High-frequency trading & volatility

Impact of high-frequency trading during COVID-19 pandemic

Otto Nakari

Master's Thesis
International Business Management
2023
Degree Thesis
Otto Nakari
High-frequency trading & volatility. Impact of high-frequency trading during COVID-19 pandemic
Identification number:
24711
Commissioned by:
N/A

Abstract:
This thesis originally started with a regulatory angle, which was to find out if high-frequency trading (HFT) or algorithmic trading can be regulated and if so, how it has worked this far. The research was aimed to Australia, as many news sources were writing articles about how the regulator down there has a demand to control HFT. It ended up as a question of where to find material for this research. Later, it became clear that it is not possible to find a reliable source for this information. One reason for that is that the trades are not publicly available online or were they executed by HFT. Therefore, the topic and the hypothesis of the research were changed to be able to find more information from open sources. The hypothesis was actually changed several times from a statistical F-Test to a less statistical to get results which could analysed in a clear way. The research questions were “Does high-frequency trading amplify market volatility? If yes, how much has high-frequency trading amplified the market during the COVID-19 pandemic?” The analysis was conducted by comparing the assumed portion of HFT (30%) with the previous year’s volume and rationalizing it with the previous research done on this topic. The results match the existing market share of HFT, but it is still not undisputable that HFT has amplified the market volatility during the COVID-19 pandemic.

Keywords: High frequency trading, algorithmic trading, COVID-19, equities, market volatility
# Contents

**Figures** ................................................................................................................................. 4

**Table of abbreviations** ........................................................................................................ 5

1 **Introduction** ......................................................................................................................... 6

1.1 Aim of the research and research questions ..................................................................... 8

1.2 Relevance ............................................................................................................................ 8

1.3 COVID-19 Pandemic ........................................................................................................... 10

1.4 Structure of the thesis ......................................................................................................... 11

2 **Literature review** ................................................................................................................ 12

2.1 Efficient markets ................................................................................................................ 12

2.2 High-frequency trading ..................................................................................................... 13

2.2.1 Existing research on high-frequency trading ................................................................. 14

2.2.2 Regulation regarding HFT in Australia ........................................................................ 19

3 **Methodology** ..................................................................................................................... 21

3.1 Research methods ............................................................................................................. 26

3.2 Data Gathering ................................................................................................................ 26

3.2.1 Bloomberg terminal .................................................................................................. 28

3.3 Data analysis ................................................................................................................... 29

4 **Results** ............................................................................................................................. 31

4.1 Statistical tests .................................................................................................................. 33

4.1.1 F-Test ...................................................................................................................... 34

4.2 ASX 300 (Years 2012-2020) ......................................................................................... 35

4.3 FTSE 100 (Years 2012-2020) ......................................................................................... 38

4.4 DAX index (Years 2012-2020) ....................................................................................... 39

4.5 CAC Index (Years 2012-2020) ....................................................................................... 41

4.6 OMX 30 Index (Years 2012-2020) ................................................................................ 42

4.7 S&P Europe 350 (Years 2012-2020) ............................................................................. 43

5 **Discussion** ......................................................................................................................... 44

6 **Conclusions** ...................................................................................................................... 47

**References** ............................................................................................................................. 50
Figures

Figure 1: Volume levels of indices in the financial markets during January 2020 to April 2020 (Bloomberg terminal, 2021)

Figure 2: Clarification of the logic behind the hypothesis

Figure 3: Average volumes of all indices from 2012 to 2020 (Bloomberg terminal, 2021)

Figure 4: Average standard deviation of volumes in all indices (Bloomberg terminal, 2021)

Figure 5: Average volume of ASX 300 index from 2012 to 2020 (Bloomberg terminal, 2021)

Figure 6: ASX 300 index, mean and variance from 2012 to 2020. (Bloomberg terminal, 2021)

Figure 7: ASX 300 index, comparison of the possible portion of HFT to the previous year’s average (Bloomberg terminal, 2021)

Figure 8: FTSE 100, Average volume 2012-2020 (Bloomberg terminal, 2021)

Figure 9: FTSE 100 index, comparison of the possible portion of HFT to the previous year’s average (Bloomberg terminal, 2021)

Figure 10: DAX index, average volume in 2012-2020 (Bloomberg terminal, 2021)

Figure 11: DAX index, comparison of the possible portion of HFT to the previous year’s average (Bloomberg terminal, 2021)

Figure 12: CAC index, average volume in 2012-2020 (Bloomberg terminal, 2021)

Figure 13: CAC index, comparison of the possible portion of HFT to the previous year’s average (Bloomberg terminal, 2021)

Figure 14: OMX 30 index, average volume in 2012-2020 (Bloomberg terminal, 2021)

Figure 15: OMX 30 index, comparison of the possible portion of HFT to the previous year’s average (Bloomberg terminal, 2021)

Figure 16: S&P Europe 350 index, average volume in 2012-2020 (Bloomberg terminal, 2021)

Figure 17: S&P Europe 350 index, comparison of the possible portion of HFT to the previous year’s average (Bloomberg terminal, 2021)

Figure 18: The Conclusion of all results (Bloomberg terminal, 2021)

Figure 19: Global HFT servers markets share, by application in 2020 (grandviewresearch.com, 2023)

Picture 1: The Research “onion” (Saunders, Lewis & Thornhill, 2018) with highlighted points of this research

Picture 2: Left Tailed Test (Cuemath.com, 2022)

Picture 3: Right Tailed Test (Cuemath.com, 2022)

Picture 4: Two Tailed Test (Cuemath.com, 2022)
# Table of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Algo</em></td>
<td>Algorithm. A process or set of rules to be followed in calculations or other problem-solving operations, especially by a computer.</td>
</tr>
<tr>
<td><em>Bloomberg Terminal</em></td>
<td>Computer software system provided by the financial data vendor Bloomberg L.P.</td>
</tr>
<tr>
<td><em>COVID-19</em></td>
<td>Global pandemic caused by a virus in December 2019</td>
</tr>
<tr>
<td><em>ESMA</em></td>
<td>European Securities and Markets Authority</td>
</tr>
<tr>
<td><em>EU</em></td>
<td>European Union</td>
</tr>
<tr>
<td><em>Equity</em></td>
<td>Ownership of a company publicly listed or private</td>
</tr>
<tr>
<td><em>F-Test</em></td>
<td>Statistical test in which the test statistic has an F-distribution under the null hypothesis.</td>
</tr>
<tr>
<td><em>HFT</em></td>
<td>High-Frequency Trading</td>
</tr>
<tr>
<td><em>INDEX</em></td>
<td>Group of equities which measures the price performance</td>
</tr>
<tr>
<td><em>IPO</em></td>
<td>Initial Public Offering</td>
</tr>
<tr>
<td><em>Kill Switch</em></td>
<td>A backup system which stops the trading immediately</td>
</tr>
<tr>
<td><em>NASDAQ</em></td>
<td>National Association of Securities Dealers Automatic Quotation System</td>
</tr>
<tr>
<td><em>NYSE</em></td>
<td>New York Stock Exchange</td>
</tr>
</tbody>
</table>
1 Introduction

In the modern world, all the corporates and countries require capital to operate. Capital is usually gained through a bank or more specifically from the stock markets. Stock markets is sometimes used as a rather broad term, and many times it used to refer to the whole market of financial instruments, including bonds, futures and even some of the instruments which are considered in financial terms as “exotic”.

History knows many situations in different times where the markets have declined many days, even weeks in a row. Those times are called “market crashes” or “stock market crashes”, depending on the context and who is referring to them. Like Garber (1989) describes on Journal of political economy, the first market crash known in history is the year 1637 “Tulip Mania Bubble” where the high demand of tulip bulbs followed by speculative investing raised the price of the bulbs to unexpectedly high figures. The market crashed because of a mass-sell off. After that there have been several market crashes, which most of them have been caused by human interaction. Before the time of computers, all the bidding in stock markets was done by humans who were shouting on the trading floor. (Macey, Mitchell & Netter, 1988) The situation was like an auction. Later, the stockbrokers got assistance methods to get their bids from their clients. One of the most important methods was of course the telephone but definitely the most efficient method was the computers. The first computer-kind of quotation systems were Ultronics and Quotron. They were introduced in the stock markets in 1960’s. NASDAQ stock exchange was founded in the United States of America in 1972, during the 1970’s and in the beginning of 1980’s all of the stock exchanges started to use electronical trades which are considered to be the modern version of trading by computers.

In 1987 the New York Stock Exchange faced a market crash which was later named as “Black Monday”. That can be considered as the first market crash where an algorithm used by a computer was partly the cause of destruction. Now in 2022, many of the trades are executed by computers automatically by a method called High-Frequency Trading (HFT). In HFT, the computer uses complex algorithms to execute large number of orders in fractions of seconds. The volatility of the markets and price changes generates the profits for the owner of the computer or algorithm.
In November 2019, the COVID-19 pandemic started from China in a city called Wuhan. The pandemic grew global in the beginning of 2020 and caused a lot of market volatility.

