ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics
Bachelor Thesis (Major Financial Economics)

THE EFFECT OF INVESTOR SENTIMENT ON VOLATILITY: A COVID-19 CASE STUDY

Name student: Wouter Vermeij
Student ID number: 508128

Supervisor: Amar Soebhag
Second assessor: Jan Lemmen
Date final version: 27-10-2021

Abstract

Although a significant amount of research is conducted on the relation between investor sentiment and volatility, the findings are still somewhat inconclusive. This research paper aims to find significant patterns in the effect of investor sentiment on volatility. To accomplish this, two proxies for investor sentiment and two proxies for volatility are used. The data used, consisting of 209 observations, covers the period 2017-2020. This means that the covid crash on the stock markets is included in the data. This enables the opportunity to study a potential effect that the covid crash had on the relationship between investor sentiment and volatility. The multivariate time series dataset was studied using a simple OLS-regression, standard ARDL-Models and predictive ARDL-Models. The results seem to differ for which proxy for investor sentiment is used, but in general there is little evidence found that investor sentiment and volatility are even causal related at all.

Abstract

Although a significant amount of research is conducted on the relation between investor sentiment and volatility, the findings are still somewhat inconclusive. This research paper aims to find significant patterns in the effect of investor sentiment on volatility. To accomplish this, two proxies for investor sentiment and two proxies for volatility are used. The data used, consisting of 209 observations, covers the period 2017-2020. This means that the covid crash on the stock markets is included in the data. This enables the opportunity to study a potential effect that the covid crash had on the relationship between investor sentiment and volatility. The multivariate time series dataset was studied using a simple OLS-regression, standard ARDL-Models and predictive ARDL-Models. The results seem to differ for which proxy for investor sentiment is used, but in general there is little evidence found that investor sentiment and volatility are even causal related at all.

Table of Contents

1 Introduction **5**
1.1 Current Relevance 5
1.2 Supportive studies and conceptual model 6
1.3 Strategy and methods 7
1.4 Next chapters 7

2 Theoretical framework and hypotheses **8**
2.1 Theoretical framework 8
2.1.1 Concept of independent and dependent variables 8
2.1.2 Theoretical background: Behavioral finance 8
2.1.3 Investor sentiment and volatility 10
2.1.4 Possible changes during Covid-19 11
2.2 Hypotheses 12

3 Data **14**
3.1 Investor sentiment 14
3.1.1 The AAII-survey 14
3.1.2 The CBOE S&P500 put/call volume ratio 15
3.2 Volatility 16
3.2.1 S&P500 real data 16
3.2.2 The CBOE Volatility Index (VIX) 17
3.3 Control variables 19
3.3.1 The U.S. 10-Year Treasury Yield 19
3.3.2 The Dollar/Euro Exchange Rate 19
3.4 Summery statistics table 20

4 Methodology **21**
4.1 Hypothesis set 1: Correlation 22
4.2 Hypothesis set 2: Standard ARDL-Modelling 23
4.3 Hypothesis set 3: Predictive ARDL-Modelling 25

5 Results **27**
5.1 Results 27
5.1.1 Hypothesis set 1: Correlation 27
5.1.2 Hypothesis set 2: Standard ARDL-Modelling 30
5.1.3 Hypothesis set 3: Predictive ARDL-Modelling 36
5.2 Discussion of results 42
5.2.1 Discussion hypothesis set 1: Correlation 42
5.2.2 Discussion hypothesis set 2: Standard ARDL-Modelling 43
5.2.3 Discussion hypothesis set 3: Predictive ARDL-Modelling 45
5.2.4 Overview hypotheses and discussions 46

6 Conclusion **49**
6.1 Conclusion 49
6.2 Limitations and recommendations 51

7 References **53**

8 Appendix **58**
8.1 Chapter 4: Methodology 58
8.1.1 First observation in 95% interval of remaining observations 58
8.1.2 Multicollinearity table 59
8.1.3 Augmented Dickey-Fuller test outputs 60
8.2 Chapter 5: Results 61
8.2.1 Hypothesis set 1 output tables 61
8.2.2 Hypothesis set 2 output tables 64
8.2.3 Hypothesis set 3 output tables 69

Chapter 1: Introduction

1.1 Current Relevance

Due to the Covid-19 outbreak in early 2020, stock markets plummeted all over the world. The S&P500 dropped about 34% between February 19th and March 23rd, the Dow-Jones dropped 12.9% (second biggest percentage loss post WWII) and the Nasdaq reached its largest percentage loss ever. It was expected that Covid-19 would plunge the world in the worst recession since WWII (The World Bank, 2020). The three main drivers of uncertainty and, therefore, the pessimistic economical view in asset markets were the epidemiological evolution, the economic outlook and the policy measures (Benigno, Canofari, Di Bartolomeo & Messori, 2020). While a positive development of these factors could result in just a small bump for the economy, expectations were not that sanguine. It was expected by The World Bank that the global economy would shrink with 5.2% and the economic activity of advanced economies to shrink with 7%, thus stock markets returning to their pre-covid levels seemed rather far away. Against these expectations of the market, the Nasdaq already hit pre-covid levels in June 2020. By November, the S&P500 and the Dow Jones reached their pre-covid levels as well. Although the pandemic was still very present, investors seemed quite positive about the future of the economy. Currently, a year after the 'covid-crash' in the stock market, stock markets are reaching all-time highs.

Apparently, many people wanted to take advantage of the rising stock markets. According to Bloomberg Intelligence via The Wall Street Journal, individual investors were responsible for 19.5% of the equity trading volume in the U.S. in 2020 which is a jump of 4 percentage points compared with 2019 (Osipovich, 2020). This might be a signal that individual investors seem confident about a recovery of the pandemic. Following the evidence from multiple studies, one might argue that the high returns during the recovery could lead to investor confidence and higher investor sentiment (Statman, Thorley & Vorkink, 2006; Daniel & Hirshleifer, 2005). This could lead to higher volatility (Daniel, Hirshleifer & Subrahmanyam, 2005). Contrarily, Lee, Jiang and Indro (2002) demonstrate that bullish changes in investor sentiment leads to downward revisions in volatility. Since the evidence seems rather puzzling, the question is whether stock markets become more volatile if investor sentiment rises and whether we see this during the economic recovery from the Covid-19 pandemic. In this way, this research contributes to the existing literature. Furthermore, the fast changes in investor sentiment caused by the pandemic provide an interesting case study.

1.2 Supportive studies and conceptual model

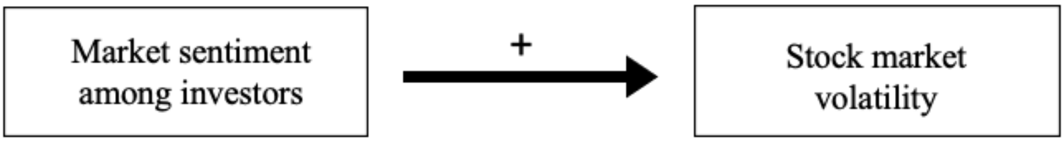
Verma and Verma (2016) show in their research that positive and negative sentiment in the market have asymmetric on excess return volatility. Furthermore, Haridas and Rishad (2020) show that (irrational) investor sentiment causes excess market volatility. Although these studies are conducted in India, the impact of investor sentiment is also found in other countries. Research on the relationship between investor sentiment and volatility has also been conducted in China by Qiang & Shu-e (2009). They found that investor sentiment is a systematic factor in forming stock prices. In addition, they find that the impact of changes in negative or positive investor sentiment is different. Moreover, Chi et al. (2012) found that investor sentiment has a great impact on stock returns. Later research by Chiu et al. (2018) even shows that transitory volatility is not related to macroeconomic fundamentals but instead is associated with changes in investor sentiment. Also, Yang and Copeland (2014) show that bullish believes among investors result in higher volatility in the short run. All these studies taking into account a positive effect of investor sentiment on volatility might be expected.

As a result of the increased number of individual investors mentioned earlier on, it might be interesting to compare the situation before and after the epidemic outbreak. As shown by Foucault, Sraer and Thesmar (2011) with French data, uninformed individual investors generate volatility in the valuations of publicly listed companies. This might be increased in times of high market returns since individual security turnover is positively related with lagged security and market returns (Statman, Thorley & Vorkink, 2006). Times of high market returns may also lead to more overconfidence (Daniel & Hirshleifer, 2005) among individual investors, but also among experts and professionals like investment bankers (Glaser, Langer and Weber, 2013) and CFO's (Ben-David, Graham and Harvey, 2013). This overconfidence leads to under- and overreaction on stock markets and excess volatility (Daniel, Hirshleifer & Subrahmanyam, 2005). The higher overconfidence also results in excessive trading (Trinugroho & Sembel, 2011), which is already present among individual investors which brokerage accounts (Odean, 1999). Excessive trading might further increase under- and overreaction, leading to even higher volatility. Therefore, the effect of the sentiment might be higher during the economic recovery of the covid pandemic.

The research question of this paper is: "Has investor sentiment a significant impact on stock market volatility and did this change during the economic recovery from the covid pandemic?". The independent variable *investor sentiment* can be defined as the overall believe in economic prosperity and the believe that stock prices overall will go up. Investor sentiment is widely described as the combination of the reaction of investors as a result of the current market situation and unjustified expectation of the future cashflows (Baker & Wurgler, 2006; 2007). The dependent variable *volatility* is mostly described as the standard deviation or variance of returns of a certain asset within a certain time period. Taking the previous mentioned literature into account is would make the most sense to expect a positive relationship. Figure 1.1 shows the straightforward model that can be expected for the mentioned studies.

Figure 1.1: The expected relationship

A very straight forward representation of the relationship that is expected to be found when conducting the research.



1.3 Strategy and methods

In order to examine the proposed relationship between investor sentiment and volatility a multivariate time series study is done. This is considered the most pertinent to the relationship because it allows to examine the relationship between both the variables through time. Furthermore, this provides the opportunity to add lags. The multivariate data is analyzed by performing an OLS-regressions and multiple ARDL-models that are also structured to examine predictive power of the independent variable on the dependent variable. Where OLS is performed once, the ARDL models are used mulitple times for different time periods. This allows to study the potential relationship before covid and after the crash. In order to increase robustness, for both the independent and dependent variable two proxies were used. Lastly, control variables are added to increase validity.

1.4 Next chapters

The following chapter reviews the theories regarding both behavioral finance in general and the relationship between investor sentiment and volatility. At the end of this chapter, the examined hypotheses are provided. Next, in chapter 3, the used data will be discussed, including a motivation and the time periods. Subsequently, chapter 4 contains the different rmethods used. Chapter 5 will present the outputs of the tests and models. This will include a separate discussion of the results for each hypothesis. Finally, chapter 6 contains the conclusion based on the results and findings, which comes together with a discussion of the limitations of the study, and future research recommendations.

Chapter 2: Theoretical Framework and Hypotheses

2.1 Theoretical Framework

In order to understand where the potential relationship between investor sentiment and stock market volatility is coming from the concepts, the background of behavioral finance and the more relationship specific literature is discussed. Lastly, based on data and established literature potential relationships are discussed that would argue in favor of changes in the effect through the covid pandemic.

2.1.1 Concept of Independent and Dependent Variable

Investor Sentiment

Investor sentiment is broadly defined as the expectations and beliefs about investment risk and potential cashflows of particular securities or financial markets that is not justified by the facts in hand of the public (Baker & Wurgler, 2007). When investor sentiment is high overall beliefs about future cash flows and risks are positive and when investor sentiment is low the expectations of future cashflows and risks are rather gloomy.

Volatility

Poon & Granger (2005) defines volatility as “the spread of all likely outcomes of an uncertain variable”. When looking at financial markets the focus of volatility is on the spread of asset returns. According to Figlewski (1997) and Poon volatility is generally undesirable since it is viewed as an indicator that asset prices are unreliable and the financial markets are not functioning efficiently.

2.1.2 Theoretical Background: Behavioral Finance

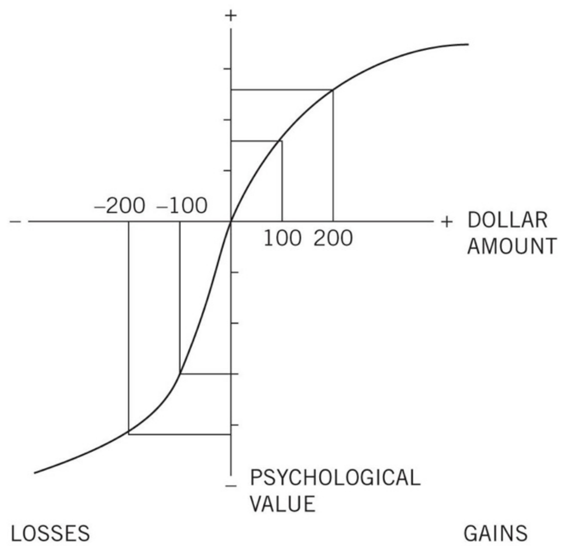
In standard finance it is assumed that the Efficient Hypothesis (EHM) holds. It proposes that investors are rational and that they consider all information when making choices considering their portfolio (Joo & Durri, 2018). Furthermore, it proposes that prices reflect all available information and that only new information should lead to price changes. This information efficiency means that participants cannot outperform the market on a consistent basis (Fama, 1965). Over the years, these assumptions have been examined by many studies. These studies have shown that EHM does not hold at all and investors are prone to multiple decision-making biases. Investors make systematic and sub-optimal decisions as a result of heuristic simplifications, which are systematic errors in judgement (Chen et al., 2004).

Behavioral finance combines behavioral (economic) insights and finance and tries to identify and explain the systematic decision errors. The two most important topics in behavioral finance regarding stock markets will shortly be discussed in the following subparts.

The first topic is the Prospect Theory of Daniel Kahneman and Amos Tversky (1979). In their research they study how people make choices and value uncertain outcomes. The develop a simple model (see Figure 2.1) that incorporates many behavioral thraits like reference dependence, diminishing sensitivity and loss aversion. Their model clearly shows that people are more sensitive to reductions in wealth than to increases. This leads to investors selling winning stocks too early and keeping losing stocks for too long since they don't want to realize the losses. This is called the *disposition effect* (Shefrin & Statman, 1985; Odean, 1998). Barberis, Huang and Santor (1999) apply the prospect theory to financial markets. They use the idea that prior losses make investors more risk averse and prior gains make them less risk averse (Thaler & Johnson, 1990), called the *house money effect*. This results in time-varying sentiments towards risk, allowing for asset prices to change accordingly.

Figure 2.1: Prospect Theory model

The simple graphical representation of the Prospect Theory that clearly shows that people are loss avers and have diminishing sensitivity.



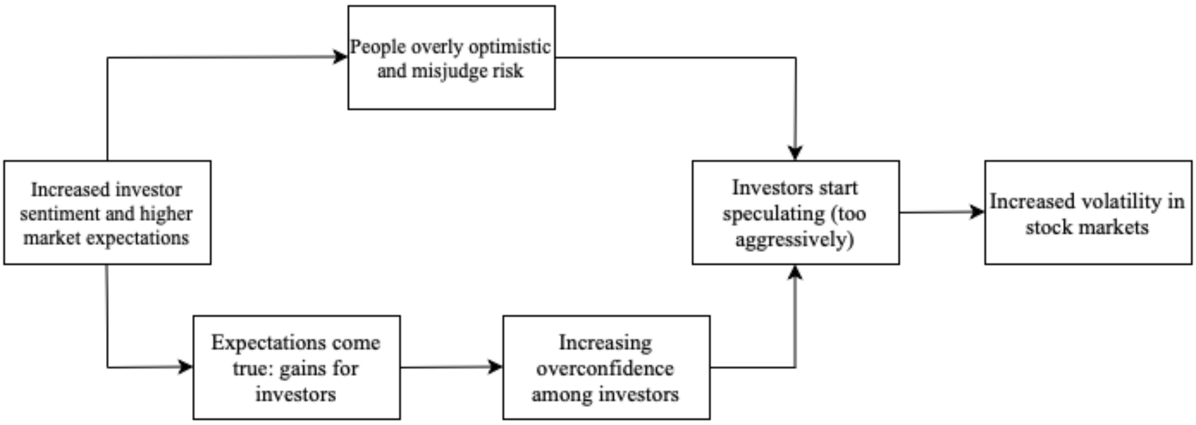
Another important aspect within behavioral finance are so-called *noise traders*. Noise traders are trading although they would be better of not trading (Black, 1986). Possibly they think they are trading on information which is in fact noise. Or maybe they just enjoy trading. Noise trading leads to excess trading and is a source for speculation (Vitale, 200). As a result of noise traders, asset prices can diverge significantly from their fundamental values (De Long et al., 1990). De Long et al also show that noise traders can send prices in certain directions while all other traders remain rational. Moreover, the presence of noise traders is partly responsible for the volatility on stock exchanges. Most typically noise traders overinflate asset prices during bullish times and excessively deflates them during bearish times. Furthermore, noise traders are likely to go with the hype (CFI, 2020) which can be relevant when looking at the covid recovery. In most cases noise traders do not have a professional background.

2.1.3 Investor Sentiment and Volatility

Relationships between investor sentiment and volatility is one of the most important subjects within behavioral finance (He, Zhu and Gu, 2020). It is clear that investors are influenced by behavioral biases and that these behavioral biases are (partly) responsible for excess volatility in the stock markets. We will now move closer to the studied relationship and look into established research concerning investor sentiment and volatility.

