ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Bachelor Thesis (IBEB)

Forecasting stock market volatility using online search queries

Mirnesa Ibišević
408590

Supervisor: Esad Smajlbegovic
Second assessor: Aleksandar Andonov

July 26, 2017

Abstract

This paper investigates the explanatory and predictive power of the Google Trends search volume index (SVI) and the number of Wikipedia page views as proxies for investor attention. Using data from the Amsterdam Exchange (AEX) index, I fit several models to estimate and forecast historical and implied volatility and subsequently test the implications of my forecasts in a real-life trading example. I find that the SVI has significant predictive power in explaining the variations of historical volatility, but not for implied volatility. Furthermore, I find (weak) evidence of Wikipedia page views as a predictor for implied volatility, but not for historical volatility. As for historical volatility, combining both metrics yields slight improvements of the volatility forecasts, but this is mainly driven by the significance of the SVI. Nevertheless, the results provide no evidence for the predictive power of the SVI for historical out-of-sample volatility forecasts of the 25 individuals stocks that are part of the AEX index. Moreover, the investment portfolios based on the volatility forecasts of the SVI forecasting model yield no higher returns compared to the benchmark.

Keywords: investor attention, Google Trends, Wikipedia, volatility forecasts

Table of Contents

1. Introduction

Traditional asset pricing models assume that all available information in the market is incorporated into the stock buying process. This also means that investors have a complete overview of all the available stocks and can distribute their attention accordingly. In reality, this attention is limited and restricted, as pointed out by Kahneman (2011) among others, as investors have limited resources available to evaluate all the available stocks. Recent studies find that these differences in the investors' attention have a significant influence on asset prices and other financial measures. This is the scope of this paper in which I explicitly focus on the predictive power of Google Trends and Wikipedia as proxies for investor attention. I use these proxies to forecast stock market volatility and subsequently test the implications of my forecasts in a real trading example.

Investor attention can be measured in different ways. A common indicator is the news exposure that a stock receives, measured in both amount and quality of sources. Literature on the influence of news on stock prices has a long track record (see e.g. Pearce & Roley, 1984; Li, Xie, Chen, Wang, & Deng, 2014; Boudoukh, Feldman, Kogan, & Richardson, 2013; Nikkinen & Sahlström, 2001). Lumsdaine (2009), for example, finds that greater news readership is associated with higher volatility of returns. In addition, McQueen & Roley (2015) examine the effect of macroeconomic news on stock prices during different stages of the business cycle. Their findings suggest that news concerning high economic activity during economic booms leads to lower stock prices.

Other metrics of investor attention include extreme returns and unusual trading volume (Barber & Odean, 2008). The authors argue that stocks with abnormal trading volume and one-day returns will more likely be noticed by investors. These extreme returns also become available in the news, driving the attention of other investors as well. They conclude that investors tend to buy the stocks that caught their attention at first. Kim & Kim (2014) consider the internet message postings on the Yahoo! Finance message board as proxy for attention, but find no significant results for stock return forecasts. Furthermore, Vozlyublennaia (2014) considers search probabilities as another measure of attention, finding a significant effect of investor attention on index returns.

An important caveat of these indirect measures is that they may not give accurate estimations of true investor attention as they do not reveal whether the investors are truly paying more attention to the stocks or whether it is just general increased interest. As pointed out by Barber & Odean (2008) as a limitation of their proxies, stocks may also be present in the news because of general press releases and not because of increased investor attention. Moreover, returns can be influenced by several other factors, including liquidity crises. Many authors have therefore turned to search queries as direct proxies; if investors increasingly search for stocks online, they obviously pay more attention to them, as opposed to simply opening a news site and clicking on the headline on the front page.

Many authors consider the Google Trends search volume index (SVI) as a proxy for attention since Google has the biggest market share in the global market of search engines[1], hence making it a representative measure of internet search behavior. The Google Trends SVI has already proven to improve forecasts in a variety of fields, such as tourist inflows (Park, Lee, & Song, 2017), youth unemployment (Fondeur & Karamé, 2013), automobile sales and consumer confidence (Choi & Varian, 2012), and trading behavior (Preis, Moat, & Stanley, 2013). A large body of the literature investigates the effect of the SVI in relation to financial measures. Bank, Larch, and Peter (2011) find that the SVI reflects overall firm recognizition and stock market investor attention. Their results show that the SVI is positively correlated with trading volume, stock liquidity and future returns of German stocks. Aouadi, Arouri, and Teulon (2013) find similar results for the French stock market, even after controlling for the financial crisis effect. Moreover, Da, Engelberg, and Gao (2011) conclude that a higher SVI is associated with a positive price pressure in the following weeks, which is reversed within the year. This supports the attention theory of Barber & Odean (2008).

Several researchers also investigate the effect of Google search queries on market volatility and examine its predictive power. Forecasting market volatility is relevant for various financial investment decisions. Volatility reflects the amount of risk that is associated with a certain stock or index. Hence, in order to value a certain financial instrument, an investor should make some predictions regarding the future volatility (Claessen & Mittnik, 2002); it is therefore desirable to predict volatility as accurately as possible. Because higher investor attention may increase the volatility in the market (since then investors are able to incorporate

[1] Source: Statista (https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/)

the information that is available in the market), it may be a useful predictor in volatility forecasts. Vlastakis & Markellos (2012) study information demand and supply for stocks traded on NYSE and NASDAQ and find that information demand at the market level is positively correlated to historical and implied volatility. Vozlyublennaia (2014) suggest that investor attention can predict index returns and volatility. Moreover, Dimpfl & Jank (2016) and Hamid & Heiden (2015) consider both in-sample and out-of-sample forecasts using daily and weekly Google search data, respectively. They find that search queries can help improve volatility forecasts, especially during periods of high volatility.

Investors are concerned with volatility forecasts in order to build their optimal portfolios. When constructing an investment portfolio, several methods can be used to determine the weight of each stock in the portfolio based on expected return and risk. Some examples of these allocation strategies are naïve diversification, mean-variance optimization and the risk parity approach. These investment strategies have been widely discussed in academic literature in order to discuss what strategy may bring the highest portfolio return. The risk parity approach differs from other strategies because it is does not focus on the optimal allocation of capital, but the optimal allocation of risk instead. Chaves, Hsu, Li, & Shakernia, (2011) compare risk parity portfolios to various asset allocation strategies and find that the risk parity portfolio outperform minimum variance and mean-variance efficient portfolios. In addition, it provides a better allocation of risk compared to the equal weighting strategy. Clarke, Silva, & Thorley (2013) also construct various risk parity portfolios and report high Sharpe ratios. In this paper, I will generate volatility forecasts for different stocks using two different forecasting models, one of them using the SVI as a predictor of volatility. Subsequently, I will use the risk parity approach (risk measured by volatility) to construct two investment portfolios to investigate the implications of these volatility forecasts in a real trading example.

Another direct measure of investor attention is the number of Wikipedia page views. Wikipedia is the largest competitor in the encyclopedia market[2], providing extensive information about a variety of subjects, including stocks and indices. Considering the fact that Wikipedia may serve the same purpose as Google web searches of providing the investors with information about the index, it may also be a valuable measure of investor

[2] Source: Econsultancy (https://econsultancy.com/blog/3185-wikipedia-has-97-of-the-encyclopedia-market-online/)

attention. The number of Wikipedia page views as proxy for investor attention has not been widely discussed in academic literature yet, however. Ungeheuer (2017) considers hourly Wikipedia page views of various firms to investigate the effect of stock returns on investor attention. His findings suggest that stocks that are ranked as daily winners and losers have large increases in investor attention.

This paper builds on existing literature on the influence of investor attention by investigating the predictive power of both the Google search volume index and the Wikipedia page views in forecasting historical and implied volatility. I apply my volatility forecast models on the Amsterdam Exchange (AEX) index. Recent studies mainly focus on the S&P 500 or Dow Jones index using daily or weekly data. Although European and American financial systems are closely linked, differences between the stock markets still exist, so it remains of interest to investigate volatility forecasts of European indices. In addition, I will perform an in-depth forecasting analysis of the individual stocks that are part of the AEX index and test the implications of those forecasts in an investment setting using the volatility (risk) parity approach. The research question boils down to:

"To what extent do the Google search volume index and the Wikipedia page views contain explanatory and predictive power for stock market volatility in the Netherlands?"

The results show that the SVI does have significant predictive power for historical volatility forecasts as opposed to the benchmark model, which only includes lags of the volatility and the control variables. On the other hand, the SVI does not improve the forecasting performance of implied volatility. Furthermore, I find opposite results for the Wikipedia page views. In this case, the search query does not provide better forecasts of historical volatility, but the forecasting models of implied volatility seem to perform slightly better compared to the benchmark model. Significant combined predictive power of both search queries is found in the case of historical volatility, but this result is mainly driven by the significance of the SVI. Nevertheless, the results provide no evidence for the predictive power of the SVI for historical out-of-sample volatility forecasts of the 25 individuals stocks that are part of the AEX index. Moreover, the investment portfolios based on the volatility forecasts of the SVI forecasting model yield no higher returns compared to the benchmark.

The remainder of this paper is organized as follows. The following section provides a review of the literature on investor attention and the significance of using search queries in relation to financial measures. Section 3 describes the data that I used for this analysis. I explain the statistical methods in Section 4. The results are discussed in Section 5. Finally, Section 6 concludes and provides suggestions for further research.

2. Literature Review

Previous literature has examined the relationship between different proxies of investor attention and market volatility and has found significant positive correlations. I investigate the correlation between the Google SVI and the number of Wikipedia page views on one hand, and the historical and implied volatility on the other hand, using different control variables. In addition, various authors conclude that Google searches have significicant predictive power in stock market volatility forecasts. I will add to this existing literature by performing in-sample and out-of-sample forecasts of volatility using models including both the Google search volume index and Wikipedia page views. Hence, I will be able to compare the forecasting performances of the Google SVI with the Wikipedia data, and to examine the combined predictive power of these two search queries as well.

Much empirical literature focuses on investor attention and financial performance measures, using different measures of attention. Barber & Odean (2008) investigate the trading behavior of investors with attention-grabbing stocks. They use three different proxies for investor attention: news, extreme returns and unusual trading volume. They find that individual investors tend to be net buyers of stocks that had higher attention levels.

However, these proxies are more indirect measures of investor attention; if the stocks are being reported in the news, it does not necessarily imply that investors are also paying more attention to them. Recent studies therefore consider direct measures of investor attention using search queries. Search queries are a “revealed attention measure” (Da et al., 2011). If an investor is searching for a stock online, he is certainly paying attention to it. Chen, Liu, & Tang (2016) look at the Baidu index in relation to scheduled macroeconomic announcements on the Chinese Stock Index Futures market. The Baidu index was launched by Baidu, Inc., which is the largest search engine of China. The authors find that investor attention has an impact on future market reactions to Consumer Price Index (CPI) announcements. Fan, Yuan, Zhuang & Jin (2016) consider the Baidu index of the Shanghai Stock Exchange 50

Index component stocks. Their results suggest that abnormal investor attention is positively correlated with trading volume and volatility.

Similarly, most authors consider the search volume index provided by Google Trends[3] to proxy for investor attention, since Google is the dominant player in the market of search engines, making it a representative measure of internet search behavior. Several studies investigate the effect of Google search data on the underpricing and long-term underperformance of initial public offerings (IPOs). Da et al. (2011) find that Google search data are an accurate measure for individual and retail investors in particular. In addition, they support the attention hypothesis of Barber & Odean (2008). They argue that this increase in attention should lead to temporary positive price pressure. Within this framework, higher search volume indices result in higher stock prices in the short run and price reversals in the long run. The authors provide evidence that investor attention has an effect on IPO returns in the short run and on the return reversal following the initial price pressure. Vakrman & Kristoufek (2015) show the same results using a different sample. Moreover, using data from IPOs in the United States, Colaco, De Cesari, & Hegde (2016) suggest that higher Google search volumes are identified with higher initial valuations, as measured by different market multiples.

