-driven Algorithms 14

1.3.2 High-frequency Trading 16

Impact of High-frequency Trading on the Markets **20**

2.1 A Review of the Available Literature 21

2.1.1 Event Studies on the 2010 Flash Crash 22

2.1.2 Methodology Analysis of Empirical Studies on HFT 28

2.2 Takeaways From the Literature Review 31

Empirical study on HFT **32**

3.1 Collecting the Data 32

3.1.1 Description of the Sample 32

3.1.2 Cleaning the Data 33

3.2 Data Analysis 35

3.2.1 Hypothesis Statement 36

3.2.2 Hypothesis Testing 37

3.2.3 Evaluation of the Linear Model 38

3.3 Results and Final Remarks 40

Conclusions **42**

Bibliography **44**

ABSTRACT

This thesis examines the evolution of program trading, with a particular focus on the development and implications of high-frequency trading (HFT) in modern financial markets. Over the last five decades, financial markets have been profoundly transformed by the rapid digitalization and globalization that characterize the contemporary era. The transition from manual trading methods to highly automated systems has resulted in the widespread adoption of algorithmic and high-frequency trading practices, which now account for a major share of the trading volumes of leading global markets.

The study begins by tracing the historical evolution of trading technologies, starting from the introduction of electronic trading systems in the 1970s and the subsequent rise of algorithmic trading. It then delves into the mechanisms and strategies underlying HFT, exploring how these practices have reshaped market dynamics by enhancing liquidity, reducing transaction costs, and increasing the speed of price discovery. However, the research also highlights the potential risks associated with HFT, particularly its role in amplifying market volatility during moments of market distress, as evidenced by events like the 1987 Black Monday and the Flash Crash of 2010.

To assess the impact of HFT on market quality, an empirical analysis is conducted using trade data from a sample of U.S. stocks. The study employs the Cancel-to-Trade Ratio (CTR) as a proxy for HFT activity and investigates its relationship with short-term market volatility. The findings reveal a statistically significant inverse correlation, suggesting that higher levels of HFT activity tend to correspond with lower short-term volatility. Despite this, the study acknowledges that in times of market turmoil, HFT can contribute to liquidity dry-ups and exacerbate systemic risks, highlighting the dual-edged nature of these trading technologies.

This thesis contributes to the ongoing academic debate on the benefits and drawbacks of HFT, providing empirical evidence that supports the notion of HFT as a stabilizing force under normal market conditions, while also recognizing its potential to destabilize markets in times of crisis. The research concludes by emphasizing the importance of continued monitoring and regulation of HFT practices to ensure that the financial system can harness the benefits of technological advancements while mitigating associated risks.

INTRODUCTION

Over the last 50 years, global financial markets have undergone major structural changes, ultimately driven by the rapid technological advancements and globalization forces that characterized the contemporary age. During this time, the global economy witnessed an unprecedented increase in real GDP per capita, with the global markets' capitalization going from $1.15 trillion in 1975 to $93.7 trillion in 2022 (Tytell et al., 2016; World Bank, 2023). This remarkable growth was primarily made possible by the digitalization process of the financial industry, began in 1971 with the creation of the NASDAQ. The advent of computers in the financial industry represented indeed a catalytic event that fostered informational and transactional efficiency, improving the allocation of financial resources in the global economy. Furthermore, digitalization laid the groundwork for the development of more sophisticated ways of trading, such as algorithmic, and later high-frequency trading, practices that were rapidly adopted by big banks and institutional investors, arriving to account today for 60-75% of the overall volume traded on the biggest global markets.

This thesis will then retrace this technological evolution from its beginnings until today, presenting all the most relevant information to highlight the opportunities and threats related to algorithmic and high-frequency trading.

DIGITALIZATION OF FINANCIAL MARKETS

1.1 FROM THE TICKER TAPE TO THE COMPUTER

Before financial markets went digital, information transmission and reception were slow. Until 1960, all stock market information was transmitted over telegraph lines and punched out on a ticker tape containing trade volumes and prices, with all price movements being registered and graphed by hand. Once the information about the market price was sent outside via ticker tape, the individual interested in making a trade could send a written order or place a call to communicate the details of their order to their broker, who negotiated the trade on their behalf in the trading pit via open outcry method (Chronohistoria, 2022).

Open outcry was the main and only way trades were placed before the advent of electronic trading systems. Brokers met and negotiated trades face-to-face in dedicated areas of the trading floors of the exchange called trading pits, by shouting their prices and using hand signals. This method served as a transparent way of price discovery as participants could compete for orders making their best price in real-time, adapting to changing market conditions.

Although the physical proximity of buyers and sellers provided for dynamic and accurate pricing of securities on the floor, the process of getting information out, clearing, and registering transactions was still slow and inefficient.

Toward the end of the 1960s, a sharp increase in trading volumes saw the number of shares trading per day rise from 4.9 million to 14.9 million, flooding back offices on Wall Street with paperwork, almost collapsing the exchanges under the weight of their own order flows, in what was later referred to as the “Paperwork crisis”.
The crisis put pressure on broker-dealer firms, which incurred substantial losses trying to cope with the volume of trades while seeing most of them never being settled.
Between 1967 and 1968, about 160 NYSE member firms went out of business (SEC Historical, 2022; SEC, 2004).

These events brought the industry’s attention to the matter of automatization, forcing firms to invest into the development of computerized data systems. In 1971, the National Agency of Securities Dealers (NASD) created the first computerized quotation system, founding the National Agency of Securities Dealers Automated Quotation (NASDAQ). The NASDAQ began its operations on February 8th as the world’s first electronic stock market, providing brokers with the latest competitive price quotes via a computer terminal.

Later in 1972, the NYSE and AMEX jointly owned subsidiary, the Securities Industry Automation Corporation (SIAC), was commissioned to develop a digitalized order processing system. Four years later the Designated Order Turnaround System (DOT) was born, allowing investors to route their orders directly to a specialist on the exchange's floor who executed the trade at the prevailing market price, increasing processing efficiency while limiting errors.
This system was later updated in 1984 with the introduction of the Super-DOT and kept in use until 2009.

The computerization of the financial industry started to pick up during the 80s, however, all the successful electronic processing systems until this point were merely giving live quotations, and electronically sending orders to the central order books of the exchanges where they were manually executed by the specialists.

The first to try automating the execution process as well was the Institutional Network Corporation, later renamed Instinet.
Instinet allowed broker-dealers and institutional traders to make orders for buying or selling securities, which were then matched together through an order-matching system that guaranteed the execution of the trade at the best available price. Instinet would then process and report executions, while not publicly disclosing trades, protecting the identities of the counterparties in the trade.
First founded in 1969, Instinet initially struggled to take off during the 1970s, finally being linked to major exchanges only in 1983, when it was connected to the NASDAQ and the London Stock Exchange (LSE) (SEC Historical, 2022).

As the 1970s paved the way with the introduction of electronic exchanges, during the 1980s more Electronic Trading Systems (ETS) that included order-matching started to make their appearance in the market. An example of this was the MAX system launched by the Chicago Stock Exchange in 1982, which became one of the first stock exchanges to provide fully automated order execution (Encyclopedia, 2022).
This last step in the computerization process of markets now allowed investors to quickly access information, send orders, and trade securities in real-time from a computer terminal, leading to a major increase in the number of daily orders received by the exchanges, while decreasing the time needed to process and clear the orders.

1.2 THE ROLE OF PROGRAM TRADING IN THE OCTOBER 1987 BLACK MONDAY

On October 19th, 1987, also known as Black Monday, financial markets confronted for the first time the potential threats associated with automated trading systems. This day saw the unthinkable happen when worldwide equity markets registered the biggest single-day percentage drop ever recorded in history.
On that day 19 out of the 23 world's largest markets dropped by more than 20%, resulting in an estimated loss of $1 trillion (Brady, 1988).

The elements that favored the crash were many: overinflated stock prices, rising interest rates, trade and budget deficits, and a decline in the value of the dollar are among the most cited. However, later reports revealed that what could have been just a moment of market correction became instead a full-on crash because of downward pressure on prices exerted by portfolio insurance and index arbitrage activities.

Portfolio insurance and index arbitrage represent early examples of trading strategies developed around algorithms that allowed investors to buy or sell securities automatically without human intervention, by capturing price signals, or by finding and exploiting arbitrage opportunities.

Portfolio insurance was a hedging technique developed to limit losses investors could incur on stock indexes in moments of uncertainty or price decline, which involved the buying and selling of index futures.

This practice had been largely adopted by institutional investors by the time of the crash, and it has been deemed to be the initial cause of downward pressure on prices, which along with the adoption of stop-loss orders helped create a negative feedback loop that exacerbated the crash.

Indeed, in the trading week preceding the crash, a higher-than-expected U.S trading deficit, along with a decline in the value of the dollar and a concurrent increase in short-term interest rates by the Fed, induced a correction in equity markets which saw the S&P 500 close on a 9% reduction on Friday 16th. The initial decline in prices, attributed to sales of equities by institutional investors, pushed portfolio insurers to sell in the futures markets to hedge against the loss in value of their stock holdings. In turn, the increased sales of futures contracts created discrepancies between the values of indexes in the futures markets and the NYSE, which were exploited by index arbitrageurs by buying futures and selling stocks, increasing the downward pressure on equities prices.

Reports show that when the markets opened on Black Monday, 30% of both the S&P Index and the Dow-Jones Industrial Average (DJIA) stocks opened for trading one hour late due to temporary suspensions operated by NYSE specialists, in an attempt to deal with the disproportionate amount of sell orders which piled up over the weekend and keep balance in their order books.

While NYSE quotations were stale at Friday’s prices, futures markets on the other hand opened on time with prices still in decline, creating more opportunities for index arbitrage and sending prices in a downward spiral once specialists resumed trading.

The erosion of price levels additionally posed a threat to market liquidity by triggering an unprecedented number of margin calls on Tuesday, amounting to three times the average number and ten times the average size of daily calls for a total of $2.1 billion.
To be able to meet the margin calls, most of the CME clearinghouse member firms had to rely on credit line extensions provided by settlement banks, which had a crucial role in providing liquidity to brokers and dealers firms despite exposure size concerns, allowing institutions to meet their margin requirements without disruptions to trading.
On the retail side, on the other hand, investors had to liquidate quickly part of their holdings to meet margin requirements or liquidate their option positions if unable to do so, which likely added more pressure on the markets (Brady, 1988).

On the morning of October 20th, the Federal Reserve timely acted to prevent a liquidity dry-up and restore public confidence in the financial system by issuing a public statement indicating that it would have supported market liquidity by carrying out open market operations, lowering short-term interest rates and providing banks and securities firms with credit and guidance to support funding needs of brokers and dealers.
The Fed's efforts in response to the crash revealed to be effective and vital to restoring public confidence, improving the functioning of markets and reducing volatility in the following weeks (Fed History, 2013).

Our focus on the 1987 Black Monday highlights how financial innovations, such as electronic and automated trading, can extensively alter the way financial markets behave and react to shocks, sometimes revealing some hidden related weaknesses. In that instance, the crash revealed how electronic trading routes and algorithmic trading increased the degree of interconnectedness between different markets and exchanges, along with faster ways of trading and information dissemination, heightening the creation of systemic risks.
It also points out to us how oftentimes innovation in this industry runs faster than the regulatory framework which is supposed to ensure the system's soundness, as it was seen that the absence of circuit-breakers in the futures markets compared to the ones employed by the NYSE, for instance, caused markets to function poorly and unevenly, increasing information asymmetries and fostering uncertainty.

1.3 THE RISE OF AUTOMATED TRADING PRACTICES

Automated trading finds its origins in the 1980s, when the emergence of Electronic Communication Networks (ECNs), such as Instinet, provided investors from both the buy and the sell side of the market with Direct Market Access (DMA), enabling them to use a broker's or another financial institution's electronic infrastructure to directly place orders on central order books of exchanges without a broker's intervention. Soon enough, DMA providers also began to provide trading algorithm services to their clients such as order-slicing tools aimed at limiting the price impact of larger orders on the market. During the 1990s, increased industry competition also led buy-side entities with DMA, such as hedge funds, pension funds, and mutual funds, to employ proprietary algorithms to hedge risks and increase profits (Kim, 2007).

On top of DMA, in the later years between the end of the 1990s and the beginning of the 2000s, several changes in the financial landscape supported the adoption and development of more advanced and complex forms of algorithmic trading.

First in the order came the widespread adoption of the Financial Information Exchange Protocol (FIX). First introduced in 1992, the FIX protocol was an electronic communication protocol used for the real-time exchange of securities transaction information, designed to improve communications and transfers of information between market participants in the attempt to standardize communication across different trading platforms and venues. Nowadays the FIX is the world's leading communication standard for securities trading, ensuring communication efficiency and transparency (FIX, 2010).