De Long et al. (1990) explain that the presence of noise traders with a strong bullish sentiment leads to higher divergence between fundamental values and stock prices. When sentiment among investors is positive, their expectations of the returns is also positive (Haridas and Rishad, 2020). This may lead to more (aggressive) speculation among investors. Naive noise traders in their eager to exploit the situation misjudge potential risk and are overly optimistic. This willingness to trade and overly optimistic view can be strengthened by overconfidence in case of the expected positive market returns. Investors suffering the *self-attribution bias* think they have investment skills when experiencing gains leading to overconfidence (Czaja & Röder, 2020), even when their gains are not exceeding the market returns. Barber and Odean (2000) claim that overconfidence leads to overestimation of the precision of the information at hand. Therefore, investors will overestimate their potential gains. This effects not just a few individual investors, but it applies market wide (Statman, Thorley and Vorkink, 2006; Daniel & Hirshleifer, 2005). Consequently, noise traders and overconfident investors keep prices diverging even more which leads to more volatility in the market. In this way a decrease in sentiment would mean that people be more careful and rational while trading. This should also lead to a decrease in volatility. The expected positive relationship between investor sentiment and volatility explained by the literature above has been backed by established empirical studies (Fang et al., 2018; Yang & Copeland, 2014). Figure 2.2 on the next page provides a clear overview of how higher investor sentiment potentially increases the volatility in stock markets.

Figure 2.2: How investor sentiment effects volatility
A clear and simple overview demonstrating how increases in investor sentiment should lead (in two ways) to increases in stock market volatility



2.1.4 Possible changes during Covid-19

With the covid-19 outbreak in the beginning of 2020 and the quick recovery of the stock markets later on, 2020 has been a very exceptional year. The crash and the comeback from the crash were seen by many as an opportunity to make money on the pandemic. The fact that the market plummeted should have led to big changes in investor sentiment. Somewhere in the beginning of the recovery has to be a shift from a negative sentiment to a positive sentiment as stock markets started to recover fast and provided the opportunity to take advantage of the unique situation.

This in combination with the rise of free trading apps boosted the trading by individual investors. Recall that in 2020 almost 20% of the trading volume in U.S. equities was done by individual investors, which is an all-time high (Osipovich, 2020). This increase in investors was mainly due to young and inexperienced individual investors (Fitzgerald, 2020). According to Michael Krause, chief investor officer at Counterpoint Mutual Funds, these new investors make all the classic mistakes in the short run. By investing in stocks that are high volatile, had lottery ticket like high payoffs or had low recent price momentum they can make a profit out of the situation back then, but it won't be profitable on the long term. Tim Welsh, founder/CEO of Nexus Strategy compared the investing style of new investors gambling. Something that is very observable from the fact that during the covid recovery some stocks have been used as speculative assets only (e.g. AMC and GameStop).

Since most noise traders are non-professionals, the increase in individual investors which are mostly inexperienced might lead to bigger effects of the previously mentioned biases. These non-professional investors are more subject to sentiment than professionals (Baker &

Wurgler, 2007). Furthermore, the increase in noise traders and speculators that are willing to go with the hype could lead to bigger hypes and even higher divergence in stock prices and their fundamental values. Consequently, the speculative nature of the investors who entered the market during the recovery of the covid pandemic could have potentially led to a stronger positive effect of investor sentiment on volatility.

2.2 Hypotheses

Based on the literature discussed in the subsections and the resulting theoretical reasoning it might be expected that the research question follows these presumptions. While a positive relationship is indeed expected, the empirical evidence does not give a one-sided answer. Therefore, it might be for the better to not exclude a potential negative effect form the research question. Consequently, the main research question of this paper is:

Has investor sentiment a significant impact on stock market volatility and did this change during the economic recovery from the covid pandemic in 2020?

In order to analyze this question, the research is divided into separate hypothesis that are used to analyze the topics using different methods. First of all, to look at the impact of investor sentiment on volatility, it is examined whether they are correlated. Correlation does not imply causation, but it could offer initial orientation on the relationship focused on. Recall that based on established literature a positive effect and thus a positive correlation is expected. Since one of the goals of this research is to examine whether a potential effect has changed, the test includes two periods that are explained in Chapter 4. Lastly, results are analyzed to examine whether the correlation has changed. Therefore, the first hypotheses are:

H1A: *Investor sentiment is positively correlated with volatility before the covid-crash*

H1B: *Investor sentiment is positively correlated with volatility during the covid recovery until 2021*

H1C: *The correlation between investor sentiment and volatility has significantly changed during the covid recovery until 2021 compared with before*

Secondly, it is tested whether there is more of a causal and modelling relationship between investor sentiment and volatility. This will be done by taking lags of both the independent and dependent variable into account. This will be performed two times to compare both the effect before and after the covid-crash. Just as with the first set of hypotheses, the results of the different timespans are compared to see if the result differ. Once again, based on the formed theory, a positive effect is expected. Hence, the second set of hypotheses are stated as follows:

H2A: *When adding lagged values of both the investor sentiment and volatility, investor sentiment has a significant positive relationship with volatility before the covid-crash*

H2B: *When adding lagged values of both the investor sentiment and volatility, investor sentiment has a significant positive relationship with volatility during the covid recovery until 2021*

H2C: *The relationship between investor sentiment and volatility has significantly changed during the covid recovery until 2021 compared with before*

Lastly, it is tested whether past and current values of investor sentiment and volatility can help forecasting the future value of volatility. This will be done by only using lags of both the independent and dependent variable into account. Just as with the hypotheses stated above, this too will be performed two times to compare both the effect before and after the covid-crash. Once again, results of the different timespans are compared to see if the result differ. Hence, the second set of hypotheses are stated as follows:

H3A: *When using only lagged values of both the investor sentiment and volatility, investor sentiment positively significantly predict volatility before the covid-crash*

H3B: *When using only lagged values of both the investor sentiment and volatility, investor sentiment positively significantly predict volatility during the covid recovery until 2021*

H3C: *The predictability of volatility based on investor sentiment has changed during the covid recovery until 2021 compared with before*

Chapter 3: Data

This research uses a multivariate time series dataset. In order to come to the multivariate time series dataset used for the analysis, data had to be retrieved from different sources. Moreover, some of the data was not directly applicable and had to be transformed in data that would be useful to the analysis. In the following subsections the source, most basic statistics and considerations for using the data are touched upon.

3.1 Investor Sentiment

3.1.1 The AAII-Survey

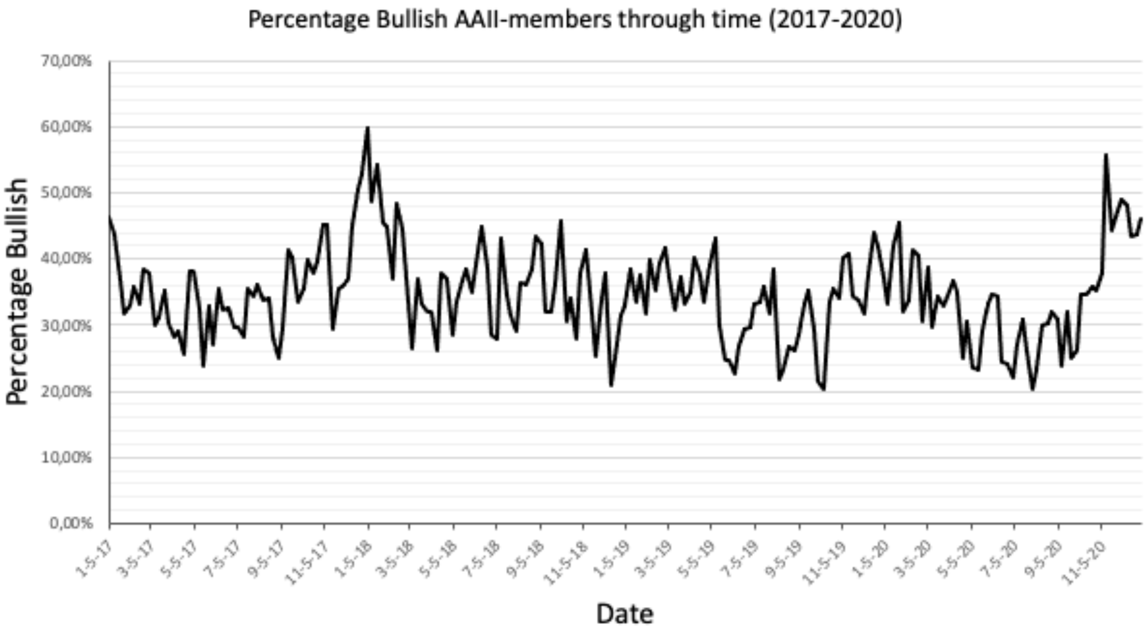
In order to get a solid answer on the research question multiple measures of investor sentiment are used. First of all, the AAII (American Association of Individual Investors) Investor Sentiment Survey is used. In this survey AAII-members (over 160.000) are asked whether they are bullish, neutral or bearish on the stock market for the next six months. The data of this survey is publicly available and easy to use. The outcomes of this survey are in percentages which enable the opportunity to analyze the data from different point of views. Brown and Cliff (2005) argue that the data coming from this survey is a fair indicator for the real sentiment in the market.

Nevertheless, individual investors don't even make up 20% of the trading volume (Osipovich, 2020). So, despite the claims of Brown and Cliff the data might not be (exactly) representative for the sentiment in the whole market which could be a major drawback. Another drawback is the fact that there is just one observation per week which is used to measure the sentiment during that week. Since volatility is measured as an average within a week, the observations are not made simultaneously. This may result in deviations from the optimal situation where observations in investor sentiment and volatility are made simultaneously.

To compare pre-covid investor sentiment with the investor sentiment during the covid-19 recovery the data from January 2017 until December 2020 is used. Since the data is weekly, this leaves us with 209 observations. One of the indicators for higher investor sentiment we can make out of this data is the percentage of people who think the stock market is going up the next six months. Looking at this variable we get an average of 34,69% and a standard deviation of 7,06%. Furthermore, to give a simple overview of this number through time, the graph is provided on the next page in Figure 3.1.

Figure 3.1: The AAII-Survey through time

The percentage bullish respondents to the AAII-Survey through time. A steep increase in the investor sentiment is observable during the recovery from the pandemic.



3.1.2 The CBOE S&P500 put/call volume ratio

In order to get more robust results another proxy for investor sentiment is used, namely the Put/Call volumes ratio of the S&P500. The put/call volume ratio is defined by the number of put options sold divided by the number of call options sold. Since people should be more likely to buy call options instead of put options when they expect markets to go up, a low Put/Call volume ratio should mean that sentiment in the market is relatively high. Similarly, if the Put/Call volume ratio is high, investors are more betting on the market to go down. For the purpose of this research the S&P500 Index options are used since the volatility variables are also scoping the S&P500. These daily ratios are retrieved from the database of the Chicago Board Option Exchange (CBOE). In order to make it correspond with the weekly data of other variables, the daily ratios are grouped for each week to compute the average ratio within that week.

The put/call volume ratio is used and verified in established research and literature (Yang & Copeland, 2014; Bandopadhyaya & Jones, 2011), but when using logical reasoning it definitely is not flawless. Since prices of options are not constant, a certain volume ratio at different observations does not mean the investor sentiment is exactly the same. The volume traded is very dependent of moneyness and expiration dates. The put/call volume ratio from the CBOE covers different expiration dates which can affect the ratio as well. Although the same values of the ratio might present different levels of investor sentiment, taking the first differences of the ratio might be a solution. In this way changes in the ratio can be compared with changes in volatility which could still help forming a more robust answer on the research question.

For the purpose of this research the S&P500 Index options are used since the volatility variables are also scoping the S&P500. The timespan is once again 2017 until 2020 so that it matches with the other data. In order to make it correspond with the weekly data of other variables, the daily ratios are grouped for each week to compute the average ratio within that week. This means that there are also 209 observations. The graph below gives a simple overview of the data.

Figure 3.2: The put/call volume ratio through time

The average weekly put/call volume ratio on the S&P500 through time. A steep decrease in the ratio is visible at the covid-19 crash around March 2020.

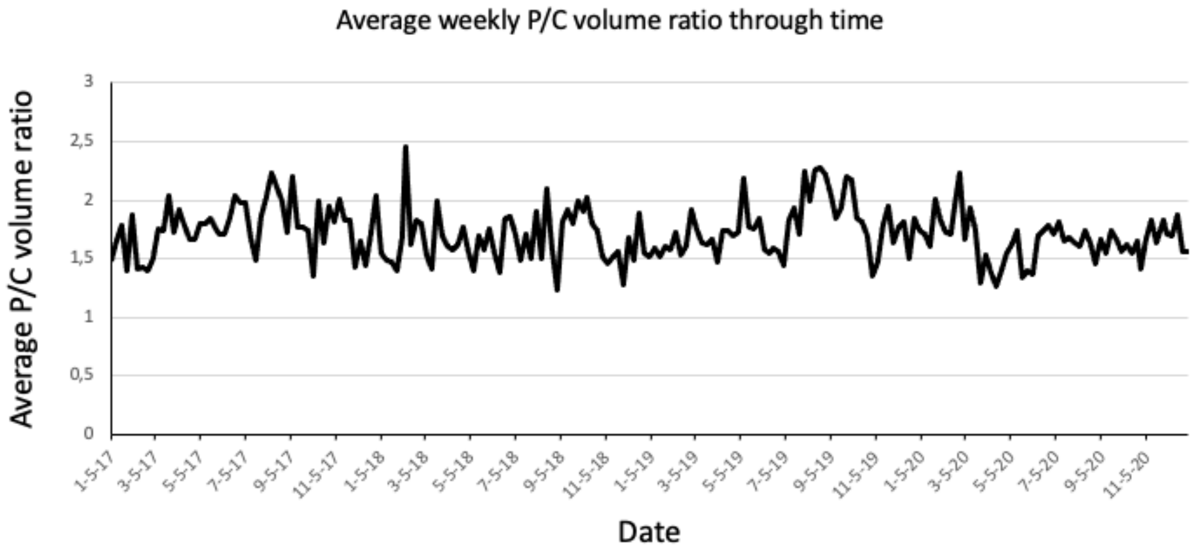


Figure 3.2: The average weekly P/C volume ratio on the S&P500 through time. A steep decrease in the ratio is visible at the covid-19 crash around March 2020.

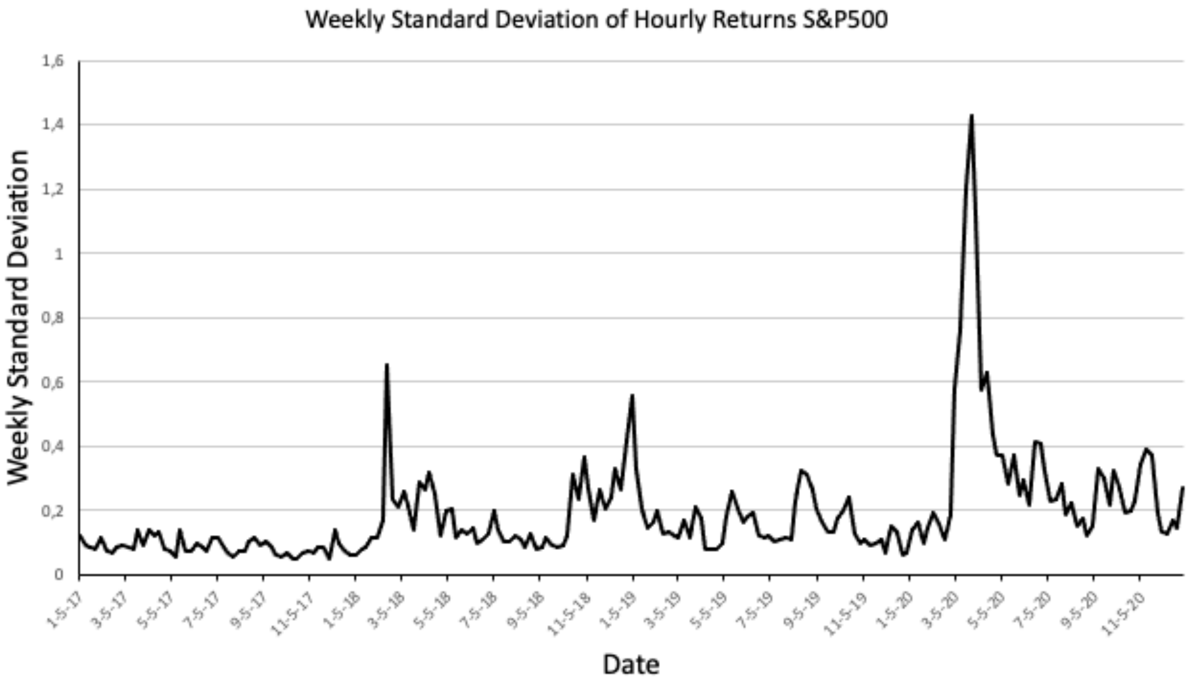
3.2 Volatility

3.2.1 S&P500 real data

The first variable to analyze stock market volatility the data of the S&P500 is used. In order to give a weekly volatility indicator, the hourly index of the S&P500 is retrieved from Tick Data. Since we want to have an indicator of the volatility of a certain week the standard deviation of hourly stock returns within that week is used. To come up with this the hourly returns are calculated and divided in groups corresponding with the weeks. Lastly, the hourly standard deviation per week is calculated. The big advantage of this is that in this way, a weekly volatility index is formed using the real values of the S&P500. Therefore, the data won't allow for measurement errors. On the other hand, the drawback of using this data is that S&P500 is just an index. Therefore, it does not or barely react to high volatility of single stocks and stocks that are not in the S&P500. The data retrieved is from January 2017 until December 2020, corresponding with the data of the AAII on investor sentiment. Consequently, also this data exists of 209 observations and has an average weekly standard deviation of hourly returns of 0.191. The graph below gives a simple overview of the data through time.