“On 11 March, 2020, the World Health Organization (WHO) officially declared the coronavirus (COVID-19) outbreak to be a global pandemic. As of 27 March, 2020, the number of confirmed cases surpassed 500,000, and it continues to rise (WHO, 2020). Over 170 countries are affected, with the US having the most confirmed cases. At the same time, the US stock market hit the circuit breaker mechanism four times in ten days. Since its inception in 1987, the breaker has only ever been triggered once, in 1997. Together with the US crash, stock markets in Europe and Asia have also plunged. FTSE, the UK's main index, dropped more than 10% on 12 March, 2020, in its worst day since 1987. The stock market in Japan plunged more than 20% from its highest position in December 2019. Central banks and authorities responded immediately by throwing their policy instruments into the market.” (Zhang, Hu & Ji, 2020)

There are examples in history when algorithmic trading has “malfunctioned” in unexpected situations in the markets, for example the “Flash crash” in New York stock exchange in 2010. According to Zhang (2010), algorithmic trading has reached almost 85% of dollar trade volumes. Some researchers say that algorithmic trading has made the markets more liquid and that is why we need them. (Hendershott & Jones, 2011) But in special situations, like COVID-19 pandemic, do these algorithms amplify the negative impact in markets, and therefore make it more volatile?

This master’s thesis is trying to find the answer to the previous questions by analyzing data of different global indices and comparing the hypothetical portion of high frequency trading defined by previous research with the volume of the previous year.
1.1 Aim of the research and research questions

The aim of this research is firstly to find out has the COVID-19 pandemic has increased the trading volumes and secondly, has algorithmic trading amplified the volatility of the financial markets even more.

In the beginning, this research was going to find out if regulatory measures as “Kill switch” would help to stabilize the market. The Australian market regulators is one of the few who has made a mandatory rule to the operators in their market to implement a “Kill switch” to their systems. The act came effective in 2012. The market regulator can also exclude certain operators from the market if they suspect that they will harm market stability.

After more than six months of online research, it became evident that there is no data available from the Australian regulator which would tell how many times the “Kill switch” measure has been historically used. That is why the aim of the research and questions were changed from the previous topic to find out if the COVID-19 pandemic has caused volatility on the financial markets, and secondly has the algorithmic trading (high-frequency trading) amplified the possible volatility.

Based on the previous research, the final research questions were formed in following way:

- Is high-frequency trading harmful to the financial markets?
- Did the COVID-19 pandemic create volatility in the financial markets?
- Does HFT (high-frequency trading) amplify the market volatility?

The main question where this research aims to have a clear answer is “Does high-frequency trading amplify market volatility? If yes, how much has high-frequency trading amplified the market during the COVID-19 pandemic?”

1.2 Relevance

According to the vast majority of empirical research, high-frequency trading (HFT) improves market liquidity, reduces costs and makes the financial markets more efficient in general. Most of the used strategies are not new, they are just improved to fit in to the automated environment. Basically, most of the issues regarding HFT already existed when the trades were manually executed (Jones, 2013).
When technology is evolving faster, it also allows faster trades. This has led to a discussion about should the HFT be more regulated. Many countries have not implemented “Kill Switch” functions to their stock exchanges like Australia has, but in most cases, the financial institutions who are using HFT have their own “Kill Switches” to prevent losses and to comply with the local or global regulation and standard. New technologies allow even faster trades than today, and it is important to understand the consequences of limiting the financial markets from both angles as “Market efficiency vs. Market stability”. (Weller, 2018)

According to the statistics 60-73% of the trades in the US are done by utilizing algorithmic trading. In general, the volumes in high-frequency trading are constantly growing around the world. Therefore, it is an interesting topic to know more about. According to the previous research, high-frequency trading has provably caused few “spikes” in the financial markets and have been amplifying the volatility. (Breckenfelder, 2020)

When looking at the numbers of volumes from 2019 to the end of 2020, it is possible to see the huge growth in the trading volumes. One explanation is the price drop in March 2020, caused by the COVID-19 pandemic but has algorithmic trading or high-frequency trading amplified this event even more?

In the past, some of the market crashes have been caused by algorithmic trading or high frequency trading. In this case, it is quite sure that those did not create the markets to decline or plunge in volatility, but did it amplify the effect? Many of the factors regarding algorithmic trading are unknown, by going through existing research it is possible to determine the market share of each continent. (Breckenfelder, 2020) Also, even the market shares are estimates, so it is impossible to be 100% sure of them. In forex trading algorithmic trading is nowadays almost like a standard procedure and it is very difficult to be an operator in the financial markets without using it. By conducting tests on the data gathered from Bloomberg it is possible to find a rather neutral view on this argument. When looking at different articles and research, it can be seen that the researchers are not completely unanimous about the real market share of HFT. That is why this research uses as hypothesis 30% market share, as the European market’s share is considered to be between 20% to 40%. The number depends on which country we are looking at and when, as well as how the research is conducted.
1.3 COVID-19 Pandemic

The global COVID-19 pandemic started in November 2019 from Wuhan China and spread fast around the world. In the western world, the first peak or wave started in February to April 2020. That can be seen already very clearly from the figures for trading volumes from February 2020 onwards.

When looking at the volumes of the indices from January 2020 to April 2020 it is possible to see how the financial markets reacted to the news of COVID-19 spreading around the world. As an example, if we look at the Australian ASX 300 index in February, the historical change percentage in volume compared with January (2012-2019) has been between 19,04% (2016) and 39,75% (2018). In 2020 the same percentage was 53%, meaning that it was a record high for the last 8 years. The numbers are even higher if comparing February and March in the sample data. Therefore, when looking at the chart from January 2020 to April 2020 with all indices of this research, it is possible to state that the markets around the world reacted in a similar way to the news of spreading COVID-19. (Zhang, Hu & Ji, 2020)

Albulescu (2021) made short research about the connection between COVID-19 fatality ratio and volatility of the financial markets. A large amount of statistical research has been done regarding COVID-19 and it can be used as supportive evidence to confirm the results in this research.
When looking at March 2020 alone, it is possible to see a clear peak in the volumes of indices when compared to the level of the previous months and month after. March 2020 is the approximate moment when COVID-19 arrived in most of the western countries and started to spread. The virus caused a global pandemic which led to panic in the financial markets, meaning that people started to sell and buy. They who needed cash started to sell and of course, like in any crisis, those who had liquidity started to see the opportunity to buy. As mentioned in many guidebooks of investing, this type of crisis is the best time to buy, because in general, it is very difficult to time the financial markets otherwise. (Baker & Wurgler, 2006)

1.4 Structure of the thesis

Chapter one gives a brief introduction to the topic and forming of the research questions, as well as short description of the global COVID-19 pandemic, as it is a major factor regarding the research questions. Chapter two goes into the main points of the research methods used in this thesis. Chapter two goes deeper into algorithmic trading through the history of trading theories related to the topic. This is required to be able to understand the development in recent years. It also goes into the literature of markets and trading having the concentration in high-frequency trading (HFT) by explaining the difference between rule-based trading and more
complex HFT. Chapter three also explains the research methods, used datasets, analyzing methods, explains the used statistical test methods and the key turning points in the financial markets which are important when analyzing the data.

Analysis and the results are represented in chapter four. Discussion and conclusion can be found from chapters five and six.

2 Literature review

Since the beginning of the financial markets, they have been somehow regulated. Usually after a crisis, the governments try to find a proper solution to prevent the crisis from never happening again. Still, when it comes to the financial markets and the trade of financial instruments, usually the governments trust the basic principle of supply and demand, and do not want to limit that, except in a very critical situation.

Algorithmic trading started approximately at the same time when computers arrived in the trading floors, so in the 1980’s. Like many new things, the algorithms or rule-based trading were not regulated at all. Now in 2023, artificial intelligence (AI) has taken huge steps forward compared to the previous decades. Breakout of the AI is concerning all industries, and many professionals agree that it should be somehow regulated. Nowadays, many market places have controls on how to detect algorithmic trading and exclude the trades made by it if needed. (Australian authorities, Nasdaq Helsinki, 2023)

This chapter has concluded the relevant previous research and theory related to algorithmic trading and how it might have an affect to the liquidity of the financial markets. The previous research is mostly about how high-frequency trading has been utilized during the situations of crisis in the markets.

2.1 Efficient markets

The efficient market hypothesis, also known as the efficient market theory, is a hypothesis that states that share prices reflect all information and consistent alpha generation is impossible. In other words, it states that share prices are never undervalued, and it is impossible to use timing and purchase shares with a lower price to outperform the market. This theory is very
controversial in the academical world. Eugene F. Fama in his research “Market efficiency, long-term returns, and behavioral finance” from 1997 has a logical argument why the researchers should not trust blindly to the efficient market theory.

The efficient market theory suggests that prices overreact to information. Fama states that is true, and the market prices move in both directions based on the information. According to Fama, this theory does not take into consideration the behavior of the other traders on the market. (Fama, 1998)

This behavioral factor is very important when looking into HFT, which is almost fully based on the available information of prices and patterns. Though, according to the research of high-frequency trading made Allen Carrion, it is difficult to make conclusions by analyzing the data and theory make a conclusion whether if high-frequency trading makes the financial markets more efficient or not.