The second enabling factor was the decimalization of price quotes commissioned by the Securities and Exchange Commission in 2001, which reduced the minimum tick size from 1/16th of a dollar ($0.0625) to $0.01. This action was taken to conform U.S. securities markets to foreign markets that had already adopted decimal pricing, and further to support healthy competition among market makers, deemed to be profiting from "artificially wide spreads" (SEC, 1994). Tighter bid-ask spreads resulted in increased competition and decreased profitability among market makers, which conversely entailed lower transaction costs for market takers. By increasing the number of price points at which stocks can be traded, decimalization additionally led to increased liquidity, reducing the price impact of large orders and volatility (SEC, 2012).

The rise of DMA, the adoption of standardized information exchange protocols, intensified competition for best-price executions, and significant advances in computational power represented the key drivers that fostered the development and widespread use of algorithmic trading strategies and later high-frequency trading. In this next section, we will explain more in-depth how AT functions, and how HFT, as a subset of the former, originates and differentiates from it.

1.3.1 ALGORITHMIC TRADING

Algorithmic trading refers to the use of computer algorithms to conduct trading decisions in line with pre-defined parameters, without any human involvement.
This is achieved by making use of computers' ability to analyze broad ranges of market data in real-time and adjust their trading decisions according to changing market conditions, by applying complex mathematical and statistical models.
The primary goal of algorithmic trading is to increase trading efficiency to maximize profits while minimizing trading costs, risks, and market impact. Algorithms do so by ensuring execution accuracy and timing beyond human capabilities, removing emotions and human errors from the equation.

The world of trading algorithms is understandably dynamic and constantly changing, as with rapid technological advancements, and new studies and data available, new kinds of strategies continuously emerge. Although this makes the classification of these strategies particularly challenging as distinctions between them are not always clear-cut, we will now analyze how some of the predominantly used AT strategies work and what they are meant to accomplish.

The first distinction between algorithms is made depending on their primary objective: execution algorithms have the main goal of ensuring optimal trade execution given the prevailing market conditions, trying to achieve execution at the best possible price, or to minimize market impact of large orders. These types of algorithms are defined as "efficiency-driven".
Systematic trading algorithms on the other hand, more complex than the former, are algorithmic strategies designed to analyze historical market data and apply complex mathematical and statistical models to make predictions about market trends in order to yield profits. They are also defined as "prediction-driven" algorithms.

1.3.1.1 EFFICIENCY-DRIVEN ALGORITHMS

Execution algorithms can be further classified depending on their specific objectives and the factors they take into account when making trading decisions. Among them, we find four main types:

Volume-Weighted Average Price (VWAP)

This type of algorithms aims at achieving optimal executions while limiting market impact by splitting large orders into smaller chunks, with the goal of executing transactions matching or even beating the Volume-Weighted Average Price benchmark. This benchmark is obtained by gathering historical and current prices and volumes for a certain stock, and calculate the weighted average with respect to order volume, following the following formula: $VWAP = \frac{\sum(size \times Price)}{tot.\ Volume}$.
Taking the example provided in *Fig. 1.1* showing the limit order book for a generic stock "x", applying our formula gives us the respective volume-weighted average bid and ask prices for the stock, also shown below.

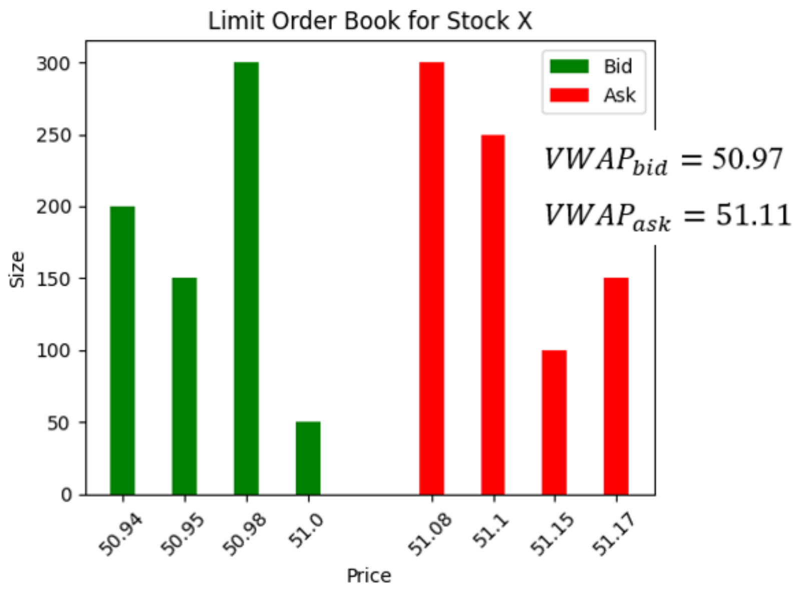

Figure 1.1 – Visualization of the VWAP benchmark for bid and ask prices on a generic stock "x". This example reports a snapshot of the benchmark price in one moment in time, which in a real-world example would be calculated by gathering prices and trading volumes observed in a selected time window.

In this presented instance, the algorithm trying to execute a large order of stock "X" would slice the order into smaller lots and will then try to execute them at a price at least equal to or higher than the

average bid price in the case of a market sell order, or lower than the average ask price in the case of a buy order.
This type of algorithm was one of the first execution algorithms largely adopted by institutional investors to limit market impact of large transactions and avoid information leakage. Indeed, the VWAP benchmark is still to this day considered a very reliable indicator of optimal trade executions by institutional investors (Papers With Backtest, 2024).

Time-Weighted Average Price (TWAP)

This strategy aims at limiting market impact by dividing large orders into smaller equal lots and executing them at regular pre-defined time intervals within a set trading time window. This strategy should also thus guarantee an average execution price in line with the average market price of the security during the interested period. However, due to its static way of trading once the strategy is set, this algorithm may be less suited for trading in highly volatile securities, in which case setting shorter trading windows or opting for better suited strategies would be advisable (PWB, 2024).

Implementation Shortfall

This type of algorithms essentially aims at reducing the "opportunity cost of trading", by executing trades so to minimize price slippage costs, namely the cost associated with the difference between the decision price (the market price at which the trade is initiated) and the actual execution price. These algorithms can be extensively useful in assuring execution effectiveness in highly liquid or volatile markets in which delays in order routing or price impact of large trades can enhance price discrepancies. This strategy employs real-time monitoring of the execution process and dynamically adjusts the pace or size of the orders to counter adverse price movements. Furthermore, this strategy also necessitates a thorough understanding of market depth, liquidity, and real-time price action to guarantee effective executions and preserve the profitability of trades (PWB, 2024).

Percentage of Volume (POV)

This algorithmic strategy allows investors to execute trades by participating in the market proportionally to the overall volume traded at any time. This algorithm receives the time window, order size, and a price band to observe as inputs, and continuously observes traded volume to forecast market activity thanks to a predictive algorithm, upon which it models its trading pattern. In practice if the algorithm is set to trade $x\%$ in each interval of the set timeframe, it will observe the absolute number of shares traded during the interval and trade $x\%$ of that number, repeating this for every interval until the order is completed or the time window expired (Kissell, 2014).

As we might notice although different in their design, execution-driven strategies all share the same basic objectives of achieving optimal executions while limiting price impact on the market by slicing large orders. For this reason, these types of algorithms are mainly integrated by institutional investors and are often offered by trading firms as a service to professional clients.

1.3.1.2 PREDICTION-DRIVEN ALGORITHMS

On the other hand, as we will now see, systematic trading algorithms are far more complex and span over a wide range of applications, which can greatly differ in goals and time scales. This type of strategies is thus categorized depending on their time horizons and trading patterns as low-frequency, medium-frequency, and high-frequency trading.

Low-Frequency Trading strategies

Low-Frequency Trading strategies (LFT) are strategies characterized by a lower number of trades with longer holding periods, laying their foundations over fundamental analysis, macroeconomic trends, and data-driven models to carry out long-term investment strategies and effective risk management.
In this category, we find strategies such as Growth and Value Investing, which leverage deep fundamental analysis to identify stocks projected to grow faster than their peers or that are undervalued based on their calculated intrinsic value.

Sentiment Analysis strategies involve the use of Natural Language Processing (NLP) models to interpret news, reports, and social media posts to predict market sentiment and direct investing decisions.

Global Macro Strategies make investment decisions by interpreting macroeconomic indicators and global events to predict long-term effects on the markets. By analyzing data such as employment, interest, and inflation rates, GDP growth, and news related to geopolitical events such as elections or trade wars, these algorithmic strategies can capture a variety of trading opportunities across different asset classes and regions, resulting in diversified portfolios and effective risk management (PWB, 2024).

Medium-Frequency Trading strategies

Medium-Frequency Trading (MFT) strategies fall in between the extremes represented by LFT and HFT respectively, focusing on strategies that involve the holding of securities from minutes to a few days. These algorithms involve a mix of fundamental and technical analysis to take advantage of short-term price movements, while keeping a more considered and less capital-intensive approach to trading than HFT, observing broader market trends and economic conditions (Indian Institute of Quantitative Finance, 2023).
Among these strategies, we find Momentum, or Trend Following trading, aimed at identifying trading opportunities by observing technical indicators such as moving averages, momentum oscillators, and trendlines. For example, a trading opportunity could be identified when a security deviates from its historical moving average, which depending on the direction of the deviation can be a signal of an upward or downward trend that can be exploited by taking a long or short position respectively.

Mean Reversion trading strategies, on the other hand, are strategies that capitalize on the assumption that a security which is now trading at a premium or a discount over its historical average price after an event, will eventually revert and converge back to its mean value. This holds onto the assumption that at times the market overreacts to news events creating temporary price anomalies.

Pairs Trading or Statistical Arbitrage is a widely adopted MFT strategy that involves taking simultaneous long and short positions in a pair of assets that are highly historically correlated to each other, with the assumption that any significant divergence in their prices should be only temporary, therefore representing a trading opportunity as illustrated in the picture below (PWB, 2024).

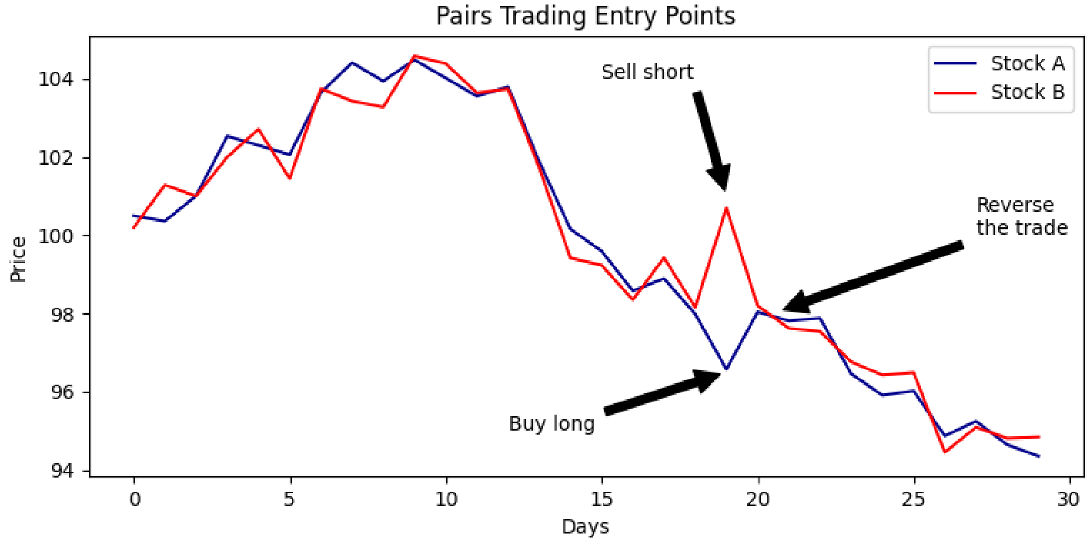

Figure 1.2 – Visualization of a Pairs Trading strategy showing price changes of two highly correlated stocks. Whenever the value of the two stocks starts to diverge significantly, an entry point opportunity arises to buy the underperforming asset and sell short the overperforming one, reversing the trade when their prices realign.

High-Frequency Trading strategies

Finally, High-Frequency Trading (HFT) is a form of algorithmic trading that leverages high-speed transactions and high-tech infrastructure to profit from short-lived price discrepancies and opportunities.

HFT particularly relies on Ultra-Low-Latency DMA (ULLDMA) infrastructures and co-location services to trade large numbers of securities in microseconds or even nanoseconds. Orders are submitted, updated, or canceled quickly, in order to consistently earn small profits from a large number of trades. The securities traded are held just for a few seconds or minutes at most, with no significant position (flat position) held at the end of the day. Hence, HFT generally focuses on highly liquid instruments (Gomber et al., 2011).

1.3.2 HIGH-FREQUENCY TRADING

High-frequency trading originated as an evolution of algorithmic trading in the 1990s, as with increasingly more firms employing algorithms for conducting their operations, improving execution speed was the next logical step to stay ahead of the competition. Resultantly, trading firms focusing on high-speed strategies started taking all the steps necessary to reduce the latency of their trades,

making large investments to improve their computing power, acquiring Sponsored Access (SA)[1], and making use of co-location services through which HFT firms could pay to locate part of their servers inside an exchange's data center, solely to gain an edge of a few nanoseconds by reducing the distance the information sent or received had to travel (McGowan, 2010).
As a consequence, many HFT strategies still in use today are no different than some AT strategies that were previously discussed but executed at higher speeds. However, split-second execution speeds also allowed for a much wider range of applications, spanning from arbitrage to market-making. Here are three exemplary applications of this:

Latency Arbitrage

This form of HFT leverages time advantages obtained from high-speed infrastructures to receive and act on price information before other market participants, or to capitalize on small, split-second price discrepancies by trading simultaneously on different markets. For these strategies, co-location services are essential where possible, although other available alternatives include microwave, satellite, or shortwave radio transmissions (PWB, 2024).

