CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

Analysis of the US Stock Market during the COVID-19 pandemic

Bachelor's thesis

Author: Adam Tůma
Study program: Economics and Finance
Supervisor: prof. PhDr. Ladislav Krištoufek, Ph.D.
Year of defense: 2021

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, May 3, 2021

Adam Tůma

Abstract

This work investigates the effect of the COVID-19 pandemic on the S&P 500 stock index and its eleven sectors. Employing the ARMA and the T-GARCH model on a time series of daily returns from 2018 until March 2021, we examine the impact on volatility, returns, and day-of-the-week effect during the stock market crash caused by the pandemic and the period after. Our main findings imply that in the case of returns, the Monday effect was more negative than the Friday effect during the market crash and vice versa in the rising market after the crash. Concluding that the calendar time hypothesis holds for the observed periods. In terms of volatility, it drastically increased across the US stock market during and even after the crash. The increase was especially noticeable for the IT and Energy sectors. We also found the U-shaped daily volume pattern changed significantly with proportionately less volume of trades happening in the first half-hour of trading and more throughout the whole day.

JEL Classification	C22, C52, C58, G01, G14
Keywords	GARCH, pandemic, stock market, trading volume, volatility

Title	Analysis of the US Stock Market during the COVID-19 pandemic

Author's e-mail	adamtuma98@gmail.com
Supervisor's e-mail	ladislav.kristoufek@fsv.cuni.cz

Abstrakt

Tato práce zkoumá vliv pandemie COVID-19 na americký akciový index S&P 500 a jeho jedenáct sektorů. Pomocí modelů ARMA a T-GARCH na časové řadě denních výnosů od roku 2018 do března 2021 zkoumáme dopad na volatilitu, výnosy a efekt dne v týdnu během krachu akciového trhu způsobeného pandemií a období po něm. Naše hlavní zjištění naznačují, že v případě výnosů byl pondělní efekt během propadu trhu víc negativní než páteční efekt a na rostoucím trhu po pádu naopak. Docházíme tedy k závěru, že pro sledovaná období platí hypotéza kalendářního času. Pokud jde o volatilitu, ta dramaticky vzrostla na celém americkém akciovém trhu jak během propadu, tak i po něm. Nárust zaznamenala zejména odvětví informačních technologií a energetiky. Zjistili jsme také, že se významně změnil denní objemový vzor obchodování ve tvaru písmene "U" s úměrně nižším objemem obchodů v první půlhodině obchodování a vyšším během zbytku dne.

Klasifikace JEL	C22, C52, C58, G01, G14
Klíčová slova	akciový trh, GARCH, objem obchodů, pandemie, volatilita

Název práce	Analýza akciových trhů v USA během pandemie COVID-19

E-mail autora	`adamtuma98@gmail.com`
E-mail vedoucího práce	`ladislav.kristoufek@fsv.cuni.cz`

Acknowledgments

I am very grateful to the supervisor prof. PhDr. Ladislav Krištoufek, Ph.D. for the valuable suggestions and comments provided during the writing of the thesis. I would also like to thank prof. Roman Horváth Ph.D. for insights provided during lectures on topics used in this work.
Mainly I am grateful towards people close to me, who supported me and gave me the opportunity to pursue university studies as I will never take this for granted.

Typeset in FSV LaTeX template with great thanks to prof. Zuzana Havrankova and prof. Tomas Havranek of Institute of Economic Studies, Faculty of Social Sciences, Charles University.

Bibliographic Record
Tůma, Adam: *Analysis of the US Stock Market during the COVID-19 pandemic.* Bachelor's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2021, pages 65. Advisor: prof. PhDr. Ladislav Krištoufek, Ph.D.

Contents

List of Tables

List of Figures

Acronyms

ACF Autocorrelation Function

ADF Augmented Dickey–Fuller

ANOVA Analysis of variance

AR Autoregressive

ARCH Autoregressive Conditional Heteroscedasticity

ARMA Autoregressive Moving-Average

EMH Efficient Market Hypothesis

GARCH Generalized Autoregressive Conditional Heteroscedasticity

iid independent and identically distributed

MA Moving-Average

NYSE New York Stock Exchange

PACF Partial Autocorrelation Function

QE quantitative easing

Bachelor's Thesis Proposal

Author	Adam Tůma
Supervisor	prof. PhDr. Ladislav Krištoufek, Ph.D.
Proposed topic	Analysis of the US Stock Market during the COVID-19 pandemic

Research question and motivation The main research question I intend to study is the effect the COVID-19 pandemic has had on the US stock market.

I would like to find out whether some new patterns can be observed on the stock market during a pandemic period, which are otherwise not present.

One of the possible patterns could be the well known day of the week effect that had been analysed many times before, in fact the first analysis was done in a journal article in the 1970s when it was shown that there is a tendency for the market to rise on Fridays and decline on Mondays (Cross, 1973). Over the years, many other analyses researching this phenomenon have been done, but I am especially interested in the possible patterns that may have arisen during the panic that was caused on the stock market due to the COVID-19 pandemic. I am keen to explore this, as my hypothesis is that the Monday effect was very strong in the first stages of the pandemic as more and more bad news were being published, however in my opinion this effect was rather short-lived as the initial panic on the market did not last very long.

During the last financial crisis in 2008, the stock market was hit especially hard in the financial sector with very high volatility (Schwert, 2011). However, that was largely due to the fact that the crisis started with the collapse of the housing markets, which was closely tied to the financial market, however, the COVID-19 crisis is very different. I am interested to test, whether some strong similarities can be seen in the behaviour of the stock market with what was researched for the Great Recession, or the market was hit much stronger in other fields. We can already see that companies in the technological sector are performing very well and I would like to explore the effects on various sectors and compare the results. I hypothesize that the market volatility was very high across all sectors in the early stages of the pandemic, and then slowly forming some patterns. With the technological sector receiving a lot of attention during the crisis, I would assume higher volatility for this sector than usual.

My thesis will cover both the possible patterns in form of the day of the week effect as well as the patterns of volatility or trading volume across multiple sectors traded on the US stock market.

Methodology I will use publicly available sources (e.g. dukascopy.com, finance.yahoo.com etc.) where I can gather data points for the whole US stock market using the Standard Poor's index of 500 largest companies traded on the US stock market, as well as multiple other stocks representing the different fields that are being traded on the stock exchange (technology, financial sector, pharmaceuticals, energy. . .).

After gathering the data, I will analyze the day of the week effect, the volume being traded and the volatility during the given period. Then I will match the results from the COVID-19 period with results for past years and compare them with previous researches done and evaluate my hypotheses.

I would like to use statistical tests to measure the day-of-the-week effect as well as some more advanced econometric models like the GARCH model for estimating the market volatility.

Expected Contribution The topic of the COVID-19 pandemic and its effects on all types of markets is very fresh with still very little literature covering it. Which gives a great opportunity for this thesis to provide some interesting data, that have not yet been covered.

The main contribution this thesis will provide is the behaviour of the investors on the stock market during a crisis like this one. It will be very interesting to provide solid data on the different sectors of the economy and the impact the crisis has had on them.

Outline

1. Introduction
    - Motivation behind the topic
    - Contribution
    - Organization of the thesis
2. Literature review and hypotheses
    - Literature on stock market analysis
    - Evidence of existing patterns on the market

- Hypotheses

3. Data and methodology
   - Description of the data
   - Description of the methods used in the analysis
   - Explaining how I perform the analysis
4. Results
   - Interpreting the results
   - Evaluating on hypotheses
5. Conclusion
   - Broader interpretation and commenting on the results
   - Implications for practice
   - Topics for further research

Core bibliography

Bhownik, R., Wang, S. 2020. Stock Market Volatility and Return Analysis: A Systematic Literature Review. International Journal of Economics and Management Systems, 5: 217-232

Cross, F. 1973. The Behaviour of Stock Prices on Fridays and Mondays. Financial Analysts Journal, 29: 67-69

Draper, P., Paudyal, K. 2002. Explaining Monday Returns. Journal of Financial Research, 25 (4): 507-520

French, K. R. 1980. Stock returns and the weekend effect. Journal of Financial Economics, 8 (1): 55-69

Schwert, G. W. 2011. Stock Volatility During the Recent Financial Crisis. European Financial Management, 17: 789-805

Tsay, R. S. 2010. Analysis of Financial Time Series. 3rd ed. Hoboken, New Jersey: John Wiley Sons

Author | Supervisor

Chapter 1

Introduction

On March 16, 2020, the New York Stock Exchange (NYSE) shut down all trading for 15 minutes for the third time within eight days. The trading halt is done automatically whenever the S&P 500 index loses 7% compared to the previous day's closing price. The last time the circuit breaker was triggered before the pandemic crash was during the financial crisis in 2008. While the novel coronavirus was first discovered in Wuhan, China, at the end of 2019, and the first confirmed case in the USA was on January 21, 2020, there was no considerable attention regarding the situation visible on the US stock market as it still peaked on February 19, 2020. However, once the prices started to fall, the market quickly turned into a bear market[1]. The uncertainty of the evolving pandemic situation caused stock markets all around the world to crash, the S&P 500 specifically plummeted more than 34% by March 23, 2020, to 2,237 index points, price levels last seen at the end of 2016. On March 23, 2020, FED[2] announced extensive measures to support the American economy with unlimited quantitative easing (QE). In matter of weeks, FED's balance sheet increased from around 4 trillion USD before the pandemic to more than 7 trillion USD by June 2020. The announcement of unlimited QE proved to be the turning point at which the US stock market started regaining its losses and turned into a bull market once again. The S&P 500 rebounded to a new all-time high on August 18, 2020, closing at 3,389 points. As of the thesis' writing, on April 16, 2021, the index closed at 4,185 points, up 86% from the lowest point over a year ago. And while it is obvious the pandemic was the cause of the stock market crash in 2020, the negative effect was relatively short-lived as

[1] Bear market describes the state of the market in which prices are declining and is often more volatile as opposed to bull market in which the prices are rising and volatility declines.

[2] Short for Federal Reserve is the central banking system in the USA.

the pandemic still affects millions of people worldwide and has killed over half a million people in the USA alone. On the other hand, this opens opportunities for analysis of both periods, firstly the market crash between February and March of 2020, and secondly the period of high growth after that.

The objective of this thesis is to look for inefficiencies on the US stock market that may have arisen during the pandemic, thus observe some predictable patterns on the market. We will analyse the day-of-the-week effect, namely the Monday and Friday effect using conditional heteroscedastic models as the uncertainty of the evolving pandemic situation may have largely impacted investor's behaviour around weekends. We use data from S&P 500 index and also data for eleven sectors of the S&P 500 to measure the impact across sectors of the economy. As we mentioned, the pandemic caused the stock market to turn into a bear market only for about two months, which a strong bull market has followed ever since then. The effects in these two states of the market can differ significantly, which is why we will analyse the effects for the bear and bull market separately and get seperate results for the different time periods using dummy variables. We will also explore the intraday volume pattern, observing how the trading activity splits out during the day and how the pandemic affected it.

The thesis is further divided into the following chapters. Chapter 2 reviews the available literature relevant for our study and includes our hypotheses. Chapter 3 provides information about the data used and the outlined time periods for the purpose of this analysis. Chapter 4 describes the methodology used both for the conditional heteroscedastic models and the intraday volume analysis. Chapter 5 provides the empirical results of the analysis. And finally, Chapter 6 summarises our findings.

Chapter 2

Literature review and hypotheses

The amount of available literature on stock markets and the US stock market is explicitly vast, but it by no means implies that there is no room for new meaningful studies. Stock markets, in general, are very dynamic as they reflect many different variables on top of the traded companies' actual valuation, whether those variables are new trends among the traded companies, change in monetary policies, macroeconomic factors, current political situation, or perhaps most recently the COVID-19 pandemic.

This section will point out the relevant studies that have been done both historically and recently regarding market anomalies such as the day of the week effect or intraday volume patterns. We also cover recent literature related to COVID-19 effects on the stock market.

2.1 Market anomalies

The day of the week effect was initially studied by Osborne (1962), who studied various patterns on the US stock market. His results showed that the market tends to decrease on Mondays more likely than on other days. Many further studies in the 20th century were done studying the phenomenon of the day of the week effect (Cross 1973; French 1980; Gibbons & Hess 1981; Jaffe & Westerfield 1985; Jaffe *et al.* 1989; Lakonishok & Maberly 1990). The studies' main conclusions include low average returns on Mondays, while Fridays tend to show higher returns than other days of the week.

Cross (1973) focused mainly on the Monday and Friday effect and concluded that the average returns on Mondays are significantly lower than on other days while also being heavily correlated to previous Friday returns. The correlation was also shown by Jaffe *et al.* (1989) whose results showed that the Monday effect virtually disappears when the market has previously risen.

French (1980) developed two hypotheses for the Monday effect. Firstly, under the calendar time hypothesis, the average returns on Mondays should be three times higher than on other days as there are two non-trading days between Friday and Monday. And secondly, under the trading time hypothesis, the average returns on Mondays should be the same as for other days as returns are being generated only during trading time. His results showed that from 1953 to 1977, average Monday returns were significantly negative, while returns for other days of the week were positive. Thus both the trading and calendar time trading hypotheses were rejected. Kiymaz & Berument (2003) used data from S&P 500 between years 1972 and 1997, to measure the day of the week effect both in terms of return and volatility. Their findings show the lowest returns on Mondays and the highest returns on Wednesdays, which are also the most volatile based on their results.

Many studies are trying to find the possible explanations for the abnormally low Monday returns. Jain & Joh (1988) found that the volume of S&P 500 stocks traded on Monday is the lowest across all days and roughly 10% lower than the average of other trading days, with Wednesday volume being the highest. Lakonishok & Maberly (1990) analysed the trading patterns of individual and institutional investors, and their results showed that individual investors tend to increase their activity on Mondays, concluding that the Monday effect might be due to the patterns in trading of individual investors.

While most of the famous studies covering the day of the week effect come from the previous century, there are still plentiful recent studies covering the topic. Zhang *et al.* (2017) employed the GARCH model on rolling samples of 500, 1000, 1500 days to show the significance of the day of the week effect on multiple stock markets. For S&P 500 index, they found that recently mainly Monday, Tuesday and Friday effect were statistically significant, especially ever since the end of the financial crisis of 2008-09, after which the US stock market has been in a very optimistic state also known as a bull market. Ülkü & Rogers (2018) employed GJR-GARCH on data from Asian stock markets split by investor identification. Their finding showed that institutions' refraining from trading on Mondays emerges as a partial explanation of the Monday effect. It further supports earlier results of Lakonishok & Maberly discussed in the previous paragraph.

Another anomaly relevant for our analysis is the intraday volume trading pattern. Earlier studies show that the average trading volume across daily trading hours differs significantly on the US stock market. The highest activity occurs during the first trading hour and declines slowly until rising again at the end of the day, thus following the classical U-shape (Jain & Joh 1988; Admati & Pfleiderer 1988). Gao *et al.* (2018) looked for intraday momentum on the US stock market, and their results show there is a significant relationship between first and last half-hour returns when predicting the latter. Interestingly, the estimator coefficient got even more vital during the 2008-09 financial crisis. Their data from the year 1993 to 2013 also show a U-shaped volume pattern. However, the average traded volume in the last half-hour is slightly higher than in the first one. Richards & Willows (2019) researched the trading activity of individual investors in the UK. Interestingly, the results showed a W-shaped trading pattern with the highest activity in the first half-hour with smaller peaks in the middle and at the end of the day.