A large body of literature examines the relationship between Google searches and volatility. Vlastakis & Markellos (2012) use the Google search index as a proxy for information demand. They conclude that this is significantly correlated to historical and implied volatility. Andrei & Hasler (2014) incorporate both investor attention and uncertainty to investigate the effect on stock return variance and risk premia. They find this relationship to be significantly positive and provide an economic explanation for this result. Given the fact that investors pay more attention to the stocks, they are able to immediately incorporate new information into the stock prices, which leads to higher return volatility. Because of the higher volatility, investors will demand a higher return to compensate for this increased risk. The results from previous literature give enough reason to believe that Google searches have a strong positive correlation with market volatility. Hence, this will be tested in the first hypothesis:

Hypothesis 1a: *The Google search volume index of the AEX is positively correlated with the market volatility in the Netherlands.*

[3] https://trends.google.nl/trends/

In addition to the Google search volume index, I will propose a novel measure of investor attention based on the number of Wikipedia page views. A few academic articles have used Wikipedia page views to proxy for attention. Kristoufek (2013) considers both Google Trends and Wikipedia search queries as proxies for investor attention in relation to the prices of the BitCoin currency. He finds a significant positive correlation and a dynamic relationship between the variables. Ungeheuer (2017) uses hourly Wikipedia firm page views to measure investor attention to study the effect of stock returns. His findings suggest that stocks that are ranked as daily winners and losers have large increases in investor attention. Other stocks that experience extreme returns do not show the same increase.

In this paper, I will investigate the effect of the weekly Wikipedia page views on stock market volatility. Since Wikipedia serves the same purpose as the Google SVI of providing the investors with information about the stock index, it is plausible to believe that the number of page views may also be positively correlated with the market volatility of the stock index, which will be investigated in the following hypothesis:

Hypothesis 1b: *The number of Wikipedia page views of the AEX is positively correlated with the market volatility in the Netherlands.*

Moreover, recent studies have investigated the market volatility forecasting performances of the Google search volume index. Smith (2012) explores the market of seven different currencies and concludes that the number of Google searches has significant predictive power beyond the GARCH model. Dimpfl & Jank (2016) investigate four stock market indices' realised volatility and the Google search volume indices. They find that search queries have additional predictive power when performing in-sample and out-of-sample volatility forecasts. Hamid & Heiden (2015) provide similar results. These past findings suggest that Google searches have predictive power in stock market volatility, leading to the second hypothesis:

Hypothesis 2: *The Google search volume index of the AEX has predictive power for the stock market volatility in the Netherlands.*

Given the fact that the number of Wikipedia page views also serve as a valid measure of investor attention, there is good reason to believe that the number of page views may also have predictive power in forecasting market volatility, which will be tested in the third hypothesis:

Hypothesis 3: *The number of Wikipedia page views of the AEX index has predictive power for the stock market volatility in the Netherlands.*

Recent literature remains ambiguous about the additional information of implied volatility in producing volatility forecasts. Canina & Figlewski (1993) argue that implied volatility has no correlation with future volatility, providing poor forecasts of consecutive realized volatility. Similarly, Bentes (2015) finds that GARCH forecasted volatility outperforms implied volatility models. On the other hand, Christensen & Prabhala (1998) show that implied volatility provides better forecasts of future volatilities compared to historical volatility. Another study by Pilbeam & Langeland (2015) shows the same results. Because of the ambiguity regarding the informational content of implied volatility, I will use both implied and historical volatility in this research. Generally, it is believed that implied volatility contains more information about the market, because it was calculcated using pricing models that incorporate market information. It may therefore give more accurate forecasts of volatility.

3. Data

3.1 Investor attention proxies

This paper focuses on investor attention for the Amsterdam Exchange (AEX) index[4]. The Google SVI of the "AEX" search term and the number of Wikipedia page views of the AEX index are used as proxies for individual investor attention. Figure 1 plots the SVI and the number of Wikipedia page views over time. Both search queries show high peaks around similar dates. I marked some of those similar peaks in the graph in order to illustrate that both search queries are able to capture investor attention. The first two peaks take place at the end of 2008 in the weeks following the collapse of the investment bank Lehman Brothers, which can be recognized as the start of the financial crisis. Similarly, the search queries show an outlier in August 2011. On the 8th of August 2011, also known as Black Monday, stock markets in the United States and Europe crashed and market shares fell (Yeomans, 2015). The search queries seem to behave consistently with developments in the financial crisis, since investors will pay more attention to the stock markets in times of crises. Furthermore,

[4] The AEX index is the most important index in the Netherlands. It includes the 25 largest stocks, based on market capitalization, of Euronext Amsterdam.

the SVI and the Wikipedia page views also show a peak in June 2016. This can be explained by the outcome of the referendum that was taking place in the United Kingdom concerning the Brexit, which disrupted the financial markets worldwide (Chapman, 2016). Hence, the search queries seem to respond, in terms of higher search volumes, to events that have an impact on the financial world. These facts suggest that both search queries can serve as reliable proxies for investor attention of the AEX index.

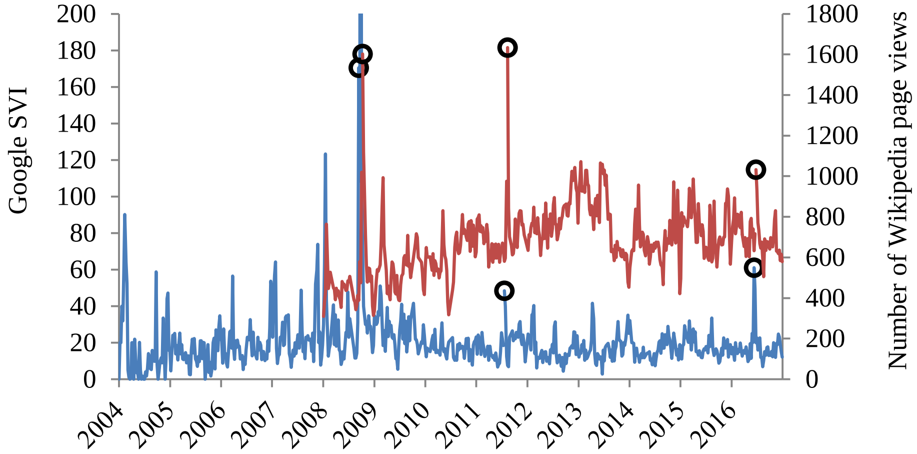

Figure 1. Google and Wikipedia. The graph shows the Google SVI (blue) and the number of Wikipedia page views (red) over time. The black circles mark some of the important peaks of both search queries.

3.2 Google Search Volume Index

The Google Trends search volume index is provided by Google Inc. Google has a market share of 93.8% in the market of search engines in the Netherlands[5], making it a representative measure of internet search behavior. Weekly Google Web search volume indices for the search word "AEX" are collected from Google Trends. A justification of the chosen search term can be found in Appendix A. The data are retrieved for the period from January 2004 to December 2016, since Google Trends does not provide any information before that period. The search volume index shows how often people entered the search-term relative to the total search volume over a particular time period. The searches are restricted to only finance related searches, to better capture investor attention for the index. Because particular words can have different meanings in other countries, only search queries in the Netherlands are considered.

[5] Source: StatCounter Global Stats (http://gs.statcounter.com/search-engine-market-share/all/netherlands)

Google Trends provides only monthly observations for larger time periods; the weekly indices are just available for short periods of time (e.g. one year). In order to get the weekly indices for the period from 2004 to 2016, I follow the method used by Risteski & Davcev (2014). I obtain weekly indices for each year seperately and I collect monthly indices over the entire period. Using these data, I calculate the adjustment factor, which is the monthly SVI divided by the weekly SVI of the first week of the month. The same adjustment factor is used for the other weeks of the month. Finally, the weekly SVI is multiplied by this adjustment factor to get the weekly adjusted values. The dataset contains two observations for which the weekly SVI of the first week of the month had a value of zero. In that case, the adjustment factor was calculated by dividing the monthly SVI by the weekly SVI of the second week of the month.

Although this is a valid method of obtaining weekly data, it has one disadvantage when considering volatility forecasts. Since the adjustment factor is calculated as dividing the monthly value by the weekly value, the adjusted weekly values of the first weeks of the month contain some future information about the search index. Hence, it needs to be taken into account that the lags of the search index actually contain some information about the present values. The present and future values are therefore not predicted using completely clean values from the past, which can lead to inconsistent forecasts. However, given the limited options that are available, this is still a reasonable method to consider, but one needs to be careful when interpreting the results.

Descriptive statistics of the search volume index are given in the first row of Table 1. The SVI shows some outliers in the data, which is evident from the graph (Figure 1). Therefore a 99% winsorization of the SVI variable was applied. From the skewness and kurtosis measures it is observed that the search volume index observations are slightly skewed. This is also evident from the histogram (Appendix B, Figure B1.a). In order to solve this, natural logarithm transformations of the search volume indices are considered, as done by Da et al. (2011) and Dimpfl & Jank (2016). The fourth row of Table 1 reports summary statistics of the logarithm of the winsorized search volume indices. Skewness and kurtosis have been reduced. The histogram also shows that the observations are closer to a normal distribution in the case of the log transformations (Appendix B, Figure B1.b).

Table 1. Descriptive statistics. *The table gives the descriptive statistics of (the log of) the SVI, Wikipedia page views and historical and implied volatility. Historical and implied volatility are denoted as HistVol and ImpVol, respectively. The variable for the number of Wikipedia page views is called Wiki. Log_X indicates the natural logarithm of variable X.*

Variable	Number of observations	Mean	Median	Standard Deviation	Min	Max	Skewness	Kurtosis
SVI	679	19.57	17.04	11.89	0.00	73.81	2.00	8.64
HistVol	680	19.75	18.10	8.51	9.86	45.85	1.71	5.46
ImpVol	574	20.38	18.01	9.12	10.66	104.80	3.72	26.12
Wiki	444	683.46	673.00	167.28	309.00	1634.00	0.85	6.60
Log_SVI	669	2.84	2.85	0.57	-0.69	4.30	-0.80	7.34
Log_HistVol	680	2.91	2.90	0.36	2.29	3.83	0.85	3.34
Log_ImpVol	574	2.95	2.89	0.33	2.37	4.65	1.23	5.50
Log_Wiki	444	6.50	6.51	0.25	5.73	7.40	-0.34	3.76

3.3 Wikipedia page views

Next, I obtain the number of daily page views of the Wikipedia page of the AEX index from Wikipedia's page view statistics[6]. I consider the AEX index page in the English language for two reasons. First, the AEX index also attracts foreign investors, who do not speak the Dutch language, and hence they will search for information, regarding the stocks, that is written in English rather than Dutch. In addition, it may be plausible to believe that many Dutch investors would also read about the index in the English language. The number of page views of the English Wikipedia page will therefore better capture the attention of all investors. In this case, the fact that "AEX" may have different meanings in other countries is not a problem, since people will only visit the Wikipedia page if they are interested in the stock index. Secondly, when comparing the page view statistics of the Dutch and English page, it is evident that the English page consistently has a higher number of page views. Since the English page is visited more often, it will give a more accurate estimate of investors' attention. Because of limited data availability, the page views are collected for the period from January 2008 to December 2016. In order to allow for fair comparisons with the Google Trends data, the daily data are transformed to weekly observations by simply adding the number of daily page views of every day of the week.

[6] Source: http://stats.grok.se/

Descriptive statistics are given in the fourth row of Table 1. The skewness measures and the histogram (Appendix B, Figure B1.c) do not reveal that the variable is highly skewed. However, in order to allow for similar interpretations as the Google search index and to handle the outlier in the data, I also consider the log of the Wikipedia page views. As evident from the summary statistics (final row of Table 1) the data is less volatile; the minimum and maximum values are closer to each other and the standard deviation is low. The outlier has also been reduced (Appendix B, Figure B1.d).