2.2 High-frequency trading

High frequency trading, also known as HFT, is a trading method which uses computers calculation power by complex algorithms in ultra-fast trading. The trades can be executed within fractions of a second. Usually, algorithmic trading is a programmed code which can make mathematical calculations and decisions based on those very quickly. But it is also possible to program a computer to make slower rule-based decisions. Slower ones are not of course capable of utilizing small deviations in prices like the faster and complex ones, which can utilize fractions of cents in their decisions. When utilizing fractions of cents, it also requires a lot of capital to work with to be able to gain sufficient profits for the operator.

Electronic high-frequency trading has come under global inspection since it was blamed for the “flash crash” in the Dow Jones Industrial Average in May 2010 when the index sank 1000 points, or 9% per cent, and regained most of those losses in less than 20 minutes. According to Wall Street, algorithmic trading is already involved in 60-73% of all trades done in the US market. (Chaboud, Chiquoine, Hjalmarsson & Vega, 2014)
2.2.1 Existing research on high-frequency trading

In 2010’s there has been a lot of discussion should the stock exchanges have the right of using a method called “Kill Switch” for HFT to prevent market crashes and be able to maintain the stability in the markets. In 2012, Australian government introduced their rule of using “Kill Switch” in HFT to stable the financial markets. It is mandatory for all traders who use algorithms to have a “kill switch” in their operating systems. The “kill switch” shall be used in a case of a market flash crash. USA has been planning to implement a similar solution and give the right of execution to the stock exchanges as well, in addition to the companies who use HFT.

Charles Jones from Columbia Business School writes about high-frequency trading and explains what HFT is, and what are the benefits and risks in using it. Appears that this is version 3.4 which means that he has done research about the topic already before. He also refers to them in the article. He uses examples of market crashes and supports them with theories in finance and other research related to HFT. The main research question is “Does HFT make markets more fragile?” which is very similar to the question represented in this master’s thesis. “Does HFT amplify the market volatility?”. Jones writes about “Kill Switch” mechanisms but because the United States Securities and Exchange Commission (SEC) has not implemented that kind of mechanism, he can only write about the discussion between the operators and SEC.

The data used in the research is secondary data from reports in another research. By analyzing the data, Jones uses support of existing theories as from market liquidity. Part of the indices are the same as been used in this thesis. Jones uses comparison of bid-ask spreads on Dow Jones stocks from 1900 to 2000 and Median bid-ask spreads on S&P 500 stocks from 2003 to 2009. Other sources of data are E-mini S&P 500 futures prices during the flash crash of May 6, 2010. The rest of his sources are previous research which he has collected regarding the topic. Based on the research, Jones gives suggestions of implementation to the financial regulators of the United States of America and raises up issues regarding HFT in the beginning of the article in the “Executive summary” section.

Kirilenko, Kyle, Samadi & Tuzun,(2017): “The flash crash: High-frequency trading in an electronic market” is partly based on Charles Jones research. As Jones concentrated on the HFT as a whole, Kirilenko, Kyle, Samadi & Tuzun,(2017) have researched only a certain time
period which is the flash crash of May 6, 2010. The flash crash of May is expected to be caused by HFT. First, the article represents the timeline of the day of the flash crash. It is from a formal report and not very clear to read but the article explains the events also in a written form and illustrates it by a diagram.

In the data, the focus is on E-mini S&P 500 future. Traders in the E-mini, including those who buy and sell throughout a trading day, do not possess formal designations such as market makers, dealers, or specialists. To classify accounts as intraday intermediaries, Kirilenko, Kyle & Samadi (2017) adopt a data-driven approach based on trading activity and inventory patterns. Their definition of intraday intermediaries is intended to capture traders who follow a strategy of consistently purchasing and selling throughout a trading day while maintaining low levels of inventory. To be classified as an intraday intermediary, a trader denoted by $j$ must meet criteria (i) with respect to its daily trading volume ($V_{olj,d}$), where $d$ denotes a trading day, (ii) with respect to its end-of-day position ($NP_{j,d,t=405}$) relative to its daily trading volume, where $t$ denotes each minute during a trading day, and (III) with respect to its intraday minute-by-minute inventory ($NP_{j,d,t}$) pattern. (Kirilenko, Kyle & Samadi, 2017)

In other words, they use mathematical formulas to categorize the traders with three-step criteria:

1. An account must trade 10 or more contracts on at least one of the three days prior to the Flash Crash (May 3, 4, and 5, 2010).

   According to the data, this volume cutoff is a prudent approach to removing accounts that do not trade an economically significant amount before categorizing intraday 12 intermediaries.

2. The three-day average of the absolute value of the ratio of the account’s end-of-day net position to its daily trading volume must not exceed 5%.

   Specifically, compute the daily ratio of a trader’s end-of-day position to its daily trading volume on May 3, 4, and 5, compute the absolute value of the ratios for each day, and calculate the three-day average of the absolute values of the ratio.
3. The three-day average of the square root of the account’s daily mean of squared end-of-minute net position deviations from its end-of-day net position over its daily trading volume must not exceed 0.5%. (Kirilenko, Kyle & Samadi, 2017, p. 12-13)

The 16 most active accounts who had the highest number of trades during the flash crash, they classify as “high-frequency traders”. The other intraday intermediary accounts are classified as Market Makers. All other traders are classified as “small traders, “fundamental buyers”, “fundamental sellers” and the rest as “opportunistic traders”.

The criterion for the classification is following:

- **Small trader** = Less than 10 contracts during the day
- **Fundamental buyer** = 10 contracts or more and accumulates a net long end-of-day position equal to at least 15% of its total trading volume for the day.
- **Fundamental seller** = Trades 10 contracts or more and the absolute value of its net short position at the end of the day is at least 15% of its total trading volume for the day.
- **Opportunistic trader** = move in and out of positions throughout the day but adjust their net holdings with significantly larger fluctuations and lower frequency than intraday intermediaries. Opportunistic Traders may follow a variety of arbitrage trading strategies, including cross-market arbitrage (for example, long futures/short securities), statistical arbitrage, and news arbitrage (buy if the news indicators are positive/sell if the news indicators are negative). Opportunistic Traders may also engage in providing intermediation across days or weeks rather than intraday.

“Classification methodology is based entirely on directly observed individual inventory and trading volume patterns of traders.” (Kirilenko, Kyle & Samadi, 2017, p. 15)

Kirilenko, Kyle & Samadi (2017) do not give any direct suggestions based on their study. They just state that this knowledge gained during the research can assist the regulator to adjust to the “new normal” which is high-frequency trading. Frank Zhang examines the impact of high-frequency trading (HFT) on the U.S. capital market in research named “High-frequency trading, stock volatility, and price discovery”. His research questions are following: “(1) Does HFT decrease or increase stock price volatility? and (2) Does HFT aid or hinder the market’s incorporation of news about firm fundamentals into stock prices?” These questions are very
similar to the research questions as in this research paper and therefore can be used as a benchmark for this research.

The sample contains all stocks covered by the CRSP and Thomson Reuters Institutional Holdings databases between the first quarter of 1985 and the second quarter of 2009. Stocks which are less than 1$ are deleted from the sample. “As the Thomson Reuters Institutional Holding database contains only data at the quarterly level, the sample is composed of firm-quarter observations. Quarterly stock turnover, which is defined as trading volume divided by 12 outstanding shares, is calculated from CRSP. To account for the double-counting of dealer trades for Nasdaq firms (Gould and Kleidon 1994), Nasdaq trading volume is divided by two.” (Zhang, 2010, p. 11-12)

As Kirilenko, Kyle & Samadi, (2017) and Zhang (2010) also categorize the trader into three categories: institutional investors, individual investors, and high-frequency traders. He uses mathematical formulas to get the share of each group. The second research question he answers by another formula. As a conclusion he states large amount of the trades in the US market are made by HFT companies and that shares that are affected by HFT tend to over-react regarding the market movements.

Also, Allen Carrion has interesting research about high frequency trading, which is quite close to this one, even though it is on a bit more granular level. The research is from 2013 and done by picking 120 equities from NYSE & NASDAQ stock exchanges in the United States of America, 60 equities from both stock exchanges. The sample in the data is from 2008 to 2009 and then covers one week from 2010. The similarity which this research is that in this research there is also a global crisis going on at some part of the timeline of the data. In Carrion’s data the crisis is the global financial crisis in 2008 when in this research is COVID-19 pandemic in 2020. This is a similar event in a way, as the financial markets had a huge drop in both situations and started to grow after it more than ever before.