As we can see, market anomalies have been studied many times before, and the results of the studies can sometimes differ as the stock market is constantly evolving thanks to the development of new technologies and by receiving more attention from the public. We believe analysing the mentioned anomalies on recent data from the US stock market affected by the pandemic is currently missing in the literature, allowing us to obtain new interesting results.

2.2 Stock market literature related to COVID-19

Since the beginning of the pandemic, countless pieces of literature have been published regarding the COVID-19 pandemic and stock markets. Baker *et al.* (2020) provide supporting evidence based on newspaper articles that the pandemic was the driving factor of a very turbulent period on the US stock market in the early stages of the pandemic at the end of February through April 2020. Their results also claim that no other pandemic in history had such an impact on the US stock market.

Zhang *et al.* (2020) compared data from February and March 2020 to measure the effect the pandemic has had on the volatility of various stock market returns. Their results for the US showed that the standard deviation in February was at 0.0069 and almost quadrupled to 0.0268 in March, the highest increase in volatility levels among the measured stock markets. Onali (2020) studied the impact of COVID-19 cases and deaths in the US and other six countries majorly affected by the pandemic on the US stock market returns.

Only the numbers reported in China showed a significant impact on the US market returns using the GARCH model. Uddin *et al.* (2021) used data from 34 developed and emerging markets and found that various economic factors can reduce the magnitude of volatility in uncertain time like a global pandemic.

Smales (2021) utilised google search volume for the "coronavirus" keyword as a proxy for investor attention to measure its correlation with stock returns. His results showed a significant negative relationship. The peak of coronavirus search volume in the US happened one day before S&P 500 dipped to its lowest level during the pandemic. Okorie & Lin (2021) employed the martingale difference spectral test to evaluate the adaptive Efficient Market Hypothesis (EMH) and found no substantial change in the level of information efficiency for the US stock market.

The majority of studies only focus on the whole US stock market by taking data from either S&P 500 or DJIA indices. There are not many studies illustrating the impact across different industries of the market. However, Baek *et al.* (2020) took a closer look into the industry-level impact of COVID-19 on the idiosyncratic and systematic risk by utilising Markov switching regime model. Their results show a significant increase of total and idiosyncratic risk across all 30 examined industries. An increase in systematic risk is present for telecommunications or utilities, while industries such as automobiles and business equipment experienced a decrease in systematic risk. Just & Echaust (2020) also used Markov switching approach to determine the relationship between S&P 500 returns and various market indicators during the pandemic. Their results show close dependence between returns and both implied volatility and implied correlation. Choi (2021) tested the EMH by applying multifractal detrended fluctuation analysis on US stock market's individual sectors, his finding shows highest efficiency for the consumer discretionary sector and lowest for the utilities sector during the pandemic crash.

2.3 Hypotheses

We believe a closer look at the day of the week effect during the COVID-19 pandemic is missing in the current literature. It can be highly valuable as the market's dynamic was primarily affected by the uncertainty of the evolving situation and the constant inflow of both good and bad news. Possible patterns that might arise from the results would greatly benefit investors' decisions for future crashes on the stock market.

As we analyse three different periods, in two of which the market experienced strong growth and one that saw the S&P 500 tumble by more than a third in a matter of days, it will be interesting to see the development of day of the week effect both in terms of average returns and volatility. We believe that we will see the Monday effect being the most significant in all cases. However, we expect a strongly negative relationship for the crash period instead of a positive one for the bull period, as the inflow of bad news regarding the pandemic development will cause investors to sell largely on Mondays. The other scenario is the one closer to the EMH and more specifically to the assumption stating, that investors are rational and always take the right decision (Fama 1970), which would suggest investors realising this weekend development beforehand and reflecting it on Friday returns, thus negating the Monday effect. We expect the day of the week effects in terms of returns to be the same for all eleven sectors of the S&P 500 across all periods.

In terms of volatility, we anticipate more interesting cross-sector results. We expect volatility increase mainly for Information Technology, Health Care and Financials sectors. In contrast, Consumer Staples or Utilities sectors should not be much affected as the included companies mainly profit from necessities. The overall market volatility during the crash period will undoubtedly increase.

We also believe investigating the intraday volume patterns will bring significant results even with a small number of observations. Historically, we can see a heavily U-shaped trading pattern in terms of volume during the day. It is believed to be caused mainly due to all earnings and significant macroeconomic news being released before the market opens, causing this information to be processed heavily during the first 30 minutes of trading, after which the market cools off (Gao *et al.* 2018). While on the other end of the day, institutional traders put enormous emphasis on closing prices as they use them to calculate portfolio returns and other various financial indicators. At the same time, market makers often want to avoid overnight risk and close their positions (Cushing & Madhavan 2000). However, during the COVID-19 pandemic bear market, these factors probably did not play such a significant role. We expect flattening of the U-shape curve and mid-day trading volumes accounting for more of the overall daily volume.

Chapter 3

Data

This section will cover information about the data used in the research. We will also outline the time periods for the proposed analysis. We want to understand how certain behaviours changed in response to the pandemic directly and in the period after the crash when the market started rising again.

The information and data from Section 3.1 were obtained from S&P Global wesbite[1]. The data for second part of our analysis mentioned in Section 3.2 were downloaded from Dukascopy[2].

3.1 Standard and Poor's 500 index (S&P 500)

As we analyse the effect on the US stock market, the S&P 500 is an excellent representation of the whole market. The index comprises 500 leading US companies on the stock market, and it covers approximately 80% of available market capitalisation. This indicator's history goes back to 1957, when it was created as the first US market-cap-weighted stock market index. In 2005, the calculation of the index was changed to float-adjusted market-cap-weighted index, which better reflects the value being traded by investors as only so-called free-float[3] shares are weighted in, meaning that the weight of each stock is calculated based on the following formula:

$$Weight_i = \frac{Float\ Adjusted\ Market\ Cap_i}{\sum_{n=1}^{500} Float\ Adjusted\ Market\ Cap_n}$$

[1] `https://www.spglobal.com/spdji/en/landing/investment-themes/sectors/`

[2] `https://www.dukascopy.com/swiss/cz/marketwatch/historical/`

[3] shares which can be publicly traded and are not restricted, i.e. not held by company management

Our analysis will be working with the more common price return version of the S&P 500 index, which only reflects the stock prices, excluding the paid out dividends. We downloaded daily close index values from the beginning of 2018 until March 15, 2021. In our model we are working with dummy variables focusing on three different periods outlined as follows:

1. **Before:** Starting from the earliest observation of the dataset on January 2, 2018, until January 31, 2020, when the Trump administration declared a public health emergency.

2. **During:** The main pandemic crash period after the declared emergency with first observation on February 3, 2020 until March 20, 2020, when massive QE with positive effect got announced the weekend after the last observation.

3. **After:** Period largely affected by endless QE from March 23, 2020 to March 15, 2021.

Figure 3.1: S&P 500 daily index values, main index and top 5 sectors

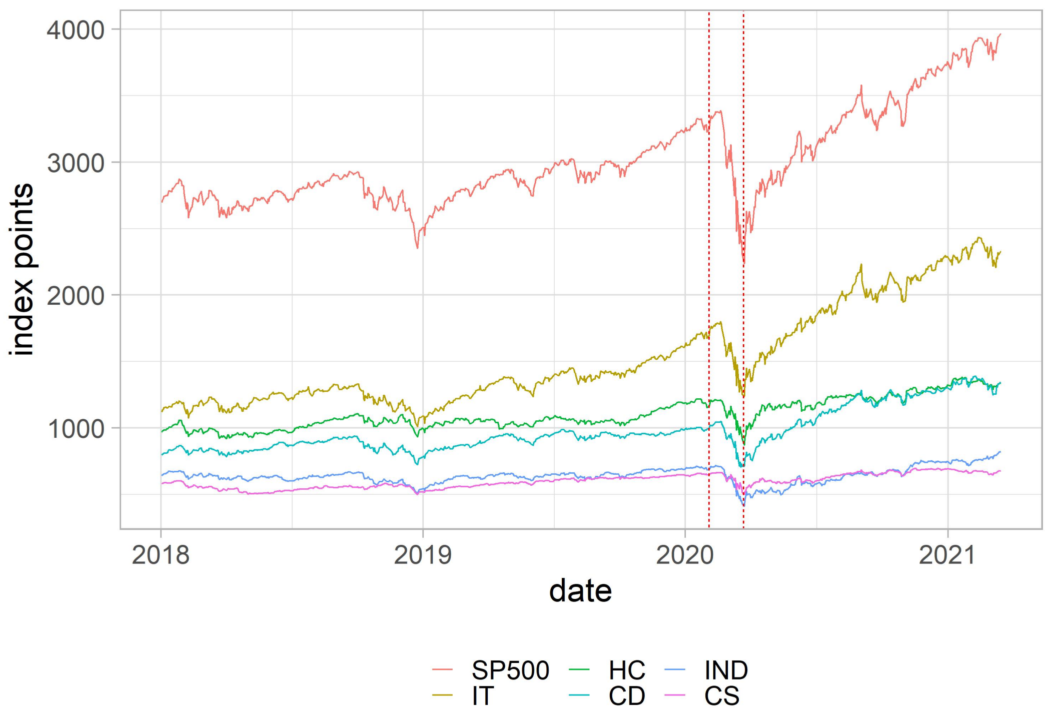
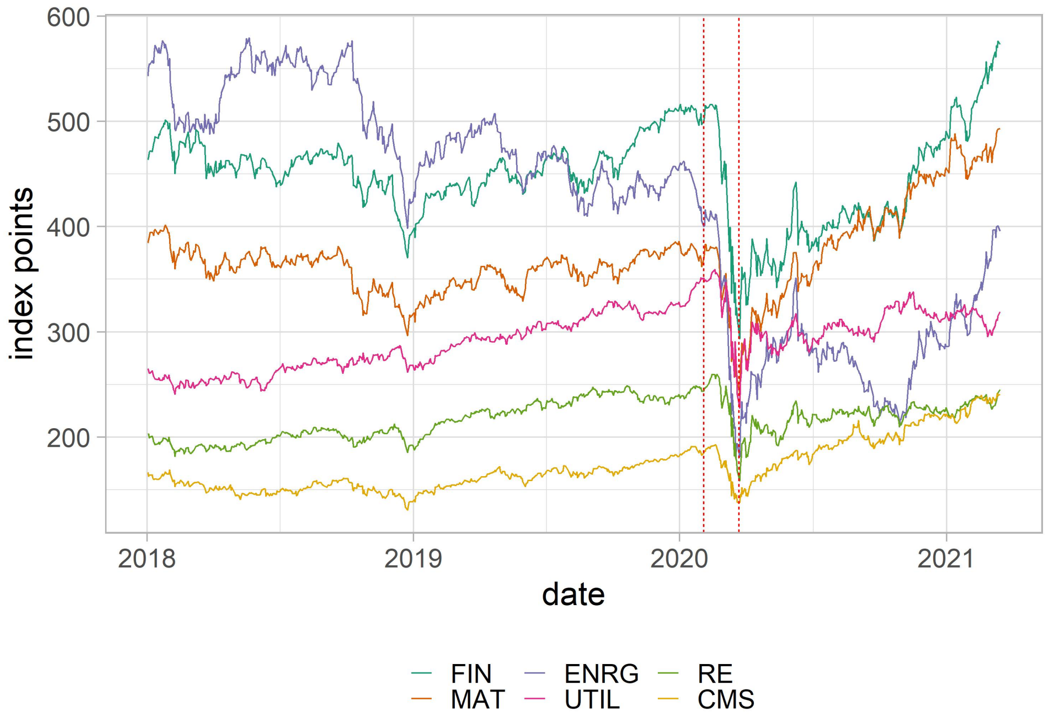


Figure 3.2: S&P 500 daily index values, bottom 6 sectors



3.1.1 Sectoral breakdown of the S&P 500 index

The 500 stocks included in the S&P 500 are broken down into 11 different sectors. The first part of our analysis is conducted on all of the 11 sectors as well as the S&P 500, to be able to measure the impact on each part of the economy.

The obtained datasets are graphically visualized in Figure 3.1 and Figure 3.2 with different periods outlined by the dotted line.

In the following part, we will give a brief description of the sectors and list the top 5 constituents by index weight for each of them (as of this thesis's writing) (S&P Global 2021). A broad understanding of the differences between specific sectors is crucial in interpreting the results in our analysis.

Information Technology - IT

The Information Technology sector includes 76 companies that offer software and information technology consulting, data processing, and also manufacturers and distributors of technology hardware and equipment such as cell phones, computers, semiconductors etc...

Top 5 constituents: Apple Inc. (AAPL), Microsoft Corp (MSFT), Visa Inc A (V), Nvidia Corp (NVDA), Mastercard Inc A (MA)

Health Care - HC

A total of 63 constituents in the Health Care sector, such as healthcare services and providers, manufacturers of healthcare equipment and supplies, technology companies related to healthcare, and biotechnology and pharmaceutical companies.

Top 5 constituents: Johnson & Johnson (JNJ), Unitedhealth Group Inc (UNH), Abbott Laboratories (ABT), AbbVie Inc. (ABBV), Pfizer Inc (PFE)

Consumer Discretionary - CD

This sector includes 61 companies closely operating in the retail sector, such as automobile manufacturers, household durables, leisure equipment or apparel stores. Hotels, restaurants and other leisure facilities are also included.

Top 5 constituents: Amazon.com Inc (AMZN), Tesla Inc (TSLA), Home Depot Inc (HD), Nike Inc B (NKE), McDonald's Corp (MCD)

Industrials - IND

The Industrials sector is comprised of 73 businesses that manufacture and distribute capital goods such as machinery, electrical equipment, building products, aircraft and defence. The sector also includes commercial services providers such as construction and engineering, printing, logistics, human resources, or consulting services.

Top 5 constituents: Honeywell Intl Inc (HON), Union Pacific Corp (UNP), Caterpillar Inc (CAT), United Parcel Service Inc B (UPS), Boeing Co (BA)

Consumer Staples - CS

Consumer staples sector is similarly focused on retail as the Consumer Discretionary sector, however, this sector includes 32 companies that produce more non-durable goods such as manufacturers and distributors of food and beverages, tobacco, non-durable household goods and personal products.

Top 5 constituents: Walmart Inc (WMT), Procter & Gamble Co (PG), Coca-cola Co (KO), Pepsico Inc (PEP), Costco Wholesale Corp (COST)

Financials - FIN

The Financials sector consists of 65 businesses such as banks, thrifts, insurance companies, brokers, asset management companies and other types of businesses closely tied to finances.

Top 5 constituents: Berkshire Hathaway B (BRK.B), JP Morgan Chase & Co (JPM), Bank of America Corp (BAC), Wells Fargo & Co (WFC), Citigroup Inc (C)

Materials - MAT

This sector includes 28 companies that manufacture construction materials, glass, paper, chemicals, forest products, packaging products. Mining companies and producers of steel are also included.

