Norwegian School of Economics
Bergen, Spring 2021

Market sentiment and its predictive abilities in the stock market

Empirical study of leading indicators derived from market sentiment

Daniel Sandal Skiftesvik

Øystein Kvalvik Vasshus

Supervisor: Svein-Arne Persson

Master´s thesis in Economics and Business Administration,

Financial economics

NORWEGIAN SCHOOL OF ECONOMICS

Abstract

The purpose of this thesis is to investigate the relationship between multiple possible leading indicators and the stock market (S&P500 index). A leading indicator can be defined as a piece of data that corresponds to the future movements in a variable of interest. Thus, considering the information they contain, one may be able to anticipate the movement of the stock market. The indicators to examine are representations of market sentiment, reflecting the risk tolerance of investors in the financial markets. They have been determined based on existing literature, in addition to reflections of the authors of this thesis. For example, the Put/Call ratio (Pan and Poteshman, 2003) and the Gold/Platinum ratio (Huang and Kilic, 2019) has been shown to have predictive abilities towards the stock market. Previous research has mainly focused on the potential leading indicators individually. To explore their abilities combined is therefore of interest.

The results appear to favorize the inclusion of multiple indicators in this analysis. However, it shows conflicting evidence for the predictive abilities of the indicators toward the stock index. Parts of the analysis appear to show predictive capabilities, while others are less conclusive, which complicates the utilization of the results. The patterns seem to change as different periods are examined, but when changes in the economic environment are considered many of these observations seem reasonable.

A vector autoregressive model (VAR-model) is facilitating the use of both the present and past notations of the indicators in a system. Multiple models were created, one where the whole sample period was included, in addition to several subsets to display the potential changes that occur. The model including the whole sample period (2005-2020) is reporting an adjusted R^2 of 12,0%, where the subsets range from 18,1% to 28,4%, suggesting that the importance of the indicators varies through time, as different pattern emerges.

The results from the VAR-model are evaluated using granger causality test and impulse response analysis. The indicators appear to be granger causal throughout the analysis, with only one exemption in one of the subsets. In the impulsive response functions, a theoretical shock is performed on the indicators. It illustrates that whenever there is an increase in indicators like the Put/Call ratio, the U.S. Dollar index (the relative strength of the dollar) and

the credit spread (Baa-Aaa), the stock index appears to consistently react negatively to this. The Gold/Platinum and the Volume Weighted Moving Average looks to result in rather inconsistent reactions. The decomposition of the variance shows that most of the variance comes from the shocks in the stock index itself, which is the recurring observation throughout the subsets. This is to be expected in the beginning of the simulated period. Later, more of the variance is explained by the other indicators. Nonetheless, the increase should have been larger to claim it supports the hypothesis of the predictive abilities.

Preface

This thesis was written as a conclusion to our Master of Science in Economics and Business administration at the Norwegian School of Economics (NHH). Our interest in the financial markets was the driving force steering us towards the subject of the thesis. An interest that has grown during our time as students at the school.

The experience of writing this thesis has been educational, as it has proved both challenging and inspirational. Utilizing the knowledge and understanding of finance acquired during the master's degree at NHH gave us great pleasure. The process also gave us the opportunity to further extend our understanding of the subject.

We would like to show our appreciation to Svein-Arne Persson for his guidance and constructive feedback during this educational and rewarding process.

Bergen, June 2021

Norwegian School of Economics

Daniel Sandal Skiftesvik

Øystein Kvalvik Vasshus

Contents

List of figures

List of tables

1. Introduction

Numerous studies have been conducted exploring the relationship between possible leading indicators and the stock market, where many of them are demonstrations of market sentiment. Brown and Cliff (2001) define market sentiment as a representation of the expectations of the market participants relative to a norm. A bullish (bearish) investor expects returns to be above (below) average, whatever "average" may be. The market sentiment indicates the attitude of the market participants towards securities or financial markets. Through the price movement of the traded securities, one can reveal the crowd psychology, or the feelings of the market.

The return of the stock market is strongly linked to the willingness of investors to undertake risk, which is what the price movements unveil. To ensure information to be obtained from a wide range of the economy, the price movements of various financial markets can be studied, giving a perception of how widespread the risk acceptance is. The risk aversion may unfold at different times in various markets, making it valuable information to consider when timing the entry/exit in securities or financial markets.

Research has shown[1] several indicators to have predictive abilities towards the stock market, calculated from various financial markets. As existing literature focus on leading indicators individually, the possibility of combining them in a statistical model is of great interest. That is the approach of this thesis, where the aim is to reveal the reaction of the S&P500 index following a significant change in the hypothesized leading indicators.

The indicators chosen are calculated from the commodity market, the option market, the currency market and the bond market. By including these markets, the idea is to capture the information if the risk aversion is unfolding at different times in the financial market. If one can identify a change in risk tolerance in a financial market that precedes the stock market, this information can be utilized in an investment strategy.

[1] (Huang and Kilic 2019), (Johnson and So, 2012), (Pan, Poteshman, 2003), (Blau, Nguyen and Whitby, 2013)

1.1 Research question

Can indicators hypothesized to have predictive abilities towards the stock market (S&P500 index) be combined in a vector autoregressive model in order to explain the historical price movements? – And to what extent do their predictive abilities change through time?

1.2 Thesis structure

To answer the research question, this thesis is divided into seven chapters. Following the introduction, the thesis starts with presenting various financial definitions and theoretical fundamentals to provide a brief introduction to the efficient market hypothesis, market sentiment and a presentation of the potential leading indicators. Thereafter, econometric methodologies are introduced and used in analyzing the research question. The findings are presented along external events that might have an impact. Next, we discuss the possible utilization of the results in combination with the limiting elements of this analysis, before we provide a conclusion to the research question.

1.3 Limitations

This thesis studies the return of the S&P500 index on a daily frequency. As every indicator is reported on a daily basis and are assumed to impact the stock index rather quickly, the daily frequency is considered advantageous. The indicators can vary extensively within a short time span, meaning less frequent data (weekly, monthly) could result in significant loss of information. However, a negative consequence of this decision is a higher likelihood of violating the desired whiteness of the residuals, a term used for diagnostics of the residuals. It requires the residuals to have expected value equal to zero, no autocorrelation, homoscedasticity and to be normally distributed.

Lütkepohl (2007, pp. 157) argues that the importance of whiteness in the residuals is dependent on the intended use of the model. For instance, when forecasting is the main objective, it may not be of prime importance. He further claims of higher importance of the whiteness whenever the model order (lags) is not chosen by statistical methods but for example on the basis of some economic theory.

Several methods for correcting the residuals were tried unsuccessfully, resulting in the presence of heteroscedasticity, autocorrelation and non-normality. Following Lütkepohls arguments, we decided to disregard the whiteness of the residuals to a certain extent. It is however important to be aware of this as it may affect the interpretation and the conclusion of the results. Reflections about this matter is further elaborated in section 6.2.

2. Theory and litterature review

As this thesis strive to create a model explaining the movements of the stock market based on historical notations, it would contradict the efficient market hypothesis. Therefore, this theory, as well as previously shown anomalies, is of relevance. It is also important to understand how the market sentiment is created and how it is affected by basic elements of the human psychology.

2.1 Efficient market hypothesis

The efficient market hypothesis (EMH) states that stock prices reflect all available information. Every stock trade at their fair value, and subsequently a consistent alpha is impossible to generate. The EMH are split into three levels: weak form, semi-strong form and strong form. According to the weak form, the market reflects all historical prices and the information they contain. Semi-strong form states that the prices reflect all publicly available information. While the strong form includes all information, both public and private (Fama, 1970).

Fama (1970) concludes that both weak and semi-strong form of the efficient market hypothesis is supported in the US capital markets. He also states that the strong form is best viewed as a benchmark, from which deviations from market efficiency can be judged. Fama revisited EMH in 1997 when recent financial literature seemed to produce multiple long-term return anomalies. The paper concludes that the anomalies are fragile. When reasonable changes are made in the way abnormal returns are calculated, they tend to disappear. Therefore, the efficient market hypothesis should not be abandoned. Empirical evidence and consensus among academics appear to agree with this conclusion, yet there are still anomalies challenging it.

2.1.1 Financial market anomalies

Tversky and Kahneman (1986) defined market anomalies as “a deviation from the presently accepted paradigms that is too widespread to be ignored, too systematic to be dismissed as random error, and too fundamental to be accommodated by relaxing the normative system”. Different types of anomalies occur across the market. Some are related to specific time periods, like the weekend effect (Smirlock & Starks, 1986), showing the likelihood of stock prices to fall on Mondays. Others are bound to fundamentals, exemplified by Fama & French (1988), who found that stocks with high dividend yield outperform the market. Another group of anomalies are connected to technical analysis, which is of great relevance to this thesis. If the market holds weak form of efficiency, it entails that an investor cannot earn abnormal returns on the basis of technical analysis, which cannot explain this anomaly.

Technical analysis is used to forecast future price movements of stocks on the basis of historical prices and information. Included in such an analysis is strategies like moving averages and trading range breaks (resistance/support). By utilizing a dataset from the Dow Jones Industrial Average Index from 1897 to 1986, Brock, Lakonishok and LeBaron (1992) found on a consistent basis, returns acquired from buy (sell) signals produce returns that are higher (lower) than “normal” returns. Furthermore, they found that there is also less volatility in the returns following buy signals than sell signals.

Further evidence of technical anomalies is demonstrated by Hon & Tonks (2001). They investigated the presence of abnormal returns by utilizing a momentum strategy on the UK stock in the years of 1955-1996. Their findings show that investors can gain an advantage that cannot be accounted for by a simple adjustment for beta-risk. In the period of 1977-1996 one could gain abnormal returns by buying past winners and selling past losers. However, this was not apparent in the period of 1955-1976, where they could not find such a presence. Implicating that momentum is not a general feature, but only appears over certain periods of time.

Models have been created trying to explain the anomalous behavior of the assets. Wouters (2006) categorizes investors into two groups, rationalists and behavioralists. She defines rationalists as those who believe abnormal returns are either due to common risk factors being ignored, or by luck. They believe the markets are efficient. She explains that behavioralists makes their investment decisions based on sentiment. Their paradigm is that only a small number of the market participants are required to be rationale in order to drive the whole market. The result is market anomalies through mispricing of securities, and the cause is the sentiment of the investors.

2.2 Market sentiment

Hui and Li (2014) finds that there are two key points in defining market sentiment: "One is expectations, which is the investors' believes and judgements about the future trend; the other one is errors in expectations, which means the expectation may be biased." They further elaborate that this gives the idea that there are two kinds of sentiment in the market. The optimism (pessimism) of the fundamentals should already be priced in, while the sentiments of noise traders who are bullish (bearish) can also affect the price.

There are several indicators and measures of market sentiment, and they can be measured directly or indirectly from economic variables. One can calculate indicators directly from the financial markets, as well as use survey-based indicators. They contain information about how optimistic or pessimistic the market participants are about the current market and can be used by investors in their decision making. The main focus of this thesis is indicators calculated with basis in the financial markets. Examples of such indicators are the Put/Call ratio, the Volatility index, the Gold/Platinum ratio, moving averages etc. The theory of how each one explains sentiment follows later in this chapter.

2.2.1 How market sentiment affects market prices

The sentiment of the market participants drives the demand and supply of securities, and consequently moves prices. A combination of sentiment indicators and trading frameworks are often used by traders in order to define entry and exit signals. The key in this regard is to interpret the information correctly and act on it fast to maximize returns.

Economic models and finance theory are often heavily based upon two assumptions, the beforementioned market efficiency and rationality. Such theory portrays humans as rational beings who always attempt to maximize utility. This is challenged by the proponents of the behavioral finance theory who believe that numerous factors drive investors behavior and decision-making, including both rational thinking and irrational thinking. As a consequence, they believe that the market price is not always a fair estimate of the underlying fundamentals, as the psychology of investors can drive fundamental values and market prices very far apart (Shefrin, 2000, pp.4-11).

The decision makers' satisfice (March and Simon, 1958, pp.99-101) says the judgement of people is generally confined in their rationality. They will waive the best solution in favor of an acceptable or reasonable one. The judgement of investors is affected by specific systematic biases (Kahneman and Tversky, 1974).

The Behavioural Finance Theory

Tversky and Kahneman displays evidence to suggest that investors inhabit psychologically grounded irrationality. Their decisions often contradict fundamental rules and are guided by their own perceptions of the world and/or influenced by other investors´ actions. Investors have cognitive and emotional biases that determine their behavior in the marketplace, resulting in a deviation from the rational investor. Examples of such are disproportionate reliance on historical performance, difficulties modifying their views to new information and overconfidence in their ability to forecast future price movements. (Tversky & Kahneman, 1974)

The Animal Spirit Theory

The animal spirit theory by John Maynard Keynes explains how people arrive at financial decisions in times of uncertainty and economic distress. He assumes cognitive biases where individuals under uncertainty are dominated by their instincts. Consequently, their actions are a result of their sentiment, rather than a thorough analysis. Investors flock to the market as it is surging, expecting the trend to continue. Eventually, the inexorable downturn follows, and the psychology of the market turn progressively pessimistic. Remarkably, investors hold on to their risky portfolios in order to avoid capitalizing losses. Thus, the market sentiment is strongly linked to herd behavior which may allow for irrational enthusiasm (Keynes, 1936, pp.161-162).

2.3 Literature review

This section will present a review of the literature examining potential leading indicators on the stock market in addition to the rationale for including these indicators in this analysis. In advance, a search through multiple reliable databases was conducted to locate relevant literature. It is important to note that this is not nearly all the research conducted on this subject and there may be similar research giving contradicting conclusions to those presented below. However, these articles contain rational economic contexts that are interesting to include in further analysis.

The basis for the choice of indicators is the possibility of other markets having predictive abilities towards the stock market. As previously mentioned, the markets considered in this thesis is the commodity market, the option market, the currency market and the bond market. Gold/Platinum ratio is derived from the commodity market, and the option market is represented by the Put/Call ratio, the Option to Stock ratio and the Implied Volatility index (VIX). The currency market is reflected using an index showing the relative strength of the U.S. dollar. Lastly, to gain information from the bond market different yield spreads and credit spreads are considered.

Every figure presented in this chapter are originals created from the data used for the following analysis. They are visualized at different timespans to make the interpretation easier.

2.3.1 The commodity market

The Gold/Platinum ratio

A study conducted by Huang and Kilic (2019) examined the relationship between daily notations of the prices of the commodities gold and platinum and the equity market. While both commodities functions as consumption goods in jewelry among other, only gold is also considered a financial collateral. Therefore, when studying the ratio between gold and platinum one isolates the shocks from consumption of jewelry from the store hold of wealth. It thereby reveals variations in risk tolerance, and thus functions as a proxy for economic state and risk tolerance. The authors of the article displays that the Gold/Platinum ratio is a strong predictor of future stock market return. A one standard deviation increase in the Gold/Platinum ratio predicts a 6,4% increase in the US stock market excess returns over the following year. Huang and Kilic ran multiple regressions examining the predictive power of the Gold/Platinum ratio on mainly the Center for Research in Security Pricing (CRSP) value weighted index and MSCI World Index as the dependent variables. However, the predictability of the Gold/Platinum ratio on the MSCI World Index is smaller than for the US returns.

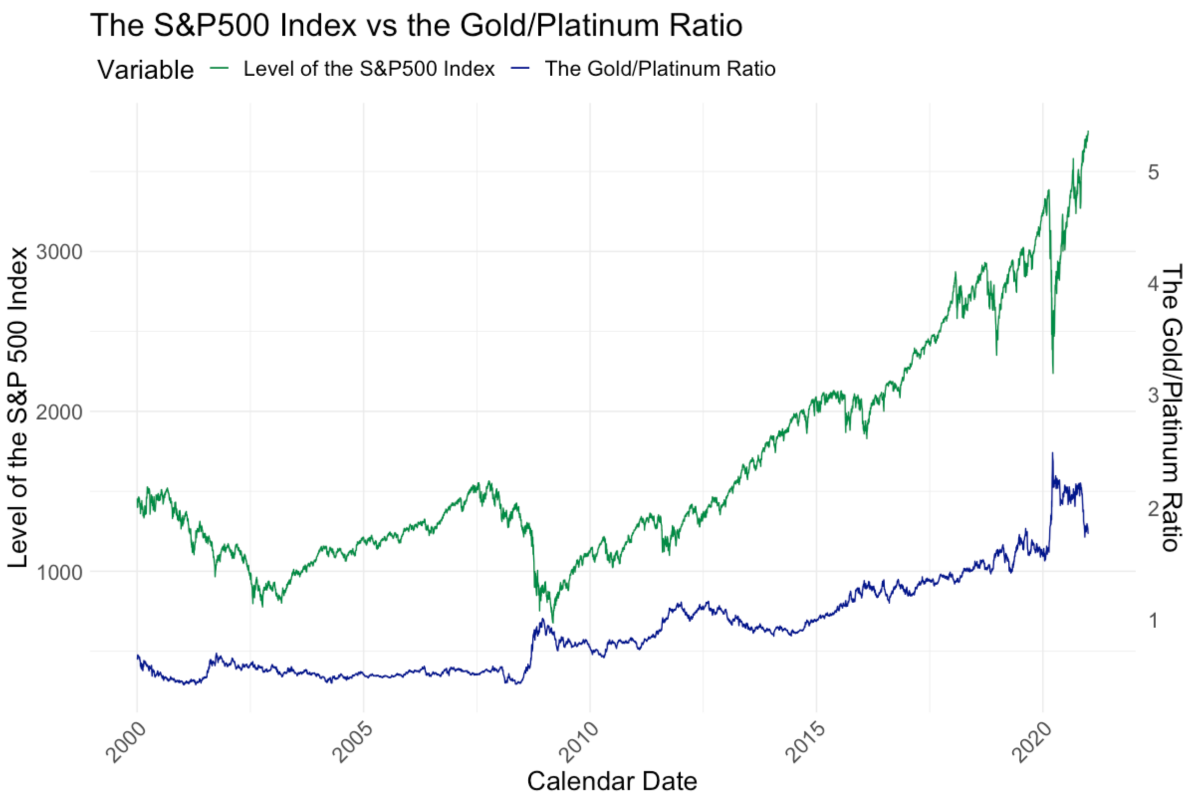


Figure 1: The S&P500 index vs the Gold/Platinum Ratio (2000-2020)

The relationship between the variables is most visible in the major corrections of the S&P500 index, where the Gold/Platinum ratio increases rapidly, supporting the hypothesis that it can serve as a proxy for investors willingness to hold risky assets. The figure also illustrates that the variables seem to reach their extremes at different times, with the Gold/Platinum ratio preceding the S&P500 index, substantiating the argument of leading abilities. However, this pattern appears clearer in the times of distress. Meaning when the stock market crashes, the Gold/Platinum ratio peaks, and might function as a buy signal. The troughs of the Gold/Platinum ratio are harder to interpret, leading to an understanding that it might not function as well as a sell signal in times of prosperity.

2.3.2 The options market

The Put/Call ratio

Black (1975) argues that the option market provides informed investors with an additional platform to trade. The possibility of higher leverage opportunities and lower downside risk will drive informed investors towards options as opposed to the underlying security itself.

According to Pan and Poteshman (2003) there is strong evidence that option trading volume holds information about future changes in underlying stock prices. They determined that increases in the daily Put/Call ratios are negatively related to the next day returns and therefore argues that the ratio has predictive abilities, by containing information about the future spot price in the stock market. They present an investment strategy buying/selling stocks with respective low/high Put/Call ratios, resulting in a return of 40 basis points per day. The Put/Call ratio is defined in this article as the put volume divided by the put plus call volume. The conclusion of Pan and Poteshman (2003) is reached by using proprietary trade data obtained directly from the Chicago Board of Options Exchange (CBOE) to construct Put/Call ratios by volume initiated by buyers to open new positions. The rationale is; if an investor with positive (negative) information about a stock decides to trade in the option market, the easiest trade is to open a new call (put) position, which offers leverage with limited liability.

As the database used by Pan and Poteshman (2003) is not available to the public, in this thesis, we are not able to distinguish whether the trades are new openings of the options or not. A later paper by Blau, Nguyen and Whitby (2013) conclude that unsigned (not proprietary data) also have predictive powers. Their dataset was limited to the total put volume divided by the total put plus call volume. However, the CBOE Put/Call ratio used in this thesis is the total put volume divided by the total call volume.

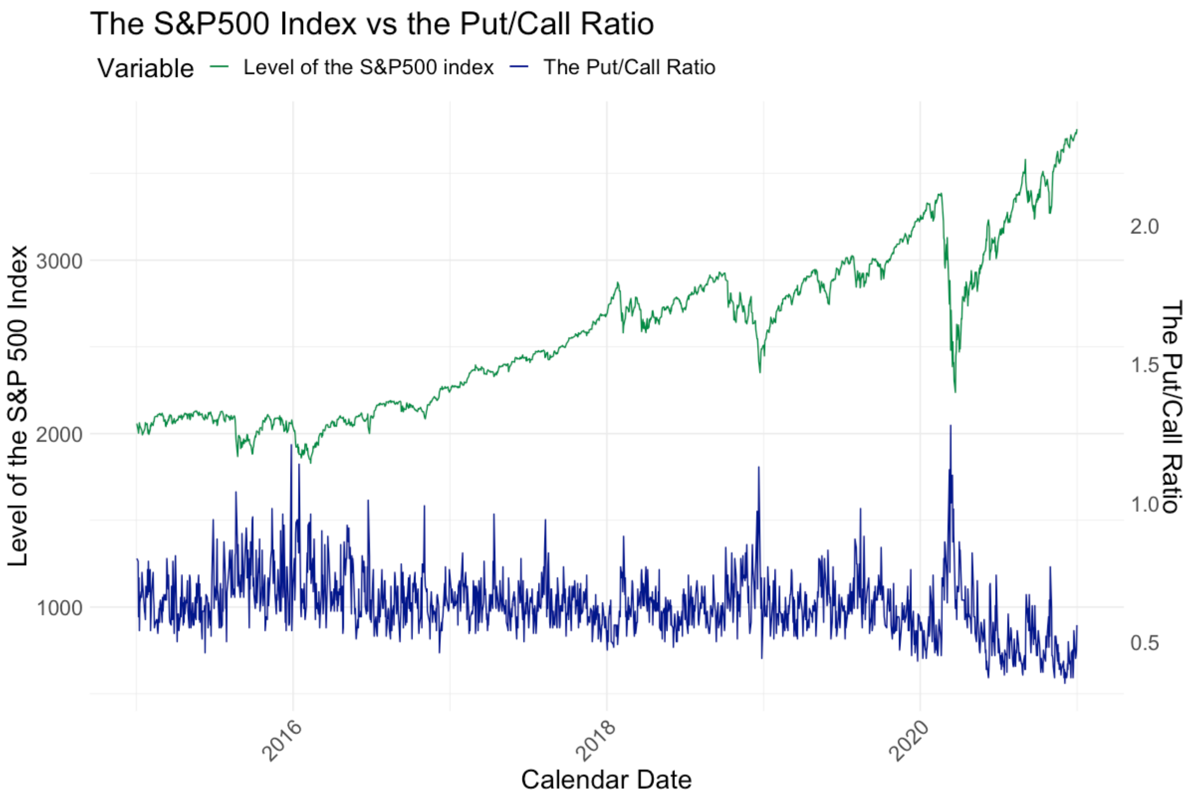


Figure 2: The S&P500 index vs the Put/Call Ratio (2015-2020)

With every correction of the S&P500 index, we can see the Put/Call ratio rising, showing the negative correlation between the variables. Each major spike of the ratio appears to be followed by the S&P500 index bottoming shortly after. Indicating it might be of use in timing when to enter the market and subsequently giving the ratio leading abilities. It is important to note that there are some false breakouts in the ratio where the S&P500 index does not have corresponding downturns. The Put/Call ratio also appear to have its lowest notations when the S&P500 index approaches the highs, though not as evident as the previously mentioned spikes, making it harder to utilize as a sell signal.

The Option to Stock volume

A different measure emerging from the options market is the Option to Stock volume. Johnson and So (2012) examines the informational content of options and equity volumes when direction is unobserved. They conclude that the amount of trading in the options markets relative to the equity markets is a negative cross-sectional signal of private information. The impact from the ratio is negative due to short-sale costs, as investors trade more frequently in the option markets relative to equity markets when they possess information giving a negative signal, rather than the contrary. In their analysis firms were split into deciles based on their Option to Stock ratios. The firms in the lowest decile outperformed the highest decile by 0,34% each week (19,3% annualized).

The ratio is calculated using the total volume of put plus call options divided by the total volume of the underlying assets.