Figure 3.3: The weekly SD of hourly S&P500 returns through time

The weekly standard deviation of hourly returns of the S&P500 through time. The enormous spike marks the covid crash in the beginning of 2020.



3.2.2 The CBOE Volatility Index (VIX)

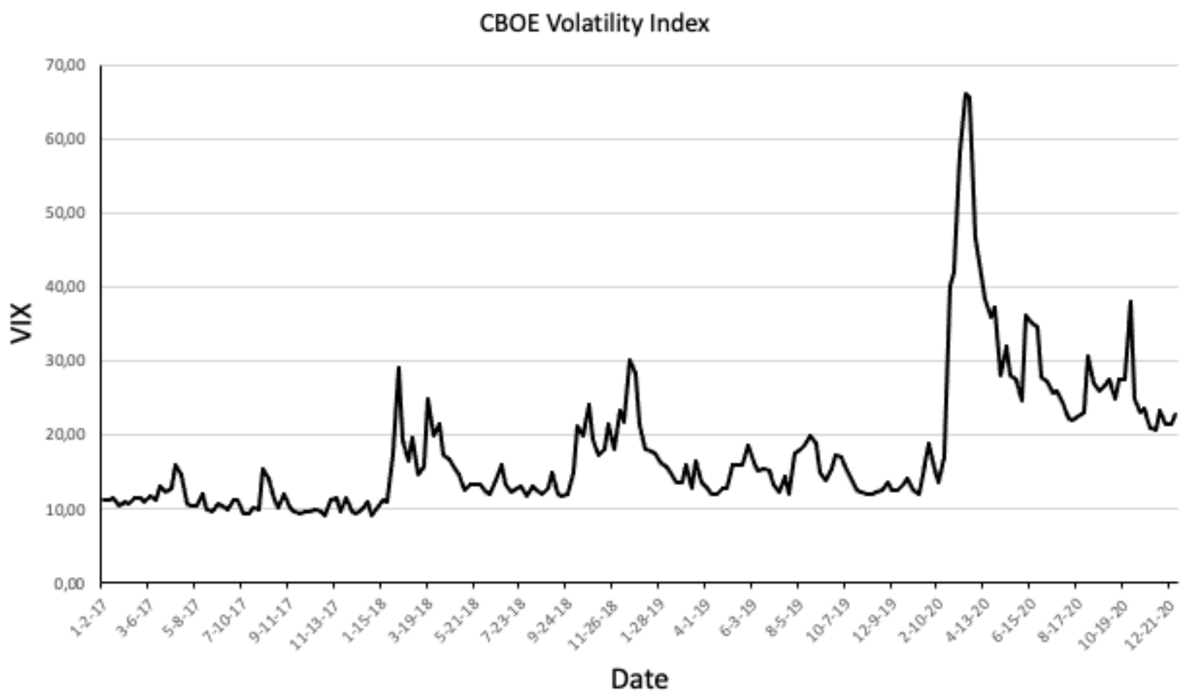
Another measure used to analyze stock market volatility in the U.S. is the CBOE Volatility Index, better known as the VIX. The VIX is a benchmark index that is designed to track the volatility of the S&P500. In contrast with the real data of the S&P500, the VIX is more of an expectation of the volatility in the short term since it is calculated using prices of S&P500 index options. For this expected volatility the prices of options expiring between 23 and 37 days are used. The data of the VIX is retrieved from the database of Yahoo Finance.

An empirical study of Chow, Jiang and Li (2020) pointed out that the VIX in general understates the true volatility so one might argue that using the VIX while the real data is available seems illogical. Although this makes a good point, using the VIX provides an interesting opportunity to compare estimations with real data when investor sentiment is higher/lower. The fact that people increasingly misjudge risk when sentiment is higher could allow for a different potential correlation in comparison with the true volatility.

Yahoo Finance offers the opportunity to retrieve the VIX weekly. On the one hand this is an advantage because this means that the data does not had to be adjusted to match the other variables. On the other hand, the observations are made on Monday where the AAII survey is conducted on a Thursday and the real S&P500 data covers the whole week. Of course, this can be seen as a drawback, but this does also provide an extra opportunity to look more directly to potential causality. The graph below gives a simple overview of the VIX through time.

Figure 3.4: The CBOE VIX through time

The CBOE Volatility Index through time. The enormous pike marks the covid crash in the beginning of 2020. During the recovery, the VIX stays relatively high (above 20) whereas the real volatility of the S&P500 Index seems to return to 'normal' levels.



3.3 Control Variables

Including control variables in the analyzes means that a greater proportion of variation in the dependent variable will be explained by the variables that are in the model. Therefore, the effect of the independent variable on the dependent variable focused on is better isolated. Therefore, the internal validity of the study is enhanced by using control variables.

3.3.1 The U.S. 10-Year Treasury Yield

The first control variable that is used are changes in the U.S. 10-year treasury yield. This yield is the return in percentages on the debt obligations of the U.S. government with a time to maturity of 10 years. In the discount function, which is one of the basic building blocks of finance, risk-free rate that is used is often derived from the U.S. treasury curve (Gürkaynak, Sack and Wright, 2007). Following the Capital Asset Pricing Model this risk-free rate has important role in determining the prices of assets. If the treasury yield, which is considered the risk-free rate, goes up prices will be discounted at a higher factor resulting in lower assets prices (Fama & French, 2004). The same works vice versa. Therefore, changes in the treasury yield result in changes in stock prices creating volatility. Consequently, by controlling for changes in the treasury yield, the changes in volatility caused by changes in treasury yield will be filtered out. In this way the validity will be improved by reducing potential omitted variable bias.

3.3.2 The Dollar/Euro Exchange Rate

The second variable that will be controlled for are changes in the dollar/euro exchange rate. Many conducted studies have showed that changes in exchange rates cause increases in the volatility of stock prices (Özbey, Can and Tra, 2016). Also established research on the S&P500 stock prices confirm that exchange rates have a significant impact (Kim, 2003). Changes in exchange rates allow for changes in the international competitiveness and trade balance. This results in changes in real economic variables, such as real income and output (Dornbusch and Fischer, 1980). An appreciation of a domestic currency will be a burden for exporters affecting its share price (Tian & Ma, 2010). Controlling for changes in the exchange rate between the U.S. and Europe could potentially improve the validity of the research by lowering potential omitted variable bias.

3.4 Summery Statistics Table

Below a table with descriptive statistics about the data used is given in order to give a better idea about what the data looks like. The final data set consist of 209 weekly observations from 2017 until 2020.

Table 3.1: Table with summery statistics for the used data

Table contains the summery statistics for both the dependent and independent variables. Also the control variables are included.

Statistic	Mean	St. Dev.	Min	Max
Dependent variables				
S&P500 St. Dev. Hourly Returns	0.1912	0.1744	0.0479	1.4299
CBOE Volatility Index	17.7966	9.2432	9.14	66.04
Independent variables				
AAII Bullish Individual Investors (proportion)	0.3469	0.0706	0.2023	0.5975
CBOE Weekly Average Put/Call Volume Ratio	1.7158	0.22	1.232	2.452
Control variables				
Changes in U.S. 10-Year Treasury Yield	0.0668	0.0679	0.001	0.583
Changes in USD/EUR Exchange Rate	0.0088	0.0073	0.0001	0.0472

Chapter 4: Methodology

As mentioned in the hypotheses section the relationship and potential changes in the relationship will be tested by multiple approaches and methods. All the performed tests are run by the statistical software Stata/MP 15.0. Since the different hypotheses correspond with the different methods, the most logical way to discuss the used methods is by addressing them for every set of hypotheses.

4.1 Hypothesis set 1: Correlation

During this study it will be investigated whether the expectations based on addressed literature are true; does higher investor sentiment bring higher volatility? For the first hypothesis it will be investigated whether higher investor sentiment is correlated with higher volatility. This will provide first insights about how investor sentiment and volatility move relative to each other. Furthermore, this correlation will be studied using both the data of the AAII-survey Bullish Individual Investors and the CBOE Put/Call Volume Ratio as proxies for investor sentiment. In the same way, both the Weekly St. Dev. of S&P500 Hourly Returns and the CBOE Volatility Index will be used as proxies for volatility. Last but not least, the control variables will be added to increase validity.

Since this study aims to investigate whether the potential effect of investor sentiment on volatility has changed as a result of the covid-19 pandemic, it is necessary to split the data in one way or another to compare results in the different conditions. To do this, three conditions were created that were incorporated in the dataset by adding a dummy variable for each condition. This is supposed to make it easier to test for changes in the investigated relationship. The first period *Before Covid* was used as some sort of control period where things are regarded as relatively 'normal' market conditions. This period ends on February 19th 2020; the day that the S&P500 start dropping with 34% (Jackwerth, 2020). This marks the beginning of the period that will be called *During Crash*. Although the steep slope downward ended on March 23rd 2020, an additional period is added in which volatility is still going down. The end of the crash will be marked at the first observation that falls in the 95% confidence interval of the remaining observations in terms of the first dependent variable; St. Dev. of real S&P500 (see Appendix 8.1.1 for explanation and tests). This results from May 21st 2020 being the first observation of the last period; *After Crash*. This resulted in the following conditions:

Before Covid: Observation 1 until 163; January 2nd 2017 until February 19th 2020

During Crash: Observation 164 until 176; February 20th 2020 until May 20th 2020

After Crash: Observation 177 until 209; May 21st 2020 until December 31st 2020

In order to test for the correlation of investor sentiment on volatility, the dummy variables will come in handy. In the regression equations they represented by *BC, DC* and *AC* respectively. By using interaction effects in the model, it is possible to analyze changes in the correlation in an easier way since they are presented in the same output. Because two proxies for investor sentiment are used and two proxies for volatility, it is possible to analyze the effect in four different regressions. Performing these regressions can allows to create one big table that provides the opportunity to potential effects and changes in a simple overview.

First, the weekly standard deviation of hourly S&P500 returns (*SDS&P*) will be regressed on the proportion bullish investors in the AAII survey (*PERBULLISH*). Secondly, the CBOE Volatility Index (*VIX*) will be regressed on the proportion bullish investors in the AAII survey. As third, the weekly standard deviation of hourly S&P500 returns will be regressed on the CBOE weekly average put/call volume ratio (*PCRATIO*). And lastly, the CBOE Volatility Index will be regressed on the CBOE weekly average put/call volume ratio. By adding both the changes in U.S. 10-Year Treasury Yield (***Δ****10YEARYIELD*) and the changes in USD/EUR exchange rate (***Δ****USD/EUR*) as control variables, the following regression equations were created:

(1) $SDS\&P_t = \beta_0 + \beta_1 * PERBULLISH_t * BC + \beta_2 * PERBULLISH_t * DC + \beta_3 * PERBULLISH_t * AC + \beta_4 * \Delta 10YEARYIELD_t + \beta_5 * \Delta USD/EUR_t + u_t$

(2) $VIX_t = \beta_0 + \beta_1 * PERBULLISH_t * BC + \beta_2 * PERBULLISH_t * DC + \beta_3 * PERBULLISH_t * AC + \beta_4 * \Delta 10YEARYIELD_t + \beta_5 * \Delta USD/EUR_t + u_t$

(3) $SDS\&P_t = \beta_0 + \beta_1 * PCRATIO_t * BC + \beta_2 * PCRATIO_t * DC + \beta_3 * PCRATIO_t * AC + \beta_4 * \Delta 10YEARYIELD_t + \beta_5 * \Delta USD/EUR_t + u_t$

(4) $VIX_t = \beta_0 + \beta_1 * PCRATIO_t * BC + \beta_2 * PCRATIO_t * DC + \beta_3 * PCRATIO_t * AC + \beta_4 * \Delta 10YEARYIELD_t + \beta_5 * \Delta USD/EUR_t + u_i$

After performing the regressions, β_1, β_2 and β_3 will denote the interaction effects of the proxies for investor sentiment and the specific period while β_0, β_4 and β_5 will denote the effect of the constant and control variables respectively. An effect is regarded as significant when the corresponding p-value is below the 0.05 significance level, which is formulated and proposed by Pesaran et al. (2001). After conducting the regressions and having presented the outcomes in a clear overview, the results will be discussed in order to accept/reject the hypotheses.

4.2 Hypothesis set 2: Standard ARDL Modelling

The first hypothesis aims to find correlation between the proxies for investor sentiment and volatility. Since potential correlation does not imply causation, it is important to use other research methods in order to find whether there is a predictive causal effect. Even when no correlation was found, using lags of both the independent variable and the dependent variable itself could shed new light on the relationship. This lagged relationship will be studied using both the data of the AAII-survey Bullish Individual Investors and the CBOE Put/Call Volume Ratio as proxies for investor sentiment as well. In the same way, both the Weekly St. Dev. of S&P500 Hourly Returns and the CBOE Volatility Index will be used as proxies for volatility. Lastly, the control variables will be added to increase validity.

The lags of both the independent and dependent variable will be added for two reasons. First of all, the usage of lags of the independent variable could imply more of a causal relationship. If an increase in investor sentiment in observation *t-1* leads to an increase in volatility in observation *t*, it is much more logical to suspect a causal relationship than when there is only correlation since the correlation could be caused by reverse causality. Secondly, the usage of lags of the dependent variables functions as some sort of control variable. If variation in volatility in observation *t* was significantly affected by the level of volatility in observation *t-1*, it is important to filter out this effect since it is not caused by increases or decreases in the independent variable. In order to test for these effects, an Autoregressive Distributed Lag (ARDL) Model is used since this model uses both lags of the independent and dependent variable as regressors (Greene, 2008). This model is suggested for time series economic modelling by Pesaran and Shin (1997).

In order to make use of an ARDL Model, some assumptions needed to be tested. The most important issues will shortly be touched upon. First of all, multicollinearity should be avoided since this could lead to disadvantageous effects on estimated coefficients in a multiple regression analysis (Mansfield & Helms, 1982). From the table in the Appendix 8.1.2 it can be concluded that there is no strong correlation between any variables used during *Before Covid* and *After Crash*, except for the weekly standard deviation of hourly S&P500 returns and the CBOE Volatility Index. But since these are both proxies for volatility and will not be used in the same regression, this will not influence results. The second assumption is stationarity of the data when integrated of order I(0) or I(1). By using Augmented Dickey-Fuller (ADF) Tests all data is tested for stationarity. As can be concluded from the output tables in Appendix 8.1.3, all data during *Before Covid* and *After Crash* appears to be stationary since all tests have a strong enough negative ADF-statistic and p-values of below the rejection level of 0.05 (Greene, 2008). Therefore, the ARDL-model will give valid estimates of relationships during these periods. Unfortunately, the data in period *During Crash* did not pass the test for both multicollinearity and stationarity, making in not useful to include this period in the analysis. Lastly, it is important to note that the period *After Crash* has 33 observations, which is just over the minimum of 30 supported by Agresti and Min (2002).

To make sure that the effects of investor sentiment on volatility are as valid as possible, the periods will be analyzed separately. This will allow for possible different lags for the control variables. Once again, the two proxies for investor sentiment and two proxies for volatility are used. Together with the fact that two separate periods will be used results in eight tests. These eight tests will be presented in two tables corresponding with the different periods. By comparing the different results in the tables, it can be examined whether the effect of investor sentiment on volatility has changed. Using the same names for the variables as in hypothesis 1, this leaves us with the following 4 regression equations that will be tested in two time periods:

(1) $SDS\&P_t = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-k}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

(2) $VIX_t = \beta_0 + \Sigma(\Psi_i * VIX_{t-k}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

(3) $SDS\&P_t = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-k}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

(4) $VIX_t = \beta_0 + \Sigma(\Psi_i * VIX_{t-k}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

After performing the regressions, Ψ_i will denote the autoregressive effects of the lags of the proxy for volatility. In the same way, $þ_i$ will denote the effect of (the lags of) the proxy for investor sentiment. The constant will be noted by β_0 and the effect of (the lags of) the control variables will be presented by $ø_i$ and $ð_i$. An effect is regarded as significant when the corresponding p-value is below the 0.05 significance level, as formulated and proposed by Pesaran et al. in 2001. After conducting the regressions and having presented the outcomes in a clear overview, the results will be discussed in order to accept/reject the hypotheses.

4.3 Hypothesis set 3: Predictive ARDL Modelling

Since hypothesis 2 already gives an insight whether there is some sort of lagged relationship that should imply causality. In hypothesis 3 it is examined whether investor sentiment has a predictive power for volatility. This predictive relationship will once again be studied using both the data of the AAII-survey Bullish Individual Investors and the CBOE Put/Call Volume Ratio as proxies for investor sentiment. Obviously, both the Weekly St. Dev. of S&P500 Hourly Returns and the CBOE Volatility Index will be used as proxies for volatility. Lastly, the control variables will be added to increase validity.

Just as for hypothesis 2, for hypothesis 3 an adjusted ARDL model will be used. The difference between the ARDL models in hypothesis 2 and the ARDL models used for hypothesis 3 is the fact that it that the latter aim to predict volatility in *t+1* instead of *t*. This means that only priorly known values are used to predict the next value of volatility. The ARDL will once again be using lags of both the dependent variable and the independent variable, and the control variables. Just as with hypothesis 2, the usage of lags of the dependent variables functions as some sort of control variable. If variation in volatility in observation *t+1* was significantly affected by the level of volatility in observation *t*, it is important to filter out this effect since it is not caused by increases or decreases in the independent variable.