Top 5 constituents: Sherwin-Williams Co (SHW), Ecolab Inc (ECL), Air Products & Chemicals Inc (APD), Freeport-McMoran Inc (FCX), Newmont Corporation (NEM)

Energy - ENRG

The Energy sector includes 23 companies operating in exploration, production, refining, storage and transportation of oil and gas, coal and consumable fuels.

Top 5 constituents: Exxon Mobil Corporation (XOM), Chevron Corporation (CVX), ConocoPhilips (COP), EOG Resources Inc (EOG), Schlumberger Limited (SLB)

Utilities - UTIL

The Utilities sector includes 28 companies such as gas, electric and water utilities, and independent power producers and energy traders. The sector also includes companies that generate and distribute electricity obtained from renewable resources.

Top 5 constituents: Nextera Energy Inc (NEE), Duke Energy Corp (DUK), Southern Co. (SO), Dominion Energy Inc (D), Sempra Energy (SRE)

Real Estate - RE

This sector includes 30 companies active in the development and management of real estate as well as equity real estate investment trusts (REITs), however, mortgage REITs are included in the Financials sector.

Top 5 constituents: American Tower Corp (AMT), Prologic Inc (PLD), Crown Castle International Corp (CCI), Equinix Inc (EQIX), Public Storage (PSA)

Communication Services - CMS

Communications services sector includes 26 companies that provide content such as entertainment, news and social media primarily on the internet, also includes providers of internet, broadband, cellular, cable and landlines.

Top 5 constituents: Alphabet Inc Class C/A (GOOG/GOOGL), Facebook Inc (FB), Walt Disney Company (DIS), Comcast Corp (CMCSA), Verizon Communications (VZ)

3.2 USA500 intraday data

For the intraday volume analysis, we obtained 30-minute data from Dukascopy. The downloaded index is called USA500, and according to the methodology provided on the website, it includes top 500 stocks on the US stock market. We could not use high-frequency data of the S&P 500 as the index is not directly traded on the stock market and providers who recalculate the data typically charge high fees for their datasets. However, the index values are almost identical to the S&P 500 as they should include the same companies. The correlation of daily close prices between the USA500 and S&P 500 is 0.999992.

The observed period for the analysis includes data from January 1, 2019, to February 26, 2021. The downloaded dataset was continuous and included pre-market[4] trading hours and non-trading hours during weekends and holidays. To clean the dataset of such observations, we first excluded all observations outside the regular trading hours of the major US stock exchanges between 9:30 a.m. and 4:00 p.m. EST. After dropping weekend and holiday observations, we still had to exclude shortened trading days as those observations would influence

[4]Pre-market trading is trading that occurs between 4 a.m. and 9:30 a.m. EST

our results due to the trading days not lasting the entire 6.5 hours. A list of all excluded holidays and shortened trading days can be found in Table A.2.

For this analysis, we divided the observed timeline into three periods similarly as in the previous section:

1. **Before:** Starting from the earliest observation of the dataset on January 2, 2019, until January 31, 2020, when the Trump administration declared a public health emergency.

2. **During:** The main pandemic crash period after the declared emergency with first observation on February 3, 2020, until March 20, 2020, when massive QE with positive effect got announced the weekend after the last observation.

3. **After:** Period largely affected by endless QE from March 23, 2020 to February 26, 2021.

Table 3.1: Summary statistics of volume traded of the USA 500 index, in millions

	Before		During		After	
half hour	*mean*	*STD*	*mean*	*STD*	*mean*	*STD*
1	5.04	(3.01)	12.80	(10.81)	12.30	(6.75)
2	3.77	(2.48)	12.87	(13.05)	9.87	(6.23)
3	3.03	(1.99)	11.06	(11.25)	8.90	(6.38)
4	2.68	(2.09)	10.65	(11.04)	7.98	(6.12)
5	2.35	(1.79)	10.20	(11.41)	6.84	(5.37)
6	1.89	(1.50)	10.65	(10.98)	6.17	(5.24)
7	1.64	(1.41)	9.96	(11.35)	5.65	(5.11)
8	1.52	(1.42)	10.09	(11.83)	5.37	(5.12)
9	1.52	(1.60)	10.43	(11.90)	5.23	(4.88)
10	1.83	(1.87)	10.53	(11.46)	6.27	(5.15)
11	1.83	(2.09)	10.91	(12.28)	6.37	(5.41)
12	1.94	(1.98)	10.62	(12.27)	7.47	(5.79)
13	2.84	(2.06)	12.61	(14.26)	9.59	(6.34)

Chapter 4

Methodology

This section's main objective is to provide sufficient theoretical background into topics relevant for our analysis. Namely, we will introduce conditional heteroscedastic models that will be used in the empirical research and the methodology about how to find the best fit of the model. We will also cover the Analysis of variance (ANOVA) method, which we use to test the statistical significance of changes in the intraday volume patterns during the pandemic. For this chapter's purpose, two well-acknowledged books by Tsay (2005) and Bartoszyński & Niewiadomska-Bugaj (2008) were used to provide the necessary theory behind the subjects.

4.1 Financial time series analysis

4.1.1 Stationarity

When performing a time series analysis, it is important to have stationary data as many models and statistical tools rely on its implications of time-invariant properties. For this analysis, it is sufficient to understand only the weak stationarity as the strict one is hard to verify empirically and not needed for the models used. A time series $\{r_t\}$ is weakly stationary when $E(r_t) = \mu$ is constant for all observations and $Cov(r_t, r_{t-\ell}) = \gamma_\ell$ only depends on ℓ, which is an arbitrary integer.

4.1.2 Return series

The stationarity assumption has an important implication for financial time series analysis, as we are working with data where the weak stationarity rarely holds true. From Figure 3.1 we can see that expected value is very likely to be time-varying for our time series. To overcome this fact we will use the return series, and more specifically the log return series $\{r_t\}$ as its statistical properties are more tractable:

$$r_t = ln\frac{P_t}{P_{t-1}} = ln(P_t) - ln(P_{t-1}) \tag{4.1}$$

Where P_t stands for the price of the stock at time t, in our case the index points value.

4.1.3 ADF test

For testing whether the return series $\{r_t\}$ is stationary, we will be using the Augmented version of the Dickey-Fuller test (Dickey & Fuller 1979). Suppose we have the following model:

$$r_t = \alpha_0 + \alpha r_{t-1} + \phi_1 \Delta r_{t-1} + \ldots + \phi_L \Delta r_{t-L} + \epsilon_t$$

Where α_0 is a constant, ϵ_t is an error term, $\Delta r_{t-\ell}$ is the first difference at lag ℓ and α, ϕ_ℓ are estimated coefficients.

In the ADF test, we are testing whether there is a unit root process present. Thus the null hypothesis assumes that $\alpha = 1$. Ultimately we want to reject the null hypothesis, accepting the alternative one meaning the return series is stationary.

4.1.4 KPSS test

Another way to test for stationarity in our data is the KPSS test, the name of the test comes from the initials of Kwiatkowski, Phillips, Schmidt, & Shin (1992). Suppose the following model for our return series $\{r_t\}$:

$$r_t = \beta_t + d_t + \epsilon_t, \quad d_t = d_{t-1} + u_t$$

Where we have deterministic trend β_1, random walk d_t, stationary error ϵ_t and u_t are normal independent and identically distributed (iid) variables with a zero mean and variance σ_u^2. The initial value of d_0 is fixed intercept.

Unlike the ADF test, here we have the null hypothesis $H_0 : \sigma_u^2 = 0$ stating that the series is stationary against the alternative one $H_1 : \sigma_u^2 \neq 0$, which assumes unit root presence. In the analysis, we will be using both the ADF and the KPSS test, which will derive more powerful results.

4.1.5 Auto-correlation function

When working with a weakly stationary return series $\{r_t\}$ the Autocorrelation Function (ACF) is a great first indicator of the series's behaviour and will help us find the right order of moving-average part of the Autoregressive Moving-Average (ARMA) model which we will introduce later in this section. The correlation coefficient between r_t and $r_{t-\ell}$ under a weak stationarity assumption is a function of ℓ only and we call it the lag-ℓ autocorrelation of r_t denoted as ρ_ℓ:

$$\rho_\ell = \frac{Cov(r_t, r_{t-\ell})}{\sqrt{Var(r_t)Var(r_{t-\ell})}} = \frac{Cov(r_t, r_{t-\ell})}{Var(r_t)} = \frac{\gamma_\ell}{\gamma_0}$$

Where we used the property of weakly stationary time series $\{r_t\}$ that $Var(r_t) = Var(r_{t-\ell})$. If $\rho_\ell = 0$ for all $\ell > 0$, we can say that a weakly stationary time series is not serially correlated.

4.1.6 Partial auto-correlation function

The Partial Autocorrelation Function (PACF) is a great tool for determining the right order of autoregressive part of the ARMA model. Unlike the ACF, the PACF estimates the coefficients while taking into account lower lag orders of the PACF. An effective way to describe this is to consider the following Autoregressive models:

$$\begin{aligned}
r_t &= \phi_{0,1} + \phi_{1,1} r_{t-1} + e_{1t} \\
r_t &= \phi_{0,2} + \phi_{1,2} r_{t-1} + \phi_{2,2} r_{t-2} + e_{2t} \\
r_t &= \phi_{0,3} + \phi_{1,3} r_{t-1} + \phi_{2,3} r_{t-2} + \phi_{3,3} r_{t-3} + e_{3t} \\
&\vdots \quad \vdots
\end{aligned}$$

Where $\phi_{0,j}$ is a constant term, $\phi_{i,j}$ is the coefficient of r_{t-i} and $\{e_{jt}\}$ is the error term of an AR(j) model. These models can be estimated using the least squares method, the estimates of $\phi_{\ell,\ell}$ are then the $lag - \ell$ sample PACF of r_t.

4.1.7 Portmanteau statistic and Ljung-Box test

In financial econometrics it is important and often required to test jointly, whether several autocorrelations of r_t are equal to 0 to show if there is a dependence in the return series or not. Box & Pierce (1970) proposed the Portmanteau test statistic denoted as $Q^*(m)$ for the null hypothesis $H_0 : \rho_1 = \ldots = \rho_m = 0$, in words if the series is independent, against $H_1 : \rho_i \neq 0$ for $i \in \{1, \ldots, m\}$ suggesting dependant series. Under the assumption that $\{r_t\}$ is an iid sequence, $Q^*(m)$ is assymptotically a chi-squared random variable with m degrees of freedom.

$$Q^*(m) = T \sum_{\ell=1}^{m} \hat{\rho}_\ell^2$$

T is the number of observations in the sample and $\hat{\rho}_\ell$ is a biased estimator of ρ_ℓ, where the bias is in the order of $1/T$, which in our case is minimal, as we have over 800 observations for our return series.

Ljung & Box (1978) modified the Portmanteau statistic to increase the power of of the test in finite samples.

$$Q(m) = T(T+2) \sum_{\ell=1}^{m} \frac{\hat{\rho}_\ell^2}{T-\ell}$$

In our empirical research we will be using this modified version to jointly test for several autocorrelations. The selection of m can affect the performance of the $Q(m)$ statistic heavily, we will be sticking to the suggested (Tsay 2005) $m \approx ln(T)$ which in our case is $ln(805) \approx 7$.

4.1.8 ARMA model

After finding some dependancy in our dataset, we can employ the ARMA model to predict $\{r_t\}$. ARMA is a combination of Autoregressive (AR) model and Moving-Average (MA) model, which together create a more complex model, so that the number of parameters are kept small. A general ARMA(p,q) is defined in the form:

$$r_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \theta_i a_{t-i} + a_t \tag{4.2}$$

Where $\phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i}$ is the auto-regressive part of the model, and $\sum_{i=1}^{q} \theta_i a_{t-i}$ is the moving-average part. $\{a_t\}$ are white noise residuals, ϕ_0 is constant, ϕ_i and β_i are the estimated coefficients of the AR and MA processes, respectively.

However, the ARMA model is rarely used for forecasting future values in financial time series, as it is relying on knowing the present (t) value of $\{a_t\}$. Under the assumption that $\{a_t\}$ is a white noise series, meaning it is a sequence of iid random variables with finite mean and variance, the ARMA model is perhaps the most reliable estimate. However, in the case of financial time series, $\{a_t\}$ is rarely a white noise series, as the residuals hardly ever follow a normal distribution. What we observe in financial time series, and especially in stock prices, is so-called volatility clustering, firstly noted by Mandelbrot (1963). Volatility clustering suggests that in times of high price variation, the price variation is expected to stay high, and vice versa.

4.1.9 Conditional heteroscedastic models

The problem with volatility modelling is that it is not directly observable. In our case of daily log returns of the S&P 500 indices, we only have one observation in a trading day. We would need high-frequency data for more consistent estimates of volatility, which is often harder to obtain from publicly available sources, especially for stock indices. However, in our analysis, the daily returns should be sufficient if we apply the correct conditional heteroscedastic models.

When studying volatility, it is important to assume that the series $\{r_t\}$ is either serially uncorrelated or with minor lower order serial correlations, but it is dependant. We will also consider the conditional mean and conditional variance of $\{r_t\}$ given F_{t-1}:

$$\mu_t = E(r_t|F_{t-1}), \quad \sigma_t^2 = Var(r_t|F_{t-1}) = E[(r_t - \mu_t)^2|F_{t-1}] \tag{4.3}$$

Where F_{t-1} denotes the information set available at time $t-1$. In conditional heteroscedastic models, we still assume ARMA(p,q) model for the mean μ_t, however we take away the series $\{a_t\}$, which will be modeled seperately by the conditional heteroscedasticity models. We can see this in the following equations.

$$r_t = \mu_t + a_t, \quad \mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \theta_i a_{t-i} \tag{4.4}$$

When combining equations (4.3) and (4.4) we derive

$$\sigma^2 = Var(r_t|F_{t-1}) = Var(a_t|F_{t-1}).$$

The models which we will introduce next are concerned with the evolution of σ_t^2 over time.

4.1.10 ARCH

Autoregressive Conditional Heteroscedasticity (ARCH) model, developed by Engle (1982), is the first model that allows for systematic modelling of volatility in financial time series. General ARCH(m) model assumes that

$$a_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \ldots + \alpha_m a_{t-m}^2, \tag{4.5}$$

where $\{\epsilon_t\}$ is a sequence of iid random variables with zero mean and variance equal to 1, $\alpha_0 > 0$, and $\alpha_i \geq 0$ for $i > 0$. The coefficients α_i must satisfy some regularity conditions in order to obtain finite unconditional variance of a_t.

From the model specification, we can see that large past squared shocks $\{a_{t-i}^2\}$ imply a large conditional variance σ_t^2 for a_t, which in practice suggests that large shocks are expected to be followed by another large shock. Thus ARCH model is a great tool for modelling volatility of stock price returns, where we can observe volatility clustering.

One of the weaknesses of ARCH model is that it is very mechanical way to describe the behavior of the conditional variance. It does not provide us with any insights for understanding the source of variations of a financial time series. The model also assumes that positive and negative shocks have the same effect on volatility. In practice, financial time series often respond differently to positive and negative shocks.