Market Making

Market making is a crucial component of the financial ecosystem, ensuring liquidity and facilitating trade execution for a vast array of securities. Market making involves the simultaneous quoting of buy and sell prices for a financial instrument, intending to profit from the bid-ask spread.
Traditionally, market making was preeminently manual, with traders setting bid and ask prices based on their assessment of market conditions. However, with the rise of HFT and AT this process has been largely automated as algorithms can rapidly adjust quotes in response to changing market conditions, manage multiple securities simultaneously, and execute trades in fractions of a second, thus providing liquidity to the market in a more efficient manner.

High-frequency market-making strategies can take two distinct forms: spread-capturing or rebate-driven.

[1] Sponsored Access is a form of Direct Market Access through which Broker-Dealers firms grant their clients direct access to the Exchange trading system without routing it through their own trading system first. This arrangement aims at facilitating low-latency trading and preserving the confidentiality of clients' sophisticated, proprietary trading strategies (NSE International Exchange, n.d.). For more see https://www.nseix.com/technology/direct-market-access.

Spread-capturing strategies simply profit by posting limit buy or sell orders (passive trading), providing liquidity and capturing the spread between the higher (ask) price at which they sell the securities to the market-takers (active buyers and sellers), and the lower (bid) price at which they are willing to buy them from them.

Rebate-driven strategies, on the other hand, rely on the asymmetric fee structure employed by most trading venues nowadays, which charge a lower transaction fee or provide a rebate to market-makers to incentivize liquidity provision, while charging a higher fee to market-takers. The rationale behind this structure is that the incentive to provide liquidity enhances competition between market-makers, which causes spreads (thus their profit) to shrink, benefiting buyers and price stability. The reduction in profit is then supposed to be offset by the rebate or lower fee charged (McGowan, 2010).

An example of this consideration towards market-making activities particularly pertinent in this context was given when Italy in 2013 became the first country to introduce a tax specifically targeting HFT, charging a 0.02% tax on equity transactions lasting less than 0.5 seconds, except for market-making activities as it reads:

"The transactions effected on the Italian financial market are subject to a tax on high-frequency trading relating to shares, participating financial instruments, securities representing equity investment and transferable securities…

… with the exclusion of those used: (1) for the performance of the market-making activity…"

(Italian Minister of Economy and Finance, 2013, Art.12)

Liquidity detection

Liquidity detection strategies (often referred to as sniffing/sniping) are HFT algorithms designed to discern the patterns left by other market participants and adjust their trading accordingly. Liquidity detectors are able to "sniff out" the activity of order-slicing algorithms employed to cover the tracks of large orders, and thus be able to predict small price movements produced by the entrance of the order on the market.

A clever application of this liquidity detection is used for instance in the "Sub-penny jumping" strategy, illustrated in the picture below taking the example of a large limit buy order entering the market.

Indeed, in this strategy, when a large order is detected entering the market, the algorithm immediately responds by placing a limit (buy) order of its own, at a price slightly higher $p + \varepsilon$, effectively frontrunning the human trader (*Fig. 1.3b)*. If the HFT order is matched by the market and the trader buys the security at $p + \varepsilon$, the algorithm places a limit sell order to sell the securities at a higher price (*Fig. 1.3c)*.

Figure 1.3a

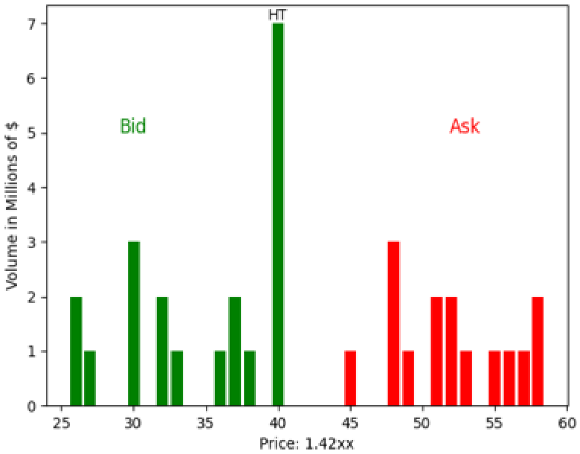

Figure 1.3b

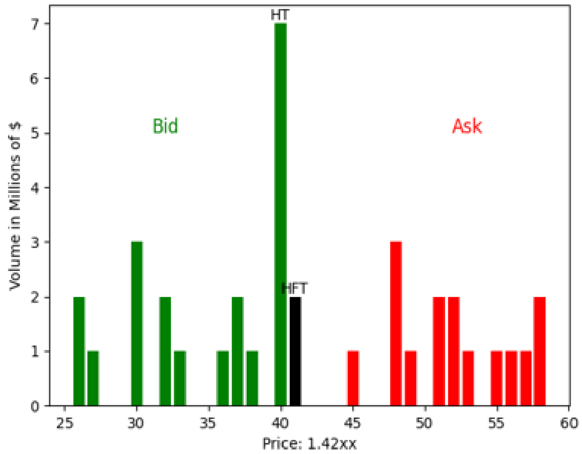

Figure 1.3c

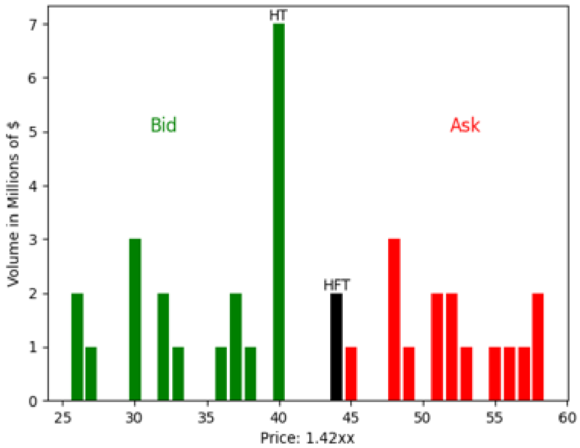

Figure 1.3d

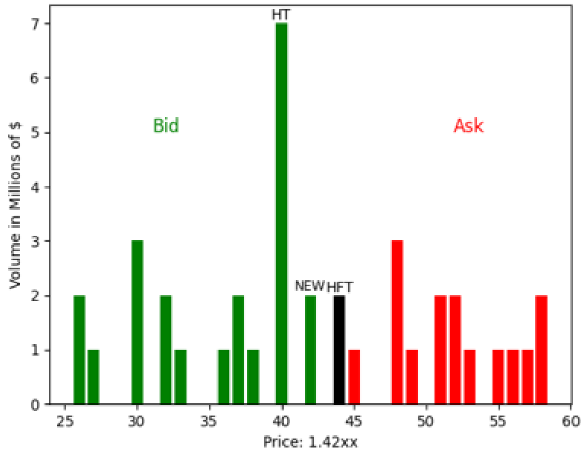

Figure 1.3 – Graphical representation of the sub-penny jumping strategy implemented by HFTs after detecting a large order entering the market. Image re-adapted from Mahmoodzadeha & Gençay (2017).

At this point, three possibilities emerge:

- The HFT sell order can be bought by another trader and the HFT trader earns the spread.
- A new market buy order at a price $p_{new} > p + \varepsilon$ is entered, the algorithm then cancels the old sell order and places a new one to match the buy order's price. The HFT trader earns the difference between the two prices (*Fig. 1.3d*).
- If the algorithm detects that the market is drifting away from its position, it cancels the sell order and matches the order of the human trader in *Fig. 1.3a*, incurring a small loss of $\varepsilon \times vol$.

(Mahmoodzadeha & Gençay, 2017).

This type of practice has been largely controversial as it seems to bring no clear benefit to the markets and may rather undermine the ability of human traders to see their orders being timely executed. Nonetheless, this type of activities is still loosely regulated.

Now that we have outlined the main strategies and characteristics associated with algorithmic and high-frequency trading, the next chapter will instead focus on the impact that these practices have on the smooth functioning of financial markets. Our analysis will then proceed diving into the academic literature addressing the most relevant issues related to the low latency trading environment, thoroughly discussing event studies and empirical studies, and assessing the most commonly employed methodologies to research these topics.

IMPACT OF HIGH-FREQUENCY TRADING ON THE MARKETS

Since the start of their adoption in the 1990s, trading algorithms have been rapidly and steadily growing their presence in worldwide financial markets, playing a major role in improving the way our financial system allocates financial resources, and promoting economic efficiency.
Additionally, in more recent years this growth has been boosted by technological achievements obtained in the realms of artificial intelligence and machine learning, allowing trading programs to account for 60-75% of the overall traded volume in the U.S, European, and Asian markets, amounting to a total market size of $15.6 billion in 2023, expected to more than double by 2032 (IMARC Group, 2023).
Yet, being a very technologically and capital-intensive sector, the main players in HFT are institutional investors such as investment banks, hedge funds, and pension funds, having the resources to invest millions of dollars annually to maintain and improve their IT infrastructures and retain a competitive position.
Therefore, considering how heavily our financial system and institutions lean on these technologies every day, automated trading has long been under the microscope of the academic community and has been extensively studied in an attempt to determine if, and to what extent, these tools might benefit or harm the proper functioning of financial markets.

2.1 A REVIEW OF THE AVAILABLE LITERATURE

The available academic literature on high-frequency trading predominantly focuses on its impact on market quality, primarily examining how it affects market liquidity, depth, and volatility, during both periods of stability and times of market distress.
According to data collected by Carè and Cumming (2024), a total of 863 papers had been published on these topics by 2022, with a significant increase in publications from 2010 onwards (see *Fig. 2.1*). This surge in scholarly attention is largely attributed to the events of the Global Financial Crisis of 2008, which prompted a reassessment of trading practices and risk management strategies, and to the Flash Crash of May 6, 2010. The Flash Crash in particular largely caught the attention of the academic community involved in studying automated trading practices, as analogously to the 1987 Black Monday, this market crash showed how the interplay of algorithmic and high-frequency trading strategies concurred to absorb liquidity and increase volatility in the markets in moments of market turmoil.

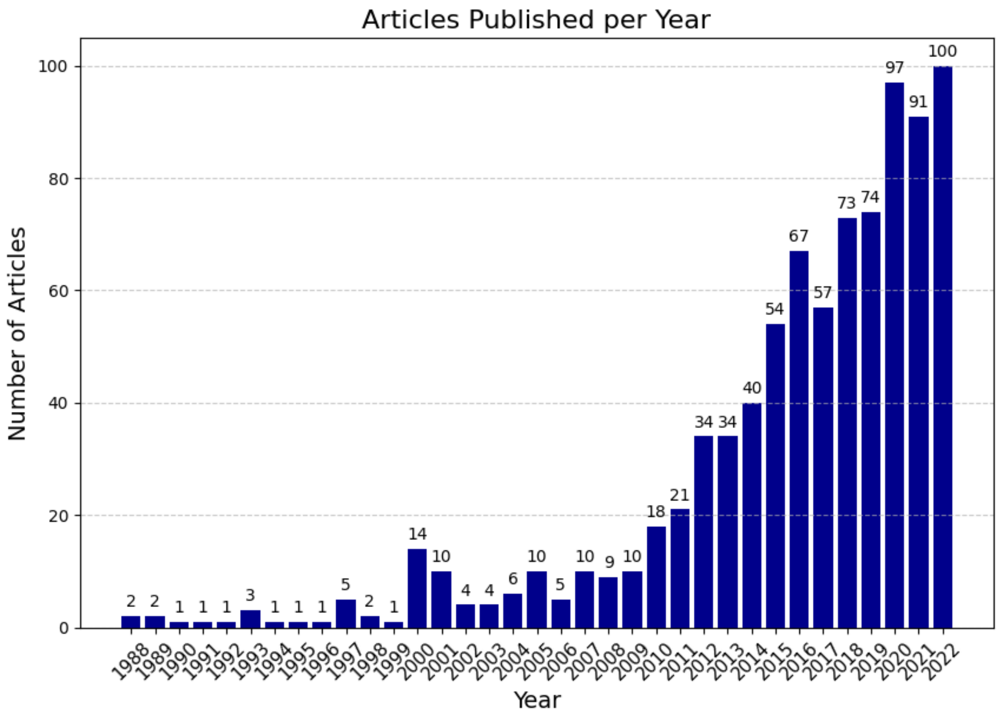

Figure 2.1 – Articles on AT and HFT published by year 1988-2022. Graph re-adapted from Carè and Cumming (2024).

Indeed, as Brogaard (2010) remarks, all the empirical works related to HFT study this phenomenon either through event studies about the 2010 crash or by making use of indirect proxies.