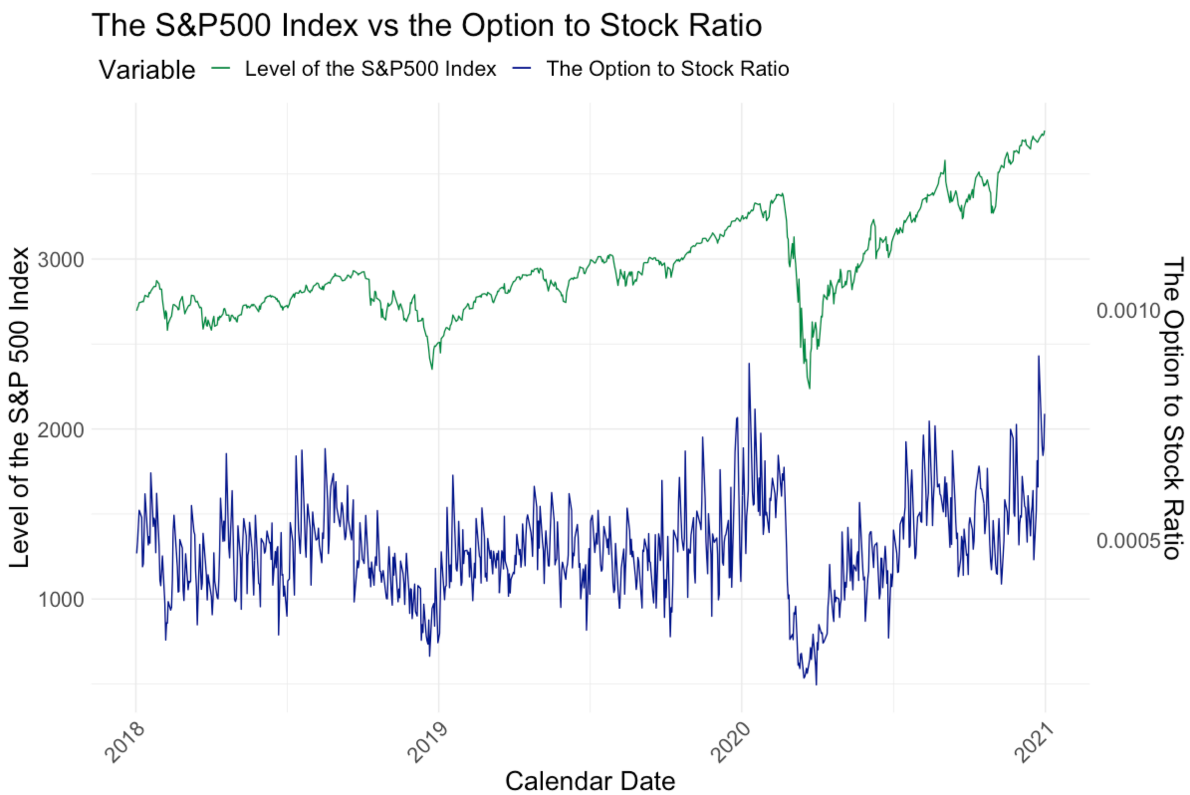


Figure 3: The S&P500 index vs the Option to Stock Ratio (2018-2020)

The Option to Stock ratio has a positive correlation to the S&P500 index. Figure 3 appears to show that the peaks and throughs comes earlier in the Option to Stock ratio, implying its possible foretelling powers. The trend is easily visible in the major corrections, but not equally evident in times of less volatility. However, in the periods of low volatility, there are a couple of spikes in the ratio that does not coincide with a reaction in the S&P500 index.

The CBOE Implied Volatility Index (VIX)

The CBOE implied volatility index provides a measure of the market risk and sentiment through the price inputs of the S&P500 index options. It is a real-time market index displaying the market participants expected volatility over the coming 30 days (Kuepper, 2020). Implied in this definition is an understanding that the index could have predictive abilities in the short term, where a sudden increase in the volatility index could signify a coming correction in the markets.

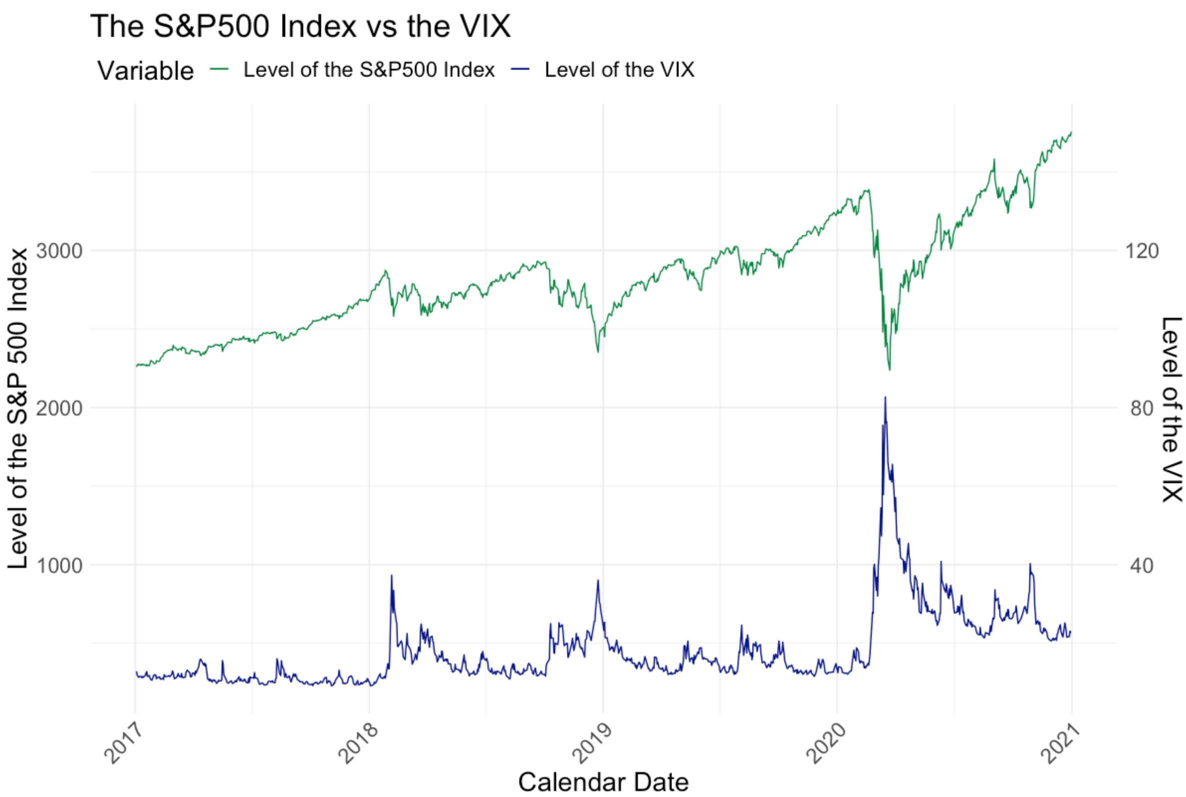


Figure 4: The S&P500 index vs the Implied Volatility index (2017-2020)

The volatility index is low in times of prosperity but increases rapidly whenever there is rising uncertainty in the market. This is visualized by a spike every time the S&P500 index decreases significantly. The spikes appear to precede the lows of the stock index. Thus, having leading properties giving reason to believe it can function as a buy signal at the peak. Contrary to the previously mentioned indicators derived from the options market, it also appears to have few false breakouts. Every time the ratio spikes, a corresponding correction seem to be evident in the stock index, implying that it can be a reliable indicator. One can interpret this increase as a sell signal.

2.3.3 The currency market

The U.S. dollar Index

The U.S. dollar index is a measure of the relative strength of the US dollar compared to a basket of foreign currencies. It is a weighted geometric mean of the value relative to the following currencies: Euro (EUR, 57,6%), Japanese yen (JPY, 13,6%), Pound sterling (GBP, 11.9%), Canadian dollar (CAD, 9,1%), Swedish krona (SEK, 4,2%) and the Swiss franc (CHF, 3,6%) (Chen, 2020).

The dollar is viewed as a minimal risk currency of storing wealth. It has the status as the reserve currency and is therefore often referred to as a *safe haven*. In times of uncertainty, a risk averse investor will allocate a larger part of the portfolio in less risky assets. American government bonds are considered safe in this environment due to the liquidity in the bonds and the solvency of the American government. Along with the size and liquidity of the American securities market, this might increase the foreign activity in these markets, and thus strengthen the currency through a higher demand for it.

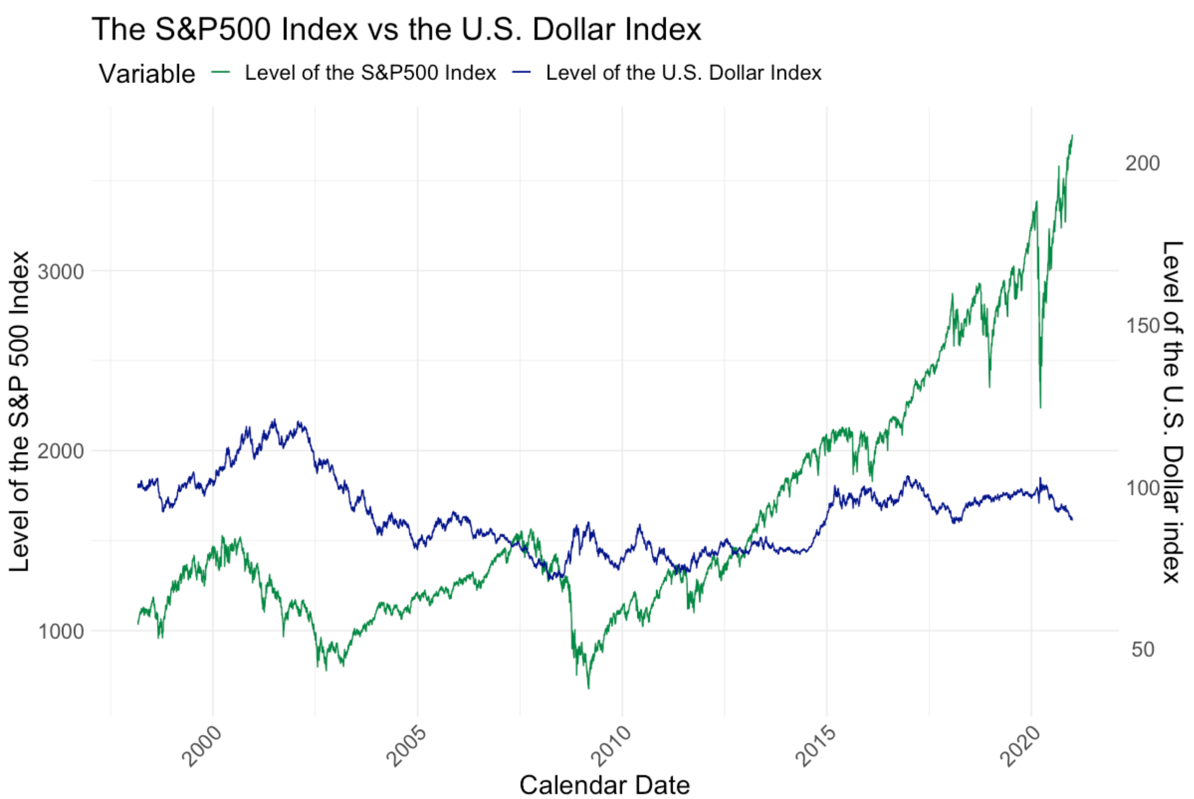


Figure 5: The S&P500 index vs the U.S. Dollar Index (2000-2020)

Safe haven effects can to a certain extent be identified in the figure 5. One can see a slight decrease in the strength in times of economic prosperity and somewhat of an increase in times of distress. These effects are more visible in years before the Great Recession, and less visible in the years after. The monetary policy performed after the Great Recession with the printing of money may affect this relationship as an increase in the supply can counteract the effect an increase of demand would normally have on the price.

2.3.4 The bond market

The yield spread

Yield and credit spreads are generally acknowledged as a strong indicator for business cycles and the likelihood of entering a recession or a recovery. Yield spreads are calculated by the difference between a long-term bond and a short-term bond, for instance the U.S. Treasuries. The difference between the 10-year U.S. Treasury bond and the 3-month U.S. Treasury bill is most commonly used for this purpose. If disregarding the risk premium, the yield curve reflects the markets expectations of future short-term rates, implying that one can deduce the market participants expectation of the economic development. As economic conditions worsen, the interest rates are expected to be lowered. Accordingly, the longer-term rates will decrease as the markets expectations of future turmoil increase.

Through monetary policy, the central banks are able to affect the yield curve, but their ability to impact the market rates is greater in the short range than the long range. A contraction by the central bank will increase the nominal short-term interest rates and vice versa. The room to maneuver for the central banks diminishes as interest rates move towards zero. Therefore, when interest rates are historically low, central banks around the world can perform monetary policies where they increase the demand for the long-term bonds and thus lowering the yield for long-term funding. This action illustrates to the market participants that the interest rates will remain low and thereby encourages investors to increase the economic activity. This may affect the predictive abilities of the yield curve as it is no longer solely based on the expectations of the markets, but also reflects governmental interventions.

A *twisted yield-curve* has been a reliable sign that the economy is expected to experience turbulence in the near future, historically. The term refers to a situation where the long-term interest rates are lower than the short-term rates. In figure 6, this phenomenon is illustrated by the yield spread crossing the x-axis. In this context it is important to state that the stock market is not the economy. There are various factors affecting the stock market, and in the short run the trend of the stock market can deviate significantly from the trend of economic productivity. However, in the long run, they are bound to intertwine.

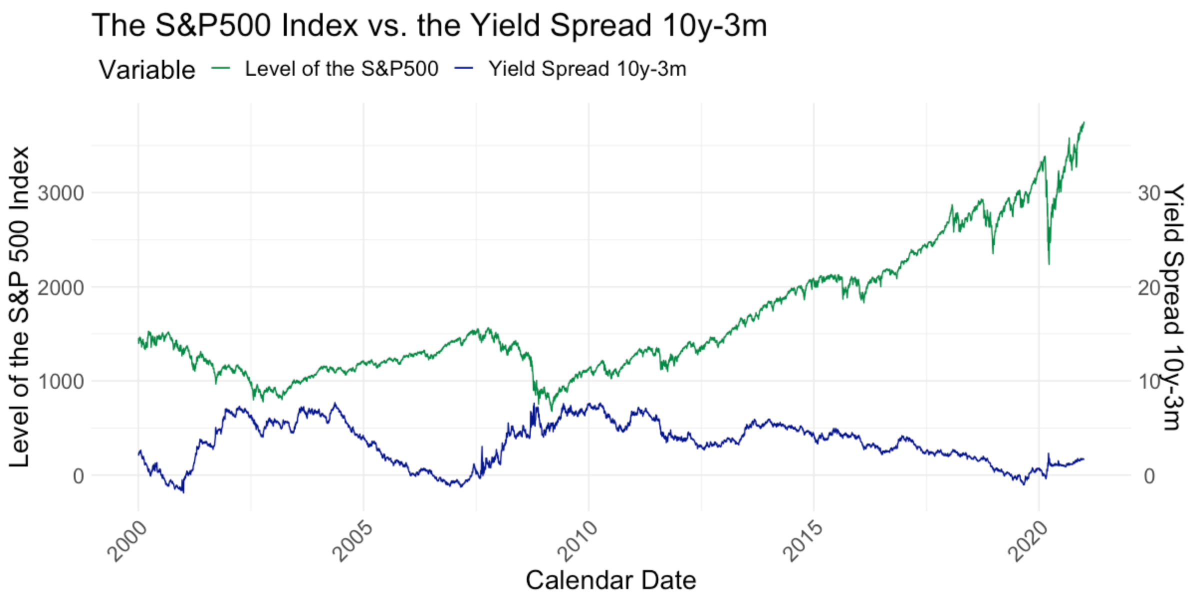


Figure 6: The S&P500 index vs the Yield Spread 10y-3m (2000-2020)

The twisting of the yield curve is observed on four occasions within this timespan, in the early 2000s, before the Great Recession, in the mid 2019 and in March of 2020. These observations can be related to times of significant uncertainty. It appears that the initial twisting of the spread precedes the top of the stock market index, supporting the hypothesis of its predictive powers. The variables also seem to have opposed trends, where the stock market trends upwards whenever the yield spread trends in the opposite direction, and vice versa.

The credit spread

The credit spread is often defined as the difference between two bonds of the same maturity but differentiated by the credit rating of the bond issuer. Moody's is an American financial services company providing investors with credit ratings and risk analysis. A well-known spread to analyze is the difference between the corporate bonds rated as Baa versus Aaa. Bonds rated Aaa are judged to be of the highest quality with minimal risk, whereas Baa are subject to moderate credit risk. As they are considered medium-grade they may possess speculative characteristics[2].

In times of financial distress, investors willingness to take risks decrease. The demand for risky bonds falls, leading to a lower price. Thus, bondholders demand a higher yield for holding riskier bonds with more default risk. As a result, the spread increases with the level of risk aversion in the market.

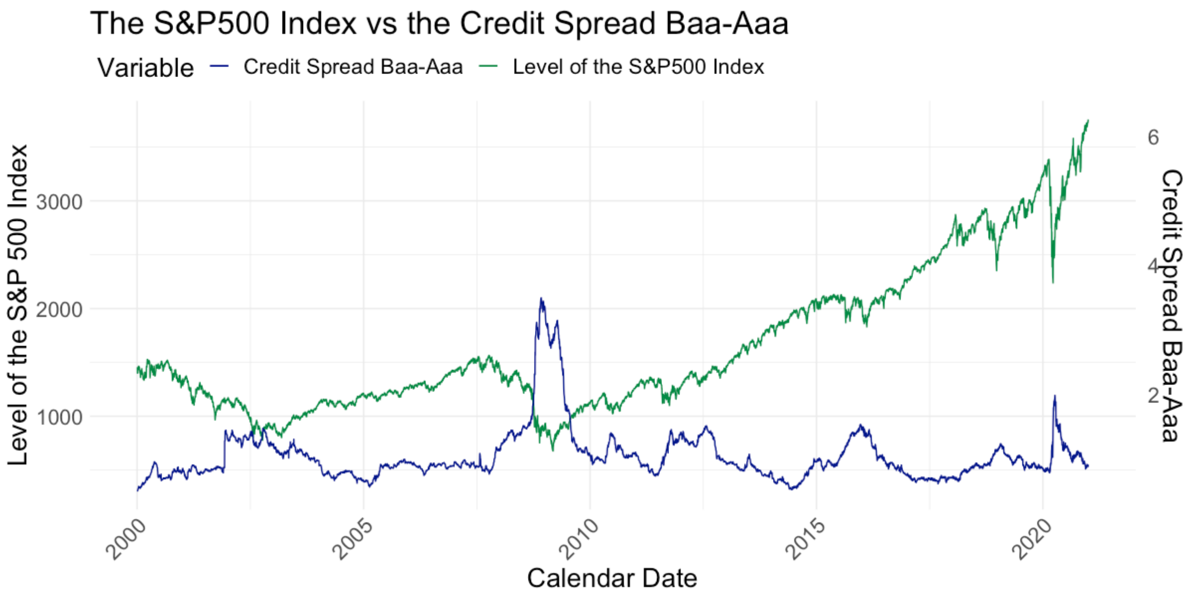


Figure 7: The S&P500 index vs the Credit Spread Baa-Aaa (2000-2020)

Figure 7 illustrates the relationship were the credit spread decreases in times of prosperity and increases when uncertainty enters the markets. The spread appears to rise whenever the S&P500 index stagnates or trends downwards, with the peaks of the spread coinciding with trough of the index. This relationship also holds true when the variables move in the opposite direction. The initial rise/fall in the credit spread looks to precede the change of direction of the stock index.

The credit spread has several peaks during this period. The most significant of all is the increase in the spread during the Great Recession. The difference between the two corporate

[2]Moody's description of the rated bonds
https://www.moodys.com/viewresearchdoc.aspx?docid=PBC_79004

bonds rating peaked at approximately 3,5% higher yield in the Baa rated bond than in the Aaa rated bond. The difference of increase in the credit spread during the financial crisis versus the recent corona-crisis were likely because the Federal Reserve quickly announced interventions in the credit market[3], not allowing the spread to increase further.

2.3.5 Technical indicator

The Volume Weighted Moving Average (VWMA)

The volume weighted moving average (hereafter VWMA) is a metric actively used by traders to determine trends in the stock market. Due to mean reversion, assets do not deviate from their moving average for an extended period of time relative to the moving average considered, leading to it being commonly used in trading. Volume is also considered in this metric because whenever the volume is low it takes fewer transactions to affect the price. Therefore, by using this indicator one takes the changes in volume into account as well as the changes in price. It is given by this formula; (Fernando, 2021)

$$VWMA = \frac{\sum Price * Volume}{\sum Volume}. \tag{2.1}$$

A security's deviation from its VWMA can be considered a measure of sentiment as it visualizes the crowd psychology. It shows how the participants often are piling in/out of the security due to the beforementioned irrationality. Consequently, strengthening the ongoing momentum. As the deviation grows in either direction, so the reason to believe there is a presence of euphoria/dysphoria and a reversion towards the mean can be expected.

[3]Press release from the Federal Reserve Bank
https://www.federalreserve.gov/newsevents/pressreleases/monetary20200323b.htm

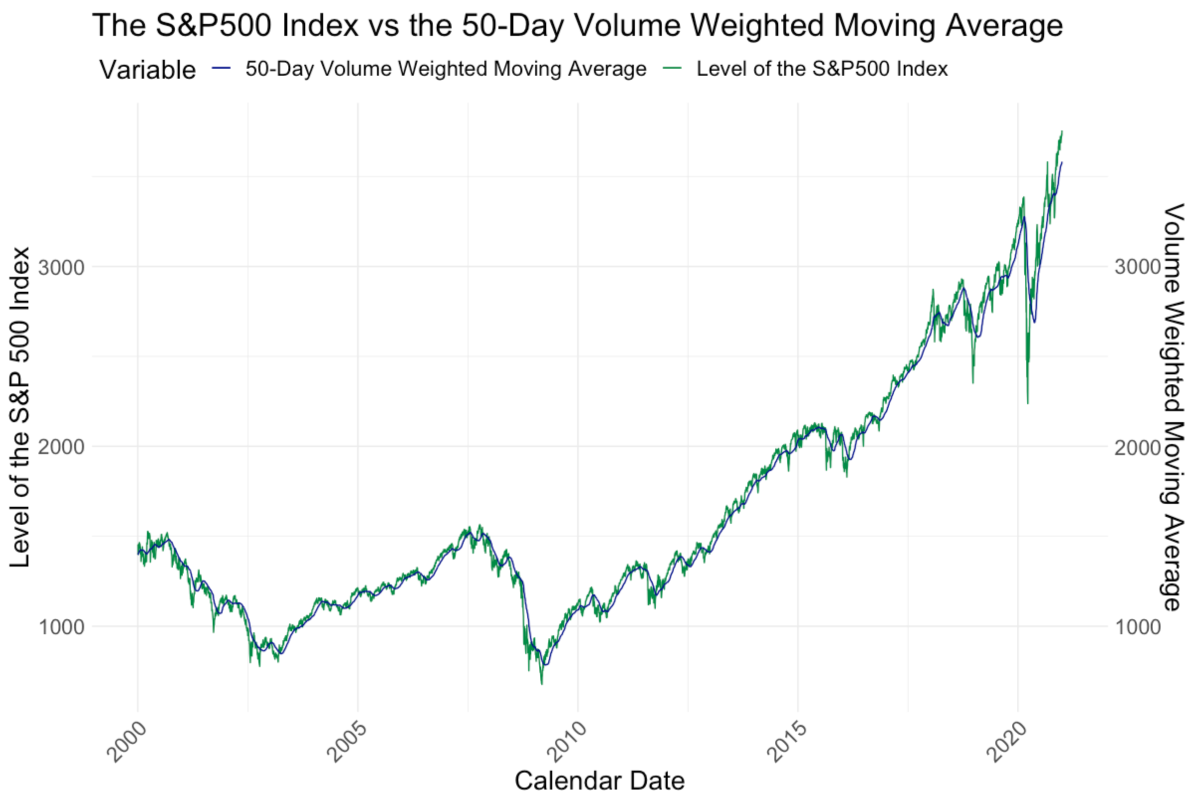


Figure 8: The S&P 500 index vs 50-day volume weighted moving average (2000-2020)

When price is trending upwards, the VWMA will in general be lagging, meaning staying below the price. The opposite is true when the price is trending downwards. This is visualized in figure 8, exemplified by the 50-day VWMA. The leading abilities comes from the expectation that the price cannot deviate from its own moving average for an extended period of time due to mean reversion. With every major deviation from the VWMA, the price corrects fairly quickly as the figure illustrates. The difference between the VWMA and the S&P500 during the market turmoil in 2020 is striking.

2.3.6 Correlation between the variables

To quantify the connections and to gain intuition on the relationship between the previously presented figures, a correlation matrix (Table 1) is presented below. There are huge variations in the correlation coefficients of the indicators to the S&P500 index. The VWMA is created directly from the index and therefore has a high correlation of 0.995. The Gold/Platinum is also strongly correlated with the index, with a coefficient of 0.872. The yield spread presented have a correlation ≈ -0.45 to the stock index. The correlation of the remaining indicators to the S&P500 index are for the most part negative and rather low.

Additionally, the matrix is used to ensure that the indicators are not different measures that are representing the same relationships. For instance, if two indicators are highly correlated, a model would not necessarily benefit from including both indicators in order to explain the movements of the stock index as this would likely cause multicollinearity. Disregarding the S&P500 index, there are few correlations that are noteworthy. The coefficient of 0.682 between the Credit spread Baa-Aaa and the VIX is significant. So is the relationship between the VWMA and the Gold/Platinum ratio giving a correlation of 0.875. As the data is to be further processed before it is utilized, we see no reason to eliminate indicators on the basis of this.