4.1.11 GARCH

The simplicity of ARCH model can sometimes lead to high amount of parameters needed for quality description of the volatility. Bollerslev (1986) proposes an extension of the model known as Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model. A mean corrected log return series a_t follows a GARCH(m,s) model if

$$a_t = \sigma_t \epsilon_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2, \tag{4.6}$$

where $\{\epsilon_t\}$ is a sequence of iid random variables with zero mean and variance equal to 1, $\alpha_0 > 0$, and $\alpha_i \geq 0$, $\beta_j \geq 0$ and $\sum_{i=1}^{max(m,s)}(\alpha_i + \beta_i) < 1$, which implies that the unconditional variance of a_t is finite, whereas the conditional variance of σ_t^2 is evolving over time. We can see that equation (4.6) reduces down to (4.5) when parameter $s = 0$. Also if $m = s = 0$, $\{a_t\}$ is only white noise and we cannot observe any ARCH effect.

The difference between ARCH and GARCH is that the latter allows for lagged conditional variances σ^2_{t-j} to enter the model as visible from the equation (4.6). Consequently, this leads to more simple models, where we often find that GARCH(1,1) is the best fit. We will explain more about finding the best model fit in Subsection 4.1.13.

4.1.12 T-GARCH

We often find asymmetries in financial time series data, and as we already mentioned, markets tend to react differently to positive and negative shocks. The standard ARCH and GARCH models have no tool to deal with that. Zakoian (1994) proposed an implementation that can deal with asymmetry, the threshold GARCH model (TGARCH). TGARCH(m,s) model that we will be using in our empirical research is defined as follows:

$$a_t = \sigma_t \epsilon_t, \quad \sigma_t = \alpha_0 + \sum_{i=1}^{m} \alpha_i \sigma_{t-i}(|a_{t-i}| - \eta_i a_{t-i}) + \sum_{j=1}^{s} \beta_j \sigma_{t-j}, \tag{4.7}$$

where η_i is the coefficient in the leverage term and $|\eta_i| \leq 1$. We should also note that in the case of TGARCH, we are modelling σ_t instead of σ_t^2. This is a key difference to the earlier proposed threshold model known as GJR-GARCH (Glosten *et al.* 1993). Modelling conditional standard deviation instead of conditional variance allows us to have no positivity constraints. This is especially useful when adding dummy variables to the model, which will be a key part of the empirical analysis, as the coefficients can turn out negative, telling us additional information.

4.1.13 Model fitting based in information criteria

When working with conditional heteroscedasticity models, it is crucial to specify the model correctly, both for conditional mean and variance. In our empirical research we will be minimising the information criteria to determine the best number of parameters in our model. All information criteria are likelihood based. We will be using the Akaike Information Criterion (Akaike 1973) defined as

$$AIC = -2ln(\hat{L}) + 2k \tag{4.8}$$

Where $\hat{L}$ is the likelihood function evaluated at the maximum likelihood estimates and k is the number of estimated parameters in the model.

4.2 Intraday volume analysis

4.2.1 Daily volume traded in percentage

In the analysis of the daily volume patterns on the US stock market during the COVID-19 pandemic crash, we are interested in the pattern changes in percentage volumes rather than absolute volumes. While it is pretty apparent that the volume of trades increased during the big sellout on the market, it is not so clear whether the proportion of trades in a given day changed significantly. Thus we will use the following formula to transform our absolute volume data into proportionate volume data.

$$P_{id} = \frac{V_{id}}{\sum_{j=1}^{13} V_{jd}} \tag{4.9}$$

Where V_{id} is the volume traded on day d in half-hour $i \in \{1, \ldots, 13\}$. From this formula, we can see that $\sum_{i=1}^{13} P_{id}$ must equal 1.

4.2.2 ANOVA

ANOVA is a great tool for testing whether means across multiple groups differ significantly. The main objective is to test

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_k \quad \text{against} \quad H_1 : \mu_i \neq \mu_{i'} \text{ for some } i, i'$$

Suppose we have k groups each with n_i observations where $i \in \{1, \ldots, k\}$. Then first we calculate the sum of squares for between group variability:

$$SS_{between} = n_1(\bar{X_1} - \bar{X_G})^2 + \ldots + n_k(\bar{X_k} - \bar{X_G})^2$$

Where $\bar{X_i}$ is the mean for group i and $\bar{X_G}$ is the total mean across all groups. Then we can derive the mean square for between group variability:

$$MS_{between} = \frac{SS_{between}}{k-1}$$

Next part is to calculate the sum of squares for within group variability, from which we can obtain the mean square as well.

$$SS_{within} = \sum_{j=1}^{n_1}(x_{j1} - \bar{X_1})^2 + \ldots + \sum_{j=1}^{n_k}(x_{jk} - \bar{X_k})^2$$

$$MS_{within} = \frac{SS_{within}}{N - k}$$

Where $\{x_{ji}\}$ are observations of group $i \in \{1, \ldots, k\}$, $j \in \{1, \ldots, n_i\}$, and N stands for the total number of observations across all groups.

Lastly we calculate the F-test statistic. We reject the hypothesis $H_0 : \mu_1 = \mu_2 = \ldots = \mu_k$ at level of significance α if:

$$F = \frac{MS_{between}}{MS_{within}} > F_{\alpha, k-1, N-k}$$

It is important to note that in ANOVA, we always use a one-sided critical region since any violation of H_0 tends to increase the numerator of the F statistic without affecting the denominator.

Chapter 5

Empirical research and results

5.1 Descriptive statistics and return series

5.1.1 Log-return series and stationarity

We obtained the log-return data by following the Equation 4.1. In Figure 5.1 are plotted the log returns for each of the sectors. At first glance, we can see that the Consumer Staples and Health Care sectors have one of the lowest (but still pretty big) spikes during the pandemic market crash compared to other sectors. It implicates that those sectors were less affected by the pandemic. However, we will get much more precise results from the conditional heteroscedasticity models later in this chapter.

As mentioned in the methodology, the stationarity assumption is critical for working with time series data. We employed the ADF and the KPSS tests mentioned in the methodology to test for stationarity. The results in Table A.1 show that the log-return series is stationary for all sectors.

5.1.2 Descriptive statistics

From the descriptive statistics in Table 5.1 we can immediately observe that the variance is relatively high compared to the mean, implicating volatile series, and especially for the Energy sector, which is the only sector that decreased over the observed period as we can see from the negative mean. The high volatility and decreasing trend for the Energy sector might be explained due to the increasing popularity of green fuels, which can repel investors from believing in the future of consumable fuels which dominate the Energy sector.

The negative skewness across all sectors implies the returns tend to generate

Figure 5.1: Daily log returns from 2018 to March 15, 2021

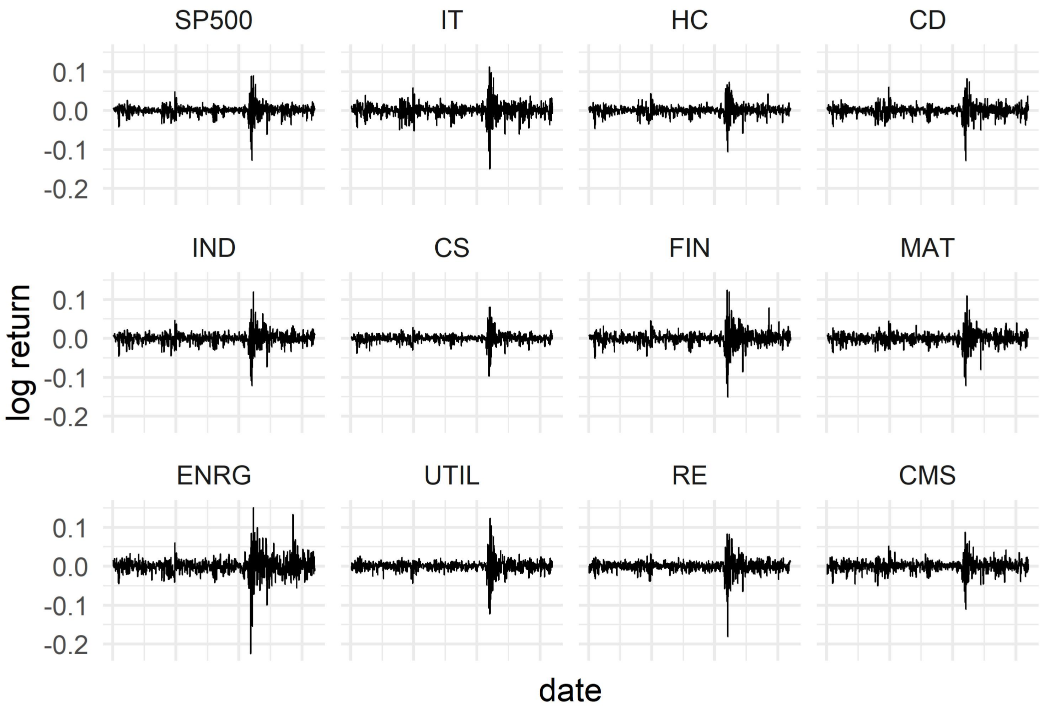


Table 5.1: Descriptive statistics of the S&P 500 log returns

	N	Mean	Variance	Skewness	Kurtosis	Min	Max
SP500	805	0.00049	0.00021	-1.02	19.02	-0.128	0.090
IT	805	0.00093	0.00033	-0.66	13.83	-0.150	0.113
HC	805	0.00042	0.00018	-0.48	13.19	-0.105	0.073
CD	805	0.00067	0.00023	-1.08	14.71	-0.129	0.083
IND	805	0.00032	0.00028	-0.70	14.88	-0.122	0.120
CS	805	0.00018	0.00014	-0.38	18.02	-0.097	0.081
FIN	805	0.00026	0.00035	-0.69	17.24	-0.151	0.124
MAT	805	0.00033	0.00027	-0.69	13.48	-0.121	0.110
ENRG	805	-0.00037	0.00060	-1.03	18.88	-0.224	0.151
UTIL	805	0.00022	0.00024	-0.20	20.72	-0.123	0.123
RE	805	0.00023	0.00027	-1.74	26.83	-0.181	0.083
CMS	805	0.00046	0.00023	-0.69	11.10	-0.110	0.088

primarily small positive returns. In contrast, the negative returns tend to be more extreme as the distribution's left-tail is longer. This is also visible from the minimum values being greater than the maximum values in absolute terms. High positive kurtosis is an indicator of heavy tails and more outliers in our series than we would find in a normal distribution.

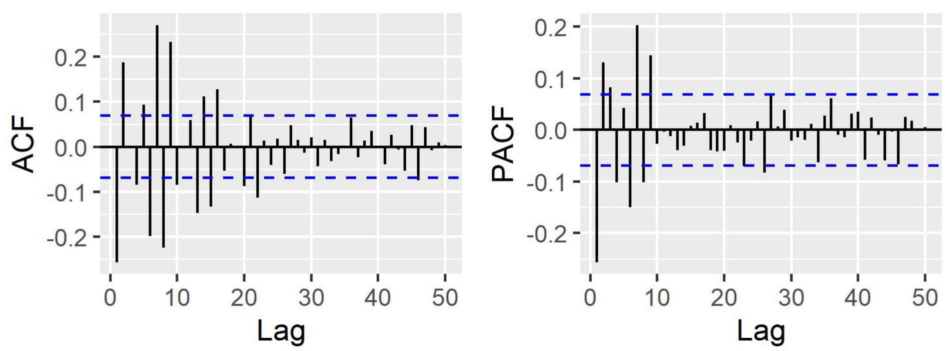
5.2 S&P 500 return series analysis

This section will be going through the steps mentioned in the methodology for the S&P 500 index. The same steps will be applied to the individual sectors as well. However, we will not cover them separately and instead provide only the final results later in this chapter. Additional results for the individual sectors can be found in Appendix.

5.2.1 ARMA model

From the ACF in Figure 5.2 it is clear that there is a heavy autocorrelation structure in the returns. This should not be the case under the EMH, however, our data is heavily affected by the pandemic period, so we may observe some inefficiencies in terms of linear dependency. When computing the Ljung-Box test for seven lags, we can easily reject the null hypothesis as the p-value is lower than $2.2 * 10^{-16}$. Thus we accept the alternative hypothesis that there is dependence in the data. The PACF also shows a significant result for the first lag. From both the ACF and PACF, we can expect both the autoregressive and moving-average parts of the ARMA(p,q) model to be present, meaning that neither p nor q will equal 0.

Figure 5.2: ACF and PACF of S&P 500 daily log returns



Before moving on to minimising the information criteria and finding the correct ARMA(p,q), we need to introduce the dummy variables, that will be included in the model. The dummy variables can result in different estimates of the ARMA(p,q) and it is crucial to include them beforehand. The ARMA(p,q) model for the return series $\{r_t\}$ is defined as:

$$\begin{gathered} r_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \theta_i a_{t-i} + a_t + \gamma_1 I_{DUR*Mon} + \gamma_2 I_{AFTR*Mon} + \\ + \gamma_3 I_{DUR*Fri} \gamma_4 I_{AFTR*Fri} + \gamma_5 D_{Mon} + \gamma_6 D_{Fri} + \gamma_7 D_{DUR} + \gamma_8 D_{AFTR} \end{gathered} \tag{5.1}$$

Where I are interaction dummy variables, which equal 1 when both dummies are satisfied, i.e. $I_{DUR*Mon}$ equals 1 on Mondays in period *During* outlined in Section 3.1. D dummy variables equal 1 when only the one condition in the index is satisfied. The rest of the equation is the same as in Equation 4.6.

Dummy variables for other days of the week than Monday and Friday are not included as this analysis's main interest is to observe the behaviour around weekends. From the existing literature, it is also apparent that Mondays and Fridays have had the most significant effects on the stock market. Given the low amount of observations in the analysis's focused periods, it is better not to over-parametrise the model.

After minimising the information criteria, the best fit of the model is ARMA(3,1) with no intercept. However, the results shown in Table 5.2 do not follow the results from the autocorrelation function. We see that a sizeable moving-average coefficient cancels out the autoregressive coefficients, meaning that if there is an unexpected return of 1%, the AR coefficients will expect the next day to return to be about -1%. In contrast, the MA coefficient will negate the shock by adding back the unexpected return of 1%. Thus the results are hardly telling us any information. To obtain results easier to interpret, we will only optimise the AR(p) model with a maximum value of p up to 5 instead. The best model fit for the return series is AR(4) model with results shown in Table 5.2. We decided to include an intercept in the model to get rid of some potential linear trend in the series. The AR(4) model results are more in line with the autocorrelation function and will be easier to interpret. Also, the information criterion has changed marginally, meaning the likelihood of both models is very close. However, we will not be elaborating on the estimated coefficients of the ARMA models as the conditional heteroscedastic models will change those estimates once we account for the volatility modelling.