This is due to a variety of reasons: Zaharudin et al. (2022) for example indicate how the absence of a universally recognized definition of HFT among scholars and institutional bodies makes it difficult to delineate which activities fall into the category, and hence to properly assess and quantify its presence. Furthermore, as it is not possible to directly observe which trades are initiated by a computer program, researchers have to rely on indirect proxies to detect the presence of High-frequency traders (HFTs) (Hendershott et al., 2011).
On the other hand, the study of the 2010 Flash Crash comes less of a challenge as the magnitude of the event pushed the Commodity Futures Trading Commission (CFTC) and the Securities Exchange Commission to publish a joint report to investigate the events leading to the crash, thoroughly reporting the facts and relevant data.

Our literature review will now proceed by examining the most influential pieces of research pertaining to each of the two approaches, in order to gain an exhaustive understanding of the methodologies employed and compare the conclusions obtained.

2.1.1 EVENT STUDIES ON THE 2010 FLASH CRASH

The vast majority of the academic literature investigating the role of HFT in the 2010 Flash Crash takes as its core source the CFTC-SEC joint report (2010) and opens by retracing the report's description of the events on May 6, 2010.

As the report describes indeed, on that morning the U.S. markets opened for trading to unsettling news coming from Europe concerning the sovereign debt crisis, with the Euro registering a sharp decline against the U.S. Dollar and the Japanese Yen.
At 1:00 p.m. the build-up of negative market sentiment started to move prices on a number of securities on the NYSE, which started to register higher-than-average LRP trading pauses[2].
By 2:30 p.m. the S&P 500 volatility index (VIX) was up 22.5% since the opening, while selling pressure had brought the DJIA down 2.5%.
The same selling pressure on the S&P 500 ETF (SPY) and the E-mini S&P 500 future contracts reduced the buy-side market depth down 20% and 55% respectively, enlarging spreads and depressing prices.

[2] Liquidity Replenishment Points (LRPs) essentially are trading slowdowns implemented to act as a "speedbump" and dampen volatility whenever a NYSE stock's price abruptly fluctuates above a certain threshold CFTC-SEC (2010).

Yet, the actual trigger event started at 2:32 p.m., when amongst thinning liquidity and declining prices a large mutual fund complex (later proven to be *Waddell & Reed Financial Inc*.) entered a sell program to sell the equivalent of $4.1 billion in E-mini contracts to hedge against an equity position. The execution program used was a Percentage of Volume (POV) algorithm set to disregard price and time to ensure immediacy, which resulted in the order quickly being discharged in only 20 minutes, gravely adding pressure onto the market.

This selling pressure was initially absorbed by HFTs, fundamental buyers[3] in the futures market, and cross-market arbitrageurs who transferred this pressure to the equities market by buying E-mini contracts while selling S&P 500 index products like the SPY.

However, since HFT strategies involve trading a high number of securities with very short holding periods, maintaining no substantial long or short position, in the four minutes spanning from 2:41 p.m. to 2:44 p.m., HFTs aggressively started to sell E-mini contracts in the attempt to decrease their accumulated long positions, which with a lack of demand from fundamental buyers, were bought back and forth by other HFTs. As a result of this "hot potato" effect, 27.000 contracts were traded in a time span of 14 seconds, with prices and liquidity plunging as these rapid movements scared fundamental investors away from the market. In the meantime, the activity of cross-market arbitrageurs intensified, exacerbating the downward pressure on prices across more than 300 securities, which in turn triggered many retail stop-loss orders creating a negative feedback loop.

The result of these events at the height of the crash brought both the E-mini and the SPY down more than 7% from their opening prices, while the DJIA experienced a record intraday loss of 1.000 points (-9%), amounting to an estimated worldwide loss of $1 trillion.

The markets finally took a breath at 2:45 p.m. when trading on the E-mini paused for 5 seconds when the CME Stop Logic Functionality[4] was triggered to prevent further price declines. During the pause, the selling pressure on the E-mini significantly decreased, and buy interest resumed, which led prices to stabilize once trading resumed and eventually recover, later followed by the SPY and the DJIA.

[3] CFTC-SEC (2010) defines fundamental buyers and sellers as:" market participants who are trading to accumulate or reduce a net long or short position" to gain long-term exposure to a market or to hedge existing exposures in related markets.

[4] "Stop Logic functionality is a CME Group proprietary functionality that serves to mitigate artificial and disruptive market spikes which can occur because of the continuous triggering of stop loss orders in an illiquid market condition." (CME Group, 2016)

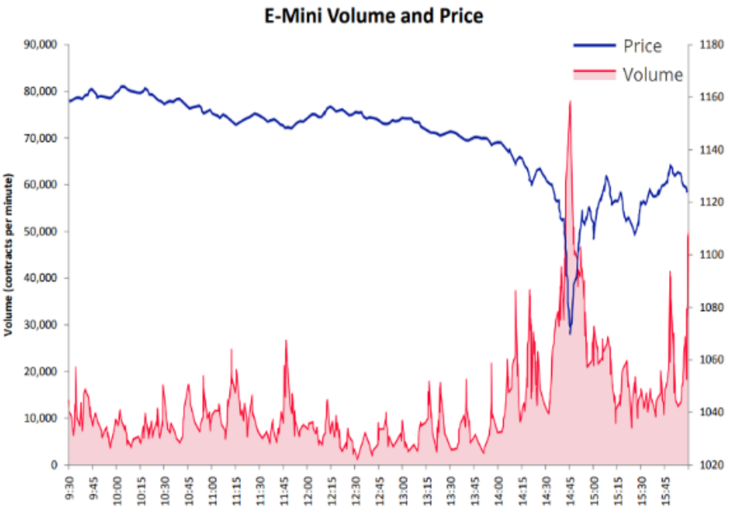

Figure 2.2a - E-mini S&P 500 volume and price graph on May 6, 2010. Graph re-adapted from CFTC-SEC (2010)

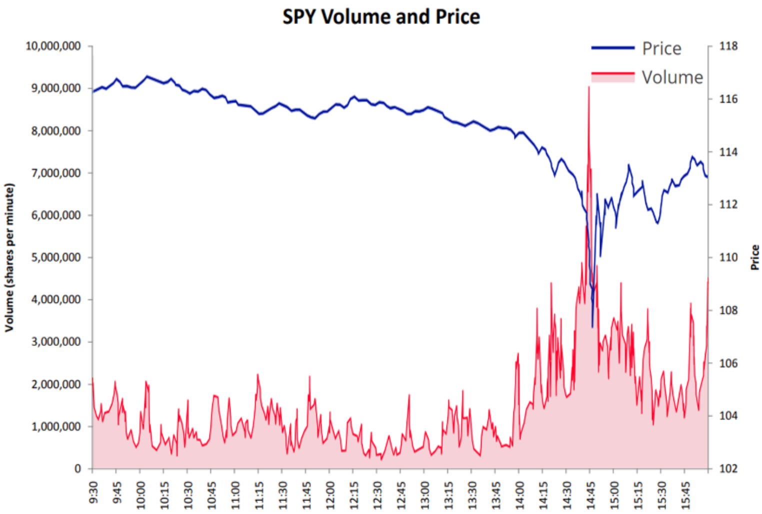

Figure 2.2b - S&P 500 volume and price graph on May 6, 2010. Graph re-adapted from CFTC-SEC (2010)

The CFTC-SEC report describes in detail how the markets moved through the analysis of prices, volatility, traded volumes, and order imbalance in the days prior and on the day of the crash, listing a number of enabling factors. However, the sole objective of the report was to explain and communicate the events of the crash and raise talking points for regulators to prevent such events from happening again, without pointing any fingers at specific direct causes.

Thus, in the following years, numerous scholars focused their research on exploring the impact of HFT on market microstructure, investigating whether events like the 2010 Flash Crash could be caused by it and whether such incidents might become a new norm.
The two most accredited theories were brought forward in a series of research papers by Easley et al. (2010a, 2010b) and Kirilenko et al. (2014, 2017) respectively, which share slightly different approaches to explain how HFT might have adversely influenced liquidity in a moment of market stress exacerbating the crash, although resting on similar assumptions.

Kirilenko et al. (2014) examine transaction-level "audit-trail" data for June's 2010 E-mini S&P 500 futures contract[5], for the four trading days spanning from May 3-6, 2010, provided by the CME to the CFTC.

[5] Kirilenko et al. (2014) explain that the E-mini S&P 500 futures contracts come with four possible expiration dates: March, June, September, or December of each year. Consequently, the contracts that attract the highest trading volumes are usually those with the closest expiration date, named the "front month" contracts. The choice of June contracts data for the analysis is then explained by June being the front month at the time of the crash.

The paper begins by carefully categorizing the market participants as either High-Frequency Traders, Market Makers (MMs), Fundamental Buyers, Fundamental Sellers, Opportunistic Traders, or Small Traders based on each account's trading pattern observing factors as traded volume, intraday, and end-of-day inventory balance.

The authors then address the issue of order book imbalance, explaining how the 75,000 contracts sell program resulted in a spike in order volume as the market "digested" the imbalance created. This spike occurs as the entrance of a large order in the market generates a cascade of intermediation trades in which traders sharing different strategies and holding periods (e.g. trend following, mean reversion) pass contracts around until the imbalance is absorbed and a price adjustment is achieved. Regressing second-by-second changes in the inventories of market makers and HFTs during the crash, the paper finds this consistent with the assumption made in Huang and Wang (2008), which explains how a large order imbalance can trigger a market crash even in the absence of fundamental shocks as it overwhelms the limited risk-bearing capacity of intermediaries. The main finding is indeed that during the normal course of business, HFTs exploit their latency advantage over MMs by aggressively trading contracts at the best bid/ask and then placing new limit orders ahead of other MMs by a tick size (similar to the sub-penny jumping strategy illustrated in chap. 1.2.2).
This practice, defined as immediacy absorption[6], tends to accelerate price changes (although not resulting in directional price moves under normal market conditions), placing an immediacy absorption cost on all slower traders who are not fast enough to cancel their orders in the event of an imminent price move. This is confirmed by the inventory data regression observed in the paper, which shows that HFTs during the crash anticipated and traded in the direction of the price move while slower MMs got run over, trading after the price move. This enhanced cost on slower liquidity providers leads MMs to decrease their market presence by reducing their acceptable inventory capacity to levels that are too small to eventually offset temporary order imbalances as during the E-mini crash.

The paper's final take is that HFT did not cause the Flash Crash of 2010. HFTs did not opportunistically change their trading patterns in any way, but rather just increased their activity to adjust to the increasing trading volumes in response to a large order imbalance. However, their aggressive trading behavior with immediacy absorbing orders accelerated the price decline, pushing MMs and fundamental traders away from the market, and vanishing liquidity. The subsequent five-

[6] The term derives from the liquidity provision activity of market makers, which through limit orders stand ready to transact at any moment, providing immediacy to market takers.

second trading pause triggered by the CME Stop Logic Functionality at 2:45 p.m. eventually allowed fundamental buyers and MMs to re-enter the market and provide liquidity, as the now lower prices would more than compensate the risk tied to the market-making activity.

Easley et al. (2010b) on the other hand, takes on a more structural approach to the analysis of the crash, looking at this event not as an isolated incident, but rather as symptomatic evidence of the change financial markets' microstructure has undergone with the advent of low-latency forms of trading.

The paper indeed begins by assessing how markets' microstructure has fundamentally changed since the advent of HFT, addressing a number of criticalities stemming from it.

First, due to the intricacies of HFT, the main players in this realm are usually a low number of highly specialized firms, which preeminently engage in market-making strategies, rapidly trading a very high number of securities and capturing spreads. Coupled with a heavy decrease in the number of retail investors after the market downturn of the global financial crisis of 2008, the result obtained is that the majority of the volume of securities traded today has been largely concentrated in the hands of a small number of individuals, which are mostly responsible for the liquidity provision.

The paper addresses this as a critical change in the market microstructure, since unlike Designated Market Makers (DMMs)[7], HFTs engaging in market-making strategies are typically under no obligation to provide liquidity to both sides of the markets at all times, with the possibility of turning into market-takers when it suits their strategy. According to the paper, this means that HFTs are usually willing to provide liquidity when the market is not moving directionally and the probability of incurring a loss on their position is low. Conversely, when the market starts to be permeated by more informed traders[8] who trade directionally based on the information they possess, an order volume imbalance will emerge, moving prices against the positions of HFTs and DMMs, which will start to accumulate or lose inventory on the wrong side of the market and incur losses.

[7] "A designated market maker (DMM) is an intermediary who has been contracted by a trading venue to stand ready to trade a financial security or contract against its own inventory. While other market participants, known as market makers (MMs), may also carry out this role on a voluntary basis, DMMs do so in a formal capacity and in a contractually agreed way." (Bank of England, 2012)

[8] By "informed trader" is intended a type of trader who has access to relevant new information (or can process new information faster than other traders) and consequently trades directionally on the basis of this new information, See Easley et al. (2012, p.12-13) for more.