The correlation coefficient between the SVI and the Wikipedia page views is -0.12, which indicates a weak negative correlation. When more investors search for the stock online in Google, fewer investors read about the stock index on the Wikipedia page. This is a surprising result, since most people tend to visit the Wikipedia page via Google searches, which would imply a positive correlation.

3.4 Historical and implied volatility

Weekly volatilities, both historical and implied, of the AEX index are obtained from Bloomberg. Historical volatilities, calculated as the average volatility over the past 30 days, are retrieved from January 2004 to December 2016. Implied volatilities are only available from January 2006 onwards. Bloomberg calculates the implied volatilities at a fixed level of moneyness of 100% with a maturity of three months.

The graph (Appendix B, Figure B2) illustrates that historical volatility shows some weak trend over time, with a relatively large increase in volatility during the time of the crisis. Implied volatility shows more variation. An outlier is also present right before the start of the financial crisis. Summary statistics are given in the second and third rows of Table 1. The measures suggest that implied volatility is slightly skewed. Therefore, I will also use the natural logs of both measures of volatility in order to allow for similar interpretations. The last two rows of Table 1 show that skewness and kurtosis are reduced. Outliers also do not seem to be a problem anymore. The histograms of the variables also confirm this (Appendix B, Figure B1.f and B1.h).

The correlation coefficient between historical and implied volatility is 0.64, which suggest a strong positive correlation between the two variables. This is also apparent from the graph

(Appendix B, Figure B2). The search volume index is weakly positively correlated with volatility; it has a correlation coefficient of 0.11 and 0.20 with historical and implied volatility, respectively. The Wikipedia page views are also weakly, but negatively correlated with historical (-0.29) and implied (-0.16) volatility.

3.5 Control variables

Several articles have also investigated the effect of different economic variables on stock volatility. For example, Vlastakis & Markellos (2012) use market returns as a control variable in their models and find that returns have a significant effect on volatility. Andersen (1996) concludes that trading volume plays an important role in modelling volatility. Therefore, I also collect data on the returns and trading volumes of the index. I retrieve the trading volume (expressed in billions of Euros)[7], the opening price (P_o) and the adjusted closing price[8] (P_c) of each week from Yahoo! Finance[9]. The return of week t is then calculated as follows:

$$Ret_t = \frac{P_{c,t} - P_{o,t}}{P_{o,t}} \; x \; 100\ \% \tag{1}$$

3.6 Stationarity

In order to ensure unbiased and efficient forecasts, it is important to investigate whether the given variables are stationary. An augmented Dickey-Fuller (ADF) test will be performed to analyze this, as it is the most commonly used measure for tests of non-stationarity. The null hypothesis of this test is that the model has a unit root. If the hypothesis is rejected, the model is stationary. The results are given in Table C1 of Appendix C. The null hypothesis is rejected at the 1% level for both the log of the SVI and implied volatility, the returns and the trading volume. However, for the logarithm of the historical volatility the null hypothesis cannot be rejected, indicating non-stationarity. To eliminate this, first differences ($logHistVol_t - logHistVol_{t-1}$) are taken. When performing the ADF test again for this new variable, the null hypothesis is rejected, which suggests that the variable is stationary. Although the log of implied volatility is stationary, first differences are also taken for this variable

[7] Data on trading volume is only available from October 2004 onwards.
[8] Close price adjusted for dividends and splits.
[9] Yahoo! Finance is a media property that collects data on financial news and commentary, such as stock quotes and financial reports (source: https://finance.yahoo.com/).

($logImpVol_t - logImpVol_{t-1}$) in order to allow for the same interpretations in the two models of volatility.

3.7 The investment portfolios

Moreover, an in-depth analysis of the individual stocks of the AEX index is also provided in order to use the volatility forecasts to construct investment portfolios based on the risk parity approach, since investors are mainly concerned with building their optimal investment portfolios. The AEX index is an index that consists of a maximum of the 25 largest stocks, based on market capitalization, that are listed on Euronext Amsterdam. Most investors only buy shares of the individual stocks. They are unable to buy shares of the AEX index itself, unless they buy an exchange-traded fund (ETF) share or a future on the index. Therefore, the AEX search term may only represent a portion of the total trading activity of the index, since most investors search for online information about the individual stocks that they are interested in. Hence, in order to investigate whether the same conclusions regarding the SVI can be drawn for those individual stocks, I perform an in-depth analysis for the 25 individuals stocks that are part of the AEX index as of July 2017[10]. An overview of the companies included in the analysis can be found in Table D1 of appendix D.

The SVI search terms are the company name and the ticker symbol of each respective stock. The search volume index data have been collected from Google Trends and are restricted to finance-related searches only. For simplicity, the weekly SVI have been collected directly from Google Trends from July 2012 to June 2017 since Google Trends does not provide direct weekly data before that time period. The weekly historical volatilities of the individual stocks over the above mentioned time period are retrieved from Bloomberg. All variables used in this analysis proved to be stationary based on the ADF test.

4. Methodology

4.1 Linear regressions

The log of the search volume index in week t is denoted as log_SVI_t. The log of the Wikipedia page views is denoted as log_Wiki_t. In addition, the first difference of the log of

[10] Source: https://www.aex.nl/products/indices/NL0000000107-XAMS/market-information

the historical volatility and implied volatilities are defined as $dlog\_HistVol_t$ and $dlog\_ImpVol_t$, respectively. The first model that is investigated is a simple linear regression of volatility on the search volume index and the Wikipedia page views. The regression coefficients are defined as β_i, where β_0 denotes the constant. I control for the returns and trading volume, defined as Ret_t and $TradV_t$ respectively, in the following models:

$$dlog\_HistVol_t = \beta_0 + \beta_1 log\_SVI_t + \beta_2 log\_Wiki_t + \beta_3 Ret_t + \beta_4 TradV_t + \varepsilon_t \tag{2}$$

$$dlog\_ImpVol_t = \beta_0 + \beta_1 log\_SVI_t + \beta_2 log\_Wiki_t + \beta_3 Ret_t + \beta_4 TradV_t + \varepsilon_t \tag{3}$$

4.2 Forecasting models

Previous literature shows that ARCH models and simple linear regression models provide better forecasts of volatility than other, more complex, models such as the moving average model, historical mean model and exponential smoothing model (Brailsford & Faff, 1996). Therefore, simple autoregressive distributed lag (ARDL) models are created to construct volatility forecasts, controlling for the other variables. The following models serve as the benchmark models in this case:

$$dlog\_HistVol_t = \beta_0 + \sum_{i=1}^{p} \beta_i dlog\_HistVol_{t-i} + \sum_{i=1}^{q} \beta_{p+i} Ret_{t-i} + \sum_{i=1}^{r} \beta_{p+q+i} TradV_{t-i} + \varepsilon_t \tag{4}$$

$$dlog\_ImpVol_t = \beta_0 + \sum_{i=1}^{p} \beta_i dlog\_ImpVol_{t-i} + \sum_{i=1}^{q} \beta_{p+i} Ret_{t-i} + \sum_{i=1}^{r} \beta_{p+q+i} TradV_{t-i} + \varepsilon_t \tag{5}$$

In order to investigate the predictive power of the Google SVI, these models are extended with lags of the SVI:

$$dlog\_HistVol_t = \beta_0 + \sum_{i=1}^{p} \beta_i dlog\_HistVol_{t-i} + \sum_{i=1}^{q} \beta_{p+i} Ret_{t-i} + \sum_{i=1}^{r} \beta_{p+q+i} TradV_{t-i} + \sum_{i=1}^{s} \beta_{p+q+r+i} log\_SVI_{t-i} + \varepsilon_t \tag{6}$$

$$dlog\_ImpVol_t = \beta_0 + \sum_{i=1}^{p} \beta_i dlog\_ImpVol_{t-i} + \sum_{i=1}^{q} \beta_{p+i} Ret_{t-i} + \sum_{i=1}^{r} \beta_{p+q+i} TradV_{t-i} \sum_{i=1}^{s} \beta_{p+q+r+i} log\_SVI_{t-i} + \varepsilon_t \tag{7}$$

Finally, the predictive power of the Wikipedia page views are investigated by extending the ARDL models of equations (6) and (7) to include lags of the log of the Wikipedia page views, using the same control variables. These extended models are compared to the benchmark models to examine the predictive power of the search queries.

An information criterion is used in order to determine the appropriate lag length of the different models. More specifically, I choose the model that has the lowest Bayesian information criterion (BIC), since this criterion penalizes complex models more heavily. When including more lags, more parameters need to be estimated, hence increasing the uncertainty of the model. Each type of model is estimated using the same number of observations when determining the lag length to ensure fair comparisons between the models.

Once the appropriate lag length of all the variables in each model have been determined, both in-sample and out-of-sample forecasts are considered. For the in-sample forecasts, all observations are used to estimate the regression. The fitted values from these regressions are compared to the actual values using statistical measures, as discussed later in this section. Regarding the pseudo out-of-sample forecasts, I generate forecasts for a number of observations P. Let $s = T - P$ be the estimation window, where T is the total number of observations in the dataset. Then the forecasting regression is estimated using the shortened data for $t = 1, \dots, s$. The first estimation window contains all the observations of the first two years of the dataset, following Dimpfl & Jank (2016). Notice that this does not mean that all estimations are based on the same time period, since implied volatility and Wikipedia page views data are only available from 2006 and 2008 onwards, respectively. The two years contain enough observations and show enough variation to ensure reliable estimations. The forecasted value for the first period beyond this shortened sample, $s + 1$, is denoted as $\tilde{Y}_{s+1|s}$ and is be compared to the actual value of volatility in that period, Y_{s+1}. These steps are repeated for the remaining dates from $s = T - P + 1$ to T, hence increasing the estimation window for every new forecast, incorporating more information to produce more accurate estimates.

4.3 Forecasting performance measures

The performance of the forecasting models are compared using the mean squared forecast error (MSFE) and the quasi-likelihood loss function (QL), which are robust to noise in the volatility proxy (Patton, 2011). In addition, they are one of the most widely used measures in empirical literature. They are defined as follows:

$$MSFE = \left(Y_{s+1} - \tilde{Y}_{s+1|s}\right)^2 \tag{8}$$

$$QL = \frac{Y_{s+1}}{\tilde{Y}_{s+1|s}} - \log\left(\frac{Y_{s+1}}{\tilde{Y}_{s+1|s}}\right) - 1 \tag{9}$$

The MSFE measures the magnitude of a typical mistake using the model. In both cases, the measures equal zero whenever the forecasted value is equal to the actual value of volatility, or in other words, the volatility has been forecasted perfectly. Hence, lower MSFE and QL scores indicate a better forecasting performance. The MSFE and QL are calculated for every observation for which a forecasted value has been generated. The MSFE and QL scores (in each model) are then summed up and divided by the number of observations. This is necessary for a fair comparison between the models, since they all have a different number of observations.

In addition, Mincer & Zarnowitz (1969) regressions are performed for every model. The R^2 of these regressions are evaluated to examine forecasting performances, where a higher R^2 implies that the forecasted value explains more of the variation in actual volatility, i.e. a better forecast of volatility. The regression is defined as follows:

$$Y_{s+1} = \beta_0 + \beta_1 \tilde{Y}_{s+1|s} + \varepsilon_t \tag{10}$$

Finally, Granger causality tests are performed for the in-sample forecasts to test whether the lags of the log of the search queries have predictive content, beyond the other regressors in the model. The null hypothesis of this test is that the coefficients on all lags of the log of the Google index and Wikipedia page views are zero. This is tested using the F-statistic. If the null hypothesis is rejected, it indicates that the SVI or the Wikipedia page views Granger cause the volatilities, or in other words, the variables are a useful predictor of volatility, given the other control variables.

4.4 The volatility parity approach

As an additional in-depth analysis, I use different volatility forecasts based on different forecasting models to construct an investment portfolio using the risk (volatility) parity approach to investigate the implications of the forecasts in a trading example. An investment portfolio should consist of several assets that each are assigned a certain weight. In this particular case, I construct a portfolio of the 25 individual stocks that are part of the AEX index and determine the weight by the volatility forecasts of the different assets.