Table 5.2: ARMA models of the S&P 500

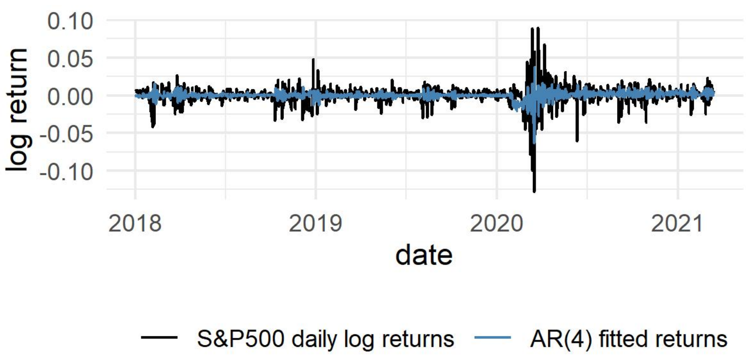
	ARMA(3,1)		ARMA(4,0)	
ϕ_0	-	-	0.0006	(0.001)
ϕ_1	-1.1269	(0.041)	-0.2614	(0.036)
ϕ_2	-0.0959	(0.053)	0.1233	(0.037)
ϕ_3	0.2087	(0.036)	0.0223	(0.037)
ϕ_4	-	-	-0.1315	(0.035)
θ_1	0.9160	(0.025)	-	-
γ_1	-0.0084	(0.006)	-0.0094	(0.006)
γ_2	0.0053	(0.003)	0.0057	(0.003)
γ_3	-0.0089	(0.006)	-0.0057	(0.006)
γ_4	-0.0014	(0.003)	-0.0017	(0.003)
γ_5	-0.0007	(0.001)	-0.0011	(0.002)
γ_6	0.0005	(0.001)	-0.0003	(0.001)
γ_7	-0.0071	(0.003)	-0.0076	(0.003)
γ_8	0.0016	(0.001)	0.0010	(0.001)
AIC	-4667		-4623	

Note: Standard errors of estimates are in parenthesis.

5.2.2 ARCH effect

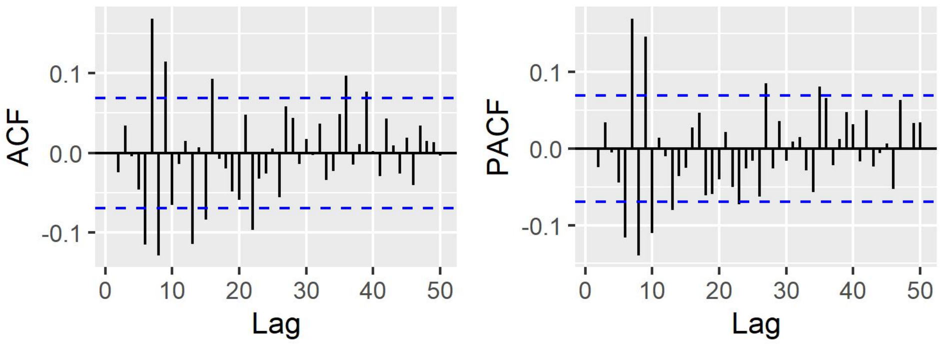
In methodology, we talked about the inefficiencies of using the ARMA model in financial time series. This subsection shows how that applies to our data before we move on to estimating the T-GARCH model. Figure 5.3 is an excellent visualisation of the fact that the AR model's fitted values are much less volatile than the actual values. This is because the autoregressive model does not estimate the disturbance $\{a_t\}$ and only assumes it is a white noise series.

Figure 5.3: Signal-noise ratio of the S&P 500 daily log returns



However, some autocorrelation is still left in the data, as we can see from the ACF and PACF of the residuals in Figure 5.4. The fact that there is still some dependency in the residuals is essential. It means that the AR(4) is still leaving some information unused, and better results will be obtainable when employing conditional heteroscedastic models. If the residuals were just a white noise series, we would not observe any ARCH effect, and the AR model would probably be sufficient.

Figure 5.4: ACF and PACF of the S&P 500 AR model fit residuals



To quickly test whether the data is suitable for applying ARCH models Engle (1982) proposed a way to measure for so-called ARCH effect in the data. The ARCH effect can be significantly approved by rejecting null hypothesis of homoscedastic residuals and accepting the alternative one of heteroscedastic residuals. One can measure the test by calculating the Portmanteau Q test statistic for squared residuals. In our case, we reject the null hypothesis as the Ljung-Box test at 7 lags returns p-value lower than $2.2 * 10^{-16}$. Concluding the presence of ARCH effect in the residuals, so we can move on to estimating the T-GARCH model.

5.2.3 T-GARCH model specification

We will be using the T-GARCH(m,s) model for estimating the volatility. We can see from the skewness in Table 5.1 there is a certain asymmetry in the data, so we think the Threshold GARCH model will be the best fit. It will allow us to find additional information regarding how significant the asymmetry is across sectors. For this analysis, the standard T-GARCH(1,1) will be sufficient as the main focus is on comparing the data across sectors. Economists also often use variations of the standard GARCH(1,1) model as it is regarded satisfactory with no need for higher orders of m or s when working with financial data (Brooks & Burke 2003).

We will also be adding two dummy variables for measuring the volatility. Thus for the return series $\{r_t\}$ we will be using the following set of equations to estimate the T-GARCH(1,1):

$$r_t = \mu_t + a_t \tag{5.2}$$

$$\begin{aligned}\mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + \gamma_1 I_{DUR*Mon} + \gamma_2 I_{AFTR*Mon} + \gamma_3 I_{DUR*Fri} + \\ + \gamma_4 I_{AFTR*Fri} + \gamma_5 D_{Mon} + \gamma_6 D_{Fri} + \gamma_7 D_{DUR} + \gamma_8 D_{AFTR}\end{aligned} \tag{5.3}$$

$$a_t = \sigma_t \epsilon_t \tag{5.4}$$

$$\sigma_t = \alpha_0 + \alpha_1 \sigma_{t-1}(|a_{t-1}| - \eta_1 a_{t-1}) + \beta_1 \sigma_{t-1} + \delta_1 D_{DUR} + \delta_2 D_{AFTR} \tag{5.5}$$

Where p equals 4 for the S&P 500 return series, but different orders of p up to 5 will be used across sectors based on the information criterion minimisation. We added two dummy variables for the standard deviation equation. These will provide insight into the effects of the pandemic crash and the following period of quantitative easing brought on the market.

5.2.4 S&P 500 T-GARCH results

In Table 5.3 are the estimated coefficients for the T-GARCH(1,1) model with AR(4) as the included mean model. From the mean model results, the autoregressive coefficient ϕ_1 is significant and slightly negative with an estimate of -0.071. This result tells us that the market tends to contradict previous observation very slightly. If, for example, the market growths by 1% within a day, then the market is expected to return -0.07% the next day based on the ϕ_1 estimate. However, we still have to account for all of the other variables in the model.

Looking at the added dummy variables, we only obtained two significant results. The first one for γ_2 shows that the returns on Mondays in the period from March 23, 2020, onwards have been significantly higher. This is very interesting, and it shows that the calendar time hypothesis (French 1980) holds for this bull period on the US stock market. On the other hand, the Monday returns during the crash period (γ_1) show a much stronger negative relationship than the Friday returns in the same period (γ_3). Unfortunately, both of these results are statistically insignificant due to the low number of observations.

Table 5.3: S&P 500 T-GARCH estimates

Mean model results			STD model results		
ϕ_0	7.4E-4**	(3.5E-4)	α_0	0.0006***	(0.000)
ϕ_1	-0.071**	(0.034)	α_1	0.153***	(0.033)
ϕ_2	0.024	(0.034)	β_1	0.811***	(0.031)
ϕ_3	0.018	(0.040)	η_1	0.875***	(0.198)
ϕ_4	-0.009	(0.034)	δ_1	0.0007	(0.0006)
γ_1	-0.0068	(0.0053)	δ_2	0.0006***	(0.0001)
γ_2	0.0031***	(0.0009)			
γ_3	-0.0041	(0.0050)			
γ_4	-0.0004	(0.0019)			
γ_5	-0.0007	(0.0005)			
γ_6	0.0004	(0.0007)			
γ_7	0.0020***	(0.0007)			
γ_8	-0.0006	(0.0006)			

Note: *, **, *** stand for 10%, 5% and 1% level of significance respectively. Standard errors of estimates are in parenthesis

However, we can still be reasonably confident that the Monday returns were lower when looking at the standard deviations. It confirms the hypothesis that investors did not prepare for the inflow of bad news during the weekends beforehand, resulting in lower returns on Mondays.

The second significant dummy γ_7 shows a positive relationship for the period during the crash. This is very interesting, as we would probably expect a negative relationship instead. This can be because many negative returns possibly happened on Mondays and Fridays during the crash, which coefficients still have much stronger negative relationship estimates. Alternatively, the negative returns are reflected in the mean and volatility model so well that we do not see a change of pattern, which would result in a significantly negative estimate of γ_7. Something similar is the case for the period after the crash as the estimate of γ_8 is negative and insignificant, though we would expect it to be positive and significant.

We also observe a positive significant intercept $\phi_0 = 0.00074$, showing that there is a positive trend in the series. It comes as no surprise, as the S&P 500 has gained in value over the observed period.

From the volatility model, we can see a very significant intercept α_0, which is not telling us much information on its own. However, when comparing it with the dummy variables δ_1 and δ_2, it is rather interesting as we can see that the volatility doubled in the period after the crash, which also has a significant coefficient. For δ_1, we can see something similar as in the mean model. While we know that the volatility during the crash increased dramatically, the observed period is too short, and the information is again mainly included in the T-GARCH(1,1) model coefficients. As we learned, GARCH models do a great job of modelling volatility clustering, so the estimate of δ_1 does not necessarily have to be higher than δ_2.

The estimate of η_1 is suggesting a significant presence of asymmetry in the return series. It confirms that bad shocks have a much stronger effect on the stock market than positive ones. The sum of α_1 and β_1 is about 0.96, this number is often regarded as the persistence of the model and is generally close to 1, implying that the volatility is persistent and is likely to be estimated very closely to the previous observation. This is mainly driven by a much higher estimate of β_1 (0.81), which is more tied to the persistence of volatility as it estimates based on previous observation directly. The α_1 (0.15) estimate tells us more about the impact of new information in the form of disturbance a_t, and in our case, it also includes the information about asymmetry.

5.3 Sectoral T-GARCH results

In this section, we applied the same procedure as in Section 5.2 to all eleven sectors of the S&P 500 index, and we will evaluate the results of both the included mean AR(p) model and the volatility T-GARCH(1,1) model itself. In Appendix, the reader can find ACF for the sectors' return series.

5.3.1 Mean model results

Table 5.4 shows the results for the included AR(p) mean model in the T-GARCH model. The order of p was chosen by minimising the information criteria for each of the sector. The majority of the sectors have four or even five autoregressive coefficients included in the model. However, for the IT, CMS and ENRG sectors, we only have a simple AR(1) model. In the case of IT and CMS sectors, we see slightly negative significant estimates similarly as for the whole S&P 500. Energy sector shows an insignificant estimate of 0.008 for ϕ_1. This does not provide any information and shows that the sector is tough to predict. Similarly, HC, IND, CD and MAT sectors do not have any significant estimates of the autoregressive coefficients.

Table 5.4: T-GARCH mean model results

	ϕ_0	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7	γ_8
SP500	7.4E-4**	-0.071**	0.024	0.018	-0.009	-	-0.0068	0.0031***	-0.0041	-0.0004	-0.0007	0.0004	0.0020***	-0.0006
	(3.5E-4)	(0.034)	(0.034)	(0.040)	(0.034)	-	(0.0053)	(0.0009)	(0.0050)	(0.0019)	(0.0005)	(0.0007)	(0.0007)	(0.0006)
IT	1.3E-3***	-0.108***	-	-	-	-	-0.0043	0.0059***	-0.0089	-0.0004	-0.0010	0.0001	0.0022	-0.0022***
	(4.3E-4)	(0.037)	-	-	-	-	(0.0101)	(0.0012)	(0.0103)	(0.0027)	(0.0007)	(0.0011)	(0.0050)	(0.0008)
HC	4.1E-4	-0.007	0.047	-0.016	-0.038	-	-0.0049	0.0023***	-0.0015	-0.0002	-0.0012**	0.0012	0.0004	-0.0004
	(3.7E-4)	(0.027)	(0.037)	(0.032)	(0.026)	-	(0.0084)	(0.0008)	(0.0086)	(0.0018)	(0.0005)	(0.0007)	(0.0038)	(0.0007)
CD	9.0E-4	-0.042	0.001	-	-	-	-0.0026	0.0040	-0.0064***	-0.0020	-0.0006	-0.0001	0.0025**	-0.0001
	(8.0E-4)	(0.035)	(0.034)	-	-	-	(0.0065)	(0.0025)	(0.0015)	(0.0018)	(0.0011)	(0.0013)	(0.0012)	(0.0013)
IND	-3.2E-4	0.022	0.007	-0.006	-0.023	-	-0.0086	0.0018	-0.0036	0.0023	-0.0002	0.0019**	0.0015	0.0002
	(3.5E-4)	(0.036)	(0.035)	(0.037)	(0.036)	-	(0.0090)	(0.0026)	(0.0098)	(0.0024)	(0.0010)	(0.0008)	(0.0052)	(0.0011)
CS	1.8E-4	-0.079**	-0.011	0.023	-0.066*	-	-0.0022	0.0020	0.0005	-0.0001	-0.0009	0.0012	0.0003	-0.0002
	(3.7E-4)	(0.038)	(0.037)	(0.036)	(0.037)	-	(0.0060)	(0.0016)	(0.0017)	(0.0017)	(0.0008)	(0.0008)	(0.0014)	(0.0007)
FIN	-7.0E-4	0.004	0.073**	0.014	-0.010	-	-0.0096	0.0005	-0.0035	-0.0013	0.0009	0.0017***	0.0006	0.0009
	(4.4E-4)	(0.032)	(0.029)	(0.034)	(0.032)	-	(0.0070)	(0.0027)	(0.0060)	(0.0026)	(0.0009)	(0.0006)	(0.0011)	(0.0014)
MAT	-6.5E-5	-0.009	0.028	0.001	-0.040	-	-0.0019	0.0038	-0.0079	0.0020	-0.0011	0.0005	0.0007	0.0003
	(5.2E-4)	(0.037)	(0.038)	(0.037)	(0.037)	-	(0.0103)	(0.0028)	(0.0113)	(0.0034)	(0.0010)	(0.0012)	(0.0056)	(0.0015)
ENRG	-1.1E-3*	-0.008	-	-	-	-	-0.0300**	0.0037	-0.0058	0.0005	0.0027***	0.0004	0.0000	0.0011
	(6.0E-4)	(0.037)	-	-	-	-	(0.0152)	(0.0048)	(0.0171)	(0.0048)	(0.0010)	(0.0013)	(0.0084)	(0.0024)
UTIL	1.0E-3**	-0.013	-0.101***	0.015	-0.065*	0.084**	0.0079	0.0049*	0.0010	0.0007	-0.0030***	0.0003	-0.0045	-0.0014
	(4.4E-4)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.0089)	(0.0026)	(0.0121)	(0.0025)	(0.0010)	(0.0009)	(0.0049)	(0.0011)
RE	6.7E-4	-0.035	0.001	0.016	-0.080**	0.035	0.0000	0.0030	0.0009	0.0023	-0.0018*	-0.0002	0.0006	-0.0009
	(4.5E-4)	(0.037)	(0.037)	(0.036)	(0.036)	(0.037)	(0.0092)	(0.0024)	(0.0108)	(0.0024)	(0.0010)	(0.0009)	(0.0048)	(0.0011)
CMS	8.3E-5	-0.092**	-	-	-	-	-0.0023	0.0017	0.0002	-0.0016*	-0.0005	0.0008	0.0000	0.0008
	(4.3E-4)	(0.036)	-	-	-	-	(0.0092)	(0.0024)	(0.0056)	(0.0010)	(0.0011)	(0.0008)	(0.0044)	(0.0007)

Notes: *, **, *** stand for 10%, 5% and 1% level of significance respectively. Standard errors of estimates are in parenthesis.
γ_i dummies stand for: 1 - DUR*Mon, 2 - AFTR*Mon, 3 - DUR*Fri, 4 - AFTR*Fri, 5 - Mon, 6 - Fri, 7 - DUR, 8 - AFTR

The intercepts show us a significant positive linear trend for the IT and UTIL sectors, suggesting that they have enjoyed a growing trend since 2018. IT sector has been a hot topic dominating the US stock market in the past year, so this result comes as no suprise. The Energy sector, on the other hand, has a significantly negative intercept.