To make sure that the effects of investor sentiment on volatility are as valid as possible, the periods will be analyzed separately again. This will allow for possible different lags for the control variables. Once again, the two proxies for investor sentiment and two proxies for volatility are used. Together with the fact that two separate periods will be used results in eight tests. These eight tests will be presented in two tables corresponding with the different periods. By comparing the different results in the tables, it can be examined whether the effect of investor sentiment on volatility has changed. Using the same names for the variables as in hypothesis 1, this leaves us with the following 4 regression equations that will be tested in two time periods:

(1) $SDS\&P_{t+1} = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-i}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

(2) $VIX_{t+1} = \beta_0 + \Sigma(\Psi_i * VIX_{t-i}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

(3) $SDS\&P_{t+1} = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-i}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

(4) $VIX_{t+1} = \beta_0 + \Sigma(\Psi_i * VIX_{t-i}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

After performing the regressions, Ψ_i will denote the autoregressive effects (of the lags) of the proxy for volatility on the next value of the volatility proxy. In the same way, $þ_i$ will denote the effect of (the lags of) the proxy for investor sentiment. The constant will be noted by β_0 and the effect of (the lags of) the control variables will be presented by $ø_i$ and $ð_i$. An effect is regarded as significant when the corresponding p-value is below the 0.05 significance level, as formulated and proposed by Pesaran et al. in 2001. After conducting the regressions and having presented the outcomes in a clear overview, the results will be discussed in order to accept/reject the hypotheses. Lastly, the period *After Crash* will be based on 32 observations, which is once again just over the minimum of 30 as supported by Agresti and Min (2002).

Chapter 5: Results

This chapter will discuss the results of the different OLS-regressions and ARDL-models. First the results per hypothesis will be presented in clear overviews and shortly discussed. Next, the results will be interpreted and discussed more in-depth. This will include conclusions about the hypotheses based on the results. Lastly, this chapter will be closed by a short overview about the acceptance and rejections of the hypothesis.

5.1 Results

5.1.1 Hypothesis set 1: Correlation

The results of the regressions in Table 5.1 on page 28 quantify the relationship between the proxies for investor sentiment and volatility. In general, the OLS regressions do find statistically significant correlations but the results differ for the proxies for investor sentiment. Furthermore, the control variables appear to be significant in every regression, whereas the constant is not significant in just one of the four performed regressions.

The statistically significant estimated coefficient of AAII-survey Bullish Individual Investors before the pandemic in the OLS equation (1) regression model is -0.365 ($p = 0.001$). This implies that if the proportion bullish individual investors increases with 1 percent point, the Weekly St. Dev. of S&P500 Hourly Returns decreases with 0.365. After the covid crash, this coefficient has changed to 0.093 ($p = 0.450$). This p-value implies that, using a 0.05 significance level, only the latter coefficient is not significant. Remarkably, it is found that while using the first to proxies for investor sentiment and volatility, a rather negative correlation is present instead of the expected positive correlation. Secondly, there is no significant effect anymore after the covid crash.

In OLS equation (2) the statistically significant estimated coefficient of the AAII-survey Bullish Individual Investors before the pandemic is -27.953 ($p = 0.000$). This implies that if the proportion bullish individual investors increases with 1 percent point, the CBOE Volatility Index decreases with 27.953. After the covid crash, this coefficient has changed to 4.532 ($p = 0.404$). This p-value implies that, using a 0.05 significance level, only the latter coefficient is not significant. Remarkably, it is found that while using the AAII-Survey and VIX for investor sentiment and volatility respectively, a rather negative correlation is present instead of the expected positive correlation. Secondly, just as with OLS equation (1) there is no significant effect anymore after the covid crash.

In the third OLS equation (3) the statistically significant estimated coefficient of the S&P500 Put/Call Volume Ratio before the pandemic is 0.032 (*p = 0.398*). This implies that if the put/call volume ratio increases with 1, the Weekly St. Dev. of S&P500 Hourly Returns increases with 0.032. After the covid crash, this coefficient has changed to 0.092 (*p = 0.027*). This p-value implies that, using a 0.05 significance level, only the latter coefficient is significant. In contrast to the AAII-survey, this significant result is positive, as suspected. Secondly, just as with OLS equations (1) & (2) the correlation seems to have changed; before covid, there was no significant correlation, but after the crash there is.

In the last OLS equation (4) for hypothesis 1, the statistically significant estimated coefficient of the S&P500 Put/Call Volume Ratio before the pandemic is 0.820 (*p = 0.626*). This implies that if the put/call volume ratio increases with 1, the CBOE Volatility Index increases with 0.820. After the covid crash, this coefficient has changed to 6.349 (*p = 0.001*). This p-value implies that, using a 0.05 significance level, only the latter coefficient is significant. Once again, in contrast to the AAII-survey, this significant result is positive, as suspected. Secondly, just as with OLS equations (1), (2) and (3), the correlation seems to have changed; before covid, there was no significant correlation, but after the crash there is.

Table 5.1: Regression estimates of the relationship between investor sentiment and volatility

Superscripts *, **, *** represent statistical significance at the 10%, 5% and 1% level, respectively. Standard errors are reported in parentheses below the coefficient estimates. For the full regression results Appendix 8.2.1 can be consulted.

	(1)	(2)	(3)	(4)
	Coeff.	Coeff.	Coeff.	Coeff.
β_1	-0.365*** (0.111)	-27.953*** (4.886)	0.032 (0.038)	-0.820 (1.679)
β_2	0.838*** (0.164)	44.795*** (7.203)	0.262*** (0.044)	13.241*** (1.974)
β_3	-0.093 (0.123)	4.532 (5.424)	0.092*** (0.041)	6.349*** (1.848)
β_4	0.722*** (0.119)	21.434*** (5.247)	0.700*** (0.125)	22.055*** (5.597)
β_5	3.624*** (1.170)	148.434*** (51.470)	4.405*** (1.188)	175.253*** (53.095)
β_0	0.199*** (0.041)	21.578*** (1.793)	0.0119 (0.067)	12.900*** (3.000)
N	209	209	209	209
R-squared	0.6031	0.7265	0.578	0.699
Adj. R-squared	0.5933	0.7198	0.567	0.692
F-statistic	61.69***	107.84***	55.50***	94.43***
Root MSE	0.111	4.893	0.115	5.130

These results were found using the following regression equations:

(1) $SDS\&P_t = \beta_0 + \beta_1 * PERBULLISH_t * BC + \beta_2 * PERBULLISH_t * DC + \beta_3 * PERBULLISH_t * AC + \beta_4 * \Delta 10YEARYIELD_t + \beta_5 * \Delta USD/EUR_t + u_t$

(2) $VIX_t = \beta_0 + \beta_1 * PERBULLISH_t * BC + \beta_2 * PERBULLISH_t * DC + \beta_3 * PERBULLISH_t * AC + \beta_4 * \Delta 10YEARYIELD_t + \beta_5 * \Delta USD/EUR_t + u_t$

(3) $SDS\&P_t = \beta_0 + \beta_1 * PCRATIO_t * BC + \beta_2 * PCRATIO_t * DC + \beta_3 * PCRATIO_t * AC + \beta_4 * \Delta 10YEARYIELD_t + \beta_5 * \Delta USD/EUR_t + u_t$

(4) $VIX_t = \beta_0 + \beta_1 * PCRATIO_t * BC + \beta_2 * PCRATIO_t * DC + \beta_3 * PCRATIO_t * AC + \beta_4 * \Delta 10YEARYIELD_t + \beta_5 * \Delta USD/EUR_t + u_i$

5.1.2 Hypothesis set 2: Standard ARDL-Modelling

Before Covid

The results of the ARDL-Models in Table 5.2 on page 31 quantify the relationship between the proxies for investor sentiment and volatility when adding lags of both the dependent and independent variable before Covid. In general, the ARDL-Models do find statistically significant relationships between (lags of) investor sentiment and volatility. The results differ per proxy for investor sentiment and per lag. One thing that is valid for every model is that the coefficient for the lag of the dependent variable is significantly positive at a significance level of 5%. Furthermore, the control variables appear to be insignificant in every model, whereas the constant is significant in just one of the four models.

In ARDL-Model (1) of table 5.2 the coefficient of lagged value of dependent variable is *0.635* ($p = 0.000$). This means that if the previous value of the Weekly St. Dev. of S&P500 Hourly Returns has a gone up with 1, the current value will go up with *0.635*. When looking at the coefficients for the AAII-survey Bullish Individual Investors, it turns out that the coefficients for the current value, the first lag, the second lag and the third lag are *-0.384* ($p = 0.000$), *0.315* ($p = 0.005$), *-0.203* ($p = 0.058$) and *0.333* ($p = 0.000$) respectively. Using a 5%-significance level, only the third coefficient is not significant. Remarkably, the sign of the coefficient's changes for each lag.

In the next ARDL-Model (2) the coefficient of lagged value of dependent variable is *0.762* ($p = 0.000$). This means that if the previous value of the CBOE Volatility Index has a gone up with 1, the current value will go up with *0.762*. When looking at the coefficients for the AAII-survey Bullish Individual Investors, it turns out that no coefficient is significant, using a 5%-significance level. The only coefficient is *-2.833* with a *p*-value of *0.365*. Remarkably, this is a big contrast with the previous model while both predict volatility. Furthermore, it is observable that once again the control variables are not significant, and the constant is significant with a value of *4.815* (p = 0.002).

The third ARDL-Model (3) uses the Put/Call Volume Ratio as proxy for investor sentiment. The coefficient of lagged value of dependent variable is *0.696* ($p = 0.000$). This means that if the previous value of the Weekly St. Dev. of S&P500 Hourly Returns has a gone up with 1, the current value will go up with *0.696*. When looking at the coefficients for the S&P500 Put/Call Volume Ratio, it turns out that the coefficients for the current value and the first lag are *0.120* ($p = 0.000$) and *-0.070* ($p = 0.005$) respectively. Using a 5%-significance level, both the current value and the first lag are significant. Note that current value of the proxy for investor sentiment has a positive coefficient where this coefficient was negative when the percentage bullish individual investors was used. Lastly, it turns out that neither the control variables nor the constant is significant.

Also, the fourth ARDL-Model (4) uses the Put/Call Volume Ratio as proxy for investor sentiment. The coefficient of lagged value of dependent variable is *0.839* (*p =0.000*). This means that if the previous value of the CBOE Volatility Index has a gone up with 1, the current value will go up with *0.839*. When looking at the coefficients for the S&P500 Put/Call Volume Ratio, it turns out that the coefficients for the current value, the first lag and the second lag are *5.490* (*p = 0.000)*, *2.980* (*p = 0.001)* and *-1.903* (*p = 0.022*) respectively. Using a 5%-significance level, all three these coefficients are significant. Note that current value of the proxy for investor sentiment has a positive coefficient where this coefficient was negative when the percentage bullish individual investors was used. Lastly, it turns out that neither the control variables nor the constant is significant.

Table 5.2: Standard ARDL-Model Outputs Before Covid

Superscripts *, **, *** represent statistical significance at the 10%, 5% and 1% level, respectively. Standard errors are reported in parentheses below the coefficient estimates. For the full regression results Appendix 8.2.2 can be consulted.

	(1)	(2)	(3)	(4)
Before Covid	Coeff.	Coeff.	Coeff.	Coeff.
Volatility Pr.				
L1	0.635*** (0.060)	0.762*** (0.052)	0.696*** (0.059)	0.839*** (0.045)
Investor Sentiment Pr.				
--	-0.384*** (0.097)	-2.833 (3.122)	0.120*** (0.025)	5.490*** (0.852)
L1	0.315*** (0.111)		-0.070*** (0.024)	-2.980*** (0.858)
L2	-0.203* (0.107)			-1.903** (0.824)
L3	0.333*** (0.093)			
$\Delta 10YEARYIELD$	0.050 (0.100)	1.001 (3.818)	0.089 (0.099)	-0.045 (3.380)
$\Delta USD/EUR$	-1.004 (0.932)	-62.096* (35.371)	-1.363 (0.913)	-29.254 (31.423)
Constant	0.037 (0.040)	4.815*** (1.522)	-0.038 (0.049)	1.473 (1.910)
N	159	159	159	159
R-squared	0.515	0.626	0.512	0.713
Adj. R-squared	0.492	0.617	0.500	0.702
F-statistic	22.90***	64.55***	32.13***	63.06***
Root MSE	0.064	2.454	0.063	2.163

These results were found using the following regression equations:

(1) $SDS\&P_t = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-k}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

(2) $VIX_t = \beta_0 + \Sigma(\Psi_i * VIX_{t-k}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

(3) $SDS\&P_t = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-k}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

(4) $VIX_t = \beta_0 + \Sigma(\Psi_i * VIX_{t-k}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

After Crash

The results of the ARDL-Models in Table 5.3 on page 34 should present a potential relationship between the proxies for investor sentiment and volatility when adding lags of both the dependent and independent variable after the Covid crash of 2020. Remarkably, for none of the models, a significant coefficient was found for the proxies for investor sentiment. Only the first lag of the dependent variable and changes in the 10-year treasury yield seem to have a consistent significant coefficient. For the second control variable, changes in US Dollar/Euro exchange rate, and the constant only one of the models produces a significant coefficient when using a 5% significance level.

In ARDL-Model (1) of table 5.3 the coefficient of lagged value of dependent variable is *0.395* (*p = 0.011*). This means that if the previous value of the Weekly St. Dev. of S&P500 Hourly Returns has a gone up with 1, the current value will go up with *0.395*. When looking at the coefficients for the AAII-survey Bullish Individual Investors, it turns out that neither coefficients for the current value, nor for lagged values are significant. Stata only gives the coefficient for the current value, which is *-0.147* (*p = 0.*295). Since this value is not significant using a 5%-significance level, it is assumed that there are no lags of the AAII-survey Bullish Individual Investors that hold a significant coefficient. Furthermore, it is found that both current values of the control variables are significant at a 5%-significance level while the constant is significant at 10%-significant level.

In the next ARDL-Model (2) the coefficient of lagged value of dependent variable is *0.364* (*p = 0.024*). This means that if the previous value of the CBOE Volatility Index has a gone up with 1, the current value will go up with *0.364*. When looking at the coefficients for the AAII-survey Bullish Individual Investors, it turns out that no coefficient is significant, using a 5%-significance level. The only coefficient given by Stata is *-11.124* with a *p*-value of *0.115*. When looking at the control variables, it is found that the changes in USD/EUR Exchange Rate are not significant while changes in the 10-year Treasury Yield are significant at a 1%-significance level. Lastly, the constant is *18.235* with *p*-value of *0.003*.

The third ARDL-Model (3) of table 5.3 uses the Put/Call Volume Ratio as proxy for investor sentiment. The coefficient of lagged value of dependent variable is *0.427* (*p =0.006*). This means that if the previous value of the Weekly St. Dev. of S&P500 Hourly Returns has a gone up with 1, the current value will go up with *0.427*. When looking at the coefficients for the S&P500 Put/Call Volume Ratio, it turns out that once again no significant coefficients were found. The only coefficient given is that of the current value, which is *-0.005 (p = 0.956)*. Using a 5%-significance level, this coefficient is not significant and since this is the only coefficient given, it is assumed that there are no significant coefficients for lags of the S&P500 Put/Call Volume Ratio. Furthermore, it turns out that the constant is not significant at a 5%-significance level. And lastly, when looking at the control variables, only changes in the 10-year Treasury Yield have a significant coefficient when using a 5%-significance level.

Lastly, the fourth ARDL-Model (4) also uses the Put/Call Volume Ratio as proxy for investor sentiment. The coefficient of lagged value of dependent variable is *0.464* (*p =0.002*). This means that if the previous value of the CBOE Volatility Index has a gone up with 1, the current value will go up with *0.464*. When looking at the coefficients for the S&P500 Put/Call Volume Ratio, it turns out that also in this last model, no significant coefficients were found. The coefficient for the current value is *-6.219* with a p-value of *0.*192. Using a 5%-significance level, this is of course not a significant value. Since this is the only coefficient provided, it is assumed that there are also no significant coefficients for the lagged values. Lastly, it turns out that the changes in the 10-year Treasury Yield and the constant have a significant coefficient when using a 5%-significance level while the coefficient for the changes in the USD/EUR Exchange Rate is not significant.

Table 5.3: Standard ARDL-Model Outputs After Crash

Superscripts *, **, *** represent statistical significance at the 10%, 5% and 1% level, respectively. Standard errors are reported in parentheses below the coefficient estimates. For the full regression results Appendix 8.2.2 can be consulted.

	(1)	(2)	(3)	(4)
After Crash	Coeff.	Coeff.	Coeff.	Coeff.
Volatility Pr.				
L1	0.395** (0.144)	0.364** (0.152)	0.427*** (0.144)	0.464*** (0.134)
Investor Sentiment Pr.				
--	-0.147 (0.137)	-11.124 (7.602)	-0.005 (0.095)	-6.219 (4.648)
$\Delta 10YEARYIELD$	0.498** (0.215)	32.339*** (10.604)	0.516** (0.222)	35.419*** (10.831)
$\Delta USD/EUR$	3.659** (1.752)	44.727 (86.247)	3.398* (1.797)	6.510 (87.000)
Constant	0.133* (0.066)	18.235*** (5.711)	0.085 (0.162)	22.202** (8.615)
N	33	33	33	33
R-squared	0.465	0.496	0.443	0.490
Adj. R-squared	0.388	0.424	0.364	0.417
F-statistic	6.08***	6.89***	5.57***	6.73***
Root MSE	0.068	3.365	0.070	3.385

These results were found using the following regression equations:

(1) $SDS\&P_t = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-k}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

(2) $VIX_t = \beta_0 + \Sigma(\Psi_i * VIX_{t-k}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

(3) $SDS\&P_t = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-k}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

(4) $VIX_t = \beta_0 + \Sigma(\Psi_i * VIX_{t-k}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

5.1.3 Hypothesis set 3: Predictive ARDL-Modelling

Before Covid

The results of the ARDL-Models in Table 5.4 on page 37 quantify the predictive relationship between the proxies for investor sentiment and volatility before covid. In general, the ARDL-Models do not find statistically significant predictive relationships between investor sentiment and volatility. One thing that is valid for every model is that the coefficient for the lag of the dependent variable is significantly positive at a significance level of 5%. Furthermore, the control variables appear to be insignificant in every model, whereas the constant is significant in just one of the four models.