If the losses become too big, eventually HFTs will start trading directionally aggressively or leave the market altogether, while DMMs will resort to stub quotes to avoid further losses[9], vanishing liquidity, and exacerbating the imbalance with the consequent directional move of the market.
This theory was first developed and published by the paper's authors in 1996 in the "Journal of Finance", with the intent of explaining how a high level of "order flow toxicity" can be predictive of the risk of a liquidity crash.
Order flow is considered to be "toxic" when it is generated by a high number of informed traders, who are likely to trade directionally upon price signals and generate order imbalances. When this happens, as we said, market-makers incur the risk of providing liquidity at a loss and are thus likely to leave the market. For this reason, in Easley et al. (1996) the authors developed the Probability of Informed Trading (PIN) measure, later re-elaborated in Easley et al. (2010a) as Volume-ynchronized Probability of Informed Trading (VPIN) to better fit the high-frequency environment, with the intent of providing a reliable measure of order flow toxicity.
The model was then put to the test in Easley et al. (2010b), observing the VPIN values in the E-mini S&P 500, during the weeks before and after the Flash Crash of 2010. The observations see the VPIN metric reach drastically high levels in the week preceding the crash, to then reach its highest levels recorded to date on the E-mini during the crash as illustrated in *Fig. 2.3*.

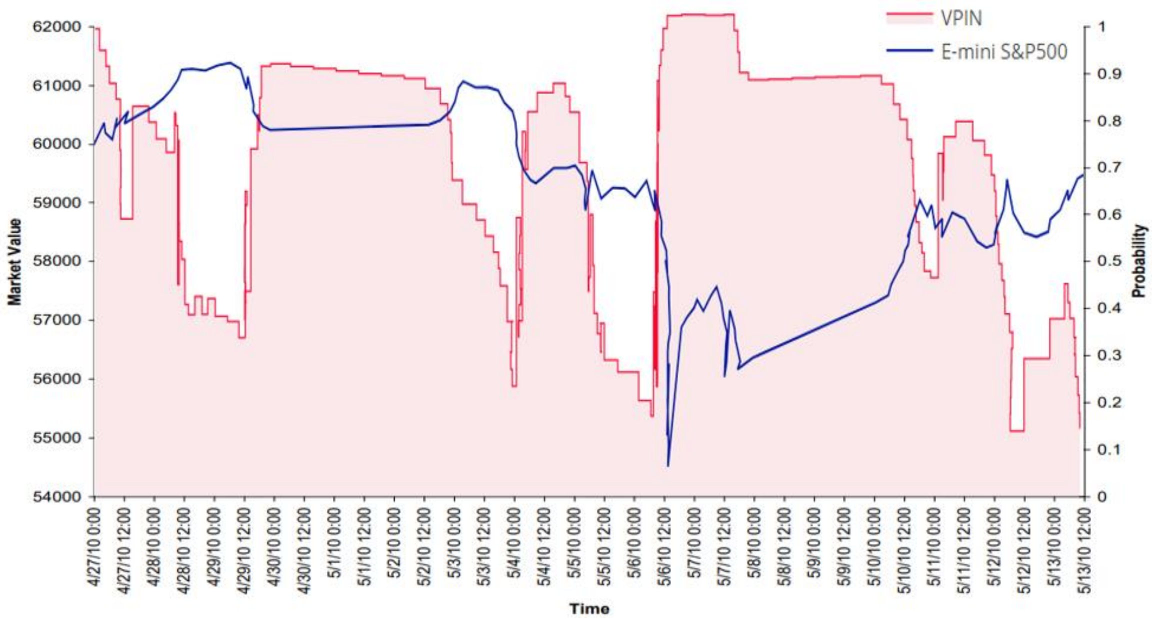

Figure 2.3 – Graphical visualization of the interaction between the VPIN metric and price on the E-mini S&P 500 one week before and after the Flash Crash. Graph re-adapted from Easley et al. (2010b).

[9] Stub quotes are: "quotes at unrealistically low or high prices that fulfill a market maker's obligation to provide continuous bids and offers, but at levels that the market maker does not expect to be reached under ordinary market conditions." (SEC, 2010, p.38), for more see SEC (2010, p.63-64)

These historical observations would indeed seem to confirm that the VPIN metric may be useful to effectively anticipate the unfolding of a liquidity scenario like the one we have previously discussed. This result seems consistent with the assumption that extreme levels of order toxicity can indeed turn liquidity providers into liquidity consumers. Nonetheless, the authors are careful in concluding that the VPIN can be looked at as a predictor of a liquidity-induced crash, rather pointing out that at relatively normal levels this measure simply works as a measure of order flow toxicity. On the other hand, at abnormally high levels, this metric may be also able to give us an indication of the likelihood of a liquidity-induced crash.

The paper additionally investigates the properties of the VPIN metric, by computing its correlation with price volatility on the E-mini, obtaining a value of 0.1596, therefore implying a positive but moderate impact of order toxicity on price changes.

The results of this research give the authors reason to conclude that the 2010 Flash Crash was a liquidity crash retraceable to an extreme surge in order flow toxicity, which caused MMs and HFTs to turn into liquidity consumers or exit the market to avoid providing liquidity at a loss to better-informed traders. The article additionally presents a measure for order flow toxicity, the VPIN metric, which from historical observations seems to be a reliable warning sign for liquidity events when extreme values are reached.

The literature reviewed in this section investigating the event of the 2010 Flash Crash provides us with useful insights into how HFT has influenced markets' microstructure, deeply shaping the dynamics of our financial system, bringing new possibilities, but some challenges as well.

This next section will instead focus on pieces of academic literature studying HFT outside of the context of the Flash Crash, directly investigating whether its impact on overall market quality is beneficial or damaging.

2.1.2 METHODOLOGY ANALYSIS OF EMPIRICAL STUDIES ON HFT

As noted earlier, conducting empirical research on HFT poses a significant challenge for researchers, influenced by several factors. Firstly, this is due to the relatively recent nature of HFT and the speed at which these technologies have evolved over time, making this phenomenon challenging to frame. But most importantly, this is because oftentimes the data required to conduct empirical studies on HFT is either scarce, not publicly available, or costly to obtain.

SEC (2014) particularly stresses this point, pointing out in their literature review how all the empirical studies available on HFT had to rely on only two possible types of datasets.
The first type is direct datasets in which the trading activity traced is directly flagged as either HFT or non-HFT by the institution providing the datasets (usually exchanges or single trading firms).
For instance, in this category, we find the CFTC-SEC datasets covering the Flash Crash we mentioned in the previous section. Another widely used example of direct dataset, available only to selected researchers, are a group of NASDAQ datasets directly tracking the aggregated activity of 26 HFT firms over a total of 120 U.S stocks, evenly split into big-cap, mid-cap, and small-cap. These datasets provide tick order book information[10] related to best bid and ask prices, trade durations, order cancellations, and order-to-trade ratios, categorizing orders as either passive (liquidity-providing) or aggressive (liquidity-taking). However, one presented limitation of the NASDAQ datasets is that they leave out a sizable amount of HFT activity deriving from proprietary desks of broker-dealer firms, or from smaller firms who rely on bigger firms for their market access, as these activities are not clearly quantifiable.

Conversely, the second type of datasets employed by researchers generally contain a wide range of data used for the manual identification of HFT activity, through the construction of coherent indirect proxies.
The use of proxy datasets to study HFT is a relatively popular choice among researchers, as it provides a broader range of options to choose from in collecting the data, adapting the proxies employed based on the data available to the researcher. At the same time, two major challenges arise from conducting proxy-based studies on HFT: first, proxy-based studies require initial data to construct the proxies, and then additional data to test the hypotheses made, making the finding of complete datasets harder. Second, the choice of proxies is critical, as they must coherently align with the assumptions being tested. Proxy studies also incur the possibility of obtaining conflicting results among similar studies as different proxies may identify HFT activity in different ways, creating inconsistencies.

Among the literature employing the indirect approach, the three most prominent methodologies used to identify HFT are intraday inventory management methods, lifetime of order methods, and message traffic methods (ESMA, 2014).

10

Intraday inventory management aims at spotting HFTs by observing single accounts' intraday trading volumes and changes in inventory holdings, paying particular attention to end-of-day inventories and positions held overnight. An account may indeed be flagged as a HFT when it shows higher-than-normal trading volumes, with inventories never particularly committed to one specific side of the market, and fluctuating net holdings often crossing the zero line[11].
End-of-day inventories and overnight positions, instead, are screened under the assumption that HFTs usually have neutral positions by the end of trading and rarely keep any positions overnight.
This identification method was used for example to categorize market participants and single out HFT accounts in Kirilenko et al. (2014). This method is viewed as rather reliable, although it requires large amounts of data to be employed and predominantly works to identify market-making strategies by HFTs, possibly leaving other HFT strategies unnoticed.

Lifetime of order methods detect HFT by measuring the time elapsed between an order submission and its modification or cancellation from the order book. This identification method is based on the ability of HFTs to not only trade fast, but also modify, cancel, or resubmit their orders faster than other traders. Based on this concept Hasbrouck & Saar (2013) develops the "strategic runs" measure, which captures HFT activity whenever a sequence of at least 10 linked order book messages (e.g. submissions, cancellations, executions), is sent consecutively in less than one second during a 10-minute interval. Resultantly, this metric identifies and measures the share of HFT activity out of the total trade activity in a 10-minute interval. Hasbrouck and Saar (2013) uses the results of this measure to study the impact of HFT overall market quality, finding evidence of a positive effect on liquidity provision and of a consequent decrease in short-term volatility.

Message traffic methods build proxies to identify HFT based on the number of order book messages sent, relying on empirical evidence showing that low latency trading commonly implies frequent posting, modification, and cancellation of multiple orders at very high speeds. Building proxies based on message traffic is a popular methodology as it results easier to measure than order lifetimes while still providing valuable insights. This approach has been indeed widely employed by academics, industry bodies, and regulators to study both AT and HFT. For instance, the 2013 German "High-Frequency Trading Act" used message traffic as one of the metrics to identify HFT, enforcing the law on all firms generating more than 2 messages per second, or 75,000 per trading day (BaFin, 2013).

[11] Crossing the zero line means going from net long positions to net short positions or vice-versa. For more see Kirilenko at al. (2010) figure 4 at page 39.

Examples of the most common message traffic proxies used include Order-to-Trade Ratios (OTR), Cancel-to-Trade Ratios (CTR), and Orders-per-minute (O/min) metrics.

2.2 TAKEAWAYS FROM THE LITERATURE REVIEW

From the available literature delving into the empirical study of HFT, it is rather difficult to achieve a unanimous conclusion about its effect on the markets. Easley et al. (2010a, 2010b) and Kirilenko et al. (2014, 2017) seem to agree that while the Flash Crash of May 2010 was not caused by HFT, it was exacerbated by HFTs shifting from liquidity providers to liquidity takers during a period of heightened market uncertainty. This shift consequently overwhelmed designated market makers, pushing them beyond their risk-bearing capacities and leading to a liquidity-induced crash.

At the same time, Hasbrouck & Saar (2013) with their "strategic runs" measure show evidence that low-latency trading tends to improve market conditions enhancing liquidity and reducing short-term volatility, even during moments of market stress like during the financial crisis of 2008.

Brogaard (2010) on another note, uses a NASDAQ direct sample to reach positive although mixed conclusions, proving that HFT positively adds to the price discovery process and volatility reduction, while seeming to provide liquidity only at a superficial level, providing less book depth than regular traders.

Most of the academic literature seems indeed to generally evaluate the phenomenon of HFT as beneficial for the overall market quality during stable market times while presenting some criticalities during moments of instability. However, the extent to which these features hold true is still uncertain, and the academic debate is still open for further discussion.

In light of this, the next chapter will aim to add to the existing research, by conducting an empirical study of our own.

EMPIRICAL STUDY ON HFT

The analysis conducted in our literature review presented the methodologies and findings of the most cited pieces of literature about HFT, comprehensively comparing the possible different approaches employed to empirically study its impact on the overall market quality.
However, although academic attention is starting to pick up, the limited amount of research conducted on these issues so far, and the variety of approaches available, fail to provide a clear unanimous conclusion.
For this reason, this chapter will be aimed at conducting an empirical study of our own in the attempt of adding to the academic literature and test our assumptions first-hand.

3.1 COLLECTING THE DATA

To do so, it is essential to begin by gathering precise and reliable data to construct our analysis.
As anticipated by several of the works cited, this aspect of the research process is the most critical and difficult part. An ideal dataset for the study of HFT should indeed contain granular daily (or preferably intraday) trade information, coherent with the scope of the observations carried out, and provide useful data for the construction of proxies employed to flag HFT activity.
The fragmented nature of modern electronic financial markets presents an additional challenge, as nowadays most stocks trade across multiple venues. Hence, an ideal dataset should also account for fragmentation, collecting trade data from multiple exchanges to ensure completeness of information, as HFT activity might not be homogenously distributed across exchanges.