“The first question in Carrion’s research address what are the sources of HFT profitability? That is done by investigating market timing performance. According to him, it helps characterize traders’ strategies to give insights into their motives for trading, which likely impacts market quality, and provides evidence on intraday return predictability. His second research question is what trading costs do HFTs face when executing their strategies? This
provides additional insights into the sources of their profitability, as well as their decisions on when to supply and demand liquidity. Examining the permanent price impacts of HFT trades also tests theoretical predictions that they impose high adverse selection costs on other traders when demanding liquidity and avoid being adversely selected when providing liquidity. Finally, what impact do HFTs have on market quality?” (Carrion, 2013)

Carrion concludes his research in the following way: “HFTs seem to possess intraday market timing ability, and this result is not driven solely by very short-term signals or trading at fleeting prices. The magnitude of their market timing performance suggests that there is economically significant predictability in intraday 40 prices. Trading costs are low in this market, but spreads are wider on trades where HFTs provide liquidity and tighter on trades where HFTs take liquidity. This suggests that HFTs provide liquidity when it is scarce and consume liquidity when it is plentiful. Prices incorporate information from order flow and market-wide returns more efficiently on days when HFT participation is high. This effect is driven by HFT demand-side participation, implying that HFTs improve price efficiency when demanding liquidity. This new evidence can potentially provide guidance to theoretical researchers seeking to model HFT behavior and market quality impacts. For example, the relatively low spreads earned on their liquidity providing trades, their market timing performance, and the large share of their trades that demand liquidity together suggest that one may not want to model HFTs as uniformly following market-making strategies. The HFT intraday market timing results suggest that models where HFTs solely profit from very short-term activities such as trading at fleetingly available prices may be incomplete.” (Carrion, 2013)

Peter Gomber who is a professor in e-finance at Goethe University Frankfurt has made good research commissioned by the Frankfurt stock exchange regarding HFT. Despite the research is done in 2012, it is still valid from many parts of the research. Gomber has gathered many propositions of high-frequency trading and tries to have a short answer to all of them. The propositions give a great understanding about what is high-frequency trading, and which matters related to it are myths and which are facts.

The most relevant argument or proposition in Gomber research is: “Academic literature mostly shows positive effects of HFT based strategies on market quality.” The majority of papers, focusing on HFT, do not find evidence for negative effects of HFT on market quality. On the contrary, the majority argues that HFT generally contributes to market quality and price
formation and finds positive effects on liquidity and short-term volatility. Only one paper critically points out that under certain circumstances HFT might increase an adverse selection problem and in case of the flash crash one study documents that HFT exacerbated volatility. As empirical research is restricted by a lack of accessible and reliable data, further research is highly desirable.” (Gomber, 2012)

As also Gomber explains in his research, it is very difficult to find reliable data which would show the portion of high-frequency trading in the financial market. That is why here, this research, the data concentrates on the average volumes which are rather easy to get from a Bloomberg terminal and then by creating assumptions of the share of the trades which are conducted by using HFT. The assumptions of the portion of HFT are based on few articles like the one written by Johannes Breckenfelder, senior economist from the European Central Bank in 2020. (Breckenfelder, 2020)

2.2.2 Regulation regarding HFT in Australia

For the last decade, the Australian market regulator has been quite strict when it comes to algorithmic trading, and especially high-frequency trading. 2014 the regulator created an act which is known as the “Kill-switch act”. It was implemented to prevent market crashes and increase market stability. The act states: “In response to concerns about dark liquidity and high-frequency trading, the government has introduced new Market Integrity Rules relating to dark liquidity and automated trading that are due to take effect between May 2013 and May 2014. These include:

- a price improvement requirement for dark trades, to encourage more trading to occur on lit exchange markets and support the price formation process (May 2013);
- enhancements to the market operator controls for extreme price movements, including automated trading pauses (May 2013 and 2014);
- enhancements to market participant filters and controls for automated trading, including a 'kill switch' to immediately shut down problematic algorithms (May 2014); and enhancements to the data ASIC receives to improve our market surveillance (March 2014).” (gov.Au.)
In practice, it means that every operator who is using algorithms in their trading operations, needs to have a “Kill Switch” structure implemented to their trading systems. Function of a “Kill Switch” as a structure will terminate all the trades in the operator’s system immediately. The Australian stock exchange (ASX) refers by “Kill Switch” act to both parties’ exchanges and companies who conduct trading; “A kill switch is the ability of an exchange or Participant to instantly stop trading activity by a specific client and to purge open orders.” (ASX.com)

“ASIC expects that the automated trading pauses and 'kill switches' will reduce the potential for algorithms to exacerbate market volatility and cause liquidity to disappear.” (ASIC, 2013, p. 37)

In Finland, there’s less regulation in HFT compared to for example Australia. But Nasdaq Helsinki has a right to use” Kill functionality” if they consider that a broker is disturbing the market with their actions. (Nasdaq, kaupankäyntisäännöt 2021, p. 33)

According to the rules of many global exchanges, they still usually have an option to cancel all or partly the trading in their exchange if they notice unusual market disturbance. Therefore, the “Kill Switch” act in Australia is not completely a unique function. Despite it is not a unique function, Australian government has claimed that having a mandatory “Kill Switch” function is stabilizing the local market. (FINRA, 2023)

According to the research conducted by the Australian Securities and Investments Commission (ASIC), the average holding time for securities was 42 minutes in 2013 with following breakdown:

- 0.1% of high-frequency traders held positions for less than one second
- 1.2% were held for an average of two minutes or less
- 18% were held for less than 10 minutes
- 51% for less than 30 minutes (ASIC, 2013, p. 31)

When looking at the figures of estimates done by the Australian Securities and Investment Commission, the holding times on average are quite low. Depending on the investment strategy a large portion of them can be executed automatically, but only 1.3% of the trades are held less than 2 minutes and therefore, can be categorized by timing to be high-frequency trading. The
trades with longer holding time can of course be executed by an algorithm but are not fast enough to be categorized as HFT.

The HFT already represents large portion of the trading globally, depending on the market from 24% to more than 50% (ESMA, 2014)

According to a report conducted by the European Securities and Markets Authority (ESMA) in 2014 an average of 24% of the total trades around the exchanges in Europe can be flagged as high-frequency trades.

Unfortunately, there is not so much of research regarding the regulation of the algorithmic trading in Australia. Even when looking at the latest popular research regarding algorithmic trading alone in Australia, there are not many of them. The most popular research from the latest is by Zhou & Kalev (2019) “Algorithmic and high frequency trading in Asia-Pacific, now and the future”. The article by Zhou & Kalev (2019) mostly discusses why Asia-Pacific as an area is attractive for algorithmic and high frequency trading. Shortly, because of good infrastructure, labour and it is a large economy.

3 Methodology

As Cheng & Phillips (2014) explains later on in this chapter, quantitative research process typically starts from a research question or questions and through a criterion measuring the data. It can also start from the data but in general it is a method which analyses numerical data. “The following phases are linked with a quantitative methodology and are used interchangeably: a deductive approach, an etic view, objective epistemology, a structured approach, numerically based data collection, statistical analyses, and replicable research design. In other words, quantitative studies have four main characteristics: systematic/reconstructed logic and linear path (step-by-step straight line); hard data in nature (e.g. numbers); they rely on positivist principles, they have an emphasis on measuring variables and testing hypotheses, they usually verify or falsify a relationship or hypothesis we already have in mind.” (O’Gorman, 2015)

Also, Hair, Page, & Brunsveld (2019) have a similar description of quantitative research methodology. They have also added to the description that it is important to display the data in
a graphical form. That helps the reader to understand the structure of the data, which is especially important if the amount of data is high. In their book they have examples of presenting the research to business managers, but it is similar to presenting the data to the reader and of course to the supervisor of the thesis. Hair, Page, & Brunsveld (2019) states that the research question in quantitative study should be very specific and refer to numbers. As the hypothesis is chosen to be 30% market share of HFT, it matches with the theory mentioned above.

The figure below pinpoints the methods used in this particular research. The research “onion” is created by Saunders, Lewis & Thornhill (2018) This method was chosen to be one of the research methods for this thesis, as it is very easy to understand and illustrative.

Picture 1: The Research “onion” (Saunders, Lewis & Thornhill, 2018) with highlighted points of this research

Going from the out layer of the research “onion”, this research approach is deductive which means that the theory comes before the data, as the aim is to analyze the research question first and then find out can it be confirmed by the data. The methodological choice is mono method quantitative, as only actual numbers are measured in the analysis. Therefore, because the historical data of the financial markets is gathered from Bloomberg terminal, the strategy can be called archival research.
This is just the processes on the outer rim of the research methods, but if going into the core of the analysis, as the data is gathered from the financial markets by Bloomberg, the research can be defined as “secondary analysis of existing data”. There is a lot of research and theory about this method of analysis, for example by Emma Bell & Alan Bryman (2007), but it was interesting to find out that this method is not bounded by the area of research. Cheng & Phillips (2014) describe the method very well in their scientific article “Shanghai archives of psychiatry”.

There are two general approaches for analyzing existing data: the ‘research question-driven’ approach and the ‘data-driven’ approach. In the research question approach, researchers have an a priori hypothesis or a question in mind and then look for suitable datasets to address the question. In the data-driven approach, researchers look through variables in a particular dataset and determine what kind of questions can be answered by the available data. In practice, the two approaches are often utilized together and iteratively. Researchers typically begin with a general understanding of the question or hypothesis and then search for available datasets that contain the variables needed to address the research questions of interest. If they do not find datasets that contain all variables needed, they usually modify the research question(s) or the analysis plan based on the best available data.