Two volatility forecasting models are compared. The benchmark model is the historical average volatility model, which is a standard approach used by investors to predict future volatility of the stocks in their portfolios. This model predicts future volatility by simply observing the volatility from the past. In this particular case, volatility is calculated as the average volatility of the past 30 days. The second model is constructed using past values of volatility and the SVI as a predictor of historical volatility in an out-of-sample setting, using the same control variables as before, since SVI is the most commonly used search query in volatility forecasts.

The two models are used to generate out-of-sample forecasts of historical volatility with an increasing estimation window. The estimation window for the first forecast included the observations of the first two years of the dataset, adding one extra observation for every subsequent forecast. The forecasting performances of both models are compared using the same measures as before: the mean squared forecast error (MSFE), the quasi-likelihood loss function (QL) and the R^2 of the Mincer-Zarnowitz regressions.

Subsequently, these volatility forecasts are used to construct two different investment portfolios and investigate which portfolio yields the highest return. The weight of each asset in the portfolio is calculated using the following formula:

$$w_i = \frac{(\sigma_t^i)^{-1}}{\sum_{i=1}^{25}(\sigma_t^i)^{-1}} \tag{11}$$

where w_i denotes the weight of asset i and $(\sigma_t^i)^{-1}$ denotes the inverse of the volatility of asset i at time t. Subsequently, the return of the portfolio is calculated using the formula:

$$r_{t,t+1}^{VP} = \sum_{i=1}^{25}(w_i * r_{t,t+1}^i) \tag{12}$$

Here, $r_{t,t+1}^{VP}$ denotes the return of the portfolio from t to $t+1$ constructed using the volatility parity approach and $r_{t,t+1}^{i}$ indicates the return of asset i over the period from t to $t+1$. For simplicity, only long positions are considered and leverage is not taken into account. The returns of the individual assets[11] and the total return of the portfolio are calculated as the cumulative returns over the period over which the volatility forecasts have been made (July 2014 to June 2017).

5. Results

5.1 Interdependence of variable quantities

The results of the linear regressions are presented in Table 2 and 3. When taking historical volatility as the dependent variable, the log of the SVI is significant in the first four models at the 5% level, but it loses significance after adding the control variables. The log of the Wikipedia page views is not significantly correlated with the dependent variable, however. Regarding the control variables, only the coefficient of the trading volume is significantly positive, which is consistent with past literature (Andersen, 1996).

For the models considering implied volatility, the results are somewhat different. In this case, the log of the SVI is not significant in the estimated models, whereas the coefficient of the log of the Wikipedia page views is significant and positive in all models estimated, suggesting a positive correlation between the number of page views and implied volatility. Again, the returns are not significant, as opposed to previous literature (Vlastakis & Markellos, 2012). However, since there is good reason to believe that returns may have a delayed effect on volatility, the returns will still be kept in future models. Another reason is that there still may be other variables that are correlated with the dependent and independent variables but that have not been included in these models (since the coefficient of the variable changes in the fifth model), which could cause a downward bias of the coefficients.

These results suggest that the search volume index is weakly, positively correlated with historical volatility, but the correlation with implied volatility is still ambiguous. This is in contrast to previous literature that found a significant correlation with both historical and implied volatility (Andrei & Hasler, 2014; Vlastakis & Markellos, 2012). Furthermore, the

[11] The individual asset returns are calculated by retrieving the stock prices from Bloomberg.

Wikipedia page views are positively correlated with implied volatility, also when controlling for the other variables. The results also imply that the page views have no significant correlation with historical volatility, however.

Table 2. Linear regressions on historical volatility. *The table shows the results of the linear regressions on the first difference of the log of historical volatility. The natural logarithm of the SVI and the Wikipedia page views are denoted log_SVI and log_Wiki, respectively. The control variables are the index returns (Ret) and the trading volume expressed in billions (TradV). Robust standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p <0.01*

Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
log_SVI	0.009**		0.018**	0.017**	0.010*
	(0.003)		(0.009)	(0.008)	(0.005)
log_Wiki		0.011	0.018	0.017	0.016
		(0.016)	(0.018)	(0.017)	(0.016)
Ret				-0.022	0.012
				(0.067)	(0.057)
TradV					0.051**
					(0.022)
Constant	-0.025***	-0.071	-0.163	-0.161	-0.161
	(0.009)	(0.100)	(0.137)	(0.133)	(0.125)
Observations	669	444	444	443	443
R^2	0.027	0.008	0.067	0.067	0.136
R^2 adjusted	0.025	0.005	0.063	0.061	0.128

Table 3. Linear regressions on implied volatility*. The table shows the results of the linear regressions on the first difference of the log of implied volatility. The natural logarithm of the SVI and the Wikipedia page views are denoted log_SVI and log_Wiki, respectively. The control variables are the index returns (Ret) and the trading volume expressed in billions (TradV). Robust standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p <0.01*

Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
log_SVI	0.020*		0.022	0.024*	0.004
	(0.011)		(0.015)	(0.014)	(0.012)
log_Wiki		0.050**	0.057**	0.059**	0.056**
		(0.025)	(0.026)	(0.027)	(0.024)
Ret				0.122	0.215
				(0.188)	(0.181)
TradV					0.142***
					(0.039)
Constant	-0.058*	-0.324**	-0.436**	-0.449**	-0.451**
	(0.032)	(0.161)	(0.195)	(0.195)	(0.174)
Observations	572	443	443	442	442
R^2	0.008	0.014	0.023	0.025	0.076
R^2 adjusted	0.007	0.012	0.019	0.018	0.067

5.2 Constructing forecasting models

Several forecasting models have been constructed to determine the most appropriate lag length of every variable. I did not consider models with many lags (i.e. more than five lags), because of the risk of reducing the predictive power of the models, since more parameters would need to be estimated, hence increasing the uncertainty of the model. First, I considered the forecasting model of historical volatility with just its own lags. The results are given in Table E1 of Appendix E. The BIC had the lowest value for the model including two lags, hence the AR(2) model was preferred. Both the first and second lags were significant at the 5% level.

Hereafter, lags of the control variables were also added to the model. The BIC score increased with the number of lags of both variables (Appendix E, Table E1), so only the first lag was considered. Hence, to provide forecasts of historical volatility, the benchmark model included two lags of historical volatility and one lag of the returns and trading volume.

For the next forecasting model, lags of the SVI were added to the previous model. As evident from the results (Appendix E, Table E2), the BIC score had the lowest value for the model that included only one lag of the search index. The first lag was significantly positive in all three models; other lags were not significant in the second and third model. Furthermore, lags of the Wikipedia page views were also included in the benchmark model (Appendix E, Table E3). Because of the shorter time frame of the Wikipedia data, fewer observations were taken into account. The coefficients of the lags of the page views were not significant in any model. The BIC had the lowest value for the model including the first and second lag of the page views. Finally, lags of the both the SVI and Wikipedia page views were also included. The results are given in Table E4 of Appendix E. In this case, the most appropriate forecasting model was the model including two lags of the Wikipedia page views and one lag of the SVI. None of the coefficients of the Wikipedia page views were significant, however.

The forecasting models of implied volatility looked slightly different. Again, the benchmark model only included lags of implied volatility and the two control variables. The estimated models can be found in Table E5 of appendix E. In this case, however, when including more lags of implied volatility the BIC score increased, so only the first lag was included in the model. Hence, the benchmark model for the forecasts of implied volatility included one lag of the dependent variable, and two lags of the returns and trading volume.

Hereafter, lags of the SVI were included to construct the second forecasting model. The results are given in Table E6 of Appendix E. The coefficients of the lags of the SVI were not significant in any model that was estimated. In addition, the BIC score increased when adding more lags of the SVI, hence only including the first lag of the SVI. Moreover, lags of the Wikipedia page views were added to the benchmark model (Appendix E, Table E7). Again, the coefficients were not significant in all three models that were estimated and the BIC score increased after adding more lags of the Wikipedia page views. Therefore, only the first lag was considered for the forecasting model. The final forecasting model of implied volatility included only the first lag of both the SVI and the Wikipedia page views. The coefficients of the lags of the page views were not significant in any of the three estimated models (Appendix E, Table E8). The significance of the other variables remained the same.

5.3 In-sample forecasts

After determining the appropriate lag length of each variable in every model, the volatility forecasts were made using the estimated models. These models were estimated using more observations than before. This is because previous models needed to have the same number of observations when including more lags of the variables to allow for a fair comparison between the different models. When introducing the $AR(p)$ model, p number of observations is used as a 'pre-sample', hence losing a larger number of observations in the regression estimation when introducing more lags of the variables.

The results of the forecasting models of historical volatility are given in the first three columns of Table 4. The MSFE barely decreased after adding the first lag of the SVI, but the QL showed a large decrease from 13.39 to 4.47, suggesting an improvement in the forecasting ability of the model. Moreover, the R^2 increased and the Granger causality test for the coefficients of the SVI could be rejected (Appendix F, Table F1). This suggests that the SVI has predictive power in historical volatility in-sample forecasts, which is in line with previous literature (Dimpfl & Jank, 2016; Smith, 2012). The third model showed less consistent results suggesting that the lags of the Wikipedia page views have predictive power; the MSFE slightly increased and the F-statistic of the Granger-causality test was insignificant, whereas the QL score and the R^2 of the model increased. Hence, the predictive power of the Wikipedia page views remains uncertain in this case. Nevertheless, the fourth model showed that the search queries were able to improve the volatility forecasts, since almost all measures scored better compared to previous models. The MSFE remained about the same value, but the F-statistic of the Granger causality test was significant at the 5% level for the coefficients of the lags of the SVI, and for the joint test of the lags of both variables. However, significance reduced after adding the lags of the Wikipedia data. Also, the null hypothesis of the test of the Wikipedia page views could not be rejected. Hence, the results suggest that the improvement in forecasting ability was probably mainly driven by the SVI.

Furthermore, other forecasting models have been used to generate in-sample forecasts of implied volatility as well. The results are given in the first three columns of Table 5. The MSFE of the second model remained about the same value compared to the model excluding the lag of the SVI. The other two measures also did not point to any improvement in forecasting ability. Hence, the SVI alone did not improve the forecasting performance of the model. Nevertheless, the forecasts did seem to improve after adding lags of the Wikipedia page views in the third and fourth model, as both models scored better on all three measures compared to the benchmark model. This may indicate that the SVI and Wikipedia page views have some combined predictive power in forecasting implied volatility. Notice that these results are derived descriptively by looking at the changes in the values of forecasting performance measures only. None of the Granger causality tests could be rejected in this case.

Table 4. Forecasting performances for historical volatility*. The table shows the forecasting performance of in-sample and out-of-sample forecasts of four different models: including lags of the dependent variable, returns and trading volume (1), adding lags of the SVI (2), adding lags of the Wikipedia page views (3), and adding lags of both search queries (4). The dependent variable is the first difference of the log of historical volatility. The three performance measures are the mean squared forecast error (MSFE), the quasi-likelihood loss function (QL) and the R^2 of the Mincer-Zarnowitz regressions. The bold numbers indicate the preferred model.*

	IN-SAMPLE			OUT-OF-SAMPLE		
Model	**MSFE**	**QL**	**R^2**	**MSFE**	**QL**	**R^2**
(1)	0.0008	13.3886	0.1177	0.0010	3.9363	0.0376
(2)	**0.0008**	4.4717	0.1362	0.0010	2.6409	0.0475
(3)	0.0009	6.2554	0.1779	0.0007	12.0390	0.0223
(4)	0.0008	**2.9801**	**0.2034**	**0.0007**	**2.4360**	**0.0399**

Table 5. Forecasting performances for implied volatility. *The table shows the forecasting performance of in-sample and out-of-sample forecasts of four different models: including lags of the dependent variable, returns and trading volume (1), adding lags of the SVI (2), adding lags of the Wikipedia page views (3), and adding lags of both search queries (4). The dependent variable is the first difference of the log of implied volatility. The three performance measures are the mean squared forecast error (MSFE), the quasi-likelihood loss function (QL) and the R^2 of the Mincer-Zarnowitz regressions. The bold numbers indicate the preferred model.*

	IN-SAMPLE			OUT-OF-SAMPLE		
Model	**MSFE**	**QL**	**R^2**	**MSFE**	**QL**	**R^2**
(1)	0.0055	1.9998	0.4856	0.0060	2.1335	0.4311
(2)	0.0055	3.3615	0.4862	0.0061	**1.6959**	0.4287
(3)	0.0049	1.2470	0.5374	**0.0041**	1.8221	**0.6127**
(4)	**0.0049**	**1.0682**	**0.5375**	0.0041	3.5100	0.6110

In addition, the results suggest that forecasting models of historical volatility outperform the implied volatility models when considering the MSFE. In all cases, the MSFE was lower for the historical volatility as compared to implied volatility. However, based on the QL and R^2 scores, the implied volatility models seemed to provide consistently better forecasts.