3.1.1 DESCRIPTION OF THE SAMPLE

After extensive research to locate the ideal data source for our objectives, we gain institutional access to the SEC Market Information Data Analytics System (MIDAS) through the Wharton Research Data Services (WRDS) database.
We identify the “Metrics by Individual Securities” dataset, which contains daily trade information for over 8.300 stocks and ETFs gathered across 12 different U.S.-based exchanges, updated quarterly.
The dataset contains 26 different metrics for each security, including several useful metrics to construct proxies for the identification of HFT activity (WRDS, 2019).

Since the world of trading algorithms is rapidly and constantly evolving, we strongly value the freshness of the data we collect to ensure that our results are up to date. For this reason, we collect the most recent sample attainable from our database, which is currently updated up to December 31st, 2021, collecting the data related to the previous trading month starting on November 30th, 2021.
Our sample traces trade order information for 1,519 randomly selected stocks for the 23 trading days of December, listing each stock's daily traded volume, number of order cancellations, actual trades executed, and computes the Cancel-to-Trade Ratio (CTR). For each stock, the sample also provides daily updated volatility deciles computed comparing the stock's daily volatility with the rest of the market.

This initial sample perfectly fits our needs as it is large[12], granular, comprehensive of data gathered from multiple exchanges, and contains useful message traffic metrics we can employ to imply the presence of HFT on that particular stock.
Our next step will then be to start filtering and cleaning the data in order to prepare it for our observations.

3.1.2 CLEANING THE DATA

The first step of our data-cleaning process will be simply to filter out all those stocks that present missing observations on any given day, to ensure proper comparability between observations. We proceed to remove a total of 78 stocks for which the observed values do not cover the entire 23 trading days sampled, leaving us with 1,441 stocks remaining.

Since our goal is to create an environment to study HFT, we filter our data to better reflect the characteristics of a HFT sample. As we noted earlier, 60-75% of the trading volume in U.S. markets is generated by algorithms, thus we can start by assuming that our sample of randomly chosen U.S. stocks intrinsically captures AT activity to an extent. In order to capture HFT as a particular subset of this AT activity, we can follow the premise that one of the distinctive characteristics of HFT that separates it from AT is its focus on highly liquid instruments (Gomber et al. 2011). To do this we can use traded volume as a measure of liquidity focusing our observations on the stocks with the highest average trading volumes, under the assumption that they will be the ones with the highest share of

[12] Hasbrouck and Saar (2013) for instance use a sample of 351 stocks from the S&P500 Index, Hossain (2022) uses 149 LSE stocks, Brogaard (2010) and Brogaard et al. (2012) use a direct NASDAQ dataset of 120 stocks.

HFT activity. This assumption is also reinforced by the fact that HFT entails trading securities at high volumes, making those securities on which HFT activity is present more likely to show higher-than-average trading volumes.
To observe these stocks, we will then filter the sample for volume by computing each stock's average daily traded volume, using only the ones belonging to the highest quartile (AVG_vol > 75th percentile). This step additionally narrows our sample down to the 360 most liquid stocks in our sample.

At this point, we compute our proxy measure of HFT, the average Cancel-to-Trade Ratio (AVG_CTR), averaging out the daily CTR on each of the 23 days forecasted for every stock in our remaining sample.
We then visualize this data to better observe its properties and eventually identify and remove outliers that might tamper with the accuracy of our analysis.
For this purpose, we use a boxplot to visualize each stock's AVG_CTR to see if there are any extreme values worth noting.

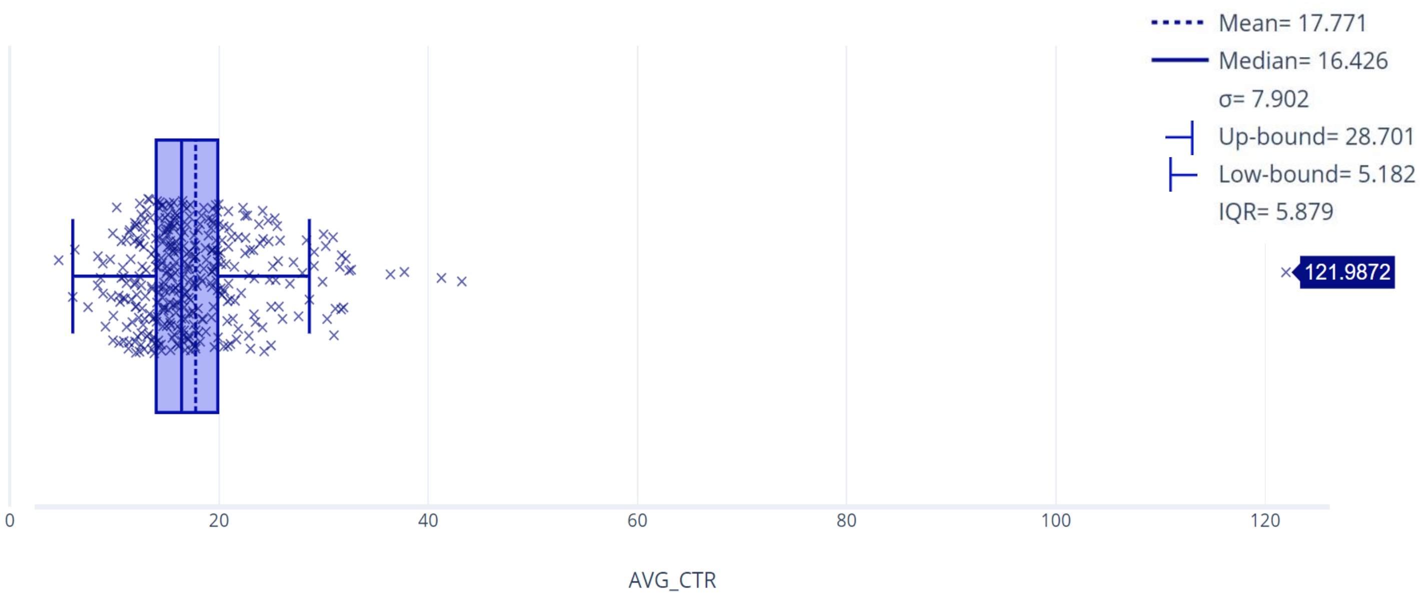

Figure 3.1 – Boxplot of the 360 stocks' AVG_CTR showing upper and lower bounds according to the Inter-Quartile Range (IQR) method to expose extreme values.

The graph indeed shows the presence of some extreme values outside of the upper and lower bounds calculated with the IQR, with one particular outlier standing at more than 13 standard deviations from the mean. This value, relative to the stock "Gran Tierra Energy Inc." (GTE), although seemingly not erroneous from closer inspection, is too abnormally distant from the rest of the values to be kept in as it might bias the results of our analysis, thus we decide to take it out of the sample.

Our final sample of stocks is now identified, containing detailed data for the 359 most liquid stocks out of the 1519 from the initial sample.
We now proceed to compute the summary statistics for the sample, reported in *Table 3.1*.

N = 359	Average CTR (AVG_CTR)	Average Volatility Decile (AVG_Vol_Dec)	Average daily traded volume	Average daily # of trade messages	Average daily order cancellations
Mean	17.481	5.385	1,919,695.794	20,791.823	360,326.704
Median	16.404	5.391	971,086.696	13,869.087	211,718.000
Std	5.619	2.668	3,442,345.707	30,242.334	530,358.854
Lowest value	4.687	1	431,088.087	1,795.913	9,087.870
Highest value	43.210	10	39,110,093.43	432,981.870	7,064,550.348
Skewness	1.242	0.049	6.678	8.354	6.841
Kurtosis	2.481	-1.423	57.325	100.537	73.628

Table 3.1 – Summary statistics for the 359 stocks in the sample.

Our summary statistics for the average CTR, volatility deciles, and average daily traded volume provide us with key insights into the dataset's characteristics. The average CTR with a mean value of 17.481 and a standard deviation of 5.619 suggests moderate variation across the stocks in the sample, with a slight right skewness (1.242), indicating a tendency to reach higher values in some instances. This means that on average, for each trade that goes through, an average stock in our sample experiences around 17 order submissions and consequent cancellations.
The volatility deciles on the other hand, averaging 5.385, exhibit minimal skewness (0.049), suggesting a fairly balanced distribution of stocks' volatility in our sample, although the negative kurtosis (-1.423) indicates a slightly flatter distribution than a normal curve.

3.2 DATA ANALYSIS

Now that the sample data is ready, we can start building our statistical model for the analysis.
Our goal will be the one to empirically show whether the degree of HFT activity on a particular stock is tied to a decrease in its short-term volatility in a statistically significant way.

To do so we use the average CTR measure as a proxy for HFT, a measure often used by researchers (e.g. Hasbrouck and Saar, 2013; Karkowska & Palczewski, 2023) in alternative to the Order-to-Trade Ratio (OTR) to imply the presence of low-latency traders as their strategies are characterized by a large number of rapid order submissions and frequent cancellations.

3.2.1 HYPOTHESIS STATEMENT

The academic literature reports that with the increase of traders' speed the average number of order cancellations in the market is constantly increasing (Karkowska & Palczewski, 2023). We then use this notion to imply that an increase of the average CTR for a security must be reflective of an increased share of HFT activity on that instrument. Hence, in order to assess whether an increase in HFT activity is followed by a decrease in short term volatility, our objective will be to test if it exists a linear relationship between each stock's average daily volatility and average daily CTR.

If such a relationship exists, the regression line we expect to see should have the following form:

$$Y_i = \beta_0 - \beta_1 \times avg\_CTR_i + \epsilon_i$$

Where:

- Y_i is the daily average volatility percentile of stock i, the dependent variable;
- avg_CTR_i is the daily average Cancel-to-Trade Ratio of stock i;
- The slope coefficient β_1 is negative, proving the inverse linear relationship we are trying to demonstrate.

We then begin investigating this relationship by formalizing our null and alternative hypothesis about the slope coefficient as follows:

$$\begin{cases} H_0: \beta_1 = 0 & \rightarrow \text{ Null hypothesis: a stock's avg\_CTR does not influence its short-term volatility} \\ H_1: \beta_1 < 0 & \rightarrow \text{ Alternative hypothesis: avg\_CTR negatively affects short-term volatility} \end{cases}$$

If our hypothesis holds, we should be able to reject the null hypothesis, thus confirming that an inverse linear relationship exists between the implied presence of HFT activity and short-term volatility.

3.2.2 HYPOTHESIS TESTING

To test our hypothesis, we use the Excel regression function to estimate the slope and intercept of our line of best fit, for which we obtain the following estimations:

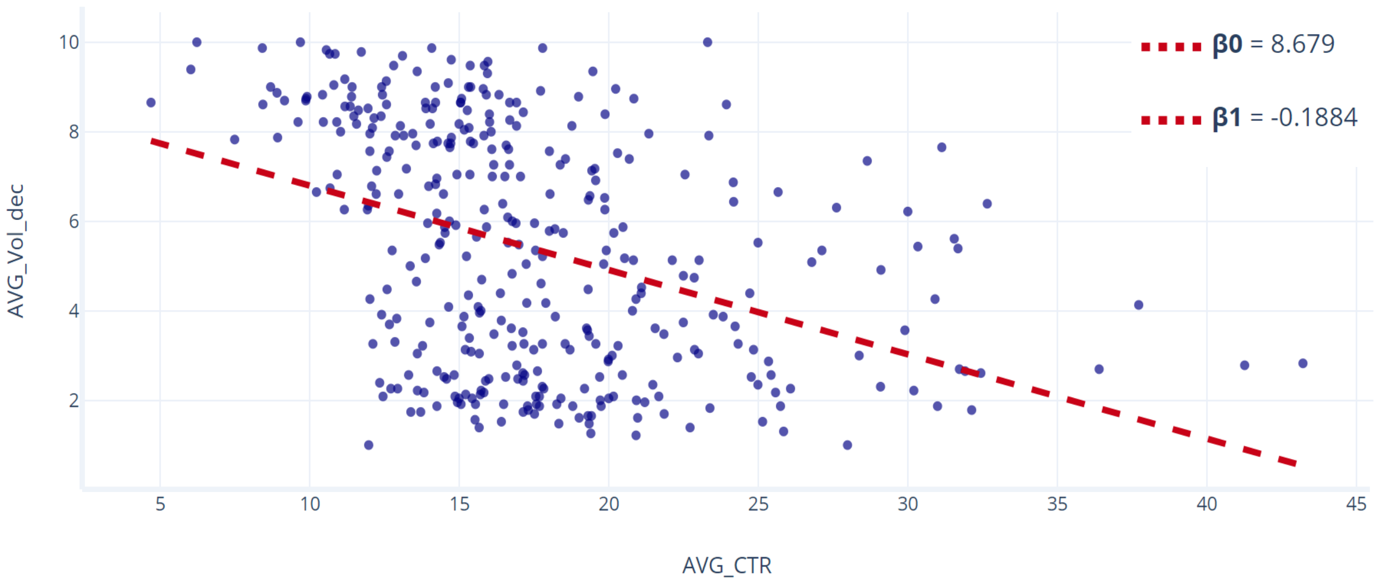

Figure 3.2 – Scatter plot of the 359 stocks in the sample showing the estimated regression line of best fit.

