SIGVE DALSBØ LOHNE (9114) & SINDRE M. A. SALTVIK (9027)
SUPERVISOR: PETER MOLNÁR

The Invasion of Ukraine and Stock Returns

Research Question:

"Is Google searches for Ukraine correlated with stock returns on the Oslo Stock Exchange?"

Master thesis, 2023

Master of Science in Business Administration

University of Stavanger Business School

Specialization: Applied Finance

Acknowledgements

We would like to express our gratitude to everyone who has contributed to the completion of this master's thesis. Especially to Peter Molnár, our dedicated supervisor, for his guidance, expertise, and support throughout the entire research process. His insightful feedback and constructive suggestions have significantly shaped the direction and quality of this work.

We would also like to acknowledge our deep interest in the stock markets and our curiosity regarding the potential impact the invasion of Ukraine has had on global stock markets. This fascination led us to explore the dynamics of our national stock exchange, the Oslo Stock Exchange, as a case study. By investigating the relationship between geopolitical events and stock market behaviour, we aimed to contribute to understanding financial market dynamics during periods of heightened geopolitical tension.

Abstract

This master's thesis examines how Norway's stock market volatility is affected by the continuing invasion of Ukraine. Our study attempts to offer useful insights for investing strategies and decision-making by investigating the relationship between investor attention and market dynamics during times of international war. We examine weekly changes in stock returns using data from the financial markets, and we applied an abnormal search volume index based on the volume of searches for the term "Ukraina" in Google. The research adds to the body of knowledge, advances our knowledge of how geopolitical crises can impact financial markets, and has applications for investors, analysts, scholars and the general public. This study emphasizes the significance of considering geopolitical issues when analysing market risks and provides methods for mitigating investments in uncertain times. We aim to provide a framework for evaluating the effects of geopolitical events on financial markets, aiding stakeholders in navigating risks associated with geopolitical tensions. We can conclude that there is a correlation between stock returns for the 20 largest companies on the Oslo Stock Exchange and Google search results for the word Ukraine. Therefore, Google Trends search volume may be used as a proxy to determine stock returns in times of geopolitical conflict.

Contents

Acknowledgements i

Abstract i

1 Introduction 1

2 Literature Review 2

3 Data Description 3

4 Methodology 5

4.1 Regression Input 5

4.2 Stock Volatility 6

4.3 Abnormal Search Volume Index 7

4.4 Extension of the Analysis 7

4.5 Choice of Effects 8

4.6 Regression Diagnostics 9

4.7 Limitations 11

5 Results 13

5.1 Panel Data Regression 13

5.2 Applying Lagged Variables 15

5.3 Geographical Differences 16

5.4 Energy-Specific Results 17

5.5 Observation Frequency Differences 18

6 Conclusion 20

References 22

1 Introduction

In recent years, the field of finance has witnessed a surge in research exploring the relationship between online search activity and financial markets. With the advent of digital technologies and the availability of vast amounts of online data, scholars have sought to understand the information content and predictive power of internet search behaviour in financial decision-making. One area of interest in this domain is the impact of geopolitical events on stock market volatility (Zhang et al., 2023).

The ongoing invasion of Ukraine by external forces has unfolded as a significant geopolitical event, capturing global attention, and evoking widespread concern. This development has prompted researchers to delve into the potential repercussions of the crisis on financial markets, seeking to unravel the relationship between investor sentiment and market dynamics in the face of such international conflicts (Sandoval & Franca, 2012).

This study aims to contribute to the existing literature by examining the impact of the Ukraine invasion on stock market volatility in Norway. As a nation, Norway has substantial economic ties to many European countries and an active stock exchange, providing a suitable context for investigating the link between geopolitical events and financial market behaviour. In particular, we employ a novel approach by utilizing Google Trends data, specifically the search volume for the term "Ukraina," as a proxy for investor attention and sentiment related to the invasion.

The methodology section of this paper outlines the key steps involved in conducting our analysis. Section (4.1) details the regression input, where we employ a panel data regression model to investigate the relationship between weekly changes in stock prices and Google Trends search volume. Section (4.2) elaborates on the calculation of stock volatility as our dependent variable, emphasizing the use of the Garman-Klass (1980) volatility estimator. Additionally, section (4.3) explains the methodology behind the creation of the Abnormal Search Volume Index (ASVI) to capture investor attention. Section (4.4) discusses potential extensions to enhance the robustness of our findings, including the incorporation of lagged explanatory variables, exploring geographical factors, considering additional energy price indices, and analysing daily data. Subsequently, sections (4.5) and (4.6) delve into the choice of effects in the regression model and the diagnostic tests conducted to ensure the validity of our results. Finally, section (4.7) addresses the possible limitations of our research, emphasizing the need for cautious interpretation.