5.4 Out-of-sample forecasts

Now I focus on the out-of-sample forecasts with an increasing estimation window. The results of the historical volatility forecasts are given in the last columns of Table 4. When adding the lags of the SVI to the benchmark model, the value of the MSFE was almost unchanged, but the QL and R^2 scores improved, which may indicate a slight improvement in the forecasting performance. These results are similar to the in-sample forecasts, which are also consistent with previous literature that also considered both in-sample and out-of-sample forecasts (Dimpfl & Jank, 2016; Hamid & Heiden, 2015). However, the results of the third model including lags of the Wikipedia page views are somewhat inconsistent as the model seemed to improve the forecasting performance based on the MSFE, but worsened it when looking at the other two measures. Especially the QL score showed an extreme result of 12.04. One possible explanation for the differences in the QL and MSFE could be the way in which the two scores are calculated. The MSFE tends to punish extreme outliers more heavily, whereas the QL measure is likely to be higher when there are a larger number of small errors. When adding lags of both search queries, the forecasting performance seemed to improve based on all three measures. Therefore, the results provide some evidence for the combined predictive power of the search queries, but the forecasting ability of the Wikipedia page views separately remains ambiguous in the case of the out-of-sample historical volatility forecasts. It appears that again the SVI has more predictive power compared to the Wikipedia page views, but the evidence is weak.

The results for the implied volatility out-of-sample forecasts can be found in the last columns of Table 5. The second model, including only the lags of the SVI, did not perform any better compared to the benchmark model, based on the MSFE and the R^2. However, the forecasts did improve, based on these two measures, when adding the lags of the Wikipedia page views. The QL scores suggested the opposite result in both cases. This is in contrast to past literature that found significant and consistent results of improvements in forecasting ability across all measures (Dimpfl & Jank, 2016). Moreover, the results were inconsistent in the fourth model. In this case, the forecasting performance improved when including lags of both search queries, based on the MSFE and the R^2, but the QL score increased compared to the three previous models. Hence, the results are different compared to the in-sample forecasts. The results suggest that only the model including the lags of the Wikipedia page views generated better forecasts, which may indicate that the Wikipedia data have some predictive power in forecasting implied volatility, but there is no strong evidence that supports this hypothesis.

Moreover, as in the case of the in-sample forecasts, the historical volatility models seemed to provide better forecasts compared to the implied volatility models, based on the MSFE. The QL and R^2 scores suggested that (almost) all implied volatility models outperformed the historical volatility models.

5.5 The investment portfolios

The following part of the analysis uses the volatility forecasts of the individual stocks of the index in order to construct investment portfolios. The results for the out-of-sample forecasts are given in Table 6. Based on the QL and the R^2 measures, the benchmark model seemed to outperform the SVI model for almost every company. The differences are small, however. The benchmark model also seemed to outperform the SVI model for most companies based on the MSFE, but the results are less consistent. Nevertheless, the SVI forecasting models of some companies were able to outperform the benchmark. Some of these companies, including the ING Group, ASML and Unibail-Rodamco, have a relatively large index weighting, hence representing a larger part of the AEX index. It may therefore be the case that the SVI forecasting model can still prove to be useful when constructing the investment portfolios. All things considered, it is not possible to conclude that the volatility forecasting models including lags of the SVI search terms are able to outperform the benchmark models of all companies listed on the AEX index. Hence, further research is required to determine the predictive power of the SVI in forecasting stock market volatility.

Using the volatility forecasts of both models, the weights of the individual assets were calculated. These weights have been used to calculate the total return of both investment portfolios. The total returns of the individual stocks have also been calculated. The results are given in Table 7. Regarding the cumulative returns of the individual stocks, the volatility forecasts with SVI as one of its predictors yielded higher returns for most companies, even though the volatility forecasts of most stocks performed worse compared to the benchmark based on the three performance measures. However, the cumulative returns of the total investment portfolio based on the historical average volatilities had a slightly higher return (51.54%) as opposed to the portfolio that was based on the forecasting models including the SVI (50.86%), which is a surprising result. These results suggest that the SVI has no additional predictive power in volatility forecasts of individual stocks that can be used by investors to determine their optimal portfolio. It may, however, be a useful predictor when considering investments in one particular stock. Further research would need to investigate this.

Table 6. Out-of-sample forecast performance*. The table shows the forecasting performance of the historical volatility models for each company separately. The benchmark model (1) is the historical average volatility model. The second model (2) considers the SVI search terms of the company name and the ticker symbol, in addition to the control variables. The three performance measures are the mean squared forecast error (MSFE), the quasi-likelihood loss function (QL) and the* R^2 *of the Mincer-Zarnowitz regressions. The bold numbers indicate the preferred model.*

	MSFE		**QL**		$\mathbf{R^2}$	
Ticker symbol	(1)	(2)	(1)	(2)	(1)	(2)
AALB	7.423	**7.194**	**0.051**	0.054	0.858	**0.860**
ABN	**8.631**	18.851	**0.050**	0.071	**0.457**	0.315
AGN	25.395	**25.205**	0.058	0.058	**0.818**	0.813
ADRNY	**10.262**	11.964	**0.059**	0.062	**0.751**	0.700
AKZA	15.585	**15.108**	**0.059**	0.060	**0.740**	0.734
ATC	35.868	**38.660**	**0.051**	0.054	**0.892**	0.882
MT	**25.745**	27.793	**0.049**	0.052	**0.927**	0.921
ASML	18.614	**17.978**	**0.055**	0.058	**0.784**	0.779
BOKA	**28.081**	30.592	**0.062**	0.065	**0.796**	0.769
DSM	**8.124**	8.306	**0.050**	0.053	**0.820**	0.808
GLPG	**48.132**	54.687	**0.060**	0.069	**0.871**	0.852
GTO	38.758	**38.660**	**0.068**	0.077	**0.672**	0.645
HEIA	**5.572**	5.802	**0.049**	0.051	**0.855**	0.844
INGA	20.943	**20.480**	**0.049**	0.056	**0.835**	0.832
KPN	**11.224**	15.979	**0.052**	0.065	**0.760**	0.648
NN	**14.529**	15.254	**0.058**	0.064	**0.769**	0.751
PHIA	**8.251**	8.253	**0.053**	0.057	**0.799**	0.790
RAND	**19.305**	20.696	**0.046**	0.051	**0.816**	0.803
REN	**4.526**	4.726	**0.049**	0.051	0.880	**0.891**
RDSA	**9.113**	10.698	**0.046**	0.053	**0.912**	0.899
SBMO	**38.747**	41.714	**0.053**	0.067	**0.800**	0.773
UL	6.247	**5.983**	**0.046**	0.047	0.851	0.851
UNA	**14.138**	14.234	**0.056**	0.060	**0.784**	0.773
VPK	**31.506**	32.351	**0.068**	0.080	**0.676**	0.636
WKL	**5.252**	5.345	**0.050**	0.051	**0.859**	0.851

***Table 7. Portfolio returns**. The table given an overview of the total cumulative returns of the individual stocks as well as the total cumulative returns of both portfolios. The historical average volatility portfolio is based on forecasts from the historical average volatility model (the benchmark). The SVI portfolio is based on volatility forecasts using the forecasting model including lags of the SVI. The bold numbers indicate the higher return.*

Ticker Symbol	Historical Average Volatility	SVI
AALB	2.93%	**2.98%**
ABN	0.39%	**0.40%**
AGN	0.03%	**0.10%**
ADRNY	1.25%	**1.40%**
AKZA	2.30%	**2.38%**
ATC	0.80%	**0.84%**
MT	**0.79%**	0.77%
ASML	3.33%	**3.40%**
BOKA	-0.49%	**-0.36%**
DSM	2.03%	**2.12%**
GLPG	6.53%	**6.58%**
GTO	**-0.20%**	-0.33%
HEIA	3.02%	**3.20%**
INGA	2.71%	**2.86%**
KPN	**1.43%**	1.40%
NN	1.00%	**1.01%**
PHIA	2.52%	**2.58%**
RAND	2.30%	**2.37%**
REN	**3.25%**	1.22%
RDSA	**-0.40%**	-0.42%
SBMO	2.57%	**2.66%**
UL	0.73%	**0.78%**
UNA	2.52%	**2.60%**
VPK	1.30%	**1.47%**
WKL	3.41%	**3.56%**
PORTFOLIO	**51.54%**	50.86%

6. Conclusion

This paper investigates the explanatory and predictive power of the Google search volume index and the number of Wikipedia page views as proxies for investor attention in forecasting stock market historical and implied volatility of the AEX index. The volatility reflects the amount of risk associated with a certain stock or index and is therefore an important consideration when making financial investment decisions. Previous literature considers different indirect and direct measures of investor attention and finds significant correlations with asset prices and other financial indicators. These findings can be explained by the belief that investors only have limited resources available to evaluate the stocks in the market. The attention theory states that investors therefore tend to buy stocks that recently caught their attention, leading to positive price pressure in the short run.

The results suggest that the SVI is positively correlated with historical volatility, but the evidence becomes weaker after adding more control variables. Also, no significant correlation is found with respect to implied volatility. On the contrary, the number of Wikipedia page views is found to be significantly and positively correlated with implied volatility, but not with historical volatility.

Furthermore, I construct various forecasting models to examine the predictive power of both search queries in an in-sample and out-of-sample setting. Past literature finds significant evidence that the models including lags of the SVI are able to improve the forecasting performances of the models. However, in my study, I find no consistent evidence that the search queries have additional predictive power, comparing the forecasting performances of the models based on the MSFE, QL and the R^2 of the Mincer-Zarnowitz regressions. The results suggest that SVI has significant predictive power in forecasting historical volatility, whereas the Wikipedia page views show no consistent improvement in forecasting historical volatility as opposed to the benchmark model. This is true for both the in-sample and out-of-sample forecasts. Evidence was found for the combined predictive power of both variables, but this result was probably mainly driven by the significance of the SVI.

As for implied volatility, the models including the lags of the SVI fail to outperform relative to the benchmark model, considering both in-sample and out-of-sample forecasts. The models including Wikipedia page views slightly improve the in-sample and out-of-sample forecasts, but further research is required to explain the predictive power in more detail. Moreover, the results suggest that the search queries have combined predictive power for implied volatility forecasts in the case of the in-sample forecasts, but it remains ambiguous as for the out-of-sample forecasts.

As an additional in-depth analysis, I construct a forecasting model using past values of the SVI as a predictor of historical volatility in an out-of-sample setting to generate volatility forecasts of the 25 individual stocks that are part of the AEX index. This forecasting model is compared to the historical average volatility model which serves as the benchmark. I use these volatility forecasts to construct two different investment portfolios using the volatility parity approach, since financial investors are mainly concerned with forecasting stock volatility in order to construct their optimal investment portfolios. The results suggest that SVI forecasting models of the volatility of the 25 individual stocks are not able to outperform the benchmark model for all companies. Furthermore, the portfolio that is constructed using weights that are based on the historical average volatility forecasts yields higher returns compared to the portfolio that was based on the volatility forecasts of the SVI forecasting model. Hence, the SVI does not prove to have predictive power in forecasting historical volatility of the individual stocks.