In the first model (1) of table 5.4 the coefficient of current value of dependent variable is *0.675* ($p = 0.000$). This means that if the current value of the Weekly St. Dev. of S&P500 Hourly Returns has a gone up with 1, the next value will go up with *0.675*. When looking at the coefficients for the AAII-survey Bullish Individual Investors, it turns out that no significant coefficients could be found. Stata only gives the coefficient for the current value, which is *0.089* ($p = 0.308$). Since this value is not significant using a 5%-significance level, it is assumed that there are no lags of the AAII-survey Bullish Individual Investors that hold a significant coefficient. Furthermore, it is found that both the current values of the control variables and the constant are insignificant at a 5%-significance level.

In the second predictive ARDL-Model (2) of table 5.4 the coefficient of the current value of the dependent variable is *0.808* ($p = 0.000$). This means that if the current value of the CBOE Volatility Index has a gone up with 1, the next value will go up with *0.808*. When looking at the coefficients for the AAII-survey Bullish Individual Investors, it turns out that no significant coefficients could be found in this case as well. Stata only provides the coefficient for the current value, which is *3.937* ($p = 0.212$). Since this value is not significant using a 5%-significance level, it is assumed that there are no lags of the AAII-survey Bullish Individual Investors that hold a significant coefficient. Lastly, it is found that both the current values of the control variables and the constant are insignificant at a 5%-significance level.

In the next model (3) the coefficient of the current value of the dependent variable is *0.667* ($p = 0.000$). This means that if the current value of the Weekly St. Dev. of S&P500 Hourly Returns has a gone up with 1, the next value will go up with *0.667*. When looking at the coefficients for the S&P500 Put/Call Volume Ratio, it turns out that no significant coefficients could be found. Only the coefficient for the current value is given, which is *-0.019* ($p = 0.453$). Since this value is not significant using a 5%-significance level, it is assumed that there are no lags of the S&P500 Put/Call Volume Ratio that hold a significant coefficient. Furthermore, it turns out that both the current values of the control variables and the constant are insignificant at a 5%-significance level.

In the last predictive model (4) for the period *before covid* that is used, the coefficient of the current value of the dependent variable is *0.802* (*p = 0.000*). This means that if the current value of the CBOE Volatility Index has a gone up with 1, the next value will go up with *0.802*. When looking at the coefficients for the S&P500 Put/Call Volume Ratio, it turns out that no significant coefficients could be found. Only the coefficient for the current value is provided, which is *-1.614* (*p = 0*.0.71). Since this value is not significant using a 5%-significance level, it is assumed that there are no lags of the S&P500 Put/Call Volume Ratio that hold a significant coefficient. Furthermore, it turns out that both the current values of the control variables are not significant at a 5%-significance level. Lastly, the constant is *5.528* and turns out to be significant at a 5%-significance level with a *p*-value of *0.001*.

Table 5.4: Predictive ARDL-Model Outputs Before Covid

Superscripts *, **, *** represent statistical significance at the 10%, 5% and 1% level, respectively. Standard errors are reported in parentheses below the coefficient estimates. For the full regression results Appendix 8.2.3 can be consulted.

	(1)	(2)	(3)	(4)
After Crash	Coeff.	Coeff.	Coeff.	Coeff.
$Volatility\ Pr.$				
--	0.675***	0.808***	0.667***	0.802***
	(0.065)	(0.052)	(0.064)	(0.051)
$Investor\ Sentiment\ Pr.$				
--	0.089	3.937	-0.019	-1.614*
	(0.087)	(3.141)	(0.025)	(0.887)
$\Delta 10YEARYIELD$				
--	-0.043	-1.239	-0.489	-1.150
	(0.107)	(3.860)	(0.107)	(3.826)
$\Delta USD/EUR$				
--	1.485	26.553	1.525	25.170
	(1.000)	(35.994)	(0.998)	(35.758)
$Constant$	0.008	1.249	0.073	5.528***
	(0.036)	(1.545)	(0.044)	(1.681)
N	158	158	158	158
R-squared	0.425	0.620	0.423	0.624
Adj. R-squared	0.410	0.610	0.408	0.615
F-statistic	28.28***	62.42***	28.07***	63.56***
Root MSE	0.069	2.477	0.069	2.463

These results were found using the following regression equations:

(1) $SDS\&P_{t+1} = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-i}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

(2) $VIX_{t+1} = \beta_0 + \Sigma(\Psi_i * VIX_{t-i}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

(3) $SDS\&P_{t+1} = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-i}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

(4) $VIX_{t+1} = \beta_0 + \Sigma(\Psi_i * VIX_{t-i}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

After Crash

The results of the ARDL-Models in Table 5.5 on page 40 present the predictive relationship during the covid recovery in 2020 between the proxies for investor sentiment and volatility. In general, the ARDL-Models do not find consistent statistically significant predictive relationships between investor sentiment and volatility. Although all first lags of the proxies for investor sentiment are negative, none of these is statistically significant. When analyzing the different models, there is also no pattern that can be found between the models. Furthermore, the control variables appear to be insignificant in every model, whereas the constant is significant in all the four models. Lastly, it is remarkable that the r-squared statistics are relatively low, especially for the models using the AAII-Survey as the proxy for investor sentiment.

In the first predictive ARDL-Model (1) of table 5.5 the coefficient of current value of dependent variable is *0.437* (*p = 0.045*) and is significant when using a 5%-significance level. This means that if the current value of the Weekly St. Dev. of S&P500 Hourly Returns has a gone up with 1, the next value will go up with *0.437*. When looking at the coefficients for the AAII-survey Bullish Individual Investors, it turns out that no significant coefficients could be found. The only coefficient given is that for the current value, which is *-0.123* (*p = 0.*480). Since this value is not significant using a 5%-significance level, it is assumed that there are no lags of the AAII-survey Bullish Individual Investors that hold a significant coefficient. Furthermore, it is found that both the current values of the control variables are insignificant at a 5%-significance level. Lastly, the constant is 0.166 with a *p*-value of *0.044* which is significant at a 5%-significance level.

In the next model (2) of table 5.5 the coefficient of the current value of the dependent variable is *0.808* (*p = 0.077*). Since the *p*-value is below the 5%-significance level, this coefficient is considered to be insignificant. When looking at the coefficients for the AAII-survey Bullish Individual Investors, it turns out that no significant coefficients could be found in this case as well. Stata only provides the coefficient for the current value, which is *-5.780* (*p = 0.*535). Since this value is not significant using a 5%-significance level, it is assumed that there are no lags of the AAII-survey Bullish Individual Investors that hold a significant coefficient. It is also found that both the current values of the control variables are insignificant at a 5%-significance level. The constant turns out to be *17.149* with a *p*-value of *0.020* and is therefore considered significant.

In the third model (3) the coefficient of the current value of the dependent variable is *0.667* (*p = 0.037*). This means that if the current value of the Weekly St. Dev. of S&P500

Hourly Returns has a gone up with 1, the next value will go up with *0.379*. When looking at the coefficients for the S&P500 Put/Call Volume Ratio, it turns out that the coefficient for the current value is not significant. Remarkably, the coefficient for the first lag is *0.321* and turns out to be significant ($p = 0.005$). The control variables used seem to be insignificant while the constant is 0.879 and significant using a 5%-significance level ($p = 0.001$).

In the last predictive ARDL-Model (4) for the period *after covid* that is used, the coefficient of the current value of the dependent variable is *0.280* ($p = 0.119$). Since the *p*-value is below the 5%-significance level, this coefficient is considered to be insignificant. When looking at the coefficients for the S&P500 Put/Call Volume Ratio, a remarkable pattern can be observed. The coefficients for the current value and the first two lags are provided where only the second lag has a significant coefficient. This coefficient is *-13.206* with a p-value of *0.040*. Furthermore, it turns out that both the current values of the control variables are not significant at a 5%-significance level. And lastly, the constant is *61.460* and turns out to be significant at a 5%-significance level with a *p*-value of *0.001*.

Table 5.5: Predictive ARDL-Model Outputs After Covid

Superscripts *, **, *** represent statistical significance at the 10%, 5% and 1% level, respectively. Standard errors are reported in parentheses below the coefficient estimates. For the full regression results Appendix 8.2.3 can be consulted.

	(1)	(2)	(3)	(4)
After Crash	Coeff.	Coeff.	Coeff.	Coeff.
Volatility Pr.				
--	0.437** (0.065)	0.379* (0.206)	0.379** (0.173)	0.280 (0.173)
Investor Sentiment Pr.				
--	-0.123 (0.172)	-5.780 (9.206)	-0.113 (0.098)	-4.207 (5.311)
L1			0.321*** (0.105)	-8.066 (5.434)
L2				-13.206** (6.092)
$\Delta 10YEARYIELD$				
--	0.133 (0.281)	14.570 (14.473)	0.117 (0.243)	6.748 (14.098)
$\Delta USD/EUR$				
--	0.274 (2.300)	20.696 (107.109)	-2.352 (1.990)	-118.534 (102.308)
Constant	0.166** (0.078)	17.149** (1.545)	0.879*** (0.224)	61.410*** (16.066)
N	32	32	32	32
R-squared	0.251	0.301	0.487	0.488
Adj. R-squared	0.140	0.197	0.389	0.365
F-statistic	2.26***	2.90**	4.94***	3.97***
Root MSE	0.082	4.024	0.069	3.578

These results were found using the following regression equations:

(1) $SDS\&P_{t+1} = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-i}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

(2) $VIX_{t+1} = \beta_0 + \Sigma(\Psi_i * VIX_{t-i}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

(3) $SDS\&P_{t+1} = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-i}) + \Sigma(\text{þ}_i * PCRATIO_{t-i}) + \Sigma(\text{ø}_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(\text{ð}_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

(4) $VIX_{t+1} = \beta_0 + \Sigma(\Psi_i * VIX_{t-i}) + \Sigma(\text{þ}_i * PCRATIO_{t-i}) + \Sigma(\text{ø}_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(\text{ð}_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

5.2 Discussion of results

5.2.1 Discussion hypothesis set 1: Correlation

Hypothesis 1A

The first set of hypotheses focused on the expected positive correlation between investor sentiment and volatility. Hypothesis 1A stated that investor sentiment is positively correlated with volatility before the covid-crash. This hypothesis was tested using four OLS-regressions with interaction effects. The results of these analyzes show that the hypothesis can be rejected. In the first two regressions, the proportion bullish individual investors in the AAII-survey was used. This resulted in a significant negative correlation for both the proxies for volatility. These findings are in line with the research conducted by Lee, Jiang and Indro (2002). Moreover, the last two regressions, using the put/call volume ratio of S&P500 options gave very small, non-significant coefficients. Neither of the proxies for investor sentiment is positively correlated with volatility, resulting in a rejection of hypothesis 1A.

Hypothesis 1B

After looking at the period before covid, the covid recovery during 2020 was examined. Hypothesis 1B stated that investor sentiment was positively correlated with volatility during the covid recovery until 2021. This hypothesis was tested using the same four OLS-regressions with interaction effects as in hypothesis 1A. Following the results of the regressions, hypothesis 1B should be rejected. Using the proportion bullish individual investors in the AAII-survey there was no significant result founded. Therefore, using this proxy for investor sentiment, the hypothesis would be rejected. When using the put/call volume ratio for S&P500 options as a proxy for investor sentiment, a significant positive correlation is found. Since a high put/call volume ratio is an indicator for low investor sentiment, the correlation that was hypothesized should give a negative coefficient. In other words, the correlation that is found induces a negative correlation between investor sentiment and volatility. As a result, hypothesis 1B is rejected using the put/call ratio as well.

Hypothesis 1C

The last hypothesis of this set stated that the correlation between investor sentiment and volatility has changed significantly during the covid recovery. This is done by looking at the results of the same regressions as the previous hypotheses. When comparing the different coefficients in the four regressions it is very obvious that in each regression the coefficient changed significantly. In the first two regressions, that used the proportion bullish individual investors in the AAII-survey, the coefficient changed from significant negative to insignificant. Therefore, there was found a negative correlation first and during the covid recovery there was no correlation at all. In the last two regressions, that used the put/call volume ratio of S&P500 options, the exact opposite was true. Before, there was no significant correlation and during the recovery in 2020, there was a significant positive coefficient. This significant positive coefficient means, of course, that when investor sentiment goes up, volatility decreases. Following these results, hypothesis 1C is accepted.

5.2.2 Discussion hypothesis set 2: Standard ARDL-Modelling

Hypothesis 2A

In the second set of hypotheses, lagged values were added to the model to create an ARDL-model. In this set of hypotheses, the first hypotheses stated that when lagged values of both the dependent and independent variable are added to the model, investor sentiment has a significant positive effect on volatility before covid. This hypothesis was tested using four ARDL-Models that used different combinations for the proxies for investor sentiment and volatility. When analyzing the results of the four models together, it is not possible to find a certain positive or negative relationship. Therefore, the hypothesis could be rejected. In the first model, the current value as well as the first three lagged values had significant coefficients. However, the signs of these coefficients are not all the same. Since none of the coefficients stands out it in a matter of size, no conclusions about a particular relationship can be drawn from this model. The second model does not even have a significant coefficient for the proxy for investor sentiment. Therefore, this model does not give any more information about a potential significant relationship as well. Since these models both don’t seem very informative and they both differ quite a bit, no significant pattern is found. As a result, hypothesis 2A would be rejected based on the models that use the data of the AAII-survey.

Where the first two models do not provide any information about a relationship, the last two seem to give an analogous idea of the relationship. For both models, the current values of investor sentiment have a positive coefficient that is strongly significant. Moreover, the coefficient for the first lag of the proxy for investor sentiment is significantly negative. Even more remarkable is that the sizes of the coefficient of the lagged value in comparison with the coefficient of the current value are almost the same in both the models (58.3% for the third model vs. 54.3% for the fourth model). Following the pattern found in the last two models, it might be said that a relationship can be found when using the put/call volume ratio of S&P500 options. Recall that a coefficient needs to be negative instead of positive for the put/call volume ratio. Since this pattern involves a positive coefficient that is stronger than the negative coefficient, the relationship between investor sentiment and volatility is considered to be negative. As a result, the hypothesis is rejected.

Hypothesis 2B

The second hypothesis in the second set of hypotheses stated that stated that when lagged values of both the dependent and independent variable are added to the model, investor sentiment has a significant positive effect on volatility during the covid recovery in 2020. This hypothesis uses another time period then previous hypothesis but was tested using the same four ARDL-Models. When analyzing the results of these four models, it is certainly not possible to find a certain positive or negative relationship. Neither of the models find a significant coefficient for the used proxies for investor sentiment. Therefore, hypothesis 2B is rejected.

Hypothesis 2C

The last hypothesis of this set stated that the relationship between investor sentiment and volatility has changed significantly during the covid recovery. This is done by looking at the outputs and discussions for hypotheses 2A and 2B. When comparing these outcomes and discussions it becomes clear that it might come in handy to separate the discussion by proxies for investor sentiment since the results differ for both the proxies. First, we focus on the ARDL-Models that used the AAII-survey as the proxy for investor sentiment. In the period before covid, the models do not show an analogous pattern and therefore no certain relationship is assumed. In the second time span, during the covid recovery in 2020, both models did not give one single significant coefficient for investor sentiment. Therefore, it assumed that there is no relationship between the percentage bullish investors in the AAII-survey and volatility present during the covid recovery as well. Since no relationship is concluded in both the time periods, the relationship has not changed as well. As a result, the hypothesis is rejected when using the AAII-survey as a proxy for investor sentiment.

When looking at the ARDL-Models that use the put/call ratio for S&P500 options a different story is found. The two ARDL-Models that were performed for the period before covid show a remarkable resembling pattern were the value of volatility increases when investor sentiment decreases but increases when the first lag of investor sentiment increases. Since the positive coefficient is stronger, a negative relationship was concluded. This may sound counterintuitive but recall that the put/call ratio should go down when investor sentiment increases. When focusing on the period that presents the covid recovery in 2020, no relationship was found. Both ARDL-Models do not provide any significant coefficients for the put/call volume ratio of S&P500 options. As a result, no relationship was concluded during the covid recovery in 2020. Since there was a relationship concluded before covid but not during the covid recovery, the hypothesis is accepted when using the put/call volume ratio of S&P500 options as a proxy for investor sentiment.

5.2.3 Discussion hypothesis set 3: Predictive ARDL-Modelling

Hypothesis 3A

In the last set of hypotheses, ARDL-Models were used to predict the next value of the proxies for investor sentiment. In this way, a potential predictive pattern could be observed, and it would be observable whether investor sentiment could be a good predictor for volatility. The first hypothesis in this set of hypotheses stated that when only lagged values were used in the model, investor sentiment positively significantly predicts volatility before covid. This hypothesis was tested using ARDL-Models that used varying combinations of proxies for investor sentiment and volatility. When analyzing the outputs for the models, it becomes clear that this hypothesis should be rejected. Not one of the models find a statistically significant coefficient for any lags of the proxies for investor sentiment. As a result, hypothesis 3A is rejected.