When conducting either research question-driven or data-driven approaches to the analysis of existing data, researchers must follow the same basic methods.

a) There must be an analytic plan that includes the specific variables to be considered and the types of analyses to be conducted. In the research question-driven approach this is determined before the researchers look at the actual data available in the dataset; in the data-driven approach this is determined after the researchers look through the dataset. (Cheng & Phillips, 2014)

As the thesis started by formulating the research questions, it follows this “research question-driven approach” and it determines how the figures gathered by Bloomberg are analysed.

b) Researchers must have a thorough understanding of the strengths and weaknesses of the dataset. This involves obtaining detailed information about the population under study, sampling scheme and strategy, time frame of data collection, assessment tools, response levels,
and quality control measures. To the extent possible, researchers must acquire and study in detail all survey instruments, codebooks, guidebooks, and any other documentation provided for users of the databases. These documents should provide sufficient information to assess the internal and external validity of the data and allow researchers to determine whether or not there are enough cases in the dataset to generate meaningful estimates about the topic(s) of interest. (Cheng & Phillips, 2014)

Part b) takes into account all types of research methods, but all of them have some weaknesses and weaknesses. In this case, the financial markets are different on different continents, as well as the considered market share of high frequency trading.

“(c) Before conducting the analysis, researchers need to generate operational definitions of the exposure variable(s), outcome variable(s), covariates, and confounding variables that will be considered in the analysis.” (Cheng & Phillips, 2014)

Even though, in this research method guideline it is said that “original data set should never be altered”, in this case the bank or market holidays needed to be removed to be able to analyse the data set in a same way. Otherwise, other samples would have a different number of market days available in the analysis than the others.

d) The first step in the analysis is to perform frequency tables and cross-tabulations of all variables that will be included in the main analysis. This provides information on the use of the coding pattern for each variable and the profile of missing data for each variable. Due attention should be paid to avoiding patterns, which can result in large numbers of missing values for certain variables. In comprehensive surveys that take a long time to complete, skipping a group of questions that are not relevant for a particular respondent (i.e., ‘skips’) is a common method used to reduce interviewee burden and to avoid interviewee burn-out. For example, in a survey about alcohol-related problems, the survey module typically starts with questions about whether the interviewee has ever drunk alcohol. If the answer is negative, all questions about drinking behaviors and related problems are skipped because it is safe to assume that this interviewee does not have any such problems. Prior to conducting the full analysis, these types of missing values (which indicate that a particular condition is not relevant for the respondent) need to be distinguished from missing values for which the data is, in fact, missing (which
indicate that the status of the individual related to the variable is unknown). Researchers should be aware of these skips in order to make a strategic decision about the coding of these variables.

(e) Finally, the researcher should recode the original variables in order to properly handle missing values and, if necessary, to transform the distribution of the variables so that they meet the assumptions of the statistical model to be used in the intended analysis. The recoded variables should be stored in a new dataset and all syntax for the recoding of variables (and for the analysis itself) should be documented. The original dataset should NEVER be altered in any way.

(f) When using data from longitudinal surveys or when using data stored in different datasets, it is critical to check the accuracy of the identifier variable(s) to ensure that the data from different time periods or from different datasets is matched correctly when merging the datasets.

g) For longitudinal studies, the assessment methods and the coding methods for key variables can alter over time. Furthermore, close examination of the survey questionnaires and codebooks are essential to ensure that each variable in the combined dataset has a uniform interpretation throughout the study. This may require the creation of separate uniform variables that are constructed in different ways at different points in time throughout the study.

(h) Many population-based surveys, particularly those focused on assessing the prevalence of relatively uncommon conditions such as schizophrenia, employ multi-stage sampling strategies to enrich the sample. In this case, the data set usually includes design variables for each case (including sampling weight, strata, and primary sampling unit) that are needed to adjust the analysis of interest (such as the prevalence of a condition, odds ratios, mean differences, etc.). Researchers who conduct secondary analysis of existing data should consider the design variables used in the original study and apply these variables appropriately in their own analyses in order to generate less biased estimates.” (Cheng & Phillips, 2014)

Many parts of the step-by-step research guidelines of secondary analysis by Cheng & Phillips (2014) are meant for a questionnaire or for a multi-method research where both numerical and qualitative data is being used. As this research is not using qualitative data, some of the steps can be skipped or altered.
3.1 Research methods

The main research question is “Does high-frequency trading amplify market volatility? If yes, how much has high-frequency trading amplified the market during the COVID-19 pandemic?” This research tries to find out the answer by simple variation tests like comparing if the portion of HFT is equal to previous year’s average volume to measure the changes in volatility in the financial markets.

The data for the research is gathered from Bloomberg terminal. The data consists of volumes of six different stock exchange indices around the world. The research concentrates on the volumes of the indices and is seeking an answer and confirmation for the research questions mentioned in chapter one.

The data is analyzed by testing the portion with different percentages and comparing it to the previous year’s average volume and rationalize it by the previous research in this topic.

![0-52 = Average is 26%](image)

*Figure 2: Clarification of the logic behind the hypothesis*

3.2 Data Gathering

To be able to gather data, the right datasets are needed. As said before, in this research, the data is gathered from Bloomberg terminal, and the comparison is made between these indices: Australian S&P/ASX 300, French CAC, German DAX, Swedish OMX 30 and the British FTSE 100 Index.

“S&P/ASX 300 is extensively used as a performance benchmark index. The index is highly liquid, float-adjusted and includes up to 300 of Australia's largest securities by float-adjusted market capitalization. The S&P/ASX 300 index includes the large cap, mid cap and small cap components of the S&P/ASX index family. The index was launched in April 2000.” (Bloomberg)
As Bloomberg states S&P/ASX 300 index is commonly used to understand the state of the Australian markets. That is why in this research the volatility of the index is used to measure the stability of the Australian market.

“The CAC 40® is a free float market capitalization weighted index that reflects the performance of the 40 largest and most actively traded shares listed on Euronext Paris and is the most widely used indicator of the Paris stock market. The index serves as an underlying for structured products, funds, exchange traded funds, options, and futures. It is operated by Euronext, the pan-European exchange.” (Bloomberg, 2023)

“The German Stock Index is a total return index of 40 selected German blue-chip stocks traded on the Frankfurt Stock Exchange. The equities use free float shares in the index calculation. The DAX has a base value of 1,000 as of December 31, 1987. As of June 18, 1999 only XETRA equity prices are used to calculate all DAX indices.” (Bloomberg, 2023)

“The OMX Stockholm 30 Index consists of the 30 most actively traded stocks on the Stockholm Stock Exchange and is a market weighted price index. The composition of the OMXS30 index is revised twice a year. The index was developed with a base level of 125 as of September 30, 1986. Effective on April 27, 1998 there was a 4-1 split of the index value.” (Bloomberg.com, 2023)

“The FTSE 100 Index is a capitalization-weighted index of the 100 most highly capitalized companies traded on the London Stock Exchange. The equities use an investibility weighting in the index calculation. The index was developed with a base level of 1000 as of December 30, 1983. * Please see UKEDA100 Index and FTPTP100 Index for the official FTSE 100 Index Dividend Yield and P/E Ratio*” (Bloomberg.com, 2023)

As Bloomberg states, The FTSE 100 Index follows the 100 most highly capitalized companies in the London Stock Exchange. This index is also a good benchmark when comparing the stability of the Australian S&P/ASX 300 index.

The timeframe is from 2012 onwards, so it is possible to understand how the financial markets react in normal situations and how the volumes usually change. The dataset also includes the
data from 2020 which was a downturn globally in the financial markets because of the COVID-19 pandemic.

Bloomberg terminal is chosen to be the source of data, as it is considered to be one of the most trustworthy and used by many professionals in the financial industry. Also, because it gathers data from several sources. For a reader who does not have access to Bloomberg, some of the data can be confirmed by using Google search or the search option in Yahoo Finance. As mentioned before, the target of this thesis is to find out has HFT amplified the effect of COVID-19 pandemic regarding market volatility. The research relies on the available financial market data of different indices.

Data was collected in January 2021, so it was possible to have the whole year of 2020. It was collected by downloading the data from a commercial Bloomberg terminal and stored in a Excel spreadsheet. When gathering the data for this research, it needed to be slightly modified before it was ready for testing. The only modification which needed to be made was to remove the market holidays from the data set which did not have any value. The market holidays were not identical in all sample countries. Therefore, there is a chance that it has affected the results in some way. Even though there is an effect, it would be insignificant.

After removing the market holidays of each index, the average volumes of them were calculated and compared with the next year’s hypothetical portion of HFT (30%). As an example, portion of HFT in 2020 vs. average volume of 2019.

All the indices mentioned below provide a picture of the British, Scandinavian, Center-European and Australian equity markets. American indices as S&P 500 are limited out, as the volumes are significantly higher than in Europe and Australia. As a note, according to several studies, the portion of high-frequency trading is higher in the US than in EU or Australia.

3.2.1 Bloomberg terminal

Bloomberg terminal is a computer software which gathers real-time financial data, news, and feeds from different online data sources. It is also possible to chat with other users of the terminal globally, as well as use the platform for trading. Because of the rather high subscription price (24 000$/year in 2022), most of the users are professionals as portfolio
managers and analysts who are working in the financial industry. The terminal gathers information such as prices, corporate actions, and trading volumes of trading objects (stocks, bonds, ETF’s etc.). It also gathers financial news of different companies, this option is known to be used among some algorithms in trading (Bloomberg.com, 2017). So, algorithms can pick up key words in the financial news which can be used as a trigger for buy or sell transactions in the operators trading system. Of course, a sophisticated algorithm should also contain other factors before triggering the final decision.

Regarding the data of Bloomberg, this research concentrates on the volumes of different global indices. The data is gathered in 2021 by importing the data set of the indices from Bloomberg terminal to Excel. The data can be verified by using websites such as Yahoo Finance if Bloomberg terminal is not accessible.