Correlation matrix of the indicators

	Level of the S&P500 index	Gold/Platinum ratio	Put/Call Ratio	Option/Stock Ratio	The VIX	Yield spread 10y-3m	Credit spread Baa-Aaa	U.S. dollar index	Volume Weighted Moving Average
Level of the S&P500 index	1	0.872	-0.161	-0.243	-0.271	-0.470	-0.273	0.158	0.995
Gold/Platinum ratio	0.872	1	-0.046	-0.453	0.021	-0.250	0.044	0.058	0.875
Put/Call Ratio	-0.161	-0.046	1	-0.271	0.274	0.130	0.255	-0.116	-0.117
Option/Stock Ratio	-0.243	-0.453	-0.271	1	-0.089	-0.185	-0.293	0.257	-0.256
The VIX	-0.271	0.021	0.274	-0.089	1	0.202	0.682	0.105	-0.223
Yield spread 10y-3m	-0.470	-0.250	0.130	-0.185	0.202	1	0.266	-0.236	-0.476
Credit spread Baa-Aaa	-0.273	0.044	0.255	-0.293	0.682	0.266	1	-0.122	-0.261
U.S. dollar index	0.158	0.058	-0.116	0.257	0.105	-0.236	-0.122	1	0.170
Volume Weighted Moving Average	0.995	0.875	-0.117	-0.256	-0.223	-0.476	-0.261	0.170	1

Table 1: Correlation between the indicators (2000-2020)

3. Methodology

A common econometric approach is to use models with lagged values in order to test for predictive powers in prior observations of the data. This section will start with a presentation of the autoregressive model to give a brief introduction to the vector autoregressive model, which is the main model in this thesis. Criterions and tests to meet the requirements of a VAR-model follows thereafter, and lastly a presentation of the framework for analyzing the results.

3.1 Introduction to methodology and priliminary requirements

3.1.1 The Autoregressive model

The autoregressive model (AR-model) is an established model used in financial econometrics to analyze economic data. The intention of the model is to study a variable y on its past values to predict current and future values of y. An AR(p) model can be denoted as a linear model, where p represents the number of lagged values included in the model,

$$y_t = \alpha_0 + ß_1 y_{t-1} + ß_2 y_{t-2} + ß_3 y_{t-3} + \ldots + ß_p y_{t-p} + \varepsilon_t. \quad 3.1$$

As equation 3.1 illustrates, y_t is regressed on lagged values, till p numbers of lags. It is normal assume that the error-term (ε_t) has an expected value of zero, $E(\varepsilon_t)$= 0, and variance, E $(\varepsilon_t^2) = \sigma^2$.

When examining the AR-model, one can test the dependency on a value to prior values of the time series, or else the time series follows a random walk. This economic concept is a stochastic and random process and implies that the time series does not follow a determined path driven by trend or seasonality, for instance. Enders (2010, pp.184) denotes a random walk as:

$$y_t = y_{t-1} + \varepsilon_t. \quad 3.2$$

A random walk means that y at time t is obtained by starting with the previous variable y_{t-1} and adding a zero mean random variable that is independent of y_{t-1}. If y_t is independent of time t, the process can be defined as a random walk. Assumptions in a random walk is that it is independent and identically distributed (i.i.d) (ε_t: t = 1, 2, ….) with mean zero and constant variance (σ_ε^2).

3.1.2 Vector autoregressive model

A vector autoregressive model (VAR-model) is constructed by bringing together several variables in terms of their own past values, in comparison to an AR-model which only model a single variable on its past values. Enders (2010, pp.297-298) denotes a VAR-model with two series, y and z, and one lagged value as:

$$y_t = b_{10} - b_{12}z_t + \gamma_{11}y_{t-1} + \gamma_{12}z_{t-1} + \varepsilon_{yt} \tag{3.3}$$

$$z_t = b_{20} - b_{21}y_t + \gamma_{21}y_{t-1} + \gamma_{22}z_{t-1} + \varepsilon_{yt}. \tag{3.4}$$

Where it is assumed that (1) both y_t and z_t are stationary; (2) ε_{yt} and ε_{yt} are white-noise disturbance with standard deviations of σ_y and σ_z, respectively; and (3) error terms are uncorrelated white-noise disturbance. The system incorporates feedback because y_t and z_t are allowed to affect each other. For example, -b_{21} is the contemporaneous effect of a unit change of y_t on z_t, and γ_{12} will report the effect of a unit change in z_{t-1} on y_t.

Equation 3.3 and 3.4 can be presented in matrix-form as follows:

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}\begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}. \tag{3.5}$$

The VAR-model is useful in forecasting purposes and when testing for leading variables in time series. If the regression returns coefficients that statistically significantly deviate from zero ($\neq 0$), one might conclude that a time series have leading characteristics toward another.

The coefficients in a VAR-model are hard to interpret and therefore other means are used to evaluate the results. The major interest is to obtain the largest adjusted R-squared.

3.1.3 Stationarity

The data is required to be stationary when utilizing a VAR-models. A stationary time series is not dependent on time, trends or seasonality. Time series that contain trends and seasonality will affect the value of the series depending on the time period being studied.

A covariance-stationary process is given by Watsham and Parramore (1997, pp.231) as the following:

$$E(X_t) = \mu \tag{3.6}$$

$$E(X_t^2) = \sigma^2 \tag{3.7}$$

$$cov(X_{t1}, X_{t2}) = \gamma_{t1,t2} = \gamma_\tau . \tag{3.8}$$

These equations illustrates that the process has a constant expected value (3.6) and variance (3.7), simultaneously as the covariance is dependent of the interval $\tau,\ (t1 - t2) = \tau$ (3.8).

3.1.4 Stationarity test – test for unit roots

Dickey-Fuller test

In order to examine if a time series is stationary one can perform unit root tests. This test investigates if a time series is non-stationary and possess a unit root. The null hypothesis is defined by the presence of a unit root. Dickey and Fuller (1979) introduced three different regression equations and Enders (2010, pp.206) describes their method as:

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t \tag{3.9}$$

$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + \varepsilon_t \tag{3.10}$$

$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + \alpha_2 \mathrm{t} + \varepsilon_t . \tag{3.11}$$

The parameter of interest is γ. Equation 3.9, under the condition that $\gamma = 0$ is defined as a "pure random walk" model. Where equation 3.10 adds an intercept (α_0) or a *drift* term, equation 3.11 includes both a *drift* and a *linear time trend.* Dickey and Fullers test involves estimating the equations of best fit in order to obtain the estimated value of γ and its associated standard error. This allows the researcher to compute a t-statistic using Dickey-Fuller tables to determine whether to accept or reject the null hypothesis.

Augmented Dickey-Fuller test

As an extension of the original model, Dickey and Fuller formulated the Augmented Dickey-Fuller test (ADF-test). The ADF-test enables testing of a higher-order equation. Not every time series in economic data can be well represented in the first-order autoregressive process and the presence of autocorrelated error terms is not unusual. Hence, the ADF-test is normally used on economic data.

The test is described in Enders (2010, pp.215) as:

$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + \sum_{i=2}^{p} ß_i \, \Delta y_{t-i+1} + \varepsilon_t. \tag{3.12}$$

Including p numbers of lags allows for performing the test for a higher-order autoregressive equation and handles the concern for autocorrelation. To decide the optimal lag length the researcher needs to examine different methods of information criterion which is further described in the next section. To determine a critical value, calculate a t-value and state the conclusion of stationarity in the series, the ADF-test uses similar techniques as the beforementioned.

3.1.5 Information Criterion

Another important preliminary step in building a VAR-model is to select the optimal length of lagged values in the model. There are various information criterions that represents the tradeoff between achieving the best fitted model and losing degrees of freedom. Examples of such are the AIC (Akaike IC), HQ (Hannan & Quinn IC), SC (Schwartz IC), FPE (Akaike´s Final Prediction Error). These methods run a VAR-model for different numbers of lags and reports the lag length that minimizes the information criterions "value".

When there is conflicting results, Ivanov and Killian (2001) states that the AIC-method is more beneficial on monthly data, while HQ and SC is more beneficial on quarterly data. The FPE method was not mentioned in this article. The number of lags suggested increases as the frequency of the data increases. We deduce from this that the AIC-method generally is better with a higher frequency in the data. Thus, being the most advantageous method on daily data.

Akaike Information Criterion (AIC) method is illustrated with following formula (Enders, 2010, pp.317):

$$AIC = T * \ln(sum\ of\ squared\ residuals) + 2n, \qquad 3.13$$

where n is numbers of parameters estimated, and T is number of usable observations. The method will select the model that returns the lowest AIC value and note the associated numbers of lags.

3.1.6 Stability-test

Stability tests are used in order to detect structural changes in linear regression relationships. By fitting a model to the given data, the generalized fluctuation test derives an empirical process capturing the fluctuation in either the residuals or in the estimates. Boundaries to determine structural breaks can be computed as the limiting process of these empirical processes are known. The fluctuation is improbable if the path of the empirical process crosses these boundaries. Hence the null hypothesis of no structural breaks should be rejected (at significance level α).

Empirical fluctuation process (The CUSUM process)

The CUSUM process contains the cumulative sum of standardized residuals. As suggested by Ploberger, Krämer (1992), the structural cumulative change test can be based on the common OLS residuals. The OLS-CUSUM empirical fluctuation process is given by

$$W_n^0(t) = \frac{1}{\hat{\sigma}\sqrt{n}} \sum_{i=1}^{|nt|} \hat{\mu}_i \qquad (0 \le t \le 1), \qquad 3.14$$

where $\hat{\sigma}$ is the estimated standard deviation, n is the number of observations, $\hat{\mu}_i$ is the estimated residuals and t is a random variable given the constraints. The limiting process for $W_n^0(t)$ is the standard Brownian bridge.

Boundaries

The null hypothesis of no structural change should be rejected when the fluctuation of the empirical process *efp(t)* becomes improbably large compared to the fluctuation of the limiting process. For the residual-based process the comparison is executed by a boundary *b(t)*, that is crossed by the limiting process by a probability α. Hence, if either *b(t)* or *-b(t)* is crossed by *efp(t)* for any t, the conclusion should be that the fluctuation is improbably large, and the null hypothesis should be rejected at confidence level α.

The previously mentioned limiting process of the Brownian bridge is not stationary. Therefore, the CUSUM processes use boundaries that are proportional to the standard deviation function of this process, and is given by:

$$b(t) = \lambda * \sqrt{t(1-t)}, \qquad 3.15$$

where the confidence level is represented by λ.

3.2 Framework of the analysis

3.2.1 Granger Causality

Granger-causality was introduced by Granger (1969) as an attempt to define *real-causality*. Granger's method was further modified by Sims (1972) where causality was measured by examining prior values of one time series and including them in an AR model with prior values of the other time series.

$$y_t = \alpha_0 + ß_1 y_{t-1} + \cdots + ß_p y_{t-p}. \qquad 3.16$$

Equation 3.16 illustrates an AR(p) model including p-lags to predict value of y at time t by the past values of y. The intuition behind the Granger-causality test is to examine if this AR-model perform better in explaining the value of y at time t, including other explanatory (lagged) variables into the equation.

$$y_t = \alpha_0 + ß_1 y_{t-1} + \cdots + ß_p y_{t-p} + \delta_1 x_{t-1} + \cdots + \delta_p x_{t-p}. \tag{3.17}$$

When including the variable *x* to equation, one can use traditional statistical testing (t-test, F-test) to examine if the model 3.17 gives a better explanatory power than model 3.16. If the explanatory power becomes larger, δ coefficient is jointly significant; Granger argues that *x* "Granger causes" *y* at time *t*.

Null hypothesis: *x* ratio does not granger cause *y* at time *t*.

$$H0 = \delta_1 = \delta_2 = \ldots = \delta_p = 0 \tag{3.18}$$

$$H1 = At\ least\ one\ \delta_p \neq 0\,. \tag{3.19}$$

Similarly, the granger causality can be tested the other way around in a VAR-model.
One can examine if *y* "granger causes" *x*.

Critics of Granger Causality

Critics of granger causality has pointed at the fact that the method is not really estimating the true causality. *Precedence* has been proposed as an alternative expression to (Granger) *causality*. Therefore, the results of this method need to be handled with caution. This is a highly theoretical approach in order to find relations in economic data, which might differ from the real economic world.

Given the granger causality method, the researcher cannot be fully certain that one time series causes the other, however the usage of the method helps predicting the desired time series. To further analyze the results, one can present a correlation matrix to assess the strength of the correlation in the lagged values as a control for the granger causality results.

3.2.2 Impulsive Responsive Function (IRF)

The impulsive responsive function is a practical way to study the responsiveness of one dependent variable to a changes in error terms in the VAR-model. The coefficients of a VAR-model are generally hard to interpret, therefore impulsive responsive function is a helpful method to visualize the impact one variables has in forecasting another. Enders (2010, pp.307-308) illustrates this considering the following VAR-model:

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}. \qquad 3.20$$

Then, the VAR-model is redefined into the terms of the $\{\varepsilon_{1t}\}$ and $\{\varepsilon_{2t}\}$ sequences:

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{yt-i} \\ \varepsilon_{zt-i} \end{bmatrix}. \qquad 3.21$$

Equation 3.21 examines the interaction between y_t and z_t, where the ϕ coefficients note how a one-unit shock in the error term affect the entire time paths of y_t and z_t. These four coefficients, $\phi_{11}(i), \phi_{12}(i), \phi_{21}(i), \phi_{22}(i)$, are called the impulsive response functions. For instance, coefficient $\phi_{12}(0)$ notes an instantaneous impact in ε_{zt} on y_t. In the same way, the coefficients $\phi_{21}(1), \phi_{22}(1)$ are one-period responses of unit changes in ε_{yt-1} and ε_{zt-1} on z_t.

3.2.3 Forecast error variance decomposition

The forecast error variance decomposition can tell what fraction of the movement in the variables are due to their own shocks, versus shocks from other variables in the VAR-system. If at all horizons, a shock in ε_{zt} explain none of forecast error variance of y_t, one can conclude that the y_t sequence is exogenous, and vice versa. In this instance, the y_t sequence moves independently of shocks in the z_t sequence.

Enders (2010, pp.313-314) denotes the n-period forecast error as

$$x_{t+n} - E_t x_{t+n} = \sum_{i=0}^{n-1} \phi_i \, \varepsilon_{t+n-i}. \qquad 3.22$$

If only the y_t sequence is considered, the n-step-ahead forecast error is given by

$$\begin{aligned} y_{t+n} - E_t y_{t+n} = {} & \phi_{11}(0)\varepsilon_{yt+n} + \phi_{11}(1)\varepsilon_{yt+n-1} + \ldots + \phi_{11}(n-1)\varepsilon_{yt+1} \\ & + \phi_{12}(0)\varepsilon_{zt+n} + \phi_{12}(1)\varepsilon_{zt+n-1} + \ldots + \phi_{11}(n-1)\varepsilon_{zt+1}. \end{aligned} \qquad 3.23$$

When the forecast error variance of y_{t+n} is denoted as $\sigma_y(n)^2$ it gives

$$\sigma_y(n)^2 = \sigma_y^2[\phi_{11}(0)^2 + \phi_{11}(1)^2 + \ldots + \phi_{11}(n-1)^2] \\ + \sigma_z^2[\phi_{12}(0)^2 + \phi_{12}(1)^2 + \ldots + \phi_{12}(n-1)^2]. \qquad 3.24$$

The forecast error increases as the forecast horizon n increases as a result of all $\phi_{jk}(i)^2$ necessarily being nonnegative. Finally, the fraction of $\sigma_y(n)^2$ caused by shocks in ε_{yt} and ε_{zt} is expressed by respectively

$$\frac{\sigma_y^2[\phi_{11}(0)^2 + \phi_{11}(1)^2 + \ldots + \phi_{11}(n-1)^2]}{\sigma_y(n)^2} \qquad 3.25$$

and

$$\frac{\sigma_z^2[\phi_{12}(0)^2 + \phi_{12}(1)^2 + \ldots + \phi_{12}(n-1)^2]}{\sigma_y(n)^2}. \qquad 3.26$$

Enders (2010, pp.314) states that in applied research it is normal for a variable to explain close to all of its forecast errors in the short term, and a smaller proportion at a longer horizon. This is expected if ε_{zt} had little effect on y_t but affected the y_t sequence with a lag.

4. Data

4.1 Data collection

Most of the data collected is publicly available information that does not require any additional access. Data such as notations of traded securities like the S&P500 index, gold and platinum prices are obtained directly from Yahoo Finance. The CBOE volatility index is also available for the public through such notations. All data regarding interest rates are retrieved directly from the Federal Reserve Economic Database (FRED). Access to the options data and the U.S. dollar index is gained through a Bloomberg Terminal.

4.2 Data preparation

This section will provide a review of the data preparation performed to obtain the necessary requirements as stated in the methodology section. The variables used in this thesis were either returns, ratios, spreads or dummies.

A vector autoregressive model requires all data to be stationary. In order to achieve this, a computation of the percentage change was done for all the indicators by the use of this formula:

$$\frac{x_t - x_{t-1}}{x_{t-1}}. \qquad 4.1$$

Data in economic analysis are often denoted in this term, and therefore makes the interpretation of the analysis easier. Even though some of the indicators such as ratios and spreads were stationary originally, a calculation of percentage change was done, to check whether the predictability of these indicators was affected.

A dummy is often included in a model to combat structural breaks which may occur in a time series. This technique is an advantageous tool to meet the stability requirements stated in the methodology section. We created the dummy *crisis* with two secondary dummies to control for such structural breaks during the dataset. By comparing the moving 52 week high of the S&P500 index to each day's level * 0,8. Meaning if the market is lower- or equal to 80% of the previous 52 week high. If this is true, the dummy *bear* returns 1. This is indicating a bear market where the volatility is high. The second argument is indicating the start of a new bull market. If the current level of the S&P500 index is higher than the moving 52 week low *1,2, the *bull* dummy is 1. Normally, the volatility is lower in such a phase. The crisis dummy returns 1 when bear = 1 and bull = 0. This results in a variable indicating times of high volatility, correcting for structural breaks when this occurs in the stock market.

The VWMA of the S&P500 index was calculated using the formula presented in chapter 2.3.5. Different lengths of moving averages were calculated and tried, ranging from 10-day to 100-day. The moving averages from the lower part of the range visualize a short-term trend of the index, where the longer moving averages are less responsive to changes in the underlying and visualize a longer-term trend.

The VWMAs were subtracted from the level of the S&P500 index to include it without violating the stationarity criterion. Resulting in a variable showing the deviation from the level of the index, which is stationary due to mean reversion. If choosing a moving average in the lower part of the range presented, one would expect the mean reversion to occur faster, which could be beneficial as the simulation period in this analysis is rather short. Due to violations of the VAR-model criterions, shorter than 50-day moving averages proved impossible. Nevertheless, a 50-day moving average is still relatively short but may complicate the interpretation of the impulse response analysis.

This is the only indicator in the analysis not denoted as a percentage change. The assumption is that the actual deviation of the VWMA from the level of the S&P500 has a higher predictive power than the percentage change of it. When the deviation is high, we are likely to see a reversion in the nearest future. By using the actual deviation, we ensure that the model treats a 100% increase in the deviation from 1 to 2 different from a 100% increase from 30 to 60, as the likelihood of a reversion following these examples are not the same.

4.2.1 Considerations regarding the indicator inclusion

To satisfy the VAR-model requirements, some of the indicators mentioned in chapter two were excluded from the model. Nevertheless, at least one indicator from each financial market were included. The indicators *Option to Stock* and the *Yield spread* are not included in the regression model. These indicators created structural breaks in the model, which could not be corrected for. Therefore, they were excluded to fulfill the stability criterion. The yield spread also failed the Granger-causality test.

The indicators included in the model is presented in the Table 2 below.

Summary statistics of daily data in the period of 2005-2020

	Mean	Standard Deviation	Minimum	Maximum	Observations
Return of the S&P500 index	0.0003	0.0123	-0.1198	0.1079	3,992.0000
Percentage change in the Put/Call ratio	0.0120	0.1537	-0.5537	1.1228	3,992.0000
Percentage change in the Gold/Platinum ratio	0.0004	0.0117	-0.0561	0.1343	3,992.0000
Percentage change in the VIX	0.0031	0.0792	-0.2957	1.1560	3,992.0000
Percentage change in the U.S. dollar index	0.0000	0.0048	-0.0269	0.0255	3,992.0000
Percentage change in the credit spread Baa- Aaa	0.0003	0.0211	-0.1890	0.4563	3,992.0000
Crisis dummy	0.0376	0.1902	0.0000	1.0000	3,992.0000
Deviation of Volume Weighted Moving Average	17.4064	74.1749	-743.261	370.7127	3,992.0000

Table 2: Summary statistics table (2005-2020)

Correlations in the processed data

Table 6 in the Appendix 9.1.2 notes the correlation between the indicators included in the model. The correlation coefficients differ from table 1 as the indicators in the model is calculated at a percentage change. In table 1 the indicators are unprocessed data. The majority of the correlation coefficients in the matrix are close to zero. An outcome like this is expected as the correlation is based on returns and subsequently making high correlation coefficients unlikely. The S&P500 index is included in most coefficient that can be considered rather high. Interestingly, some indicators correlation coefficients shift from positive to negative as this calculation is done. Exemplified by the return of the Gold/Platinum ratio to the return of S&P500 index.

As previously mentioned, if independent indicators are highly correlated, they may be explaining the same relationship, causing multicollinearity. Consequently, it may be beneficial to exclude one of the indicators. When excluding the S&P500 index from the correlation matrix, only one coefficient stands out. The return of the Put/Call ratio and the return of the VIX has the highest correlation of 0.316. Both represent the option market, but the correlation is not considered to be high enough to omit any of the indicators. Thus, no omissions are necessary based on the correlation between the indicators.

4.2.2 Subsets of the dataperiod

Based on the data obtained, the maximum period possible would be 1997-2020. In order to satisfy the stability criterion, we could go no further back than 2005. The structural breaks in the dataset prior to 2005 was not corrected for with the usage of a dummy variable. Therefore, the sample period starts in 2005.

To analyze the development in the importance of the variables through time, the data and model was divided into multiple periods. One with the whole sample period of 2005-2020, in addition to a model per each subset of 5 years within this period, initially. In the period of 2010-2015 there were no crises according to the crisis dummy, meaning it never deviated from zero. This resulted in singularities in the regression, invalidating this period. Therefore, the five-year period was not possible and was extended to 2010-2020. The approach of creating the model of each subset is identical. The only differentiating factor is the time span considered.

Summary statistics of the periodic subsets are presented in the appendix, section 9.2.1, table 7-9.

5. Results

This section will first discuss the different regression models created and the changes observed throughout the subsets, as well as provide possible explanations grounded in an economic rationale. As previously mentioned, the coefficients of a VAR model are many and hard to interpret separately. Therefore, we will focus on the Impulsive Reponses Functions for the relevant indicators and reflect upon the results. An example of the VAR-system is given in appendix 9.2.2 for the whole sample period (2005-2020) denoted in mathematical terms. Additionally, the results are compared to the results of partial models where the same approach is used with each explanatory variable, individually. A forecast error variance decomposition is presented at the end.

5.1 The regression models

As previously revealed, we created multiple regression models of subsets within the dataset. If every variable satisfies the different criterions of a VAR model, TRUE is returned. Otherwise, FALSE is returned. A summary of the tests required is visualized in Table 3[4].

Summary of the regression models

	2005-2020	2005-2010	2010-2020	2015-2020
Stationarity of in-data	TRUE	TRUE	TRUE	TRUE
Number of lags	26	28	22	18
Stabil model	TRUE	TRUE	TRUE	TRUE
Granger Causal	TRUE	TRUE	TRUE	FALSE
Adjusted R^2	0.120	0.215	0.181	0.284

Table 3: Summary of regression models

[4] All of the respective tests are presented for each indicator individually in the appendix in section 9.2.3/9.2.4

Stationarity

The stationarity tests are presented in table 10 in the appendix 9.2.3. Some of the indicators were stationary unprocessed, and otherwise the calculation of the percentage difference ensured a stationary variable. There was only one occasion where this criterion was of concern, when we initially planned a subset of 2010-2015 rather than 2010-2020. In the five-year period from 2010, the dummy variable created never changed from 0. Thus, the stationarity could not be calculated.

Number of lags

As the different methods of computing optimal lag length reported conflicting lag length, we utilized the AIC-method following the Ivanov and Kilian (2001) article considering it the better fit with higher frequency data. To control for this, we compared the information criterion value of the four different methods. All resulted in the AIC-method returning the lowest value, indicating it being the best methods to use for our data.

When computing the optimal number of lags, a maximum number must be set for the calculation. Most of the indicators were hypothesized to impact the S&P500 index rather quickly, while the VWMA was thought to benefit from a higher number of lags as a longer period is often needed for the mean reversion to occur. Considering this difference, we decided to limit the maximum the number of lags at 50 to ensure that the VWMA were not given too much emphasis, as our intent was to examine the predictive abilities of the indicators jointly.

The optimal number of lags recommended by the AIC method was used, with an exemption made in the period of 2010-2020, due to stability concerns. The recommended amount was 18 and not 22 as presented in table 3.

Stability

The reason for deviating from the recommended number of lags in the beforementioned period was due to a structural break, which was avoided by adding four more lags. The stability criterion was also the one breached whenever we attempted VWMAs based on shorter periods.

As previously mentioned, some of the unprocessed indicators were stationary. However, when they were utilized unprocessed, they often resulted in a breach of stability. Subsequently, no unprocessed indicators were used. The stability of the models (at 95% confidence interval) is visualized in figures 16-19 in the appendix 9.2.5.