The included dummy variables show very few significant results, which are very similar across sectors. For the Monday effect during the stock market crash, we see a strong negative relationship for most sectors. However, the estimate is significant only for the Energy sector, which shows an estimate of -0.03. The Friday effect during the pandemic also shows negative insignificant results for the majority of sectors, but the relationship is much weaker. The one sector where we see something completely different is the Consumer Discretionary sector, which has a significant estimate of -0.0064 for the Friday effect during the pandemic, as opposed to a insignificant estimate of -0.0026 for Mondays. It implies that the inflow of bad news during weekends might not have affected the CD sector as much as the other sectors.

The overall Monday and Friday effects for the whole timeline show supporting evidence to the existing literature as the Monday effect is mostly negative. In contrast, the Friday effect is slightly positive. Even though most of the estimates are insignificant, we can definitely still observe some inefficiency on the market, especially for the IND and FIN sectors, which have significant estimates of 0.0019 and 0.0017, respectively, for the Friday effect.

The Information Technology sector shows very high and significant Monday returns for the period after the crash. Since the whole S&P 500 index also shows big positive Monday returns for the given period, it is not surprising as the IT sector enjoyed a very successful year in 2020 on the US stock market. Surprisingly though, no other sector has a significant (at 5% level of significance) result for Monday returns in that period, but the results are still positive for all of them.

5.3.2 STD model results

The results in Table 5.5 were obtained by estimating the T-GARCH(1,1) model. Each sector has a specific included mean model based on the results in the previous subsection. Apart from the estimated coefficients, we can also find the Akaike information criteria in the last column, which tells us how good of a fit the model is. Equation 4.8 shows the standard calculation for AIC, however,

Table 5.5: T-GARCH volatility model results

	α_0	α_1	β_1	η_1	δ_1	δ_2	AIC
SP500	0.0006***	0.153***	0.811***	0.875***	0.0007	0.0006***	-6.41
	(0.000)	(0.033)	(0.031)	(0.198)	(0.0006)	(0.0001)	
IT	0.0009***	0.113***	0.836***	0.881***	0.0015*	0.0008***	-5.71
	(0.000)	(0.030)	(0.029)	(0.252)	(0.0009)	(0.0002)	
HC	0.0006***	0.093***	0.864***	1.000***	0.0013*	0.0002	-6.32
	(0.000)	(0.027)	(0.029)	(0.279)	(0.0006)	(0.0001)	
CD	0.0005***	0.106***	0.862***	0.737***	0.0012**	0.0005***	-6.09
	(0.000)	(0.023)	(0.024)	(0.197)	(0.0005)	(0.0001)	
IND	0.0005***	0.086**	0.888***	0.922**	0.0011*	0.0002	-5.93
	(0.000)	(0.036)	(0.042)	(0.372)	(0.0006)	(0.0001)	
CS	0.0005***	0.090***	0.862***	0.863***	0.0013***	0.0001	-6.66
	(0.000)	(0.029)	(0.033)	(0.305)	(0.0005)	(0.0001)	
FIN	0.0007***	0.127***	0.845***	1.000***	0.0007	0.0004**	-5.86
	(0.000)	(0.027)	(0.028)	(0.210)	(0.0006)	(0.0002)	
MAT	0.0006	0.074	0.886***	1.000	0.0017**	0.0003	-5.90
	(0.000)	(0.058)	(0.079)	(0.705)	(0.0008)	(0.0002)	
ENRG	0.0007**	0.053**	0.905***	0.940**	0.0029***	0.0008**	-5.25
	(0.000)	(0.027)	(0.038)	(0.410)	(0.0010)	(0.0004)	
UTIL	0.0003***	0.026*	0.941***	1.000	0.0022***	0.0001**	-6.31
	(0.000)	(0.013)	(0.014)	(0.713)	(0.0005)	(0.0001)	
RE	0.0003***	0.043*	0.934***	1.000	0.0018***	0.0001	-6.16
	(0.000)	(0.025)	(0.027)	(0.633)	(0.0005)	(0.0001)	
CMS	0.0007***	0.059**	0.891***	1.000**	0.0012**	0.0003**	-5.89
	(0.000)	(0.024)	(0.032)	(0.447)	(0.0005)	(0.0001)	

Notes: *, **, *** stand for 10%, 5% and 1% level of significance respectively. Standard errors of estimates are in parenthesis.
AIC is the Akaike information criterion. δ_i dummies stand for: 1 - DUR, 2 - AFTR

in this case, the AIC is calculated based on the following formula:

$$AIC = \frac{-2ln(\hat{L})}{N} + \frac{2k}{N}$$

Where $\hat{L}$ is the likelihood function evaluated at the maximum likelihood estimates, k is the number of estimated parameters in the model, and N is the total number of observations.

The estimated coefficients of the model are mostly significant. Even most of the included dummy variables show at least a 10% level of significance, which is an excellent sign implying that the estimated models are very well specified, and we can rely on the results. The AIC on its own is not telling much, as it is mostly used to evaluate when comparing competing models. For example, models like standard GARCH(1,1) or GJR-GARCH(1,1) return AIC of -6.36 and -6.39, respectively, for the S&P 500 index. Since we are minimising the AIC, we know the T-GARCH(1,1) with a value of -6.41 is marginally better. The purpose of the AIC in Table 5.5 is to illustrate that there is no sector for which we would see a struggling model with two times lower AIC than a different one.

From the α_0 estimates, we can see the intercept of the volatility equation. The Information Technology sector has the highest estimate of 0.0009, showing that the expected volatility has generally been higher than for other sectors. Interestingly, sectors with lower intercept α_0 have a much greater estimate of δ_1, reflecting the volatility rise during the pandemic crash period. It shows that even sectors with moderately lower unconditional volatility have seen a dramatic increase in volatility during the pandemic. The most significant effect is seen for the Energy sector. Its volatility increased dramatically, with an intercept of 0.0007 rising an additional 0.0029 during the pandemic and 0.0008 in the period after. The very wild ride of the Energy sector on the US stock market in 2020 was driven mainly by the unstable oil industry situation[1].

While we do not have a significant estimate of δ_1 for the S&P 500, we can see the individual sectors' impact with 10/11 significant estimates. It shows that investors can significantly decrease the risk during turbulent periods on the US stock market by broadening the portfolio. The Financials sector, which did not return a significant estimate δ_1 is the third largest sector in terms of the number of constituents. On the other hand, sectors like ENRG, MAT, UTIL or CMS, all with less than 30 constituents, have very high and significant increases in volatility.

The estimates of δ_1 generally show that all sectors were affected in terms of volatility during the crash period. However, the estimates of δ_2 tell us a completely different story as some sectors continued to have increased volatility and some did not. The IND, CS, UTIL, RE, MAT, and HC sectors show very little increase in volatility in the period after the crash, it strikes as especially interesting for the Health Care sector as the pandemic is continuing throughout the whole period, and the industry was a big talking point as investors were looking to invest in healthcare companies, which will help bring the world out of

[1]Saudi Arabia announced unexpected oil price discounts on March 8, 2020, which caused the oil prices to plummet dramatically. On April 20, some oil prices even went below zero.

the pandemic. On the other hand, IT, FIN, CD, CMS and ENRG sectors have seen a substantial increase in volatility even after the crash period. If we look at what are the individual sectors composed of, this significant rise in volatility makes sense. IT companies have seen an extreme increase in demand due to home-office or home-schooling contrasted by the shortage of semiconductors, which understandably supports the increased volatility on the market. Banks in the financial sector are providing more unsecured loans to support affected businesses. Hotels and restaurants in the CD sector were strongly restricted in the past year. People spend more time on Netflix, YouTube and other platforms, which reflect results in the CMS sector. These are all factors why we see persistent volatility increase among those sectors.

The coefficients α_1 and β_1 are very similar across sectors, showing that excess volatility modelling does not differ much across sectors after we account for intercepts. We see that IT and FIN sectors have higher estimates of α_1 at 0.11 and 0.13, respectively, showing that those sectors are more driven by past shocks than other sectors. On the other hand, MAT, UTIL and RE sectors have very low and even insignificant (at 5% level of significance) estimates of α_1 showing that these sectors are mainly driven by previous observations of volatility, included in the β_1 estimate, and not driven by past shocks. Consequently, the estimates of η_1 for these sectors are also insignificant as they are closely tied to α_1. In contrast, other sectors show high significant asymmetry with η_1 close to one, concluding that the market tends to react more strongly to negative shocks.

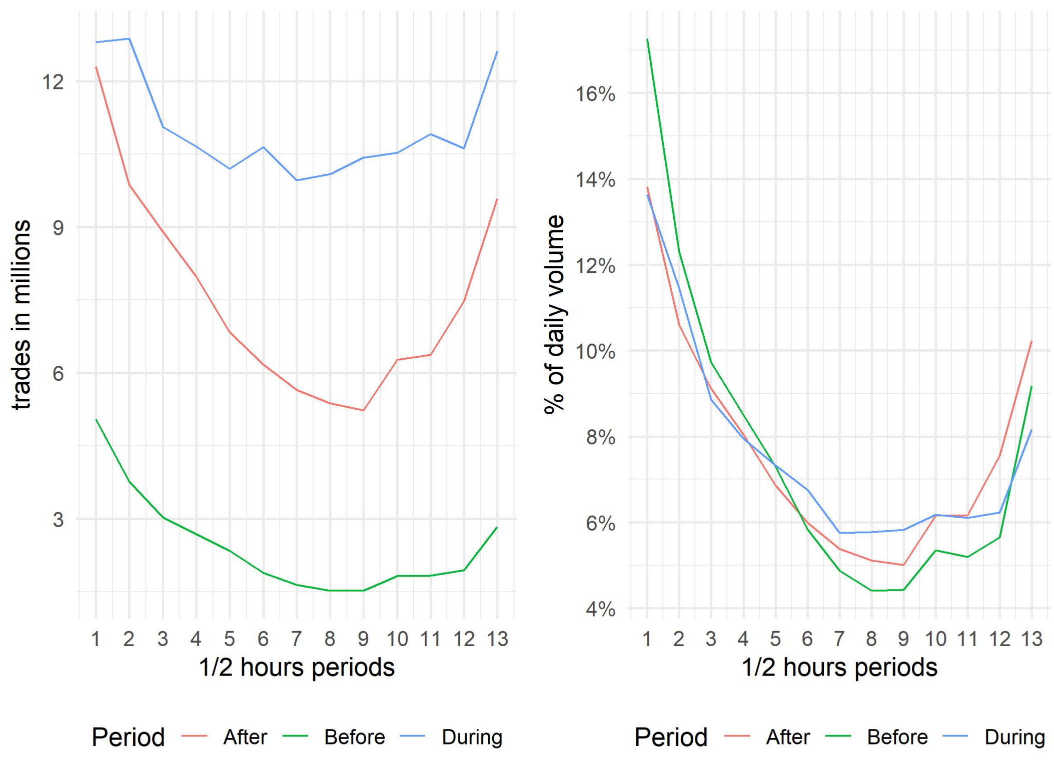
5.4 Intraday volume patterns

From Figure 5.5, we can see the volume traded in absolute terms on the left and daily proportions on the right. In both cases, the numbers are calculated as means for the given time period. The events triggered by the COVID-19 pandemic certainly attracted a lot of investors' attention, which caused the volume of trades to increase dramatically and stay high even after the situation on the US stock market calmed down a bit after the crash.

In this part of the analysis, we are interested in the changes to daily volume proportions patterns in 30-minute intervals. These changes are apparent from the right plot of Figure 5.5. However, we provide further analysis to test which intervals' differences are significant by employing the ANOVA method.

The change of daily pattern might be valuable information for models which

Figure 5.5: Volume traded of the USA500 index



implement volume values to forecast returns. It allows such models to change accordingly in times of unexpected events affecting the stock market.

5.4.1 ANOVA results

Table 5.6 shows the mean values for daily volume proportions calculated based on Equation 4.9. The values are grouped by days of the week and by time period. Values below the double-line titled "Average" are only grouped by time period. The ANOVA is employed on all groups in all 30-minute intervals.

We are interested if there is any difference between the means of the three time periods (as outlined in Section 3.2), so the null hypothesis we are trying to reject is $H_0 : \mu_{before} = \mu_{during} = \mu_{after}$. It is important to note that we cannot conclude that one period differs significantly from another one by rejecting the null hypothesis. Instead, we conclude that the three periods together differ significantly.

Looking at the results for an average day, we can see the proportions changed significantly at all 30-minute intervals, except between 11 am-12 pm, with at least a 0.05 level of significance. The most notable change is seen in the first half hour of trading between 9:30-10 am, where we observe up to 20% of daily volume being fulfilled on Mondays and 17% on an average day in the period before the pandemic, dropping to 13% and 14% respectively in the period during the pandemic crash. The volume in the first half hour remained lower in the period after the crash as well with 14%, on the other hand, this period has the highest volume of trades being fulfilled in the last 30 minutes of trading at 10% of the total daily volume.

Overall, we see the patterns changed significantly, mainly for a trading day's early and late hours. The most significant changes are on Mondays, which is understandable for the early hours as the market responds to news happening during weekends. However, interestingly the changes are also very significant throughout the whole day. On the other hand, we see only one significant change on Wednesday.

From the results, we can confidently conclude that the daily volume U-shaped curve flattened during the pandemic, as the stock market's attention was spread out more evenly. It would be interesting to obtain volume data for specific types of investors to see better what causes this change in pattern. It might very well be caused by individual investors, who perhaps do not pay much attention to their trades' timings. Whereas institutional investors often realise the trades in the first and last minutes of trading.