Because the indices in this research are from different countries, there are few gaps where the data is not available. That is because of bank/market holidays in the specific country. Still, the data is comparable, as the number of days is almost the same in each data set.

3.3 Data analysis

The research uses certain statistical testing (standard deviation) to understand the data of different countries. This criterion is based on market crashes or slumps which can indicate differences between the markets.

Data analysis and design involves measuring variables which can be dependent or independent. Dependent variable is what the researcher think will be affected by another variable, while the independent variable(s) is what will be assumingly affected by the dependent variable. These will be identified directly from the research question. Sampling will not be valid for the research of this thesis, as the data is secondary data from Bloomberg which contains the values of the financial indices.

“As measurement types O’Gorman & MacIntosh (2015) divide them to two classifications:

- Nonmetric measurement scales which describe differences by indicating the presence or absence of characteristics
- *Metric measurement scales* which are used when subjects differ in a degree on a particular attribute” (O’Gorman & MacIntosh, 2015, p. 163)

These measurement scales are especially used in questionnaires, as most of the quantitative example’s O’Gorman & MacIntosh (2015) use.

Hair, Page & Brunsveld (2019) gives more relevant data analysis methods regarding this thesis research and explains the steps in an efficient way.

“Quantitative data consists of measurements in which numbers are used to directly represent the properties of phenomena. To be useful, the data needs to be analyzed and interpreted. Data analysis in quantitative research involves the following steps:

1. Review conceptual framework and proposed relationships.
2. Prepare data for analysis.
3. Determine whether research involves descriptive analysis or hypothesis testing.
4. Conduct analysis
5. Evaluate findings to assess whether they are meaningful.

Descriptive statistics such as graphics and charts give the researcher a better understanding of the data. Measures of central tendency enable researchers to summarize and condense information to better understand it. The mean, median, and mode are measures of central tendency. Measures of central tendency locate the center of the distribution as well as other useful information. Measures of dispersion describe the tendency for responses to depart from the central tendency (mean, median, and mode). Calculating the dispersion of the data, or how the responses vary from the mean, is another means of summarizing the data. Typical measures of dispersion used to describe the variability in a distribution of numbers include the range, variance, standard deviation, skewness, and kurtosis.” (Hair, Page & Brunsveld. 2019)

The structure of the analysis follows closely how Hair, Page & Brunsveld (2019) describes the data analysis in their book “Essentials of Business Research Methods”.

How do you measure market stability? A measure of volatility provides ample information for an investor to base his decisions. The time to enter, buy or sell are critical decisions that depend
on market moods. Shrewd investors find volatile markets or crashes an ideal time for picking value stocks at bargain prices. Since these are purchased at discounted rates, it gives a sufficient cushion for the lo long-term investor. Though scouting for bargain stocks may appear a lucrative proposition, one must not grab a poorly performing stock that is heading downwards. (Edwards & Zhang, 1998)

4 Results

The data in the results is gathered from a Bloomberg terminal. The results include five indices: ASX 300 (Australia), CAC (France), DAC (Germany), OMX 30 (Sweden) and FTSE 100 (United Kingdom).

When looking at the numbers as graphs, all the indices have similar trends visible between the years 2019 and 2020. That is quite natural, as the trading volumes increased significantly when the COVID-19 pandemic started. Most of the investors, mainly individuals, started panicking and selling their assets when the prices started to decrease. They were afraid and didn’t want or couldn’t afford to lose the value of their assets suddenly. There can be many reasons, one is for example if you have invested with debt when the equity prices have been on a higher level. But the most common reason for sudden sale is simply just fear.

There is not much data about the traders or operators in the sample financial markets. That is why it is rather difficult to understand which trades are conducted by operators using HFT. Even the information regarding the holding time is rather difficult to have access to.
Already, when analyzing the data and calculating the average changes in volumes of the indices, it is possible to understand that the year 2020 when the COVID-19 pandemic started, the change has been huge compared to the previous years and especially when comparing 2020 to 2019. Still, the “peak” of volumes in 2020 is surprisingly low, and actually lower than the level in 2016. But the increase of volumes is higher from 2019 to 2020 than in any other part of the sample.

As we can see from the diagram above, S&P Europe 350 has the highest portion in volumes. That is natural of course, as it contains volumes from more than one country in Europe. The British FTSE 100 and the Australian ASX 300 are almost equal in size when it comes to volumes. The other indices as OMX 30 (Sweden), DAX (Germany) and CAC (France) are country specific, and therefore much smaller in size.
The change in the previous diagram comes even more clear when looking into the standard deviation of all volumes in the indices. From this data it is of course not possible to state that it is caused or even amplified by HFT, but this is a certain indicator of uncertain times in the global financial markets.

4.1 Statistical tests

In the data of this research, the volumes of the different country indices are not directly comparable with each other, they need to be measured by something. Statistical testing can be used for this purpose because it treats the final numbers equally despite their high or low value in volumes.

Having collected several tests, it is necessary to examine what can be tested, and we include here some very general remarks about the general issue of hypothesis testing. Broadly speaking, there are two basic concepts to comprehend before commencing. First, the tests are designed not to prove nor to disprove hypotheses. We never set out to prove anything; our aim is to show that an idea is unsatisfactory as it leads to an unsatisfactorily small probability. Second, the hypothesis we are trying to disprove is always chosen to be the one in which there
is no change; for example, there is no difference between the two population means, between the two samples, etc. Therefore, it is usually referred to as the null hypothesis, H0. If these concepts were firmly held in mind, we believe that the subject of hypothesis testing would lose a lot of its mystique. (Kenji, 2006)

### 4.1.1 F-Test

F-Test in statistics is used to compare two variances and to find out are they equal or not. In this specific research F-Test with critical value is used. In F-Test with critical value the correct alpha needs to be determined. Alpha in this test describes the probability, so if the probability is 35%, the alpha used in the test should be 0.65 (1.00-0.35). According to research by Johannes Breckenfelder the portion of high-frequency trading in Europe is between 24% and 43%. In the US it is even more, around 50%. (Breckenfelder, 2020)

F test is a statistical test that is used in hypothesis testing to check whether the variances of two populations or two samples are equal or not. In an f test, the data follows an f distribution. This test uses the f statistic to compare two variances by dividing them. An f test can either be one-tailed or two-tailed depending upon the parameters of the problem.

To conduct an f test, the population should follow an f distribution and the samples must be independent events. On conducting the hypothesis test, if the results of the f test are statistically significant then the null hypothesis can be rejected otherwise it cannot be rejected. (Markowski & Markowski, 1990)

The f test formula for different hypothesis tests is given as follows:

**Left Tailed Test:**

Null Hypothesis: $H_0 : \sigma_1^2 = \sigma_2^2$

Alternate Hypothesis: $H_1 : \sigma_1^2 < \sigma_2^2$

Decision Criteria: If the f statistic < f critical value then reject the null hypothesis

*Picture 2: Left Tailed Test (Cuemath.com, 2022)*
### Right Tailed Test:

Null Hypothesis: $H_0 : \sigma_1^2 = \sigma_2^2$

Alternate Hypothesis: $H_1 : \sigma_1^2 > \sigma_2^2$

Decision Criteria: If the $f$ test statistic $> f$ test critical value then reject the null hypothesis

*Picture 3: Right Tailed Test (Cuemath.com, 2022)*

### Two Tailed Test:

Null Hypothesis: $H_0 : \sigma_1^2 = \sigma_2^2$

Alternate Hypothesis: $H_1 : \sigma_1^2 \neq \sigma_2^2$

Decision Criteria: If the $f$ test statistic $> f$ test critical value then the null hypothesis is rejected

*Picture 4: Two Tailed Test (Cuemath.com, 2022)*

In Excel this formula is a bit different, but the basic functionality is the same. During the research it also became clear that it is not possible to use many of these statistical tests to confirm an unsure portion of HFT. The portion of HFT cannot be confirmed anywhere and there are only estimates based on the holding times of single equities. The data is usually given by the local stock exchange operators like Nasdaq. Of course, if the holding time is less than one minute, we can be almost certain that the trade is not executed by a human and instead an algorithm or ruled based automatization has been utilized.

### 4.2 ASX 300 (Years 2012-2020)

“The S&P/ASX 300 is a stock market index managed jointly by Standard & Poors and the Australian Stock Exchange. It gives investors a benchmark for Australian companies with market capitalizations above AUD 100 million. It contains small, mid and large cap firms and includes approximately 80% of all Australian equities.” (Capital.com, 2020)

In general, ASX 300 has quite high volumes when compared to its peers in other countries of same size. What makes it interesting, the volumes were actually decreasing instead of
increasing, until 2020 the COVID-19 pandemic happened. For that, there can be many reasons, such as the Australian market has not necessarily been so attractive as the other markets, but those reasons are not analyzed in this research.

The same trend can be seen from both average volume, and variance of ASX 300 index. In the previous chapters of this thesis, it can be seen as a global trend during the COVID-19 pandemic. Therefore, it can be stated that Australia was not isolated from that trend of the financial markets in 2020.