Hypothesis 3B

The second hypothesis in this set of hypotheses stated that when only lagged values were used in the model, investor sentiment positively significantly predicts volatility during the covid recovery in 2020. This hypothesis was tested as well using ARDL-Models that used varying combinations of proxies for investor sentiment and volatility. The results of these models are not consistent, which makes it harder to form a conclusion about the hypothesis. Once again, the results differ quite much when using other proxies for investor sentiment. Using the AAII-Survey as a proxy for investor sentiment no patterns was found in either of the models. The first two models do not find any statistically significant coefficients for investor sentiment. Therefore, the conclusion based on the AAII-Survey is clear; the hypothesis should be rejected. Unfortunately, a conclusion based on the Put/Call Volume Ratio for S&P500 options seems harder. Both models do find some statistically significant coefficients. But both the models give totally different coefficients. The third models finds that the lagged value for the Put/Call Volume Ratio positively influences the future value of volatility while the fourth model finds that the second lagged value for the Put/Call Ratio negatively influences the future value of volatility. Since no consistent significant pattern can be concluded from these outcomes and the R squared value of the models is not even 0.5, the hypothesis that investor sentiment is a positive predictor for volatility is rejected.

Hypothesis 3C

The last hypothesis stated that the predictability of volatility based on investor sentiment has changed during the covid recovery in 2020 compared with before covid. This is done by looking at the results and conclusions for the previous hypotheses as well. Hypothesis 3A focused on the data and situation before covid. From the performed ARDL-Models, it was concluded that investor sentiment had no predicting power on volatility. For the ARDL-Models that were performed on the data from during the covid recovery, approximately the same was true. Using the AAII-Survey, it was clear that no predicting power was present. Although the inconsistent but significant appearing outcomes from the last two models made it a bit more complicated, no significant predicting power of investor sentiment on volatility could be concluded as well. In both the period before covid, as well as in the period during the covid recovery in 2020, there was no predictability of volatility based on investor sentiment. Since no changes in this predictability could be found, hypothesis 3C is rejected.

5.2.4 Overview hypotheses and discussions

In the overview in table 5.6, the hypotheses and their conclusions can be seen in a simple and clear way. While analyzing and discussing the outputs and results of the performed regressions, it became clear that the conclusions for the hypotheses may be very dependent on what proxy to use for investor sentiment. Since the conclusions may be dependent on the proxy for investor sentiment, they are separated in the table below.

Table 5.6: Table with Overview of Hypotheses and their Conclusions

This table shows in a simple way whether each hypothesis is accepted or rejected, depending on what proxy for investor sentiment is used. Explanations for the conclusions concerning the hypotheses are given in the previous section.

Hypothesis	**AAII-Survey**	**Put/Call Volume Ratio**
H1A: *Investor sentiment is positively correlated with volatility before the covid-crash*	**Rejected**	**Rejected**
H1B: *Investor sentiment is positively correlated with volatility during the covid recovery until 2021*	**Rejected**	**Rejected**
H1C: *The correlation between investor sentiment and volatility has significantly changed during the covid recovery until 2021 compared with before*	**Accepted**	**Accepted**
H2A: *When adding lagged values of both the investor sentiment and volatility, investor sentiment has a significant positive relationship with volatility before the covid-crash*	**Rejected**	**Rejected**
H2B: *When adding lagged values of both the investor sentiment and volatility, investor sentiment has a significant positive relationship with volatility during the covid recovery until 2021*	**Rejected**	**Rejected**

H2C: *The relationship between investor sentiment and volatility has significantly changed during the covid recovery until 2021 compared with before*	**Rejected**	**Accepted**
H3A: *When using only lagged values of both the investor sentiment and volatility, investor sentiment positively significantly predict volatility before the covid-crash*	**Rejected**	**Rejected**
H3B: *When using only lagged values of both the investor sentiment and volatility, investor sentiment positively significantly predict volatility during the covid recovery until 2021*	**Rejected**	**Rejected**
H3C: *The predictability of volatility based on investor sentiment has changed during the covid recovery until 2021 compared with before*	**Rejected**	**Rejected**

Chapter 6: Conclusion

In this research, the potential effect of investor sentiment on volatility was investigated. Moreover, potential changes in this effect as a result of the Covid-19 related stock market crash were explored. As academic literature suggested, a positive effect was hypothesized. In order to investigate this potential effect, several econometric models were used. These included interaction effect regressions, standard ARDL-Models and predictive ARDL-Models, where the latter tried to predict the next value of volatility. To increase the validity of the models, control variables were added. The used control variables were changes in the 10-year treasury rate and changes in the dollar/euro exchange rate. This chapter contains the conclusions based on the discussion on the hypotheses in chapter 5.2. Lastly, the limitations of this research are focused on after which recommendations about future research are raised.

6.1 Conclusion

As stated in section 1.1, the academic literature concerning the effect of investor sentiment on volatility is rather puzzling. Where Daniel, Hirshleifer and Subrahmanyam (2005) show that increased investor sentiment could lead to higher volatility, Lee, Jiang and Indro (2002) show quite the opposite. This research strived to clear up this relationship. In addition, the big swings in investor sentiment and volatility during the pandemic provide an opportunity for an interesting case study. As a result, the research question is: "*Has investor sentiment a significant impact on stock market volatility and did this change during the economic recovery from the covid pandemic in 2020?"*. In order to investigate this topic and increase validity, two proxies for investor sentiment and two proxies for volatility were used. As the discussion on results from section 5.2 suggests, conclusions might differ slightly depending on what proxy for investor sentiment is used.

The first set of hypotheses concerned correlation between investor sentiment and volatility. Where the AAII-Survey found a negative correlation, the put/call volume ratio for the S&P500 did not find any correlation. Since a positive correlation was hypothesized, hypothesis 1A was rejected. After focusing on the period before the stock market crash, the period of stock market recovery in 2020 was focused on. The different proxies for investor sentiment did again find different results. Using the AAII-Survey, no significant correlation could be found while the put/call volume ratio for S&P500 options do find significant positive coefficient. This positive correlation means a negative correlation between investor sentiment and volatility. As a result, hypothesis 1B is rejected for both the proxies for investor sentiment as well. Although none of the models did find any support for hypotheses 1A and 1B, 1C could not be rejected. Using the AAII-survey, the correlation changed from significantly negative to insignificant. Besides, using the put/call ratio, the correlation changed from insignificant to significantly positive. In other words, the correlation changed for both the proxies for investor sentiment during the pandemic. Although some correlation is found, it is important to note that correlation does not imply causation.

The second set of hypotheses, standard ARDL-Models were used to model the relationship between investor sentiment and volatility before and after the stock market crash. In the period before the stock market crash the result differ again for the two different proxies for investor sentiment. Once again, using the AAII-survey no significant pattern could be found and therefore hypothesis 2A is rejected when the AAII-Survey is used. On the other hand, the put/call volume ratio for S&P500 options did find a significant pattern. This pattern is characterized by a positive coefficient being the strongest. This implies more of a negative relationship between investor sentiment and volatility. As a result, hypothesis 2A was rejected using the S&P500 put/call volume ratio as well. Moving on to the period after the stock market crash, no significant pattern could be found using either of the proxies for investor sentiment. As a result, hypothesis 2B is rejected. Since no pattern is found in either of the time spans when using the AAII-Survey, no changes have taken place as a result of the stock market crash. Therefore, hypothesis 2C is rejected when using the AAII-survey. When using the put/call volume ratio for S&P500 options a pattern was found before covid but not after the stock market crash meaning that hypothesis 2C is accepted using the put/call volume ratio for S&P500 options. In conclusion, the AAII-Survey do not provide any patterns in either of the periods, while the put/call volume ratio for S&P500 options did find a positive pattern before covid. This pattern disappeared when looking at the time period after the Covid-19 related stock market crash. A critical note that has to be added to this conclusion is that the significant positive coefficient that creates this pattern, is that of the current value. This means that causation cannot be concluded since this does not deal with the issue of reverse causality.

The third and last set of hypotheses, predictive ARDL-Models were used to model the relationship between investor sentiment and volatility before and after the stock market crash. In the period before the stock market crash the results of both the proxies for investor sentiment are insignificant. No significant pattern could be found, resulting in the rejection of hypothesis 3A. The same is true for the period after the stock market crash. No pattern was observed when analyzing the outputs of the models. As a result, hypothesis 3B is rejected as well. Since no patterns have been found at all using predictive modelling and both hypotheses 3A and 3B are rejected, the rejection of hypothesis 3C is straightforward. Using predictive ARDL-Modelling no evidence for a relationship between investor sentiment and volatility was found.

In essence, this paper focused on detecting a significant effect of investor sentiment on volatility and the potential changes as a result of the changed circumstances due to the Covid pandemic. The results of the performed tests and formed models do not suggest any causal effect at all. Using the AAII-Survey, no significant pattern or effect was noted at all. Using the put/call volume ratio for S&P500 options does find some significant patterns, but these suffer from issues that keep them from being a causal effect. In the interaction regressions, correlation is found, but since correlation is not causation, no causal effect can be concluded. In the second set of hypotheses, the put/call volume ratio for S&P500 options does have a significant positive pattern. However, since the most important coefficient represents the current value for investor sentiment, reverse causality is not ruled out. As a result, no causal effect can be concluded in this case as well.

6.2 Limitations and recommendations

It should be noted that this research has its limitations. First of all, no causality could be established. Although creating an experiment that could detect causality is very hard, this research attempted to find any patterns in the association between investor sentiment and volatility. As shortly mentioned in previous section, many of the models used could not find any causal relations since reverse causality could not be ruled out. Since only the third set of hypotheses uses models that predict future values, only this model could in essence find any association that would not suffer reverse causality. This means that it could not be ascertained that any upward or downward movements in investor sentiment lead to changes in volatility, instead of the reverse in the models used in the first two sets of hypotheses. As a result, the contribution to existing literature is limited.

Besides the issue of reverse causality, another common issue in econometrics might play a big role. Namely omitted variable bias. This means that other variables that could have increased the explanatory power of the model are left out. Although it was attempted to decrease possible omitted variable bias by adding control variables to the model, the R-squared values for most of the models can be considered quite low. This is especially true for the models using the AAII-Survey as a proxy for investor sentiment. As a result, omitted variable bias seems rather probable. The fact that this research did not find an association between investor sentiment and volatility in general combined with the idea that omitted variable bias might be present, does not mean that there is no association at all. When omitted variables could be added to the models, this could also result in filtering out noise that might prevented this research from finding the real existing association. In other words, adding the omitted variables to the models could lead to the finding of an association between the independent and dependent variable. Another way to deal with omitted variable bias is conducting a well-designed experiment with a large number of participants that represent investors. In this ideal experiment, the investors trade in simulated stock markets based on a stream of specially design news announcements. By conducting this experiment two times, one with positive macro-economic news announcements and one with negative ones, all other forces except for the sentiment created with the news announcements will be removed in this simulation.

The next issue that needs to be addressed is the possibility that since both proxies for investor sentiment give significantly different outputs, one of the two (or even both) might be a bad variable. Although both the AAII-Survey (Brown and Cliff, 2005) and put/call volume ratio's (Yang & Copeland, 2014; Bandopadhyaya & Jones, 2011) are supported by established research, it must not be ruled out that one of the two or both are not applicable for the used time spans or this particular case study. This paper used relatively simple and straightforward proxies for investor sentiment, but some other academic literature proxies are combined to create an index that is could turn out to be a better predictor. For example, for the prediction of cross-sectionals of stock returns (Baker, 2007). For future research, it might be interesting to see how volatility could be affected by investor sentiment using a combined index of investor sentiment proxies.

The last matter that should be noted is the fact that the ARDL-Models that are used for the period that represents the covid recovery in 2020 use just over the minimum number of observations. The models used have either 33 or 32 observations where the minimum number of observations is considered to be 30 by Agresti and Min (2002). Although the minimum number of observations are met in the models, the relatively small number of observations might cause deviations in the outputs. The small number of observations could allow for relatively big deviations from the unobserved true coefficients to have a significant impact on the observed coefficients. This would not have been the case when the number of observations was bigger. Therefore, coefficients might be heavier influenced by outliers than would be the case if the number of observations was higher. Future research may focus on using larger amounts of data. For example, the time period from 2009 until 2019. In this period without large spikes or crashes might be a better lapse of time to study macro-economic relationships.

Chapter 7: References

Agresti, A., & Min, Y. (2002). On sample size guidelines for teaching inference about the binomial parameter in introductory statistics. *Department of Statistics University of Florida. Gainesville, Florida*, *32611*, 38545. http://users.stat.ufl.edu/~aa/ci_proportion.pdf

Baker, M. P., & Wurgler, J. A. (2003). Investor Sentiment and the Cross-section of Stock Returns. *SSRN Electronic Journal*. Published. https://doi.org/10.2139/ssrn.464843

Baker, M., & Wurgler, J. (2007). Investor Sentiment in the Stock Market. *Journal of Economic Perspectives*, *21*(2), 129–151. https://doi.org/10.1257/jep.21.2.129

Bandopadhyaya, A., & Jones, A. L. (2011). Measures Of Investor Sentiment: A Comparative Analysis Put-Call Ratio Vs. Volatility Index. *Journal of Business & Economics Research (JBER)*, *6*(8). https://doi.org/10.19030/jber.v6i8.2458

Barber, B. M., & Odean, T. (2000). Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors. *The Journal of Finance*, *55*(2), 773–806. https://doi.org/10.1111/0022-1082.00226

Barberis, N., Huang, M., & Santos, J. T. (1999). Prospect Theory and Asset Prices. *SSRN Electronic Journal*. Published. https://doi.org/10.2139/ssrn.169790

Ben-David, I., Graham, J. R., & Harvey, C. R. (2013). Managerial Miscalibration. *The Quarterly Journal of Economics*, *128*(4), 1547–1584. https://doi.org/10.1093/qje/qjt023

Benigno, P., Canofari, P., Di Bartolomeo, G., & Messori, M. (2020, november). *Uncertainty and the Pandemic Shocks*. European Parliament Monetary Dialogue Papers. https://www.europarl.europa.eu/cmsdata/215208/02.LUISS_final.pdf

Black, F. (1986). Noise. *The Journal of Finance*, *41*(3), 528–543. https://doi.org/10.1111/j.1540-6261.1986.tb04513.x

Brown, G. W., & Cliff, M. T. (2005). Investor Sentiment and Asset Valuation. *The Journal of Business*, *78*(2), 405–440. https://doi.org/10.1086/427633

Chen, G. M., Kim, K. A., Nofsinger, J. R., & Rui, O. M. (2004). *Behavior and Performance of Emerging Market Investors: Evidence from China*. Academia.edu. https://d1wqtxts1xzle7.cloudfront.net/49965167/Behavior_and_performance_of_emerging_mar20161029-6333-1mb9fq1.pdf

Chiu, C. W., Harris, R. D., Stoja, E., & Chin, M. (2018). Financial market Volatility, macroeconomic fundamentals and investor Sentiment. *Journal of Banking & Finance*, *92*, 130–145. https://doi.org/10.1016/j.jbankfin.2018.05.003

Chow, K. V., Jiang, W., & Li, J. (2020). Does VIX Truly Measure Return Volatility? *Handbook of Financial Econometrics, Mathematics, Statistics, and Machine Learning*, *2*(14), 1533–1559. https://doi.org/10.1142/9789811202391_0040

Corporate Finance Institute. (2021). *Noise Trader - Defenition, Current Numbers, Impact*. https://corporatefinanceinstitute.com/resources/knowledge/trading-investing/noise-trader/

Czaja, D., & Röder, F. (2020). Self-attribution bias and overconfidence among nonprofessional traders. *The Quarterly Review of Economics and Finance*, *78*, 186–198. https://doi.org/10.1016/j.qref.2020.02.003

Daniel, K., & Hirshleifer, D. (2015). Overconfident Investors, Predictable Returns, and Excessive Trading. *Journal of Economic Perspectives*, *29*(4), 61–88. https://doi.org/10.1257/jep.29.4.61

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (2005). Investor Psychology and Security Market Under- and Overreaction. *Advances in Behavioral Finance, Volume II*, 460–501. https://doi.org/10.1515/9781400829125-016

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise Trader Risk in Financial Markets. *Journal of Political Economy*, *98*(4), 703–738. https://doi.org/10.1086/261703

Dornbusch, R., & Fischer, S. (1980). Exchange Rates and the Current Account. *American Economic Association*, *70*(5), 960–971.

Fama, E. F. (1965). The Behavior of Stock Prices. *Journal of Business*, *38*(1). https://doi.org/10.1086/294743

Fama, E. F., & French, K. R. (2004). The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspectives*, *18*(3), 25–46. https://doi.org/10.1257/0895330042162430

Fang, L., Yu, H., & Huang, Y. (2018). The role of investor sentiment in the long-term correlation between U.S. stock and bond markets. *International Review of Economics & Finance*, *58*, 127–139. https://doi.org/10.1016/j.iref.2018.03.005

Felsenthal, M. (2020, 8 juni). *Global Economic Prospects, June 2020*. World Bank. https://www.worldbank.org/en/news/press-release/2020/06/08/covid-19-to-plunge-global-economy-into-worst-recession-since-world-war-ii

Field, A. (2018). *BUNDLE: Field: Discovering Statistics using IBM SPSS Statistics 5e + SPSS 24*. SAGE Publications.