Granger causality

The granger causality tests are presented in table 11 in the appendix 9.2.4. Each indicator passes this test except once in a subset. In the period of 2015-2020, the percentage change in the U.S. dollar index is no longer granger causal, resulting in the summary reporting FALSE on granger causality. Changes to this indicator and a possible reason for this observation is discussed later in 5.2.3.

Even though the tests claim causality it is important to be conscious of the difference between “real” causality and estimated causality. One cannot draw definite conclusions of true causality using these methods. Even if the results are corrected with respect to econometric criterions there is always a likelihood of the real-world context to differ from this. To account for this distinction, we discuss the results of the impulsive response analysis through economic theory and economic sense and present possible reasons for the observed changes, derived from changes in the economic environment.

In section 3.2.1 granger causality, we mentioned that the critics of the granger causality argued for another method to strengthen the causality argument, by an additional consideration of the correlation coefficient between the indicators and the S&P500 index. In the Appendix, section 9.1.1, table 5, the correlations of the percentage change in the indicators to the return of the S&P500 index are presented in a matrix ranging from t-7 (leading) to t+7 (lagging) in the whole sample period (2005-2020). However, the usefulness of the correlation coefficients is limited as the coefficients appear to be spurious and for the most part close to zero. This is reasonable as they are based on the daily returns and high correlations are unlikely as a result. Consequently, we were not able to take advantage of this method to strengthen any claims of causality, strengthening the importance of the beforementioned economic reasoning.

Explanatory rate

Interestingly, the explanatory rate of the models (Adjusted R^2) increases through time, towards the present day. The whole period gives an explanatory rate of 12,0%, while the first and last subset gives respectively 21,5% and 28,4%. It is also noteworthy that this increase comes with a decline in the number of lags. With fewer lags, a lower explanatory rate is expected, ceteris paribus. Given this, there is reason to believe that the predictability of the indicators included in the model has grown from 2005 through 2020. Consequently, highlighting their combined rising importance.

The crisis dummy signifying times of high volatility seems to have a huge importance in the model. Whenever it is excluded, the optimal number of lags gets lower (to about half), resulting in a significantly lower explanatory rate. It can be exemplified by the period of 2015-2020. If the crisis dummy is excluded the optimal lag length is 9, giving an explanatory rate of 19.1%. The rate increases to 23.9% if the indicator is included keeping the same lag length (9). When the optimal lag length including the indicator is selected (18), the explanatory rate further improves to 28.4%. The optimal number of lags increases when the indicator is included because there is information in the indicator in the days between $_{t-9}$ and $_{t-18}$ where the reward of increasing the number of lags outweighs the penalty of the additional parameters.

5.2 Impulsive Response Functions (IRF)

This section will present the results of the VAR-model with an illustration of how the return of the S&P500 index responds to shocks to the error term of each variable included in the models. The unit-root shocks are obtained from the regression and presented in the tables with the respective indicator. The impulsive response function estimates the mean return for the following days based on 100 simulations, with a confidence interval chosen to be 95%. When the confidence interval is significantly deviating from zero, one can conclude the direction of the market return as a result of this shock. Otherwise, a direction of the stock index cannot be stated based on the IRFs. To simplify the interpretation of this part of the analysis insignificant observations are considered zero. In the real world, however, this is not true. Implications from this is further discussed in section 6.1, where we discuss the practical utilization of the results.

For each variable being shocked, four impulsive response functions are presented, one for the whole sample period, in addition to the three subsets. It is important to be attentive to the changes in the Y-axis of the figures. The patterns often look very similar, but the changes are often larger than what is initially perceived.

The impulsive response functions of the VIX did not return any significant results in any of the periods, meaning at no simulated day could a conclusion about the direction of the stock index be reached. This may indicate that the VIX is more a coinciding indicator rather than a leading one. Therefore, this indicator as well as the dummy variable crisis will not be illustrated in the following section. As their contribution improved the accuracy of the models, they were included despite this.

5.2.1 The Gold/Platinum ratio

When modelling the whole period, it shows a significant positive return in the S&P500 index three days after the increase in the ratio. The following period from 2005-2010 shows a slightly significant positive return in day three. The two last periods illustrate a negative return in the second and the fifth day, with the last period being substantially more negative.

Unit root shock

	2005-2020	2005-2010	2010-2020	2015-2020
Percentage change in the Gold/Platinum ratio	0.01119	0.01227	0.01000	0.01090

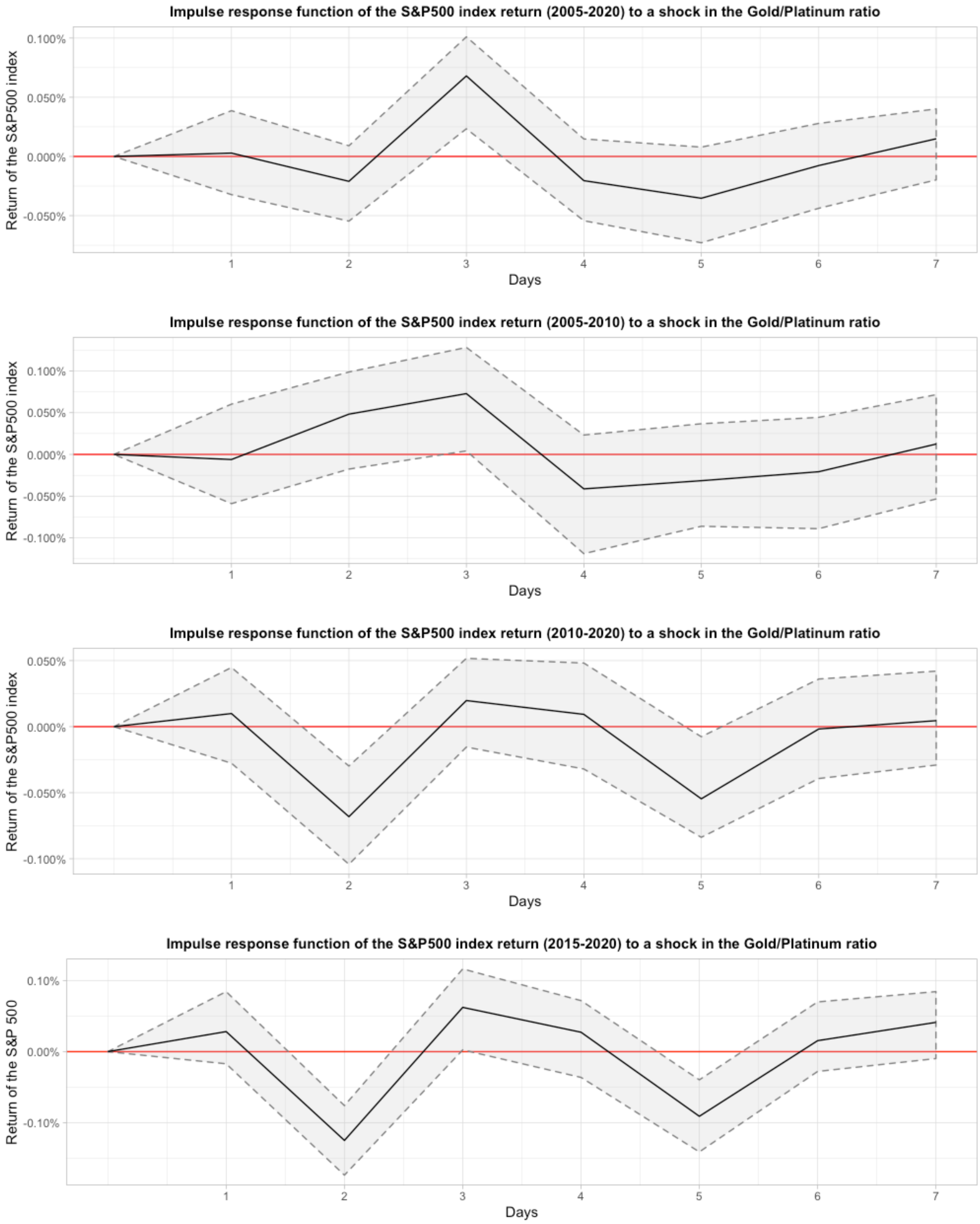


Figure 9: IRF of the S&P500 index return to a shock in the Gold/Platinum ratio

The IRFs are visualizing that the reaction in the S&P500 index following a shock in the ratio are inconsistent through time. The two latter subperiods (2010-2020 and 2015-2020) appear to be fairly similar, whereas the first subperiod (2005-2010) displays a different pattern. As the whole period is considered, the earliest subperiod (2005-2010) seems to be of most importance, as the only remaining significant day is the one from that subset and the other loses their significance. The inconsistency makes it hard to apply the information from the IRFs directly in decision making, whereas there does not seem to be any pronounced changes in the financial markets explaining it.

One of the major findings of Huang and Kilic (2019) were an increase in US stock return of 6,4% the following year after a standard deviation shock to the Gold/Platinum ratio. A similar conclusion of positive returns cannot be drawn based on the impulsive response functions above. While the article investigated the results the following year, this thesis examines data at a shorter time horizon, which may complicate the comparison. From an economic perspective, where the Gold/Platinum ratio is used as a proxy for the risk tolerance of investors, an increase in this ratio would result in an expectation of lower returns the following days as investors gets more risk averse. In the periods of 2010-2020 and 2015-2020 the impulsive response functions directions reflect this rationale.

The mechanism of how the ratio changes

Conventional wisdom often refers to gold as a hedge in times of distress. Implying that the increase in the Gold/Platinum ratio is a result of higher gold prices. However, as Huang and Kilic (2019) questioned; gold is not really a hedge to the stock market, as prices do not go up during bad times. This is visualized in the figure 10, created from the original data, showing that the gold price normally stays flat or decrease as uncertainty enters the markets. However, gold does not lose as much value as the stock index.

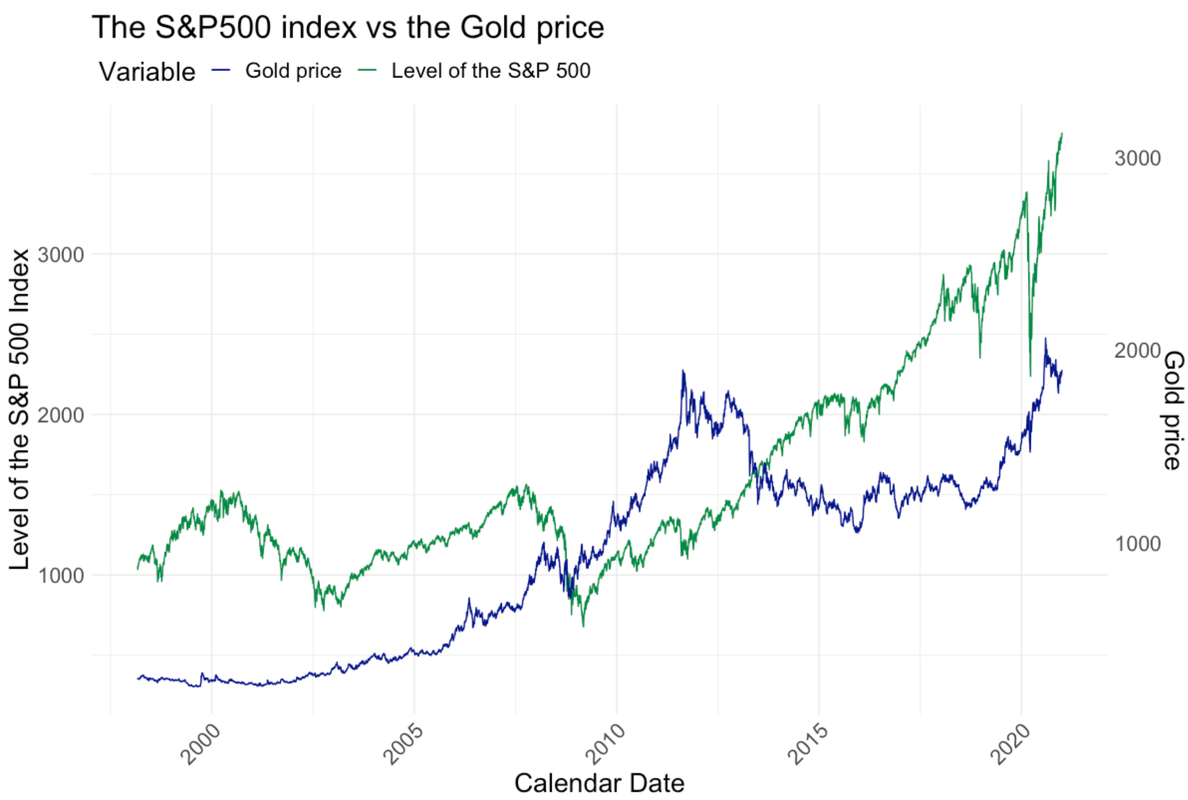


Figure 10: The S&P500 vs the Gold price (1998-2020)

The increase in the Gold/Platinum ratio in uncertain times is a result of gold prices falling less than platinum prices. Due to a fall in consumption whenever uncertainty enters the market, both gold and platinum prices are likely to drop. But as gold has an additional use of storing wealth it is likely to keep more of its value relative to platinum.

5.2.2 The Put/Call ratio

A significant change in the predictive ability of the ratio through time is shown. Within the period of 2005-2010 one cannot state that the impact is either positive or negative. In the two later periods (2010-2020, 2015-2020), a negative impact can be stated, as for the whole sample. The whole sample period of 2005-2020 also detect a significant negative return in day four, whereas the subperiod of 2010-2020 implies a corrective pattern with significant positive returns in the second day.

Unit root shock

	2005-2020	2005-2010	2010-2020	2015-2020
Percentage change in the Put/Call ratio	0.13293	0.13901	0.13008	0.13009

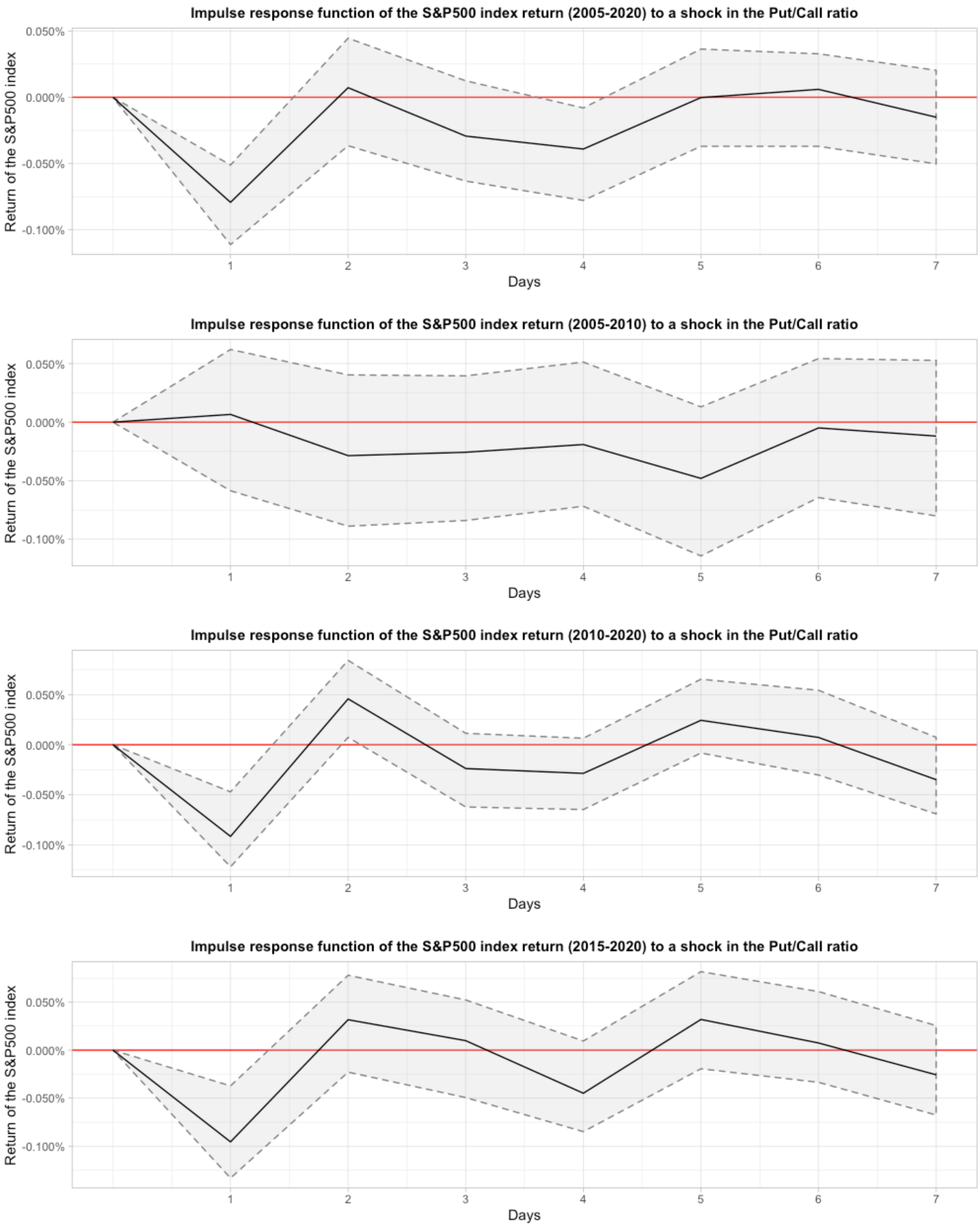


Figure 11: IRF of the S&P500 index return to a shock in the Put/Call ratio

Excluding the earliest subset (2005-2010), the IRFs illustrates rather consistent reactions in the S&P500 index following the increased risk aversion. The fact that there appear to be a consistent negative reaction the following day makes this information useful in investing. An increase in the ratio is a sign that one may consider selling the market. Such information can probably also be applied as the ratio experience a significant withdrawal, indicating that it may be a good time to buy.

These results are complying to Pan and Poteshmans conclusion where an increased Put/Call ratio relates to negative returns. They were able to create excess returns by selling stocks with high Put/Call ratios and buying stocks with a low ratio. The IRFs show the effect of a shock in the Put/Call ratio to be negative, leading to negative returns of the S&P500 index the following day, which also seems to get increasingly negative in the recent years.

The equity options market

According to the IRFs there has been an increased predictability of the Put/Call ratio in recent years. A possible reason can be that the US equity options market has grown significantly through the same period, visualized in figure 12. The figure includes indexes representing respectively the total number of calls and puts traded on the Opra exchanges, retrieved from Bloomberg. Exchanges included are the Amex, Philadelphia, Pacific, NYSE, CBOE, ISE, Boston, Nasdaq, Bats, C2, Nasdaq OMX BX and Miami. The options are also added together to show the total volume. Figure 12 is an original created from this data.

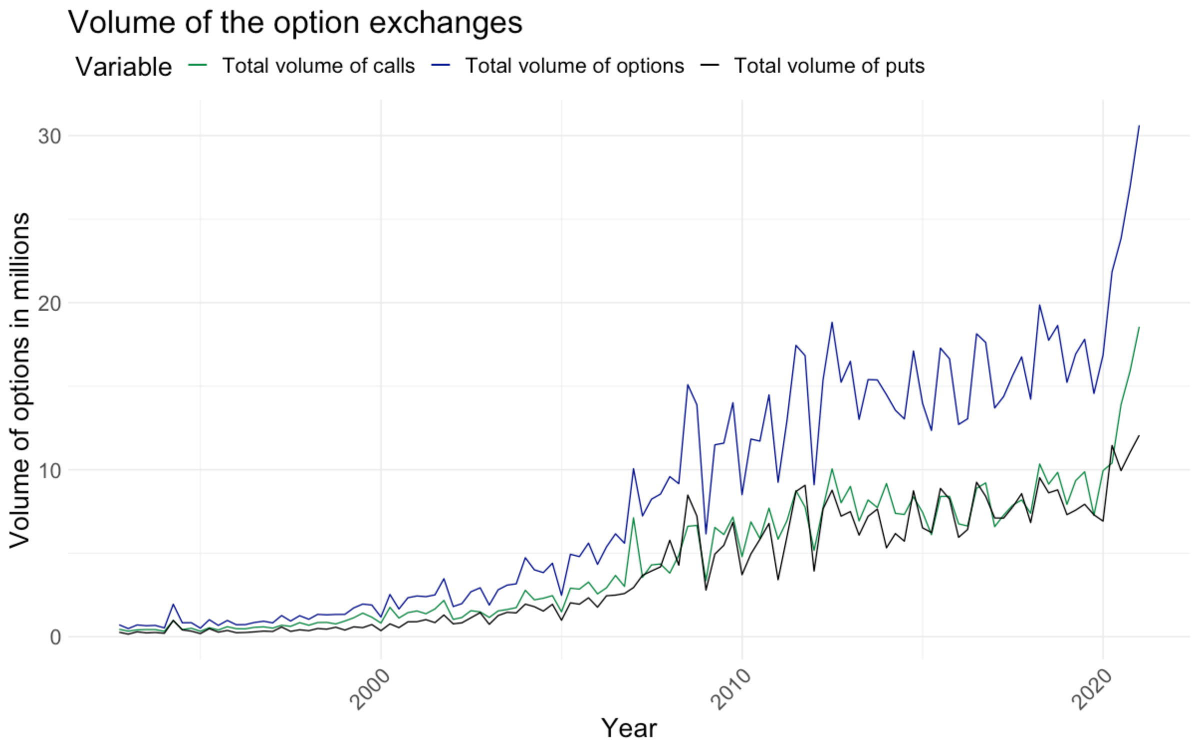


Figure 12: Options volume of the Opra exchanges (1993-2020)

In the early 2000s the options volume was somewhat stable at approximately 2.5 – 4 million contracts. A trend change in the volume is visible around 2005, coinciding with the start of our sample. There is also a huge spike in 2020 one cannot fail to notice with a 50% increase in the total options volume. It is reasonable to assume that an increase of that magnitude will have a positive effect on the predictive power of the options market, and hereby improve the model when the period is constrained towards 2020.

Possible reasons for the increasing volume in the options market

There are several possible reasons for the large increase in options trading volume in recent years. According to Tenn (2020) here has been a significant increase in retail participation in both the stock market and the options market, especially from 2020 and forth, which might explain the extraordinary spike and the record setting trading volume. Multiple factors have contributed, like lockdowns, stimuli checks, zero-commission trading and the readily available trading platforms.

Buying options are a good alternative when investing short term as it requires less capital with an opportunity of higher returns. From figure 12, we can see an exponential increase in the sale of options. Mostly coming from a huge escalation in the call volume, which entails that investors overall had a very bullish view on a relatively short term. The put volume also increased but not nearly to the same extent. This is also the first time the options volume deviate considerably from each other, where they appear to have followed the same trend previously.

Does option selling have a direct impact on the price of the underlying asset?
Options are often sold by market makers; whose function is to provide liquidity to the market. Generally, they do not want to take active bets of stocks going in either direction. This is avoided by hedging the options they are selling. Mainly, this is done by being delta neutral[5]. This hedges against small movements in the underlying assets price. As the delta of the option varies with the price of the underlying asset, so does the requirements to the holdings of the market makers in regard to staying hedged. For example, to stay delta neutral with the sale of a call option, the market makers have to buy the underlying asset when the price of it increases. Thus, the delta hedging done by the market makers reinforce the ongoing momentum of the asset. The opening of new options positions might therefore impact the price directly, where the seller of the options will either buy or sell the underlying asset depending on the option. In order to stay delta neutral with the sale of one call option contract with a delta of 0.50, the seller must buy the equivalent of 50 shares of the underlying asset $(0.50\,(delta) * 1(contract) * 100(shares\ per\ contract) = 50)$. This substantiates the argument of the options predictive power as they may directly affect the price of the asset.

5.2.3 The currency effect (U.S. dollar index)

The IRF shows a significant negative return of S&P500 index in day two and six studying the whole sample period. The following subperiod (2005-2010) the function remains significant in the same days, but the effect is shown to be even more negative. The remaining periods has no significant observations during the first seven days after a shock.

[5] See appendix for more information about this strategy.