Table 5.6: Results of the ANOVA for 30-minute interval average proportions of daily volume

		9:30-10:00	10:00-10:30	10:30-11:00	11:00-11:30	11:30-12:00	12:00-12:30	12:30-13:00	13:00-13:30	13:30-14:00	14:00-14:30	14:30-15:00	15:00-15:30	15:30-16:00
Monday	Before	0.20 (0.05)	0.13 (0.03)	0.10 (0.02)	0.09 (0.02)	0.07 (0.02)	0.06 (0.01)	0.05 (0.01)	0.04 (0.01)	0.04 (0.02)	0.04 (0.02)	0.04 (0.02)	0.05 (0.02)	0.09 (0.03)
	During	0.13 (0.07)	0.15 (0.1)	0.08 (0.03)	0.08 (0.03)	0.08 (0.03)	0.08 (0.02)	0.06 (0.02)	0.06 (0.02)	0.05 (0.02)	0.06 (0.02)	0.05 (0.03)	0.05 (0.02)	0.07 (0.01)
	After	0.15 (0.05)	0.11 (0.03)	0.09 (0.02)	0.08 (0.02)	0.07 (0.02)	0.06 (0.02)	0.05 (0.01)	0.05 (0.01)	0.05 (0.01)	0.06 (0.02)	0.06 (0.02)	0.07 (0.02)	0.10 (0.03)
	F-Test	13.9***	4.9***	3.3**	3.6**	0.8	5.4***	3.7**	7.8***	2.8*	8.6***	9.6***	18.6***	5***
Tuesday	Before	0.17 (0.06)	0.12 (0.03)	0.10 (0.03)	0.08 (0.03)	0.07 (0.02)	0.06 (0.02)	0.05 (0.02)	0.05 (0.02)	0.05 (0.03)	0.05 (0.02)	0.05 (0.03)	0.06 (0.02)	0.09 (0.03)
	During	0.14 (0.09)	0.11 (0.08)	0.09 (0.05)	0.10 (0.05)	0.07 (0.03)	0.07 (0.03)	0.05 (0.03)	0.05 (0.04)	0.05 (0.03)	0.06 (0.03)	0.06 (0.03)	0.07 (0.03)	0.09 (0.04)
	After	0.14 (0.03)	0.11 (0.02)	0.09 (0.02)	0.09 (0.02)	0.07 (0.02)	0.06 (0.02)	0.05 (0.01)	0.05 (0.01)	0.05 (0.01)	0.05 (0.01)	0.06 (0.02)	0.08 (0.03)	0.11 (0.03)
	F-Test	6.3***	2.6*	0.5	0.6	0.5	0.4	0.2	0.5	0	1.4	2.2	9.2***	4.2**
Wednesday	Before	0.16 (0.05)	0.12 (0.05)	0.09 (0.03)	0.08 (0.03)	0.07 (0.03)	0.05 (0.02)	0.05 (0.02)	0.04 (0.02)	0.04 (0.02)	0.08 (0.06)	0.07 (0.06)	0.07 (0.04)	0.09 (0.03)
	During	0.13 (0.07)	0.10 (0.05)	0.09 (0.05)	0.07 (0.03)	0.06 (0.02)	0.07 (0.03)	0.05 (0.01)	0.05 (0.02)	0.07 (0.04)	0.07 (0.03)	0.07 (0.04)	0.07 (0.03)	0.10 (0.02)
	After	0.14 (0.04)	0.10 (0.03)	0.09 (0.03)	0.07 (0.02)	0.06 (0.02)	0.06 (0.02)	0.05 (0.02)	0.05 (0.02)	0.05 (0.02)	0.07 (0.03)	0.07 (0.03)	0.08 (0.03)	0.11 (0.03)
	F-Test	2.4*	0.8	0.4	0.2	1	0.8	0.5	1.7	5.3***	0.1	0.1	1.6	2
Thursday	Before	0.17 (0.05)	0.12 (0.04)	0.10 (0.03)	0.09 (0.03)	0.07 (0.02)	0.06 (0.02)	0.05 (0.02)	0.04 (0.02)	0.05 (0.03)	0.05 (0.02)	0.06 (0.04)	0.06 (0.03)	0.09 (0.03)
	During	0.10 (0.06)	0.10 (0.06)	0.08 (0.04)	0.07 (0.03)	0.10 (0.07)	0.08 (0.04)	0.08 (0.03)	0.08 (0.03)	0.07 (0.04)	0.06 (0.04)	0.06 (0.03)	0.06 (0.02)	0.06 (0.02)
	After	0.13 (0.04)	0.10 (0.03)	0.09 (0.02)	0.08 (0.02)	0.07 (0.01)	0.06 (0.01)	0.06 (0.02)	0.06 (0.02)	0.05 (0.02)	0.06 (0.02)	0.06 (0.02)	0.07 (0.02)	0.10 (0.03)
	F-Test	13.1***	4.1**	1.9	1.1	4.4**	3.8**	5.8***	13.3***	2.7*	3.6**	0.4	6.9***	4.9***
Friday	Before	0.17 (0.05)	0.13 (0.04)	0.10 (0.03)	0.09 (0.02)	0.08 (0.03)	0.06 (0.02)	0.05 (0.02)	0.05 (0.03)	0.04 (0.02)	0.05 (0.02)	0.04 (0.02)	0.05 (0.02)	0.10 (0.03)
	During	0.17 (0.07)	0.12 (0.04)	0.10 (0.04)	0.08 (0.03)	0.06 (0.02)	0.05 (0.02)	0.05 (0.02)	0.05 (0.02)	0.05 (0.02)	0.06 (0.02)	0.06 (0.02)	0.06 (0.02)	0.09 (0.02)
	After	0.14 (0.04)	0.11 (0.03)	0.09 (0.02)	0.08 (0.02)	0.07 (0.02)	0.06 (0.02)	0.06 (0.02)	0.05 (0.02)	0.05 (0.02)	0.06 (0.02)	0.06 (0.02)	0.07 (0.02)	0.09 (0.02)
	F-Test	3.9**	4.7**	0.1	0.9	2.8*	1.4	1.6	0.5	2.4*	4.8***	9.2***	9***	0.3
Average	Before	0.17 (0.06)	0.12 (0.04)	0.10 (0.03)	0.08 (0.03)	0.07 (0.02)	0.06 (0.02)	0.05 (0.02)	0.04 (0.02)	0.04 (0.02)	0.05 (0.04)	0.05 (0.04)	0.06 (0.03)	0.09 (0.03)
	During	0.14 (0.08)	0.11 (0.07)	0.09 (0.04)	0.08 (0.03)	0.07 (0.04)	0.07 (0.03)	0.06 (0.03)	0.06 (0.03)	0.06 (0.03)	0.06 (0.03)	0.06 (0.03)	0.06 (0.03)	0.08 (0.03)
	After	0.14 (0.04)	0.11 (0.03)	0.09 (0.02)	0.08 (0.02)	0.07 (0.02)	0.06 (0.02)	0.05 (0.02)	0.05 (0.02)	0.05 (0.02)	0.06 (0.02)	0.06 (0.02)	0.08 (0.03)	0.10 (0.03)
	F-Test	31.3***	12.6***	3.9**	2.4*	2.2	3.3**	6.3***	11.4***	9.2***	4.8***	6.9***	31.2***	11***

Notes: F-Test value corresponds to the outcome of ANOVA testing method mentioned in the methodology. Standard deviation is in parenthesis.
*, **, *** stand for 10%, 5% and 1% level of significance respectively, at which we reject the $H_0 : \mu_1 = \mu_2 = \mu_3$

Chapter 6

Conclusion

This work brings interesting results of pattern changes on the US stock market as a response to the COVID-19 pandemic, which has caused the market to crash heavily in March 2020. While the pandemic has been evolving every second, the US stock market is open for trading only during certain hours a day and is closed on weekends. This discrepancy of no trading during weekends, while there is a ton of new information being released, is the primary motivation why we focus mainly on the Monday and Friday effect. We employ the T-GARCH(1,1) model on a return series of the S&P 500 index and its eleven sectors. Our dataset runs from the beginning of 2018 until March 15, 2021, and by using dummy variables, we divide the timeline into three periods. The main period during the pandemic crash runs from February 3, 2020, just after the Trump administration declared a national public health emergency, to March 20, 2020, as the last observation before the FED announced unlimited quantitative easing.

Our results show significantly higher values of volatility during the crash across all sectors. Interestingly, some sectors show very little increase in volatility for the period after the crash, while heavily affected sectors by the pandemic like the Information technology, Financials, Consumer discretionary, Energy and Communication services, which have seen some major disruptions caused by the pandemic[1], show lasting impact in volatility levels. It is an excellent representation of how the pandemic affected various types of businesses in a long-term perspective and provides a great insight for investors, who can use this knowledge to better calculate the risks associated with businesses operating in the affected areas of the economy. Unfortunately, the results regarding Mon-

[1] IT - shortage of semiconductors, FIN - more unsecured loans to affected businesses, CD - heavy restrictions on hotels and restaurants, ENRG - oil prices falling below zero at one point, CMS - people spending more time online

day and Friday effects are mostly insignificant. However, we can still observe a pattern that the Monday returns during the crash were lower than Friday ones. On the other hand, the results are reversed for the bull market after the crash, where the estimates for Monday returns are generally higher than for Friday returns. This finding supports the calendar time hypothesis (French 1980), where the Monday returns include the returns of the two non-trading days and thus are higher than for other days of the week.

The thesis also analyses the investors' behaviour in terms of intraday volume trading patterns. Using 30-minute data, we observed the proportions of daily trading happening in half-hour intervals of a trading day. Using statistical analysis, we can confidently conclude that the trading pattern has changed across the observed three periods. The results show that before the pandemic crash, almost one-fifth of all daily trades happened in the first half-hour of trading between 9:30-10:00 AM. However, in the periods during and after the crash, the proportions evened out by a considerable margin. The standard U-shape pattern for daily trading has flattened during the pandemic crash as more of the volume took place throughout the whole day, but the biggest proportion of trades still happened in the first half-hour of trading for all periods.

Finally, we would like to point out topics for further research based on our analysis. As the results of the day-of-the-week effect are not significant enough to make firm conclusions, it would be interesting to see if that would change with more observations for the crash period. Unfortunately, the market quickly started recovering, and there is no way to obtain more daily observations. Nevertheless, working with high-frequency data might bring some more precise results for the Monday and Friday effects. For next research, it might also be interesting to look deeper into the intraday volume analysis as having data on the investor's type might tell us more about the causes of the pattern changes. Knowing whether institutional investors or small investors drove the change might help us better understand the effects of the COVID-19 pandemic on the US stock market.

Bibliography

Admati, A. R. & P. Pfleiderer (1988): "A theory of intraday patterns: Volume and price variability." *The Review of Financial Studies* **1(1)**: pp. 3–40.

Akaike, H. (1973): *Information Theory and an Extension of the Maximum Likelihood Principle*, pp. 199–213. New York, NY: Springer New York.

Baek, S., S. K. Mohanty, & M. Glambosky (2020): "Covid-19 and stock market volatility: An industry level analysis." *Finance Research Letters* **37**: p. 101748.

Baker, S. R., N. Bloom, S. J. Davis, K. J. Kost, M. C. Sammon, & T. Viratyosin (2020): "The unprecedented stock market impact of covid-19." *Working Paper 26945*, National Bureau of Economic Research.

Bartoszyński, R. & M. Niewiadomska-Bugaj (2008): *Probability and Statistical Interference*. Wiley series in probability and statistics. Hoboken, NJ: John Wiley & Sons, 2nd edition.

Bollerslev, T. (1986): "Generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics* **31(3)**: pp. 307–327.

Box, G. E. P. & D. A. Pierce (1970): "Distribution of residual autocorrelations in autoregressive-integrated moving average time series models." *Journal of the American Statistical Association* **65(332)**: pp. 1509–1526.

Brooks, C. & S. P. Burke (2003): "Information criteria for garch model selection." *The European Journal of Finance* **9(6)**: pp. 557–580.

Choi, S.-Y. (2021): "Analysis of stock market efficiency during crisis periods in the us stock market: Differences between the global financial crisis and covid-19 pandemic." *Physica A: Statistical Mechanics and its Applications* **574**: p. 125988.

Cross, F. (1973): "The behavior of stock prices on fridays and mondays." *Financial Analysts Journal* **29**: pp. 67–69.

Cushing, D. & A. Madhavan (2000): "Stock returns and trading at the close." *Journal of Financial Markets* **3(1)**: pp. 45–67.

Dickey, D. A. & W. A. Fuller (1979): "Distribution of the estimators for autoregressive time series with a unit root." *Journal of the American Statistical Association* **74(366)**: pp. 427–431.

Engle, R. F. (1982): "Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation." *Econometrica* **50(4)**: pp. 987–1007.

Fama, E. F. (1970): "Efficient capital markets: A review of theory and empirical work." *The Journal of Finance* **25(2)**: pp. 383–417.

French, K. R. (1980): "Stock returns and the weekend effect." *Journal of Financial Economics* **8(1)**: pp. 55–69.

Gao, L., Y. Han, S. Zhengzi Li, & G. Zhou (2018): "Market intraday momentum." *Journal of Financial Economics* **129(2)**: pp. 394–414.

Gibbons, M. R. & P. Hess (1981): "Day of the week effects and asset returns." *The Journal of Business* **54(4)**: pp. 579–596.

Glosten, L. R., R. Jagannathan, & D. E. Runkle (1993): "On the relation between the expected value and the volatility of the nominal excess return on stocks." *The Journal of Finance* **48(5)**: pp. 1779–1801.

Jaffe, J. & R. Westerfield (1985): "The week-end effect in common stock returns: The international evidence." *The Journal of Finance* **40(2)**: pp. 433–454.

Jaffe, J. F., R. Westerfield, & C. Ma (1989): "A twist on the monday effect in stock prices: Evidence from the u.s. and foreign stock markets." *Journal of Banking Finance* **13(4)**: pp. 641–650.

Jain, P. C. & G.-H. Joh (1988): "The dependence between hourly prices and trading volume." *The Journal of Financial and Quantitative Analysis* **23(3)**: pp. 269–283.

Just, M. & K. Echaust (2020): "Stock market returns, volatility, correlation and liquidity during the covid-19 crisis: Evidence from the markov switching approach." *Finance Research Letters* **37**: p. 101775.

Kiymaz, H. & H. Berument (2003): "The day of the week effect on stock market volatility and volume: International evidence." *Review of Financial Economics* **12(4)**: pp. 363–380.

Kwiatkowski, D., P. C. Phillips, P. Schmidt, & Y. Shin (1992): "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?" *Journal of Econometrics* **54(1)**: pp. 159–178.

Lakonishok, J. & E. Maberly (1990): "The weekend effect: Trading patterns of individual and institutional investors." *The Journal of Finance* **45(1)**: pp. 231–243.

Ljung, G. M. & G. E. P. Box (1978): "On a measure of lack of fit in time series models." *Biometrika* **65(2)**: pp. 297–303.

Mandelbrot, B. (1963): "The variation of certain speculative prices." *The Journal of Business* **36(4)**: pp. 394–419.

Okorie, D. I. & B. Lin (2021): "Adaptive market hypothesis: The story of the stock markets and covid-19 pandemic." *The North American Journal of Economics and Finance* p. 101397.