The significance of our study lies in its potential to shed light on the impact of a geopolitical crisis on stock market volatility, specifically in the context of the Ukraine invasion. By analysing the relationship between Google Trends search volume and stock prices of the 20 largest companies on the Oslo Stock Exchange, we aim to provide valuable insights into investor sentiment and market behaviour during times of heightened geopolitical tension. Our research not only contributes to the academic literature on the topic but also offers practical implications for further research regarding similar topics.

We aim to provide insights into the impact of the Ukraine invasion on stock market volatility in Norway. By utilizing Google Trends data and employing a robust regression analysis, we seek to uncover the relationship between investor sentiment captured by search volume and stock price fluctuations. The findings of this research have the potential to deepen the understanding of the dynamics between geopolitical events and financial markets.

2 Literature Review

We started by looking into an article by Bijl et al. (2016) that examines the use of Google Trends search volume to forecast stock returns. They state that previous studies have shown that high Google search volume can predict high stock returns in the coming two weeks. In addition to this, Heiden & Hamid (2015) wrote an article where they tried to forecast stock market volatility using Google Trends. This ties together with an article by Wang et al. (2018), that is examining if it is possible to use Google Trends to predict the direction of stock markets. We decided to base our study on these papers where we want to examine if there is a correlation between the Norwegian Google search volume for “Ukraina” and stock returns for the 20 largest companies on the Oslo Stock Exchange (OSE).

Moreover, Liu et al. (2019) writes an article regarding how Google Trends can be used to analyze the impact of disaster events on stock prices. Nguyen & Bui (2019), build on the same Google Trends theories and investigate how search results can affect stock market volatility in a developing economy.

Bijl et al. (2016) cite an article by Preis et al. (2010) which studies the correlation between search volume for company names and stock returns for the given company. However, Preis et al. (2010) did not find a significant correlation. Instead, they found that search volume could be utilized to predict future trading volume. Preis et al. (2010) further explain that several other factors affect returns, which could be the reason for the lack of significant results. They mention that the volume and number of transactions also affect stock market fluctuations and that this research requires a larger amount of data. In a later article, Preis et al. (2013) investigated if general financial search terms could be used to predict market movements. They tested a strategy where a market portfolio is bought or sold based on search volumes for certain relevant terms, which outperformed the market index by 310 % over seven years.

In addition to this, Fama (1965a) states that historic stock price patterns tend to reoccur in the future. He further claims that stocks are driven by information, and since new information is unknown, future stock prices will be random or unpredictable. Many articles agree that stock prices are somewhat unpredictable, and Salisu et al. (2020) present what is called good media news and bad media news, and how it affects returns. Moreover, they add how Google Trends also influences investors' decision-making when it comes to making financial investments. According to Salisu et al. (2020) single- and multi-factor predictive models will overall outperform what Fama (1965b) states as the "random walk". This is a statement against what Fama presented in his article saying that at the end of the day, stocks are random and cannot be predicted.

To elaborate further on how search volume can affect decision making as stated by Salisu et al. (2020), Swamy et al. (2019) wrote an article about investor attention and Google Search Volume Index. Their research is based on the theory that market prices are affected by the available information. In addition, they examined how search intensity can generate useful information for investors. The article cites Joseph et al. (2011) in their research paper, they studied the search volume for distinct stock tickers and compared the results with the price of that specific stock.

3 Data Description

The data used in this study consists of the weekly stock prices of the 20 largest companies listed on the OSE, based on market capitalization per February 1st, 2023. We have collected stock data dating back to April 7th, 2019, as the newest company in the data pool had their initial public offering then. The collected data encompasses the period up until December 31st, 2022, allowing for a comprehensive analysis within this timeframe.

To test the effects of the invasion of Ukraine on the OSE this study comprises the Google Trends search volume for the term "Ukraina" from the Norwegian database. Due to the nature of the Google Trends and the OSE data, we decided to use weekly data based on the research done by Jara-Diaz & Rosales-Salas (2015). This is because Google Trends can only provide daily data within a timeframe of 9 months, which means the data must be extracted individually for each period and thereafter be merged. On the other hand, the OSE is only open between Monday and Friday and therefore it can be problematic to compare it to Google Trends which has available data from Monday to Sunday. By extracting weekly data, we can better align the data from Google Trends with the time frame of the stock data.

The Google Trends data is presented as a search volume index (SVI) from 1 to 100, indicating the search volume. For any given timeframe, whether a day, a week or a month, the highest search volume is represented by the number 100. This indicates that the corresponding period had the highest search volume among the selected timeframes.

We opted to focus on the 20 largest companies listed on the OSE and selected a comprehensive sample of companies from different sectors to enable us to examine potential variations in the effects of the conflict across these sectors. Many of the larger companies in our sample are multinational corporations with operations across various European countries, which may have exposed them to a greater impact of the invasion than domestic firms.