Unit root shock

	2005-2020	2005-2010	2010-2020	2015-2020
Percentage change in the U.S. dollar index	0.00472	0.00516	0.00443	0.00422

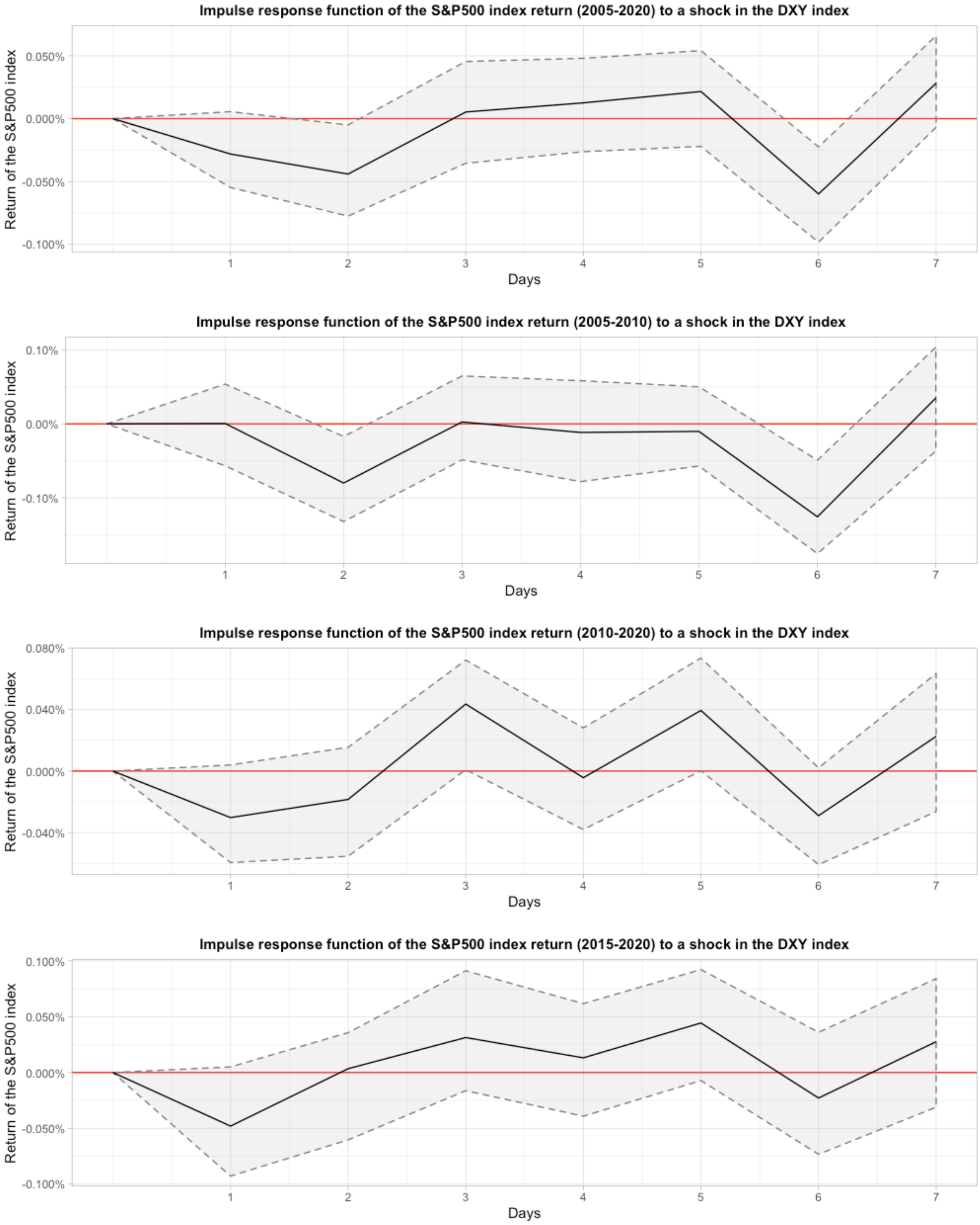


Figure 13: IRF of the S&P500 index return to a shock in the U.S. dollar index

In uncertainty reveals itself, investors tend to prefer currencies unphased by the turmoil. As the current reserve currency, the dollar is attractive to investors in this environment, resulting in a shift in demand. This will lead to a strengthening of the currency, ceteris paribus. A significant reaction following a shock in the U.S. dollar index could signify such conditions being present. The IRFs indicates a significant negative return of the S&P500 index for the whole sample period, as well as the in the first subset. Thus, supporting the hypothesis that a sudden increase in the relative strength of the dollar caused by uncertainty leads to negative returns in the stock market. Even though the whole period still has significant days, the expected returns on those days are cut to about half, compared to the earliest subset (2005-2010).

The variability the IRFs are visualizing through time is not unexpected in this indicator. Interest rates have been trending downwards for the last several decades. The effect of central banks lowering interest rates is diminishing the closer it gets to zero. Hence, they perform other policies to stimulate the economy. To combat this issued Federal Reserve has expanded their balance sheet by quantitative easing, resulting in a higher supply of the dollar. The expansion of the balance sheet started during the financial crisis and have increased steadily till the corona crisis, where the amount of printed money increased significantly[6]. When interpreting the IRFs of the last two subsets it is reasonable to believe that this intervention in the market might affect the patterns that had previously emerged. Clear patterns are hard to establish as such an increase of the supply is bound to have an effect on the price of the dollar. Quantitative easing is also thought to be the reason why the U.S. dollar index is no longer granger causal in the latest subperiod modelled (2015-2020).

5.2.4 The credit spread

After a shock in the credit spread, there is a slightly significant observation in day five for the period of 2005-2020 and a significant, negative return in day two for 2005-2010 period. In the two last periods, the IRFs illustrates a negative return in S&P500 index in the third and fifth day for the ten-year period (2010-2020), but only day five remain significant in the last five-year period (2015-2020).

[6] The Federal Reserve has expanded their balance sheet extensively in recent years https://www.federalreserve.gov/monetarypolicy/bst_recenttrends.htm

Unit root shock

	2005-2020	2005-2010	2010-2020	2015-2020
Percentage change in the credit spread Baa- Aaa	0.01943	0.01595	0.01958	0.02067

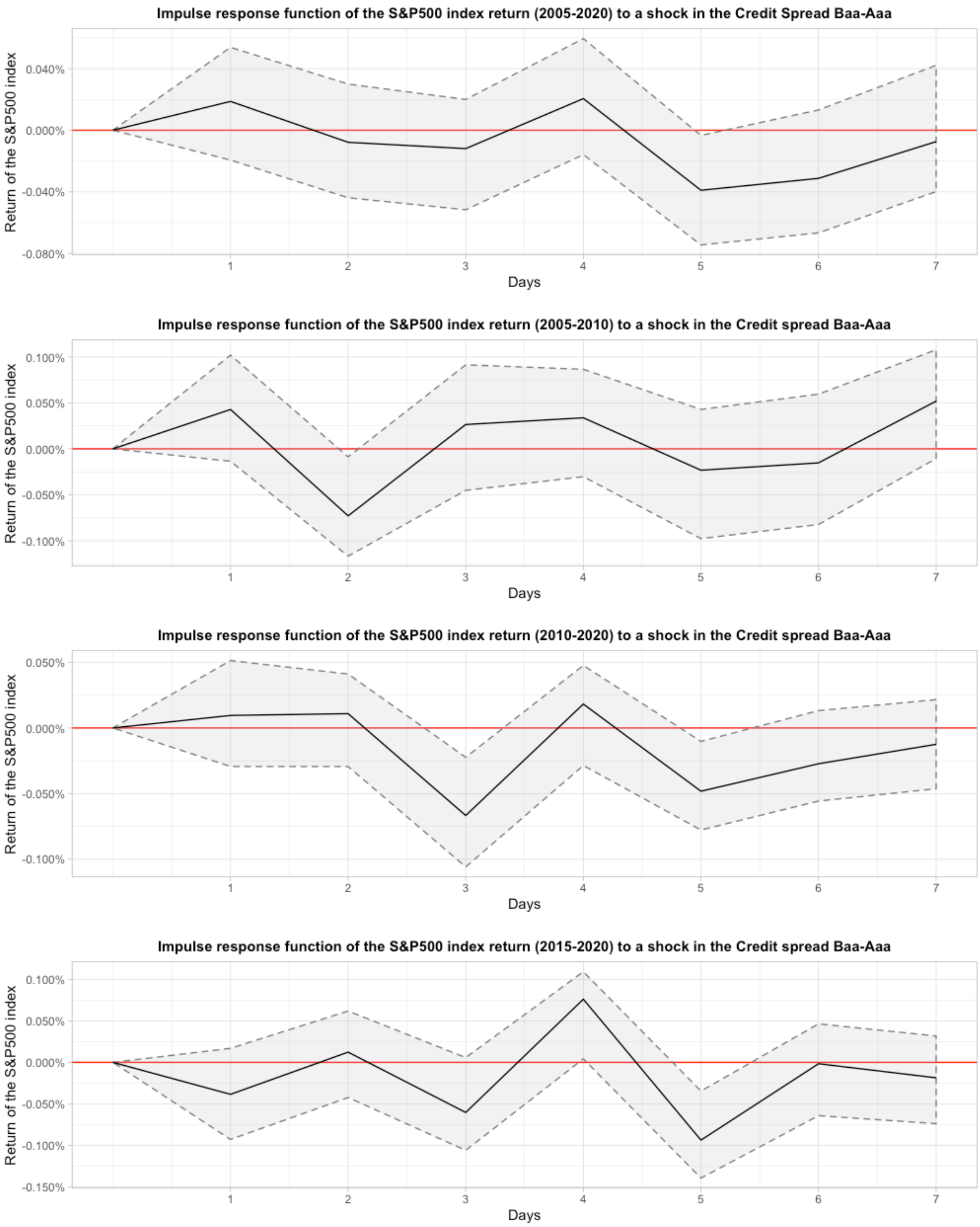


Figure 14: IRF of the S&P500 index return to a shock in the Credit spread Baa-Aaa

Since the credit spread is a proxy reflecting the willingness of investors to take on risks, it entails an expectation of negative returns following a significant increase in the credit spread. This is observed in different days across the subsets. Additionally, the last IRF (2015-2020) demonstrates a result in conflict to the negative return argument as there is a slightly significant positive return in the fourth day following the shock. Disregarding this one positive observation, the results appear to consistently be negative returns in the stock market following an increase in the credit spread.

Unlike the Put/Call ratio and the U.S. dollar index, there are no apparent reason why there would be a substantial change in the IRFs of the credit spread. The changes observed in this variable through time are minor and appears to revolve around the magnitude of the impact as well as the day of significant impact. This makes it hard to apply the information from the IRFs directly in an investment strategy. As most of the significant days are shows negative returns one can infer that there is an opposing relationship between the variables. However, one cannot be able to pinpoint the exact day of the reaction. Leading to an understanding that the variable can be used to a certain extent as a leading indicator.

5.2.5 Deviation from the Volume Weighted Moving Average

The IRFs of a shock in the deviation from the VWMA returns significant results for every sample period. Compared to the other indicators in the model, they are also significantly different from zero for a longer time-horizon. Whereas a longer time-horizon than seven days after the shock have not returned significant results, previously. For the VWMA, the period is extended to 12 days before the effect is no longer visible.

When considering the whole sample period, the impact of a shock shows a significant negative return the first day, a positive return in the second before it turns insignificant in the third and fourth day. From day 5 through 9 every other day is significantly positive and negative, before the remaining turns insignificant.

The beginning of the earliest subset (2005-2010) has similarities to the beforementioned, while the later subsets return insignificant results early in the simulated period. A recurring pattern throughout the IRFs is one where every other day is significantly positive and negative.

Unit root shock

	2005-2020	2005-2010	2010-2020	2015-2020
Deviation of the Volume Weighted Moving Average	20.25052	13.91466	21.27999	26.06056

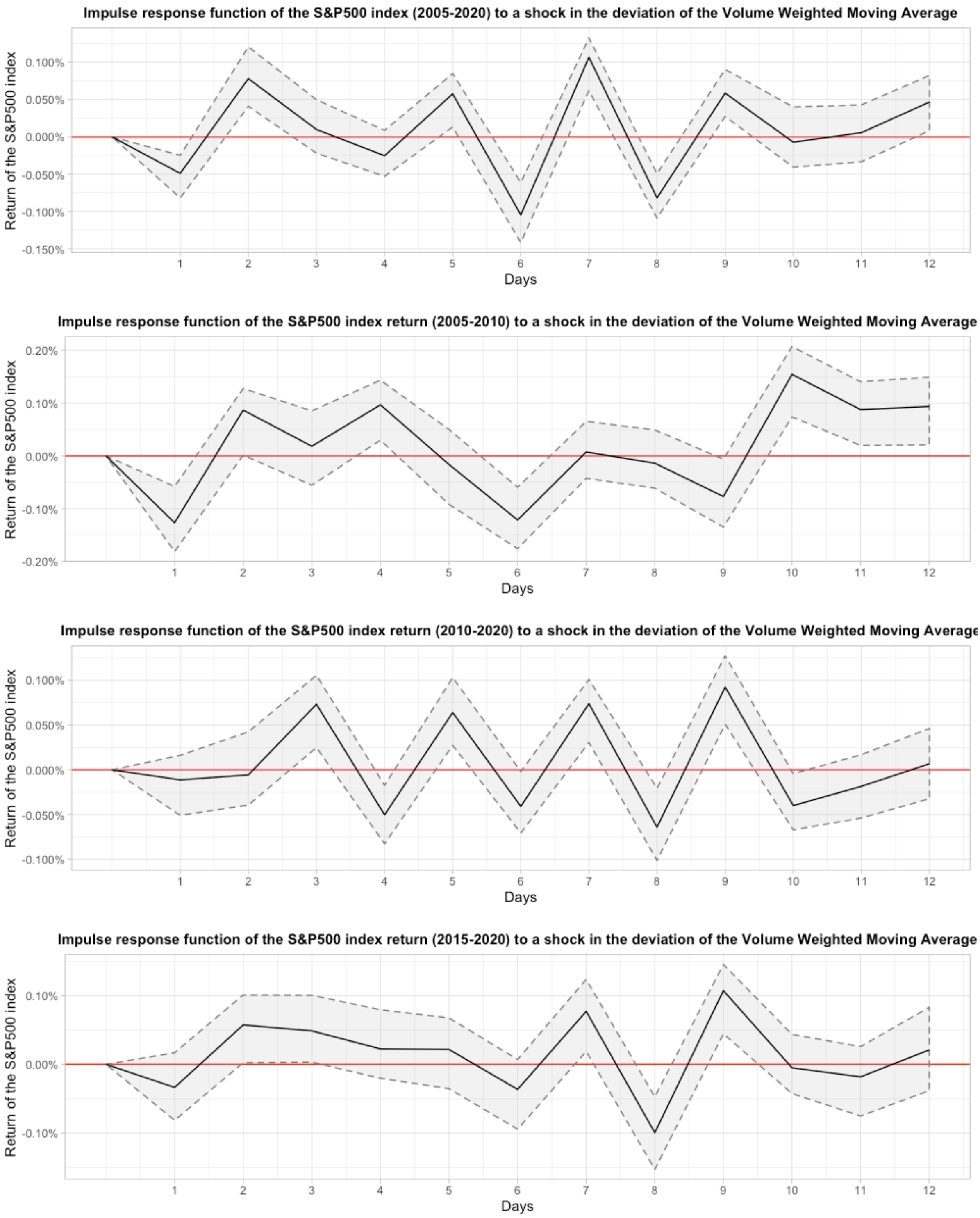


Figure 15: IRF of the S&P500 index return to a shock in the deviation of the Volume Weighted MA

Even though the first two IRFs shows the S&P500 index to have a negative return the following day from a shock which is aligned with the expectation of mean reversion, the same effect is not visible in the later periods. Nor are there any other visible patterns recurring to support any claim of mean reversion within the simulated period, as the reactions appear rather sporadic.

Nevertheless, there is no indication from our results implicating a possibility of refuting the hypothesis of predictive powers. The expectation is that the price of the S&P500 index corrects itself whenever the deviation from the VWMA is substantial for an extended period of time. That there are no visible patterns within the simulated period gives reason to believe that the correction normally takes longer than 12 days. This is very reasonable as the moving average is calculated on the basis of 50 days. Considering this, a conclusion regarding this matter cannot be stated from the IRFs.

5.2.6 Partial models comparison

Appendix 9.4 includes partial models where one-by-one explanatory indicators are included individually in a VAR-model with the S&P500 index return to study the effects isolated. It is of interest to compare the impulsive response functions from the main models to the restricted models to investigate if the main model increases the predictability when including multiple indicators in the VAR-system. This comparison will be based on the period of 2005-2020. The conclusion regarding the other periods is the same as this, and it would therefore be redundant to include all.

When the same approach is used in creating the restricted models, approximately similar patterns are seen from shocks performed in the indicators. In the IRF for the Put/Call ratio, the main model reports a negative return the first day at 0,075% for the S&P500 index, while the partial model reports a slightly more negative return at 0,1%. Similarly, the IRF of Gold/Platinum ratio in the partial model seems to report a positive return (day three) at 0,075% while it is slightly below this point in the unrestricted model. This is the recurring observation in these comparisons, where the significant results are closer to zero as more variables are taken into account, but the patterns still remain. Generally, there are few large deviations in the restricted models from the patterns presented in section 5.2, although some significant results turn insignificant as more variables are included.

The adjusted explanatory rate of the restricted models naturally shifts as a result of fewer indicators included. The adjustment can also result in a lower suggested optimal lag length, further decreasing the explanatory power. Even though some indicators, such as the U.S. dollar index, the credit spread, and the VWMA has optimal lag lengths close to or higher than suggested in the main models (respectively 16, 27 and 48 number of lags), the adjusted R squared is lower in comparison.

The partial models pass most of the VAR-model specific tests, similarly to the main models. Nonetheless, there are a few exceptions noted in the corresponding summary statistics. The partial model with the S&P500 index and the VWMA has structural breaks. Most of the partial models do also suffer non-white residuals.

The system created by the VAR-model becomes more complicated when more indicators are taken into account, resulting fewer significant results. Thus, when it is significantly different from zero, one can be more certain of this outcome. The combination also leads to models explaining more of the movement of the stock index according to the adjusted R^2.

5.3 Forecast error variance decomposition

The forecast error variance decompositions are performed for each of the subsets. The results are very similar, with only some minor differences. Therefore, only the decomposition of the whole period (2005-2020) is presented in this section. The remaining can be found in the appendix 9.5, table 15, 16 and 17.

For every variable, most of the variance can be explained by shocks in their own error terms. As mentioned in section 3.2.3, Enders explained this is to be expected as variables tend to experience this in the short-term. The simulation period is kept consistent with previous parts of the analysis, at seven days. At this time horizon no major changes occur in the variance explained. If the horizon is expanded to for instance 20 days, significant changes arise. However, this is not considered to be advantageous to the analysis. When we increase the horizon, it is less likely that the past values of S&P500 explain its own variance as more noise is added to the equation. This results in the other explanatory variables being given more importance without giving additional support to our hypothesis of predictive powers in the indicators.

In the seven days simulated (2005-2020) the explained variance stays fairly persistent throughout the simulation period. The S&P500 index is the variable with the most influence on the others. Naturally, the VWMA is mostly influences by the returns of the index. The VIX and Put/Call ratio also have large contributions from the S&P500 index with respectively ≈ 54% and ≈ 19% of their variance explained through the index. The remaining indicators only have noteworthy influence from shocks to their own error term.

As the decompositions of the subset move towards 2020, a similar pattern to the IRFs emerges. The indicators appear to gain significance, though not equally evident. Early in the simulation the variance remains explained by shocks to their own error term but as one approach the end of the period, the other indicators gain significance. The observed increase in explained variance is not considered to be high enough, resulting in the forecast error variance decomposition not giving any additional support to the argument of predictive abilities in the selected variables.

Forecast error variance decomposition 2005-2020

Days	Return of the S&P500 index	Percentage change in the Put/Call ratio	Percentage change in the Gold/Platinum ratio	Percentage change in the VIX	Percentage change in the U.S. dollar index	Percentage change in the credit spread Baa-Aaa	Crisis dummy	Deviation from the VWMA
Return of the S&P500 index								
1	100.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%
2	99.197%	0.457%	0.001%	0.044%	0.058%	0.026%	0.045%	0.174%
3	98.553%	0.457%	0.032%	0.064%	0.197%	0.030%	0.052%	0.615%
4	98.121%	0.516%	0.365%	0.063%	0.198%	0.040%	0.078%	0.619%
5	97.806%	0.622%	0.392%	0.098%	0.208%	0.070%	0.144%	0.660%
6	97.337%	0.618%	0.478%	0.108%	0.240%	0.177%	0.148%	0.893%
7	96.266%	0.614%	0.477%	0.108%	0.489%	0.244%	0.152%	1.650%
% change in the Put/Call ratio								
1	21.947%	78.053%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%
2	18.598%	80.914%	0.062%	0.255%	0.065%	0.000%	0.097%	0.009%
3	19.143%	80.149%	0.070%	0.344%	0.074%	0.011%	0.158%	0.051%
4	19.169%	80.074%	0.115%	0.343%	0.074%	0.014%	0.159%	0.052%
5	19.219%	79.967%	0.114%	0.367%	0.104%	0.016%	0.161%	0.052%
6	19.192%	79.856%	0.126%	0.400%	0.105%	0.029%	0.231%	0.061%
7	19.155%	79.583%	0.243%	0.408%	0.193%	0.075%	0.277%	0.065%
% change in the Gold/Platinum ratio								
1	8.908%	0.163%	90.929%	0.000%	0.000%	0.000%	0.000%	0.000%
2	9.352%	0.406%	89.573%	0.003%	0.191%	0.005%	0.048%	0.421%
3	9.334%	0.408%	88.881%	0.311%	0.220%	0.039%	0.363%	0.444%
4	9.536%	0.435%	88.308%	0.327%	0.221%	0.168%	0.423%	0.583%
5	9.415%	0.596%	87.277%	0.322%	0.254%	0.231%	1.290%	0.615%
6	9.331%	0.590%	86.496%	0.344%	0.267%	0.394%	1.355%	1.224%
7	9.323%	0.608%	86.236%	0.366%	0.268%	0.403%	1.482%	1.316%
% change in the VIX								
1	54.308%	1.089%	0.031%	44.572%	0.000%	0.000%	0.000%	0.000%
2	54.378%	1.270%	0.144%	44.106%	0.015%	0.040%	0.000%	0.048%
3	54.205%	1.265%	0.259%	44.066%	0.016%	0.064%	0.008%	0.116%
4	53.903%	1.264%	0.535%	43.825%	0.053%	0.161%	0.084%	0.176%
5	53.793%	1.294%	0.534%	43.806%	0.067%	0.235%	0.084%	0.187%
6	53.572%	1.540%	0.582%	43.583%	0.140%	0.277%	0.096%	0.210%
7	53.379%	1.535%	0.667%	43.481%	0.199%	0.292%	0.102%	0.346%
% change in the U.S. dollar index								
1	2.425%	0.029%	0.017%	0.338%	97.192%	0.000%	0.000%	0.000%
2	3.552%	0.102%	0.061%	0.646%	95.600%	0.000%	0.023%	0.016%
3	3.546%	0.102%	0.061%	0.647%	95.484%	0.034%	0.063%	0.064%
4	3.588%	0.120%	0.113%	0.694%	95.062%	0.070%	0.121%	0.232%
5	3.588%	0.168%	0.123%	0.752%	94.786%	0.102%	0.248%	0.233%
6	3.647%	0.169%	0.177%	0.763%	94.652%	0.106%	0.250%	0.236%
7	3.633%	0.170%	0.178%	0.771%	94.293%	0.289%	0.389%	0.276%
% change in the credit spread Baa-Aaa								
1	0.096%	0.002%	0.008%	0.018%	0.106%	99.769%	0.000%	0.000%
2	0.435%	0.024%	0.105%	0.018%	0.103%	99.286%	0.000%	0.028%
3	1.012%	0.024%	0.136%	0.048%	0.140%	98.545%	0.001%	0.094%
4	2.218%	0.026%	0.165%	0.165%	0.139%	97.135%	0.004%	0.149%
5	2.209%	0.054%	0.191%	0.205%	0.276%	96.785%	0.039%	0.241%
6	2.982%	0.053%	0.505%	0.217%	0.319%	94.254%	0.559%	1.112%
7	2.969%	0.069%	0.527%	0.255%	0.396%	93.671%	0.672%	1.441%

Forecast error variance decomposition 2005-2020

Days	Return of the S&P500 index	Percentage change in the Put/Call ratio	Percentage change in the Gold/Platinum ratio	Percentage change in the VIX	Percentage change in the U.S. dollar index	Percentage change in the credit spread Baa-Aaa	Crisis dummy	Deviation from the VWMA
Crisis dummy								
1	4.220%	0.488%	0.186%	0.982%	0.001%	0.015%	94.109%	0.000%
2	5.248%	0.555%	0.257%	1.178%	0.036%	0.043%	92.635%	0.047%
3	5.681%	0.524%	0.360%	1.152%	0.113%	0.036%	92.069%	0.066%
4	7.715%	0.525%	0.381%	1.290%	0.109%	0.093%	89.422%	0.465%
5	8.481%	0.488%	0.352%	1.343%	0.113%	0.095%	88.573%	0.554%
6	8.406%	0.446%	0.509%	1.236%	0.243%	0.114%	87.897%	1.148%
7	9.648%	0.427%	0.545%	1.321%	0.267%	0.112%	86.584%	1.097%
Deviation from the VWMA								
1	80.394%	0.090%	0.202%	1.152%	0.172%	0.022%	0.003%	17.966%
2	80.165%	0.731%	0.273%	1.669%	0.099%	0.036%	0.102%	16.925%
3	78.094%	0.824%	0.656%	2.014%	0.087%	0.034%	0.083%	18.207%
4	76.656%	1.013%	0.579%	2.464%	0.067%	0.026%	0.067%	19.127%
5	75.687%	1.384%	0.557%	3.036%	0.056%	0.044%	0.060%	19.176%
6	74.338%	1.605%	0.654%	3.298%	0.067%	0.045%	0.051%	19.941%
7	73.914%	1.801%	0.707%	3.661%	0.061%	0.080%	0.048%	19.729%

Table 4: Forecast error variance decomposition (2005-2020)

6. Discussion

In this part, we would like to shed a light on certain aspects of the analysis. We reflect upon the applicability of the results, the choice of econometric model, as well as clarify limitations of the model with relevance to the analysis. We also suggest areas of interest for future analysis.

6.1 Reflections about the results

Dividing the data into five and ten-year periods, illustrated interesting developments in the predictability of the indicators across time. Most of the variations visualized are logical considering the changes the financial markets have seen in the corresponding time span. Given the substantial increase in activity in the options market in recent years, it is reasonable to relate the Put/Call ratio seemingly gaining importance to this fact. Similarly, when quantitative easing is more commonly used by central banks, it can affect the pricing of currencies. As a consequence, their predictive abilities may be impaired, visualized for the dollar in the results.

Our results are mostly in compliance with Pan and Poteshman (2006), who introduced the predictive ability of the Put/Call ratio on stocks. Overall, our analysis shows this to be true for the ratio, even though the earliest subset (2005-2010) is inconclusive. While their findings resulted in an investment strategy that gave 40 basis points return per day, our results show a decline of the S&P500 index the following day after a performed shock to the ratio. When considering the Gold/Platinum ratio by comparing our result to the results of Huang & Kilic (2019), the outcome is less coherent. Firstly, the results from our analysis regarding this indicator are not consistent through time. Secondly, the difference in the time span examined in this thesis and the article further complicates this comparison, making it hard to state whether the results contradict each other.

Even though the results indicate the direction of how the stock market would react to changes in the indicators, in most of the indicators we were not able to consistently pinpoint the exact day of the reaction, as it varied through the subsets. In hindsight, one could argue for the use of data with a lower frequency could result in more consistent findings. For example, if the data was analyzed on a weekly basis, the returns would be cumulated disregarding the issue of identifying the exact day, but rather emphasize whether there is a significant reaction within that timeframe. This could be beneficial for the inconsistent readings, while equally disadvantageous for the consistent ones.

Practical utilization of the results

A simplification was made in the interpretation of the IRFs to be able to visualize the changes of significance the indicators appear to experience through time, where the days one could not significantly state the direction was considered to be zero. To apply the information of the IRFs directly in a trading strategy is difficult, as this is not true in the real world. The only indicator giving consistent results regarding the reaction of the shock is the Put/Call ratio, where the reaction is expected the next day. One can potentially monetize rapidly on shorting the market whenever the Put/Call ratio increases substantially. Otherwise, the days of the significant reactions from shocks in the indicators are highly inconsistent and therefore hard to utilize in a similar manner. Disregarding which day the significant reactions occur, the direction of the stock index after shocks in the indicators appear to stay fairly consistent throughout the subperiods. Giving an understanding about the relationship between the indicators and the stock index, where a change of the risk tolerance in the respective financial markets often seem to affect the return of the stock index in the coming period.

The returns shown in the IRFs as a result of shocks in the indicators are ranging from -0,1% to 0,1%. If transaction costs are accounted for, the potential profit would likely be smaller/disappear. Nevertheless, as the results indicates predictive powers of the indicators, there are reasons to believe they can be utilized at their more extreme notations. The returns provided from daily trading would likely not exceed the transaction costs, but a possible utilization could be as buy/sell signals with a longer time perspective or in times of turmoil. If the percentage change in the indicators surpass one standard deviation, it entails that the returns of the S&P500 index would likely be higher/lower than presented in the IRFs, meaning the indicators could be useful to interpret in times where they experience high volatility.

As the economic environment changes, it appears so does the predictive abilities of the indicators. Implying that as the information of the indicators is used in a predictive manner, one has to be considerate of the times, as well as the impact outside forces may have on the indicators, potentially distorting the interpretation of them. The relationships do not appear to be constant through time, leading to a necessity of a dynamic mindset if the information in the indicators is to be used in an investment approach.

It is also important to note that the models of which the IRFs are based upon does not explain the whole return of the S&P500 index. As previously mentioned, the models explanatory rate ranges from 12,0% to 28,4%, and this must be considered if the results are to be used in decision making.

6.2 Reflections on the choice of model and the implications of non-white residuals

The choice of model has implications for the outcome of the thesis and its results. The intention of this thesis was to investigate the relationship between multiple possible leading indicators and the S&P500 index, where both the present and past notations of the indicators were considered. With this in mind, a reasonable choice was the vector autoregressive model. However, in the databases at our disposal, we could not find any research where a VAR-model had previously been used similarly, based on daily data. For the most part it was used on monthly, quarterly and yearly data, often in relations with analysis of business cycles. We therefore found it interesting to use the VAR-method to investigate the research question of this thesis. In principle, the model should be applicable independent of the timescale used. In hindsight however, it has impacted the analysis, where the assumptions for whiteness of the residuals are not satisfied.

As mentioned in chapter 1.3, we have certain inadequacies in the model. A summary of the tests with relation to this can be found in Appendix 9.2.6. There is a presence of heteroscedasticity, autocorrelation, and non-normally distributed residuals. Typical techniques of handling such errors were attempted unsuccessfully. Each variable was tried in natural logarithmic form to combat heteroscedasticity, and a correction for heteroscedastic-consistent standard errors was not feasible in the VAR framework utilized in this analysis. Outliers were also omitted without it having any effect on either the heteroscedasticity or the non-normally distributed residuals. However, one could argue that the extreme values of the indicators are the most important to this analysis, as they are the ones showing the largest change in risk aversion. Therefore, since the omission of them did not help the state of the residuals, they are kept in the sample.

Increasing the number of lags in the model is also a method of combating issues related to the residuals, especially autocorrelation. When the optimal lag length (suggested by the information criterions) is utilized, the assumptions for whiteness of residuals are not satisfied, for all four periods. Heteroscedasticity is removed when optimal lags are multiplied by 10, and autocorrelation is removed when multiplied by 20 (Appendix 9.2.6). As the numbers of lags increases, the likelihood of obtaining spurious results increases. When the numbers of lags need to be multiplied by 20 to do this correction, we found it unreasonable to do this correction and rather interpret the result considering these limitations.

The presence of heteroscedasticity, autocorrelation and non-normally distributed residuals affects the conclusion that can be drawn from the results. Heteroscedasticity affects the estimated standard errors used in hypothesis testing (t-statistics and F-statistics), in computing confidence intervals and in the impulse response analysis. Non-normality also affects the hypothesis tests that assumes normal distribution. When there is a presence of autocorrelation, one cannot characterize the ordinary least square methodology as the best linear unbiased estimate since the OLS estimators becomes inefficient. In addition, the estimated variance becomes biased and inconsistent which often leads to an overestimation of the R^2. Consequently, we need to draw conclusions based on this and be aware that the significancy is affected through the lack of econometric least squares assumptions being fully satisfied.

Considering the beforementioned elements and issues one can argue that the VAR-method was not the best way to approach this analysis. The use of ordinary least squares methods normally is better at explaining general features rather than special features. This is a consequence of the desire to create the best linear unbiased estimates, whereas extensive measures are done to ensure this, for example omitting outliers. As mentioned, many of proxies for market sentiment chosen in this thesis is thought to have the most reliable information whenever they are at their extreme notations when they are illustrating the largest change in risk tolerance. As omitting them would cause a severe loss of information it is clearly an argument for selecting a different way to analyze the significance of the indicators. The use of daily data in this manner is also an argument against the choice of the VAR-model, as the likelihood of non-white residuals increase, resulting in reliability of the results being affected.

6.3 Suggestions for further analysis

Since the VAR-method does not appear to be the best way to analyze the predictive abilities of the indicators, it would be interesting to analyze them using a different approach. There seem to exist patterns binding some of the variables, which can be visualized more reliably by other methods. It would be interesting to see a similar analysis being done using machine learning, for example.

Another element briefly touched upon in this thesis is the effect of delta hedging on the underlying asset price. This can also be a subject for further analysis, where one could investigate the extent of this impact. In an environment where the options market is growing at a high rate, this can be highly relevant.

7. Conclusion

By analyzing the predictive abilities using a vector autoregressive model, we found conflicting evidence for the predictive abilities of the indicators towards the S&P500 index based on the commonly used methods of interpreting the model. Parts of the analysis appear to show predictive abilities, while others are less conclusive. Models were created for the whole sample period (2005-2020) as well as for subsets within the period. The explanatory power of the indicators in the period of 2005-2020 is 12,0%, while the subsets range from 18,1% (2010-2020) to 28,4% (2015-2020). Implying that the indicators is gaining significance, or that the patterns are clearer in the constrained periods.

The granger causality tests show a causal relationship between the indicators throughout the sample period, with one exemption. The U.S. dollar index (relative strength of the dollar) lose its causality in the latest subset (2015-2020). This is seen as a logical response to the extensive interventions by the Federal Reserve in recent years, consequently affecting the pricing of the dollar through a massive increase in the supply.

The impulsive response functions also illustrate foretelling capabilities in the chosen indicators. The direction of the S&P500 index due to shocks performed on the indicators is determined within a 95% confidence interval. Indicators such as the Put/Call ratio, the U.S. dollar index and the credit spread (Baa-Aaa) illustrated a negative reaction to the performed shock, while indicators like Gold/Platinum and the volume weighted moving average reported inconsistency in their reactions. Some indicators appear to gain importance, while others become less important throughout the subsets. For instance, the Put/Call ratio gave no conclusive direction in the earliest subset (2005-2010), but showed consistent negative returns the following day in the other subsets. Similarly, the U.S. dollar index showed significant negative returns in the earliest subperiod, whereas the effect is no longer significant later, which can be linked to the loss of granger causality. Since the Put/Call ratio was the only indicator illustrating a consistency in which day the reaction from the shock occurred throughout the sample subsets, and the other variables were less consistent, it complicates the potential utilization of the results.

The results are compared to partial models where the same approach is used in creating models with each indicator individually. There are no large deviations in the patterns in this comparison. However, as more variables are included in the model and the system explains more of the variance of the stock index, the patterns get more reliable. It leads to fewer significant observations and more conclusive results.

The forecast error variance decomposition seems to have a similar pattern as the IRFs where the indicators gain significance in the later periods, though not equally evident.
The recurring observation in the forecast error variance decomposition throughout the subsets is that the S&P500 index mostly explain its own variance, with only minor contributions from the other variables. This is expected in the beginning of the seven-day simulated period, but one would expect the indicators to explain more as one approach the later part of the simulation. The variance decomposition rather shows that the stock index explains a bigger portion of the variance of the other variables. Consequently, not supporting the hypothesized foretelling abilities of the indicators towards the stock index.

One can question whether the choice of the VAR-model was optimal when researching the relationship between the variables at this frequency in the data. It complicated the desire to abide by the general econometric requirements. The fact that they have been disregarded have reduced the reliability of the results. Despite this, interesting patterns in the reaction of the S&P500 index have been displayed which can be related to recent changes in the economic environment. These patterns can be linked to the expected response from changes in the market sentiment and the risk tolerance of the market participants, but in order to base an investment strategy on the result one may desire more reliable results.

8. References

Black, F., (1975). Fact and fantasy in the use of options. Financial Analysts Journal 31 (4), 36–72.

Blau, B., Nguyen, N. and Whitby, R., (2013). The Information Content of Option Ratios. *SSRN Electronic Journal*,.

Brock, W., Lakonishok, J. and LeBaron, B., (1992). Simple Technical Trading Rules and the Stochastic Properties of Stock Returns. The Journal of Finance, 47(5), pp.1731-1764.

Brown, G. and Cliff, M., (2001). Investor Sentiment And The Near-Term Stock Market. SSRN Electronic Journal.

Chen, J (Updated Jan 28, 2021) Delta Hedging. *Investopedia* [Internet]. Available from: https://www.investopedia.com/terms/d/deltahedging.asp [Read Mar 24, 2021]

Chen, J (Updated Dec 28, 2020) U.S. Dollar Index. *Investopedia* [Internet]. Available from: https://www.investopedia.com/terms/u/usdx.asp [Read Mar 03, 2021]

Dickey, D., & Fuller, W. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. Journal of the American Statistical Association, pp. 427-431.

Enders, W. (2010). *Applied Econometric Time Series, 3rd. edition.* Wiley.

Fama, E., (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance, 25(2), pp.383.

Fama, E., (1997). Market Efficiency, Long-Term Returns, and Behavioral Finance. SSRN Electronic Journal.

Fama, E. and French, K., (1988). Dividend yields and expected stock returns. Journal of Financial Economics, 22(1), pp.3-25.

Fama, E. and French, K., (1989). Business conditions and expected returns on stocks and bonds. Journal of Financial Economics, 25(1), pp.23-49.

Fernando, J. (Updated Feb 24, 2021). Volume Weighted Average Price Definition. *Investopedia* [Internet]. Available from: https://www.investopedia.com/terms/v/vwap.asp [Read Mar 23, 2021].

Granger, C. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37-3, ss. 424-438.

Hon, M. and Tonks, I. (2003). Momentum in the UK stock market. Journal of Multinational Financial Management, 13(1), pp.43-70.

Huang, D. and Kilic, M. (2019) Gold, platinum, and expected stock returns. *Journal of Financial Economics*, 132(3), pp.50-75.

Hui, B. and Li, P. (2014) Does Investor Sentiment Predict Stock Returns? The Evidence From Chinese Stock Market. *J Syst Sci Complex,* 27(1), pp 130-143.

Ivanov, V., & Killian, L. (2001). A Practitioner's Guide to Lag-Order Selection for Vector Autoregressions. I CEPR Discussion Paper no. 2685. London: Centre for Economic Policy Research.

Johnson, T. and So, E. (2011) The Option to Stock Volume Ratio and Future Returns. *SSRN Electronic Journal*.

Keynes, J.M. (1936) *The General Theory of Employment, Interest and Money*, Wordsworth Editions Ltd).

Kuepper, J. (Updated Oct 30, 2020). Cboe Volatility Index (VIX) Definition. *Investopedia* [Internet]. Available from: https://www.investopedia.com/terms/v/vix.asp [Read Feb 15, 2021].

Lütkepohl, H (2007). *New Introduction to Multiple Time Series Analysis.* Springer

Maddala, G., & Kim, I.-M. (1998). *Unit Roots, Cointegration and Structural Change.* Cambridge: Cambridge University Press.

March, G.G., Simon, H.A. (1958). *Organizations*. John Wiley & Sons,

Pan, J. and Poteshman, A. (2003) The Information in Option Volume for Stock Prices. *SSRN Electronic Journal*.

Ploberger, W. and Krämer. W. (1992) The CUSUM test with OLS residuals. Econometrica, 60(2): 271–285.

Sims, C. (1972) Money, Income and Causality. *The American Economic Review, Vol 62, No.4*, ss. 540-552.

Shefrin H. (2000) *Beyond Greed and Fear: Understanding Behavioral Finance and the Psychology of Investing*. Harvard Business School Press

Smirlock, M. and Starks, L. (1986) Day-of-the-week and intraday effects in stock returns. Journal of Financial Economics, 17(1), pp.197-210.

Smith, T. (Updated Dec 8, 2020) Market Sentiment. Investopedia [Internet]. Available from: https://www.investopedia.com/terms/m/marketsentiment.asp [Read Feb 15, 2021]

Tenn, J. (Updated Nov 18, 2020) 2020 Options Market: The Rise In Young Retail Traders And New Trading Platforms [Internet] Available form : https://finance.yahoo.com/news/2020-options-market-rise-young-180432151.html [Read March 20, 2021]

Tversky, A. and Kahneman, D. (1974) Judgment under Uncertainty: Heuristics and Biases. Science, 185(4157), pp.1124-1131.

Tversky, A. and Kahneman, D. (1986) Rational Choice and the Framing of Decisions. The Journal of Business, 59(S4), pp 251.

Watsham, T. & Parramore, K. (1997) Quantitative Methods in Finance. Thomson Learning. (pp.231)

Wooldridge, J. (2016) Introductory Econometrics, a modern approach. Michigan State University, Cengage Learning

Wouters, T. (2006) Style investing: behavioral explanations of stock market anomalies, Doctor of Philosophy, University of Groningen.

9. Appendix

9.1 Correlation matrices

9.1.1 Correlation matrix of leading/lagged indicators to the S&P500 index

Correlation matrix of the indicators' daily percentage change to the S&P index return in the period of 2005-2020

	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
CBOE Put/Call	-0.022	0.033	-0.0005	0.021	0.029	0.075	0.143	-0.389	-0.007	0.013	-0.028	0.004	0.011	0.016	-0.036
Gold/Platinum	-0.011	-0.014	-0.017	-0.004	-0.052	-0.011	-0.078	-0.300	0.062	-0.036	0.057	-0.007	-0.029	-0.003	-0.015
Implied volatility index	-0.001	0.034	0.013	0.028	0.002	0.007	0.100	-0.723	0.071	-0.021	-0.027	0.025	0.015	0.014	-0.035
U.S. dollar index	0.009	0.004	-0.036	0.013	0.013	0.008	-0.117	-0.142	-0.009	-0.048	0.006	0.009	0.049	-0.052	0.015
Moodys Baa-Aaa spread	-0.017	0.005	-0.114	0.041	-0.132	-0.047	-0.064	-0.035	0.035	-0.026	0.021	-0.005	-0.020	-0.033	0.002
Crisis dummy	-0.091	-0.078	-0.061	-0.067	-0.087	-0.068	-0.068	-0.096	0.005	-0.027	-0.022	-0.025	-0.017	-0.014	-0.025
Volume Weighted Moving Average	0.167	0.141	0.167	0.166	0.190	0.191	0.177	0.226	-0.040	0.006	-0.015	-0.021	-0.005	-0.013	0.012

Table 5: Correlation matrix of leading/lagged indicators to the S&P500 index return (2005-2020)

9.1.2 Correlation between the percentage change of the indicators (2000-2020)

Correlation matrix of the percentage change of the indicators

	Return of the S&P500 index	Percentage change in the Gold/Platinum ratio	Percentage change in the Put/Call Ratio	Percentage change in the Option/Stock Ratio	Percentage change in the The VIX	Percentage change in the Yield spread 10y-3m	Percentage change in the Credit spread Baa-Aaa	Percentage change in theo U.S. dollar index	Deviation in of the Volume Weighted Moving Average
Return of the S&P500 index	1	-0.243	-0.358	0.078	-0.719	-0.003	-0.004	-0.059	0.225
Percentage change in the Gold/Platinum ratio	-0.243	1	0.090	-0.027	0.192	-0.014	-0.017	-0.009	-0.120
Percentage change in the Put/Call Ratio	-0.358	0.090	1	-0.011	0.316	0.031	-0.007	0.009	-0.030
Percentage change in the Option/Stock Ratio	0.078	-0.027	-0.011	1	-0.081	-0.015	-0.0002	0.005	0.027
Percentage change in the The VIX	-0.719	0.192	0.316	-0.081	1	-0.008	0.002	0.031	-0.134
Percentage change in the Yield spread 10y-3m	-0.003	-0.014	0.031	-0.015	-0.008	1	-0.00001	0.022	-0.002
Percentage change in the Credit spread Baa-Aaa	-0.004	-0.017	-0.007	-0.0002	0.002	-0.00001	1	-0.007	0.003
Percentage change in the U.S. dollar index	-0.059	-0.009	0.009	0.005	0.031	0.022	-0.007	1	-0.051
Deviation in of the Volume Weighted Moving Average	0.225	-0.120	-0.030	0.027	-0.134	-0.002	0.003	-0.051	1

Table 6: Correlation between the percentage change of the indicators (2000-2020)
NB, Volume Weighted Moving Average is noted in deviation, not percentage difference.

9.2 Model statistics

9.2.1 Summary statistic, subsetted periods

Summary statistics of daily data in the period of 2005-2010

	Mean	Standard Deviation	Minimum	Maximum	Observations
Return of the S&P500 index	0.0001	0.0143	-0.0903	0.1079	1,497.0000
Percentage change in the Put/Call ratio	0.0135	0.1639	-0.4955	0.8718	1,497.0000
Percentage change in the Gold/Platinum ratio	0.0004	0.0132	-0.0561	0.0983	1,497.0000
Percentage change in the VIX	0.0027	0.0700	-0.2957	0.6422	1,497.0000
Percentage change in the U.S. dollar index	0.0000	0.0055	-0.0269	0.0255	1,497.0000
Percentage change in the credit spread Baa- Aaa	0.0004	0.0170	-0.1000	0.1538	1,497.0000
Crisis dummy	0.0882	0.2836	0.0000	1.0000	1,497.0000
Deviation of Volume Weighted Moving Average	2.3602	51.7530	-295.915	118.0823	1,497.0000

Table 7: Summary statistics table (2005-2010)

Summary statistics of daily data in the period of 2010-2020

	Mean	Standard Deviation	Minimum	Maximum	Observations
Return of the S&P500 index	0.0005	0.0110	-0.1198	0.0938	2,745.0000
Percentage change in the Put/Call ratio	0.0112	0.1479	-0.5537	1.1228	2,745.0000
Percentage change in the Gold/Platinum ratio	0.0003	0.0106	-0.0500	0.1343	2,745.0000
Percentage change in the VIX	0.0032	0.0834	-0.2957	1.1560	2,745.0000
Percentage change in the U.S. dollar index	0.0001	0.0045	-0.0237	0.0205	2,745.0000
Percentage change in the credit spread Baa- Aaa	0.0002	0.0228	-0.1890	0.4563	2,745.0000
Crisis dummy	0.0066	0.0807	0.0000	1.0000	2,745.0000
Deviation of Volume Weighted Moving Average	25.1083	80.8419	-743.261	370.7127	2,745.0000

Table 8: Summary statistics table (2010-2020)

Summary statistics of daily data in the period of 2015-2020

	Mean	Standard Deviation	Minimum	Maximum	Observations
Return of the S&P500 index	0.0005	0.0118	-0.1198	0.0938	1,498.0000
Percentage change in the Put/Call ratio	0.0114	0.1481	-0.5537	1.1228	1,498.0000
Percentage change in the Gold/Platinum ratio	0.0004	0.0117	-0.0500	0.1343	1,498.0000
Percentage change in the VIX	0.0036	0.0900	-0.2591	1.1560	1,498.0000
Percentage change in the U.S. dollar index	0.0000	0.0044	-0.0237	0.0205	1,498.0000
Percentage change in the credit spread Baa- Aaa	0.0003	0.0258	-0.1890	0.4563	1,498.0000
Crisis dummy	0.0120	0.1090	0.0000	1.0000	1,498.0000
Deviation of Volume Weighted Moving Average	29.9975	102.4669	-743.261	370.7127	1,498.0000

Table 9: Summary statistics table (2015-2020)

9.2.2 VAR-representation

SP = S&P500 index return
PC = Put Call CBOE return
GP = Gold Platinum return
U.S. dollar = Currency index return (relative strength dollar to index)
CS = Credit spread (Moodys Baa – Moodys Aaa)
VWMA = Volume weighted moving average

$$SP = \alpha_0 + \sum_{i=1}^{26} ß_1 SP_{t-i} + \sum_{i=1}^{26} ß_2 PC_{t-i} + \sum_{i=1}^{26} ß_3 GP_{t-i} + \sum_{i=1}^{26} ß_4 DXY_{t-i} + \sum_{i=1}^{26} ß_5 CS_{t-i} + \sum_{i=1}^{26} ß_6 VWMA_{t-i} + \varepsilon_{1t} \tag{a}$$

$$PC = \alpha_0 + \sum_{i=1}^{26} ß_1 SP_{t-i} + \sum_{i=1}^{26} ß_2 PC_{t-i} + \sum_{i=1}^{26} ß_3 GP_{t-i} + \sum_{i=1}^{26} ß_4 DXY_{t-i} + \sum_{i=1}^{26} ß_5 CS_{t-i} + \sum_{i=1}^{26} ß_6 VWMA_{t-i} + \varepsilon_{1t} \tag{b}$$

$$GP = \alpha_0 + \sum_{i=1}^{26} ß_1 SP_{t-i} + \sum_{i=1}^{26} ß_2 PC_{t-i} + \sum_{i=1}^{26} ß_3 GP_{t-i} + \sum_{i=1}^{26} ß_4 DXY_{t-i} + \sum_{i=1}^{26} ß_5 CS_{t-i} + \sum_{i=1}^{26} ß_6 VWMA_{t-i} + \varepsilon_{1t} \tag{c}$$

$$DXY = \alpha_0 + \sum_{i=1}^{26} ß_1 SP_{t-i} + \sum_{i=1}^{26} ß_2 PC_{t-i} + \sum_{i=1}^{26} ß_3 GP_{t-i} + \sum_{i=1}^{26} ß_4 DXY_{t-i} + \sum_{i=1}^{26} ß_5 CS_{t-i} + \sum_{i=1}^{26} ß_6 VWMA_{t-i} + \varepsilon_{1t} \tag{d}$$

$$CS = \alpha_0 + \sum_{i=1}^{26} ß_1 SP_{t-i} + \sum_{i=1}^{26} ß_2 PC_{t-i} + \sum_{i=1}^{26} ß_3 GP_{t-i} + \sum_{i=1}^{26} ß_4 DXY_{t-i} + \sum_{i=1}^{26} ß_5 CS_{t-i} + \sum_{i=1}^{26} ß_6 VWMA_{t-i} + \varepsilon_{1t} \tag{e}$$

$$VWMA = \alpha_0 + \sum_{i=1}^{26} ß_1 SP_{t-i} + \sum_{i=1}^{26} ß_2 PC_{t-i} + \sum_{i=1}^{26} ß_3 GP_{t-i} + \sum_{i=1}^{26} ß_4 DXY_{t-i} + \sum_{i=1}^{26} ß_5 CS_{t-i} + \sum_{i=1}^{26} ß_6 VWMA_{t-i} + \varepsilon_{1t} \tag{f}$$

9.2.3 Stationarity of the indicators within the models

Stationarity of the indicators within the models

	2005-2020		2005-2010		2010-2020		2015-2020	
	P-Value	Critical value	P-Value	Critical value	P-Value	Critical value	P-Value	Critical value
Return of the S&P500	0.01	-4605.2947	0.01	-1667.44804	0.01	-3167.7172	0.01	-1876.51602
Put/Call ratio	0.01	-4313.7376	0.01	-1711.31790	0.01	-2922.5663	0.01	-1638.25969
Gold/Platinum ratio	0.01	-4107.8392	0.01	-1430.65510	0.01	-2908.6721	0.01	-1672.13311
The VIX	0.01	-3934.7862	0.01	-1528.83290	0.01	-2678.2078	0.01	-1490.27736
The U.S. dollar index	0.01	-3962.9528	0.01	-1491.90920	0.01	-2654.0110	0.01	-1445.75998
Credit spread Baa-Aaa	0.01	-5539.1571	0.01	-1767.41096	0.01	-3700.2266	0.01	-2085.10848
Crisis dummy	0.01	-283.7517	0.01	-94.28509	0.01	-297.5495	0.01	-151.86494
Deviation from VWMA	0.01	-181.2812	0.01	-54.67516	0.01	-127.1877	0.01	-66.74476

Table 10: Stationarity of the indicators

9.2.4 Granger Causality of the indicators within the models

Granger Causality of the indicators within the models

	2005-2020		2005-2010		2010-2020		2015-2020	
	P-Value	Critical value	P-Value	Critical value	P-Value	Critical value	P-Value	Critical value
Put/Call ratio	0.0001405	1.427932	0.0001434	1.414081	0.005254	1.317150	0.0001735	1.518853
Gold/Platinum ratio	< 2.2e-16	1.827462	0.0000037	1.410072	< 2.2e-16	2.581669	< 2.2e-16	2.592663
Credit spread Baa-Aaa	< 2.2e-16	1.827462	0.0001617	1.410072	< 2.2e-16	2.581669	< 2.2e-16	2.592663
The VIX	< 2.2e-16	2.817521	< 2.2e-16	1.864401	< 2.2e-16	2.727283	< 2.2e-16	2.164291
The U.S. dollar index	< 2.2e-16	1.722757	0.0023411	1.313651	0.004112	1.328739	0.0532073	1.213133
Crisis dummy	< 2.2e-16	3.192341	< 2.2e-16	2.253497	< 2.2e-16	10.023558	< 2.2e-16	8.659830
Deviation from VWMA	< 2.2e-16	4.745652	< 2.2e-16	2.173630	< 2.2e-16	4.571691	< 2.2e-16	2.193611

Table 11: Granger Causality of the indicators

9.2.5 Stability of the indicators within the models

Stability of the indicators in the model of 2005-2020

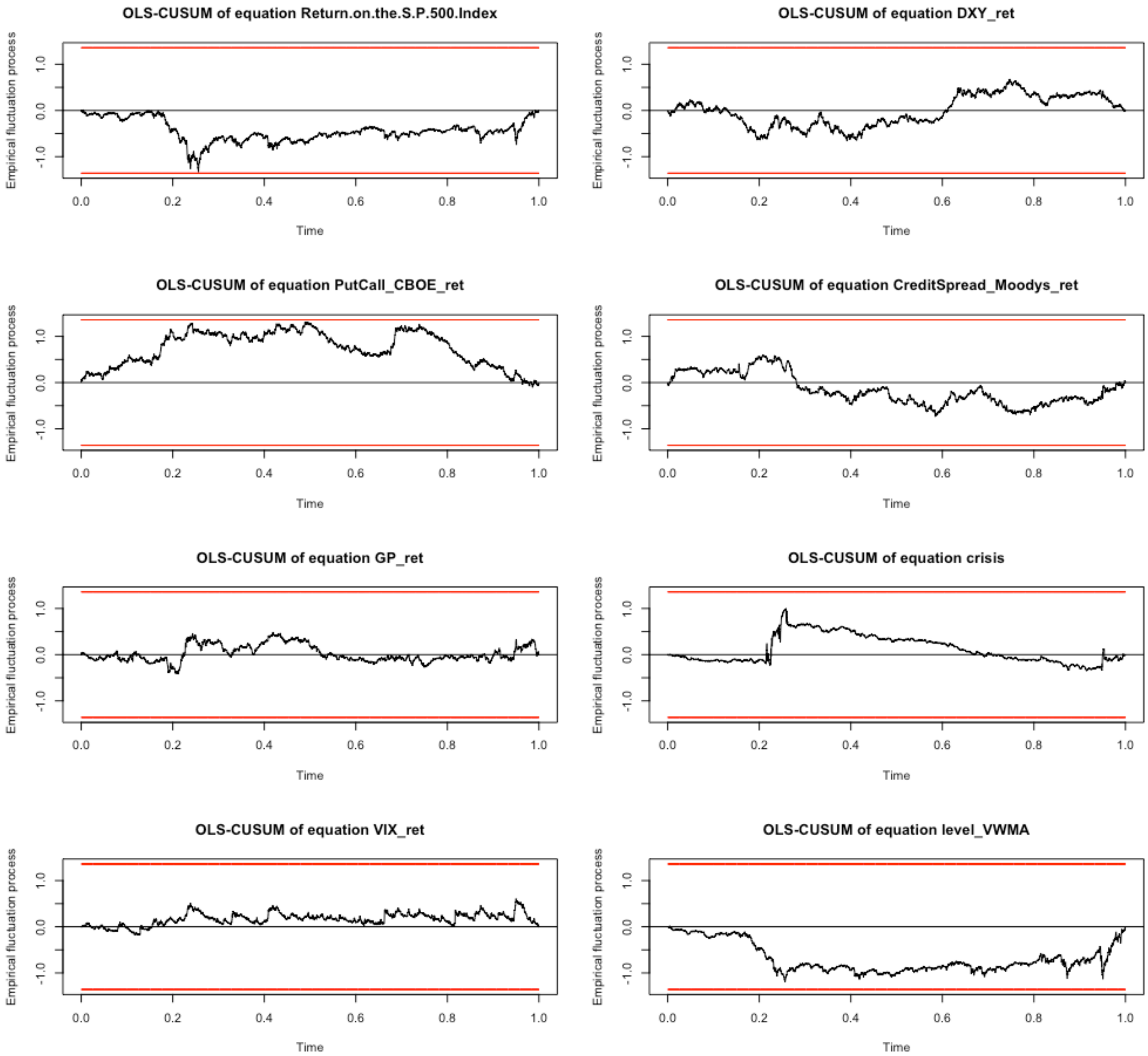


Figure 16: Stability of the indicators in the model of 2005-2020

Stability of the indicators in the model of 2005-2010

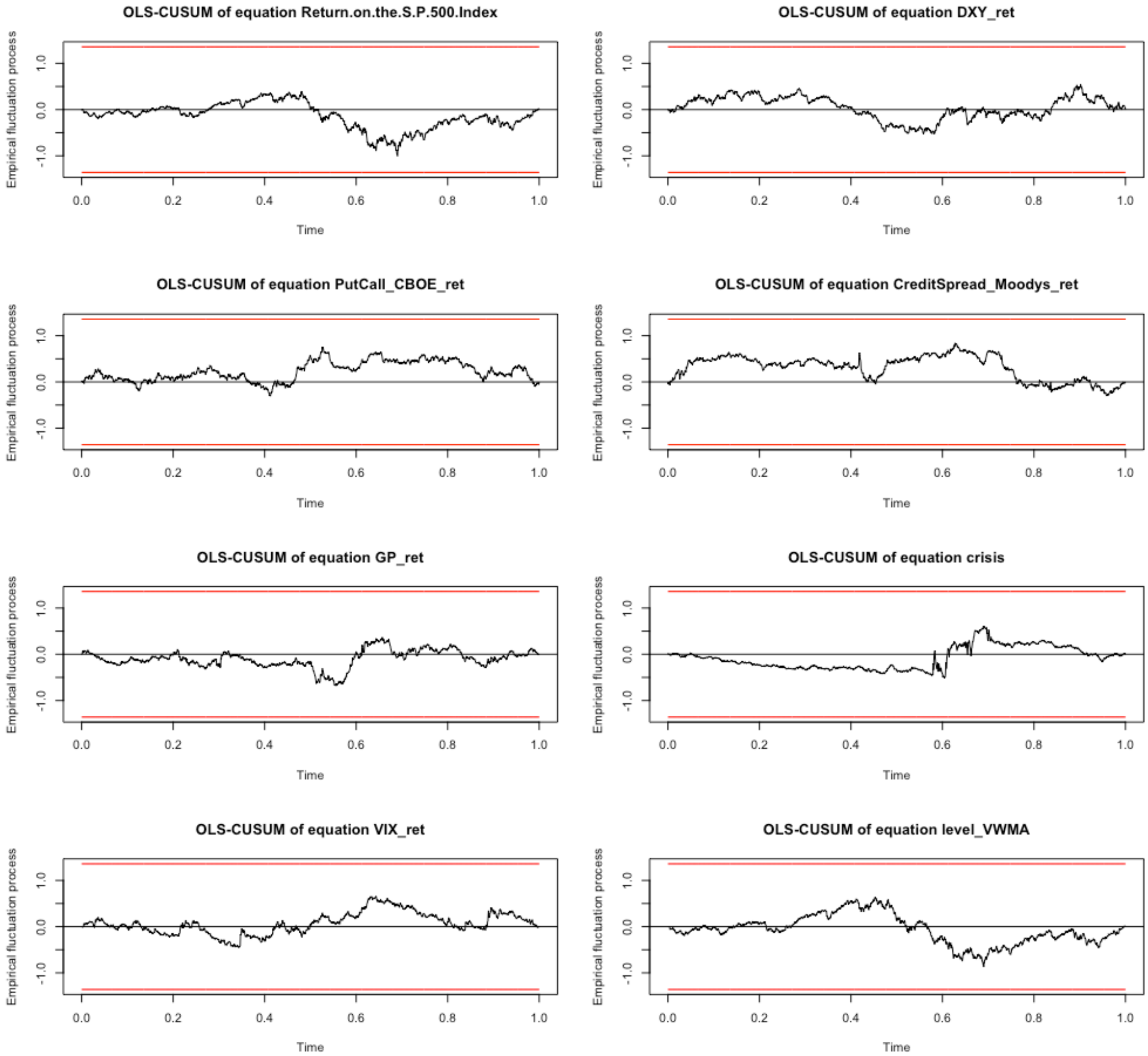


Figure 17: Stability of the indicators in the model of 2005-2010

Stability of the indicators in the model of 2010-2020

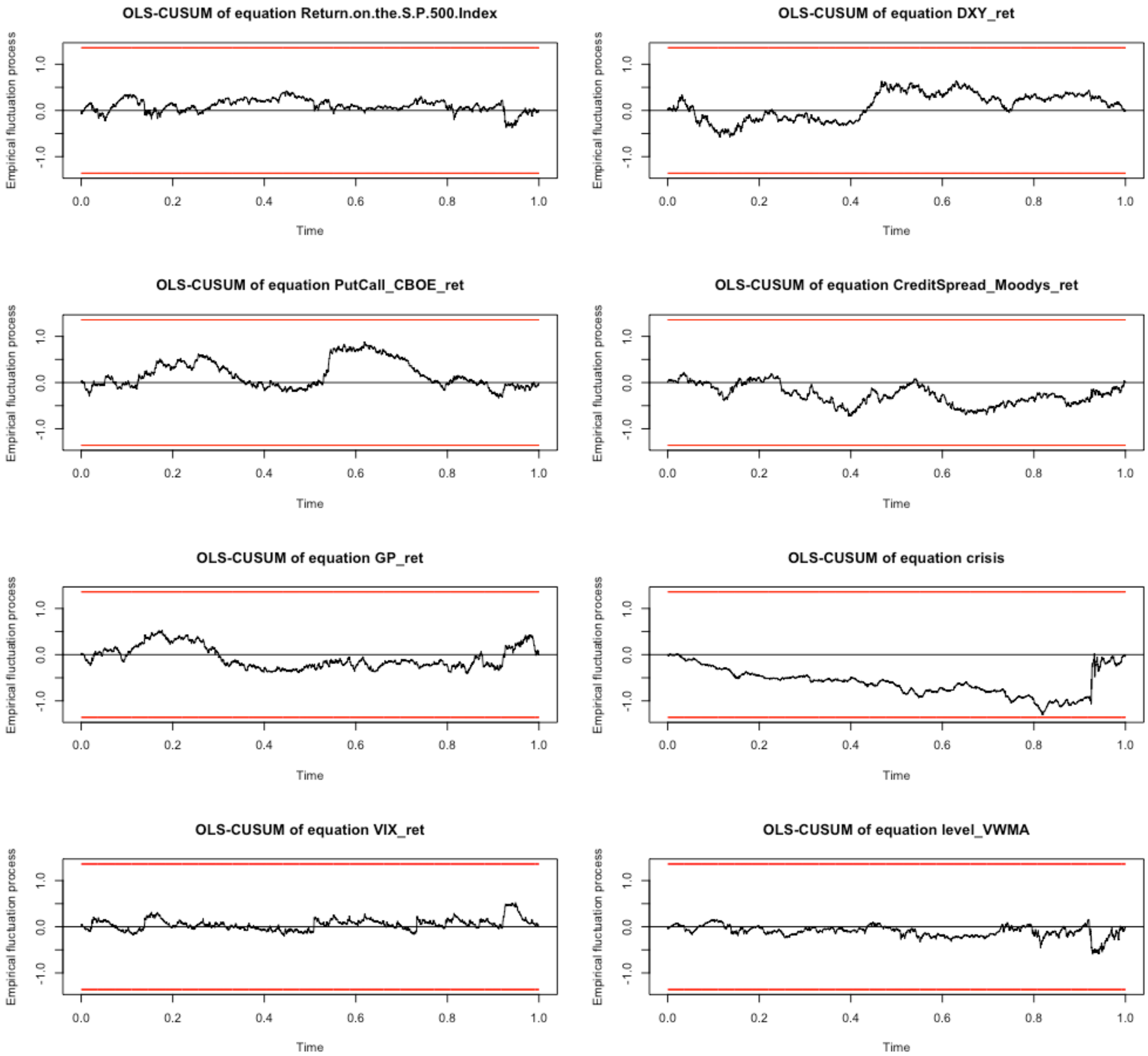


Figure 18: Stability of the indicators in the model of 2010-2020

Stability of the indicators in the model of 2015-2020

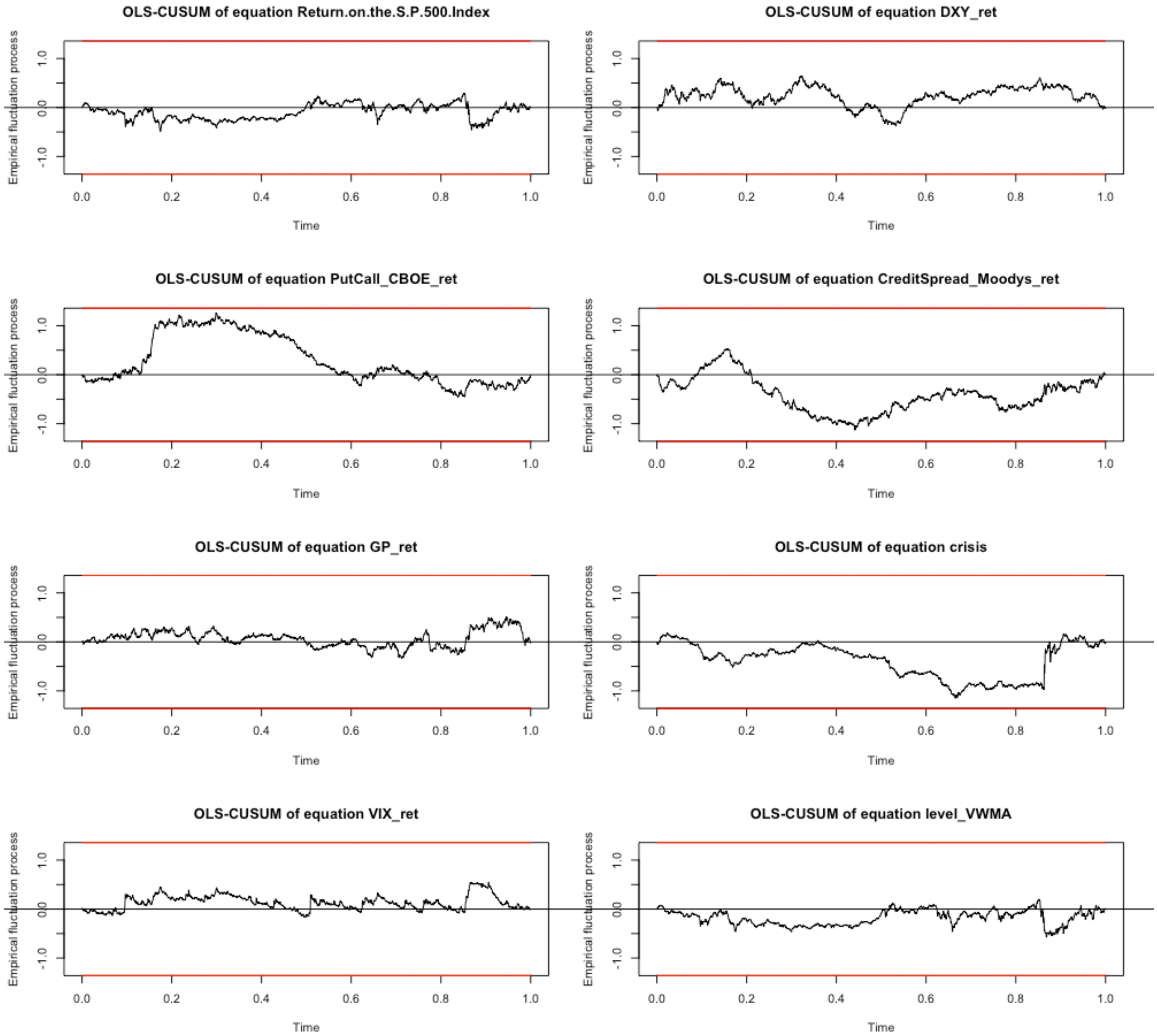


Figure 19: Stability of the indicators in the model of 2015-2020

9.2.6 Whiteness of residuals

Null hypothesis:

Autocorrelation: H_0: No autocorrelation

Heteroscedasticity: H_0: Homoscedasticity / No heteroscedasticity

Normal distributed error term: H_0: Normal distributed error terms

Whiteness of residuals test within the models

	2005-2020		2005-2010		2010-2020		2015-2020	
	P-Value	Critical value	P-Value	Critical value	P-Value	Critical value	P-Value	Critical value
Autocorrelation	< 2.2e-16	2917.2	< 2.2e-16	2902.2	< 2.2e-16	2112	< 2.2e-16	1591.1
Heteroscedasticity	< 2.2e-16	77008	< 2.2e-16	41801	< 2.2e-16	48256	< 2.2e-16	31973
Normal distributed error term	< 2.2e-16	1101383	< 2.2e-16	346240	< 2.2e-16	6103412	< 2.2e-16	637709

Table 12: Whiteness of residuals, lags = optimal

Whiteness of residuals test within the models

	2005-2020		2005-2010		2010-2020		2015-2020	
	P-Value	Critical value	P-Value	Critical value	P-Value	Critical value	P-Value	Critical value
Autocorrelation	< 2.2e-16	19235	1	11752	< 2.2e-16	15923	0.01816	11840
Heteroscedasticity	1	133416	1	42804	1	90108	1	46800

Table 13 : Whiteness of residuals, lags = optimal x10

Whiteness of residuals test within the models

	2005-2020		2005-2010		2010-2020		2015-2020	
	P-Value	Critical value	P-Value	Critical value	P-Value	Critical value	P-Value	Critical value
Autocorrelation	1	31728	1	11752	1	21784	1	11840
Heteroscedasticity	1	133416	1	32724	1	82188	1	40320

Table 14: Whiteness of residuals, lags = optimal x20

9.3 Delta-hedging

Delta hedging is an options trading strategy seeking to be directionally neutral by establishing both long and short positions in the same underlying asset neutralizing each other. The delta is a measure of the sensitivity of the change in the option price in correspondence to the price of the underlying asset, and is given by dividing the first derivative of the theoretical options price by the first derivative of the underlying asset price: (Chen, 2021)

$$\Delta = \frac{\partial V}{\partial S}$$

The delta ranges from -1 to 1, where put options have deltas from -1 to 0 and call options from 0 to 1. For example, if stock options for share X has a delta of 0.50, it implies that an increase of $1 in the underlying asset results in an increase of $0.50 in the option value per share, ceteris paribus.

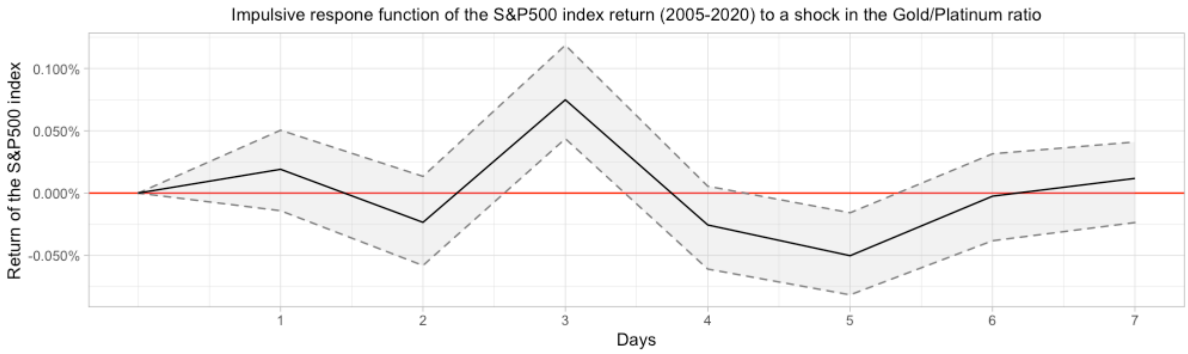
It is important to elaborate on the fact that the delta is not constant as it changes with the price of the underlying asset. The deeper in-the-money the options are, the closer the deltas are to respectively 1 and -1 depending on whether it is a call- or a put option. The same is true in the opposite direction when the options go further out-of-the-money, the delta converge towards 0. This entail that as the price of the underlying increases, so does the delta of the corresponding call options. Thus, resulting in a state where the delta hedging will amplify the ongoing momentum of the underlying asset.