Onali, E. (2020): "Covid-19 and stock market volatility."

Osborne, M. F. M. (1962): "Periodic structure in the brownian motion of stock prices." *Operations Research* **10(3)**: pp. 345–379.

Richards, D. W. & G. D. Willows (2019): "Monday mornings: Individual investor trading on days of the week and times within a day." *Journal of Behavioral and Experimental Finance* **22**: pp. 105–115.

Smales, L. (2021): "Investor attention and global market returns during the covid-19 crisis." *International Review of Financial Analysis* **73**: p. 101616.

S&P Global (2021): "S&P Sector Definitions." Available at: `https://www.spglobal.com/spdji/en/landing/investment-themes/sectors/` [Accessed: 30/3/2021].

Tsay, R. (2005): *Analysis of financial time series.* Wiley series in probability and statistics. Hoboken, NJ: John Wiley & Sons, 2nd edition.

Uddin, M., A. Chowdhury, K. Anderson, & K. Chaudhuri (2021): "The effect of covid – 19 pandemic on global stock market volatility: Can economic strength help to manage the uncertainty?" *Journal of Business Research* **128**: pp. 31–44.

Zakoian, J.-M. (1994): "Threshold heteroskedastic models." *Journal of Economic Dynamics and Control* **18(5)**: pp. 931–955.

Zhang, D., M. Hu, & Q. Ji (2020): "Financial markets under the global pandemic of covid-19." *Finance Research Letters* **36**: p. 101528.

Zhang, J., Y. Lai, & J. Lin (2017): "The day-of-the-week effects of stock markets in different countries." *Finance Research Letters* **20**: pp. 47–62.

Ülkü, N. & M. Rogers (2018): "Who drives the monday effect?" *Journal of Economic Behavior Organization* **148**: pp. 46–65.

Appendix A

Appendix A - tables

Table A.1: P-values of ADF and KPSS tests

	ADF		KPSS	
	Level	Log-return	Level	Log-return
SP500	0.34	< 0.01	< 0.01	> 0.1
IT	0.54	< 0.01	< 0.01	> 0.1
HC	0.02	< 0.01	< 0.01	> 0.1
CD	0.49	< 0.01	< 0.01	> 0.1
IND	0.49	< 0.01	< 0.01	> 0.1
CS	< 0.01	< 0.01	< 0.01	> 0.1
FIN	0.61	< 0.01	0.019	> 0.1
MAT	0.74	< 0.01	< 0.01	> 0.1
ENRG	0.46	< 0.01	< 0.01	> 0.1
UTIL	0.03	< 0.01	< 0.01	> 0.1
RE	0.06	< 0.01	< 0.01	> 0.1
CMS	0.54	< 0.01	< 0.01	> 0.1

Table A.2: List of excluded days in the intraday volume analysis

Holiday / Shortened hours	**2019**	**2020**	**2021**
New Year's Day	January 1, 2019	January 1, 2020	January 1, 2021
Martin Luther King, Jr. Day	January 21, 2019	January 20, 2020	January 18, 2021
Washington's Birthday	February 18, 2019	February 17, 2020	February 15, 2021
Good Friday	April 19, 2019	April 10, 2020	
Memorial Day	May 27, 2019	May 25, 2020	
shortened hours	July 3, 2019		
Independence Day	July 4, 2019	July 3, 2020	
Labor Day	September 2, 2019	September 7, 2020	
Thanksgiving Day	November 28, 2019	November 26, 2020	
shortened hours	November 29, 2019	November 27, 2020	
shortened hours	December 24, 2019	December 24, 2020	
Christmas Day	December 25, 2019	December 25, 2020	

Appendix B

Appendix B - figures

Figure B.1: ACF and PACF of IT sector daily log returns

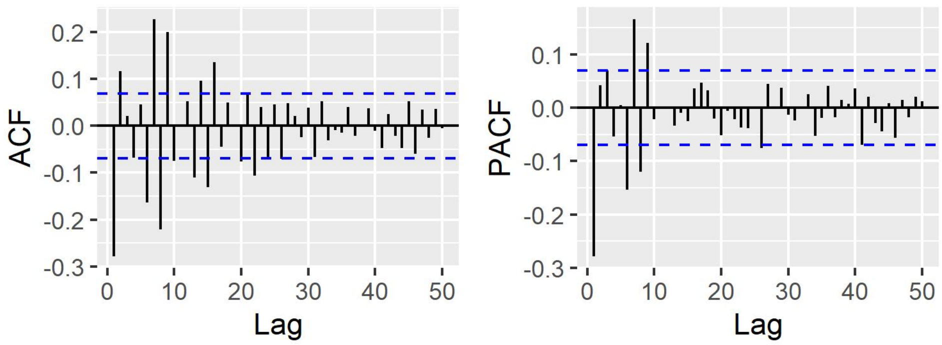


Figure B.2: ACF and PACF of HC sector daily log returns

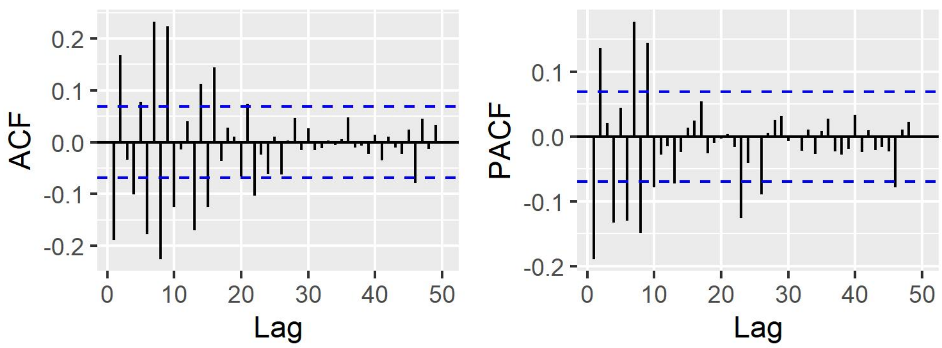


Figure B.3: ACF and PACF of CD sector daily log returns

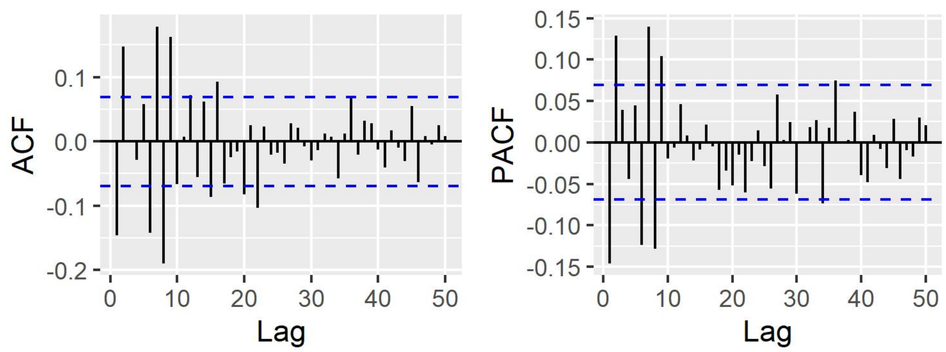


Figure B.4: ACF and PACF of IND sector daily log returns

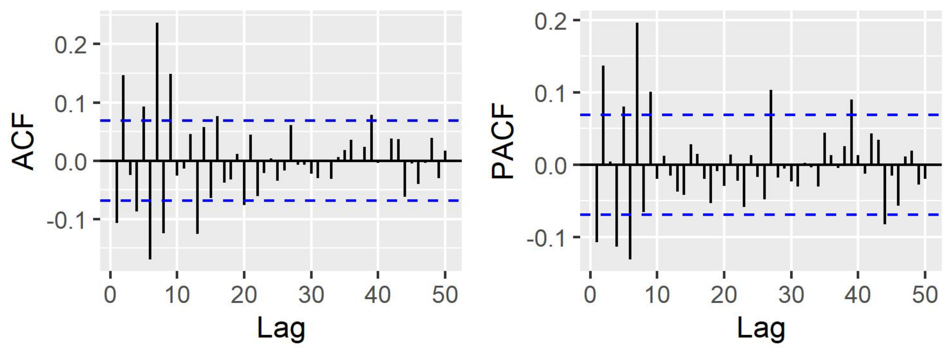


Figure B.5: ACF and PACF of CS sector daily log returns

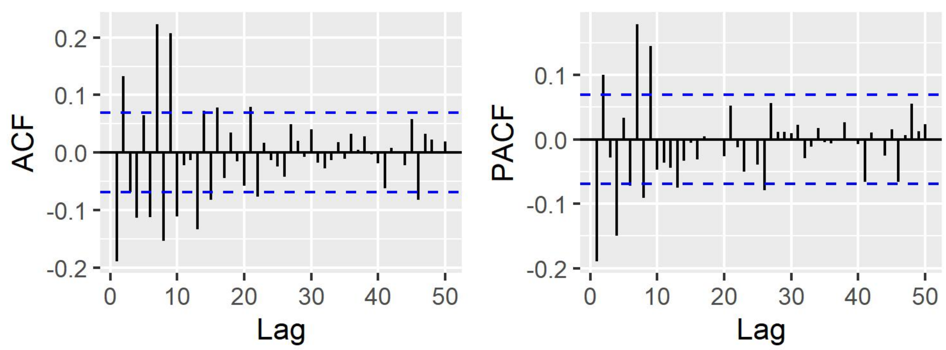


Figure B.6: ACF and PACF of FIN sector daily log returns

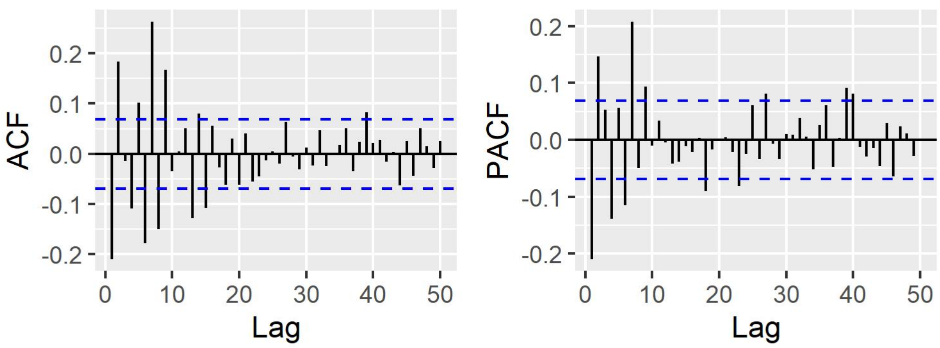


Figure B.7: ACF and PACF of MAT sector daily log returns

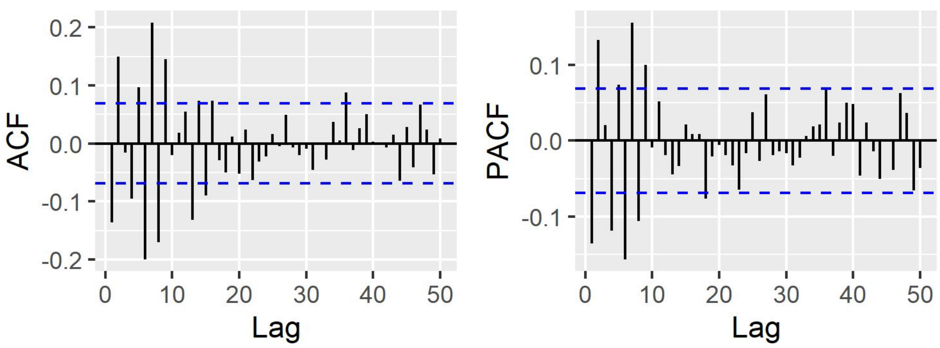


Figure B.8: ACF and PACF of ENRG sector daily log returns

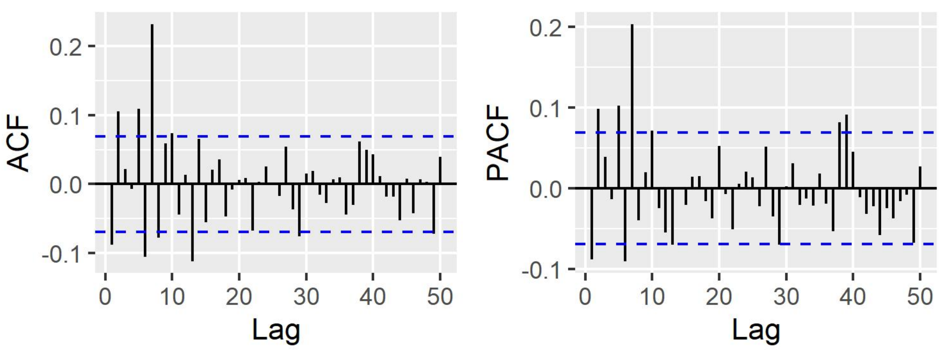


Figure B.9: ACF and PACF of UTIL sector daily log returns

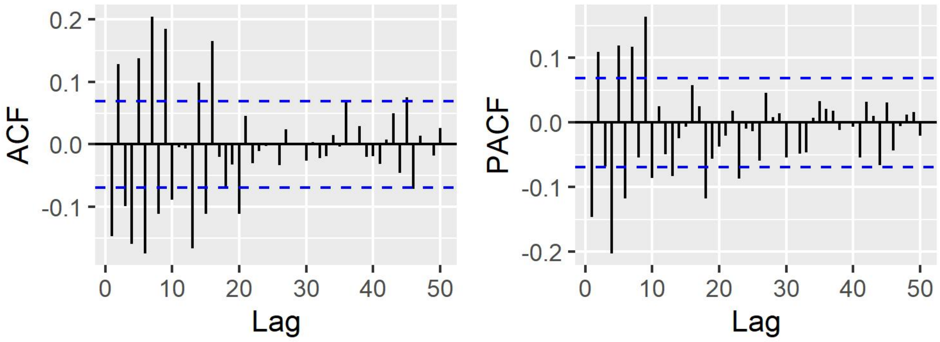


Figure B.10: ACF and PACF of RE sector daily log returns

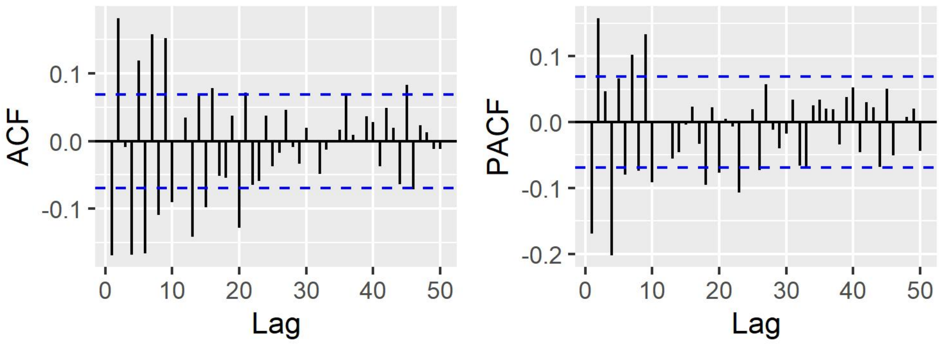


Figure B.11: ACF and PACF of CMS sector daily log returns