The decision to focus on the Norwegian market was motivated by the potential to examine the impact of the conflict on countries not directly involved in it. Moreover, given Norway's significant role in the oil, gas, and energy sectors, it was deemed worthwhile to investigate potential impacts on Norwegian firms in these industries. Hence, the Norwegian database was a natural choice for analysing Google Trends search results. Notably, we used the term "Ukraina" instead of "Ukraine" to explore search volume patterns, as the former term has a considerably higher volume of searches than the latter in Norway.

To further improve our research, we will look into the word “Ukraine” on the worldwide database in Google Trends as comparable to the results received by the Norwegian database counterpart. Due to our interest and the effect the invasion has had on energy prices across the world, we will also look at energy companies individually with the use of a dummy variable in the data set and compare the results.

All the data collected for this study is publicly available, and we did not conduct any data collection ourselves. To collect stock data, we decided to use Yahoo Finance (Yahoo Inc., 2023) because it provided the data in a standardized format that is easy to use and manipulate for our research. Additionally, it offers the option to download data in the required timeframe, which for us is weekly data. Yahoo Finance had all the data for the different stocks available, making it a comprehensive source for our study. For search data collection, we utilized Google Trends (Google LLC, 2023), which is currently the most extensive search engine available. We chose it because of its extensive search volume and popularity.

4 Methodology

In our methodology section, we will go through how we set up the regression in section (4.1). Further, we discuss how we measure stock volatility, and our return variable is calculated in section (4.2). After that, we explain how we utilize the Google Trends data in section (4.3). Additionally, we will present various approaches to extend our analysis in section (4.4). Next, we want to present our choice of effects in our regression model in section (4.5) and address the various diagnostic tests we have conducted in section (4.6). Then, we have possible limitations that could limit our research and results in section (4.7).

4.1 Regression Input

Our analysis uses a panel data regression to model the relationship between weekly changes in stock prices for the 20 largest companies on the OSE and Google Trends search volume. Specifically, we utilized the volatility of weekly stock prices for a sample of Norwegian companies as the dependent variable, with Google Trends search volume for the term "Ukraina" serving as the main explanatory variable. We attempted to identify any potential nonlinearities in the relationship between these two variables by applying a regression using volatility measures of stock and search data. This strategy is especially relevant since logarithmic returns and volume can offer a better match to the data than simple returns and volume of the stock returns and search volume (Hudson & Gregoriou, 2015).

To run our regression analysis, we have created panel data containing 5 variables: “Week”, “Stock”, “Return”, “ASVI”, and “Energy”. As explained in section (3), we have decided to use weekly data instead of daily due to a better representation of the movement of stock prices compared to search volume on Google Trends, which is represented by the "Week" variable. The "Stock" variable represents every stock included in the analysis individually, and the "Return" variable represents the weekly volatility of returns. "ASVI" stands for Abnormal Search Volume Index which will be explained in more detail later in section (4.3). Lastly, "Energy" is a dummy variable ranging from 0 to 1 indicating if a stock is an energy stock or not.

4.2 Stock Volatility

In our model, the variable "Return" is a measurement of the individual stock volatility. Molnár (2012) argues that range-based volatility estimators provide a more significant precision. After assessing various methods described by various researchers, he concludes that the work of Garman-Klass (1980) is the best estimator.

$$c_t = \ln(C_t) - \ln(O_t) \tag{1}$$

$$h_t = \ln(H_t) - \ln(O_t) \tag{2}$$

$$l_t = \ln(L_t) - \ln(O_t) \tag{3}$$

$$\sigma_t^2 = 0.5(h_t - l_t)^2 - (2\ln 2 - 1)c_t^2 \tag{4}$$

Formula (1), (2), and (3) calculate various volatility parameters based on close (C), high (H), and low (L) prices of stocks in period t with weekly observations. These will be applied in the Garman-Klass volatility estimator in formula (4). A simpler approach would have been to solely use c^2 as an estimator, but the problem with this estimator is that it would have been too noisy. Garman-Klass managed to implement volatility parameters from both high and low prices in addition to close prices, which provides additional information about the volatility of a stock (Molnár, 2012). Hence, we decided to use the Garman-Klass approach in our analysis as an estimator of stock return volatility.

4.3 Abnormal Search Volume Index

Google Trends gives us a sense of how popular various happenings and themes are based on geographical preferences. In our case, we use the search word "Ukraina" to present how popular the invasion of Ukraine is for Norwegian inhabitants. Because stocks are publicly traded, we believe that Google Trend search indices are representative since it represents the interest of every possible buyer in the region and more specifically the geographical factor, where it will be more common for Norwegians to buy stocks traded on the OSE.