Figlewski, S. (1997). Forecasting Volatility. *Financial Markets, Institutions and Instruments*, *6*(1), 1–88. https://doi.org/10.1111/1468-0416.00009

Fitzgerald, M. (2020, 13 mei). *Young investors pile into stocks, seeing "generational-buying moment" instead of risk*. CNBC. https://www.cnbc.com/2020/05/12/young-investors-pile-into-stocks-seeing-generational-buying-moment-instead-of-risk.html

Foucault, T., Sraer, D., & Thesmar, D. J. (2011). Individual Investors and Volatility. *The Journal of Finance*, *66*(4), 1369–1406. https://doi.org/10.1111/j.1540-6261.2011.01668.x

Glaser, M., Langer, T., & Weber, M. (2012). True Overconfidence in Interval Estimates: Evidence Based on a New Measure of Miscalibration. *Journal of Behavioral Decision Making*, *26*(5), 405–417. https://doi.org/10.1002/bdm.1773

Greene, W. H. (2008). *Econometric Analysis* (6th ed.). New York, NY. Pearson College Div.

Gürkaynak, R. S., Sack, B., & Wright, J. H. (2007). The U.S. Treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, *54*(8), 2291–2304. https://doi.org/10.1016/j.jmoneco.2007.06.029

Haridas, H. P., & Rishad, A. (2020). An empirical examination of investor sentiment and stock market volatility: evidence from India. *Financial Innovation*, *6*(1). https://doi.org/10.1186/s40854-020-00198-x

He, G., Zhu, S., & Gu, H. (2020). The Nonlinear Relationship between Investor Sentiment, Stock Return, and Volatility. *Discrete Dynamics in Nature and Society*, *2020*, 1–11. https://doi.org/10.1155/2020/5454625

Jackwerth, J. (2020). What Do Index Options Teach Us About COVID-19? *The Review of Asset Pricing Studies*, *10*(4), 618–634. https://doi.org/10.1093/rapstu/raaa012

Joo, B. A., & Durri, K. (2018). Comprehensive Review of Literature on Behavioral Finance. *Indian Journal of Commerce and Management Studies*, *4*(15).

Kim, K. H. (2003). Dollar exchange rate and stock price: evidence from multivariate cointegration and error correction model. *Review of Financial Economics*, *12*(3), 301–313. https://doi.org/10.1016/s1058-3300(03)00026-0

Lee, W. Y., Jiang, C. X., & Indro, D. C. (2002). Stock market volatility, excess returns, and the role of investor sentiment. *Journal of Banking & Finance*, *26*(12), 2277–2299. https://doi.org/10.1016/s0378-4266(01)00202-3

Leszczensky, L., & Wolbring, T. (2019). How to Deal With Reverse Causality Using Panel Data? Recommendations for Researchers Based on a Simulation Study. *Sociological Methods & Research*, 004912411988247. https://doi.org/10.1177/0049124119882473

Mansfield, E. R., & Helms, B. P. (1982). Detecting Multicollinearity. *The American Statistician*, *36*(3a), 158–160. https://doi.org/10.1080/00031305.1982.10482818

Moore, D. S., Notz, W. I., & Flinger, M. A. (2013). *The Basic Practice of Statistics* (6de ed.). W.H. Freeman and Company.

Naik, P. K., & Padhi, P. (2016). Investor sentiment, stock market returns and volatility: evidence from National Stock Exchange of India. *International Journal of Management Practice*, *9*(3), 213. https://doi.org/10.1504/ijmp.2016.077816

Odean, T. (1998). Are Investors Reluctant to Realize Their Losses? *SSRN Electronic Journal*. Published. https://doi.org/10.2139/ssrn.94142

Odean, T. (1999). Do Investors Trade Too Much? *American Economic Review*, *89*(5), 1279–1298. https://doi.org/10.1257/aer.89.5.1279

Osipovich, A. (2020, 31 augustus). *Individual-Investor Boom Reshapes U.S. Stock Market*. WSJ. https://www.wsj.com/articles/individual-investor-boom-reshapes-u-s-stock-market-11598866200

Özbey, F., İşcan, E., & Traş, M. F. (2016). How do Exchange Rate Movements affect Stock Prices? The case of Turkey. *Proceedings of the 22nd International Academic Conference*. Published. https://doi.org/10.20472/iac.2016.022.040

Pesaran, M. H. (1997). An Autoregressive Distributed-Lag Modelling Approach to Cointegration Analysis. *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*, 371–413. https://doi.org/10.1017/ccol0521633230.011

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, *16*(3), 289–326. https://doi.org/10.1002/jae.616

Poon, S. H., & Granger, C. (2005). Practical Issues in Forecasting Volatility. *Financial Analysts Journal*, *61*(1), 45–56. https://doi.org/10.2469/faj.v61.n1.2683

Qiang, Z., & Shu-e, Y. (2009). Noise Trading, Investor Sentiment Volatility, and Stock Returns. *Systems Engineering - Theory & Practice*, *29*(3), 40–47. https://doi.org/10.1016/s1874-8651(10)60010-5

Shefrin, H., & Statman, M. (1985). The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence. *The Journal of Finance*, *40*(3), 777–790. https://doi.org/10.1111/j.1540-6261.1985.tb05002.x

Statman, M., Thorley, S., & Vorkink, K. (2006). Investor Overconfidence and Trading Volume. *Review of Financial Studies*, *19*(4), 1531–1565. https://doi.org/10.1093/rfs/hhj032

Thaler, R. H., & Johnson, E. J. (1990). Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice. *Management Science*, *36*(6), 643–660. https://doi.org/10.1287/mnsc.36.6.643

Tian, G. G., & Ma, S. (2010). The relationship between stock returns and the foreign exchange rate: the ARDL approach. *Journal of the Asia Pacific Economy*, *15*(4), 490–508. https://doi.org/10.1080/13547860.2010.516171

Verma, R., & Verma, P. (2007). Noise Trading and Stock Market Volatility. *Journal of Multinational Financial Management*, *17*(3), 231–243. https://doi.org/10.1016/j.mulfin.2006.10.003

Vitale, P. (2000). Speculative noise trading and manipulation in the foreign exchange market. *Journal of International Money and Finance*, *19*(5), 689–712. https://doi.org/10.1016/s0261-5606(00)00025-5

Chapter 8: Appendices

8.1 Chapter 4: Methodology

8.1.1 First observation in 95% interval of remaining observations

Why?

To distinct the periods *During Crash* and *After Crash* it was chosen to start the period *After Crash* at the moment that volatility became normal again. The fact that volatility is not significantly higher than normal anymore would suggest that to some extend composure has returned to the financial markets.

In order to define 'significantly higher than normal' a period had to be chosen to represent the normal volatility. Although it might be tempting to choose the period before covid as normal, it could be more realistic to choose the period after the crash since volatility in a whole could have gone up during the presence of covid-19. When looking at the graphs of the volatility proxies on pages 15 and 16 it indeed looks like volatility may has gone up. Therefore, to test if volatility is on normal levels, the volatility level after the spike is chosen to be regarded as 'normal'.

Picking the right threshold observation

To find the first value that is inside the 95% confidence interval of the remaining observations in terms of weekly standard deviation of hourly S&P500 returns, the graphs on pages 15 and 16 were examined more precisely. This resulted in the guess that observation 177 had to be the first otherwise it had to be observation 175. For both the opportunities it was tested whether the remaining values followed a normal distribution by performing a Shapiro-Wilk test. This resulted in the following table:

Observations	Number of Observations	W	V	z	Prob>z
178 until 209	31	0.94699	1.727	1.132	0.12886
176 until 209	33	0.95068	1.684	1.084	0.13923

If the p-values are bigger than 0.05 the distribution is assumed to follow a normal distribution. A significance level of 0.05 is used as this is the most used in academic statistics (Field, 2018). Since both the p-values are above the rejection level of 0.05, for both the chosen observations of 177 and 175, the remaining values are assumed to follow a normal distribution. This means that whenever observation 175 falls in between the 95% confidence interval of the corresponding remaining values, observation 175 will be the threshold of the period *After Crash*. If observation 175 is not in the 95% confidence interval and 177 turns out to be in the 95% confidence interval of the corresponding remaining values, observation 177 will be the threshold. Therefore, the following table was constructed:

Observation	Sample Size	Sample Mean	Sample St. Dev	95% - lower bound	95 – upper bound	Value
177	31	0.2463433	0.0857478	0.216	0.277	0.250
175	33	0.2503346	0.0859546	0.217	0.276	0.2801

Following this table, it can be concluded that observation 177 lies in the 95% confidence interval of the corresponding remaining values for standard deviation of hourly S&P500 returns. Moreover, it can be concluded that observation 175 does not lie in the 95% confidence interval of the corresponding remaining values for standard deviation of hourly S&P500 returns. As a result, it is decided that the period *After Crash* will start at observation 177.

8.1.2 Multicollinearity table

In order to test for multicollinearity, a correlation table was generated for *Before Covid*, *During Crash* and *After Crash*. This resulted in the following tables respectively.

Before Covid	*PERBULLISH*	*PCRATIO*	*SDS&P*	*VIX*	***Δ****10YEARYIELD*	***Δ****USD/EUR*
PERBULLISH	1.0000					
PCRATIO	-0.2827	1.0000				
SDS&P	-0.2949	0.9098	1.0000			
VIX	-0.2036	0.1958	0.1575	1.0000		
Δ*10YEARYIELD*	-0.1622	0.1433	0.1135	0.1090	1.0000	
Δ*USD/EUR*	0.1293	-0.1342	-0.1290	-0.1077	-0.0550	1.0000

During Crash	*PERBULLISH*	*PCRATIO*	*SDS&P*	*VIX*	***Δ****10YEARYIELD*	***Δ****USD/EUR*
PERBULLISH	1.0000					
PCRATIO	0.1907	1.0000				
SDS&P	0.0900	0.9439	1.0000			
VIX	-0.0159	0.0572	-0.1442	1.0000		
Δ*10YEARYIELD*	0.0708	0.6783	0.5216	0.3898	1.0000	
Δ*USD/EUR*	0.2977	0.5755	0.6422	-0.3595	0.1954	1.0000

After Crash	*PERBULLISH*	*PCRATIO*	*SDS&P*	*VIX*	***Δ**10YEARYIELD*	***Δ**USD/EUR*
PERBULLISH	1.0000					
PCRATIO	-0.1539	1.0000				
SDS&P	-0.3994	0.7129	1.0000			
VIX	0.2014	-0.0353	-0.1562	1.0000		
***Δ**10YEARYIELD*	-0.0509	0.4267	0.4803	0.1446	1.0000	
***Δ**USD/EUR*	0.1579	0.3479	0.1265	-0.165	0.1907	1.0000

In order to determine whether correlations are present and if, in what form, the rule of thumb for interpreting strengths of relationships based on its *R*-value (Moore, Notz and Flinger, 2013):

$\lvert R\rvert < 0.3$:	No relationship or relationship is very weak
$0.3 < \lvert R\rvert < 0.5$:	Weak relationship
$0.5 < \lvert R\rvert < 0.7$:	Moderate relationship
$\lvert R\rvert > 0.7$:	Strong relationship

In *Before Covid* and *After Crash* only the weekly standard deviation of hourly S&P500 returns and the CBOE Volatility Index have a strong correlation. Since these are both proxies for volatility and will not be used in the same regression, this will not influence results. Therefore absence of multicollinearity is assumed. In the table about *During Crash*, it can be seen that multiple explanatory variables are moderately correlated. Therefore, absence of multicollinearity is not assumed and the assumption is violated.

8.1.3 Augmented Dickey-Fuller test outputs

This section contains the output results of the augmented dickey-fuller tests for stationarity that are used to conclude stationarity. Below three tables with the test outputs are presented. The first table presents the period *Before Covid* and the second presents *During Crash* and the last one presents *After Crash*.

Before Covid	Test Statistic	1% Crit. Value	5% Crit. Value	10% Crit. Value	P-Value Z(t)
PERBULLISH	-6.741	-2.350	-1.654	-1.287	0.0000
PCRATIO	-8.927	-2.350	-1.654	-1.287	0.0000
SDS&P	-5.893	-2.350	-1.654	-1.287	0.0000
VIX	-4.406	-2.350	-1.654	-1.287	0.0000
***Δ**10YEARYIELD*	-13.354	-2.350	-1.654	-1.287	0.0000
***Δ**USD/EUR*	-15.104	-2.350	-1.654	-1.287	0.0000

During Crash	Test Statistic	1% Crit. Value	5% Crit. Value	10% Crit. Value	P-Value Z(t)
PERBULLISH	-2.162	-2.718	-1.796	-1.363	0.0268
PCRATIO	-1.860	-2.718	-1.796	-1.363	0.0449
SDS&P	-1.497	-2.718	-1.796	-1.363	0.0813
VIX	-1.898	-2.718	-1.796	-1.363	0.0421
Δ*10YEARYIELD*	-1.708	-2.718	-1.796	-1.363	0.0578
Δ*USD/EUR*	-1.592	-2.718	-1.796	-1.363	0.0698

After Crash	Test Statistic	1% Crit. Value	5% Crit. Value	10% Crit. Value	P-Value Z(t)
PERBULLISH	-1.761	-2.453	--1.696	-1.309	0.0441
PCRATIO	-4.383	-2.453	--1.696	-1.309	0.0001
SDS&P	-3.367	-2.453	--1.696	-1.309	0.0011
VIX	-3.220	-2.453	--1.696	-1.309	0.0015
Δ*10YEARYIELD*	-4.352	-2.453	--1.696	-1.309	0.0001
Δ*USD/EUR*	-4.319	-2.453	--1.696	-1.309	0.0001

From the fact that the tables Before Covid and After Crash only have statistically significant p-values at the 5%-significance level, it can be concluded that all data is stationary. Meaning that the stationarity assumption holds for these two periods. For the period During Crash, not all data is stationary since not all p-values imply statistically significance at the 5%-significance level. Among the data that is not stationary are the two control variables. This means that control variables should not be included in the ARDL models since this could have adverse impact on the estimators. Since validity goes down when not using control variables, an ARDL model of the period *During Crash* would not be really informative.

8.2 Chapter 5: Results

8.2.1 Hypothesis set 1 output tables

(1) $SDS\&P_t = \beta_0 + \beta_1 * PERBULLISH_t * BC + \beta_2 * PERBULLISH_t * DC + \beta_3 * PERBULLISH_t * AC + \beta_4 * \Delta 10YEARYIELD_t + \beta_5 * \Delta USD/EUR_t + u_t$

Generates the following regression output:

Source	SS	DoF	MS
Model	3.81843902	5	0.763687804
Residual	2.51287908	203	0.012378715
Total	6.3313181	208	0.030439029

Number of Obs	= 209
F(5,203)	= 61.69
Prob > F	= 0.0000
R-Squared	= 0.6301
Adj. R-Squared	= 0.5933
Root MSE	= 0.11126

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]

$\boldsymbol{PERBULLISH * BC}$	-0.3648896	.1110906	-3.28	0.001	-0.583929	-0.1458502
$\boldsymbol{PERBULLISH * DC}$	0.8376202	.1637832	5.11	0.000	0.5146857	1.160555
$\boldsymbol{PERBULLISH * AC}$	-0.0932731	.1233277	-0.76	0.450	-0.3364406	0.1498945
$\boldsymbol{\Delta 10YEARYIELD}$	0.7217363	.1193038	6.05	0.000	0.4865026	0.9569699
$\boldsymbol{\Delta USD/EUR}$	3.623566	1.170317	3.10	0.002	1.316031	5.931101
Constant	0.1992569	.0407736	4.89	0.000	0.1188629	0.279651

(2) $VIX_t = \beta_0 + \beta_1 * PERBULLISH_t * BC + \beta_2 * PERBULLISH_t * DC + \beta_3 * PERBULLISH_t * AC + \beta_4 * \Delta 10YEARYIELD_t + \beta_5 * \Delta USD/EUR_t + u_t$

Generates the following regression output:

Source	SS	DoF	MS
Model	12910.3983	5	2582.07967
Residual	4860.4471	203	23.9430892
Total	17770.8454	208	85.4367569

Number of Obs	= 209
F(5,203)	= 107.84
Prob > F	= 0.0000
R-Squared	= 0.7265
Adj. R-Squared	= 0.7198
Root MSE	= 4.8932

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
$\boldsymbol{PERBULLISH * BC}$	-27.9531	4.885731	-3.28	0.000	-37.58639	-18.31982
$\boldsymbol{PERBULLISH * DC}$	44.79463	7.203137	5.11	0.000	30.59207	58.99719
$\boldsymbol{PERBULLISH * AC}$	4.531874	5.423915	-0.76	0.404	-6.162562	15.22631
$\boldsymbol{\Delta 10YEARYIELD}$	21.43377	5.246948	6.05	0.000	11.08827	31.77928
$\boldsymbol{\Delta USD/EUR}$	148.4341	51.47018	3.10	0.004	46.94942	249.9189
Constant	21.57842	1.79321	4.89	0.000	18.04271	25.11412

(3) $SDS\&P_t = \beta_0 + \beta_1 * PCRATIO_t * BC + \beta_2 * PCRATIO_t * DC + \beta_3 * PCRATIO_t * AC + \beta_4 * \Delta 10YEARYIELD_t + \beta_5 * \Delta USD/EUR_t + u_t$

Generates the following regression output:

Source	SS	DoF	MS

Number of Obs	= 209
F(5,203)	= 55.50
Prob > F	= 0.0000
R-Squared	= 0.5775
Adj. R-Squared	= 0.5671
Root MSE	= 0.011479