Figure 5: Average volume of ASX 300 index from 2012 to 2020 (Bloomberg terminal, 2021)

Figure 6: ASX 300 index, mean and variance from 2012 to 2020. (Bloomberg terminal, 2021)
The volumes from 2012 until 2020 have been pretty much the same, with a minor decreasing trend. From 2013 to 2019, the trading volumes have been almost static, without any significant changes in volumes.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Average volume</td>
<td>830,708.946</td>
<td>621,396.624</td>
<td>585,051.663</td>
<td>591,498.826</td>
<td>626,149.133</td>
<td>646,361.944</td>
</tr>
<tr>
<td>HFT from average</td>
<td>249,212.648</td>
<td>186,418.987</td>
<td>175,551.699</td>
<td>177,419.648</td>
<td>187,841.740</td>
<td>193,908.183</td>
</tr>
<tr>
<td>HFT Max (30%)</td>
<td>714,251.400</td>
<td>324,417.715</td>
<td>413,949.999</td>
<td>413,813.249</td>
<td>388,370.815</td>
<td>422,713.571</td>
</tr>
<tr>
<td>HFT Min (10%)</td>
<td>40,168.016</td>
<td>60,551.765</td>
<td>247,883.919</td>
<td>254,681.090</td>
<td>134,770.352</td>
<td>227,674.054</td>
</tr>
<tr>
<td>HFT Max vs. Last year average (30%)</td>
<td>112,899.178</td>
<td>71,233.948</td>
<td>77,103.826</td>
<td>72,395.844</td>
<td>59,091.725</td>
<td>77,036.053</td>
</tr>
<tr>
<td>HFT Max higher than average last year (30%)</td>
<td>155</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>HFT Max vs. Last year average (29%)</td>
<td>88,418.783</td>
<td>222,727.872</td>
<td>191,889.660</td>
<td>224,396.526</td>
<td>273,053.490</td>
<td>193,146.818</td>
</tr>
<tr>
<td>HFT Max vs. Last year average (28%)</td>
<td>155</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>HFT Max vs. Last year average (27%)</td>
<td>95</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>HFT Max vs. Last year average (26%)</td>
<td>10,469.996</td>
<td>237,735.020</td>
<td>229,643.326</td>
<td>253,917.209</td>
<td>296,918.315</td>
<td>319,318.399</td>
</tr>
<tr>
<td>HFT Max vs. Last year average (25%)</td>
<td>105</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>HFT Max vs. Last year average (24%)</td>
<td>14,889.693</td>
<td>260,209.643</td>
<td>253,223.360</td>
<td>265,777.653</td>
<td>309,380.573</td>
<td>333,419.175</td>
</tr>
<tr>
<td>HFT Max vs. Last year average (23%)</td>
<td>155</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>HFT Max vs. Last year average (22%)</td>
<td>9,486.793</td>
<td>272,702.567</td>
<td>247,002.593</td>
<td>279,638.092</td>
<td>322,802.931</td>
<td>247,509.961</td>
</tr>
</tbody>
</table>

Figure 7: ASX 300 index, comparison of the possible portion of HFT to the previous year’s average (Bloomberg terminal, 2021)

Figure 8 shows that if the portion of HFT in 2020 was at least above 26% on average during the whole year, the portion of HFT was higher than the average volume in 2019. That means that the daily variation of the portion of HFT can be between 0% to 52%.

In the case of ASX 300, when comparing the years 2012 & 2013, 2013 & 2014, 2014 & 2015, 2015 & 2016, 2016 & 2017, 2017 & 2018, it is possible to tell that the hypothesis where the hypothetical portion of HFT is 30% vs. the average volume of the previous year can be rejected. That is because the potential portion of HFT is lower than the previous year’s average.

When comparing the years 2019 and 2020, the hypothesis (30% market share of HFT vs. average volume of the previous year) cannot be rejected. That is because the possible portion is higher than the previous year’s average starting at 26%. That means that if the yearly average portion of HFT is 26%, the daily variation throughout the year should be between 0% and 52%.

When looking at the previous research of HFT, the portion of Australia is estimated to be below that. Therefore, it is very unlikely that the HFT has amplified the market during the COVID-19 pandemic significantly in Australia.
4.3 FTSE 100 (Years 2012-2020)

“The FTSE 100 Index is a capitalization-weighted index of the 100 most highly capitalized companies traded on the London Stock Exchange. The equities use an investibility weighting in the index calculation. The index was developed with a base level of 1000 as of December 30, 1983. * Please see UKEDA100 Index and FTPTP100 Index for the official FTSE 100 Index Dividend Yield and P/E Ratio*.” (Bloomberg.com, 2021)

![Average Volume of FTSE 100](image)

*Figure 8: FTSE 100, Average volume 2012-2020 (Bloomberg terminal, 2021)*

The volumes of the British FTSE 100 are moving in almost same levels as in the Australian ASX 300, but according to the diagram of average volumes, the market behaves much in a different way compared to its Australian peer. As the volumes in Australia were almost static from 2013 to 2019, in London they were more heavily increasing from 2013 until 2016. After 2016, the volumes were decreasing, almost as heavily as they were raising before that. In 2019, the volumes were just a bit higher than six years before, in 2013. However, there is one similarity with the volumes in FTSE 100 and ASX 300, the volumes in 2020 were almost on the same level as in 2016. That is a common factor with these two indices in addition to the number of scales in the volumes.
The data shows that if the portion on HFT is at least 28% on average through the whole year of 2020, the portion of HFT is higher than the previous year’s average volume. That means that the daily variation of the portion of HFT can be between 0% and 54%. 28% is not a very high average and it is very close by the estimated portion (25% to 40%) of HFT in the financial markets.

The volumes in FTSE 100 have peaked in 2016 and 2020. So, in 2016 the portion of HFT can possibly be even lower than in 2020 to reach the previous year’s average volume. There the daily variation of HFT can be as low as from 0% to 28%. Therefore, the critical point for the average daily portion of HFT is 19%.

In the other years, the hypothetical portion of HFT should be higher than 30% to be higher than the previous year’s average volume and because the portion of HFT is considered to be around 30% of the volumes in Europe, it goes outside of the hypothesis.

4.4 DAX index (Years 2012-2020)

“The German Stock Index is a total return index of 40 selected German blue-chip stocks traded on the Frankfurt Stock Exchange. The equities use free float shares in the index calculation. The DAX has a base value of 1,000 as of December 31, 1987. As of June 18, 1999, only XETRA equity prices are used to calculate all DAX indices.” (Bloomberg.com, 2021)
When testing the data of the DAX index, it can already be seen from the graph of average volumes that there are two peaks. The first peak is in 2018 and the second one in 2020.

![Average Volume of DAX](image)

**Figure 10: DAX index, average volume in 2012-2020 (Bloomberg terminal, 2021)**

During the year 2020, if the average daily volume of HFT is 23% of the total average volume, it will exceed the previous year’s average volume. That means that the daily variation of HFT can be from 0% to 46%.

In 2018, the average volume is even higher than in 2020. Therefore, the hypothetical daily portion of HFT can be even lower than in 2020. The average volume of HFT can be 22% of the total average volume and it will still exceed the previous year’s average volume. That means that the daily variation through the year can be between 0% to 44%. The other years have lower average volumes and are outside of the considered portion percentage of HFT.
4.5 CAC Index (Years 2012-2020)

“The CAC 40® is a free float market capitalization weighted index that reflects the performance of the 40 largest and most actively traded shares listed on Euronext Paris and is the most widely used indicator of the Paris stock market. The index serves as an underlying for structured products, funds, exchange traded funds, options, and futures. It is operated by Euronext, the pan-European exchange.” (Bloomberg.com, 2021)

![Figure 12: CAC index, average volume in 2012-2020 (Bloomberg terminal, 2021)](image)

From the date of CAC index, it can be seen that there are several peaks in the average volume. From 2013 to 2014 and 2014 to 2015 there are few peaks where the volumes have increased steadily. After that there is a significant decrease in the average volumes and the next higher peak can be seen in 2020. Still, the average volumes are lower than for example in 2015.

![Figure 13: CAC index, comparison of the possible portion of HFT to the previous year’s average (Bloomberg terminal, 2021)](image)
In the data set of CAC, there is only one occasion where the hypothetical portion of HFT exceeds the previous year’s average volume and that is in 2020. There, if the average portion of HFT is even 22%, it exceeds the previous year’s average volume. Therefore, the daily variation of HFT can be between 0% to 44% to create a 22% average portion.

4.6 OMX 30 Index (Years 2012-2020)

“The OMX Stockholm 30 Index consists of the 30 most actively traded stocks on the Stockholm Stock Exchange and is a market weighted price index. The composition of the OMXS30 index is revised twice a year. The index was developed with a base level of 125 as of September 30, 1986. Effective on April 27, 1998, there was a 4-1 split of the index value.” (Bloomberg.com, 2021)

![Figure 14: OMX 30 index, average volume in 2012-2020 (Bloomberg terminal, 2021)](image)

As it is possible to see from the graph, the average volumes of OMX 30 decreased rapidly from 2012 to 2014 and started to grow after that. The volumes were growing from 2014 to 2016 and started to decrease after that. Based on this information, it looks like the volumes have gone as cycles in every two years.
In the data of OMX 30 index, there is only one occasion where the hypothetical portion of HFT exceeds the previous year’s average volume and that is in 2020. The hypothetical range is very narrow and the whole year’s daily average of HFT needs to be at least 27% to exceed the previous year’s average volume. That means that the daily variation can be from 0% to 54%.