Hence, the regression line has the following equation:

$$Y = 8.679 - 0.1884 \times avg\_CTR$$

As we can observe, the value of the slope coefficient β_1 is indeed negative as expected, thus our goal now will then be to test if this value is also statistically significantly smaller than zero.

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept - β_0	8.678801791	0.421443878	20.59301899	1.15556E-62
AVG_CTR	-0.188425252	0.022952505	-8.209354621	4.10438E-15

Table 3.2 – Estimates of the regression coefficients with respective P-values.

The regression output shows a very large t statistic, from which we obtain a near-zero P-value, hence, our estimate of the regression slope is statistically significant at any confidence level, so we can thus reject the null hypothesis.

The model also reflects a correlation coefficient (R) of -0.39849, implicating a moderately large negative correlation of almost 40% between the CTR of a stock and its daily volatility decile.
These two results together seem to confirm our inference that as a stock's CTR increases (and hence the probability of HFT activity), its short-term volatility decreases to some extent.

We now observe the following table containing the summary statistics for our regression. Our parameter of interest in this instance is thecoefficient of determination R^2 found below the previously mentioned R. The R^2 coefficient tells us how much of the variance of our dependent variable (AVG_Vol_dec) is explained by the independent variable (AVG_CTR) in our model.

Regression Statistics	
Multiple R	-0.39849663
R^2	0.158799564
Adjusted R^2	0.15644326
Standard Error	2.443668795
Observations	359

Table 3.3: Summary statistics for the regression

A coefficient of 0.158 implies that the CTR explains nearly 16% of the volatility variance, which is not very substantial, revealing that CTR has only a marginal impact on a stock's volatility.
However, from an economic standpoint, it is not an indifferent result: a stock's volatility is influenced by a vast multitude of factors both fundamental and environmental, thus even a small degree of influence by our CTR measure can be viewed as significant for our scope of research.

3.2.3 EVALUATION OF THE LINEAR MODEL

Before reaching any further conclusion, we first want to check the validity of our linear model to see how well it fits our data and its reliability in predicting future values.
To do this we will have to observe other outputs of the regression and test the main assumptions at the base of our model, with particular attention to the homoscedasticity, and normality of the residuals.

The independence of observations assumption can be taken as built into our sample, as we can confidently assume that a stock's average daily volatility and CTR are not influenced by the ones of other stocks in the sample.

The linearity assumption can be checked graphically by observing the plot provided in *Fig. 3.2* and the scatter plot of residuals in *Fig. 3.3.*

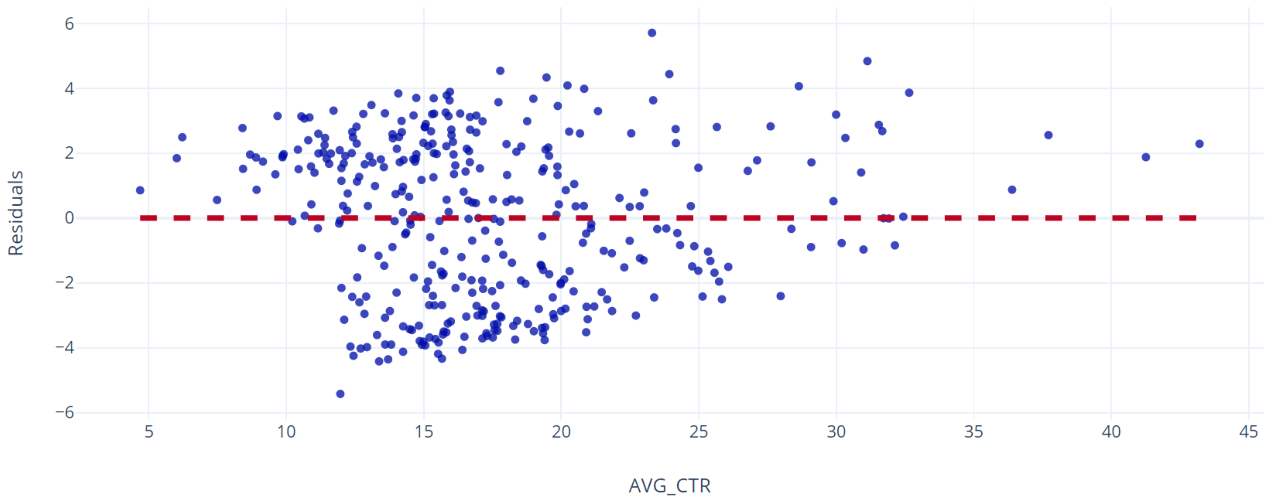

Figure 3.3 - Plot of the regression residuals against the independent variable AVG_CTR.

The plot clearly shows a random distribution of residuals around zero with no particular patterns, hence we can infer that the linearity assumption is satisfied.

For the normality assumption of residuals, on the other hand, we graphically check employing a quantile-quantile plot made by computing the normal distribution theoretical quantiles of our residuals and then plotting the result against the observed sample quantiles (*Fig. 3.4*).

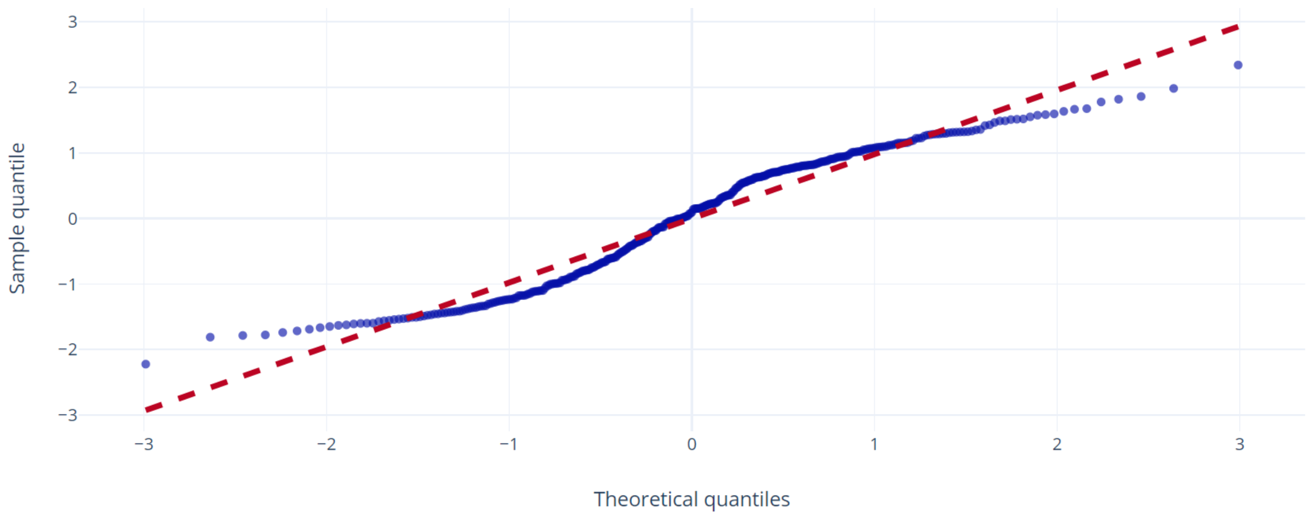

Figure 3.4 - Q-Q plot of the regression residuals.

The Q-Q plot shows a symmetrical and approximately straight distribution of the residuals, with some slight deviations along the tails, implying a slightly flatter distribution with thinner tails. We then decide to further check by computing the distribution's Kurtosis, which gives us a result of -1.169, confirming our previous statement while still being in the acceptable range of ±2 (George & Mallery, 2016). Hence, we can consider the normality condition satisfied.

Finally, we check for heteroscedasticity, for which we can observe again the residual plot in *Fig. 3.3* to check for any irregular pattern. The plot shows no particular deviation in the residual configuration, thus pointing to a homoscedastic distribution. We however decide to test directly for better identification using the Breusch-Pagan test. To do this we compute the square of our regression residuals, and we regress them against the independent variable AVG_CTR. We then compute our test statistic by multiplying the R^2 coefficient (0.001919) of this second regression for the number of observations (simply the number of stocks in the sample =359), obtaining a value of χ^2=0.689094. We then compute the P-value for the chi-square test statistic at 2-1 degrees of freedom[13] obtaining a value of 0.406473 > 0.05, hence not significant to reject the null hypothesis that our regression residuals are homoscedastic.

3.3 RESULTS AND FINAL REMARKS

In light of these results, we can conclude that our linear regression model validly fits our data and can thus correctly give us an insight into the relationship that exists between a stock's CTR and its daily average volatility.
The relationship we found was strongly statistically significant for the slope parameter, although the R^2 coefficient implied only a weak dependence of the dependent variable's variance on the dependent variable. This evidence is however not surprising and does not discredit our analysis, as a stock's price changes are influenced by a vast multitude of factors outside of the scope of our model.

From the analysis conducted, we can indeed conclude that there seems to be an inverse relationship that ties the amount of HFT activity (implied using the CTR proxy) and the daily volatility decile that the stock falls into. This study however does not include in its scope parameters such as liquidity or market depth, and thus does not draw any more generical conclusions over the impact of HFT on the overall market quality. For this reason, our findings are not to be interpreted as a direct causal link between HFT and reduced volatility, but rather as proof of strong ties between the two, since as mentioned by Biais and Foucault, (2014) causality in this realm is quite hard to prove, as models can

[13] The degrees of freedom for computation of the P-value of a chi-square distribution have value n-1, where n is the number of parameters in the auxiliary regression, in this case 2.

often incur endogeneity problems caused by the existing ties between changing market conditions and opportunistic strategies employed by HFTs.

The focus of this analysis was rather to shed some light over the widely debated effects of HFT on short-term market volatility, as events like the 2010 Flash Crash among several others often brought this matter into question and split the academic community with conflicting results.

Reconnecting the overall theoretical interpretation of our results with the literature reviewed in *chap. 2*, we may imply that the decrease in latency in the quote posting process resulting from the activity of HFTs contributes to reducing volatility by improving quote efficiency, helping securities to converge towards their mean value. The same effect, however, can conversely enhance the speed at which price corrections happen during moments of market distress, quickly deteriorating the positions of slower market makers and increasing the cost of providing liquidity beyond their tolerance, leading to liquidity dry-ups and consequent liquidity-induced crashes like in the instance of the Flash Crash as explained in Easley et al. (2010b) and Kirilenko et al. (2014).
For these reasons, continuous academic research and debate of these topics are essential to the regulatory process, to ensure that institutional bodies have the right instruments to effectively mitigate the risks stemming from an increasingly fast-evolving computerized market.

CONCLUSIONS

Over the course of the last 50 years, our financial system has experienced an unprecedented technological evolution that completely revolutionized the way capital is allocated in our society.
This revolution has completely reshaped the way securities are traded, moving this process from the exchanges' trading pits to digital screens.
In the wake of this, trading algorithms started to make their appearance on the markets, quickly being adopted by large institutions to increase efficiency and reduce the cost of their trades.
As these algorithms represent today about 60-75% of the total traded volume in the world's largest markets, a fundamental understanding of their functioning and effects are crucial.

This thesis was indeed aimed at investigating algorithmic trading, with particular attention to high-frequency trading.
During the course of our research, we proceeded to evidence how these practices first emerged and describe all the main strategies to understand their approach to trading. We then reported examples showing the possible threats presented by these practices, as it was seen how they may adversely affect market quality during moments of market turmoil, as in the case of the 1987 Black Monday or the 2010 Flash crash.
In our literature review, this thesis also highlighted the ongoing academic debate surrounding the overall impact of HFT on market quality, confronting the results obtained by different studies, and commenting on the different research methods used. While the studies conducted on the 2010 Flash Crash seem to agree that HFT seems to adversely affect market quality during unstable times, as HFTs tend to shift from liquidity providers to liquidity takers, Hasbrouck & Saar (2013) conversely concludes that HFT improves market conditions, enhancing liquidity and reducing volatility even during unstable times as during the financial crisis of 2008.

In light of the lack of a unanimous conclusion within the academic community, the main contribution of our research was to conduct an empirical study of our own, to assess first-hand the impact of HFT over short-term volatility.
This study was conducted by gathering trade data for a sample of 1519 randomly picked stocks, from the SEC MIDAS database, for the 23 trading days going from November 30th to December 31st, 2021.

Our sample was then filtered in order to keep only the most liquid stocks of the sample, leaving us with 359 stocks for which we computed the daily average volatility decile, and the daily average CTR, our proxy for HFT.
The result of our study showed that a stock's average CTR has a statistically significant inverse linear relationship with the stock's daily volatility, thus implying that a higher presence of HFT activity leads to a decrease in short-term volatility. Our model also showed a negative correlation of the two variables of nearly 40%, strengthening our results, while the determination coefficient (R^2) implied that only 16% of the volatility's variance was explained by the CTR. This second result, although seemingly low, does not discredit our conclusion, as the changes in a stock's volatility are determined by a large variety of factors beyond the CTR and outside the scope of our model.