This paper has several avenues open for further investigation. First of all, because of the limited options that were available, the weekly Google search volume index is calculated using forward-looking adjustment factors in the first part of the paper. Hence, it is unclear to which extent the significance of the results is driven by the forward-looking bias. Furthermore, future research is required to be able to take into account more observations regarding the Wikipedia page views, since these data are only available from 2008 onwards. In this way, a more fair comparison can be made between the predictive powers of both search queries. Moreover, regarding the investment portfolio construction, only long positions were considered for simplicity. Hence, it may be valuable to consider short positions for certain assets in further research to yield possibly higher returns. Leverage can also be taken into account.

Furthermore, there are other possibilities that can be explored in relation to this topic. Previous literature suggested that the Google SVI captures individual and retail attention in particular. However, it would be interesting to investigate the predictive power of institutional investor attention as well, using different proxies of attention. One possibility would be to consider the use of Bloomberg as a proxy of attention, since a majority of the Bloomberg users are institutional investors. Moreover, future research could consider using other volatility forecasting models to examine whether better forecasts can be made. For example, the ARCH class models can be used, including the GJR-GARCH. With respect to the investment portfolio construction, further research could examine other asset and risk allocation strategies construct investment portfolios, such as minimum variance and maximum diversification strategies. Other benchmark models can also be included in the analysis to investigate whether the forecasting models of the search queries are able to beat other benchmarks. Examples include implied volatility and equally weighted average statistical volatility.

References

Andersen, T. (1996). Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility. *The Journal of Finance, 51(1),* 169-204.

Andersen, T., Bollerslev, T., Diebold, F., & Labys, P. (2003). Modeling and Forecasting Realized Volatility. *Econometrica, 71 (2)*, 529-626.

Andrei, D., & Hasler, M. (2014). Investor Attention and Stock Market Volatility. *Review of Financial Studies, 28 (1)*, 33-72.

Aouadi, A., Arouri, M., & Teulon, F. (2013). Investor attention and stock market activity: Evidence from France. *Economic Modelling, 35*, 674–681.

Bank, M., Larch, M., & Peter, G. (2011). Google search volume and its influence on liquidity and returns of German stocks. *Financial Markets and Portfolio Management, 25(3),* 239–264.

Barber, B., & Odean, T. (2008). All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *Review of Financial Studies, 21 (2),* 785-818.

Bentes, S. (2015). A comparative analysis of the predictive power of implied volatility indices and GARCH forecasted volatility. *Physica A, 424*, 105-112.

Boudoukh, J., Feldman, R., Kogan, S., & Richardson, M. (2013). *Which News Moves Stock Prices? A Textual Analysis* (NBER Working Paper No. 18725). Retrieved from http://www.nber.org/papers/w18725

Brailsford, T., & Faff, R. (1996). An evaluation of volatility forecasting techniques. *Journal of Banking & Finance, 20*, 419-438.

Canina, L., & Figlewski, S. (1993). The Informational Content of Implied Volatility. *The Review of Financial Studies, 6(3),* 659-681.

Chapman, B. (2016, December 12). Brexit: Stock market listings collapse 60% in 2016 amid volatility and uncertainty. *The Independent*. Retrieved from http://www.independent.co.uk/news/business/news/brexit-stock-market-listings-drop-london-stock-exchange-a7469266.html

Chaves, D., Hsu, J., Li, F., & Shakernia, O. (2011). Risk parity portfolio vs. other asset allocation heuristic portfolios. *The Journal of Investing, 20(1),* 108-118.

Chen, J., Liu, Y., & Tang, Y. (2016). Investor Attention and Macroeconomic News Announcements: Evidence from Stock Index Futures. *Journal of Futures Markets, 36(3),* 240-266.

Choi, H., & Varian, H. (2012). Predicting the Present with Google Trends. *Economic Record, 88*, 2-9.

Christensen, B., & Prabhala, N. (1998). The relation between implied and realized volatility. *Journal of Financial Economics, 50*, 125-150.

Claessen, H., & Mittnik, S. (2002). Forecasting Stock Market Volatility and the Informational Efficiency of the DAX index Options Market. *The European Journal of Finance, 8(3),* 302-321.

Clarke, R., Silva, H., & Thorley, S. (2013). Risk parity, maximum diversification, and minimum variance: An analytic perspective. *The Journal of Portfolio Management, 39(3),* 39-53.

Colaco, H., De Cesari, A., & Hegde, H. (2016). Retail Investor Attention and IPO Valuation. *Forthcoming in European Financial Management*.

Da, Z., Engelberg, J., & Gao, P. (2011). In Search of Attention. *The Journal of Finance, 66*, 1461-1499.

Dimpfl, T., & Jank, S. (2016). Can internet search queries help to predict stock market volatility? *European Financial Management, 22*, 171-192.

Ding, R., & Hou, W. (2015). Retail investor attention and stock liquidity. *Journal of International Financial Markets, Institutions & Money, 37*, 12-26.

Fan, X., Yuan, Y., Zhuang, X., & Jin, X. (2016). Long memory of abnormal investor attention and the cross-correlations between abnormal investor attention and trading volume, volatility respectively. *Physica A, 469*, 323-333.

Fondeur, Y., & Karamé, F. (2013). Can Google data help predict French youth unemployment? *Economic Modelling, 30*, 117-125.

Hamid, A., & Heiden, M. (2015). Forecasting volatility with empirical similarity and Google Trends. *Journal of Economic Behavior & Organization, 117,* 62-81.

Kahneman, D. (2011). *Thinking, Fast and Slow*. London: Penguin Books.

Kim, S., & Kim, D. (2014). Investor sentiment from internet message postings and the predictability of stock returns. *Journal of Economic Behavior & Organization, 107,* 708-729.

Kristoufek, L. (2013). BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Scientific reports, 3,* 3415.

Li, X., Xie, H., Chen, L., Wang, J., & Deng, X. (2014). News impact on stock price return via sentiment analysis. *Knowledge-Based Systems, 69,* 14-23.

McQueen, G., & Roley, V. (2015). Stock Prices, News, and Business Conditions. *The Review of Financial Studies, 6(3)*, 683-707.

Mincer, J., & Zarnowitz, V. (1969). The evaluation of economic forecasts. *Economic forecasts and expectations: Analysis of forecasting behavior and performance*, 3-46.

Nikkinen, J., & Sahlström, P. (2001). Impact of Scheduled U.S. Macroeconomic News on Stock Market Uncertainty: A Multinational Perspecive. *Multinational Finance Journal, 5(2)*, 129-148.

Park, S., Lee, J., & Song, W. (2017). Short-term forecasting of Japanese tourist inflow to South Korea using Google trends data. *Journal of Travel & Tourism Marketing, 34*, 357-368.

Patton, A. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics, 160*, 246-256.

Pearce, D., & Roley, V. (1984). *Stock Prices and Economic News* (NBER Working Paper No. 1296). Retrieved from http://www.nber.org/papers/w1296

Pilbeam, K., & Langeland, K. (2015). Forecasting exchange rate volatility: GARCH models versus implied volatility forecasts. *International Economics and Economic Policy, 12(1)*, 127-142.

Preis, T., Moat, H., & Stanley, H. (2013). Quantifying Trading Behavior in Financial Markets Using Google Trends. *Scientific Reports, 3*, 1684.

Risteski, D., & Davcev, D. (2014). Can We Use Daily Internet Search Query Data to improve Predicting Power of EGARCH Models for Financial Time Series Volatility? *International conference on Computer Science and Information Systems.*

Smith, G. (2012). Google Internet search activity and volatility prediction in the market for foreign currency. *Finance Research Letters, 9*, 103-110.

Ungeheuer, M. (2017). Stock Returns and the Cross-Section of Investor Attention. Retrieved from SSRN: https://ssrn.com/abstract=2931547.

Vakrman, T., & Kristoufek, L. (2015). Underpricing, underperformance and overreaction in initial public offerings: Evidence from investor attention using online searches. *SpringerPlus, 4(1)*, 84.

Vlastakis, N., & Markellos, R. (2012). Information demand and stock market volatility. *Journal of Banking & Finance, 36*, 1808-1821.

Vozlyublennaia, N. (2014). Investor attention, index performance, and return predictability. *Journal of Banking & Finance, 41*, 17-35.

Yeomans, J. (2015, August 24). How did Black Monday start? *The Telegraph*. Retrieved from http://www.telegraph.co.uk/finance/china-business/11821043/How-did-Black-Monday-start.html

Appendix A: Google Search Terms

The AEX index is an index composing of a maximum of 25 different stocks listed on Euronext Amsterdam. Most investors buy shares of the individual stocks that belong to the AEX. Investors are unable to actually buy a share of the AEX itself. Higher volatility of the AEX index that followed after increased investor attention is therefore not caused by investors buying the AEX index. Hence, the search term “AEX” may only represent a small portion of the trading activity of the AEX. However, investors are able to buy an exchange-traded fund (ETF) share. When an investor buys the ETF share, the authorized participant of this investor buys a share of every stock that is part of the AEX index. In that case, the trading volume of the AEX could be correlated with the search volume of the AEX. Some may argue that investors would rather search for the ETF shares than the AEX index when searching for information online, but the ETF market has experienced major growth in the last decade[12], hence representing a larger portion of the stock market. Considering these developments, the search term “AEX” could therefore explain a portion of the investors’ interest.

In order to justify this, data have been collected for two different search terms and two different types of searches to investigate its effect on the historical volatility of the ETF iShares. The two search terms are “AEX” and “iShares ETF”. In addition, a distinction has been made between Google Web searches and Google News searches. In order to minimize the chance of biased search volumes, weekly data have been collected from July 2005 to June 2017, since Google does not directly provide weekly data before that period. The weekly historical volatility of the ETF shares has been collected from Bloomberg. Implied volatility is not available. In addition, weekly trading volumes (expressed in millions) of the ETF shares are also retrieved from Bloomberg to include as a control variable.

The descriptive statistics of the variables are given in Table A1. None of the variables seemed to be highly skewed in order to cause any problems. The Augmented Dickey-Fuller test has been performed in order to test for stationarity. Based on this test, only the historical volatility was not stationary. Therefore, first differences of the variable have been taken in order to solve the problem.

[12] Thomas Jr., L. (2016, February 22). Explaining E.T.F.s and How They Gained Their Allure. *The New York Times*. Retrieved from https://www.nytimes.com/2016/02/23/business/dealbook/good-times-for-exchange-traded-funds.html

Table A1. Descriptive statistics. *The table shows the descriptive statistics of the search volume index of "AEX" Web searches (AEX_S), "AEX" News searches (AEX_N), "iShares ETF" Web searches (ETF_S), "iShares ETF" News searches (ETF_N) and the historical volatility of the ETF shares (Hist_ETF).*

Variable	Number of observations	Mean	Median	Standard Deviation	Min	Max	Skewness	Kurtosis
AEX_S	261	42.33	41.00	8.79	29.00	100.00	2.99	17.95
AEX_N	261	44.06	43.00	12.58	19.00	100.00	0.78	4.52
ETF_S	261	34.00	58.00	22.01	0.00	100.00	0.55	3.06
ETF_N	261	59.81	30.00	13.56	33.00	100.00	0.37	2.93
Hist_ETF	257	15.36	14.09	6.00	5.55	34.95	1.06	3.70

The weekly Google search volume index (SVI) of the AEX, concerning Google Web searches, is denoted as AEX_S and the search volume index of the ETF shares is denoted ETF_S. Furthermore, the SVI variables of the Google News searches for the AEX and the ETF are called AEX_N and ETF_N, respectively. In this case, the first difference of historical volatility is denoted $dHist$ and the trading volume is referred to as ETF_Trade. In order to investigate which of these search terms has explanatory power in explaining the historical volatility, the following regression is performed:

$$dHistVol_t = \beta_0 + \beta_1 * AEX\_S_t + \beta_2 * ETF\_S_t + \beta_3 * AEX\_N_t + \beta_4 * ETF\_N_t + \beta_5 * ETF\_Trade_t + \varepsilon_t \quad (1)$$

The results are given in Table A2. The search volume index of the AEX Web searches is significantly positive in all four models estimated. None of the other variables were significant in any of the models. These results suggest that the search term “AEX” concerning Google web searches is the most suitable search term to explain the historical volatility of the ETF shares. Hence, it represents investor attention for the trading activity of the AEX index.

Table A2. Linear regression of Google search terms on historical volatility. *The table shows the results of the linear regression performed for four different search terms, including the trading volume of the ETF shares as a control variable. It includes "AEX" Web searches (AEX_S), "ETF iShares" Web searches (ETF_S), "AEX" News searches (AEX_N) and "ETF iShares" News searches (ETF_N). The dependent variable is the first difference of the historical volatility of the ETF shares. Robust standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01*

Variable	**(1)**	**(2)**	**(3)**	**(4)**
AEX_S	0.087***	0.080***	0.076***	0.065**
	(0.023)	(0.026)	(0.02)	(0.026)
ETF_S	-0.009	-0.010	-0.008	-0.008
	(0.009)	(0.009)	(0.009)	(0.009)
AEX_N			0.015	0.017
			(0.011)	(0.012)
ETF_N			0.001	0.001
			(0.004)	(0.004)
ETF_Trade		0.377		0.457
		(0.328)		(0.331)
Constant	-3.185***	-2.995***	-3.502***	-3.308***
	(0.889)	(0.958)	(0.930)	(0.986)
Observations	252	252	252	252
R^2	0.140	0.143	0.147	0.152
R^2 adjusted	0.133	0.133	0.133	0.135

Appendix B: Additional figures

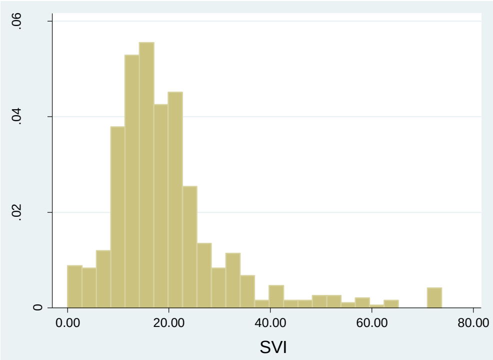

Figure B1.a. Histogram of the SVI.

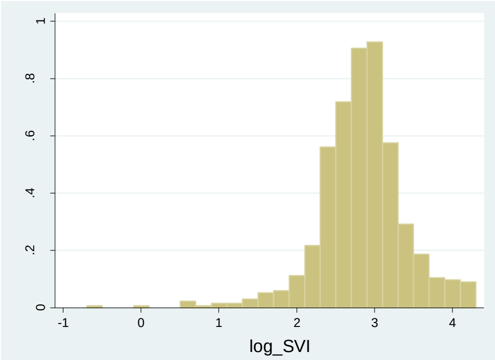

Figure B1.b. Histogram of the log of the SVI.

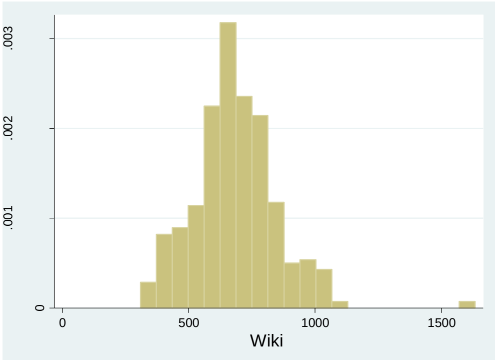

Figure B1.c. Histogram of the Wikipedia page views.

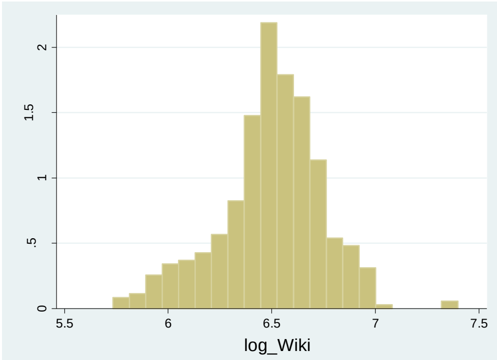

Figure B1.d. Histogram of the log of the Wikipedia page views.

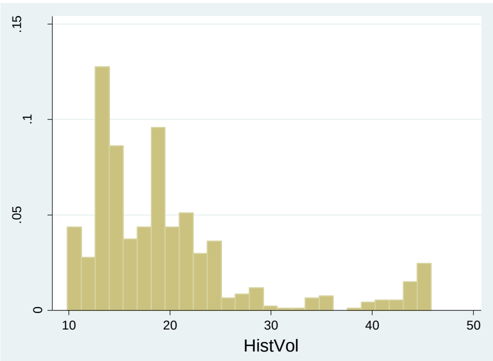

Figure B1.e. Histogram of the historical volatility.

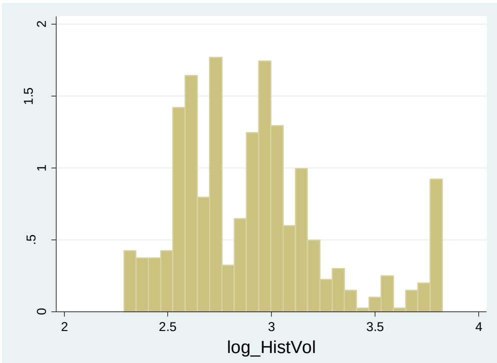

Figure B1.f. Histogram of the log of the historical volatility.

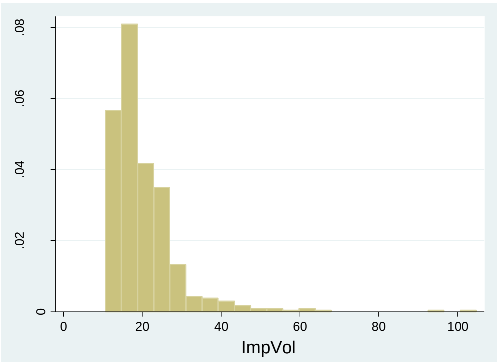

Figure B1.g. Histogram of implied volatility.

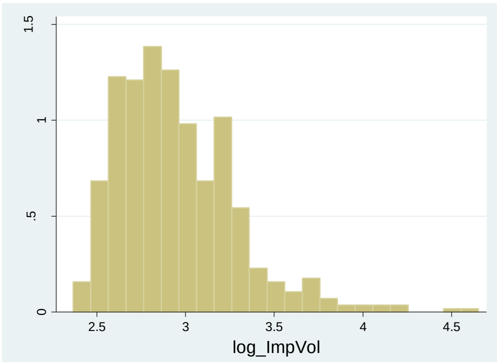

Figure B1.h. Histogram of the log of implied volatility.

Figure B1. Histograms. The figures show the histogram of the (log of the) winsorized SVI, Wikipedia page views and historical and implied volatility.

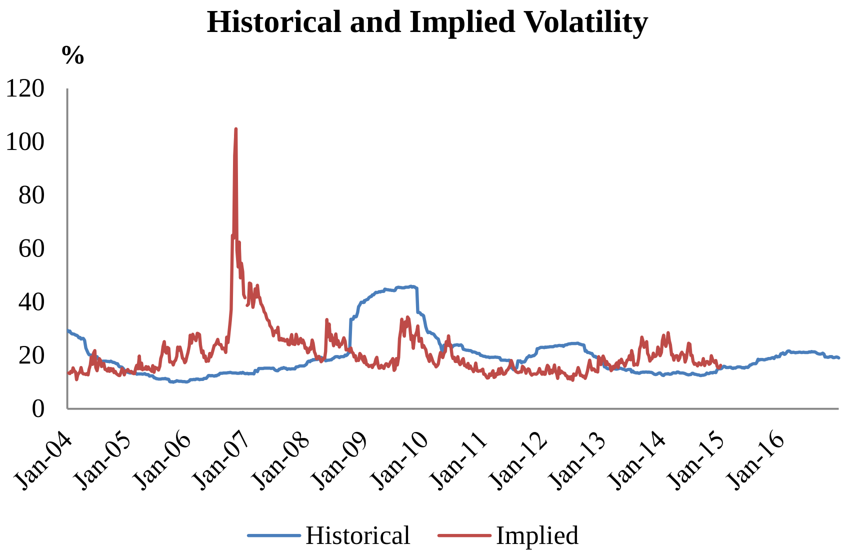

Figure B2. Historical and implied volatility. The graph shows the historical (blue) and implied (red) volatility, expressed in percentages, for the period from January 2004 to December 2016.

Appendix C: Stationarity

Table C1. Augmented Dickey-Fuller test*. The table shows the test statistics and the corresponding p-values of the of the Augmented Dickey-Fuller (ADF) test. The null hypothesis states that the variable is non-stationary. The variables are the log of the winsorized (log_SVI), the Wikipedia page views (log_Wiki), the historical (log_HistVol) and implied (log_ImpVol) volatility, the index returns (Ret) and the trading volume expressed in billions (TradV). The first difference of the log of historical and implied volatility is denoted dlog_HistVol and dlog_ImpVol, respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.001$*

Variable	Test statistic	p-value	Variable	Test statistic	p-value
log_SVI	-14.28	0.000	dlog_HistVol	-23.83	0.000
log_Wiki	-6.74	0.000	dlog_ImpVol	-27.43	0.000
log_HistVol	-1.46	0.554	Ret	-26.05	0.000
log_ImpVol	-3.79	0.003	TradV	-11.04	0.000

Appendix D: The AEX index

Table D1. The AEX index*. The table gives an overview of the 25 companies that are listed on the AEX index as of July 2017. The ticker symbol and the index weighting (in %) are also provided.*

Company	**Ticker symbol**	**Index weighting**	**Company**	**Ticker symbol**	**Index weighting**
Aalberts Industries	AALB	0.72	ING Group	INGA	12.14
ABN AMRO	ABN	1.35	KPN	KPN	1.98
Aegon	AGN	1.63	NN Group	NN	2.16
Ahold Delhaize	ADRNY	4.45	Philips	PHIA	5.99
Akzo Nobel	AKZA	3.98	Randstad Holding	RAND	1.26
Altice	ATC	2.75	RELX Group	REN	3.6
ArcelorMittal	MT	2.73	Royal Dutch Shell	RDSA	14.48
ASML	ASML	8.83	SBM Offshore	SBMO	0.5
Boskalis	BOKA	0.5	Unibail-Rodamco	UL	4.55
DSM	DSM	2.39	Unilever	UNA	15.78
Galapagos	GLPG	0.51	Vopak	VPK	0.59
Gemalto	GTO	0.84	Wolters Kluwer	WKL	2.2
Heineken	HEIA	4.07			

Appendix E: Constructing Forecasting Models

Table E1. Forecasting models for historical volatility. *The table shows the results of the AR(p) and ARDL(p,q) models on the first difference of the log of historical volatility. The variables are the first difference of the log of historical volatility (dlog_HistVol), index returns (Ret) and trading volume expressed in billions (TradV). Robust standard errors are given in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$*

Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
dlog_HistVol$_{t-1}$	0.087**	0.077**	0.067*	0.070*	0.007	-0.012
	(0.039)	(0.036)	(0.035)	(0.036)	(0.049)	(0.054)
dlog_HistVol$_{t-2}$		0.111**	0.104*	0.108*	0.066	0.060
		(0.057)	(0.0545)	(0.055)	(0.053)	(0.056)
dlog_HistVol$_{t-3}$			0.093***	0.095***		
			(0.030)	(0.031)		
dlog_HistVol$_{t-4}$				-0.039		
				(0.028)		
Ret$_{t-1}$					-0.268	-0.272*
					(0.163)	(0.164)
Ret$_{t-2}$						-0.077*
						(0.044)
TradV$_{t-1}$					0.029***	0.023***
					(0.009)	(0.008)
TradV$_{t-2}$						0.005
						(0.008)
Constant	-0.0005	-0.0005	-0.0004	-0.0004	-0.0156***	-0.0154***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.005)	(0.005)
Observations	674	674	674	674	637	637
BIC	-2812.842	-2814.738	-2814.029	-2808.541	-2714.344	-2711.572