### 4.7 S&P Europe 350 (Years 2012-2020)

S&P Europe 350 is the only index from the sample which multinational and consists of companies from different European countries. (Bloomberg.com, 2021)

Figure 15: OMX 30 index, comparison of the possible portion of HFT to the previous year’s average (Bloomberg terminal, 2021)

The average volumes of S&P Europe 350 were quite steady until the drop from 2016 to 2019.

Figure 16: S&P Europe 350 index, average volume in 2012-2020 (Bloomberg terminal, 2021)
The data of S&P Europe 350 index has two peaks, one in 2016 and another one in 2020. Still, the hypothetical portion of HFT exceeds the previous year’s average only if the average portion is at 29%. That means also that the daily variation needs to be constantly rather high to exceed that previous year’s average. If the whole year’s average is 29%, that means that the daily variation needs to be between 0% to 58%.

In 2016 the hypothetical portion can be a bit lower, as 25% on average. That still means that the daily variation of the portion in HFT needs to be between 0% to as high as 50%.

5 Discussion

The original topic for this research was to find out if the Australian stock exchange has used their “kill switch” option to control the trading by high-frequency trading (HFT) in their stock exchange. Because the data regarding that is not disclosed anywhere openly and there were no clues that it has ever been used, the topic was changed to give answers if the HFT has amplified the volumes in Australia and Europe from 2012 to 2020, and also keep in mind the hike of the volumes in the beginning of the COVID-19 pandemic.

So, fear makes people more easily press the “sell” button, but it is not so clear when it comes to automated trading or HFT. Depending on the strategy setup (programming), the algorithm will either sell or buy in this type of situation when the prices start to sink. Those investors who have more capital have probably a setup where it starts to buy instead of selling with certain thresholds.

According to several sources and research, it is very difficult to estimate the portion of high-frequency trading globally or even by each continent or country. Also, the regulation regarding HFT is not united and instead scattered. Every country has different rules when it comes to
HFT and they can even vary between different stock exchange operators in the same country. Especially in those cases if the local financial regulator has not pre-defined the rules very clearly. Some countries do not have any regulation at all, but usually the operator at least has a disclosure in their rules which allows them to stop trading if they notice something suspicious from their point of view. Because usually the word suspicious is not defined or it is defined very broadly, so it also covers HFT when considering suspicious. Preventing to harm the financial markets is also one rule which is commonly used in different stock exchanges around the world. Overall, the stock exchanges usually have covered the control of HFT in some way or another in their rules.

The sample in this research consisted of six different indices, five from Europe and one from Australia. The years in the sample were from 2012 to 2020. They were tested by using a hypothetical portion of HFT and compare will it exceed the previous year’s average volume. Why not choose any indices from the US? That is because the volumes are significantly higher in the stock exchanges of US and the portion of HFT is as well considered to be much higher than in Australia and Europe. According to several sources, the portion of HFT in the US is suspected to be even between 40% to 65% in certain exchanges which makes it rather difficult to compare with the European and Australian indices, where the portion is believed to be approximately between 25% and 35% of the total volumes.

The existing theory about the market share of HFT in Europe matches quite well when looking at the daily averages and the thresholds of needed portions of HFT used in 2020 to not reject the hypothesis (30% HFT market share vs. average volume of the previous year). Even though many of the estimates are almost 10 years old, like the 24% estimate by ESMA from 2014, it is still very close to the values needed in the hypothesis to be able analyse the results and not to reject the hypothesis.

Overall, it is very difficult to find trustworthy information regarding the volumes of high-frequency trading. One reason behind it is also that the algorithms which the companies use are part of their investment or business strategy and therefore, a very crucial part of their income. That is why the research in this master’s thesis relies heavily on estimates and hypothetical questions.
Daily variations in 2020 (average):
ASX 300 = 0% - 52% (26%)
FTSE 100 = 0% - 56% (28%)
DAX = 0% - 46% (23%)
CAC = 0% - 44% (22%)
OMX 30 = 0% - 54% (27%)
S&P Europe 350 = 0% - 58% (29%)

Because some of the indices in the sample are European indices, the companies in them are actually included in the S&P Europe 350 index. Only companies in the ASX 300 index are not included in the S&P Europe 350 index. So, it can be said that comparing some of these indices with the S&P Europe 350 index can give overlapping results. Anyway, the index is not directly compared with the other indices, instead it gives a good overall picture of the volumes across Europe, as it also includes companies from the other markets which are not a part of the thesis.

It is very difficult to find scientific research with estimates of the market share of HFT in each continent. It is possible to find commercial research but based on their funding, there is always a possibility that the research is somewhat biased by it. According to a company called Grand View Research, around 40% of the capacity of the HFT servers globally in 2020 was used to equity trading. They were almost equally used in Forex (Foreign Exchange Markets). Their research is global, and it gives some clues about which direction HFT and its market share is going to be in the future. Only the server market in 2020 is estimated to be 387.9 million US dollars. From the commercial research it is possible to see that the equity trading is estimated
to be slightly larger than the Forex market, as in the non-commercial research like Yadav’s (2015) the Forex market is estimated to hold a larger share of the HFT.

![Figure 19: Global HFT servers markets share, by application in 2020 (grandviewresearch.com, 2023)](image)

While writing this thesis, it was quite clear that most of the previous research was done only about North America. Therefore, it was challenging to find out any peer-research to verify the results. There were a few very good research articles, especially the one done by Johannes Breckenefelder from the European Central Bank (cepr.org). Also, the American research is useful in a way that usually Europe and the rest of the world follows what starts in the US.

Still, there was no research from Europe which would clearly be based on the holding times with trades executed on the stock exchanges. Therefore, it is difficult to verify the results and to be sure that the findings are correct or remarkable.

### 6 Conclusions

This thesis originally started with a regulatory angle, which was to find out if high-frequency trading or algorithmic trading can be regulated and if so, has it worked this far. The research was aimed to Australia, as many news sources were writing articles about how the regulator down there has a demand to control HFT. It ended up as a question of where to find material for this research. After several months of browsing public online sources, it became clear that
it is not possible to find a reliable source for this information. One reason for that is that the trades are not publicly available online and even if they were, it would be difficult to identify which trades were executed by using HFT or algorithmic trading. Therefore, the topic and the hypothesis of the research were changed to be able to find more information from open sources. The hypothesis on the thesis was actually changed several times from a statistical F-Test to a less statistical. That is because the F-Test did not give results which could be analyzed in a clear way.

COVID-19 pandemic started while writing this thesis and the market turmoil was very fast. That is why instead of having a very long timeframe, this research tries to concentrate more on the time of 2020 when the pandemic started and moved the trading volumes around the world. So, the research question in this master’s thesis is “Has high-frequency trading amplified the volatility of the financial markets before and during COVID-19 pandemic?” The results are interesting but cannot provide undisputable proof that HFT has amplified the market volatility. There are many reasons why that is. First of all, the portion of algorithmic trading or high-frequency trading of all trades is not proven. There are only estimates from several researchers, where most of them agree that it is around 25%-35% in the European market.

If looking at the constantly evolving technology like Chat GPT and other solutions created by artificial intelligence, it is possible to understand that this portion of trades made by different technological innovations will certainly increase in the near future. If technology is evolving fast, when looking at the history of trading, it can cause some unwanted consequences such as flash crashes in the financial markets. That is why the regulation needs to evolve on the side of technology to prevent losing control.

Back to the research question. It can be seen from the diagrams, that the COVID-19 pandemic definitely shocked the financial markets and increased the volatility in all continents globally. But after analyzing the volumes of each index measured in this research, it cannot be proven that the high-frequency trading has amplified the market volatility before or during the COVID-19 pandemic. First of all, there are no clear patterns or similarities even when looking at the indices only in Europe. Secondly, only the Australian ASX 300 index gives a result of hypothesis not rejected when comparing the years 2019 and 2020. At the same time, the same test gives a negative result of “hypothesis rejected” in all other indices. So, despite the high
volatility during the years of 2019 and 2020, it cannot be stated that the high-frequency trading has amplified the volatility even more.

The test is done by using different assumed portions of HFT. Basically, if the portion is equal to the growth of the volume compared to the previous year, it will be higher. That is why this type of comparison is definitely not flawless. Despite the possible flaws in the test, when rationally thinking it would be very unlikely that only the HFT has amplified the volatility during the COVID-19. Many financial institutions create trades in batches to minimize their trading costs. That is not directly falling into the category of what HFT is. Many financial institutions also use algorithms in their trading but usually there is a person who controls the process.

But for that can be said for sure, that if the portion of the HFT is the same every year, it will not amplify the market volatility more than it would grow without it. If the portion grows exponentially, it can of course raise the volumes significantly, as most of the trades becomes even faster and might trigger the “traditional” traders to act at the same time. In the past there have been these kinds of domino effects in the markets caused by HFT. Now in 2023, when the AI “boom” is going on, based on what we have seen before, it is very important that the regulators follow the evolution of AI and how it is related to trading. Otherwise, we might face even larger market-flash crashes than we have ever seen before.
References


Web pages:


