Master Thesis

Predicting Market Crashes Using Volatility

An empirical approach using Investor Sentiment, Risk Taking Behavior and Volatility in the US, Europe and China

Author:
Robbert van Genuchten
ANR: 609888

Supervisors:
dr. Ole Wilms
dr. Christoph Schneider

A Master Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Finance

in the

Department of Finance
School of Economics and Management

November 19, 2018

Preface

I proudly present my master thesis as the final piece of work during my time as a student. I already wrote a similar preface for my master thesis Operations Management & Logistics, but this time, it will surely be my last.

The journey that led to this moment did not go by without the help, support, and unconditional love of some people. Therefore, I want to take this opportunity to thank a number of people.

First of all, I would like to thank my university supervisor Ole Wilms. Even though we did not meet that frequently, your feedback was always to the point and constructive. During our Skype meetings you were open to discussion and would make suggestions in order to increase the quality of the thesis. Lastly, it would never take you more than just a few days to present your feedback. It was a pleasure to work with you.

While my life as a student will come to an end, I can look back at a period which has been filled with much more than just studying. I would like to thank all my friends that made my time as a student unforgettable. Special thanks to Rutger Bruijnzeel for reviewing the final report.

I want to thank my girlfriend Renée Albers for her love and support in the last couple of years. It is amazing to share so many beautiful experiences together including graduating. I am looking forward to the times to come and the places it will take us.

Finally and again, I want to express my gratitude to my parents. Without your unconditional love and encouragement I would have never be able to achieve this. You always believed in me and supported the decisions that I made throughout. I am so grateful that you gave me the opportunity to enjoy a board year, a semester in the USA, and enroll in a secondary master program. I could not have wished for a more wonderful family.

Robbert van Genuchten, November 2018

Abstract

This thesis attempts to explain and predict financial market crashes using measures of volatility, investor sentiment and risk taking in the US, Europe and China. The empirical analysis is following the reasoning of Minsky's Financial Instability Hypothesis. He states that in times of low volatility, a state of financial euphoria emerges. Investors who do not want to miss out on potential profits tend to show more risky behavior creating an unstable market. When the market starts to decline, this instability resolves into market crashes. This research is split up into two parts: a explanatory study that uses monthly increments and evaluates relations between variables in hindsight, and a quarterly study that attempts to predict market crashes using the same variables. Using logistic regression on a binary crash variable, it is demonstrated that low volatility indeed is a good explanatory and predictive variable of market crashes. If the current period shows excessively low volatility, the following period has an increased chance of facing a market crash of around 4% on average. This result revealed to be robust since significant effects are found in samples from the US, Europe and China while adding control variables to the model. Continuing to find evidence for Minsky's Hypothesis, the Credit-to-GDP Ratio appeared to be lacking the risk taking phenomenon as Minsky described it. Therefore, this thesis does not find complete support for the Financial Instability Hypothesis. Market volatility did, however, have some significant effect on investor sentiment. Implying that when volatility is high, the investor tends to be more bearish in the periods after. Concluding, this thesis provides evidence that implied volatility holds explanatory and predictive properties that helps understanding stock market crashes.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

According to the past, times of a booming economy and deep recessions have been existing interchangeably. The way banks and individuals have been investing their capital is varying over time as well. In periods of financial melancholy, investors tend to be issuing more debt in order to increase the investments and not miss out on the profits. In these times, the volatility measures that can be seen as the price that an investor is willing to pay to protect against a market crash seems to decrease. This is interesting since when more and more debt is issued for investments, the actual risk is rising because of interconnectedness in the market. So, while the market traded price of risk is declining, financial instability is actually growing. This phenomenon is described by Minsky (1977). Even though his views are often considered radical, Minsky believed in the fact that financial systems are inherently susceptible to bouts of speculation that, if they last long enough, end in crises. This implies that after periods of low interest rates and optimism, euphoria leads to more borrowing which creates an unstable economy. This is supposed to lead to market declines and eventually market crashes. Such a series of events is commonly refereed to as a 'Minsky Moment' (Lahart, 2007).

Since times of financial instability and market crashes have massive impact on the overall level of wealth, one should try to do everything in its power to prevent such an event from happening. Therefore, the question arises, if the level of market volatility can be used in order to see market instabilities and eventually market crashes coming? This, to not only be able to explain the cause in hindsight, but especially to be able to predict a prospective crash.

A recent paper written by Danielsson, Valenzuela, and Zer (2018) investigates this interesting relationship of volatility and market returns. They perform a cross-country study that ranges over 211 years of data. They conclude that low volatility has strong predictive power and is a reliable crash indicator. Additionally, they successfully conclude that periods of low volatility leads to excessive credit buildup aiding the market instability. This study uses yearly data increments to evaluate the relationships. It is believed that adding additional factors which might explain (parts of) Minsky's Hypothesis and using smaller time increments could yield additional relevant results.

This study evaluates the Minsky Hypothesis using measures for market returns, volatility, investor sentiment and risk taking. These variables are used to form an all encompassing picture. Investor sentiment is used to reveal if times of low volatility indeed trigger more euphoric investor behavior. Then, it is evaluated if this level of sentiment and volatility has an impact on the amount of risk that is taken by the investor. By studying the relationship of the amount of risk taking and market returns, it is assessed if this increase in the supposed market instability has an effect on market returns. Lastly, the direct effect of volatility on market returns (and crashes) is evaluated and tested on its explanatory and predictive power. To do so, it is decided to split the analysis up into two parts: a monthly study that tends to reveal explanatory power of the evaluated factors, and a quarterly study that assesses the predictive properties of these variables.

Seven models are used to evaluate the hypotheses that aim to confirm the Minsky hypothesis. Six data samples originating from different areas around the world are used to do so. If multiple significant relationships are found within these six samples, the results' validity is higher when one would work with only a single sample. This thesis uses a combination of OLS regression and Logistic regression. The latter to test the effect of volatility on a binary crash variable.

1.2 Project Organization

The remainder of the thesis is structured in the following way: Chapter 2 will review relevant scientific literature. This, to broaden the understanding of the subject and formulate appropriate research questions and goals. In Chapter 3 these research questions are introduced. Additionally, the thesis methodology is outlined. Some necessary scoping is done to make sure the project remains inside the bounds of a master thesis. Lastly, that chapter introduces the to be tested hypotheses. The consecutive chapter (Chapter 4) discloses the used data for the analysis. Some data transformations are required before being fit for analysis. The steps in gathering, transforming and finalizing the factors are presented there. Chapter 5 introduces the regression models that are used to test the hypotheses. The chapter consecutively reveals the results for each of the models. The chapter finishes with an overview of the found results. The results found are then translated in to conclusions and recommendations in Chapter 6. Finally, Chapter 7 will reflect on the thesis by reviewing the scientific contribution, its limitations and acknowledging areas for future research.

Chapter 2

Literature Review

2.1 Introduction

This chapter presents a brief review of existing literature. This way, the theoretical framework in which this study exists is set. Major work in the field of market crashes, bubbles and volatility are discussed. In doing so, research methods and techniques are evaluated and will help shape the approach of this thesis. Lastly, and most importantly, getting familiar with the existing scientific research will help to expose a possible gap in literature which will increase the relevance of this study.

2.2 Market Crashes

2.2.1 What are Market Crashes?

Although in some classical theoretical models financial market crashes are non-existent or are lacking important variables to determine them (Wharton, 2009), history teaches us that this is far from the truth. But what exactly is a market crash? Sornette (2003) defines a market crash as a rapid and mostly unanticipated decline in the market prices of stocks in a short period of time. Such a scenario arises when prices not only reflect the individual market participant's valuation of the asset, but also reflect an extra amount that is often called a 'bubble' (Youssefmir, Huberman, and Hogg, 1998). This mostly results in big declines in 'paper wealth' which is nowadays even more inflated because of the interconnectedness of the market (Aliber and Kindleberger, 2015).

2.2.2 Existence of Market Crashes

History teaches us that market crashes exist. Multiple articles evaluate the main financial crises that have occurred in the past. Sornette (2003) lists some of the main financial crises starting as early as 1637 with the Tulip Mania in The Netherlands. Here, price speculation and over confidence in the prospects of tulip created inflated prices that were irrational to all equivalents. Eventually, the price of the tulip became unstable when first mentioning its uncertainty up until the point that they were practically valueless. From that point onwards, multiple other bubbles and crashes have occurred. Figure 2.1 lists the ten most impactful crashes in history today.
Stock market crashes have been fascinating events to a lot of academics (Sornette, 2003). In order to understand and preferably overcome the next stock market crash, these historical events are analyzed thoroughly. Neuhauser (2015) reviews the financial crisis of 2008 extensively. She gives an overview of all research done that try to explain the tragic events. For every major financial market crash some similarities can be recognized in the period ramping-up to the bursting of the bubble: times of growing economy boost the confidence of investors, creating a state of financial melancholy, investors starting to become more risky, prices will start to get inflated and eventually collapse (Aliber and Kindleberger, 2015).

2.2.3 Financial Instability Hypothesis

While trying to find explanation and understanding of financial crises, Hyman Minsky formulated the Financial Instability Hypothesis. This economic theory, which is an interpretation of the substance of Keynes's 'General Theory' (Keynes, 1936), supports the idea of some government intervention. Minsky states that in prosperous times, when the market returns are increasing,

The big ten financial bubbles

1. The Dutch Tulip Bulb Bubble 1636
2. The South Sea Bubble 1720
3. The Mississippi Bubble 1720
4. The late 1920s stock price bubble 1927–29
5. The surge in bank loans to Mexico and other developing countries in the 1970s
6. The bubble in real estate and stocks in Japan 1985–89
7. The 1985–89 bubble in real estate and stocks in Finland, Norway, and Sweden
8. The bubble in real estate and stocks in Thailand, Malaysia, Indonesia, and several other Asian countries 1992–97 and the surge in foreign investment in Mexico 1990–94
9. The bubble in over-the-counter stocks in the United States 1995–2000
10. The bubble in real estate in the United States, Britain, Spain, Ireland, and Iceland between 2002 and 2007

Figure 2.1: The Big Ten Financial Bubbles (Aliber and Kindleberger, 2015)

a speculative euphoria develops, this increases the tendency to take bigger risks which results in big market declines (Minsky, 1977). This slow movement from stability to fragility followed by a crisis is also known as a 'Minsky Moment'. Such a moment is created by a key mechanism of the accumulation of debt by the non-government sector. According to Minsky (1992), there are three distinct income-debt relations: Hedge, Speculative and Ponzi units. Hedge financing units can fulfill all of their contractual payment obligations. Speculative finance units meet their payment commitments on an income account on their liabilities. These units need to issue new debt to meet the commitments on maturing debt (Minsky, 1992). Ponzi units cannot fulfill the repayment of principle or interest by the cash flows coming from operations. They can either sell assets or borrow. This lowers the margin of safety that it offers to the holders of its debts (Minsky, 1992). Minsky argues that financial systems that are dominated by hedging units can fall into Ponzi units as asset values start to fall. Overstretched firms need to start selling their positions taking down first Ponzi units and later the Speculative units. If this process continues, even the the Hedging borrowers will be affected (Minsky, 1992).

2.2.4 Predicting Market Crashes

Especially after the mortgage crisis of 2008, researchers have started to ask the question if it is possible to predict market crashes. Multiple different approaches have been examined. Lleo and Ziemba (2012) are using the 'Bond Stock Earnings Yield Differential' (BSEYD) Model as the predictor of market crashes. It turns out that this model is able to predict some of the major market crashes of the past. Especially the ones that had a high interest rate compared to the relative earnings (Lleo and Ziemba, 2012).

Another method that is applied multiple times to predict market crashes is the 'Log-Periodic Power Law' (LPPL) presented by Sornette, Johansen, and Bouchaud (1996) and Feigenbaum and Freund (1996). They proposed that in times of a speculative bubble, an economic index increases as a power law decorated with a log-periodic oscillation and that the ending crash is the climax of the so called Log-Periodic Power Law signature (Jacobsson, 2009). This model is able to identify signatures of near-critical behavior just before the crash, making it a predictive tool. Jiang et al. (2010) and Zhou and Sornette (2006) use this model to predict historical crashes using multiple different factors. The latter concludes that volatility played a key role in predicting the dot-com bubble. Then, lastly, as early as 1982, Blanchard and Watson (1982) already concluded that financial bubbles could be predicted. If the right information is available, such as returns, financial crises can be statistically detected. This implies that nowadays, where a lot of information is easy accessible and real time updated, financial instabilities can be foreseen and maybe even prevented.

2.3 Volatility

Volatility figures are widely available as a measure for dispersions of the returns of a security or a market index. Overall it implies that the higher the volatility the higher the risk accompanied with the relative asset or index. Scientist have been investigating market volatility and their movements. Schwert (1989) seeks answers on what determines changes in the market volatility over time. He investigates the relationship of volatility with economic activity, financial leverage, and stock trading activity. The conclusion of the article states that it is difficult to explain the fluctuations in market volatility. However, on the contrary, researchers claim that volatility indices have predictive power. For instance, the VIX (a volatility index) is commonly referred as the 'fear gauge'. Carr (2017) claims that there is substantial evidence that the VIX should be called this way. He mentions the leverage effect which implies that abnormally high VIX levels tend to be accompanied by abnormally low S&P 500 levels (Carr, 2017). Additionally, Fleming, Ostdiek, and Whaley (1995) mention that the VIX has some significant predictive properties. They notice that the VIX can forecast future volatility and, even more essential, can predict stock market returns. A counter statement is made by Bekaert and Hoerova (2014), they argue that the VIX tells very little about what the market is going to do tomorrow or the day after. They come up with this statement when they compare the predictive power of a conditional variance premium to the more often used variance premium VIX. They decompose the squared VIX into two components, the conditional variance of the stock market (CV) and the equity variance premium (VP), which is the difference between the two (Bekaert and Hoerova, 2014). This decomposed measure (CV) holds more predictive power compared to the VIX. Therefore, Bekaert and Hoerova (2014) conclude that the VIX is just a mirror of the market trend.

2.3.1 Volatility on Crashes

Unless striking views on volatility, some research has been done to find a relationship between volatility measures and market crashes. Tsuji (2003) reviews if volatility can be used as a predictor of market crashes. The findings, however, imply that market liquidity predicts market crashes way better than volatility. Despite these results, Danielsson, Valenzuela, and Zer (2018) conduct additional research that concludes that volatility itself is not a significant predictor of banking crises, but unusually high and low volatility are. They introduce a model which uses a binary crash variable that indicates if the specific year experienced a market crash. Using the one-sided Hodrick and Prescott filter (See Appendix F), they transform their volatility measure into two channels: high volatility and low volatility. They use logistic regression to evaluate 211 years of cross country data to draw their conclusion. The results of this paper seem promising. However, yearly time increments in the specific study could be viewed as too large. Even tough market crashes can have long lasting effects, an analysis using more detailed time increments could reveal a different outcome.

2.3.2 Asymmetric Volatility Phenomenon

Volatility in relation to its stock price (also known as the risk-return relationship) is one of the most researched fields within Finance although it remains a controversial topic (Al Refai and Abdelaziz Eissa, 2017). The so called 'Asymmetric Volatility Phenomenon' covers the negatively correlated relationship between market returns and the implied volatility (Bekaert, 2000). In other words, this phenomenon covers the tendency of volatility to be higher in declining markets compared to growing markets. This is an interesting given since in the classical financial models such as the CAPM model of Merton (1973) clearly formulate a positive relationship between risk and return. In this theory, that exists of risk averse investors, holding a more risky position should offer a higher return than a less risky position (Al Refai and Abdelaziz Eissa, 2017). Despite these classical models, studies found some logical explanations why the Asymmetric Volatility Phenomenon exists. Christie (1982), argues that the increase in volatility is a result of the leverage effect as defined by Black (1976). Here, a drop in the market price makes a company riskier since this triggers an increase in the debt to equity ratio and therefore increases its stock volatility. Despite this relationship, Schwert (1989) acknowledge the fact that volatility, especially in times of recession, only explains a small portion of the changes in stock volatility over time (Schwert, 1989).

There are also papers that find both relationships significant. Glosten, Jagannathan, and Runkle (1993) show that the behavior of volatility after an increase (or decrease) in stock price can vary (either increase or decrease). Nevertheless, all researched relationships focus on the behavior of volatility after a change in market returns. What is most relevant for this thesis, is the behavior of market returns depending on volatility.

2.4 Other Factors

Since volatility is not the only measure which might be relevant to evaluate if stock market crashes can be predicted, two other factors are also considered. As argued before, times of melancholy in the market might trigger a more risky behavior. Therefore, it might be relevant to include a factor for investor sentiment and risk taking.

2.4.1 Investor Sentiment

Investor sentiment tends to track the overall investor attitude towards the market. Investors can either feel bullish or bearish. A bullish investor believes that the stock price will rise over time whereas a bearish investor believes that the stock price will decline over time. Beer and Zouaoui (2011) try to find the most accurate investor sentiment measures and review a number of popular measures. While doing so, the distinction between direct and indirect investor sentiment measures is made. Direct measures are mostly coming from surveying investors asking they what they think about the current market. They can reveal beneficial information about the current state of mind of the investor without having to use sophisticated financial theory to validate them (Beer and Zouaoui, 2011). Unfortunately, Clarke and Statman (1998), among others, conclude that direct measures of investor sentiment are not always useful. They suffer from a limited sample size, respondent bias due to the tendency of answering more positive to feel better, and equal weights of each respondent regardless of the magnitude of funds managed (Beer and Zouaoui, 2011).

Indirect sentiment measures have numerous advantages above direct sentiment measures. The most important one being that they are mostly constructed based on simple market data (Beer and Zouaoui, 2011). Baker and Wurgler (2007) and Bandopadhyaya and Jones (2011) mention some of these indirect investor sentiment measures. For example: Mutual Fund Flows, Trading Volume, Dividend Premium, IPO First Day Returns, IPO Volume, and the Put-Call Volume Ratio (PCR). The latter is supposed to be a well functioning measure according to (Houlihan and Creamer, 2017). Although, particularly relevant to this specific thesis, it is occasionally claimed that the Put-Call Ratio is a predictor of volatility. Wang, Keswani, and Taylor (2006) prove that this is not the case, making the addition of the Put-Call Ratio as an indicator of investor sentiment an interesting addition to the model. This because it is believed that features of the Put-Call Ratio can improve model performance while predicting market returns (Houlihan and Creamer, 2017). Lastly, Zouaoui, Nouyrigat, and Beer (2011) review the potential of Investor Sentiment to be used to predict market crashes. They conclude that especially in countries that show more herd-like behaviour, investor sentiment is a good predictor. This because of the increased likelihood of overreaction and low institutional involvement (Zouaoui, Nouyrigat, and Beer, 2011).

2.4.2 Risk Taking

Since it is proposed that times of financial melancholy trigger riskier behavior among investors, measuring this might lead to new insights. Giese et al. (2014) review some other factors that have an impact on the macro economical environment. They provide evidence that the Credit-to-GDP ratio could be used as a signal to identify market distress because of a change in risk taking. Additionally, this measure is relevant because it entails not only money that is borrow by industrial investors but also by private investors. They do indicate that using only the Credit-to-GDP ratio as an indicator of future signaling might have some shortcomings. Adding additional measures could improve the predictive power of the Credit-to-GDP ratio. Giese et al. (2014) propose to add, among others, the Household Debt-to-Income ratio. This adds even more insight into the (risky) behavior of private investors. Furthermore, it appears that the real house price gap performs good in predicting financial uncertainties (Barrell et al., 2010). This is in line with previous mentioned risky behavior measures since house prices normally tend to rise when investors are more willing to borrow money. Using these insights, adding a measure of risk taking to this analysis might lead to a more meaningful and significant model.

2.5 Relevance

While some classic economical models do not (or cannot) recognize the existence of economical crises, history teaches us that this is certainly not the case. Since market crashes are mostly the reason for times of economic recessions and result in companies and institutions going bankrupt and an overall decline of public wealth, it would be useful to be able to predict if such an event might be coming. This way, big financial crises might be avoided in the future.

The previous section revealed that a lot of research is done regarding past crises and their causes. Lacking is research that attempt to find a common predictor of market declines and crises which is relatively easy to reproduce. Volatility seems like a promising metric, however, previous studies present some opposite views on the measure. Some view volatility measures just as a 'fear gauge' others see the predictive potential. Especially when combined with some other factors which are in line with Minsky's view on market instability, promising results might be found. Therefore, other factors that include investor sentiment and risk taking behavior should be added to the analysis. Investor sentiment is measurable using the historical Put-Call Ratio to reveal the overall confidence of the investor in the market. The Credit-to-GDP ratio shows the relative willingness to take a more risky stance in the market. Taking up more borrowed money means having trust in the economy because you are willing to take more risk. Lastly, research that is previously conducted regarding this topic is mostly focused on the US market. This leaves the questions about other global markets and their volatility measures. The following chapter (Chapter 3) will introduce the research questions to address this possible discovered gap in existing literature.

Chapter 3

Research Design

3.1 Introduction

Chapters 1 and 2 respectively introduced the problem statement and reviewed already existing literature regarding the topic. In the problem statement it became apparent that there is a need for some predictive measures that might give an early warning for market distress or crashes. The literature revealed that there is already a substantial interest in the topic of market crashes which is not surprising depending on the impact that they have on society. The financial instability hypothesis as formulated by Minsky (1977) seems to make some valid claims that should be analyzed more intensively. Scientist have been trying to predict market crashes by using different models and not without success. In this thesis it is attempted to do similar things using volatility as the main predictor. Likewise, to stay in line with Minsky's view, two important factors that expose investor behavior before and during market distress included; investor sentiment and risk taking.

This chapter will introduce the main research question that is to be answered in this thesis. Along, sub-research questions are presented that support in doing so. In Section 3.3, the overall methodology that is used to answer the research questions in this thesis is presented. Lastly, the areas that are in (or out) of scope are defined. This way, the project, analysis and solutions stay within the bounds of a master thesis and its time frame.

3.2 Research Questions

This section respectively introduces the main research question and the according sub-research questions. The problem statement and the literature review revealed that there is a need for a predictive method to see market crashes coming. This, not only for the US but also for Europe and China.

3.2.1 Main Research Question

The main research question covers three sections: 1) getting insight into the relationship between market crashes and market volatility and 2) gain additional understanding about investor behavior in these moments in time using investor sentiment and risk taking habits in 3) three major areas in the world: the US, Europe and China. This results in the following main research question:

Can Stock Market Crashes be Explained and Predicted using Volatility, Risk Taking, and Investor Sentiment Measures in the US, Europe, and China?

3.2.2 Sub Research Questions

To create a structured approach towards answering the main research question, sub-research questions are defined. This to split the main research question in manageable parts that will function as the main structure of the thesis.

1. What data is relevant in evaluating stock market crashes?
2. How can the relevant factors be transformed to be used in the analysis?
3. What are the most appropriate models to evaluate the factor's relationships?

3.3 Methodology

While doing empirical research in the field of Finance, some principles need to be included into the research design. This, to ensure, as far as possible, the validity of the result derived (Ryan et al., 2002). As is the case in this thesis, an average research project tends to answer a research question by statistically testing a (or more) hypothesis(es). By creating a good research design, the conclusions that can be drawn using the found results are usually more valid (Ryan et al., 2002).

According to Ryan et al. (2002), empirical financial research should be conducted along the following procedure:

1. Select a research design;
2. Define the variables;
3. Formulate the hypotheses;
4. Evaluate three important criteria:
    - Does the design afford the means of testing the hypotheses formulated from the research questions?
    - Is the internal validity of the design sufficiently high?
    - Is the external validity satisfactory?

3.3.1 Research Design

First, a suitable research design should be selected in order to answer the main research question. Multiple different designs are possible. However, for this specific project, the 'interrupted time series design' is selected. Here, the behavior of a (or multiple) variable(s) is evaluated over time. Mostly, it is possible to identify a significant event that is supposed to have an impact on the variable(s). Data should be available before and after this moment of interest. A big advantage of this time series approach is that it is possible to detect and remove any confounded effects upon the dependent variable (Ryan et al., 2002). By doing so, the internal validity of the study can be greatly increased.

3.3.2 Variables

While empirical studies mostly consider relationships between variables, it is important that these variables used are defined and measured in a correct and preferably standardized way. It is important to mention any possible measurement errors and variable modification (Ryan et al., 2002). The latter of which is sometimes necessary in order to create consistent measures. All used measures and their properties for this thesis are presented in more detail in Chapter 4.

3.3.3 Hypotheses

Once the research design is selected and the variables are known that will be used to answer the main research question, it is possible to formulate the study's hypotheses. These hypotheses should be constructed in a way that they can be statistically tested and that the answer aids in answering research questions. The hypotheses that are going to be tested in this thesis are presented in Section 3.5.

3.3.4 Evaluation

The last step in the research design is to ensure three important criteria. In doing so, one makes sure that conclusions drawn from the research are actually explaining the analyzed relationship. After stating correct and relevant hypotheses, the internal and external validity are crucial.

Internal Validity

The internal validity is defined as the extend in which a causal conclusion that is drawn from is study is justified. This is mostly done by minimizing the study's error. In other words, the amount of control that is achieved in the study determines the internal validity (Ryan et al., 2002). Therefore, in designing a research project, maximizing the internal validity should be the objective. There are a number of control techniques that increase and evaluate the internal validity of the study.

External Validity

What is, and should be, one of the most important objectives of every study is the level of generalizability of the drawn conclusions. If this is the case, the results of the study can be used in multiple other settings next to the exact one analyzed. The extend in which the results of the study can be generalized in other settings or samples is known as the external validity of a study (Ryan et al., 2002). While one would probably wish to optimize both internal and external validity, it is mostly believed that by increasing one, the other decreases. Where internal validity is for most studies most important, for more applied studies this might be external validity. By including multiple world regions in this study, it is attempted to increase the external validity by increasing the population studied.

3.4 Scope

Since the research question can be broadly interpreted, scoping is of great importance to make sure that the research goal is achievable within the set time frame. The scoping will define the level of detail in the used factors and the researched geographic area's.

3.4.1 Factors

While explaining and predicting market crashes could be done by numerous factors, this thesis focuses specifically on market volatility and the according investor behavioral metrics. Therefore, the number of used factors in explaining and predicting market crashes is scoped down to four. Each of these factors and the depth in which they will be used will be elaborated in this section.

Returns

Market returns will be used in order to identify times of growing and declining economy. By using stock market indices that incorporate some of the most important and influential stocks, a good representation of the overall market tendency can be made.

Volatility

One of the most important properties of stock indices is that some have according volatility indices. These are specifically important for this thesis. Therefore, for each of the selected stock index, the according volatility index is used to explain and predict market crashes.

Market Sentiment

As mentioned in the literature review (Section 2.4.1), investor sentiment can be measured in multiple different ways. For each of the selected market indices, one measure of market sentiment is selected and added to the model. Preferably the Put-Call sentiment is used. Unfortunately, these metrics (index option volume) are not always freely available. Therefore, if this is not the case, another similar and confirmed indicator of investor sentiment is used.

Risk Taking

Lastly, the relative amount of risk that is taken by investors is added to the model. This is done by using the Credit-to-GDP ratio to include not only institutional risk taking but also private risk taking.

3.4.2 Geography

Next to the fact that it is interesting to evaluate if volatility and investor behavior are indeed factors that can be used to explain and predict market crashes, it is also relevant to evaluate if there are any differences noticeable in different parts of the world. Therefore, it is decided that the world scope of this thesis consists of three major area's: US, Europe, and China. Additionally, for the US and Europe, it will be also evaluated if there is a noticeable difference within these regions by looking at multiple stock indices.

3.4.3 Result

The scoping and factor selecting results in a set of metrics that will be used in the analysis. Table 3.1 presents an overview of the used metrics in this thesis. Chapter 4 will give a more extensive description of each of the metrics.

TABLE 3.1: Scoped Research Metrics

Research Metrics					
Region	Returns Index		Volatility Index	Market Sentiment	Risk Taking
US	S&P500	(GSPC)	VIX	P/C Ratio SPX	US
	Nasdaq 100	(USNQX)	VXN	P/C Ratio NDX	US
	Dow Jones IA	(DJI)	VXD	P/C Ratio DJX	US
EU	Euro Stoxx 50	(STOXX50E)	VSTOXX	Sentix EU SMT	EU[1]
	DAX	(GDAXI)	VDAX	P/C Ratio DAX	GER
China	iShares China	(FXI)	VXFXI	SMT Index China	CN

Table 3.1 reveals that for the US, three different stock market indices will be reviewed. The according volatility indices and their P/C Ratios will be used. For Europe, two different market indices are selected. First, the Euro Stoxx 50 contains the 50 most important stocks in the euro zone. Additionally, the DAX is selected which covers the 30 biggest listed companies in Germany. Lastly, iShares China includes the 50 largest stocks in China. While most of these stock indices consist of companies in a single country, this is not the case for the Euro Stoxx 50. Therefore, to evaluate risk taking behavior the credit-to-GDP ratio for multiple countries is combined. [1] For Europe, risk taking is calculated using the Credit-to-GDP Ratio coming from Belgium, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, and The Netherlands.

3.5 Hypotheses

In this section the hypotheses are introduced that are going to be tested in the analysis. Each of the hypotheses aims to broaden the understanding of the relationship between Market Volatility, Investor Sentiment, Risk Taking and eventually Market Returns (Crashes). Figure 3.1 schematically presents the researched relationships. All proposed hypotheses examine the direct and lagged relationships between factors.

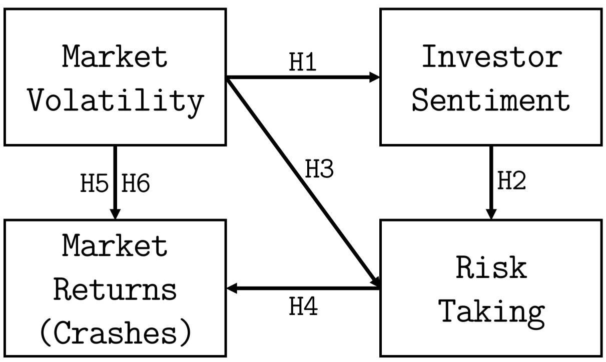


FIGURE 3.1: Tested Hypotheses

The first hypothesis tends to find a relationship between market volatility and market sentiment. Since market sentiment indicates if the investor is positive (or negative) in the future of the market and its returns, there might be relationship between the level of volatility in this period and the investor sentiment in the same or coming period(s). It is expected that when volatility is high (low), the investor sentiment will be low (high). Therefore, a negative relationship is expected.

Hypothesis 1 (H1) *Investor Sentiment has a negative relationship with Market Volatility.*

The second hypothesis focuses on the effect of investor sentiment on risk taking by the investor. It is expected that when the investor sentiment is high (low), there is a bigger (smaller) willingness to take risk by borrowing more (less). Therefore, a positive relationship between investor sentiment and risk taking is expected.

Hypothesis 2 (H2) *Investor Sentiment has a positive relationship with Risk Taking.*

The third proposed hypothesis focuses on the relationship between volatility and risk taking. It is believed (in line with Minsky's concept) that in periods of low volatility, an investor is more likely to take more risk by borrowing more. Therefore, a negative relationship is expected between Risk Taking and Market Volatility.

Hypothesis 3 (H3) *Risk Taking has a negative relationship with Market Volatility.*

The fourth hypothesis focuses on the phenomenon that borrowing more increases the amount and frequency of negative returns. This comes as a result of the introduction to more risk and interconnectedness in the market. Therefore, it is expected that Risk Taking has a negative relationship with Market Returns.

Hypothesis 4 (H4) *Market Returns have a negative relationship with Risk Taking.*

The fifth hypothesis contains the overall idea that volatility has a connection to market returns. While it is expected that during a market crash, volatility will increase and be on average high, special interest questions if the period leading to market declines is typically showing low volatility. Therefore, even though this hypothesis states that the relationship between Market Returns and Volatility is negative, it is evaluated if this lagged relationship shows adverse behavior.

Hypothesis 5 (H5) *Market Volatility has a negative relationship with Market Returns.*

To gain even a better understanding, it is tested if excessively low (high) volatility in the periods approaching a market crash are common. Therefore, Hypothesis 6 evaluates if low (high) volatility holds explanatory and predictive power on market crashes. There is a particular interest in the lagged relationship between the two variables.

Hypothesis 6 (H6) *Low (High) Volatility has a positive relationship with Market Crashes.*

Chapter 4

Data

4.1 Introduction

This chapter will focus on the exact data and the construction of the variables that are used in the empirical analysis in Chapter 5. For each of the variables, it will be elaborated how they are determined, retrieved, calculated, and which period is selected. Additionally, some descriptive statistics are presented to improve understanding of the used metrics.

While evaluating data availability, it became apparent that some of the variables are not available (or measured) as often as desired. Some of the variables are only available on a quarterly or even yearly basis. Interpolation of these series will overcome the shortcoming of some data points, however, future data is often required for such a transformation. Since this thesis not only tends to create insight in hindsight but also aims to find certain predictive qualities in used variables, it is decided to run the analysis using both monthly and quarterly intervals. Since it is believed that a quarterly analysis is too broad while evaluating crashes (since some market crashes only last a couple of days), there is a certain amount of value in analyzing monthly data. Additionally, monthly data is preferred over daily (or weekly) data since small changes in one of the factors will be ignored whereas major changes in, for instance the returns, will be noticeable in the total months return. This way, only major events will be subject to the analysis. This also sets up the analysis nicely to add an extra indicator for 'crash' months. Likewise, while evaluating stock market returns, one wishes to be able to assess possible market returns a substantial time before a market crash happens. Therefore, the benchmark of at least a month before the actual crash should be a good indicator. Note that since this monthly data uses (some) interpolated data points and therefore future data, this approach is not used as a predictive study but more as a explanatory study. Since (for the US) all used data is available on a quarterly basis, a predictive approach is justified. The presented factors in the following sections are similar for both the monthly and quarterly analysis. The latter of which merely avoids data interpolation and any other forward looking techniques.

While working and combining multiple different measures, the variable which has the smallest data range determines the data range of the complete set. Table 4.1 gives an overview of the return indices and their according variables. For each of the variables the year of first (available) measurement is presented. All of the measures have (at least) data available until 01-01-2018.

TABLE 4.1: Research Metrics First Available Year

Research Metrics Date Range					
Region	Returns Index		Volatility Index	Investor Sentiment	Risk Taking
US	S&P500	1950	1993	**1996**	1952
	Nasdaq 100	2000	**2001**	1996	1952
	Dow Jones IA	1985	**1997**	**1997**	1952
EU	Euro Stoxx 50	1986	1999	**2002**	1970
	DAX	1987	2005	**2009**	1960
China	iShares China	2004	**2011**	1997	1985

Table 4.1 reveals that for some of the stock indices not the biggest date range is available. This is mainly caused by the measure of investor sentiment. It could be concluded that in some cases this measure is left out of the analysis if turns out that the analyzed period is too small (e.g. there are no market crashes in the time frame). Furthermore, the volatility index for iShares China is only available since 2011. This relatively small period could suffer from the same issues as mentioned before.

Note that the following sections present the formation of the monthly variables used in the analysis. A similar approach is used to create/transform the quarterly data points unless stated differently. Subscript q is used to indicate that the variable uses the quarterly interval.

4.2 Market Returns

To evaluate the state of the economy and identify market crashes, price indices without dividend are used. These series are less trending compared to total returns indices and should be adequate to indicate periods of market declines. The exact indices are mentioned in Table 3.1. The data used is retrieved from the financial database of Wharton University; WRDS (Wharton University of Pennsylvania, 2018) for the relevant time period. Daily stock market quotes containing the: opening, closing, high, and the low index price. The daily closing price is used to calculate the index return. Because of several advantages (like time-additivity), log returns constructed. Hence, it is straightforward to calculate the monthly log returns using daily stock index data. Using the closing price of the day, daily returns are calculated which are then summed to monthly returns (see Equation 4.1 and 4.2).

$$R_t = ln\frac{P_t}{P_{t-1}} \tag{4.1}$$

$$R_m = \sum_{t=1}^{t=D_m} R_t \tag{4.2}$$

where:

$$\begin{aligned} t &= \text{Time in days} \\ P_t &= \text{Closing Price at time t} \\ P_{t-1} &= \text{Closing Price at time t-1} \\ R_t &= \text{Log Return at time t} \\ D_m &= \text{Last day of month m} \\ R_m &= \text{Log Return in month m} \end{aligned}$$

Appendix A reveals the monthly log returns for each of the indices. Using these monthly log returns, it can be determined when there were times of big market declines or even market crashes. By setting a benchmark for the magnitude of the market decline, a dummy variable is generated that indicates a market crash (a big market decline). Comparing these results with the time stamps of historical market crashes (like the global financial crisis of 2008, Figure 4.1), it can be evaluated if this method functions as desired.

TABLE 4.2: Crash Dummy Determination

Research Metrics Date Range							
Lowest	2.5%	5%	7.5%	10%	12.5%	15%	
Rm <	-8.92%	-6.63%	-5.62%	-4.62%	-3.51%	-3.09%	-5.00%
Amount of Crashes	9	12	19	35	44	53	31
All Crashes included?	No	No	No	Yes	Yes	Yes	Yes

Table 4.2 present the steps in determining the correct threshold to create the crash dummy. The lowest $X\%$ is gradually increased until all 'theoretical' crash months are found by the returns threshold. It appears that the lowest 10% of the observations includes all crashes. When inspecting

the performance of this method it became clear that using a returns benchmark of -4.62% was a little too broad. The dot-com bubble, the Asian Crisis, and the Flash Crisis of 2010 all showed market declines just above 5% (e.g. Dot-Com Bubble; -5.22%). Therefore, in order to only cover the most relevant crash months, it is decided to set the crash benchmark to 5%.

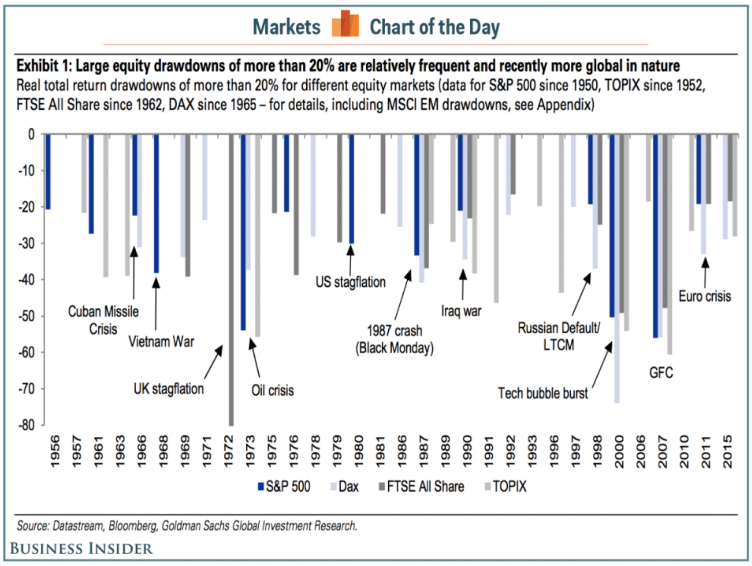


FIGURE 4.1: Market Crashes from 1956 to 2015 (Business Insider)

Equation 4.3 formulates the benchmark for a market crash. Apparently, a benchmark of a monthly average market decline of at least 5% includes all major stock market crashes of the past decade. Please acknowledge the fact that this might include days with market declines several times larger.

$$CRASH_m = \begin{cases} 1 & \text{if } R_m \leq -5\% \\ 0 & \text{else} \end{cases} \tag{4.3}$$

where:

$$\begin{aligned} CRASH_m &= \text{Dummy indicator if month m is a crash month} \\ R_m &= \text{Log return in month m} \end{aligned}$$

Figure 4.2 depicts the S&P 500 index and the identification of the months that faced a sharp index decline. In total (for the S&P series), 31 months are identified as a crash month in the period ranging from January 1990 unit December 2018.

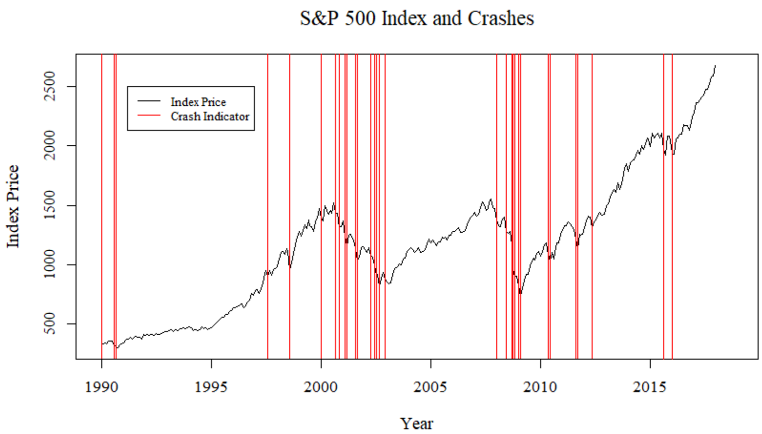


FIGURE 4.2: S&P Index and the Crash Dummy

4.3 Volatility

For each of the mentioned stock indices in Table 4.1 a corresponding volatility index is retrieved. As shown in the table (Table 3.1), implied volatility is used in this analysis. Following Minsky, these implied volatility indices carry information about the investor towards current market risk. Since it is supposed that times of euphoria create unstable markets, implied volatility forms a relevant factor. Most volatility indices are constructed by the Chicago Board Option Exchange (CBOE). Daily index quotes are obtained from their website (CBOE, 2018). The remaining volatility indices are attained using the financial database of Wharton University (Wharton University of Pennsylvania, 2018).

While there is specific interest in divergent behavior of the volatility index (i.e. abnormal high or low), the monthly average is determined. This way, it can be evaluated if months with a relatively low observed volatility are followed by months of financial crises.

$$\mu_m^V = \frac{\sum_{t=1}^{t=D_m} V_t}{D_m} \tag{4.4}$$

where:

$$\begin{aligned} t &= \text{Time in days} \\ V_t &= \text{Volatility index at time t} \\ D_m &= \text{Last day of month m} \\ \mu_m^V &= \text{Average volatility index in month m} \end{aligned}$$

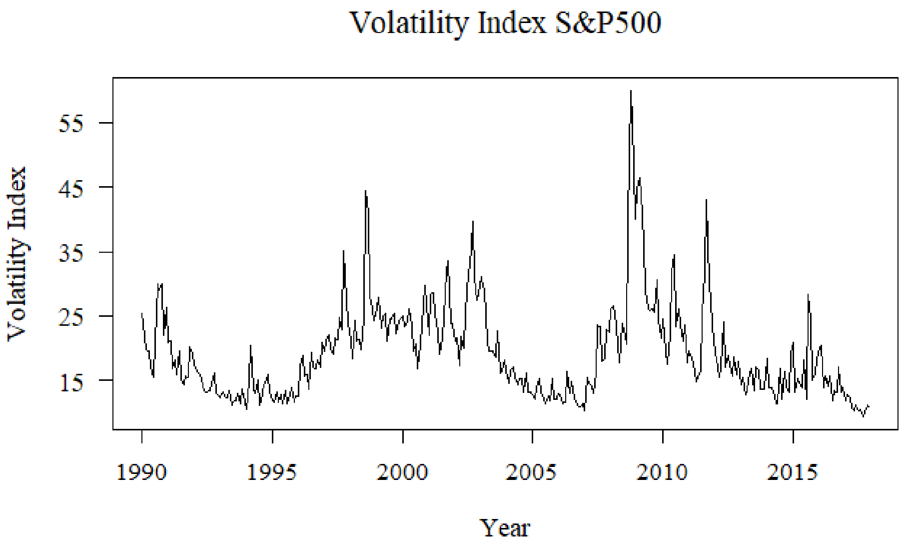


Figure 4.3: Monthly S&P Volatility

Figure 4.3 (and Appendix B.1) presents a graphical representation of the monthly volatility series. It appears that the volatility series shows some trending/seasonality behavior. To remove this seasonal trend and to be able to identify periods of abnormally high and low volatility, the Hodrick-Prescott (HP) Filter (Hodrick and Prescott, 1997) is applied to the volatility series (See Appendix F). The HP Filter recognizes the trend and divides the time series into two parts: the trend (g_t) and the cyclical (c_t) component (Equation 4.5).

$$y_t = g_t + c_t \quad \text{for } t = 1, ..., T \tag{4.5}$$

By removing the trend component, a new series arises in which only the deviations from the trend remain. This way, the times of high and low volatility are more accurately identified. Ideally, one would prefer the use the two-sided HP (HP2) filter since this filter can more accurately identify the trend. However, the two-sided HP filter uses 'future' data to determine the trend component. This implies that if one desires to use the series for forecasting and prediction, this transformation is not possible. Since the monthly series is only used explanatory and not predictive, the two-sided HP filter is used here. For the quarterly series, that is used to predict, the one-sided HP (HP1) filter is used. Even though the HP Filter seems like a appropriate approach, some academia question the accuracy of the filtering method. Hamilton (2017) wrote an article claiming that the HP filter should not be used. He raises the concern that the filter can introduces spurious dynamic relations that are purely an artifact of the filter and have no basis in the true data-generating process (Hamilton, 2017). However, this effect is mostly seen while using the HP filter on economic time series of first differences. In the case of volatility, the index level is used and therefore the HP filter is supposed to be a correct transformation.

Figure 4.4 presents the application of the HP2 filter on the S&P volatility data. As mentioned in Appendix F, when using the HP filter an appropriate frequency λ should be selected. According to Ravn and Uhlig (2002), monthly time series should use a frequency of $\lambda = 129,600$ and quarterly $\lambda = 1,600$. Using these filter frequencies, the monthly (quarterly) trend was identified and removed to construct a new time series that includes only the cyclical component. This is done for each of the volatility indices.

$$V_m = \mu_m^V - g_m^V \tag{4.6}$$

where:

$$\begin{aligned} \mu_m^V &= \text{Average volatility index in month m} \\ g_m^V &= \text{Volatility trend component in month m} \\ V_m &= \text{HP-Filtered Volatility in month m} \end{aligned}$$

According to Minsky, low volatility is often occurring in the period prior to a market crash. Therefore, following the reasoning of Danielsson, Valenzuela, and Zer (2018) and van Veen (2017), two additional variables are constructed. This, to create an even broader understanding of the impact of abnormally high or and especially low volatility. The months that the average volatility is greater than the trend component are extracted into a volatility high variable. The months that the average volatility is lower than the trend component are extracted into a volatility low variable. Equations 4.7 and 4.8 present these two variables.

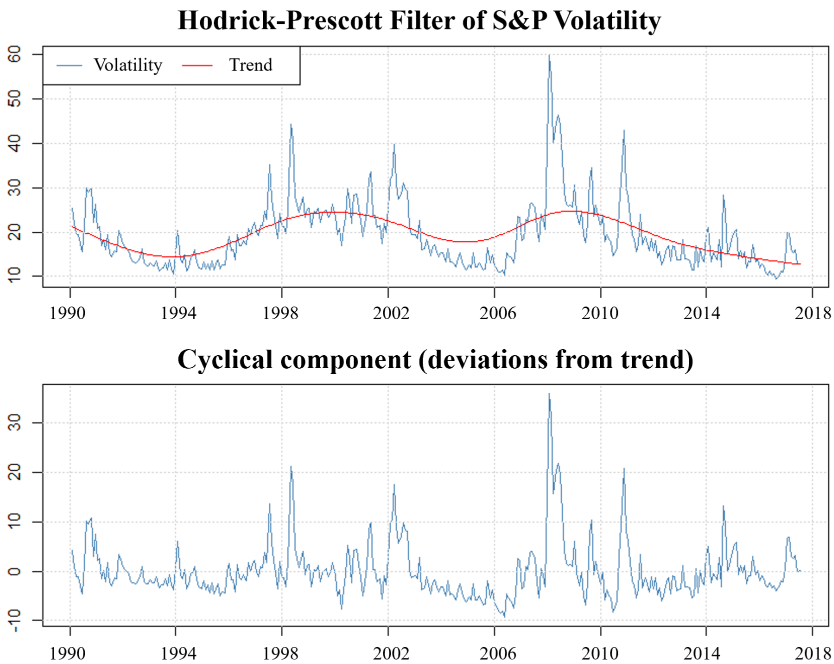
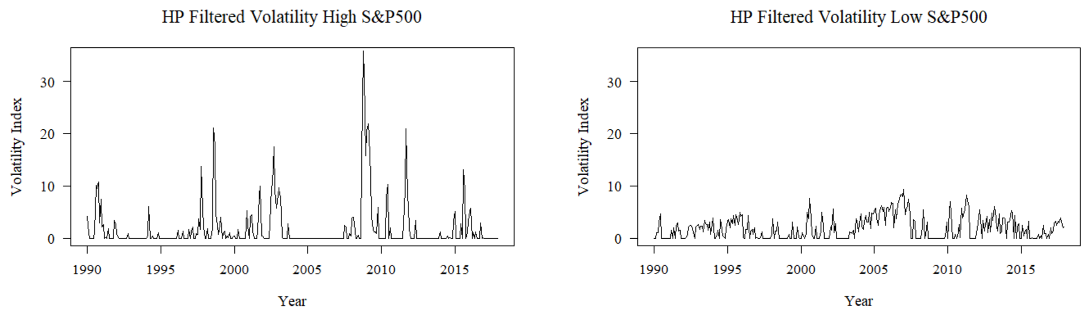


FIGURE 4.4: Hodrick-Prescott Filter of VIX Index

$$V_m^{high} = \begin{cases} V_m & \text{if } V_m > 0 \\ 0 & \text{else} \end{cases} \tag{4.7}$$

$$V_m^{low} = \begin{cases} |V_m| & \text{if } V_m \leq 0 \\ 0 & \text{else} \end{cases} \tag{4.8}$$

Appendix B.3 graphically reveals the newly created variables for each of the indices. Figure 4.5 presents these variables for the S&P volatility. What can be noticed is that the high volatility variables includes less observations while these observations overall have a bigger magnitude and vice versa.



(A) High (B) Low

FIGURE 4.5: The High and Low Volatility Variables S&P

4.4 Risk Taking

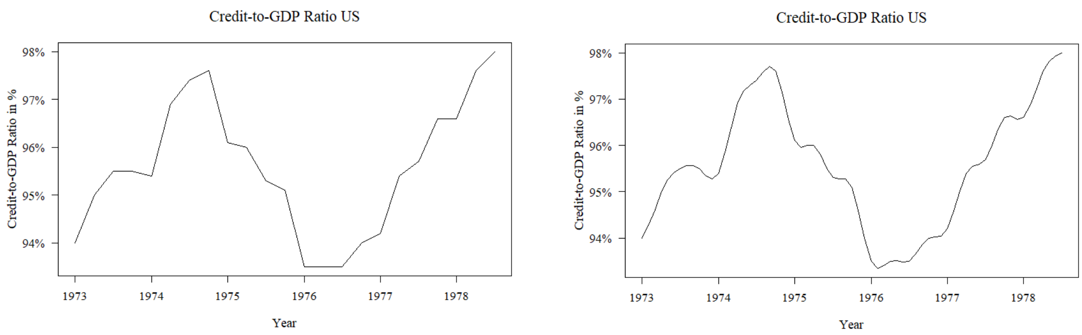
The next measure that is added to the analysis is the factor for the amount of risk taking that is done by the investors. As argued in Section 3.4.1, the Credit-to-GDP ratio (CTG Ratio) is used to evaluate this. This data is made available by the Bank for International Settlements (Bank for International Settlements, 2018). They provide a data set consisting quarterly credit-to-GDP ratio's for numerous countries going back as early as 1951. This measure essentially tracks the ratio of all debt and loans of both private and institutional borrowers of a country divided by the Gross Domestic Product of this country in a specific moment in time (see Equation 4.9).

$$CTG_q = \frac{Credit_q}{GDP_q} * 100 \tag{4.9}$$

where:

$$\begin{aligned} Credit_q &= \text{Total debt and loans in quarter q} \\ GDP_q &= \text{Gross Domestic Product in quarter q} \\ CTG_q &= \text{The Credit-to-GDP Ratio in quarter q} \end{aligned}$$

Since the data that is provided by the Bank for International Settlements in featuring quarterly time stamps only, these observations are transformed to monthly observations by using 'Cubic Spline Interpolation' (Torres-Reyna, 2014). This method is forward looking implying that future data is used to construct the missing data points. Again, since the monthly analysis is used to reveal relationships in hindsight, this is satisfactory. The quarterly series does not suffer from this problem since the data is available in these increments.



(A) Quarterly (B) Monthly

FIGURE 4.6: Credit-to-GDP Ratio in the US

Figure 4.6 reveals the effect of cubic spline interpolation. The line appears to be smoother since there are more data points per year (12 instead of 4). This approach is applied to the time series of each of the relevant countries. An important additional effect is introduced when using data points gained from interpolation; serial correlation. While the extra data points are created using existing ones, they are by definition dependent on each other. If this time series is used in OLS regression the assumption of no auto correlation is violated and therefore, the model is not accurate. Luckily, additional data transformations or the use of Newey-West standard errors (Newey and West, 1987) overcomes this concern.

As mentioned in Table 3.1, the stock indices mostly consist of companies based and listed in one specific country. For the Euro Stoxx 50, this is not the case. A new CTG Ratio measure is developed using the time series of the countries that have companies listed in this index. Hence, a new CTG Ratio is constructed by using the ratios of nine countries. This is done respectively by the magnitude of the according GDP. The historical GDP of these countries is provided by the world bank (The World Bank, 2018a). They preserve historical GDP data for numerous countries. Appendix C.1 Figure C.3a presents the newly created measure for European Credit-to-GDP Ratio that is used for the analysis of the Euro Stoxx 50 index.

The CTG series all show non-stationary behavior. This is confirmed when a Dickey-Fuller test (Dickey and Fuller, 1979) is performed on the data. A unit root is present and therefore, a variable transformation is needed before it can be used in one of the models. In an attempt to create a stationary variable, the first difference of the CTG Ratio is taken. Running the same Dickey-Fuller test on the newly created variable confirms that the series is now stationary. The property of serial correlation is, however, still present. Therefore, if this variable is used in the analysis, Newey-West standard errors are used to evaluate statistical significance.

$$RISK_m = CTG_m - CTG_{m-1} \tag{4.10}$$

Equation 4.10 presents the formula to find the first difference of the CTG Ratio. This creates the $RISK_m$ variable that is used in the analysis to evaluate risk taking behavior. Appendix C.1 presents a graphical representation of the risk taking variable for each series.

4.5 Investor Sentiment

The next factor that is added to the analysis is the measure for investor sentiment. Preferably, the Put-Call Ratio (PCR) is used. This measure provides the ratio of the volume of put options divided by the volume of call options on the according stock index. As mentioned in Section 2.4.1, this measures the level of bearish or bullish investor sentiment. Unfortunately, this data is not completely freely available. The WRDS (Wharton University of Pennsylvania, 2018) database provides the data solely for stocks that are based in the US. Therefore, the PCR for the S&P500, Nasdaq 100, and the Dow Jones Industrial Average are acquired. A website that covers news about the German stock exchange held the data for put call sentiment on the DAX index (Boersen News, 2018). PCR data is mostly provided on a daily basis, therefore, this needs to be transformed into monthly data points. Equation 4.11 presents this transformation. Appendix D contains figures of the monthly PCR for the three US stock indices and the DAX.

$$PCR_m = \frac{\sum_{t=1}^{t=D_m} O_t^P}{\sum_{t=1}^{t=D_m} O_t^C} \tag{4.11}$$

where:

$$\begin{aligned} O_t^{Put} &= \text{Put option volume on day t} \\ I_t^{Call} &= \text{Call option volume on day t} \\ D_m &= \text{Last day of month m} \\ PCR_m &= \text{Put-Call Ratio in month m} \end{aligned}$$

Unfortunately, the PCR sentiment for the Euro Stoxx 50 and the iShares China is not freely available. Therefore, for both indices, an alternative is used. It is considered that the Sentix Sentiment index is one of the most reliable sentiment measures of the European investor. This sentiment index is constructed using a monthly interview of 1600 financial analysts and institutional investors (Sentix, 2018). A relatively higher Sentix index implies a positive, or bullish, investor attitude towards the market. A lower Sentix index entails a more bearish stance of the European investor. Fortunately, this measure is already shaped in a monthly fashion. Appendix D.1.2 Figure D.5 reveals the Sentix index for the relevant period.

Then, a sentiment index for the Chinese market is desired. Because the put call volume for the specific index is not freely available an alternative is selected. Baker, Wurgler, and Yuan (2012) introduces a framework to construct a sentiment index using: dividend premium, the number of IPO's and their returns, and market turnover. Zhao, Yang, and Qian (2015) apply this method for China and prove that there is a possibility to build a similar sentiment index for the Chinese investor. Cheema, Man, and Szulczyk (2018) create an actual sentiment index for China and publish their results. This monthly sentiment index is used in this analysis. Appendix D.1.3 reveals this sentiment index over the available time period.

Similar to the volatility series, the sentiment series shows some sort of seasonal/trending behavior. To create a comparable measure, the HP filter (Appendix F) is applied to the sentiment series. Again, for the monthly series, the more desirable two-sided HP filter is used. For the quarterly series which should respect its predictive quality, the one-sided HP filter is used. Similar to the volatility factor, the monthly series uses an HP frequency of $\lambda = 129,600$ and for the quarterly series a frequency of $\lambda = 1,600$ (Ravn and Uhlig, 2002). By applying this filter to each of the sentiment series and the inverting of the Put-Call Ratio series, a new measure is created that will be positive when the average investor is bullish and will be negative when the average investor is feeling bearish. This inverting of the Put-Call Ratio series is necessary since a low Ratio indicates Bullish investor behavior and vice versa. By inverting the HP filtered series, these sentiment measures indicate bullish and bearish investor sentiment in a similar manner as the indices for Europe and China.

$$SENT_m = SMT_m - g_m^{SENT} \tag{4.12}$$

where:

$$\begin{aligned} SMT_m &= \text{Respective Sentiment Index in month m} \\ g_m^{SENT} &= \text{Sentiment trend component in month m} \\ SENT_m &= \text{HP-Filtered Sentiment Index in month m} \end{aligned}$$

Figure 4.7 presents the final sentiment variable for the S&P 500 analysis. Appendix D depicts each of the Sentiment series that is used in the analysis in the next chapter.

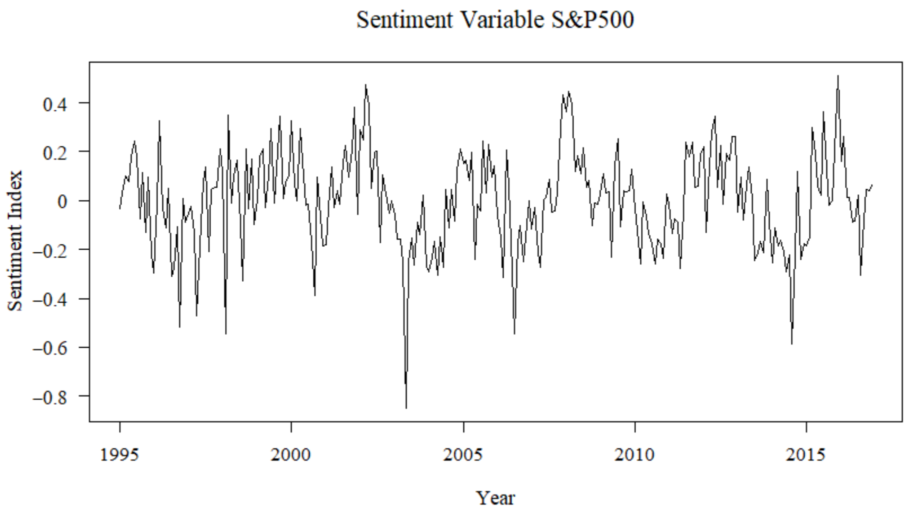


FIGURE 4.7: Sentiment Variable S&P

4.6 Control Variables

Additional to the previously presented factors for market returns, volatility, risk taking and investor sentiment, three control variables are introduced. These control variables are added to a complete model to verify that the found results are robust and not just because of macro market movements. In line with Danielsson, Valenzuela, and Zer (2018), three location specific control variables are added: GDP per capita, Inflation and the Change in Government Debt. The following sections present the exact determination of the control variables. A graphical representation of the final variables can be found in Appendix E.

4.6.1 GDP per Capita

For the US based indices, data retrieved from the Federal Reserve Economic Data (FRED) database is used (Federal Reserve Bank of st.Louis, 2018). They provide a quarterly series that will be cubic interpolated for a monthly analysis. The data for other geographic locations is available at the website of the World Bank (The World Bank, 2018b). They provide a data set covering yearly data points for several popular financial indicators for all countries in this analysis. In context of the monthly analysis, these found series are cubic interpolated. GDP per capita is typically a trending series. Therefore, the fist difference is used in this analysis. Equation 4.13 presents the formula to determine the final control variable. Here again (as for the other control variables), the measure for the Euro Stoxx index is created using all countries covered by the index relative to its yearly GDP.

$$GPC_m = GDPPC_m - GDPPC_{m-1} \tag{4.13}$$

where:

$$\begin{aligned} GDPPC_m &= \text{GDP per capita in month m} \\ GPC_m &= \text{Change in GDP per capita in month m} \end{aligned}$$

4.6.2 Inflation

Next, inflation is added the the analysis as a control variable. The Consumer Price Index (CPI) is used to do so. The CPI is defined as the change in price of goods and services that are typically purchased by specific groups of household (OECD, 2018). Again, for the US this data is available on a quarterly basis (Federal Reserve Bank of st.Louis, 2018) whereas for the other regions, yearly series is obtained (The World Bank, 2018b). Since this variable is already measuring a change in price, no transformations (except for the cubic interpolations for the monthly series) are needed. Equation 4.14 presents the final control variable.

$$INF_m = CPI_m \tag{4.14}$$

where:

$$\begin{aligned} CPI_m &= \text{Consumer Price Index in month m} \\ INF_m &= \text{Inflation control variable in month m} \end{aligned}$$

4.6.3 Government Debt

Lastly, the change in government debt is added as a control variable. Is it believed that the amount of government debt could have an impact on the probability of an economic crisis (Danielsson, Valenzuela, and Zer, 2018). Respectively, a quarterly (Federal Reserve Bank of st.Louis, 2018) and a yearly (The World Bank, 2018b) series are obtained for the US and the other regions. After cubic interpolation the first difference is used as final measure of government debt. Equation 4.15 presents the final control variable for government debt.

$$DEBT_m = DEBT_m^{gov} - DEBT_{m-1}^{gov} \tag{4.15}$$

where:

$$\begin{aligned} DEBT_m^{gov} &= \text{Government Debt in month m} \\ DEBT_m &= \text{Change in Government Debt in month m} \end{aligned}$$

Chapter 5

Analysis

5.1 Introduction

This chapter will introduce the models that are used to evaluate the stated hypotheses in Section 3.5. The data and variables that are presented in the previous chapter are used to do so. Most of the models will use OLS regression. As argued before, since the factor for risk taking ($RISK_m$) suffers from serial correlation, the models that evaluate this factor will use Newey-West standard errors to test statistical significance of the independent variables. Since the last two models use the dependent variable that signals if a month is defined as a crash month ($CRASH_m$), logistic regression is applied. This technique is often used to model a binary dependent variable, which is the case in this context. Figure 5.1 provides a overview of the thesis hypotheses and the according models. The next section introduces the models in more detail. The last section of this chapter (Section 5.3) reveals the results for each of the models.

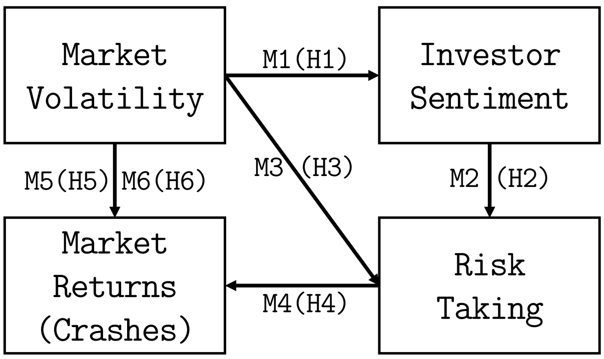


FIGURE 5.1: Tested Hypotheses

5.2 Models

This section will introduce the regression models that are used to test the hypotheses. Since there is particular interest into the explanatory and predictive power of the mentioned variables in the moments before a market crash, the models could contain numerous combinations of independent variables due to the number of lags included. Therefore, a notation with summations is used. While evaluating each of the models, the final model is selected according to which variation yields the most relevant results. Note that the presented models are using the subscript m indicating the use of the monthly variables. If m is substituted with q one finds the exact models used in the quarterly analysis.

5.2.1 Model 1: Market Volatility on Investor Sentiment

$$SENT_m = c + \sum_{l=0}^{L} \beta_l V_{m-l}^{low} + \sum_{k=0}^{K} \delta_k V_{m-k}^{high} + \epsilon_m \tag{5.1}$$

The first model evaluates the relationship of volatility with market sentiment. It is believed that low volatility will trigger positive investor sentiment since there is supposed to be less risk in the market. On the contrary, when the volatility in the market is high, investor sentiment is likely to be lower. However, when high market volatility is lagged ($l = 1, 2, 3$ and $k = 1, 2, 3$), there might be a positive relationship between the variables since an investor could believe that there will be prosperous times after times of high volatility. Running an OLS regression on Model 1 (Equation 5.1) will test Hypothesis 1.

5.2.2 Model 2: Investor Sentiment on Risk Taking

$$RISK_m = c + \sum_{l=0}^{L} \beta_l SENT_{m-l} + \epsilon_m \tag{5.2}$$

Model 2 investigates if the amount of risk that is taken by the investor has a relationship with the overall level of investor sentiment. It is believed that in times of bullish investor sentiment, an investor is more willing to take up additional leverage in order to have a bigger opportunity to make profit. This model will use Newey-West standard errors to evaluate coefficient significance due to the present serial correlation in the $RISK_m$ variable.

5.2.3 Model 3: Market Volatility on Risk Taking

$$RISK_m = c + \sum_{l=0}^{L} \beta_l V_{m-l}^{low} + \sum_{k=0}^{K} \delta_k V_{m-k}^{high} + \epsilon_t \tag{5.3}$$

Model 3 evaluates if times of low market volatility trigger an increase in risk taken by issuing additional debt. This hypothesis is in line with Minsky's believe that the increase in leverage in low volatility periods trigger a greater chance of market crashes. The combination of results of models 1 to 3 will create greater understanding if risk taking is indeed increased in these periods of investor melancholy and low volatility.

5.2.4 Model 4: Risk Taking on Market Returns

$$R_m = c + \sum_{l=0}^{L} \beta_l RISK_{m-l} + \epsilon_m \tag{5.4}$$

The previous models focused on the relation between volatility, risk taking and market sentiment. If there are significant relationships found, the connection to market returns would be of great interest. Minsky states that it is not low volatility that causes market crashes but the way the investor deals with the information of low volatility. Creating a more unstable economy when issuing more debt contributing to bigger market losses in periods of relapse resulting in a market crash. Model 4 evaluates if there is a relationship between historic debt issuance and market returns. It is expected that when there is more debt issued in the previous period(s), there is a bigger chance on negative market returns.

5.2.5 Model 5: Market Volatility on Market Returns

$$R_m = c + \sum_{l=0}^{L} \beta_l V_{m-l} + \epsilon_m \tag{5.5}$$

Since volatility might trigger more than only additional debt issuance, it is interesting to evaluate what the effect of the overall level of volatility is on market returns. It is expected that high levels of volatility in the current period should have a negative relationship with market returns, however, the relationship of the two variables in the period approaching market declines is uncertain.

5.2.6 Model 6: Low or High Volatility on Crashes

$$CRASH_m = c + \sum_{l=0}^{L} \beta_l V_{m-l}^{low} + \sum_{k=0}^{K} \delta_k V_{m-k}^{high} + \epsilon_m \tag{5.6}$$

Model 6 investigates if low (or high) volatility during and in the periods prior to a market crash have explanatory and predictive power. This model uses logistic regression since the variable $CRASH_m$ is binary (1 for crash month, 0 for non crash month). This way, the model evaluates the chance of experiencing a market crash in the coming month(s) when the volatility level is abnormally low (or high). Since this model uses logistic regression, there is no set measure of goodness-of-fit. In this case, the pseudo R^2, McFadden's R^2 (McFadden, 1973) is used to do so. If low volatility seems to be a significant predictor of future market crashes, this analysis would be partly supporting Minsky's Financial Instability Hypothesis.

5.2.7 Model 7: Complete Model

The 7$^{\text{th}}$ and last model includes all previous mentioned variables and includes the control variables as well as a lagged $CRASH$ variable. By doing so, statistical significant results for the low and high volatility channel variables are tested on robustness. It is expected that the control variables have a certain amount of statistical significance with the crash dummy variable.

$$\begin{aligned} CRASH_m = c + \sum_{l=0}^{L} \beta_l V_{m-l}^{low} + \sum_{k=0}^{K} \delta_k V_{m-k}^{high} + \sum_{j=0}^{J} \gamma_j SENT_{m-j} + \sum_{i=0}^{I} \chi_i RISK_{m-i} \\ + \sum_{p=0}^{P} \phi_p INF_{m-p} + \sum_{q=0}^{Q} \omega_q GPC_{m-q} + \sum_{w=0}^{W} \psi_w DEBT_{m-w} + \sum_{z=1}^{Z} \kappa_z CRASH_{m-z} + \epsilon_m \end{aligned} \tag{5.7}$$

5.3 Results

This section presents the results of the analysis of the introduced models. To increase the understanding of the researched relationships, the presented models are used to evaluate the six different data samples coming from different stock indices and geographical area's. Additionally, the analysis is replicated using two different time intervals; monthly and quarterly. The monthly analysis is used as a descriptive and exploratory study. The quarterly study tends to find predictive properties in the researched relationships and therefore only focuses on lagged variables. This section will discuss the monthly results first.

5.3.1 Monthly Analysis

This section will present the results of the monthly analysis. For each of the models the significant results (if any) are presented in two tables. One presenting the results of the US indices, one for the European and Chinese indices.

Model 1 Results

Model 1 evaluated the hypothesis that investor sentiment has a negative relationship with market volatility. Implying that in times of low volatility, investor sentiment is likely to be more positive and vice versa.

TABLE 5.1: Model 1 Monthly Results: US

	Dependent variable:		
	$SENT_m$		
	(S&P)	(DJIA)	(Nasdaq)
V_{m-0}^{low}	0.015**	0.015	−0.001
	(0.007)	(0.022)	(0.015)
V_{m-0}^{high}	−0.008**	0.002	0.005
	(0.004)	(0.011)	(0.006)
V_{m-1}^{low}	−0.009	0.022	−0.018
	(0.007)	(0.027)	(0.013)
V_{m-1}^{high}	0.017***	−0.030*	−0.002
	(0.004)	(0.016)	(0.006)
V_{m-2}^{low}		−0.003	
		(0.0022)	
V_{m-2}^{high}		0.025**	
		(0.011)	
Constant	−0.031	−0.068	0.042
	(0.020)	(0.054)	(0.040)
Observations	263	237	202
R^2	0.114	0.059	0.032
Adjusted R^2	0.100	0.034	0.013
Residual Std. Error	0.190	0.449	0.308
F Statistic	8.290***	2.405**	1.647
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 5.1 presents the results for the three US stock indices. Both the S&P and DJIA indices show some statistical significance. For the S&P index, the results in the current month act as expected; if volatility is low (high), the investor sentiment is high (low). The lagged relationship is however mixed. The S&P shows that if the previous month shows high volatility, investor sentiment is likely to go up in this month. The DJIA shows a contradicting, but less significant sign. When this effect is lagged for an additional month, the DJIA confirms the effect that high volatility in the months prior might increase investor sentiment. No (lagged) statistical significant relationships

are found in the Nasdaq index. According to the F-statistic, this model is overall not significant. Additionally, the adjusted R^2 for the remaining two models is relatively low implying that the fit is not optimal.

TABLE 5.2: Model 1 Monthly Results: Europe & China

	Dependent variable:		
	$SENT_m$		
	(DAX)	(EURO)	(FXI)
V^{low}_{m-0}	0.004**	0.677	−0.155**
	(0.002)	(0.520)	(0.071)
V^{high}_{m-0}	−0.003**	−0.730***	−0.030
	(0.001)	(0.249)	(0.043)
V^{low}_{m-1}	−0.003*	0.234	0.134*
	(0.002)	(0.514)	(0.070)
V^{high}_{m-1}	0.001	−0.615**	−0.060
	(0.001)	(0.261)	(0.043)
Constant	−0.000	1.067	0.256
	(0.005)	(1.703)	(0.206)
Observations	89	168	61
R^2	0.193	0.371	0.188
Adjusted R^2	0.155	0.355	0.130
Residual Std. Error	0.024	12.000	0.885
F Statistic	5.034***	23.999***	3.232**
Note:			*p<0.1; **p<0.05; ***p<0.01

The adjusted R^2 for the models that used the European and Chinese data is structurally higher. Since apparently the fit for the model using Eurostoxx data is best, these results are most relevant. This model shows statistical significance for both high volatility streams. According to the model, high volatility in the current and the previous month have a negative effect on investor sentiment. This result is conform with the hypothesis. However, no statistical significant effect is found for the low volatility variable. The models evaluating the DAX and FXI index do find significant relationships with the low volatility variable. The results do however wear contradicting signs and therefore, no strong support for this effect is found.

Overall result Combining the insights that the six models provided, there is some support for Hypothesis 1. The US models combined with the strongest performing model from Europe confirmed the statement that there is a relationship between market volatility and investor sentiment. Specifically, high volatility seems to decrease investor sentiment. The low volatility variable did not show consistent results. It could therefore only be stated that high volatility initiates bearish investor behavior.

Model 2 Results

Model 2 evaluated if investor sentiment has a positive relationship with risk taking. It is argued that when investors feel bullish, they tend to be willing to take more risk. The results, however, show no (or hardly any) significant results when running this analysis for any of the six samples. Therefore, Hypothesis 2 is rejected and it could not be concluded that a positive investor is indeed willing to take more risk.

Model 3 Results

The third model evaluates the hypothesis that low volatility stimulates risk taking. This hypothesis follows the Minsky's arguments in the Financial Instability Hypothesis.

Table 5.3: Model 3 Monthly Results: US

	Dependent variable:	
	$RISK_m$	
	(DJIA)	(Nasdaq)
V^{low}_{m-0}	−0.013	0.007
	(0.018)	(0.017)
V^{high}_{m-0}	−0.012*	−0.008*
	(0.009)	(0.008)
V^{low}_{m-1}	0.002	0.022*
	(0.022)	(0.017)
V^{high}_{m-1}	−0.003	0.003
	(0.013)	(0.008)
V^{low}_{m-2}	0.029***	
	(0.018)	
V^{high}_{m-2}	0.006	
	(0.009)	
Constant	0.108**	0.026
	(0.044)	(0.050)
Observations	241	206
R^2	0.046	0.050
Adjusted R^2	0.022	0.031
Residual Std. Error	0.376	0.389
F Statistic	1.892*	2.624**
Note:	*p<0.1; **p<0.05; ***p<0.01	

The results in the US do not seem to be as relevant as expected. The S&P series showed no statistical significance whatsoever. For the DJIA and Nasdaq series some lagged low volatility streams did show a positive relation with risk taking. Implying that if volatility is low in the previous months, an investor is more likely to have taken more risk in the months following. Likewise, if volatility is high this month, the investor is less likely to take more risk in the current month. These results are according to Hypothesis 3. Nevertheless, the Adjusted R^2 is considerably low and additionally, the models are just barely significant. These results should therefore be regarded with care.

Table 5.4: Model 3 Monthly Results: Europe & China

	Dependent variable:		
	$RISK_m$		
	(DAX)	(EURO)	(FXI)
V^{low}_{m-0}	−0.007	0.014	0.022
	(0.016)	(0.015)	(0.036)
V^{high}_{m-0}	0.010**	0.003	0.020
	(0.009)	(0.008)	(0.020)
V^{low}_{m-1}	−0.001	−0.011	
	(0.019)	(0.017)	
V^{high}_{m-1}	0.004	0.005	
	(0.012)	(0.011)	
V^{low}_{m-2}	−0.009	0.013*	
	(0.019)	(0.017)	
V^{high}_{m-2}	0.005*	0.006	
	(0.012)	(0.011)	
Constant	−0.120**	0.055	0.705***
	(0.051)	(0.057)	(0.138)
Observations	165	225	77
R^2	0.208	0.043	0.013
Adjusted R^2	0.168	0.008	−0.013
Residual Std. Error	0.308	0.408	0.672
F Statistic	5.133***	1.217	0.498
Note:	*p<0.1; **p<0.05; ***p<0.01		

The European models do not show more significant results. The model for Europe and China are not overall significant according to the F-statistic. The DAX model is, and shows a positive relationship with the zero and two month lagged high volatility stream and risk taking. This is contrary to the expectation that volatility has a negative relationship with risk taking.

Overall result Even though the two models from the US (DJIA and Nasdaq) show significant results confirming Hypothesis 3, no strong support is found in the other models. In fact, half of the models in the sample did not show any significant results at all. Therefore, it is concluded that Hypothesis 3 is rejected and that there is no support for the statement that volatility has a relationship with the amount of risk that is taken by the investor.

Model 4 Results

Then, the relationship between the amount of risk that is taken in the previous periods and market returns is evaluated. It is believed that when investors take more risk, bigger losses are realized when the market takes a turn for the worst. This is a lagged relationship where an investor takes more risk by increasing its leverage and is more likely to suffer losses in future months.

TABLE 5.5: Model 4 Monthly Results: US

	Dependent variable:
	R_m
	(S&P)
$RISK_{m-0}$	0.020
	(0.025)
$RISK_{m-1}$	0.009
	(0.044)
$RISK_{m-2}$	-0.033^{*}
	(0.025)
Constant	0.007^{**}
	(0.002)
Observations	334
R^2	0.020
Adjusted R^2	0.011
Residual Std. Error	0.041
F Statistic	2.257^{*}
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 5.5 shows that only the model of the S&P showed significant results. Risk taken two months ago, seem to have a negative relationship in the current month. This relationship is however only slightly significant. Additionally, the adjusted R^2 is considerably low and therefore questions its goodness-of-fit.

TABLE 5.6: Model 4 Monthly Results: Europe & China

	Dependent variable:
	R_m
	(EURO)
$RISK_{m-0}$	-0.002
	(0.016)
$RISK_{m-1}$	0.015
	(0.023)
$RISK_{m-2}$	0.014
	(0.023)
$RISK_{m-3}$	-0.030^{***}
	(0.016)
Constant	0.0003
	(0.004)
Observations	225
R^2	0.025
Adjusted R^2	0.007
Residual Std. Error	0.054
F Statistic	1.384
Note:	*p<0.1; **p<0.05; ***p<0.01

Reviewing the results of the European and Chinese sample, the goodness-of-fit concern is confirmed. Although the Eurostoxx model showed a significant relationship with the three months lagged RISK variable, the general significance of the model is absent according to the F-statistic. So, even though the Eurostoxx model showed a significant coefficient in line with Hypothesis 4, the overall model is not significant.

Overall result The S&P model showed results in support of Hypothesis 4. The remainder of the models, however, did not. When this finding is combined with the considerably low Adjusted R^2 of the US model, Hypothesis 4 is Rejected. Therefore, there is no support for notion that the level of risk taking in the previous months affects market returns.

Model 5 Results

Model 5 is the first model that investigates the direct link between volatility and market returns. It is expected that the two have a negative relationship in the current month but a different relationship when lagged. This model will function as the basis in understanding if there is any, and what the relationship between volatility and market returns is. Model 6 will increase this understanding by evaluating not just market returns but market crashes.

TABLE 5.7: Model 5 Monthly Results: US

	Dependent variable:		
	R_m		
	(S&P)	(DJIA)	(Nasdaq)
V_{m-0}	−0.008***	−0.008***	−0.009***
	(0.0004)	(0.001)	(0.001)
V_{m-1}	0.006***	0.008***	0.007***
	(0.001)	(0.001)	(0.001)
V_{m-2}	0.001*	−0.001	
	(0.0004)	(0.001)	
V_{m-3}		−0.002**	
		(0.001)	
V_{m-4}		0.002***	
		(0.001)	
Constant	0.006***	0.005**	0.002
	(0.002)	(0.002)	(0.004)
Observations	334	239	206
R^2	0.550	0.372	0.284
Adjusted R^2	0.546	0.358	0.277
Residual Std. Error	0.028	0.034	0.058
F Statistic	134.453***	27.598***	40.300***
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 5.7 presents the results for the US sample. As can be seen, volatility in the same period does have a negative relationship with market returns. This implies (for example) that when volatility in the current period is high, market returns are most likely low. The model does show contrary results when lagged one month. This makes the concept of Minsky, that there often is low volatility before big market declines, reasonable. If this is indeed the case will be further investigated by Model 6. Some additional significant relationships are found when even greater lags are added to the model. However, these relationships have a smaller magnitude. Lastly, the adjusted R^2 for each of the models is relatively high. Especially for the S&P model, making the found results even more relevant.

Table 5.8: Model 5 Monthly Results: Europe & China

	Dependent variable:		
	R_m		
	(DAX)	(EURO)	(FXI)
V_{m-0}	−0.009***	−0.008***	−0.007***
	(0.001)	(0.001)	(0.002)
V_{m-1}	0.008***	0.008***	0.005**
	(0.001)	(0.001)	(0.002)
V_{m-2}	0.001	−0.001**	
	(0.002)	(0.001)	
V_{m-3}	−0.004***		
	(0.001)		
V_{m-4}	0.003***		
	(0.001)		
Constant	0.007*	−0.0003	0.002
	(0.004)	(0.003)	(0.007)
Observations	164	226	76
R^2	0.389	0.381	0.134
Adjusted R^2	0.370	0.373	0.111
Residual Std. Error	0.045	0.043	0.061
F Statistic	20.118***	45.595***	5.670***
Note:	*p<0.1; **p<0.05; ***p<0.01		

As for the US sample, Europe and China show a similar relationship between volatility and market returns. In the current month, a clear negative relationship is found whereas the 1 lagged relationship is positive. Likewise, greater lags in the DAX and EURO model show some additional statistical significant coefficients. This implies that (in some cases) volatility in three or even four months prior has some explanatory power according to the model.

Overall result The overall result supports Hypothesis 5 and indicates that there might be additional predictive power withing the volatility index. All models showed similar results with a relatively high adjusted R^2.

Model 6 Results

This model tends to answer if volatility indeed can be used to explain market crashes. In particular, there is interest in if low volatility is an indicator of market crashes in the coming periods. To do so, a logistic regression is performed on the low and high volatility stream to find the relationship with the binary variable ($CRASH_m$) that marks the months that had severe market declines.

TABLE 5.9: Model 6 Monthly Results: US

	Dependent variable:		
	$CRASH_m$		
	(S&P)	(DJIA)	(Nasdaq)
V^{low}_{m-0}	−0.757**	−1.205***	−0.388**
	(0.328)	(0.370)	(0.169)
V^{high}_{m-0}	0.609***	0.357***	0.437***
	(0.141)	(0.105)	(0.103)
V^{low}_{m-1}	0.406**	0.865***	0.402**
	(0.207)	(0.251)	(0.156)
V^{high}_{m-1}	−0.318***	−0.179**	−0.288***
	(0.115)	(0.086)	(0.090)
V^{low}_{m-2}	−0.179		
	(0.195)		
V^{high}_{m-2}	−0.239*		
	(0.145)		
V^{low}_{m-3}	0.284*		
	(0.165)		
V^{high}_{m-3}	0.169*		
	(0.100)		
Constant	−4.029***	−3.277***	−2.383***
	(0.680)	(0.590)	(0.465)
Observations	333	242	206
McFadden R^2	0.556	0.432	0.294
Note:			*p<0.1; **p<0.05; ***p<0.01

The analysis of the US sample revealed some interesting results. The S&P sample found significant results up to three months of lags. All three models show the expected relationship in the current month with the high and low volatility variable. When the low volatility variable was lagged for only one month, a positive sign is found. This result implies that the chance that there will be a market crash in the next month is higher. However, when the low volatility stream is lagged even more in the S&P sample, a negative relationship is found. This relationship is not significant contrary to the additional lagged three month relationship. Again, a positive sign is found for the low volatility stream. Interestingly, the high volatility stream also shows a positive relationship when lagged three months. This high volatility streams shows a negative relationship when only lagged for one month. This implies that when the current month is showing high volatility, the probability of facing a market crash in the consecutive month is lower. This is probably due to the fact that during market crashes, volatility is usually relatively high. The data sample reveals that there are not a lot of periods that show successive months that are marked as CRASH months. According tot the McFadden R^2, all three models have a decent goodness-of-fit implying that the volatility variable has a certain extend of explanatory power when evaluating market crashes.

Table 5.10: Model 6 Monthly Results: Europe & China

	Dependent variable:		
	$CRASH_m$		
	(DAX)	(EURO)	(FXI)
V^{low}_{m-0}	−1.105***	−0.257*	−0.345
	(0.321)	(0.134)	(0.210)
V^{high}_{m-0}	0.280**	0.601***	0.254*
	(0.116)	(0.142)	(0.130)
V^{low}_{m-1}	0.800***	0.299**	0.330
	(0.213)	(0.128)	(0.211)
V^{high}_{m-1}	−0.125	−0.465***	−0.215
	(0.099)	(0.131)	(0.163)
Constant	−2.687***	−2.694***	−1.655***
	(0.562)	(0.508)	(0.625)
Observations	167	227	76
McFadden R^2	0.369	0.352	0.141
Note:			*p<0.1; **p<0.05; ***p<0.01

For the European sample and the low volatility variable, again, explanatory power is found in the one month lagged variable. The chances of a market crash in the current month when there is low volatility in the previous month are increased. For the high volatility stream, only in the EURO sample explantory power is found with a negative relationship. For the Chinese sample (which is way smaller $n = 76$), no statistical significant relationship between low volatility and market crashes is found. Only the high volatility stream shows some significance. The overall accuracy of that model is not that high either.

Overall result The overall result of the six models indicates that low volatility indeed is a good explanatory variable of market crashes in the coming months. According to the four best performing models, if the current month shows low volatility, there is an increase chance of facing market crash in the next month. As for the high volatility variable, some explanatory power was found there as well. Mainly stating that if there is high volatility in the current month, the chance of having a market crash in the next month is decreased. This is probably due to the fact that when there is high volatility in the market, there is probably already currently a market crash occurring.

Model 7 Results

The last model that is tested in the monthly analysis contains the high and low volatility stream, the investor sentiment, risk taking and control variables. The aim is to evaluate the found results by model 6. This model indicated that volatility in the months prior to a market crash hold a significant relationship. By adding control variables, it is tested if this result is actually caused by volatility and not an overall market movement. Table 5.11 presents the results of all six models.

Apart from the S&P series, the one month lagged low volatility stream remains positive and significant when the control variables are added. The S&P series does, however, show a positive and significant results for the additionally lagged low volatility variable of three months. Therefore, even though not all lags of the low volatility variables are equal, all hold explanatory power when evaluating the probability of market crash. This model also reveals that when the high volatility variable is lagged by a month, the probability of a market crash in the next month is lower. This is probably as a result of the fact that months of high volatility already have a bigger chance of containing a market crash. The returns series did not often show crash months occurring in succession. Each of the models show a relatively high McFadden R^2 indicated a good overall fit. This monthly analysis indicates that low volatility in the period prior to a market crash holds some explanatory power. Evaluating this relationship on a quarterly basis to potentially expose its predictive power is therefore definitely interesting.

TABLE 5.11: Model 7 Monthly Results

	Dependent variable:					
	$CRASH_m$					
	(S&P)	(DJIA)	(Nasdaq)	(DAX)	(EURO)	(FXI)
V^{low}_{m-0}	−0.817**	−1.334***	−0.328**	−0.553	−0.405**	−0.532
	(0.398)	(0.426)	(0.182)	(0.472)	(0.236)	(0.349)
V^{low}_{m-1}	0.215	0.917***	0.479***	1.133**	0.343*	0.784**
	(0.242)	(0.291)	(0.174)	(0.616)	(0.219)	(0.394)
V^{low}_{m-2}	−0.065					
	(0.231)					
V^{low}_{m-3}	0.305**					
	(0.208)					
V^{high}_{m-0}	0.837**	0.488**	0.475***	0.509***	0.720***	0.221
	(0.224)	(0.146)	(0.131)	(0.246)	(0.256)	(0.182)
V^{high}_{m-1}	−0.430*	−0.344**	−0.377***	−0.413	−0.621**	−0.236
	(0.167)	(0.127)	(0.112)	(0.304)	(0.256)	(0.239)
V^{high}_{m-2}	−0.576**					
	(0.261)					
V^{high}_{m-3}	0.266*					
	(0.148)					
$CRASH_{m-1}$	−0.971	−1.569	−0.148	−0.748	−0.877	−1.106
	(1.290)	(1.084)	(0.783)	(1.641)	(1.115)	(1.393)
$RISK_{m-0}$	6.417**	1.679	−4.661**	−1.706	−0.253	2.260
	(3.555)	(3.101)	(2.493)	(4.617)	(1.352)	(1.693)
$RISK_{m-1}$	−6.212**	−1.967	4.480*	−3.290	1.243	−1.990
	(3.240)	(2.834)	(2.314)	(3.801)	(1.333)	(1.767)
$SENT_{m-0}$	1.709	0.883	−0.471	−21.194*	−0.217***	−1.707**
	(2.414)	(0.794)	(0.926)	(33.116)	(0.079)	(0.842)
$SENT_{m-1}$	0.948	−0.707	0.868	−37.855***	0.203**	1.290**
	(2.394)	(0.737)	(1.129)	(26.555)	(0.080)	(0.596)
INF_{m-0}	0.714	0.817	0.358	4.608	0.203	−14.439
	(1.836)	(1.254)	(1.033)	(12.489)	(4.016)	(15.705)
INF_{m-1}	−0.092	−0.840	0.180	−5.799	−0.751	19.518
	(1.913)	(1.295)	(1.077)	(12.349)	(4.026)	(18.950)
$DEBT_{m-0}$	0.00001	0.00001*	0.00001	−10.661**	1.031	92.258**
	(0.00002)	(0.00001)	(0.00001)	(8.042)	(3.323)	(62.108)
$DEBT_{m-1}$	−0.00002	−0.00002**	−0.00001	13.402**	1.303	−74.058*
	(0.00002)	(0.00001)	(0.00001)	(9.396)	(3.272)	(54.791)
GPC_{m-0}	0.002	0.004	0.001	0.104***	0.003	−0.919***
	(0.008)	(0.006)	(0.005)	(0.056)	(0.009)	(0.450)
GPC_{m-1}	−0.018**	−0.013*	−0.012**	−0.118***	−0.004	0.750***
	(0.008)	(0.007)	(0.006)	(0.065)	(0.008)	(0.363)
Constant	−2.406**	−1.880***	−1.956**	−5.665***	−2.900***	−9.957**
	(1.199)	(0.975)	(0.918)	(2.700)	(0.909)	(4.830)
Observations	261	236	200	87	166	59
McFadden R^2	0.633	0.514	0.419	0.538	0.563	0.399
Note:						*p<0.1; **p<0.05; ***p<0.01

5.3.2 Quarterly Analysis

This section presents the analysis of the US stock indices on a quarterly basis. All data that is used to evaluate the lagged relationships is available at the moment of forecasting. This creates the possibility to make predictions using the mentioned factors if significant results are found.

Model 1 Results

The first model tends to find a relationship between the investor sentiment and market volatility. It is expected that when volatility is low in the months prior, the investor will become more bullish.

TABLE 5.12: Model 1 Quarterly Results

	Dependent variable:		
	$SENT_q$		
	(S&P)	(DJIA)	(Nasdaq)
V_{q-1}^{low}	-0.015^{**}	0.022^{*}	-0.030^{***}
	(0.006)	(0.013)	(0.011)
V_{q-1}^{high}	0.009^{**}	-0.001	-0.005
	(0.004)	(0.009)	(0.006)
Constant	0.018	-0.044	0.088^{**}
	(0.022)	(0.049)	(0.043)
Observations	85	71	59
R^2	0.187	0.047	0.115
Adjusted R^2	0.167	0.019	0.083
Residual Std. Error	0.133	0.276	0.206
F Statistic	9.409^{***}	1.686	3.622^{**}
Note:	*p<0.1; **p<0.05; ***p<0.01		

What became apparent while evaluating model 1, is that the relationships that are significant are not lagged more than 1 period. This implies that there are only short term effect of volatility on market sentiment. Unlike the monthly analysis, there is more support for the low volatility stream. The S&P and Nasdaq show a negative relationship between low volatility and market sentiment. This implies, that when there is low volatility, according to this model, market sentiment will go down three months from now. This is a surprising finding, since it is expected that low volatility has a positive effect on market sentiment. Contrarily, if the expectation is correct that low volatility month are often present in the ramp up to a market crash, this result is not that deviating. This because in times of a market crash, the investor tends to be bearish instead of bullish. The period of three months might be too big to effectively reveal the relationship between the two measures.

Model 2 Results

The second model investigates if bullish investor sentiment has a positive effect on the level of risk that an investor is willing to take. The monthly analysis did not find any significant results.

TABLE 5.13: Model 2 Quarterly Results

	Dependent variable:
	(Nasdaq)
	$RISK_q$
$SENT_{q-1}$	1.576^{***}
	(0.821)
$SENT_{q-2}$	1.836^{***}
	(0.840)
Constant	0.011
	(0.135)
Observations	57
R^2	0.288
Adjusted R^2	0.262
Residual Std. Error	1.009
F Statistic	10.939^{***}
Note:	*p<0.1; **p<0.05; ***p<0.01

Contrary to the results of the monthly analysis of model 2, some significant results are found while quarterly evaluating this relationship. The time series of the Nasdaq found significant results for the first and second lagged effect of the sentiment variable on the level of risk the investor is willing to take. Both relationships show a positive effect. This implies that when the investor sentiment is high, in the first and second quarter before the current month, it is more likely that the investor is willing to take more risk (and vice versa). The Adjusted R^2 is also decent, making the results of the analysis more relevant. The fact that the S&P and the DJIA series did not show any significant results reduce the model's credibility.

Model 3 Results

The third model evaluates if high (or low) volatility has an significant effect on the level of risk that an investor is willing to take. It is expected that when volatility is low, an investor is more likely to increase their levels of leverage and therefore increase their own risk exposure.

TABLE 5.14: Model 3 Quarterly Results

	Dependent variable:		
	$RISK_q$		
	(S&P)	(DJIA)	(Nasdaq)
V^{low}_{q-1}	−0.108	−0.259***	−0.196***
	(0.048)	(0.044)	(0.059)
V^{low}_{q-2}	−0.089*		−0.138***
	(0.048)		(0.059)
V^{high}_{q-1}	−0.038	−0.063***	−0.025
	(0.031)	(0.029)	(0.028)
V^{high}_{q-2}	−0.001		0.017
	(0.031)		(0.028)
Constant	0.697***	0.980***	0.810***
	(0.150)	(0.160)	(0.196)
Observations	107	77	64
R^2	0.169	0.316	0.456
Adjusted R^2	0.137	0.298	0.419
Residual Std. Error	0.960	0.948	0.886
F Statistic	5.193***	17.115***	12.373***
Note:			*p<0.1; **p<0.05; ***p<0.01

Interestingly, all the found significant results show a negative relationship. While most significant relationships are found in combination with the low volatility stream, the result indicates an opposite effect as expected. According to the model, if volatility is low in the quarters prior, the level of risk taking in the current quarter is lower. The monthly analysis found some positive and significant results for the low volatility stream, however, the explanatory power was relatively low. Here, the Nasdaq model entails the most explanatory power. This model shows that when volatility reaches extreme levels (high or low), risk taking is decreased in the coming periods. This relationship is echoed by the other two models. This results contradicts the expectation of Minsky.

Model 4 Results

This model covers the subsequent step in the reasoning of Minsky; additional levels of risk taking increases the amount and likelihood of market losses. Therefore, a negative relationship is expected. The monthly analysis rejected this hypothesis.

TABLE 5.15: Model 4 Quarterly Results

	Dependent variable:		
	R_q		
	(S&P)	(DJIA)	(Nasdaq)
$RISK_{q-1}$	−0.0004	−0.001	−0.009
	(0.010)	(0.012)	(0.018)
$RISK_{q-2}$	0.017	0.021	0.038
	(0.011)	(0.012)	(0.019)
$RISK_{q-3}$	−0.028***	−0.028***	−0.044*
	(0.010)	(0.011)	(0.018)
Constant	0.023***	0.016*	0.021*
	(0.007)	(0.009)	(0.015)
Observations	109	78	66
R^2	0.079	0.083	0.110
Adjusted R^2	0.052	0.046	0.067
Residual Std. Error	0.074	0.077	0.116
F Statistic	2.991**	2.228*	2.560*
Note:	*p<0.1; **p<0.05; ***p<0.01		

According to the model, there is indeed a negative relationships measurable while evaluating market returns and the level of risk taking in the quarter(s) prior. All three models show a (relatively) long term relationship where additional risk taken three quarters prior decreases market returns in the current quarter. This effect is in line with the expectation of Minsky. Even though all three models significant and negative effects, the explanatory power of each of the models is not particularly high. None of the models have an Adjusted R^2 above the 10%. Therefore, similar to the monthly analysis, there is lacking support for this relationship.

Model 5 Results

Model 5 evaluates if there is predictive power in the overall level of volatility and market returns. It is expected that the lagged relationship is positive. This follows the believe that in the quarter(s) prior to a market crash, volatility is usually low.

TABLE 5.16: Model 5 Quarterly Results

	Dependent variable:		
	R_q		
	(S&P)	(DJIA)	(Nasdaq)
V_{q-1}	0.001	−0.002	−0.001
	(0.001)	(0.002)	(0.002)
Constant	0.019***	0.015*	0.025*
	(0.007)	(0.009)	(0.014)
Observations	108	77	65
R^2	0.007	0.019	0.001
Adjusted R^2	−0.003	0.006	−0.015
Residual Std. Error	0.076	0.078	0.107
F Statistic	0.697	1.427	0.084
Note:	*p<0.1; **p<0.05; ***p<0.01		

While the monthly analysis shows perfect support for this relationship, the quarterly results are contradicting. Apparently, without using the current period in the model, the lagged relationship becomes insignificant. This implies that there is explanatory power in the overall volatility level but is lacking predictive power. It is interesting if a similar effect is noticeable while using the binary $CRASH$ variable in model 6 and 7.

Model 6 Results

This model tests if there is predictive power in the volatility level before a market crash. According to the results of the monthly analysis, there is a reasonable chance that low volatility levels in the period before a market crash can predict this event. The monthly results also showed less explanatory power contained by the high volatility stream.

TABLE 5.17: Model 6 Quarterly Results

	Dependent variable:		
	$CRASH_q$		
	(S&P)	(DJIA)	(Nasdaq)
V_{q-1}^{low}	0.278**	0.049	0.134
	(0.109)	(0.131)	(0.123)
V_{q-1}^{high}	0.169**	0.146*	0.025
	(0.078)	(0.088)	(0.063)
Constant	−2.902***	−2.026***	−1.726***
	(0.513)	(0.497)	(0.509)
Observations	108	77	65
McFadden R^2	0.186	0.117	0.145
Note:	*p<0.1; **p<0.05; ***p<0.01		

The results of the quarterly model reveal less convincing results compared to the monthly analysis. While the S&P model indicates that there is predictive power in both the high and low volatility stream, the other two models are lacking any striking results. Interestingly, both lagged volatility streams show an increase in the probability of a market crash in the following quarter. This result is similar to the findings of Danielsson, Valenzuela, and Zer (2018). All three models show a decent McFadden R^2 which implies that there is a good amount of predictive power in the significant relationships. It is interesting for evaluate the results and their significance in combination with the control variables.

Model 7 Results

The final model includes the same factors as tested in model 6 with the addition of the Risk Taking, Sentiment, and control variables. If the high and low volatility streams stay significant in this model setup, their predictive properties will seem to be respectable.

TABLE 5.18: Model 7 Quarterly Results

	Dependent variable:		
	$CRASH_q$		
	(S&P)	(DJIA)	(Nasdaq)
V^{low}_{q-1}	0.415***	0.140	0.283**
	(0.169)	(0.211)	(0.216)
V^{high}_{q-1}	0.191*	0.272	0.242*
	(0.118)	(0.146)	(0.152)
$CRASH_{q-1}$	0.580	−1.106	−1.741
	(1.011)	(1.412)	(1.616)
$RISK_{q-1}$	0.341	−0.108	0.138
	(0.383)	(0.494)	(0.579)
$SENT_{q-1}$	1.176	−1.360	0.555
	(2.522)	(1.356)	(2.588)
INF_{q-1}	1.079***	1.237**	2.798*
	(0.597)	(0.670)	(1.183)
$DEBT_{q-1}$	−0.00000	−0.00000	0.00001
	(0.00000)	(0.00000)	(0.00000)
GPC_{q-1}	−0.002	−0.002	0.003
	(0.002)	(0.002)	(0.002)
Constant	−2.905***	−2.324**	−8.388**
	(1.211)	(1.291)	(3.109)
Observations	83	70	58
McFadden R^2	0.217	0.290	0.353
Note:			*p<0.1; **p<0.05; ***p<0.01

When adding the control variables to model 6, the low volatility stream that is lagged with one quarter keeps showing a positive and significant relationship with crash variable in the S&P model. Additionally, the low volatility stream in the Nasdaq model is also showing a significant relationship. The DJIA model lost its significance as obtained in model 6. Nevertheless, the results imply that low volatility in the period prior to a market crash indeed increases the probability of containing market declines. The predictive power of the high volatility stream in the S&P model decreased. However, the Nasdaq model shows a significant result for the high volatility contrary to model 6. These two found significant relationships are holding a positive sign indicating that high volatility in the period prior could also be used to predict a market crash. Therefore, it can be concluded that high and low volatility do possess some predictive properties for market crashes but this effect is stronger for the latter. This is contrary to results of the monthly analysis but similar to the findings of Danielsson, Valenzuela, and Zer (2018) that also conclude both are. Lastly, the McFadden R^2 is relatively high for each of the models. This indicates that the model has a decent goodness-of-fit increasing the strength of volatility as a predictor of market crashes.

5.3.3 Results Overview

The previous sections presented the results of each of the models and the analysis of the monthly and quarterly time series. Table 5.19 summarizes the found results and if there is support for the hypothesis that is tested by the model.

TABLE 5.19: Overview of Analysis Results

Hypothesis Results				
Hypothesis	Dependent Variable	Independent Variable	Monthly Significant	Quarterly Significant
1	Investor Sentiment	Market Volatility	Yes	No
2	Risk Taking	Investor Sentiment	No	No
3	Risk Taking	Market Volatility	No	No
4	Market Returns	Risk Taking	No	No
5	Market Returns	Market Volatility	Yes	No
6	Market Crashes	Market Volatility	Yes	Yes

First and foremost, the analysis revealed that indeed, volatility has a strong connection to market returns and even market crashes. The monthly and quarterly data analysis revealed that there was a significant relationship between volatility and market crashes. While the explanatory monthly analysis revealed that the overall level of volatility also hold a significant relationship with market returns, the predictive property was lacking.

Unfortunately, the three models that included the risk taking variable did not show (enough) significance. This might imply that there is such relationship as expected by Minsky. It could also implicate that the used measure, Credit-to-GDP Ratio, did not represent risk taking behaviour correctly.

The monthly analysis did indicate that the investor is affected by the level of volatility in the market. Since implied volatility is by definition depending on the investor and its perception of the market this is not a surprising discovery. However, the analysis did indicate that when volatility levels are high, the investor becomes more Bearish which is in line with the expectation. There was lacking support for a similar relationship with low volatility.

Average Marginal Effect

The previous section reveals that that volatility holds a statistical significant effect with market crashes. Models 6 and 7 indicate the level of significance and the relationship sign. However, since logistic regression is used, the direct effect of the volatility channels is not interpretable from the tables. Therefore, this section introduces the Average Marginal Effect (AME) of each of the significant variables. This AME approach calculates the average effect of each of the regression independent variables and presents the effect as a probability. This way, it is possible to better understand the magnitude of the held relationship. The exact method of calculating and the adaption of this method in R is described by Leeper (2018).

Table 5.20 presents the marginal effects for model 6. The effect is calculated for each significant relationship in model 6. Additionally, the mean of all found effects is presented in the last column. The results indicate, for instance, that if there is low volatility in the current month and this is increased with one point, the chance of having a market crash in the next month is increased by 3.60% on average. Likewise, the results reveal an average result of -2.16% decrease in the chance of a market crash in the next month if the current month shows high volatility and is increased by one point. This is contrary to the findings of Danielsson, Valenzuela, and Zer (2018) who conclude that next to low volatility, high volatility is also a good predictor of market crashes.

Table 5.20: Average Marginal Effect Model 6

	Dependent variable:						
	$CRASH_m$						
	(S&P)	(DJIA)	(Nasdaq)	(DAX)	(EURO)	(FXI)	**Mean AME**
V^{low}_{m-0}	-0.0291	-0.0632	-0.0370	-0.078	-0.0207	-	**-4.56%**
	(0.0131)	(0.0193)	(0.0160)	(0.0213)	(0.0108)	-	
V^{low}_{m-1}	0.0156	0.0454	0.0384	0.0567	0.0241	-	**3.60%**
	(0.0081)	(0.0128)	(0.0148)	(0.0133)	(0.0103)	-	
V^{low}_{m-3}	0.0109	-	-	-	-	-	**1.09%**
	(0.0062)	-	-	-	-	-	
V^{high}_{m-0}	0.0234	0.0187	0.0417	0.0198	0.0483	0.0340	**3.10%**
	(0.0049)	(0.0050)	(0.0086)	(0.0077)	(0.0103)	(0.0160)	
V^{high}_{m-1}	-0.0122	-0.0094	-0.0275	-	-0.0374	-	**-2.16%**
	(0.0042)	(0.0043)	(0.0079)	-	(0.0098)	-	
V^{high}_{m-2}	-0.0092	-	-	-	-	-	**-0.92%**
	(0.0055)	-	-	-	-	-	
V^{high}_{m-3}	0.0109	-	-	-	-	-	**1.09%**
	(0.0038)	-	-	-	-	-	
V^{low}_{q-1}	0.0281	-	-				**2.81%**
	(0.0107)	-	-				
V^{high}_{q-1}	0.0171	0.0191	-				**1.81%**
	(0.0076)	(0.0111)	-				
Observations m	261	236	200	87	166	59	
Observations q	83	70	58				

While looking at the quarterly results, the effect of low volatility in the previous quarter is slightly lower. On average, if the current quarter shows low volatility and this is increased by one point, there is a increased chance of 2.81% of having a market crash in the next quarter. This result is, however, only based on the findings in the S&P model.

Table 5.21: Average Marginal Effect Model 7

	Dependent variable:						
	$CRASH_m$						
	(S&P)	(DJIA)	(Nasdaq)	(DAX)	(EURO)	(FXI)	**Mean AME**
V^{low}_{m-0}	-0.0301	-0.0601	-0.0249	-	-0.0218	-	**-3.42%**
	(0.0144)	(0.0185)	(0.0137)	-	(0.0124)	-	
V^{low}_{m-1}	-	0.0413	0.0364	0.0557	0.0185	0.0798	**4.63%**
	-	(0.0125)	(0.0130)	(0.0275)	(0.0116)	(0.0360)	
V^{low}_{m-3}	0.0112	-	-	-	-	-	**1.12%**
	(0.0074)	-	-	-	-	-	
V^{high}_{m-0}	0.0308	0.0220	0.0361	0.0250	0.0387	-	**3.05%**
	(0.0068)	(0.0057)	(0.0088)	(0.0108)	(0.0127)	-	
V^{high}_{m-1}	-0.0158	-0.0155	-0.0286	-	-0.0334	-	**-2.33%**
	(0.0055)	(0.0052)	(0.0077)	-	(0.0131)	-	
V^{high}_{m-2}	-0.0212	-	-	-	-	-	**-2.12%**
	(0.0093)	-	-	-	-	-	
V^{high}_{m-3}	0.0098	-	-	-	-	-	**0.98%**
	(0.0053)	-	-	-	-	-	
V^{low}_{q-1}	0.0462	-	0.0288				**3.75%**
	(0.0169)	-	(0.0209)				
V^{high}_{q-1}	0.0213	-	0.0246				**2.30%**
	(0.0125)	-	(0.0140)				
Observations m	261	236	200	87	166	59	
Observations q	83	70	58				

While looking at the complete model that includes control variables, highly similar results are found. The overall total effect of the one month lagged low volatility channel is increased to 4.63%

on average. This implies that when the control variables are added to the model, the relationship of the lagged low volatility stream is even stronger.

The quarterly results show an even stronger predictive relationship between low volatility in the previous quarter and the chance of facing a market crash in the current quarter. On average, the probability of having a market crash in the next quarter is increased by 3.75% if the current month shows low volatility and is decreased by 1 point. Again, high volatility in the previous quarter also increases the chance of having a market crash in the following quarter but showing a weaker relationship.

Hodrick & Prescott Filter Robustness

Lastly, since the results found, heavily rely in the Hodrick & Prescott filtering method, it is important to check robustness of these results relative to the filtering method. This is done by evaluating changes in the filter parameters and its effect on the found results. As argued in Appendix F, the HP filter uses the frequency λ to identify the trend in a time series. The frequency values as used in the presented models are according to their relevant time increments and the standard values as determined by Ravn and Uhlig (2002). For the monthly series this is equal to $\lambda = 129,600$ and for the quarterly series this is equal to $\lambda = 1600$. To test the results' sensitivity to the frequency, model 7 is reevaluated using a high and low value for λ. For the monthly series $\lambda_m^{low} = 100$ and $\lambda_m^{high} = 1,000,000$ are used. For the quarterly series $\lambda_q^{low} = 100$ and $\lambda_q^{high} = 100,000$ are used.

Appendix G presents the results of the robustness analysis. The monthly results do not suffer too much from changes in the filter frequency. Most of the originally found results are still significant. What can be seen, is the fact that an increase in the filter frequency (which creates a more flat trend line) creates slightly less significant results. This indicates that using the filter is useful because of trending behavior in the series. This is confirmed when looking at the results of the decreased frequency. Results tend to be equally significant and in some cases show an even better result.

The quarterly results seem to suffer more from changes in the filter frequency. When changing the frequency, no significant results are found. While it was already acknowledged that the overall model accuracy is decreased in the quarterly study, the robustness check confirmed the sensitivity of the found results. Therefore, the indicated predictive qualities of the volatility streams should be regarded with care.

Chapter 6

Conclusion & Recommendations

6.1 Conclusion

This thesis studies the relationship between volatility and market returns. More specifically, if volatility can be used to explain and even predict market crashes. Moreover, two additional factors are added to the analysis to evaluate what additional events create an unstable economy. These factors are in line with the reasoning of Minsky (Minsky, 1977) who states that periods of low volatility trigger an increase in risk taking by the investor because of financial melancholy. This believe is also known as the 'Financial Instability Hypothesis'. To further examine if this statement, measures for investor sentiment and risk taking are added to the analysis to evaluate the investor behaviour in times of prosperity.

To form an all encompassing analysis, six different data samples are used coming from the US, Europe, and China. Therefore, the main research question that is answered in this thesis is:

Can Stock Market Crashes be Explained and Predicted using Volatility, Risk Taking, and Investor Sentiment Measures in the US, Europe, and China?

To answer the main research question, multiple hypotheses are defined as presented in Figure 6.1. The tested hypotheses follow the reasoning about financial instability of Minsky in the following way: Market volatility is supposed to have a negative relationship with investor sentiment, which is presumed to have a positive relationship with Risk taking by the investor. An increase in risk taking might create instability in the market causing bigger declines in market returns. Additionally, the direct effect of volatility on risk taking and market returns is included in the analysis.

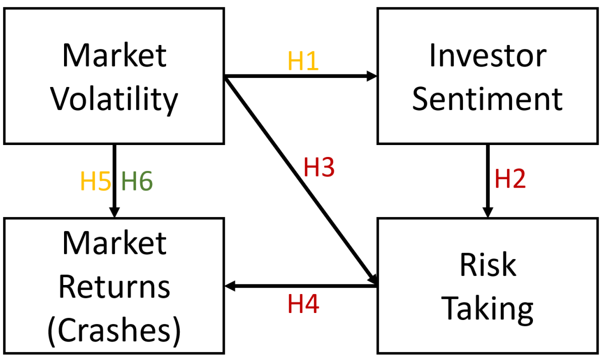


Figure 6.1: All Tested Relationships and their Significance

The conducted research is split up into two approaches: an explanatory an a predictive study. In the explanatory study, monthly time series are analyzed to find significant relationships between the proposed measures in hindsight. The predictive study used quarterly time series to evaluate the same relationships with larger increments. This way, no forward looking data was needed, implying that the analysis could be replicated in the current quarter to evaluate the following quarters.

Figure 6.1 present the results of both studies. Market Volatility seemed to have a significant relationship with Investor Sentiment (H1) in the explanatory study. This was however not the case in the predictive study. The three models that evaluated relationships with Risk Taking (H2, H3 and H4) did not show (enough) significance. The model that tested hypothesis 5 showed promising results in the explanatory study. Unfortunately, the overall level of volatility was lacking the predictive property. Model 6 (and 7), that evaluated the relation of volatility with market crashes, revealed significant results in both studies.

The most important finding in this thesis is that low volatility in the months prior to a market crash is often seen. In addition, the quarterly analysis even revealed that low volatility in the previous period is a good predictor of market crashes in most researched areas. In fact, if the current quarter shows low volatility, and this volatility is decreased by one point, there is an increase of 3.75% chance of a market crash in the next quarter on average. Additionally, the predictive study also indicated that high volatility could be used to predict market crashes with an average of 2.30%.

In order to find a broader understanding about the relationship between low volatility and market crashes, the investor sentiment assessed. According to Minsky, times of low volatility trigger euphoric investor behaviour. When evaluating this phenomenon, a significant relationship between volatility and investor sentiment was found. The explanatory study revealed that after months of high volatility, investor sentiment tends to decline. A similar relationship with low volatility was not found. Likewise, the predictive study did not find such a significant relationship. This might be caused by the fact that the quarterly increments are too large to successfully capture this effect.

While trying to expand the analysis according to the reasoning of Minsky, it was evaluated if volatility and investor sentiment indeed had an effect on the level of risk taking by the investor. This was done by using the Credit-to-GDP Ratio that evaluated the level of leverage in a country. The models did not find enough significant results to conclude that these relationships exist. Therefore, this thesis does not find support for the Financial Instability Hypothesis as a whole. Since all models that intended to find significant relationships with risk taking failed to do so, it might be the case that the Credit-to-GDP Ratio is not the most fitting measure.

Then, contrary to Danielsson, Valenzuela, and Zer (2018), this study also finds, in the monthly analysis, that the overall level of the volatility has a significant relationship with market returns. This relationship is, however not found in the quarterly predictive study. Additionally, the monthly study finds a significant relationship with the lagged high volatility stream. However, also conflicting with the results of Danielsson, Valenzuela, and Zer (2018), this relationship carries the negative sign. Implying that when the previous month showed high volatility, the likelihood of a market crash in the coming month is decreased by around 2% on average. A major difference between this study and the one of Danielsson, Valenzuela, and Zer (2018) is the time series increments. While this study uses more detailed time steps (monthly and quarterly), Danielsson, Valenzuela, and Zer (2018) opted to use a yearly time series. Since this study does find significant relationships in both studies with a small number of lags, it could be concluded that a more detailed study is relevant.

The results of this study could be used to prevent (or decrease the magnitude of) a market crash. Since implied volatility matching major stock market indices are easily accessible sources of data, one should be cautious when these indices reach low levels. Instead of believing this low level signals opportunities and indicates decreased risk, one should be aware of the fact that this might be the calm before the storm. This thesis confirms that history has shown that the latter is the case.

Concluding, this thesis evaluated Minsky's Financial Instability Hypothesis using volatility, investor sentiment and risk taking. Not enough evidence was found to completely support Misnky's hypothesis. However, it is demonstrated that low volatility successfully explains and predicts market crashes in the consecutive periods.

6.2 Recommendations

As shown in this thesis, implied volatility indices are more than just measures that indicate how much an investor is willing to pay to protect against market declines. They bear more information about the state of the economy and the overall investor behavior. Therefore, it is recommended to use this information more frequently while making investment decisions. Although some models might indicate that in times of booming economy investing is a safe option, the decrease in carefulness (as revealed by the implied volatility index) might trigger a chain of events that will start a period of financial turmoil. By actively reviewing an implied volatility index, the future state of the economy (at least on a quarterly basis) can be predicted. This should not only be used by individual investors but also by institutional investors. Additionally, it might even be desirable that policy makers use this information in order to restrict certain investments and/or debt issuance in periods of unusually low volatility.

Additionally, it is recommended to not blindly follow measures of investor and market sentiment. As this thesis indicates, investor sentiment holds a relatively low and limited amount of explanatory and predictive power. This implies that the overall investor consensus not always determines in what direction the market is headed. Using these measures in financial models should therefore be done with care.

Chapter 7

Reflection

7.1 Scientific Contribution

While economic crashes have been of interest in scientific research, research in the field of the effect of volatility on the probability of a market crash is narrow. Danielsson, Valenzuela, and Zer (2018) successfully present a study focusing on this effect. A downside of this paper is that they execute their analysis on a yearly basis. Implying the lagged effects have to be (by definition of the research design) occurring at least a year before a market crash occurs. It is believed however, that the relationship between the researched factors hold lagged effect which is bound to events that happen in a greater proximity to each other. While the peak of a market crash is mostly only experienced for a few days, it could be stated that the period surrounding is also affected. However, stating that when such an event happens, the complete year is marked as a market crash year is considered as too broad. Therefore, the monthly and quarterly approach is supposed to be a more accurate study.

7.2 Limitations & Future Research

This thesis is performed as an exploratory and predictive study in finding the potential of implied volatility to evaluate market returns. While that question can now be answered, there are some limitations in this thesis. By acknowledging these, some assumptions and the accuracy of the found results are exposed. Additionally, future research could focus on eliminating these indicated limitations in order to increase the strength of this analysis.

7.2.1 Monthly Analysis

Even though the monthly analysis certainly found interesting and useful results, these could only be used in a explanatory manner in hindsight. Because of data availability and some necessary data transformations, forward looking methods are used. As argued in this thesis, a predictive analysis on a monthly basis could form relevant insight that broadens the understanding and usage of the evaluated measures. Especially when considering some of the promising results of the quarterly predictive analysis.

7.2.2 Risk Taking

One of the main research goals of this thesis was to test Minsky's Financial Instability Hypothesis. An important part of this hypothesis is the concept that investors tend to issue more debt in times of economic euphoria. Additionally, market returns are supposed to be affected by this increase in overall leverage and risk taking behaviour. The factor that was used in this thesis to evaluate this relationship did not show the expected result. This might indicate that (on a monthly and quarterly basis) this effect does not exist. However, it could also be the case that the used factor (Credit-to-GDP Ratio) is not the most appropriate measure to evaluate this risk taking behavior. Bank level capital-to-asset ratio or non-performing loans ratio could be promising measures as mentioned by Danielsson, Valenzuela, and Zer (2018). However, retrieving this data might be difficult.

7.2.3 Time Horizon

The maximum date range in this thesis ranged from 1996 to 2018. While this is an appropriate sample size considering the analysis, extending the data rage to early 1900's could be interesting. The period between 1900 and 1996 certainly knew multiple big market declines. Evaluating if the found effect is also significant during these times would increase the overall strength of the analysis. Facing the concern of data range is specifically relevant for the analysis of the Chinese data sample. Because of the limitation of volatility data, the sample starts in 2011, missing one of the most important market crashes in recent history. The results of this analysis should therefore be treated with care.

For future research, one could opt to use realized volatility instead of implied volatility. While not relevant in this thesis since the data samples are mostly limited because of the sentiment variable, only using realized volatility and market returns could increase the sample size considerably.

7.2.4 Binary Crash Variable

To evaluate the relationship of volatility with market returns, this thesis (partly) used a binary crash variable to do so. While this approach has its benefits, it also has its limitations. Even though the benchmark of a market crash was carefully set and evaluated, creating an exact cutoff point might initiate the exclusion (or inclusion) of possible (ir)relevant events. Therefore, future research could focus on creating a non binary classification of relevant time periods. For instance, identify multiple different periods in the economic cycle and use multiclass classification as the prediction method.

7.2.5 Additional Variables

The model presented in this thesis focused on four variables and some control variables. While this is desirable when one's aim is to simplify complex market phenomena, it can be beneficial to include more explanatory variables. This, to increase the understanding about the held relationships and underlying phenomena. While this thesis confirms the suspicion about the connection of volatility and market returns, an extensive explanation about possible supplementary effects is lacking. Some examples of additional interesting variables are: interest rates, market liquidity, growth rates or bond yields.

Bibliography

Al Refai, H. and R. Abdelaziz Eissa M.and Zeitun (2017). "Asymmetric volatility and conditional expected returns". In: *International Journal of Emerging Markets* 12(2), pp. 335–351.

Aliber, R. and C. Kindleberger (2015). *Manias, panics and crashes: A history of financial crises.* 7th ed. New York, NY: Palgrave Macmillan.

Baker, M. and J. Wurgler (2007). "Investor sentiment in the stock market". In: *Journal of Economic Perspectives* 21(2), pp. 129–152.

Baker, Malcolm P., Jeffrey Wurgler, and Yu Yuan (2012). "Global, Local, and Contagious Investor Sentiment". In: *Journal of Financial Economics* 104(1), pp. 272–287.

Bandopadhyaya, A. and A. Jones (2011). "Measures of investor sentiment: A comparative analysis put-Call ratio vs. volatility index". In: *Journal of Business & Economics Research* 6(8), pp. 27–34.

Bank for International Settlements (2018). *Credit-to-GDP gaps.* URL: `https://www.bis.org/statistics/c_gaps.htm` (visited on 08/30/2018).

Barrell, R. et al. (2010). "Bank regulation, property prices and early warning systems for banking crises in OECD countries". In: *Journal of Banking & Finance* 34(9), pp. 2255–2264.

Beer, F. and M. Zouaoui (2011). "Measuring investor sentiment in the stock market". In: *SSRN Electronic Journal*, pp. 1–30.

Bekaert, G. (2000). "symmetric volatility and risk in equity markets". In: *The Review of Financial Studies* 13(1), pp. 1–42.

Bekaert, G. and M. Hoerova (2014). "The VIX, the variance premium and stock market volatility". In: *Journal of Econometrics* 183(2), pp. 181–192.

Black, F. (1976). "Studies of Stock Price Volatility Changes". In: *Proceedings of the 1976 Meeting of the Business and Economic Statistics Section*, pp. 177–181.

Blanchard, Olivier J and Mark W Watson (1982). *Bubbles, rational expectations and financial markets.*

Boersen News (2018). *PUT-CALL-SENTIMENT DAX.* URL: `https://www.boersennews.de/markt/indizes/put-call-sentiment-dax-xf0002000010/37313835/profile` (visited on 09/12/2018).

Carr, P. (2017). "Why is VIX a fear gauge?" In: *Risk and Decision Analysis* 6(2), pp. 179–185.

CBOE (2018). *CBOE Volatility Index.* URL: `http://www.cboe.com/products` (visited on 08/30/2018).

Cheema, Muhammad A, Yimei Man, and Kenneth R Szulczyk (2018). "Does Investor Sentiment Predict the Near-Term Returns of the Chinese Stock Market?" In: *International Review of Finance*, pp. 1–9.

Christie, Andrew A. (1982). "The stochastic behavior of common stock variances: Value, leverage and interest rate effects". In: *Journal of Financial Economics* 10(4), pp. 407 –432.

Clarke, Roger G. and Meir Statman (1998). "Bullish or Bearish?" In: *Financial Analysts Journal* 54(3), pp. 63–72.

Danielsson, J., M. Valenzuela, and I. Zer (2018). "Learning from history: Volatility and financial crises". In: *The Review of Financial Studies* 31(7), pp. 2774–2805.

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

Federal Reserve Bank of st.Louis (2018). *Federal Reserve Economic Data*. URL: `https://fred.stlouisfed.org/` (visited on 10/18/2018).

Feigenbaum, J. A. and P. Freund (1996). "Discrete Scaling in Stock Markets Before Crashes". In: *International Journal of Modern Physics* 10(27), pp. 3737–3745.

Fleming, J., B. Ostdiek, and R. Whaley (1995). "Predicting stock market volatility: A new measure". In: *Journal of Futures Markets* 15(3), pp. 265–303.

Giese, J. et al. (2014). "The Credit-to-GDP Gap and Complementary Indicators for Macroprudential Policy: Evidence from the UK". In: *International Journal of Finance & Economics* 19(1), pp. 25–47.

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

Hamilton, J. (2017). "Why you should never use the Hodrick-Prescott filter". In: *The Review of Economics and Statistics*, pp. 1–13.

Hodrick, R. and E. Prescott (1997). "Postwar u.s. business cycles: An empirical investigation". In: *Journal of Money, Credit and Banking* 29(1), pp. 1–16.

Houlihan, P. and G. Creamer (2017). "Can sentiment analysis and options volume anticipate future returns?" In: *Computational Economics* 50(4), pp. 669–685.

Jacobsson, E. (2009). "How to predict crashes in financial markets with the Log-Periodic Power Law". MA thesis. Department of Mathematical Statistics, Stockholm University.

Jiang, Z. et al. (2010). "Bubble diagnosis and prediction of the 2005-2007 and 2008-2009 chinese stock market bubbles". In: *Journal of Economic Behavior & Organization* 74(3), pp. 149–162.

Keynes, J. (1936). *The general theory of employment, interest and money*. Princeton paperbacks. London: Macmillan and, Limited.

Lahart, Justin (2007). *In Time of Tumult, Obscure Economist Gains Currency*. URL: `https://web.archive.org/web/20080512063221/https://online.wsj.com/public/article/SB118736585456901047.html` (visited on 10/31/2018).

Leeper, T.J. (2018). "Interpreting Regression Results using Average Marginal Effects with R's margins". In:

Lleo, S. and W. Ziemba (2012). "Stock market crashes in 2007-2009: Were we able to predict them?" In: *Quantitative Finance* 12(8), pp. 1161–1187.

McFadden, D. (1973). "Conditional logit analysis of qualitative choice behavior". In:

Merton, Robert C. (1973). "An Intertemporal Capital Asset Pricing Model". In: *Econometrica* 41(5), p. 867.

Minsky, H. P. (1977). "The financial instability hypothesis: An interpretation of keynes and an alternative to "standard" theory". In: *Nebraska Journal of Economics and Business* 16(1), pp. 5–16.

— (1992). "The Financial Instability Hypothesis". In: *The Jerome Levy Economics Institute Working Paper* 74.

Neuhauser, K. (2015). "The global financial crisis: What have we learned so far?" In: *International Journal of Managerial Finance* 11(2), pp. 134–161.

Newey, W. and K. West (1987). "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix". In: *Econometrica* 55(3), pp. 703–708.

OECD (2018). *Inflation (CPI)*. URL: `https://data.oecd.org/price/inflation-cpi.htm` (visited on 10/21/2018).

Ravn, M. and H. Uhlig (2002). "Notes: On adjusting the hodrick-prescott filter for the frequency of observations". In: *The Review of Economics and Statistics* 84(2), pp. 371–380.

Ryan, B. et al. (2002). *Research method and methodology in finance and accounting.* 2nd ed. London: Thomson.

Schwert, G. William (1989). "Why Does Stock Market Volatility Change Over Time?" In: *The Journal of Finance* 44(5), pp. 1115–1153.

Sentix (2018). *Sentix Sentiment Index.* URL: `http://www.sentix.de/index.php/en/item/sntm.html` (visited on 09/12/2018).

Sornette, D. (2003). *Why stock markets crash: Critical events in complex financial systems.* Princeton paperbacks. Princeton, N.J.: Princeton University Press.

Sornette, D., A. Johansen, and J. P. Bouchaud (1996). "Stock Market Crashes, Precursors and Replicas". In: *Journal de Physique I, EDP Sciences* 6(1), pp. 167–175.

Stock, J. and M. Watson (1999). "Forecasting inflation". In: *Journal of Monetary Economics* 44(2), pp. 293–335.

The World Bank (2018a). *GDP (current US$).* URL: `https://data.worldbank.org/` (visited on 09/10/2018).

— (2018b). *Popular Indicators.* URL: `http://databank.worldbank.org/data/indicator/NY.GDP.PCAP.CD/1ff4a498/Popular-Indicators` (visited on 10/18/2018).

Torres-Reyna, Oscar (2014). *Cubic Interpolation Using R.* Data and Statistical Services. Princeton University.

Tsuji, C. (2003). "Is volatility the best predictor of market crashes?" In: *Asia - Pacific Financial Markets* 10(2-3), pp. 163–185.

van Veen, C. (2017). "Low volatility as a Predictor of Stock Market Crashes". MA thesis. Tilburg School of Economics and Management.

Wang, Y., A. Keswani, and S. Taylor (2006). "The relationships between sentiment, returns and volatility". In: *International Journal of Forecasting* 22(1), pp. 109–123.

Wharton (2009). *Why Economists Failed to Predict the Financial Crisis.* URL: `http://knowledge.wharton.upenn.edu/article/why-economists-failed-to-predict-the-financial-crisis/`.

Wharton University of Pennsylvania (2018). *Wharton Research Data Services (WRDS).* URL: `https://wrds-web.wharton.upenn.edu/wrds/` (visited on 08/30/2018).

Youssefmir, M., B. Huberman, and T. Hogg (1998). "Bubbles and market crashes". In: *Computational Economics* 12(2), pp. 97–114.

Zhao, Yiwei, Zheng Yang, and Xiaolin Qian (2015). "Investor Sentiment and Chinese A-Share Stock Markets Anomalies". In: *International Journal of Economics and Finance* 7(9), pp. 293–312.

Zhou, W. and D. Sornette (2006). "Fundamental factors versus herding in the 2000-2005 US stock market and prediction". In: *Physica A: Statistical Mechanics and Its Applications* 360(2), pp. 459–482.

Zouaoui, M., G. Nouyrigat, and F. Beer (2011). "How does investor sentiment affect stock market crises?" In: *The Financial Review* 46(4), pp. 723–751.

Appendix A

Market Returns

A.1 Monthly Returns

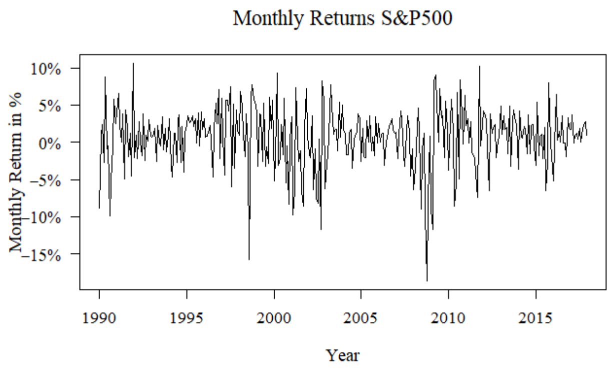


FIGURE A.1: Monthly Log Returns S&P 500

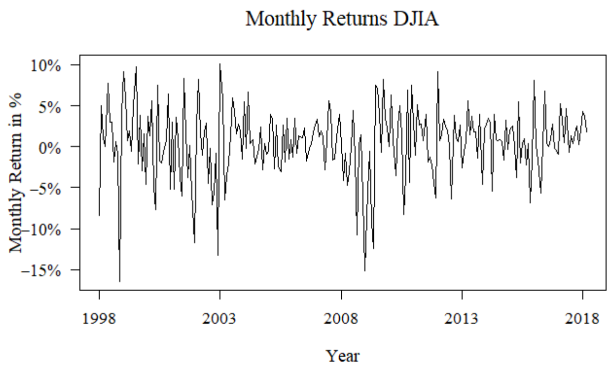


FIGURE A.2: Monthly Log Returns DJIA

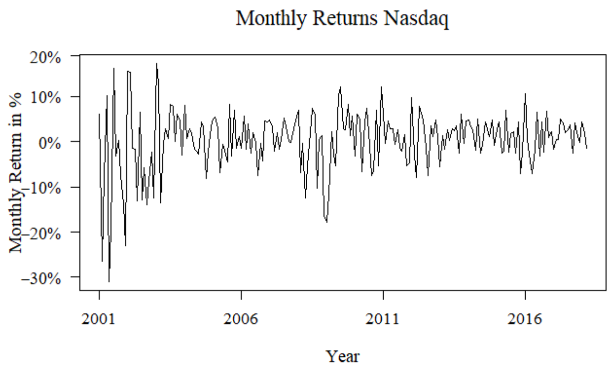


FIGURE A.3: Monthly Log Returns Nasdaq 100

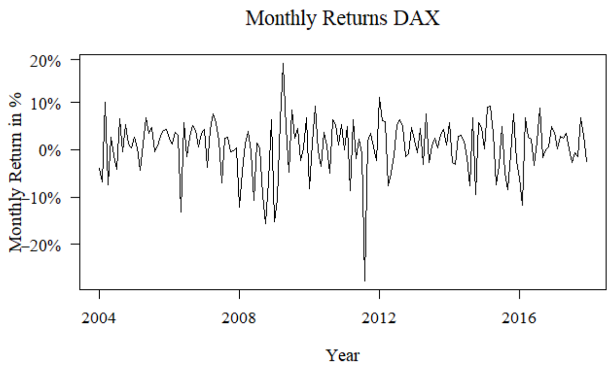


FIGURE A.4: Monthly Log Returns DAX

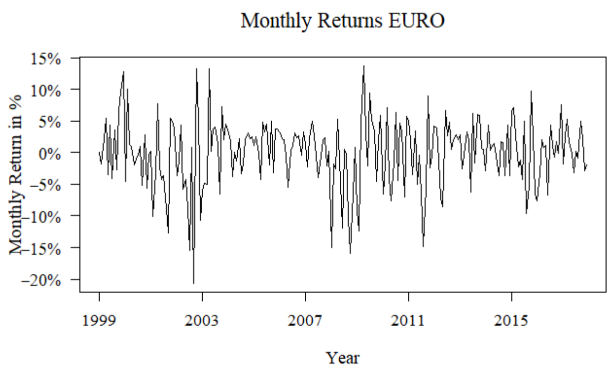


FIGURE A.5: Monthly Log Returns Euro Stoxx 50

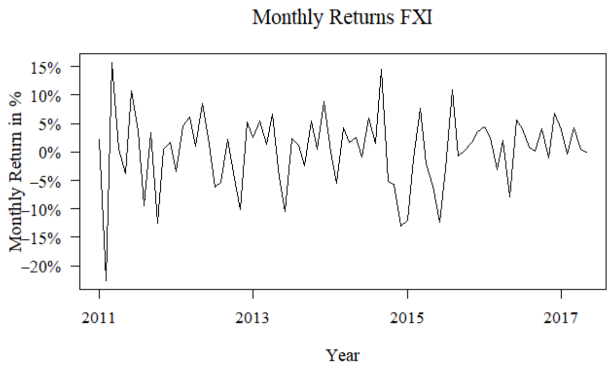


FIGURE A.6: Monthly Log Returns iShares China

A.2 Quarterly Returns

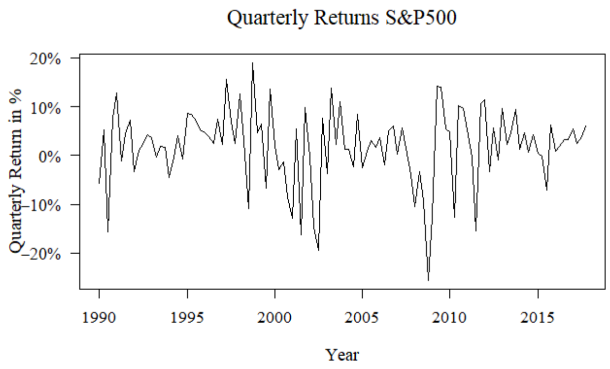


FIGURE A.7: Quarterly Log Returns S&P 500

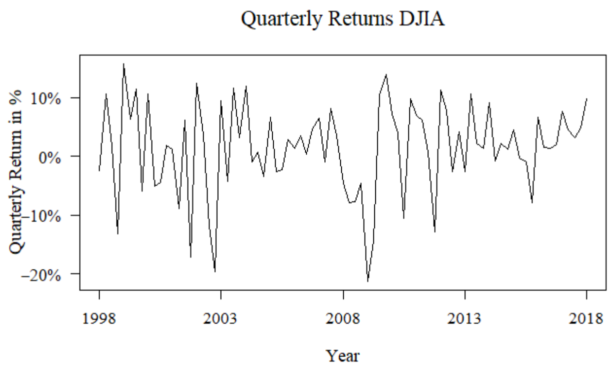


FIGURE A.8: Quarterly Log Returns DJIA

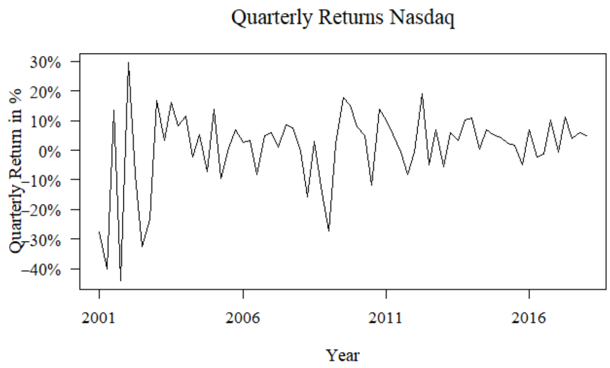


Figure A.9: Quarterly Log Returns Nasdaq 100

Appendix B

Volatility

B.1 Monthly Average Volatility

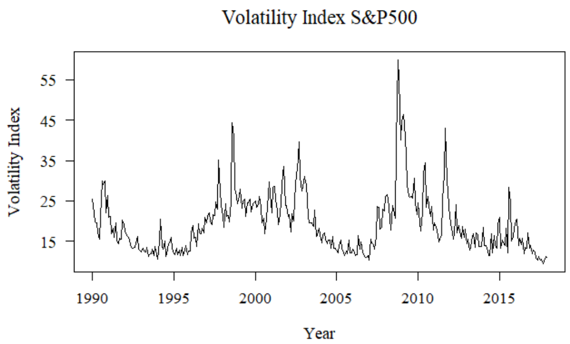


FIGURE B.1: Monthly S&P Volatility

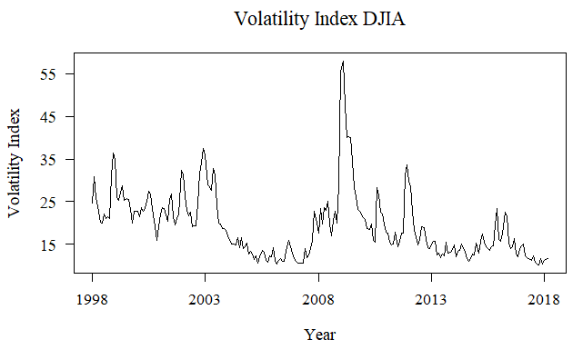


FIGURE B.2: Monthly DJIA Volatility

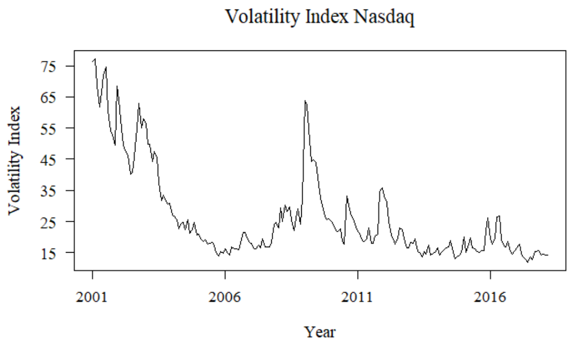


FIGURE B.3: Monthly Nasdaq Volatility

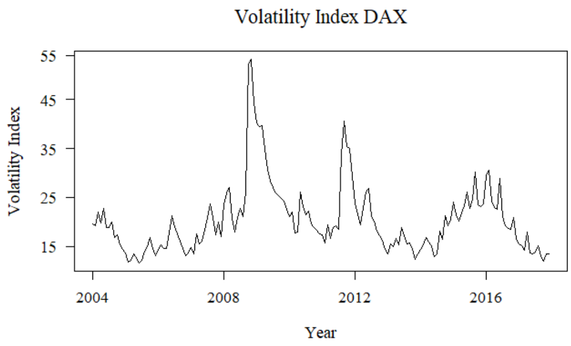


FIGURE B.4: Monthly DAX Volatility

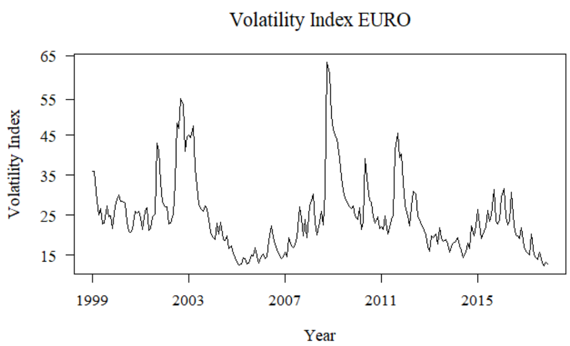


FIGURE B.5: Monthly EUROSTOXX Volatility

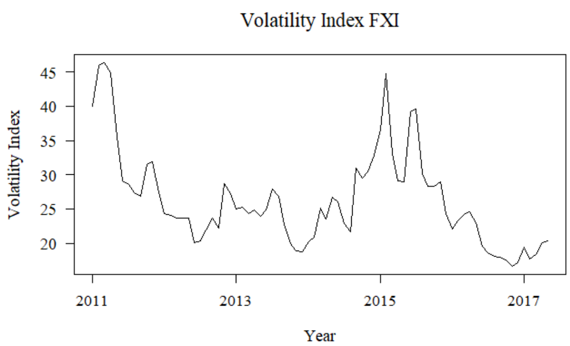


FIGURE B.6: Monthly FXI Volatility

B.2 Monthly Volatility Two-Sided HP Filtered

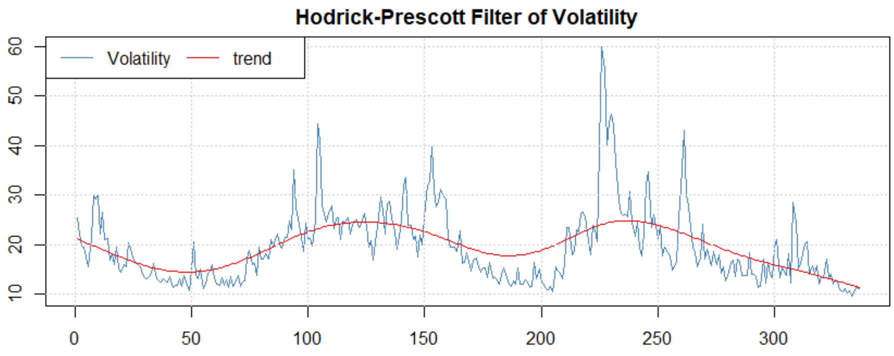


FIGURE B.7: HP2 Filter of the VIX Index

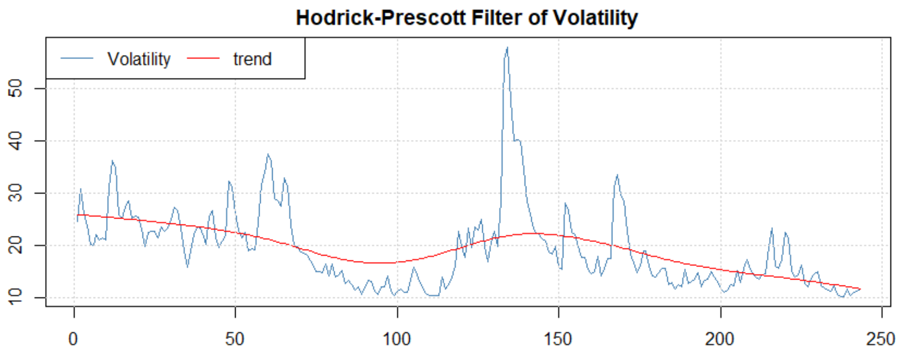


FIGURE B.8: HP2 Filter of the VXN Index

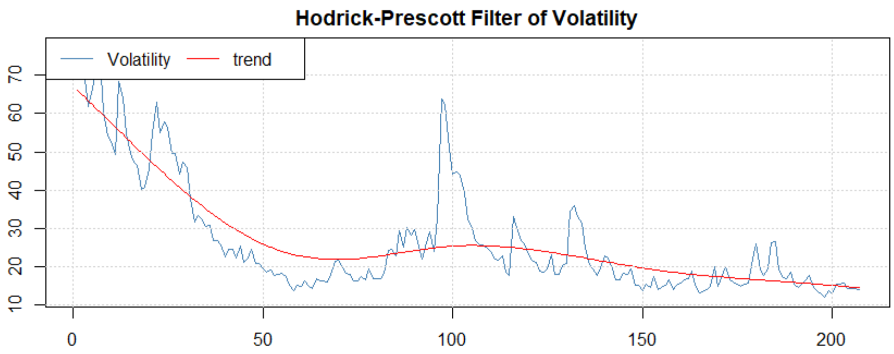


FIGURE B.9: HP2 Filter of the VXD Index

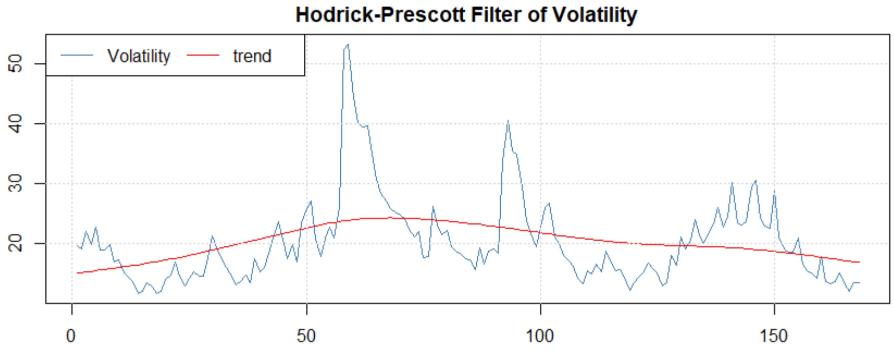


Figure B.10: HP2 Filter of the VDAX Index

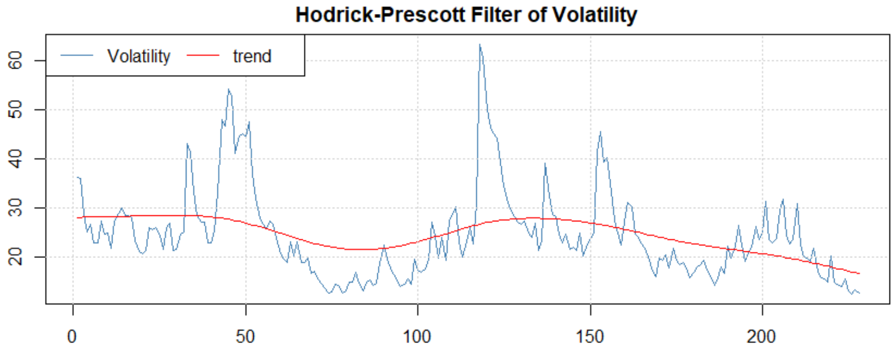


Figure B.11: HP2 Filter of the VSTOXX Index

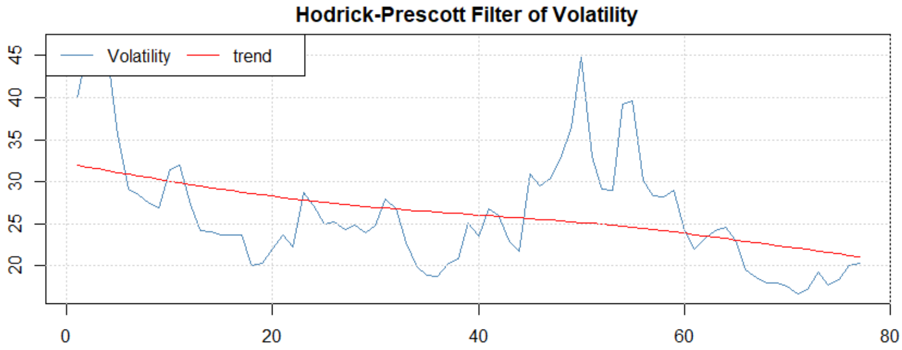
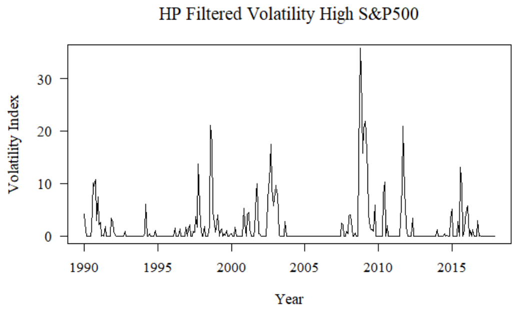
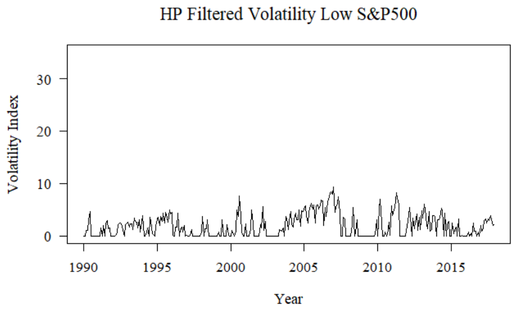


Figure B.12: HP2 Filter of the VXFXI Index

B.3 Monthly High Low Volatility

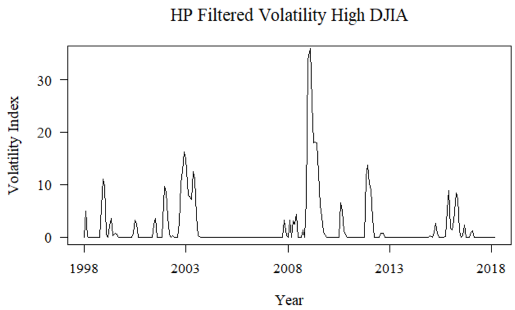


(A) High

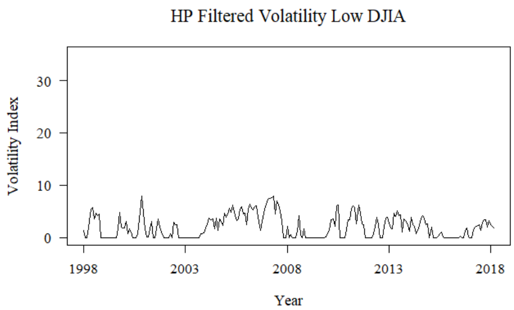


(B) Low

Figure B.13: The Monthly High and Low Volatility Variables S&P

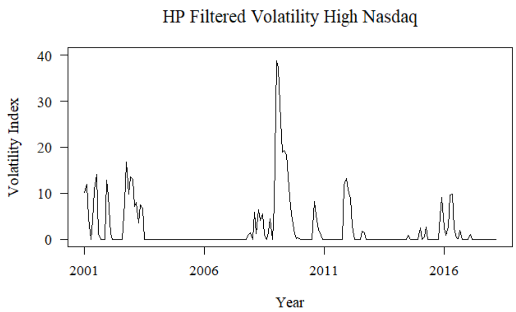


(A) High

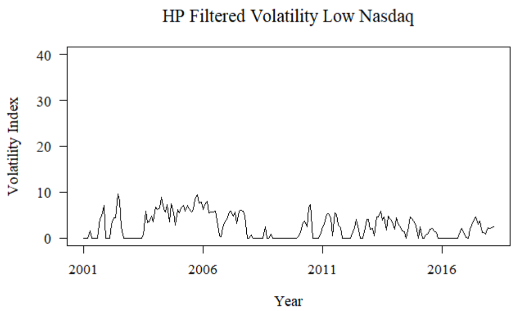


(B) Low

Figure B.14: The Monthly High and Low Volatility Variables DJIA

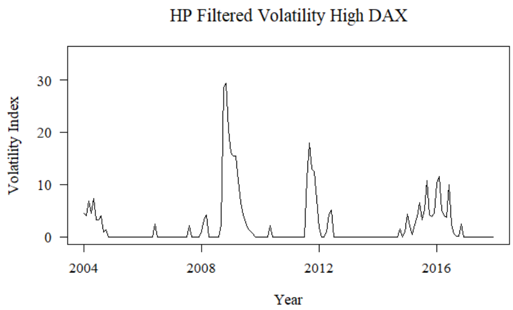


(A) High

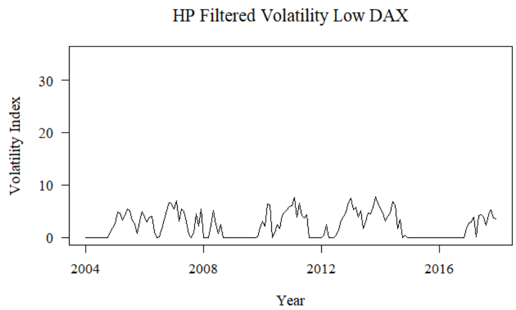


(B) Low

Figure B.15: The Monthly High and Low Volatility Variables Nasdaq

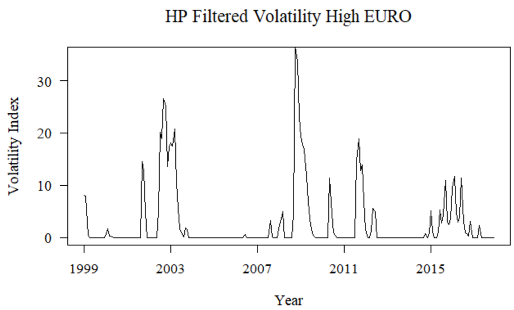


(A) High

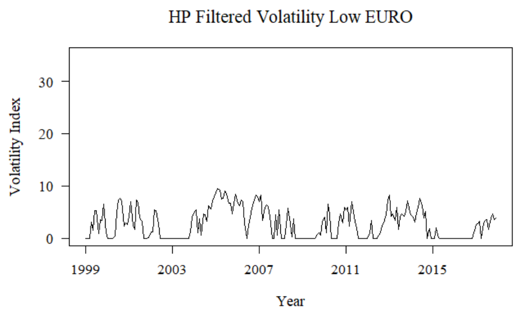


(B) Low

FIGURE B.16: The Monthly High and Low Volatility Variables DAX

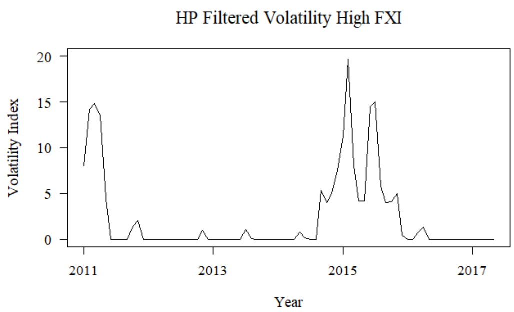


(A) High

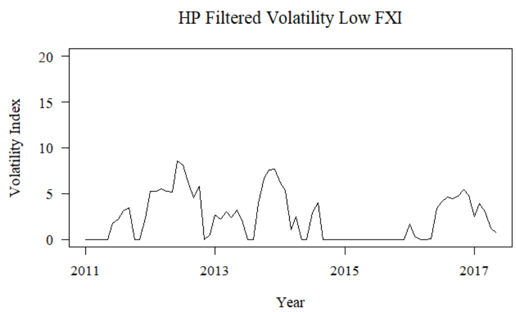


(B) Low

FIGURE B.17: The Monthly High and Low Volatility Variables EURO



(A) High



(B) Low

FIGURE B.18: The Monthly High and Low Volatility Variables FXI

B.4 Quarterly Volatility One-Sided HP Filtered

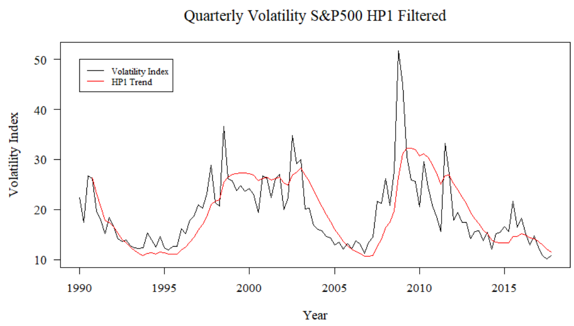


FIGURE B.19: Quarterly Volatility S&P HP1 Filtered

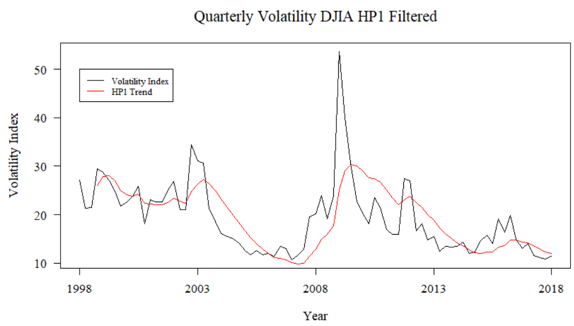


FIGURE B.20: Quarterly Volatility DJIA HP1 Filtered

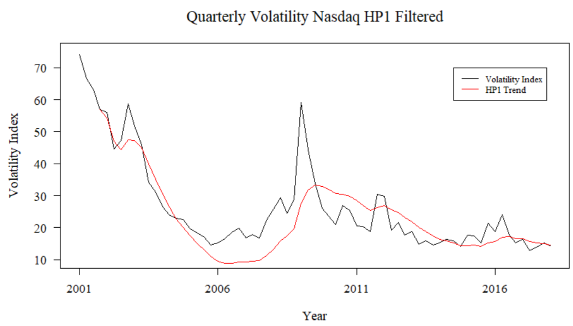
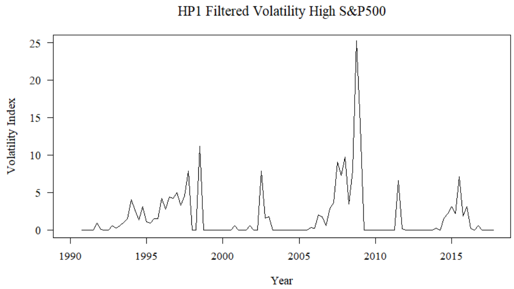
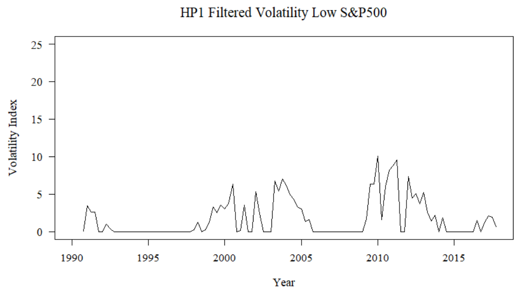


FIGURE B.21: Quarterly Volatility Nasdaq HP1 Filtered

B.5 Quarterly High Low Volatility

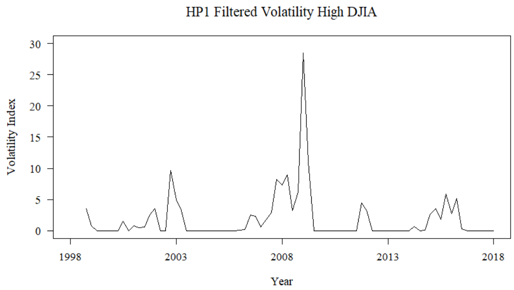


(A) High

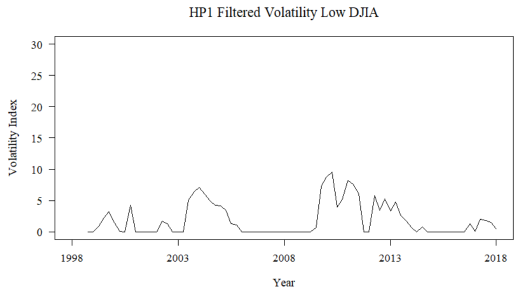


(B) Low

FIGURE B.22: The Quarterly High and Low Volatility Variables S&P

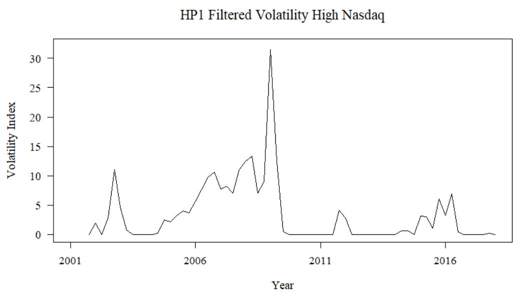


(A) High

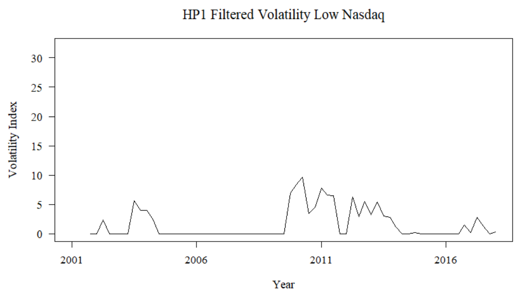


(B) Low

FIGURE B.23: The Quarterly High and Low Volatility Variables DJIA



(A) High



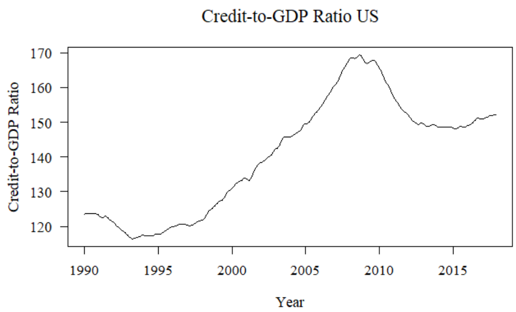
(B) Low

FIGURE B.24: The Quarterly High and Low Volatility Variables Nasdaq

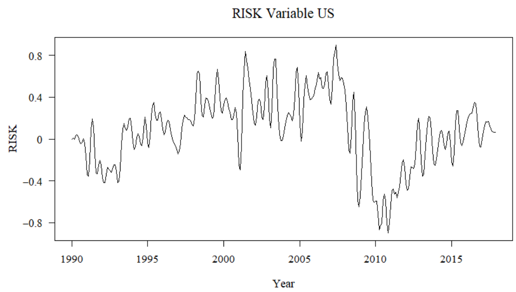
Appendix C

Risk Taking

C.1 Monthly Credit to GDP Ratio and Risk Variable

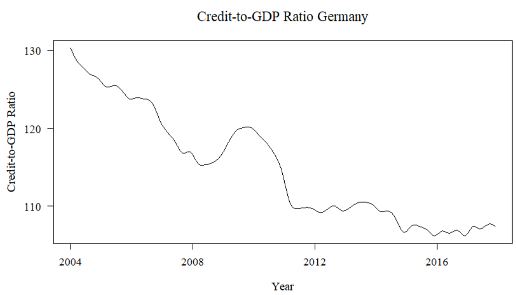


(A) Credit-to-GDP

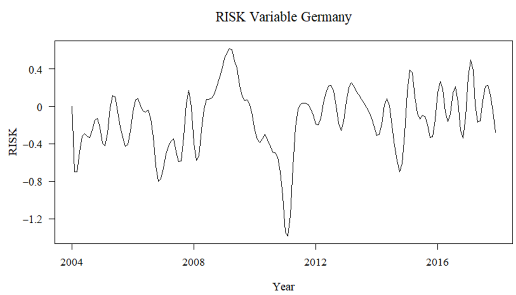


(B) RISK

Figure C.1: Credit-to-GDP Ratio and RISK Variable US

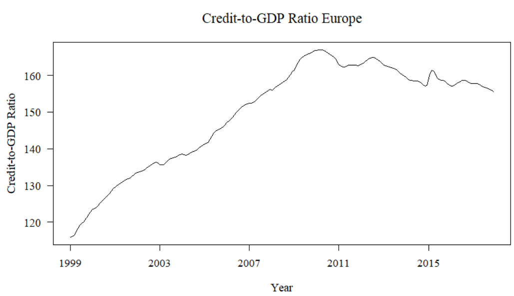


(A) Credit-to-GDP

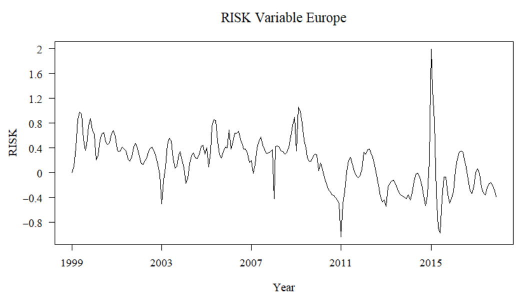


(B) RISK

Figure C.2: Credit-to-GDP Ratio and RISK Variable Germany

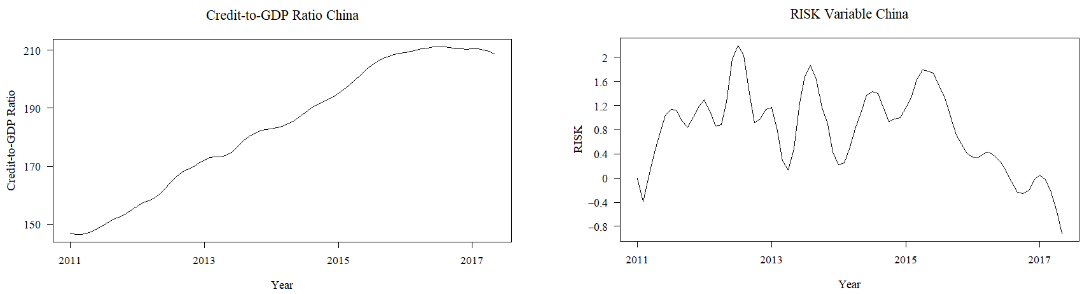


(A) Credit-to-GDP



(B) RISK

Figure C.3: Credit-to-GDP Ratio and RISK Variable Europe



(A) Credit-to-GDP (B) RISK

FIGURE C.4: Credit-to-GDP Ratio and RISK Variable China

C.2 Quarterly Risk

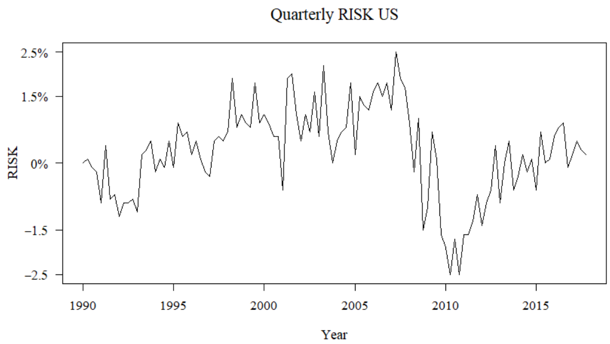


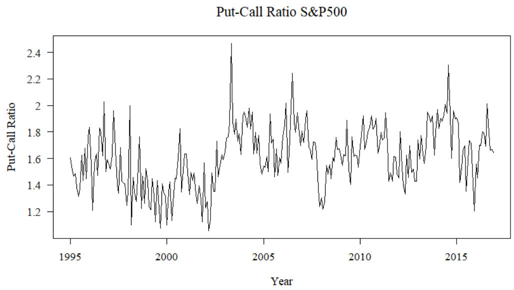
FIGURE C.5: Quarterly RISK Variable US

Appendix D

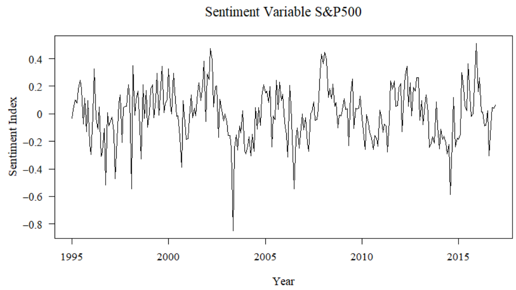
Investor Sentiment

D.1 Monthly Sentiment

D.1.1 Put-Call Ratio

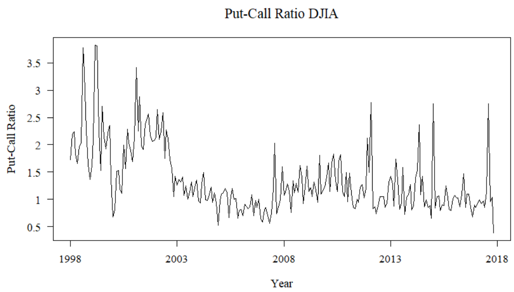


(A) Put-Call Ratio

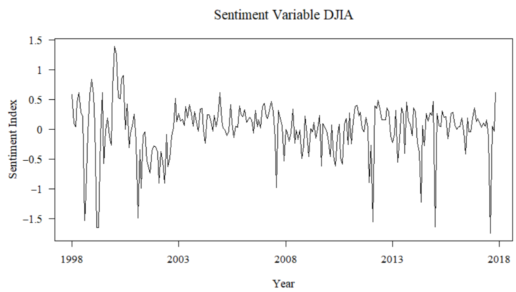


(B) Final Variable

Figure D.1: The Put-Call Ratio and Final Sentiment Variable S&P500

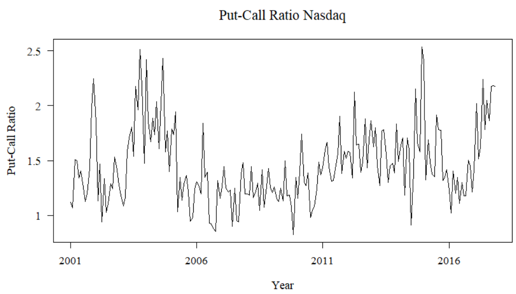


(A) Put-Call Ratio

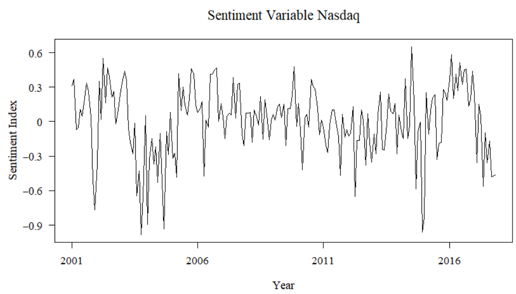


(B) Final Variable

Figure D.2: The Put-Call Ratio and Final Sentiment Variable DJIA

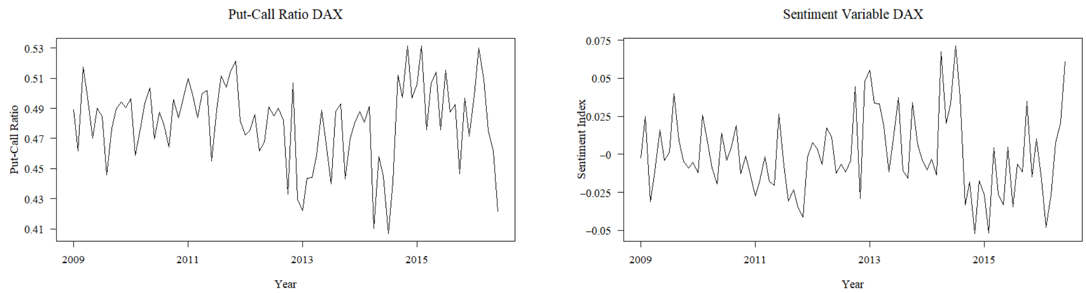


(A) Put-Call Ratio



(B) Final Variable

Figure D.3: The Put-Call Ratio and Final Sentiment Variable Nasdaq



(A) Put-Call Ratio (B) Final Variable

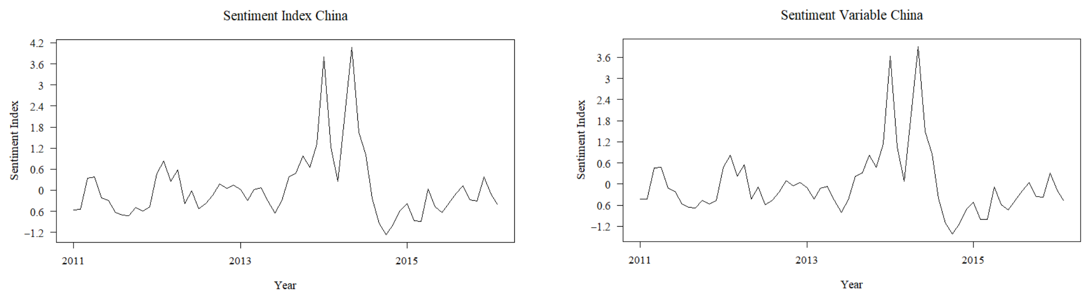
FIGURE D.4: The Put-Call Ratio and Final Sentiment Variable DAX

D.1.2 Sentix Sentiment Index

(A) Put-Call Ratio (B) Final Variable

FIGURE D.5: The Sentix and Final Sentiment Variable Europe

D.1.3 China Sentiment Index



(A) Put-Call Ratio (B) Final Variable

FIGURE D.6: The China Sentiment Index and Final Sentiment Variable China

D.2 Monthly Sentiment Two-Sided HP Filtered

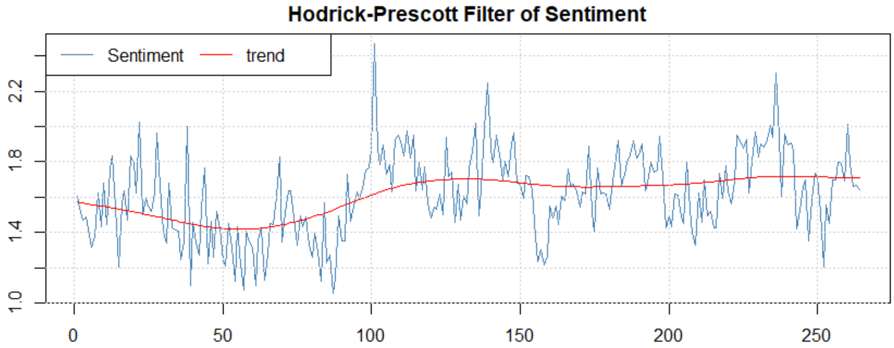


Figure D.7: HP2 Filter Sentiment S&P

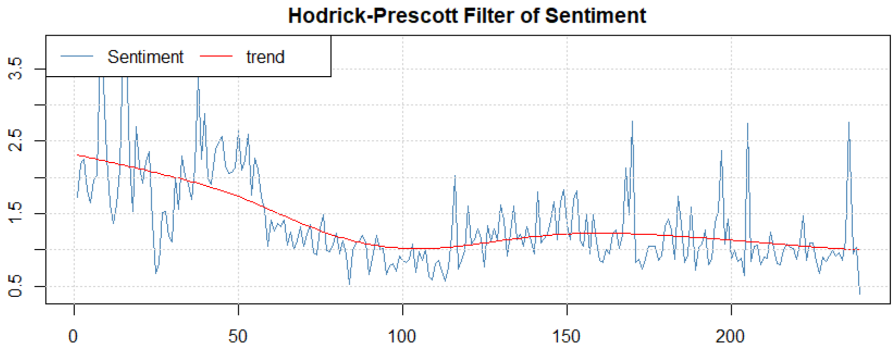


Figure D.8: HP2 Filter Sentiment DJIA

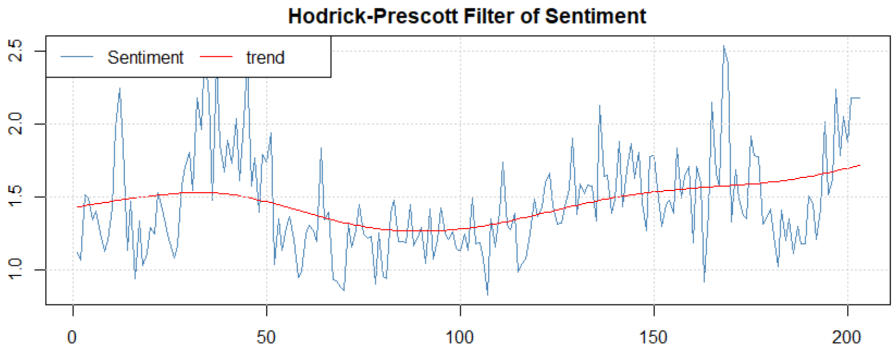


Figure D.9: HP2 Filter Sentiment Nasdaq

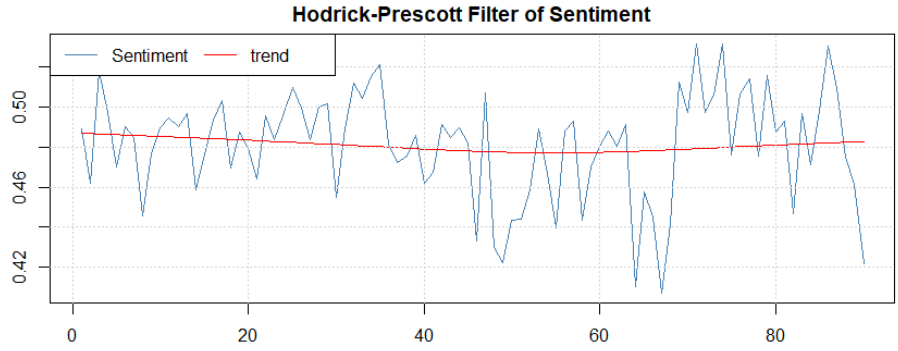


FIGURE D.10: HP2 Filter Sentiment DAX

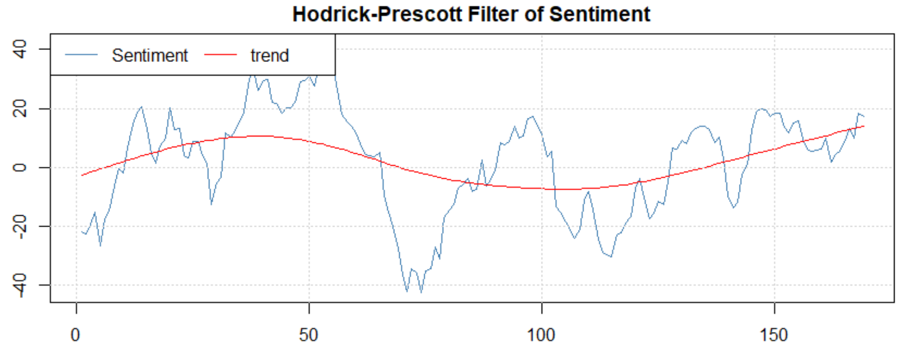


FIGURE D.11: HP2 Filter Sentiment Europe

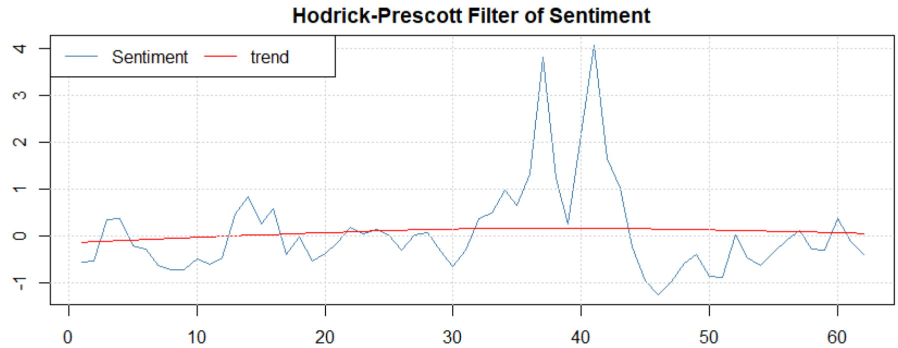
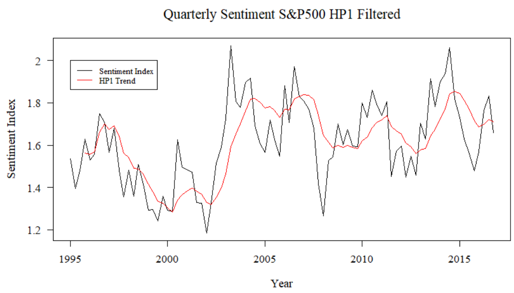
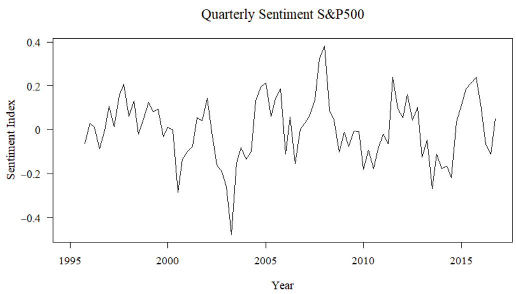


FIGURE D.12: HP2 Filter Sentiment FXI

D.3 Quarterly Sentiment

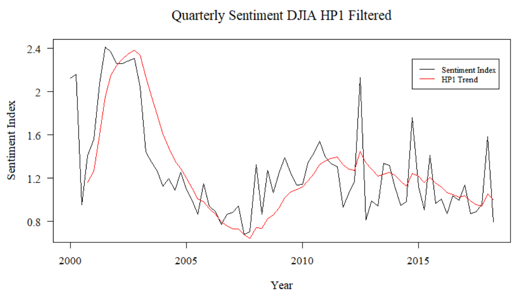


(A) HP1 Filter

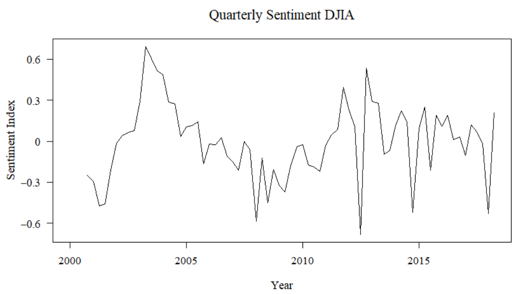


(B) Final Variable

FIGURE D.13: HP1 Filter and Final Sentiment Variable S&P

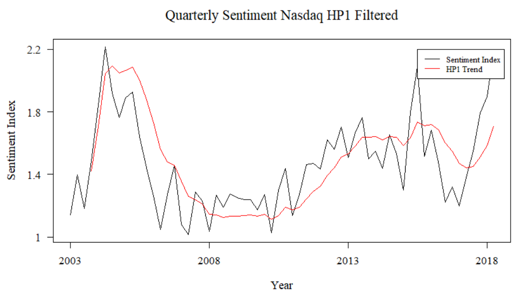


(A) HP1 Filter

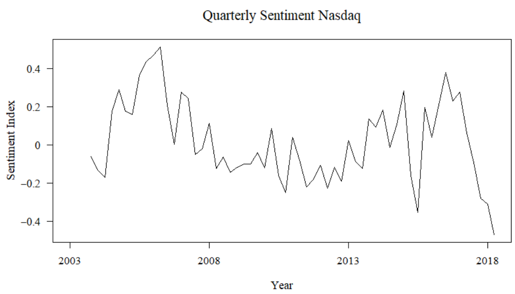


(B) Final Variable

FIGURE D.14: HP1 Filter and Final Sentiment Variable DJIA



(A) HP1 Filter



(B) Final Variable

FIGURE D.15: HP1 Filter and Final Sentiment Variable Nasdaq

Appendix E

Control Variables

E.1 GDP per Capita

E.1.1 Monthly Change GDP Per Capita

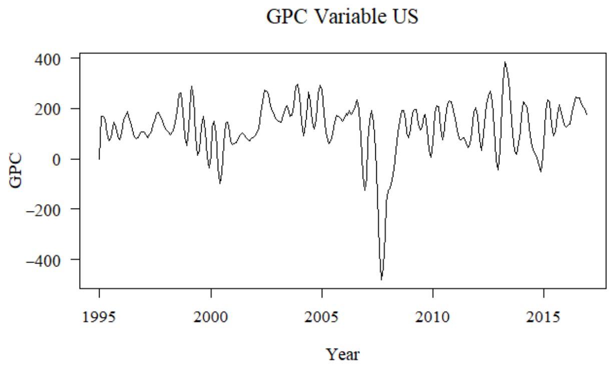


Figure E.1: Monthly Change in GDP Per Capita US

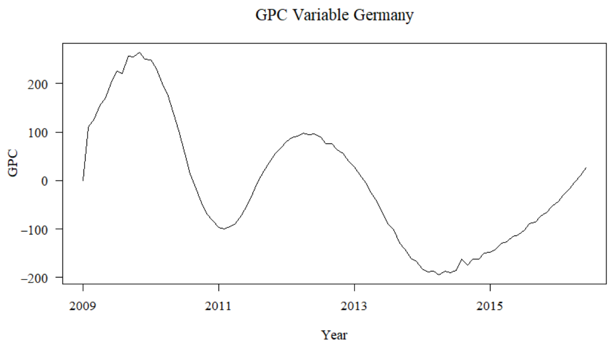


Figure E.2: Monthly Change in GDP Per Capita Germany

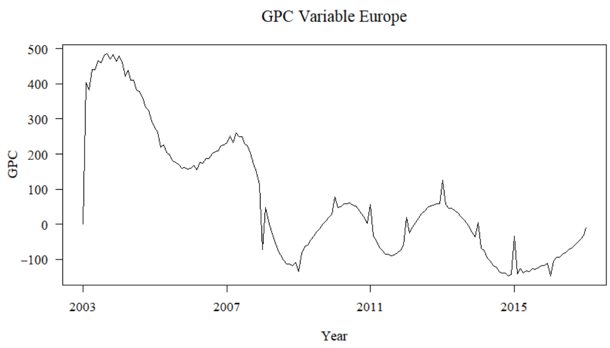


FIGURE E.3: Monthly Change in GDP Per Capita Europe

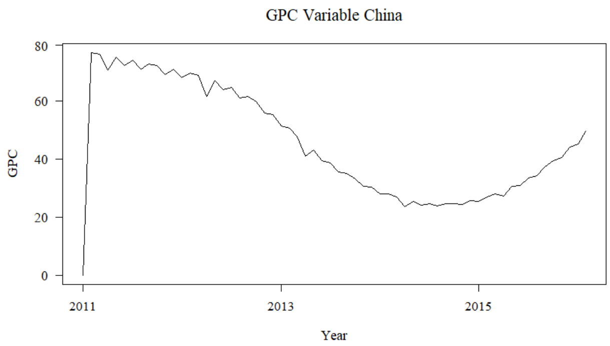


FIGURE E.4: Monthly Change in GDP Per Capita China

E.1.2 Quarterly Change GDP per Capita

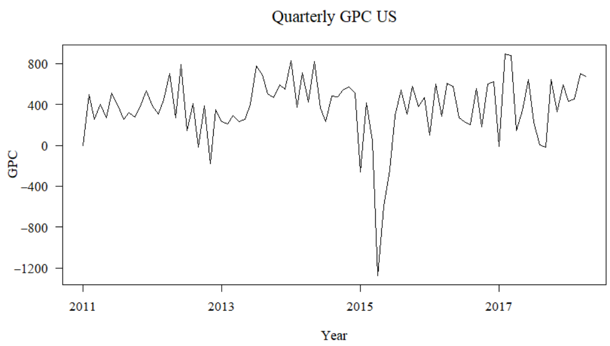


FIGURE E.5: Quarterly Change in GDP Per Capita US

E.2 Inflation

E.2.1 Monthly Inflation

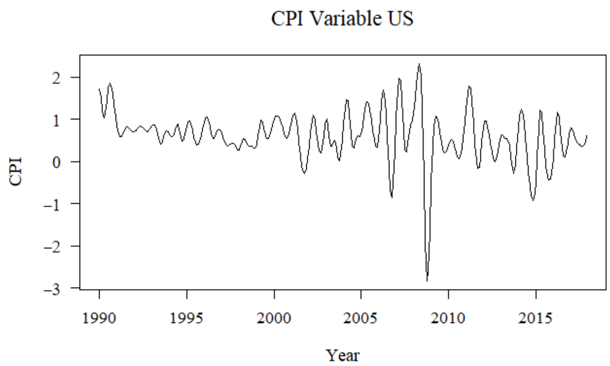


FIGURE E.6: Monthly Inflation US

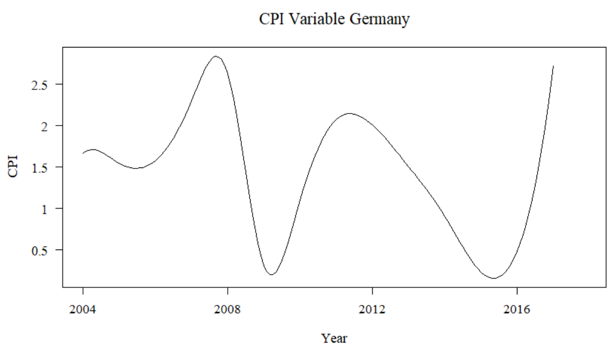


FIGURE E.7: Monthly Inflation Germany

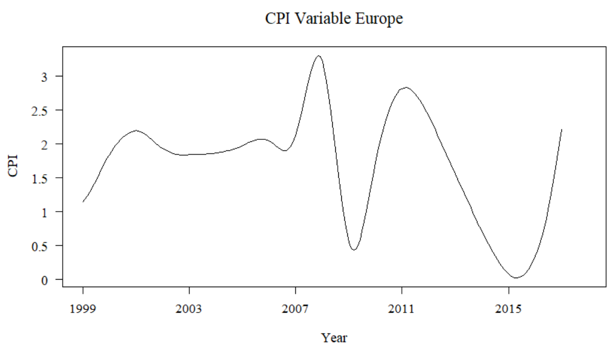


FIGURE E.8: Monthly Inflation Europe

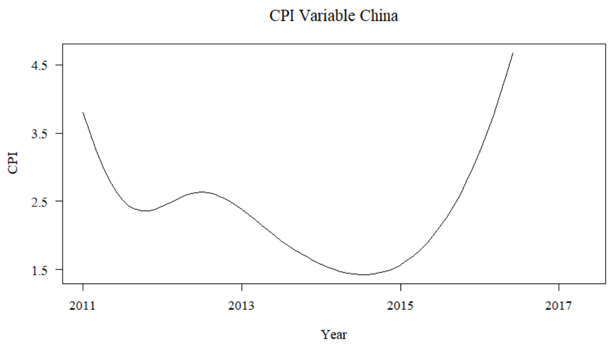


FIGURE E.9: Monthly Inflation China

E.2.2 Quarterly Inflation

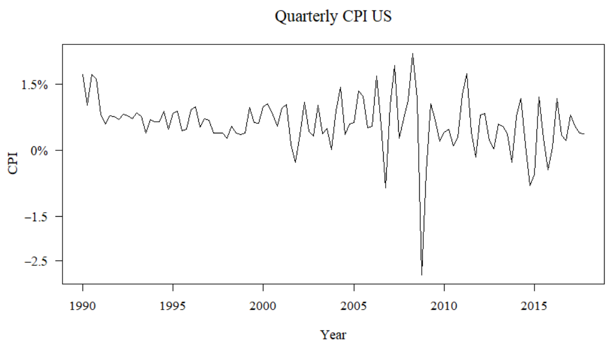


FIGURE E.10: Quarterly Inflation US

E.3 Government Debt

E.3.1 Monthly Change in Government Debt

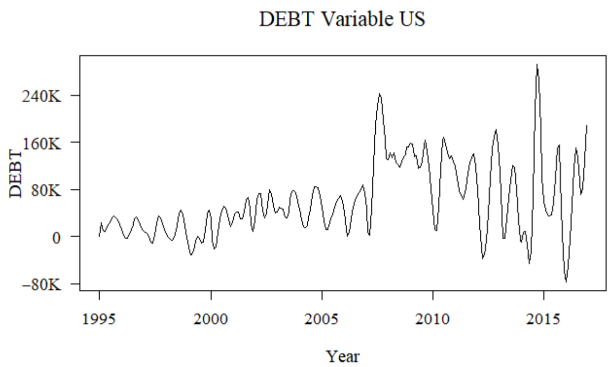


FIGURE E.11: Monthly Change in Government Debt US

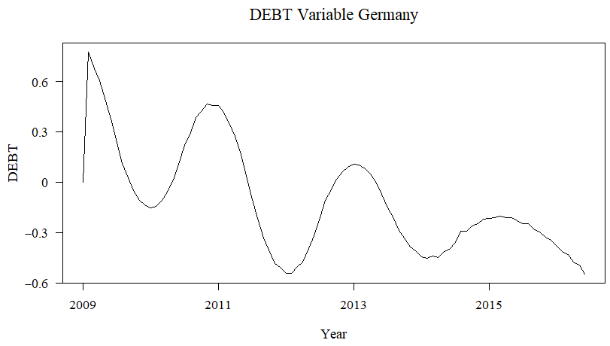


FIGURE E.12: Monthly Change in Government Debt Germany

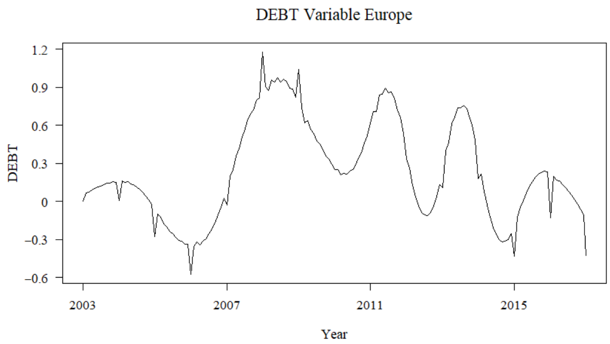


FIGURE E.13: Monthly Change in Government Debt Europe

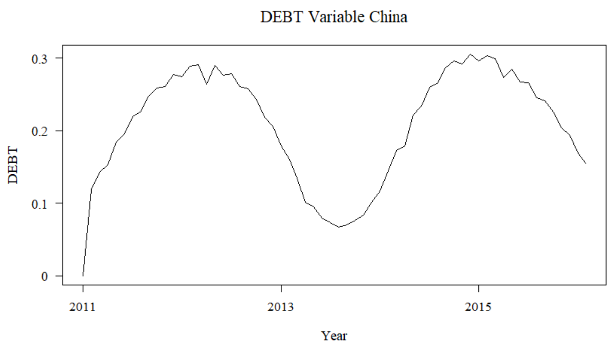


FIGURE E.14: Monthly Change in Government Debt China

E.3.2 Quarterly Change in Government Debt

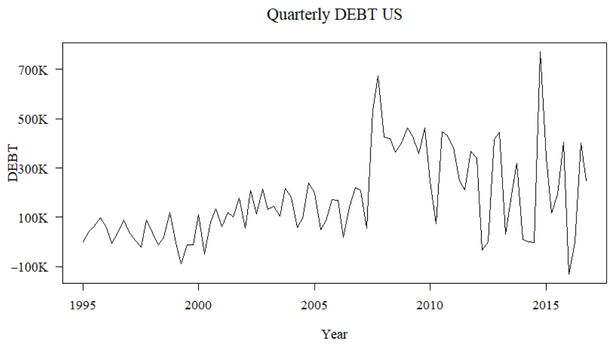


FIGURE E.15: Quarterly Change in Government Debt US

Appendix F

Hodrick and Prescott Filter

In order to be able to remove the cyclical component in time series that show trending behavior, a methodology is selected to do so. In this thesis, the Hodrick and Prescott Filter (or decomposition) is used. This is a mathematical tool often used in macroeconomics to create a smoothed time series that is more sensitive to long term fluctuations (Hodrick and Prescott, 1997).

The Hodrick-Prescott (HP) Filter recognizes the trend and divides the time series into two parts: the trend (g_t) and the cyclical (c_t) component (Equation F.1). By removing the trend component, a new series arises in which only the deviations from the trend remain.

The conventional HP Filter (regarded as the two-sided HP Filter) uses both forward and backward looking data to optimally determine the cyclical component. This approach is not always desirable since when one has the ambition to use the transformed time series for forecasting, future data is required. Therefore, an adaption of the two-sided HP filter is introduced: the one-sided HP Filter. This is done by running the standard HP filter recursively through time by using only data up to the time t such that only past and current data is used to determine the cyclical and trend components (Danielsson, Valenzuela, and Zer, 2018). According to Stock and Watson (1999), this is a good method in order to use the new time series for forecasting or prediction.

$$y_t = g_t + c_t \quad \text{for } t = 1, ..., T \tag{F.1}$$

Equation F.2 presents the HP filter formula. The first term in Equation F.2 represents the sum of the squared deviations which penalizes the cyclical component. The second term penalizes the variation in the growth rate by evaluating the sum of the squared differences of the trend. To find the correct trend component, Equation F.2 is solved using the according value for λ.

$$\min_g \left(\sum_{t=1}^{T-1} (y_t - g_t)^2 + \lambda \sum_{t=2}^{T-1} [(g_{t+1} - g_t) - (g_t - g_{t-1})]^2 \right) \tag{F.2}$$

where:

$$\begin{aligned} y_t &= \text{Time series value for t=1,2,...,T} \\ g_t &= \text{The trend component at time t} \\ c_t &= \text{The cyclical component at time t} \\ \lambda &= \text{The filter frequency} \end{aligned}$$

The level of smoothness in the final time series is depending on the frequency (λ) selected. This frequency should be chosen in a way that matches the time series intervals. According to Ravn and Uhlig (2002) a monthly time series should use $\lambda = 129,600$ and a quarterly time series $\lambda = 1,600$. The HP transformations in this thesis are executed using the frequencies as proposed by Ravn and Uhlig (2002).

Appendix G

HP Filter Robustness

G.1 Monthly Complete Models

Table G.1: HP Filter Robustness Monthly Models S&P and DJIA

	Dependent variable:					
	$CRASH_m$					
	S&P			DJIA		
	($\lambda = 100$)	($\lambda = 129600$)	($\lambda = 1M$)	($\lambda = 100$)	($\lambda = 129600$)	($\lambda = 1M$)
V^{low}_{m-0}	-0.695^{*}	-0.817^{**}	-1.559	-0.836^{**}	-1.334^{***}	-1.410^{***}
	(0.304)	(0.398)	(1.005)	(0.358)	(0.426)	(0.459)
V^{low}_{m-1}	0.674^{**}	0.215	0.163	1.096^{***}	0.917^{***}	0.881^{***}
	(0.253)	(0.242)	(0.328)	(0.281)	(0.291)	(0.299)
V^{low}_{m-2}	0.403^{**}	-0.065	-0.285			
	(0.270)	(0.231)	(0.316)			
V^{low}_{m-3}	0.719^{***}	0.305^{**}	0.282			
	(0.272)	(0.208)	(0.291)			
V^{high}_{m-0}	0.787^{**}	0.837^{**}	0.935^{**}	0.576^{***}	0.488^{**}	0.494^{***}
	(0.233)	(0.224)	(0.250)	(0.190)	(0.146)	(0.138)
V^{high}_{m-1}	-0.511^{**}	-0.430^{*}	-0.420^{*}	-0.234	-0.344^{**}	-0.371^{**}
	(0.217)	(0.167)	(0.170)	(0.155)	(0.127)	(0.122)
V^{high}_{m-2}	-0.086	-0.576^{**}	-0.772^{**}			
	(0.267)	(0.261)	(0.295)			
V^{high}_{m-3}	0.570^{***}	0.266^{*}	0.229			
	(0.197)	(0.148)	(0.154)			
$CRASH_{m-1}$	0.208	-0.971	-2.312^{*}	-1.224	-1.569^{**}	-1.681^{**}
	(1.183)	(1.290)	(1.500)	(1.069)	(1.084)	(1.086)
$RISK_{m-0}$	7.027^{***}	6.417^{**}	6.683^{**}	1.905	1.679	2.354
	(3.802)	(3.555)	(3.893)	(2.890)	(3.101)	(3.042)
$RISK_{m-1}$	-6.780^{***}	-6.212^{**}	-6.815^{**}	-1.918	-1.967	-2.581
	(3.450)	(3.240)	(3.573)	(2.665)	(2.834)	(2.846)
$SENT_{m-0}$	0.239	1.709	4.115	0.891	0.883	0.751
	(2.951)	(2.414)	(2.978)	(0.873)	(0.794)	(0.751)
$SENT_{m-1}$	-3.602^{**}	0.948	2.356	-0.042	-0.707	-0.844
	(3.371)	(2.394)	(2.798)	(0.813)	(0.737)	(0.713)
INF_{m-0}	3.152	0.714	2.037	2.319^{**}	0.817	0.762
	(2.203)	(1.836)	(2.370)	(1.375)	(1.254)	(1.262)
INF_{m-1}	-1.737	-0.092	-1.383	-2.326^{**}	-0.840	-0.846
	(2.236)	(1.913)	(2.438)	(1.375)	(1.295)	(1.306)
$DEBT_{m-0}$	0.00001	0.00001	0.00002	0.00002^{**}	0.00001^{*}	0.00001
	(0.00002)	(0.00002)	(0.00002)	(0.00001)	(0.00001)	(0.00001)
$DEBT_{m-1}$	-0.00003^{*}	-0.00002	-0.00003	-0.00002^{**}	-0.00002^{**}	-0.00002^{*}
	(0.00002)	(0.00002)	(0.00003)	(0.00001)	(0.00001)	(0.00001)
GPC_{m-0}	-0.0004	0.002	0.008	0.002	0.004	0.003
	(0.008)	(0.008)	(0.010)	(0.007)	(0.006)	(0.006)
GPC_{m-1}	-0.020^{***}	-0.018^{***}	-0.025^{***}	-0.012^{*}	-0.013^{*}	-0.013
	(0.008)	(0.008)	(0.010)	(0.007)	(0.007)	(0.007)
Constant	-5.348^{***}	-2.406^{**}	-1.987	-2.779^{***}	-1.880^{***}	-1.673^{**}
	(1.586)	(1.199)	(1.623)	(1.000)	(0.975)	(0.980)
Observations	263	263	263	238	238	238
McFadden R^2	0.670	0.633	0.692	0.483	0.514	0.519
Note:						$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

TABLE G.2: HP Filter Robustness Monthly Models Nasdaq and DAX

	Dependent variable:					
	$CRASH_m$					
	Nasdaq			DAX		
	($\lambda = 100$)	($\lambda = 129600$)	($\lambda = 1M$)	($\lambda = 100$)	($\lambda = 129600$)	($\lambda = 1M$)
V^{low}_{m-0}	−0.334**	−0.328**	−0.416***	−1.715	−0.553	−0.800
	(0.210)	(0.182)	(0.194)	(0.977)	(0.472)	(0.561)
V^{low}_{m-1}	0.725***	0.479***	0.369***	1.748**	1.133**	1.074
	(0.208)	(0.174)	(0.181)	(0.761)	(0.616)	(0.616)
V^{high}_{m-0}	0.598***	0.475***	0.432***	0.243	0.509***	0.453**
	(0.172)	(0.131)	(0.110)	(0.289)	(0.246)	(0.240)
V^{high}_{m-1}	−0.213	−0.377***	−0.347***	0.149	−0.413	−0.413
	(0.132)	(0.112)	(0.102)	(0.467)	(0.304)	(0.308)
$CRASH_{m-1}$	−0.635	−0.148	−0.423	−0.093	−0.748	−0.532
	(0.905)	(0.783)	(0.773)	(1.542)	(1.641)	(1.585)
$RISK_{m-0}$	−4.949**	−4.661**	−4.000	−3.984	−1.706	−1.940
	(2.489)	(2.493)	(2.418)	(5.540)	(4.617)	(4.601)
$RISK_{m-1}$	4.982**	4.480*	4.162*	−1.165	−3.290	−2.896
	(2.312)	(2.314)	(2.234)	(4.863)	(3.801)	(3.857)
$SENT_{m-0}$	−0.483	−0.471	−0.607	−18.717	−21.194*	−16.837
	(1.167)	(0.926)	(0.924)	(36.024)	(33.116)	(28.832)
$SENT_{m-1}$	0.830	0.868	0.479	−13.211	−37.855***	−31.039**
	(1.305)	(1.129)	(1.113)	(26.355)	(26.555)	(22.914)
INF_{m-0}	1.535*	0.358	0.436	4.719	4.608	1.615
	(1.046)	(1.033)	(0.993)	(9.967)	(12.489)	(11.111)
INF_{m-1}	−0.795	0.180	0.133	−6.735	−5.799	−2.745
	(1.026)	(1.077)	(1.016)	(10.406)	(12.349)	(11.020)
$DEBT_{m-0}$	0.00001	0.00001	0.00001	−7.013	−10.661**	−9.568**
	(0.00001)	(0.00001)	(0.00001)	(6.655)	(8.042)	(7.625)
$DEBT_{m-1}$	−0.00001	−0.00001	−0.00001	9.212	13.402**	12.026**
	(0.00001)	(0.00001)	(0.00001)	(7.484)	(9.396)	(8.582)
GPC_{m-0}	−0.0002	0.001	0.001	0.056	0.104**	0.089**
	(0.006)	(0.005)	(0.005)	(0.049)	(0.056)	(0.058)
GPC_{m-1}	−0.010*	−0.012*	−0.010*	−0.068	−0.118**	−0.100**
	(0.006)	(0.006)	(0.006)	(0.053)	(0.065)	(0.065)
Constant	−2.547***	−1.956**	−1.473*	−4.081***	−5.665***	−4.546***
	(0.916)	(0.918)	(0.805)	(2.012)	(2.700)	(2.405)
Observations	202	202	202	89	89	89
McFadden R^2	0.457	0.419	0.411	0.564	0.538	0.525
Note:					*p<0.1; **p<0.05; ***p<0.01	

TABLE G.3: HP Filter Robustness Monthly Models EURO and FXI

	Dependent variable:					
	$CRASH_m$					
	EURO			FXI		
	($\lambda = 100$)	($\lambda = 129600$)	($\lambda = 1M$)	($\lambda = 100$)	($\lambda = 129600$)	($\lambda = 1M$)
V^{low}_{m-0}	−0.981***	−0.405**	−0.427**	−0.181	−0.532	−0.569
	(0.405)	(0.236)	(0.254)	(0.525)	(0.349)	(0.363)
V^{low}_{m-1}	1.107***	0.343*	0.236	0.668**	0.784**	0.798**
	(0.411)	(0.219)	(0.219)	(0.443)	(0.394)	(0.399)
V^{high}_{m-0}	0.525***	0.720***	0.739***	0.262	0.221	0.218
	(0.205)	(0.256)	(0.255)	(0.394)	(0.182)	(0.181)
V^{high}_{m-1}	−0.277*	−0.621**	−0.666***	−1.063*	−0.236	−0.223
	(0.202)	(0.256)	(0.257)	(0.830)	(0.239)	(0.227)
$CRASH_{m-1}$	−0.223	−0.877	−0.985	0.297	−1.106	−1.154
	(1.038)	(1.115)	(1.112)	(1.476)	(1.393)	(1.402)
$RISK_{m-0}$	−0.012	−0.253	−0.125	2.713*	2.260	2.261
	(1.387)	(1.352)	(1.330)	(1.883)	(1.693)	(1.698)
$RISK_{m-1}$	1.240	1.243	1.161	−2.706**	−1.990	−1.969
	(1.292)	(1.333)	(1.324)	(2.137)	(1.767)	(1.778)
$SENT_{m-0}$	−0.215**	−0.217***	−0.239**	−1.633**	−1.707**	−1.723**
	(0.087)	(0.079)	(0.083)	(0.907)	(0.842)	(0.850)
$SENT_{m-1}$	0.202**	0.203**	0.221**	1.950**	1.290**	1.276**
	(0.093)	(0.080)	(0.083)	(0.901)	(0.596)	(0.593)
INF_{m-0}	−1.532	0.203	−0.225	−13.718	−14.439	−14.726
	(3.772)	(4.016)	(4.002)	(14.922)	(15.705)	(15.923)
INF_{m-1}	0.909	−0.751	−0.399	17.962	19.518	19.915
	(3.800)	(4.026)	(4.026)	(18.100)	(18.950)	(19.189)
$DEBT_{m-0}$	0.268	1.031	1.591	64.802*	92.258**	95.695**
	(3.268)	(3.323)	(3.437)	(53.353)	(62.108)	(62.347)
$DEBT_{m-1}$	2.491	1.303	0.475	−44.934	−74.058*	−77.592**
	(3.355)	(3.272)	(3.355)	(47.319)	(54.791)	(54.851)
GPC_{m-0}	0.005	0.003	0.002	−0.893**	−0.919***	−0.941***
	(0.008)	(0.009)	(0.009)	(0.472)	(0.450)	(0.456)
GPC_{m-1}	−0.009**	−0.004	−0.002	0.758**	0.750***	0.771***
	(0.008)	(0.008)	(0.009)	(0.401)	(0.363)	(0.368)
Constant	−3.776***	−2.900***	−2.402**	−9.791*	−9.957**	−10.148**
	(0.880)	(0.909)	(0.895)	(5.123)	(4.830)	(4.890)
Observations	168	168	168	61	61	61
McFadden R^2	0.575	0.564	0.577	0.472	0.399	0.401

Note: *p<0.1; **p<0.05; ***p<0.01

G.2 Quarterly Complete Models

Table G.4: HP Filter Robustness Quarterly Models S&P and DJIA

	Dependent variable:					
	$CRASH_q$					
	S&P			DJIA		
	($\lambda = 100$)	($\lambda = 1600$)	($\lambda = 100K$)	($\lambda = 100$)	($\lambda = 1600$)	($\lambda = 100K$)
V^{low}_{q-1}	0.248	0.415**	−0.087	−0.020	0.140	−0.247
	(0.151)	(0.169)	(0.128)	(0.173)	(0.211)	(0.238)
V^{high}_{q-1}	0.133	0.191*	0.005	0.216	0.272	0.125
	(0.170)	(0.118)	(0.089)	(0.173)	(0.146)	(0.107)
$CRASH_{q-1}$	0.514	0.580	0.106	−0.764	−1.106	−1.339
	(1.003)	(1.011)	(0.924)	(1.140)	(1.412)	(1.313)
$RISK_{q-1}$	−0.111	0.341	−0.146	−0.133	−0.108	−0.209
	(0.337)	(0.383)	(0.339)	(0.368)	(0.494)	(0.433)
$SENT_{q-1}$	2.009	1.176	0.137	−0.505	−1.360	−1.375
	(3.163)	(2.522)	(2.244)	(1.627)	(1.356)	(1.288)
INF_{q-1}	0.996**	1.079***	0.915**	1.132**	1.237**	0.857*
	(0.593)	(0.597)	(0.559)	(0.653)	(0.670)	(0.600)
$DEBT_{q-1}$	−0.00000	−0.00000	−0.00000	−0.00000	−0.00000	−0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
GPC_{q-1}	−0.002**	−0.002	−0.002*	−0.003***	−0.002*	−0.002
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Constant	−1.698**	−2.905***	−0.869	−1.530*	−2.324**	−1.051
	(0.962)	(1.211)	(1.054)	(1.092)	(1.291)	(1.234)
Observations	84	84	84	70	70	70
McFadden R^2	0.158	0.217	0.128	0.231	0.289	0.279
Note:						*p<0.1; **p<0.05; ***p<0.01

Table G.5: HP Filter Robustness Quarterly Model Nasdaq

	Dependent variable:		
	$CRASH_q$		
	Nasdaq		
	($\lambda = 100$)	($\lambda = 1600$)	($\lambda = 100K$)
V^{low}_{q-1}	−0.003	0.283**	−140.502
	(0.148)	(0.216)	(24,435.000)
V^{high}_{q-1}	0.233	0.242*	0.016
	(0.270)	(0.152)	(0.091)
$CRASH_{q-1}$	−1.857	−1.741	−1.461
	(1.522)	(1.616)	(1.348)
$RISK_{q-1}$	−0.026	0.138	0.126
	(0.363)	(0.579)	(0.460)
$SENT_{q-1}$	0.353	0.555	−1.103
	(2.662)	(2.588)	(2.251)
INF_{q-1}	3.109	2.798*	2.321**
	(1.306)	(1.183)	(1.066)
$DEBT_{q-1}$	0.00001	0.00001	0.00001
	(0.00001)	(0.00000)	(0.00000)
GPC_{q-1}	0.003	0.003	0.002
	(0.002)	(0.002)	(0.002)
Constant	−7.386	−8.388**	−5.808**
	(3.267)	(3.109)	(2.582)
Observations	58	58	58
McFadden R^2	0.294	0.353	0.314
Note:			*p<0.1; **p<0.05; ***p<0.01