Model	3.65653541	5	0.731307083
Residual	2.67478268	203	.013176269
Total	6.3313181	208	.030439029

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
$\boldsymbol{PCRATIO * BC}$	0.0318168	0.037561	0.85	0.398	-0.0422421	0.1058758
$\boldsymbol{PCRATIO * DC}$	0.2623779	0.044167	5.94	0.000	0.1752927	0.3494631
$\boldsymbol{PCRATIO * AC}$	0.0923626	0.041342	2.23	0.027	0.0108484	0.1738767
$\boldsymbol{\Delta 10YEARYIELD}$	0.6997393	0.125237	5.59	0.000	0.452807	0.9466715
$\boldsymbol{\Delta USD/EUR}$	4.4051	1.187954	3.71	0.000	2.062789	6.747411
Constant	0.0118882	0.067116	0.18	0.860	-0.1204461	0.1442225

(4) $VIX_t = \beta_0 + \beta_1 * PCRATIO_t * BC + \beta_2 * PCRATIO_t * DC + \beta_3 * PCRATIO_t * AC + \beta_4 * \Delta 10YEARYIELD_t + \beta_5 * \Delta USD/EUR_t + u_t$

Generates the following regression output:

Source	SS	DoF	MS
Model	12427.7189	5	2485.54379
Residual	5343.12649	203	26.3208201
Total	17770.8454	208	85.4367569

Number of Obs	= 209
F(5,203)	= 94.43
Prob > F	= 0.0000
R-Squared	= 0.6993
Adj. R-Squared	= 0.6919
Root MSE	= 5.1304

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
$\boldsymbol{PCRATIO * BC}$	-0.8195305	1.67875	-0.49	0.626	-4.129553	2.490493
$\boldsymbol{PCRATIO * DC}$	13.24106	1.974026	6.71	0.000	9.348835	17.13328
$\boldsymbol{PCRATIO * AC}$	6.349206	1.847742	3.44	0.001	2.705978	9.992435
$\boldsymbol{\Delta 10YEARYIELD}$	22.05468	5.5974	3.94	0.000	11.01818	33.09117

$\Delta USD/EUR$	175.2528	53.09493	3.30	0.001	70.5645	279.941
Constant	12.90027	2.999722	4.30	0.000	6.985664	18.81488

8.2.2 Hypothesis set 2 output tables

Before Covid

(1) $SDS\&P_t = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-k}) + \Sigma(\text{þ}_i * PERBULLISH_{t-i}) + \Sigma(\text{ø}_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(\text{ð}_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

Generates the following regression output:

Sample: 5 until 163

Number of Obs	= 159
F(5,203)	= 22.90
Prob > F	= 0.0000
R-Squared	= 0.5149
Adj. R-Squared	= 0.4924
Root MSE	= 0.0635

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
SDS&P						
L1	0.6350766	0.0596741	10.64	0.000	0.5171726	0.7529806
PERBULLISH						
--	-0.3836425	0.0971692	-3.95	0.000	-0.5756293	-0.1916557
L1	0.3146925	0.1109212	2.84	0.005	0.0955345	0.5338504
L2	-0.20346	0.106589	-1.91	0.058	-0.4140585	0.0071385
L3	0.3330947	0.0930264	3.58	0.000	0.1492931	0.5168962
$\Delta 10YEARYIELD$	0.0495525	0.1004853	0.49	0.623	-0.1489863	0.2480914
$\Delta USD/EUR$	-1.004097	0.9315061	-1.08	0.283	-2.844566	0.8363715
Constant	0.037018	0.0404158	0.92	0.361	-0.0428355	0.1168715

(2) $VIX_t = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-k}) + \Sigma(\text{þ}_i * PERBULLISH_{t-i}) + \Sigma(\text{ø}_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(\text{ð}_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

Generates the following regression output:

Sample: 5 until 163

Number of Obs	= 159
F(5,203)	= 22.90
Prob > F	= 0.0000
R-Squared	= 0.5149
Adj. R-Squared	= 0.4924
Root MSE	= 0.0635

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VIX						
L1	0.7618351	0.0516488	14.75	0.000	0.6598035	0.8638667
PERBULLISH						
--	-2.833454	3.12176	-0.91	0.365	-9.000453	3.333545
$\boldsymbol{\Delta 10YEARYIELD}$	1.001201	3.817944	0.26	0.793	-6.541101	8.543503
$\boldsymbol{\Delta USD/EUR}$	-62.09577	35.37072	-1.76	0.081	-131.9702	7.778657
Constant	4.815132	1.522055	3.16	0.002	1.80833	7.821933

(3) $SDS\&P_t = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-k}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

Generates the following regression output:

Sample: 5 until 163

Number of Obs	= 159
F(5,203)	= 32.13
Prob > F	= 0.0000
R-Squared	= 0.5122
Adj. R-Squared	= 0.4962
Root MSE	= 0.0633

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
SDS&P						
L1	0.6964482	0.0588856	11.83	0.000	0.5801143	0.8127821
PCRATIO						
--	0.1198739	0.0245113	4.89	0.000	0.0714497	0.1682982
L1	-0.0695609	0.0244898	-2.84	0.005	-0.1179427	-0.021179
$\boldsymbol{\Delta 10YEARYIELD}$	0.089046	0.098655	0.90	0.368	-0.105856	0.2839479
$\boldsymbol{\Delta USD/EUR}$	-1.362635	0.9126953	-1.49	0.138	-3.165747	0.4404773
Constant	-0.0381889	0.0488541	-0.78	0.436	-0.1347045	0.0583268

(4) $VIX_t = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-k}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

Generates the following regression output:

Sample: 5 until 163

Number of Obs	= 159
F(5,203)	= 22.90
Prob > F	= 0.0000
R-Squared	= 0.5149
Adj. R-Squared	= 0.4924
Root MSE	= 0.0635

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VIX						
L1	0.8394651	0.044916	18.69	0.000	0.7507248	0.9282054
PCRATIO						
--	5.485527	0.8521602	6.44	0.000	3.80192	7.169135
L1	-2.979718	0.8582295	-3.47	0.001	-4.675317	-1.284119
L2	-1.903269	0.8241218	-2.31	0.022	-3.531482	-0.2750566
$\Delta \mathbf{10YEARYIELD}$	-0.0454629	3.379643	-0.01	0.989	-6.722603	6.631677
$\Delta \mathbf{USD/EUR}$	-29.25375	31.42335	-0.93	0.353	-91.33668	32.82917
Constant	1.473448	1.91011	0.77	0.442	-2.300346	5.247241

After Crash

(1) $SDS\&P_t = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-k}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

Generates the following regression output:

Sample: 177 until 209

Number of Obs	= 33
F(5,203)	= 6.08
Prob > F	= 0.0012
R-Squared	= 0.4647
Adj. R-Squared	= 0.3883
Root MSE	= 0.0682

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
SDS&P						
L1	0.3946386	0.1444698	2.73	0.011	0.0987055	0.6905716
PERBULLISH						
--	-0.1467199	0.1374521	-1.07	0.295	-0.4282778	0.134838
$\Delta 10YEARYIELD$	0.4984549	0.2145205	2.32	0.028	0.0590296	0.9378802
$\Delta USD/EUR$	3.658632	1.751657	2.09	0.046	0.0705248	7.24674
Constant	0.1326116	0.0658572	2.01	0.054	-0.0022908	0.267514

(2) $VIX_t = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-k}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

Generates the following regression output:

Sample: 177 until 209

Number of Obs	= 33
F(5,203)	= 6.89
Prob > F	= 0.0005
R-Squared	= 0.4960
Adj. R-Squared	= 0.4240
Root MSE	= 3.3653

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VIX						
L1	0.3637652	0.1524078	2.39	0.024	0.051572	0.6759585
PERBULLISH						
--	-11.12422	7.601863	-1.46	0.155	-26.69593	4.447495
$\Delta 10YEARYIELD$	32.3388	10.60432	3.05	0.005	10.61683	54.06077
$\Delta USD/EUR$	44.72729	86.24671	0.52	0.608	-131.9411	221.3957
Constant	18.23482	5.711277	3.19	0.003	6.535801	29.93384

(3) $SDS\&P_t = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-k}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

Generates the following regression output:

Sample: 177 until 209

Number of Obs	= 33
F(5,203)	= 5.57
Prob > F	= 0.0020
R-Squared	= 0.4430
Adj. R-Squared	= 0.3635
Root MSE	= 0.0695

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
SDS&P						
L1	0.4269891	0.1440909	2.96	0.006	0.1318322	0.722146
PCRATIO						
--	-0.0053094	0.0953897	-0.06	0.956	-0.2007063	0.1900875
$\boldsymbol{\Delta 10YEARYIELD}$	0.5155343	0.2217337	2.33	0.028	0.0613334	0.9697352
$\boldsymbol{\Delta USD/EUR}$	3.397581	1.797035	1.89	0.069	-0.2834787	0.7.078641
Constant	0.0851713	0.1623205	0.52	0.604	-0.2473272	0.4176698

(4) $VIX_t = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-k}) + \Sigma(\text{þ}_i * PCRATIO_{t-i}) + \Sigma(\text{ø}_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(\text{ð}_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$ and $k \geq 1$

Generates the following regression output:

Sample: 177 until 209

Number of Obs	= 33
F(5,203)	= 6.73
Prob > F	= 0.0006
R-Squared	= 0.4901
Adj. R-Squared	= 0.4172
Root MSE	= 3.3850

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VIX						
L1	0.4643586	0.1339375	3.47	0.002	0.1900002	0.738717
PCRATIO						
--	-6.21925	4.648419	-1.34	0.192	-15.74111	3.302605
$\boldsymbol{\Delta 10YEARYIELD}$	35.41892	10.83138	3.27	0.003	13.23183	57.606
$\boldsymbol{\Delta USD/EUR}$	6.510347	86.99981	0.07	0.941	-171.7007	184.7214

Constant	22.20246	8.615242	2.58	0.016	4.554933	39.84998

8.2.3 Hypothesis set 3 output tables

Before Covid

(1) $SDS\&P_{t+1} = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-i}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

Generates the following regression output:

Sample: 5 until 162

Number of Obs	= 158
F(5,203)	= 28.28
Prob > F	= 0.0000
R-Squared	= 0.4251
Adj. R-Squared	= 0.4100
Root MSE	= 0.0687

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
SDS&P						
--	0.6756955	0.0646224	10.46	0.000	0.5480282	0.8033629
PERBULLISH						
--	0.0887077	0.0868147	1.02	0.308	-0.0828026	0.260218
Δ10YEARYIELD						
--	-0.0428893	0.1073669	-0.40	0.690	-0.2550024	0.1692237
ΔUSD/EUR						
--	1.485346	0.998486	1.49	0.139	-0.4872529	3.457946
Constant	0.0077074	0.0359783	0.21	0.831	-0.063371	0.0787858

(2) $VIX_{t+1} = \beta_0 + \Sigma(\Psi_i * VIX_{t-i}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

Generates the following regression output:

Sample: 5 until 162

Number of Obs	= 158
F(5,203)	= 62.42
Prob > F	= 0.0000
R-Squared	= 0.6200
Adj. R-Squared	= 0.6101
Root MSE	= 2.4772

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VIX						
--	0.8078708	0.0523892	15.42	0.000	0.7043711	0.9113705
PERBULLISH						
--	3.936733	3.141384	1.25	0.212	-2.269354	10.14282
$\boldsymbol{\Delta 10YEARYIELD}$						
--	-1.239236	3.86001	-0.32	0.749	-8.865035	6.386562
$\boldsymbol{\Delta USD/EUR}$						
--	26.553	35.99383	0.74	0.462	-44.55606	97.66206
Constant	1.248696	1.545443	0.81	0.420	-1.804466	4.301859

(3) $SDS\&P_{t+1} = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-i}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

Generates the following regression output:

Sample: 5 until 162

Number of Obs	= 158
F(5,203)	= 28.07
Prob > F	= 0.0000
R-Squared	= 0.4233
Adj. R-Squared	= 0.4082
Root MSE	= 0.0688

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
SDS&P						
--	0.6669684	0.0635176	10.50	0.000	0.5414837	0.7924531
PCRATIO						
--	-0.0187385	0.02492	-0.75	0.453	-0.0679702	0.0304932
$\boldsymbol{\Delta 10YEARYIELD}$						

--	-0.489396	0.1071762	-0.46	0.649	-0.2606758	0.1627966
ΔUSD/EUR						
--	1.525179	0.9984962	1.53	0.129	-0.4474407	3.497798
Constant	0.0726849	0.0444952	1.63	0.104	-0.0152194	0.1605893

(4) $VIX_{t+1} = \beta_0 + \Sigma(\Psi_i * VIX_{t-i}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

Generates the following regression output:

Sample: 5 until 162

Number of Obs	= 158
F(5,203)	= 63.56
Prob > F	= 0.0000
R-Squared	= 0.6243
Adj. R-Squared	= 0.6145
Root MSE	= 2.4634

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VIX						
--	0.8017161	0.0506385	15.83	0.000	0.7016751	0.9017571
PCRATIO						
--	-1.614069	0.8867545	-1.82	0.071	-3.365933	0.1377946
Δ10YEARYIELD						
--	-1.150105	3.826027	-0.30	0.764	-8.708766	6.408557
ΔUSD/EUR						
--	25.17036	35.75787	0.70	0.483	-45.47255	95.81327
Constant	5.5275	1.680674	3.29	0.001	2.207177	8.847823

After Crash

(1) $SDS\&P_{t+1} = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-i}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

Generates the following regression output:

Sample: 177 until 208

Number of Obs	= 32
F(5,203)	= 2.26
Prob > F	= 0.0892
R-Squared	= 0.2506
Adj. R-Squared	= 0.1396
Root MSE	= 0.0821

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
SDS&P						
--	0.4366623	0.2073279	2.11	0.045	0.0112605	0.862064
PERBULLISH						
--	-0.123466	0.1724006	-0.72	0.480	-0.4772029	0.2302709
Δ10YEARYIELD						
--	0.13225949	0.2805398	0.47	0.640	-0.4430252	0.7082149
ΔUSD/EUR						
--	0.2740903	2.299858	0.12	0.906	-4.44483	4.99301
Constant	0.1657092	0.0783301	2.12	0.044	0.0049892	0.3264293

(2) $VIX_{t+1} = \beta_0 + \Sigma(\Psi_i * VIX_{t-i}) + \Sigma(þ_i * PERBULLISH_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

Generates the following regression output:

Sample: 177 until 208

Number of Obs	= 32
F(5,203)	= 2.90
Prob > F	= 0.0405
R-Squared	= 0.3007
Adj. R-Squared	= 0.1971
Root MSE	= 4.0242

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]

VIX						
--	0.378938	0.2060504	1.84	0.077	-0.0438424	0.8017184
PERBULLISH						
--	-5.780212	9.20578	-0.63	0.535	-24.66891	13.10849
$\boldsymbol{\Delta 10YEARYIELD}$						
--	14.57024	14.4731	1.01	0.323	-15.1261	44.26658
$\boldsymbol{\Delta USD/EUR}$						
--	20.69576	107.1088	0.19	0.848	-199.0733	240.4648
Constant	17.14928	6.943548	2.47	0.020	2.902292	31.39626

(3) $SDS\&P_{t+1} = \beta_0 + \Sigma(\Psi_i * SDS\&P_{t-i}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

Generates the following regression output:

Sample: 177 until 208

Number of Obs	= 32
F(5,203)	= 4.94
Prob > F	= 0.0026
R-Squared	= 0.4874
Adj. R-Squared	= 0.3888
Root MSE	= 0.0692

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
SDS&P						
--	0.3786756	0.172724	2.19	0.037	0.0236364	0.7337148
PCRATIO						
--	-0.1132759	0.0983759	-1.15	0.260	-0.3154904	0.0889387
L1	-0.3206865	0.1051038	-3.05	0.005	-0.5367305	-0.1046426
$\boldsymbol{\Delta 10YEARYIELD}$						
--	0.1165848	0.2428707	0.48	0.635	-0.382643	0.6158126
$\boldsymbol{\Delta USD/EUR}$						

--	-2.351746	1.99018	-1.18	0.248	-6.442619	1.739127
Constant	0.8789504	0.2241086	3.92	0.001	0.4182886	1.339612

(4) $VIX_{t+1} = \beta_0 + \Sigma(\Psi_i * VIX_{t-i}) + \Sigma(þ_i * PCRATIO_{t-i}) + \Sigma(ø_i * \Delta 10YEARYIELD_{t-i}) + \Sigma(ð_i * \Delta USD/EUR_{t-i}) + u_t$ where $i \geq 0$

Generates the following regression output:

Sample: 177 until 208

Number of Obs	= 32
F(5,203)	= 3.97
Prob > F	= 0.0063
R-Squared	= 0.4880
Adj. R-Squared	= 0.3652
Root MSE	= 3.5782

	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
VIX						
--	0.2797909	0.1732902	1.61	0.119	-0.0771069	0.6366886
PCRATIO						
--	-4.207149	5.310659	-0.79	0.436	-15.14466	6.730359
L1	-8.066469	5.434263	-1.48	0.150	-19.25854	3.125605
L2	-13.20608	6.091949	-2.17	0.040	-25.75768	-0.659472
$\boldsymbol{\Delta 10YEARYIELD}$						
--	6.748127	14.09767	0.48	0.636	-22.28658	35.78283
$\boldsymbol{\Delta USD/EUR}$						
--	-118.5342	102.3082	-1.16	0.258	-329.2419	92.17357
Constant	61.41025	16.0659	3.82	0.001	28.3219	94.49859