Table E2. Forecasting models for historical volatility*. The table shows the results of the ARDL(p,q) models on the first difference of the log of historical volatility, including the lags of SVI. The variables are the first difference of the log of historical volatility (dlog_HistVol), index returns (Ret), trading volume expressed in billions (TradV) and the log of the SVI (log_SVI). Robust standard errors are given in parentheses.* $*\ p < 0.1$, $**\ p < 0.05$, $***\ p < 0.01$

Variable	**(1)**	**(2)**	**(3)**
dlog_HistVol$_{t-1}$	-0.009	-0.014	-0.019
	(0.051)	(0.052)	(0.053)
dlog_HistVol$_{t-2}$	0.058	0.053	0.046
	(0.053)	(0.054)	(0.053)
Ret$_{t-1}$	-0.260*	-0.259*	-0.260*
	(0.158)	(0.157)	(0.156)
TradV$_{t-1}$	0.024***	0.023***	0.022***
	(0.008)	(0.008)	(0.008)
log_SVI$_{t-1}$	0.008***	0.006**	0.006**
	(0.003)	(0.003)	(0.002)
log_SVI$_{t-2}$		0.004	0.002
		(0.003)	(0.003)
log_SVI$_{t-3}$			0.004
			(0.003)
Constant	-0.036***	-0.041***	-0.045***
	(0.011)	(0.014)	(0.016)
Observations	629	629	629
BIC	-2683.332	-2678.910	-2674.335

Table E3. Forecasting models for historical volatility. *The table shows the results of the ARDL(p,q) models on the first difference of the log of historical volatility, including the lags of Wikipedia page views. The variables are the first difference of the log of historical volatility (dlog_HistVol), index returns (Ret), trading volume expressed in billions (TradV) and the log of the Wikipedia page views (log_Wiki). Robust standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p <0.01*

Variable	**(1)**	**(2)**	**(3)**	**(4)**
$dlog\_HistVol_{t-1}$	-0.007	-0.011	0.007	0.011
	(0.079)	(0.080)	(0.072)	(0.069)
$dlog\_HistVol_{t-2}$	0.049	0.070	0.077	0.066
	(0.070)	(0.061)	(0.060)	(0.060)
Ret_{t-1}	-0.339	-0.328*	-0.328*	-0.321*
	(0.213)	(0.197)	(0.194)	(0.184)
$TradV_{t-1}$	0.035*	0.025*	0.030**	0.028**
	(0.019)	(0.013)	(0.015)	(0.012)
log_Wiki_{t-1}	0.003	0.033	0.032	0.034
	(0.008)	(0.025)	(0.024)	(0.027)
log_Wiki_{t-2}		-0.037	-0.056*	-0.057*
		(0.025)	(0.032)	(0.034)
log_Wiki_{t-3}			0.025*	0.035
			(0.014)	(0.026)
log_Wiki_{t-4}				-0.013
				(0.020)
Constant	-0.037	0.013	-0.021	-0.007
	(0.057)	(0.0427)	(0.0470)	(0.044)
Observations	393	393	393	393
BIC	-1594.183	-1599.341	-1598.183	-1593.341

Table E4. Forecasting models for historical volatility. *The table shows the results of the AR(p) and ARDL(p,q) models on the first difference of the log of historical volatility, including the lags of SVI and Wikipedia page views. The variables are the first difference of the log of historical volatility (dlog_HistVol), index returns (Ret), trading volume expressed in billions (TradV) and the log of the SVI (log_SVI) and the Wikipedia page views (log_Wiki). Robust standard errors are given in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$*

Variable	**(1)**	**(2)**	**(3)**	**(4)**
dlog_HistVol$_{t-1}$	-0.034	-0.036	-0.017	-0.014
	(0.085)	(0.084)	(0.076)	(0.073)
dlog_HistVol$_{t-2}$	0.038	0.059	0.066	0.058
	(0.074)	(0.064)	(0.062)	(0.062)
Ret$_{t-1}$	-0.319	-0.310*	-0.310*	-0.305*
	(0.200)	(0.186)	(0.184)	(0.175)
TradV$_{t-1}$	0.026*	0.018	0.022*	0.021*
	(0.015)	(0.011)	(0.012)	(0.011)
log_SVI$_{t-1}$	0.014**	0.013**	0.013**	0.013**
	(0.006)	(0.005)	(0.005)	(0.005)
log_Wiki$_{t-1}$	0.008	0.036	0.034	0.036
	(0.009)	(0.025)	(0.024)	(0.027)
log_Wiki$_{t-2}$		-0.034	-0.053*	-0.054*
		(0.023)	(0.031)	(0.033)
log_Wiki$_{t-3}$			0.026*	0.033
			(0.014)	(0.025)
log_Wiki$_{t-4}$				-0.009
				(0.019)
Constant	-0.106	-0.055	-0.091	-0.0780
	(0.076)	(0.054)	(0.063)	(0.055)
Observations	393	393	393	393
BIC	-1601.546	-1605.312	-1604.619	-1599.232

Table E5. Forecasting models for implied volatility. *The table shows the results of the AR(p) and ARDL(p,q) models on the first difference of the log of implied volatility. The variables are the first difference of the log of implied volatility (dlog_ImpVol), index returns (Ret) and the trading volume expressed in billions (TradV). Robust standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p <0.01*

Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
$dlog\_ImpVol_{t-1}$	-0.141***	-0.141***	-0.141***	-0.115**	-0.163	-0.155
	(0.048)	(0.051)	(0.051)	(0.054)	(0.111)	(0.115)
$dlog\_ImpVol_{t-2}$		-0.002	-0.004			
		(0.071)	(0.074)			
$dlog\_ImpVol_{t-3}$			-0.014			
			(0.056)			
Ret_{t-1}				-2.312***	-2.279***	-2.285***
				(0.177)	(0.181)	(0.171)
Ret_{t-2}					-0.218	-0.213
					(0.404)	(0.406)
Ret_{t-3}						-0.215
						(0.200)
$TradV_{t-1}$				0.029	0.082***	0.079***
				(0.019)	(0.022)	(0.021)
$TradV_{t-2}$					-0.080***	-0.96***
					(0.029)	(0.029)
$TradV_{t-3}$						0.013
						(0.034)
Constant	0.0004	0.0004	0.0004	-0.0155	-0.0001	0.0029
	(0.004)	(0.004)	(0.004)	(0.010)	(0.013)	(0.014)
Observations	562	562	562	570	570	570
BIC	-956.099	-949.769	-943.543	-1316.483	-1324.964	-1322.823

Table E6. Forecasting models for implied volatility. *The table shows the results of the ARDL(p,q) models on the first difference of the log of implied volatility, including the lags of SVI. The variables are the first difference of the log of implied volatility (dlog_ImpVol), index returns (Ret), trading volume expressed in billions (TradV) and the log of the SVI (log_SVI). Robust standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p <0.01*

Variable	**(1)**	**(2)**	**(3)**
$dlog\_ImpVol_{t-1}$	-0.165	-0.166	-0.167
	(0.112)	(0.111)	(0.112)
Ret_{t-1}	-2.272***	-2.274***	-2.273***
	(0.181)	(0.181)	(0.182)
Ret_{t-2}	-0.221	-0.225	-0.229
	(0.406)	(0.406)	(0.406)
$TradV_{t-1}$	0.079***	0.081***	0.079***
	(0.022)	(0.022)	(0.022)
$TradV_{t-2}$	-0.081***	-0.079***	-0.076***
	(0.029)	(0.028)	(0.027)
log_SVI_{t-1}	0.006	0.009	0.010
	(0.007)	(0.008)	(0.008)
log_SVI_{t-2}		-0.007	-0.005
		(0.009)	(0.009)
log_SVI_{t-3}			-0.005
			(0.009)
Constant	-0.014	-0.007	-0.002
	(0.025)	(0.030)	(0.034)
Observations	570	570	570
BIC	-1306.592	-1300.921	-1294.916

Table E7. Forecasting models for implied volatility. *The table shows the results of the ARDL(p,q) models on the first difference of the log of implied volatility, including the lags of Wikipedia page views. The variables are the first difference of the log of implied volatility (dlog_ImpVol), index returns (Ret), trading volume expressed in billions (TradV) and the log of the Wikipedia page views (log_Wiki). Robust standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p <0.01*

Variable	**(1)**	**(2)**	**(3)**
dlog_ImpVol$_{t-1}$	-0.190***	-0.191***	-0.190***
	(0.061)	(0.059)	(0.059)
Ret$_{t-1}$	-2.468***	-2.468***	-2.468***
	(0.141)	(0.142)	(0.142)
Ret$_{t-2}$	-0.105	-0.107	-0.110
	(0.225)	(0.220)	(0.220)
TradV$_{t-1}$	0.071**	0.071***	0.072***
	(0.028)	(0.027)	(0.027)
TradV$_{t-2}$	-0.071**	-0.072**	-0.067**
	(0.034)	(0.034)	(0.034)
log_Wiki$_{t-1}$	0.005	0.007	0.006
	(0.014)	(0.024)	(0.025)
log_Wiki$_{t-2}$		-0.002	-0.016
		(0.030)	(0.034)
log_Wiki$_{t-3}$			0.021
			(0.026)
Constant	-0.035	-0.032	-0.061
	(0.095)	(0.113)	(0.115)
Observations	407	407	407
BIC	-1042.706	-1036.704	-1031.446

Table E8. Forecasting models for implied volatility. *The table shows the results of the ARDL(p,q) models on the first difference of the log of implied volatility, including the lags of SVI and Wikipedia page views. The variables are the first difference of the log of implied volatility (dlog_ImpVol), index returns (Ret), trading volume expressed in billions (TradV) and the log of the SVI (log_SVI) and the Wikipedia page views (log_Wiki). Robust standard errors are given in parentheses. * p < 0.1, ** p < 0.05, *** p <0.01*

Variable	**(1)**	**(2)**	**(3)**
$dlog\_ImpVol_{t-1}$	-0.194***	-0.194***	-0.192***
	(0.062)	(0.060)	(0.060)
Ret_{t-1}	-2.456***	-2.456***	-2.455***
	(0.139)	(0.141)	(0.140)
Ret_{t-2}	-0.115	-0.115	-0.119
	(0.228)	(0.223)	(0.223)
$TradV_{t-1}$	0.065**	0.065**	0.066**
	(0.029)	(0.027)	(0.028)
$TradV_{t-2}$	-0.073**	-0.073**	-0.069**
	(0.033)	(0.033)	(0.033)
log_SVI_{t-1}	0.008	0.008	0.008
	(0.008)	(0.008)	(0.008)
log_Wiki_{t-1}	0.008	0.009	0.008
	(0.015)	(0.024)	(0.024)
log_Wiki_{t-2}		-0.001	-0.017
		(0.030)	(0.035)
log_Wiki_{t-3}			0.022
			(0.026)
Constant	-0.073	-0.072	-0.103
	(0.107)	(0.128)	(0.128)
Observations	407	407	407
BIC	-1037.765	-1031.758	-1026.570

Appendix F: Granger-Causality Test

Table F1. Granger-causality test*. The table shows the F-statistics of the Granger-causality test of the model including lags of the winsorized SVI (1), including lags of the Wikipedia page views (2) and including lags of both search queries (3). The variables are the log of the SVI (log_SVI) and the Wikipedia page views (log_Wiki). It tests whether the coefficients of the lags of the search queries are jointly significant. The null hypothesis states that the variables have no predictive power.*
** $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$*

Variable	HISTORICAL VOLATILITY			IMPLIED VOLATILITY		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
log_SVI	7.8***		6.95***	0.64		0.14
log_Wiki		1.67	1.17		0.38	0.43
log_SVI + log_Wiki			2.91**			0.23