One of the key insights derived from our research is the dual nature of HFT's impact on market stability. On one hand, the decrease in latency and improvement in quote efficiency driven by HFT activity can reduce volatility by enabling faster convergence of security prices towards their mean value. This effect is particularly beneficial during stable market conditions, where HFT helps to maintain liquidity and smooth out price fluctuations. On the other hand, the same mechanisms that contribute to market stability during normal conditions can exacerbate instability during periods of market stress. For example, during events like the Flash Crash, HFT firms, which typically act as liquidity providers, may quickly shift to liquidity takers to avoid losses thereby intensifying market disruptions and leading to liquidity dry ups.

In conclusion, this thesis adds to the available academic literature by providing empirical evidence on the complex relationship between HFT activity and market volatility. Our findings suggest that while HFT can play a role in reducing market volatility under normal conditions, it may also pose significant risks during periods of market instability. These insights underline the importance of continuous monitoring and regulation of HFT, to ensure that the benefits of low-latency trading are maximized while minimizing the potential risks to market stability.
It is therefore crucial to continue to investigate the relationship between HFT and market quality, as only a deeper knowledge of these practices can promote the development of sound regulatory frameworks to address the challenges posed by a continuously evolving financial ecosystem.

BIBLIOGRAPHY

Agenzia delle Entrate (2013) *Financial Transaction Tax (FTT) Decree.* Rome: Agenzia delle Entrate. Available at: https://www.agenziaentrate.gov.it/portale/documents/20143/254575/FTT+Decree_DecretoFTT_11_3_2013_EN.pdf/032b90a1-4d1e-fc89-9b51-f167e7a80b38 [Accessed 22 August 2024].

BaFin (2013) 'Information on the German High-frequency Trading Act', *Xetra*. Available at: https://www.xetra.com/resource/blob/253202/61b3ede0e339cddf80f0591b7d71925a/data/Information-of-bafin-on-hft.pdf [Accessed 1 September 2024].

Bank of England (2012) 'The role of designated market makers in the new trading landscape', *Quarterly Bulletin*. Available at: https://www.bankofengland.co.uk/-/media/boe/files/quarterly-bulletin/2012/the-role-of-designated-market-makers-in-the-new-trading-landscape.pdf [Accessed: 5 September 2024].

Biais, B. and Foucault, T. (2014). 'HFT and market quality'. *Bankers, Markets & Investors*, (128), pp.5-19. Available at: https://www.revue-banque.fr/binrepository/0000000000000000000_RB00476_MG108076969.pdf [Accessed 10 Sep. 2024].

Brady, N. (1988) *Report of the Presidential Task Force on Market Mechanisms*. Washington, D.C.: U.S. Department of the Treasury. Available at: https://dn720204.ca.archive.org/0/items/reportofpresiden01unit/reportofpresiden01unit.pdf [Accessed 5 September 2024].

Brogaard, J.A. (2010) 'High-frequency trading and its impact on market quality', *Northwestern University Kellogg School of Management*. Available at: https://conference.nber.org/confer/2010/MMf10/Brogaard.pdf [Accessed 22 August 2024].

Brogaard, J., Hendershott, T. and Riordan, R. (2014) 'High-frequency trading and price discovery', *The Review of Financial Studies*, 27(8), pp. 2267–2306. Available at: https://doi.org/10.1093/rfs/hhu032 [Accessed 1 September 2024].

Carè, R. and Cumming, D. (2024) 'Technology and automation in financial trading: A bibliometric review', *Research in International Business and Finance*, 71. Available at: https://doi.org/10.1016/j.ribaf.2024.102471 [Accessed 22 August 2024].

Chronohistoria (2022) ‘Stock Trading Before Computers: How It Was Done’. [online] Available at: https://chronohistoria.com/stock-trading-before-computers-how-it-was-done/ [Accessed 21 Aug. 2024].

CME Group (2016) Response to ESMA Consultation Paper: Guidelines on calibration, publication and reporting of trading halts, *Annex 1*. Available at: https://www.esma.europa.eu/sites/default/files/consultations/2016/12/esma_mifid_gth_cme_europe_annex1.docx [Accessed: 5 September 2024].

Commodity Futures Trading Commission and U.S. Securities and Exchange Commission (2010) *Findings regarding the market events of May 6, 2010: Report of the staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory Issues*. Washington, D.C.: U.S. Commodity Futures Trading Commission and U.S. Securities and Exchange Commission. Available at: https://www.sec.gov/files/marketevents-report.pdf [Accessed: 1 September 2024].

Easley, D., Kiefer, N.M., O'Hara, M. and Paperman, J.B. (1996) ‘Liquidity, Information, and Infrequently Traded Stocks’, *The Journal of Finance*, 51(4), pp. 1405-1436. Available at: http://www.jstor.org/stable/2329399 [Accessed: 1 September 2024].

Easley, D., López de Prado, M. and O’Hara, M. (2010a) ‘Measuring flow toxicity in a high-frequency world’, *SSRN Working Paper Series*. Available at: http://papers.ssrn.com/abstract=1695596 [Accessed: 1 September 2024].

Easley, D., López de Prado, M. and O’Hara, M. (2010b) ‘The microstructure of the “flash crash”: Flow toxicity, liquidity crashes, and the probability of informed trading’, *SSRN Working Paper Series*. Available at: https://papers.ssrn.com/abstract=1695041 [Accessed: 1 September 2024].

Encyclopedia (2022) *Chicago Stock Exchange*. Available at: https://encyclopedia.pub/entry/36073#:~:text=Founded%20on%20March%2021%2C%201882,Midwest%20Stock%20Exchange%20in%201949 [Accessed: 14 September 2024].

European Securities and Markets Authority (2014) *High-frequency trading activity in EU equity markets*. Available at: https://www.esma.europa.eu/sites/default/files/library/2015/11/esma20141_-_hft_activity_in_eu_equity_markets.pdf [Accessed: 14 September 2024].

Federal Reserve History (2013) ‘Stock market crash of 1987’. *Federal Reserve History*. Available at: https://www.federalreservehistory.org/essays/stock-market-crash-of-1987 [Accessed: 1 September 2024].

FIX Protocol Ltd. (2010) 'What is FIX?'. Available at: https://www.fixevents.com/Archives/2010/about.html [Accessed: 5 September 2024].

George, D. and Mallery, P. (2016). *IBM SPSS Statistics 23 Step by Step: A Simple Guide and Reference.* 14th ed. New York: Routledge. Available at: https://knowledgezone.home.blog/wp-content/uploads/2019/05/wp-1558033893715.pdf [Accessed 22 Aug. 2024].

Gomber, P., Arndt, B., Lutat, M. and Uhle, T. (2011) 'High-frequency trading', *SSRN Working Paper Series*. Available at: https://papers.ssrn.com/abstract_id=1858626 [Accessed: 1 September 2024].

Hasbrouck, J. and Saar, G. (2013) 'Low-latency trading', *Journal of Financial Markets*, 16(4), pp. 646-679. Available at: https://doi.org/10.1016/j.finmar.2013.05.003 [Accessed: 1 September 2024].

Hendershott, T., Jones, C.M. and Menkveld, A.J. (2011) 'Does Algorithmic Trading Improve Liquidity?', *Journal of Finance*, 66(1), pp.1-33. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1100635 [Accessed 21 August 2024].

Hossain, S. (2022) 'High-frequency trading (HFT) and market quality research: An evaluation of the alternative HFT proxies', *Journal of Risk and Financial Management*, 15, p. 54. Available at: https://doi.org/10.3390/jrfm15020054 [Accessed 1 September 2024].

Huang, J. and Wang, J. (2010) 'Market liquidity, asset prices, and welfare', *Journal of Financial Economics*, 95(1), pp. 107-127. Available at: https://doi.org/10.1016/j.jfineco.2008.08.008 [Accessed: 5 September 2024].

IMARC Group (2023) *Algorithmic Trading Market: Global Industry Trends, Share, Size, Growth, Opportunity and Forecast 2024-2032*. Available at: https://www.imarcgroup.com/algorithmic-trading-market [Accessed 5 Sept. 2024].

IIQF (2023) 'Types of Algo Trading Frequency-wise'. *Indian Institute of Quantitative Finance*. Available at: https://www.iiqf.org/blogs/algo-trading/types-of-algo-trading-frequency-wise.html [Accessed: 5 September 2024].

Karkowska, R. and Palczewski, A. (2023) 'Does high-frequency trading actually improve market liquidity? A comparative study for selected models and measures', *Research in International Business and Finance*, 64, p. 101872. Available at: https://doi.org/10.1016/j.ribaf.2022.101872 [Accessed 1 September 2024].

Kim, K. (2007) *Electronic and algorithmic trading technology: The complete guide*. 1st ed. Academic Press, Elsevier. Available at: https://www.sciencedirect.com/book/9780123724915/electronic-and-algorithmic-trading-technology [Accessed: 5 September 2024].

Kissell, R. (2014) *The Science of Algorithmic Trading and Portfolio Management*. 1st ed. Elsevier Monographs, Elsevier. Available at: https://doi.org/10.1016/C2012-0-00818-6 [Accessed 5 September 2024].

Kirilenko, A., Kyle, A.S., Samadi, M. and Tuzun, T. (2014) *The Flash Crash: The Impact of High-Frequency Trading on an Electronic Market*. Washington, D.C.: U.S. Commodity Futures Trading Commission. Available at: https://www.cftc.gov/sites/default/files/idc/groups/public/@economicanalysis/documents/file/oce_flashcrash0314.pdf [Accessed: 1 September 2024].

Kirilenko, A., Kyle, A.S., Samadi, M. and Tuzun, T. (2017) 'The Flash Crash: High-Frequency Trading in an Electronic Market', *The Journal of Finance*, 72(3), pp. 967-998. Available at: https://doi.org/10.1111/jofi.12498 [Accessed 1 September 2024].

Mahmoodzadeh, S. and Gençay, R. (2017) 'Human vs. high-frequency traders, penny jumping, and tick size', *Journal of Banking & Finance*, 85, pp. 69-82. Available at: https://doi.org/10.1016/j.jbankfin.2017.08.015 [Accessed 22 August 2024].

McGowan, M. (2010) 'The Rise of Computerized High Frequency Trading: Use and Controversy', *Duke Law and Technology Review*, 16. Available at: https://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=1211&context=dltr [Accessed 5 Sept. 2024].

NSE International Exchange (n.d.) 'Direct market access'. *NSE International Exchange*. Available at: https://www.nseix.com/technology/direct-market-access [Accessed: 1 September 2024].

Papers with Backtest (2024) *Algorithmic Trading Strategies - Glossary*. Available at: https://wiki.paperswithbacktest.com/glossary/algorithmic-trading-strategies [Accessed 22 August 2024].

Securities and Exchange Commission (1994) *Market 2000: An Examination of Current Equity Market Developments*. Washington, D.C.: U.S. Securities and Exchange Commission. Available at: https://www.sec.gov/divisions/marketreg/market2000.pdf [Accessed 22 August 2024].

Securities and Exchange Commission (2004) *International Securities Settlement Conference — "The U.S. view of the role of regulation in market efficiency"*. Available at: https://www.sec.gov/news/speech/spch021004leb.htm [Accessed: 5 September 2024].

Securities and Exchange Commission (2012) *Report to Congress on Decimalization*. Washington, D.C.: U.S. Securities and Exchange Commission. Available at: https://www.sec.gov/files/decimalization-072012.pdf [Accessed 5 September 2024].

Securities and Exchange Commission (2014) *Equity market structure literature review, Part II: High-frequency trading*. Washington, D.C.: U.S. Securities and Exchange Commission. Available at: https://www.sec.gov/marketstructure/research/hft_lit_review_march_2014.pdf [Accessed 1 September 2024].

Securities and Exchange Commission (no date) 'Market Information Data Analytics System (MIDAS)'. Available at: https://wrds-www.wharton.upenn.edu/pages/get-data/intraday-indicators-wrds/wrds-sec-midas/ [Accessed: August 2, 2024].

Securities and Exchange Commission Historical Society (2022) *Transformation & Regulation: Equities Market Structure, 1934 to 2018*. Available at: https://www.sechistorical.org/museum/galleries/msr/index.php [Accessed: 5 September 2024].

Tytell, I., Emsbo-Mattingly, L. and Hofschire, D. (2016) 'Secular outlook for global growth: 2015–34', *National Association for Business Economics*, 51(1), pp. 18-26. Available at: https://doi.org/10.1057/be.2016.5 [Accessed: 1 September 2024].

Wharton Research Data Services (WRDS) (2019) 'WRDS MIDAS Manual 1.0', *University of Pennsylvania*. Available at: https://wrds-www.wharton.upenn.edu/documents/1133/WRDS_MIDAS_Manual_1.0.pdf [Accessed: 1 September 2024].

World Bank (2023) *Market capitalization of listed domestic companies (current US$)*. Available at: https://data.worldbank.org/indicator/CM.MKT.LCAP.CD?end=2022&start=1975&view=chart [Accessed: 22 August 2024].

Zaharudin, K.Z., Young, M.R. and Hsu, W.H. (2022) 'High-Frequency Trading: Definition, Implications, and Controversies', *Journal of Economic Surveys*, 36(1), pp.75-107. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/joes.12434 [Accessed: 21 August 2024].