9.4 Partial models

This section denotes results from the models where we included explanatory variables (one by one) and studied the effect they isolated has on the return of the S&P 500 index. The models are presented with their respective impulsive response functions and a statistic summarizing the regression results. Where the statistic reports "Reject" or "Fail to reject" the respective null hypothesis are as follows:

Stationary: H_0: Time series are not stationary.
Granger-causality: H_0: Variable do not Granger-cause S&P500 index return
Stability: H_0: No structural change
Autocorrelation: H_0: No autocorrelation
Heteroscedasticity: H_0: Homoscedasticity / No heteroscedasticity
Normal distributed error term: H_0: Normal distributed error terms

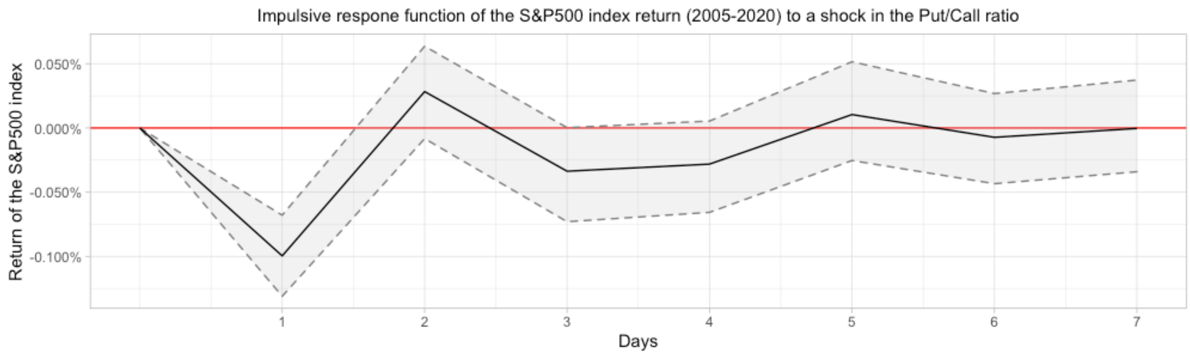
S&P500 and Gold/Platinum



Summary statistic: S&P500 and Gold/Platinum ratio

	Results	P-values	Critical values
Number of lags	27		
Adjusted R squared	6.005%		
Stationarity	Reject	0.01	-4107.839
Granger causality	Reject	< 2.2e-16	3.73
Stability	Fail to reject		
Autocorrelation	Reject	< 2.2e-16	2230.431
Heteroscedasticity	Reject	< 2.2e-16	5191.862
Normal distributed error terms	Reject	< 2.2e-16	22549.474

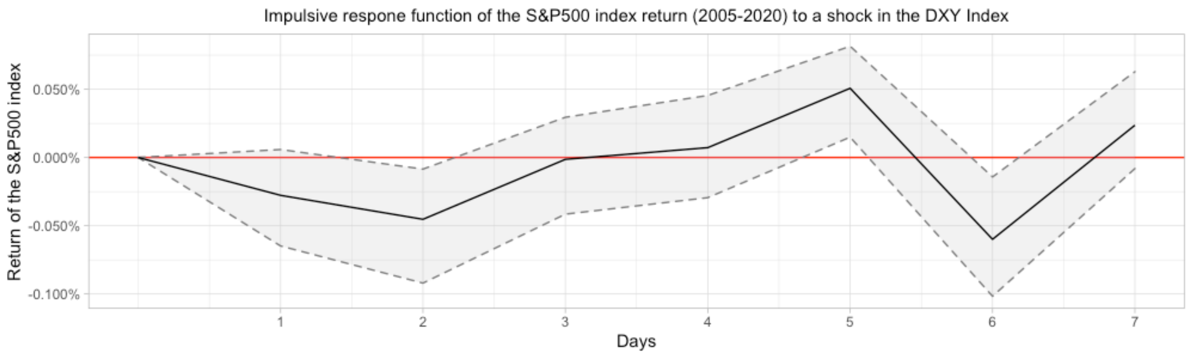
S&P500 and PutCall



Summary statistic: S&P500 and Put/Call ratio

	Results	P-values	Critical values
Number of lags	9		
Adjusted R squared	3.941%		
Stationarity	Reject	0.01	-4107.839
Granger causality	Reject	< 2.2e-16	4.049
Stability	Fail to reject		
Autocorrelation	Reject	0.023	1091.59
Heteroscedasticity	Reject	< 2.2e-16	3831.52
Normal distributed error terms	Reject	< 2.2e-16	19520.447

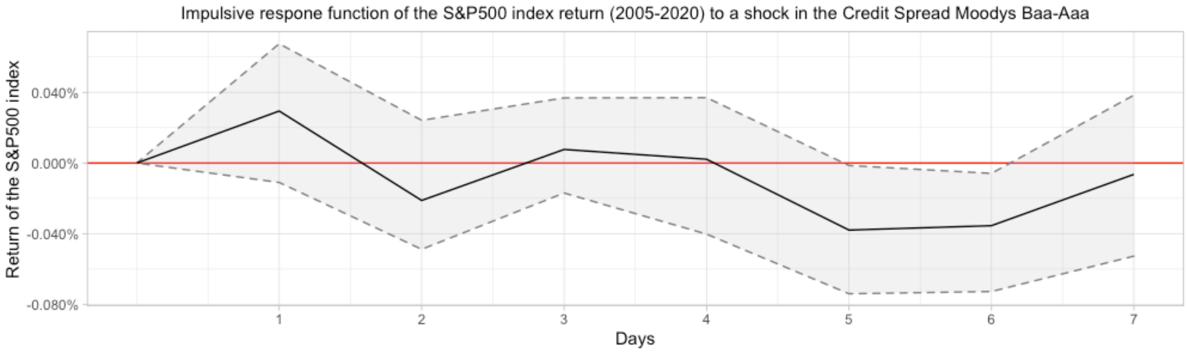
S&P500 and Currency effect (U.S. dollar)



Summary statistic: S&P500 and U.S. dollar Index

	Results	P-values	Critical values
Number of lags	16		
Adjusted R squared	4.301%		
Stationarity	Reject	0.01	-4107.839
Granger causality	Reject	< 2.2e-16	3.272
Stability	Fail to reject		
Autocorrelation	Reject	< 2.2e-16	1207.054
Heteroscedasticity	Reject	< 2.2e-16	4777.916
Normal distributed error terms	Reject	< 2.2e-16	17889.728

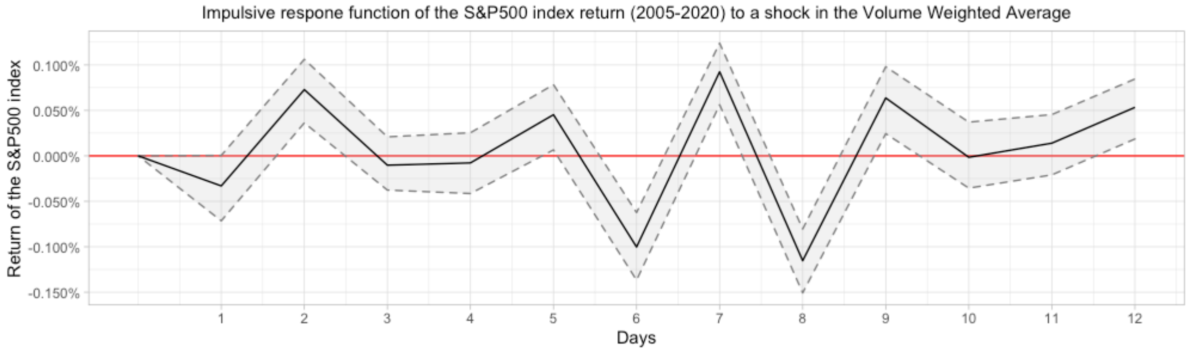
S&P500 and Credit spread



Summary statistic: S&P500 and Credit Spread Moodys

	Results	P-values	Critical values
Number of lags	27		
Adjusted R squared	5.095%		
Stationarity	Reject	0.01	-13.659
Granger causality	Reject	0.004	1.871
Stability	Fail to reject		
Autocorrelation	Fail to reject	0.25	1029.808
Heteroscedasticity	Reject	< 2.2e-16	5409.242
Normal distributed error terms	Reject	< 2.2e-16	146753.588

S&P500 and VWMA



Summary statistic: S&P500 and Volume Weighted Moving Average

	Results	P-values	Critical values
Number of lags	48		
Adjusted R squared	8.771%		
Stationarity	Reject	0.01	-9.526
Granger causality	Reject	< 2.2e-16	10.859
Stability	Reject		
Autocorrelation	Reject	< 2.2e-16	1224.788
Heteroscedasticity	Reject	< 2.2e-16	6343.238
Normal distributed error terms	Reject	< 2.2e-16	90835.166

9.5 Forecast error variance decompositions

9.5.1 FEVD (2005-2010)

Forecast error variance decomposition 2005-2010

Days	Return of the S&P500 index	Percentage change in the Put/Call ratio	Percentage change in the Gold/Platinum ratio	Percentage change in the VIX	Percentage change in the U.S. dollar index	Percentage change in the credit spread Baa-Aaa	Crisis dummy	Deviation from the VWMA
Return of the S&P500 index								
1	100.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%
2	98.854%	0.003%	0.002%	0.067%	0.000%	0.110%	0.006%	0.959%
3	97.018%	0.051%	0.139%	0.196%	0.376%	0.419%	0.420%	1.381%
4	96.611%	0.089%	0.448%	0.206%	0.374%	0.458%	0.423%	1.392%
5	95.833%	0.109%	0.543%	0.228%	0.379%	0.520%	0.468%	1.920%
6	95.502%	0.241%	0.597%	0.333%	0.383%	0.549%	0.468%	1.927%
7	93.606%	0.238%	0.610%	0.476%	1.269%	0.551%	0.533%	2.718%
% change in the Put/Call ratio								
1	19.492%	80.508%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%
2	15.446%	83.716%	0.049%	0.486%	0.086%	0.003%	0.200%	0.013%
3	16.360%	81.725%	0.064%	1.139%	0.088%	0.278%	0.328%	0.018%
4	16.315%	81.505%	0.096%	1.224%	0.113%	0.335%	0.333%	0.080%
5	16.421%	81.345%	0.096%	1.232%	0.114%	0.364%	0.348%	0.081%
6	16.330%	81.028%	0.167%	1.227%	0.118%	0.374%	0.610%	0.146%
7	16.261%	80.398%	0.271%	1.248%	0.291%	0.628%	0.756%	0.148%
% change in the Gold/Platinum ratio								
1	3.743%	0.087%	96.170%	0.000%	0.000%	0.000%	0.000%	0.000%
2	5.082%	0.710%	93.454%	0.106%	0.093%	0.015%	0.369%	0.171%
3	5.094%	0.740%	92.923%	0.119%	0.092%	0.062%	0.457%	0.512%
4	5.290%	0.735%	92.327%	0.210%	0.134%	0.065%	0.606%	0.634%
5	5.368%	1.005%	88.775%	0.206%	0.253%	0.218%	3.062%	1.113%
6	5.350%	1.088%	88.458%	0.205%	0.279%	0.227%	3.187%	1.205%
7	5.308%	1.219%	88.084%	0.217%	0.584%	0.232%	3.161%	1.195%
% change in the VIX								
1	57.238%	0.474%	0.027%	42.260%	0.000%	0.000%	0.000%	0.000%
2	56.661%	0.464%	0.338%	41.775%	0.061%	0.001%	0.009%	0.691%
3	56.224%	0.772%	0.341%	41.675%	0.231%	0.003%	0.020%	0.734%
4	55.810%	0.881%	0.625%	41.380%	0.229%	0.139%	0.156%	0.781%
5	55.153%	1.001%	0.746%	40.904%	0.258%	0.150%	0.169%	1.618%
6	54.753%	1.016%	0.788%	40.916%	0.509%	0.151%	0.260%	1.608%
7	53.956%	1.026%	0.831%	40.378%	1.197%	0.190%	0.424%	1.998%
% change in the U.S. dollar index								
1	2.420%	0.037%	0.164%	0.103%	97.276%	0.000%	0.000%	0.000%
2	4.404%	0.260%	0.329%	0.132%	94.743%	0.006%	0.001%	0.125%
3	4.389%	0.261%	0.352%	0.156%	94.452%	0.072%	0.130%	0.188%
4	4.860%	0.285%	0.538%	0.169%	92.729%	0.411%	0.196%	0.812%
5	4.931%	0.286%	0.558%	0.174%	92.466%	0.470%	0.305%	0.810%
6	5.103%	0.365%	0.595%	0.191%	91.987%	0.486%	0.334%	0.939%
7	5.013%	0.768%	0.607%	0.189%	90.561%	0.530%	0.533%	1.798%
% change in the credit spread Baa-Aaa								
1	1.271%	0.093%	0.070%	0.148%	0.129%	98.289%	0.000%	0.000%
2	1.850%	0.170%	0.252%	0.522%	0.140%	96.641%	0.001%	0.424%
3	2.270%	0.282%	0.253%	0.617%	0.156%	95.935%	0.001%	0.486%
4	3.752%	0.288%	0.251%	0.964%	0.281%	93.764%	0.077%	0.623%
5	3.788%	0.328%	0.504%	1.000%	0.329%	93.267%	0.077%	0.707%
6	3.959%	0.325%	0.542%	1.403%	0.332%	92.483%	0.207%	0.748%
7	4.285%	0.413%	0.565%	1.516%	0.484%	91.792%	0.205%	0.741%

Forecast error variance decomposition 2005-2010

Days	Return of the S&P500 index	Percentage change in the Put/Call ratio	Percentage change in the Gold/Platinum ratio	Percentage change in the VIX	Percentage change in the U.S. dollar index	Percentage change in the credit spread Baa-Aaa	Crisis dummy	Deviation from the VWMA
Crisis dummy								
1	5.683%	0.452%	0.099%	0.907%	0.013%	0.009%	92.837%	0.000%
2	7.235%	0.673%	0.333%	1.513%	0.108%	0.007%	90.016%	0.114%
3	7.645%	0.587%	0.550%	1.604%	0.161%	0.006%	89.351%	0.096%
4	8.797%	0.549%	0.639%	1.631%	0.220%	0.059%	87.853%	0.252%
5	9.137%	0.568%	0.609%	1.654%	0.222%	0.164%	87.242%	0.404%
6	8.488%	0.665%	0.628%	1.521%	0.447%	0.153%	87.216%	0.883%
7	9.736%	0.653%	0.603%	1.698%	0.501%	0.669%	85.260%	0.880%
Deviation from the VWMA								
1	93.192%	0.069%	0.001%	0.715%	0.006%	0.032%	0.001%	5.985%
2	94.474%	0.048%	0.056%	0.519%	0.009%	0.050%	0.005%	4.838%
3	93.650%	0.136%	0.041%	0.403%	0.125%	0.090%	0.170%	5.384%
4	93.187%	0.262%	0.142%	0.327%	0.199%	0.079%	0.262%	5.542%
5	92.026%	0.418%	0.120%	0.350%	0.285%	0.069%	0.245%	6.487%
6	91.237%	0.683%	0.115%	0.305%	0.314%	0.061%	0.233%	7.052%
7	90.897%	0.982%	0.160%	0.285%	0.713%	0.063%	0.208%	6.691%

Table 15: Forecast error variance decomposition (2005-2010)

9.5.2 FEVD (2010-2020)

Forecast error variance decomposition 2010-2020

Days	Return of the S&P500 index	Percentage change in the Put/Call ratio	Percentage change in the Gold/Platinum ratio	Percentage change in the VIX	Percentage change in the U.S. dollar index	Percentage change in the credit spread Baa-Aaa	Crisis dummy	Deviation from the VWMA
Return of the S&P500 index								
1	100.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%
2	97.787%	0.816%	0.010%	0.064%	0.089%	0.009%	1.214%	0.012%
3	94.834%	0.987%	0.446%	0.111%	0.118%	0.020%	3.468%	0.015%
4	93.467%	1.025%	0.476%	0.223%	0.293%	0.432%	3.572%	0.512%
5	92.967%	1.092%	0.480%	0.244%	0.292%	0.459%	3.725%	0.740%
6	91.958%	1.134%	0.746%	0.269%	0.430%	0.665%	3.695%	1.104%
7	91.361%	1.127%	0.739%	0.455%	0.502%	0.726%	3.847%	1.243%
% change in the Put/Call ratio								
1	26.612%	73.388%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%
2	23.675%	75.942%	0.056%	0.153%	0.043%	0.008%	0.001%	0.123%
3	23.892%	75.624%	0.067%	0.157%	0.074%	0.015%	0.044%	0.127%
4	23.975%	75.444%	0.069%	0.159%	0.075%	0.036%	0.105%	0.136%
5	23.949%	75.316%	0.084%	0.184%	0.147%	0.040%	0.139%	0.140%
6	23.936%	75.266%	0.087%	0.208%	0.147%	0.047%	0.152%	0.158%
7	23.975%	75.058%	0.152%	0.257%	0.200%	0.049%	0.151%	0.158%
% change in the Gold/Platinum ratio								
1	14.392%	0.148%	85.461%	0.000%	0.000%	0.000%	0.000%	0.000%
2	14.342%	0.279%	84.401%	0.122%	0.443%	0.002%	0.395%	0.016%
3	14.095%	0.327%	82.015%	0.384%	0.437%	0.034%	2.607%	0.099%
4	14.145%	0.346%	81.214%	0.385%	0.433%	0.506%	2.721%	0.250%
5	14.102%	0.381%	81.014%	0.383%	0.474%	0.519%	2.821%	0.306%
6	14.084%	0.429%	79.840%	0.378%	0.490%	0.788%	2.960%	1.031%
7	13.917%	0.430%	79.117%	0.376%	0.604%	0.816%	3.672%	1.069%

Forecast error variance decomposition 2010-2020

Days	Return of the S&P500 index	Percentage change in the Put/Call ratio	Percentage change in the Gold/Platinum ratio	Percentage change in the VIX	Percentage change in the U.S. dollar index	Percentage change in the credit spread Baa-Aaa	Crisis dummy	Deviation from the VWMA
% change in the VIX								
1	61.093%	0.441%	0.015%	38.451%	0.000%	0.000%	0.000%	0.000%
2	60.932%	0.778%	0.113%	38.047%	0.019%	0.074%	0.037%	0.000%
3	60.430%	0.854%	0.412%	37.815%	0.034%	0.104%	0.344%	0.007%
4	60.083%	0.854%	0.542%	37.575%	0.146%	0.261%	0.363%	0.175%
5	59.776%	0.871%	0.624%	37.509%	0.178%	0.401%	0.422%	0.220%
6	59.466%	1.178%	0.705%	37.244%	0.237%	0.418%	0.420%	0.331%
7	59.366%	1.174%	0.784%	37.103%	0.244%	0.437%	0.438%	0.453%
% change in the U.S. dollar index								
1	2.746%	0.012%	0.391%	0.494%	96.356%	0.000%	0.000%	0.000%
2	2.818%	0.012%	0.415%	0.646%	95.216%	0.007%	0.613%	0.273%
3	2.814%	0.034%	0.415%	0.677%	95.097%	0.015%	0.638%	0.311%
4	2.823%	0.055%	0.419%	0.809%	94.538%	0.165%	0.816%	0.375%
5	2.820%	0.072%	0.423%	0.835%	94.268%	0.308%	0.887%	0.388%
6	2.828%	0.072%	0.424%	0.864%	94.088%	0.307%	0.929%	0.487%
7	2.820%	0.161%	0.435%	0.861%	93.599%	0.713%	0.926%	0.485%
% change in the credit spread Baa-Aaa								
1	0.050%	0.034%	0.007%	0.037%	0.258%	99.614%	0.000%	0.000%
2	0.202%	0.062%	0.061%	0.037%	0.253%	99.111%	0.274%	0.001%
3	0.693%	0.085%	0.061%	0.036%	0.255%	98.533%	0.285%	0.052%
4	0.999%	0.085%	0.113%	0.040%	0.348%	98.023%	0.315%	0.075%
5	1.004%	0.169%	0.242%	0.040%	0.454%	97.538%	0.475%	0.077%
6	1.556%	0.168%	0.716%	0.068%	0.433%	91.109%	5.731%	0.220%
7	1.696%	0.170%	0.733%	0.068%	0.425%	88.915%	7.775%	0.218%
Crisis dummy								
1	3.320%	0.453%	0.993%	0.441%	0.163%	0.127%	94.503%	0.000%
2	3.695%	0.440%	0.832%	0.509%	0.192%	0.395%	93.923%	0.013%
3	3.448%	0.592%	0.962%	0.352%	0.161%	0.280%	94.141%	0.064%
4	7.433%	0.896%	0.914%	1.112%	0.377%	0.941%	86.862%	1.465%
5	7.444%	0.865%	0.837%	0.966%	0.334%	1.013%	87.271%	1.269%
6	8.545%	0.883%	0.922%	0.874%	0.353%	1.351%	85.495%	1.576%
7	8.728%	0.914%	0.858%	0.774%	0.380%	1.811%	85.050%	1.485%
Deviation from the VWMA								
1	87.139%	0.003%	0.031%	0.052%	0.473%	0.008%	0.203%	12.089%
2	86.314%	0.485%	0.021%	0.276%	0.313%	0.005%	0.480%	12.106%
3	85.508%	0.407%	0.388%	0.455%	0.217%	0.003%	0.701%	12.321%
4	84.085%	0.404%	0.405%	0.819%	0.236%	0.101%	0.592%	13.357%
5	83.353%	0.520%	0.371%	1.233%	0.250%	0.093%	0.669%	13.511%
6	82.011%	0.537%	0.557%	1.389%	0.382%	0.198%	0.674%	14.252%
7	81.179%	0.551%	0.656%	1.769%	0.408%	0.343%	0.608%	14.486%

Table 16: Forecast error variance decomposition (2010-2020)

9.5.3 FEVD (2015-2020)

Forecast error variance decomposition 2015-2020

Days	Return of the S&P500 index	Percentage change in the Put/Call ratio	Percentage change in the Gold/Platinum ratio	Percentage change in the VIX	Percentage change in the U.S. dollar index	Percentage change in the credit spread Baa-Aaa	Crisis dummy	Deviation from the VWMA
Return of the S&P500 index								
1	100.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%
2	96.509%	0.874%	0.076%	0.648%	0.221%	0.142%	1.422%	0.108%
3	91.515%	0.918%	1.487%	0.663%	0.210%	0.148%	4.658%	0.401%
4	90.274%	0.913%	1.812%	0.847%	0.295%	0.471%	4.780%	0.607%
5	89.291%	1.073%	1.843%	0.956%	0.304%	0.969%	4.924%	0.640%
6	87.486%	1.136%	2.511%	1.097%	0.466%	1.698%	4.940%	0.666%
7	87.015%	1.128%	2.502%	1.398%	0.505%	1.679%	5.001%	0.772%
% change in the Put/Call ratio								
1	25.262%	74.738%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%
2	21.897%	77.435%	0.058%	0.181%	0.208%	0.000%	0.000%	0.221%
3	22.171%	76.971%	0.074%	0.219%	0.296%	0.027%	0.003%	0.238%
4	22.226%	76.600%	0.166%	0.315%	0.294%	0.041%	0.016%	0.341%
5	22.161%	76.386%	0.165%	0.322%	0.430%	0.048%	0.137%	0.352%
6	22.074%	76.153%	0.188%	0.352%	0.466%	0.108%	0.169%	0.488%
7	22.057%	75.936%	0.208%	0.402%	0.540%	0.110%	0.173%	0.573%
% change in the Gold/Platinum ratio								
1	14.617%	0.153%	85.230%	0.000%	0.000%	0.000%	0.000%	0.000%
2	14.419%	0.235%	84.109%	0.182%	0.551%	0.167%	0.334%	0.002%
3	14.121%	0.274%	81.038%	0.484%	0.533%	0.195%	3.029%	0.326%
4	14.126%	0.272%	79.536%	0.475%	0.527%	1.481%	3.217%	0.365%
5	14.106%	0.433%	79.072%	0.472%	0.545%	1.473%	3.492%	0.406%
6	14.565%	0.765%	76.924%	0.464%	0.565%	2.240%	3.574%	0.904%
7	14.394%	0.856%	76.233%	0.463%	0.662%	2.245%	4.133%	1.014%
% change in the VIX								
1	57.634%	0.626%	0.005%	41.735%	0.000%	0.000%	0.000%	0.000%
2	57.317%	1.095%	0.042%	41.156%	0.047%	0.297%	0.039%	0.007%
3	56.373%	1.265%	0.819%	40.565%	0.275%	0.340%	0.268%	0.095%
4	55.876%	1.429%	1.324%	40.200%	0.404%	0.380%	0.289%	0.098%
5	55.469%	1.471%	1.523%	39.904%	0.441%	0.750%	0.344%	0.098%
6	54.966%	2.061%	1.718%	39.546%	0.437%	0.826%	0.348%	0.097%
7	55.005%	2.067%	1.786%	39.393%	0.454%	0.838%	0.351%	0.106%
% change in the U.S. dollar index								
1	0.010%	0.013%	1.946%	0.422%	97.609%	0.000%	0.000%	0.000%
2	0.461%	0.100%	1.925%	0.617%	95.331%	0.011%	1.345%	0.211%
3	0.465%	0.159%	1.939%	0.631%	95.189%	0.018%	1.389%	0.210%
4	0.564%	0.162%	1.926%	0.875%	94.547%	0.129%	1.550%	0.247%
5	0.569%	0.184%	2.207%	0.887%	94.026%	0.205%	1.675%	0.247%
6	0.761%	0.202%	2.208%	0.885%	93.764%	0.213%	1.683%	0.283%
7	0.774%	0.560%	2.199%	0.974%	92.843%	0.671%	1.691%	0.288%
% change in the credit spread Baa-Aaa								
1	0.003%	0.012%	0.018%	0.104%	0.000%	99.863%	0.000%	0.000%
2	0.103%	0.315%	0.246%	0.332%	0.012%	98.721%	0.200%	0.070%
3	0.621%	0.317%	0.411%	0.331%	0.047%	97.981%	0.222%	0.070%
4	0.780%	0.399%	0.469%	0.338%	0.288%	97.353%	0.265%	0.108%
5	0.820%	0.493%	0.580%	0.337%	0.301%	96.390%	0.891%	0.188%
6	2.341%	0.450%	1.467%	0.304%	0.347%	86.777%	7.752%	0.563%
7	2.394%	0.435%	1.490%	0.294%	0.342%	83.716%	10.786%	0.543%

Forecast error variance decomposition 2015-2020

Days	Return of the S&P500 index	Percentage change in the Put/Call ratio	Percentage change in the Gold/Platinum ratio	Percentage change in the VIX	Percentage change in the U.S. dollar index	Percentage change in the credit spread Baa-Aaa	Crisis dummy	Deviation from the VWMA
Crisis dummy								
1	4.498%	0.302%	1.615%	1.086%	0.072%	0.022%	92.405%	0.000%
2	4.884%	0.266%	1.398%	1.218%	0.086%	1.086%	91.049%	0.013%
3	4.496%	0.496%	1.500%	0.865%	0.151%	0.761%	91.546%	0.186%
4	10.913%	0.905%	1.313%	2.270%	0.159%	2.663%	81.466%	0.309%
5	10.690%	0.829%	1.307%	1.998%	0.129%	3.100%	81.675%	0.271%
6	12.538%	0.893%	1.224%	1.883%	0.118%	4.065%	79.004%	0.275%
7	12.753%	0.952%	1.062%	1.648%	0.156%	5.000%	78.190%	0.239%
Deviation from the VWMA								
1	96.250%	0.002%	0.005%	0.017%	0.006%	0.001%	0.000%	3.718%
2	94.543%	0.540%	0.015%	0.261%	0.154%	0.116%	0.732%	3.638%
3	92.918%	0.477%	0.565%	0.435%	0.195%	0.157%	0.614%	4.640%
4	91.522%	0.417%	0.511%	0.844%	0.153%	0.410%	0.462%	5.680%
5	90.038%	0.537%	0.432%	1.413%	0.127%	0.353%	0.431%	6.669%
6	88.669%	0.536%	0.612%	1.564%	0.150%	0.532%	0.370%	7.566%
7	87.629%	0.521%	0.669%	2.048%	0.141%	0.679%	0.333%	7.980%

Table 17: Forecast error variance decomposition (2015-2020)