Due to the nature of Google Trends data, we cannot simply apply the raw SVI to our model. As mentioned earlier, the SVI varies from 0-100, where its values depend on the given period the data is extracted from. Da et al. (2011) propose a solution for this problem with the ASVI. The formula (5) uses the logarithm of SVI on week t and subtracts the logarithmic median of the prior 8 weeks. The research concludes that it captures investor attention in a timelier fashion and measures the attention of retail investors. Due to problems with infinite values for observations where ASVI < 1 we added a constant of + 1 to every observation, therefore will the ASVI values in our analysis range from 1-101.

$$ASVI_t = \ln(SVI_t) - \ln[Med(SVI_{t-1}, \dots, SVI_{t-8})] \tag{5}$$

4.4 Extension of the Analysis

In addition to the primary analysis, there are several potential extensions that we will present in the results which will provide valuable insights and enhance the robustness of our findings. Firstly, we believe it is beneficial to incorporate lagged explanatory variables, specifically lagged values of the ASVI variable. Lagged variables capture the effect of past values on the current outcome, allowing us to examine any temporal patterns or lags in the relationship between search volume and stock returns. By including lagged variables, we can assess whether the historical search activity, as captured by the lagged ASVI, has a significant impact on the current stock returns.

Furthermore, exploring the geographical factor in our analysis could yield valuable insights. We expand our investigation by comparing the results obtained from a Norwegian search for "Ukraina" with a worldwide search for "Ukraine". This approach allows us to examine whether there are divergent effects or variations in the relationship between search volume and stock returns across different geographic regions. It provides an opportunity to assess whether the local context or global factors play a significant role in shaping the relationship.

Additionally, given our interest in energy stocks, it would be advantageous to conduct a closer examination of this sector. To support our analysis, we will include additional indices related to energy prices, such as crude oil or natural gas prices. These indices can serve as proxies for the overall energy market conditions, providing a more comprehensive understanding of the relationship between energy stocks and search volume. By incorporating energy price indices, we can account for the influence of broader market factors on the performance of energy stocks and assess their interaction with search volume.

Lastly, altering the frequency of our data from weekly to daily can offer a different perspective on the relationship between search volume and stock returns. Daily data provides a higher level of granularity, capturing more immediate and short-term fluctuations in search activity and stock prices. By analysing daily data, we can explore whether the relationship between search volume and stock returns manifests differently at a finer time resolution and investigate any intraday patterns or dynamics that may be missed when working with weekly data.

Incorporating these extensions into our analysis will enrich our understanding of the relationship between ASVI and stock returns, provide additional insights into the energy sector, examine the geographical context, and explore the impact of different data frequencies. These extensions have the potential to uncover further nuances and enhance the robustness of our findings.

4.5 Choice of Effects

In a panel data regression, researchers often have the option to choose between two main approaches: random effects (RE) and fixed effects (FE). Each approach entails a set of assumptions that should be carefully considered in the context of the dataset at hand. Clark & Linzer (2014) provide valuable insights into this decision-making process, highlighting the trade-offs involved. RE models have the advantage of reducing the variance of estimates for coefficients of interest, but they can introduce bias under certain conditions. On the other hand, FE estimates are unbiased, but they may be subject to high sample dependence. Considering these considerations, we have opted to employ the FE approach in our analysis to mitigate potential biases and effectively capture within-entity variations.

There is also a third approach to consider: pooled effects, also known as pooled Ordinary Least Squares (OLS). Collischon & Andreas (2020) argue that FE models are generally as good as, if not better than, pooled OLS models. They encourage scholars to prioritize the use of FE models whenever possible while acknowledging the limitations that should be considered. On the other hand, Basumatary & Devi (2022) conducted a study comparing the results of a regression using pooled OLS and a regression using FE. Their findings revealed contrasting impacts of various factors on the outcome variable. Additionally, they used a Wald test statistic to determine the better method, and due to its significance level being below 0.05, they preferred the FE model over pooled OLS. Given the divergent results and the recognition that both the FE model and the pooled OLS model have their merits and limitations, we believe it is crucial to run both types of regressions in our analysis. This approach will provide a comprehensive understanding of the relationship between the independent and dependent variables while considering the potential biases and effects associated with each modelling approach.

4.6 Regression Diagnostics

In this section, we present the regression diagnostics conducted to assess the validity and reliability of our regression model. The diagnostic tests aim to evaluate the presence of potential issues and to ensure the robustness of our regression results and the validity of our conclusions. We have chosen to conduct tests for the following potential issues: multicollinearity, heteroscedasticity, autocorrelation, and deviations from normality.

Multicollinearity

We started by determining whether multicollinearity existed, which is an approach for identifying correlation or linear relationships among the variables in our regression model. We may assess the potential effects on the dependability and accuracy of our results by evaluating the multicollinearity, which is the degree of dependence between two or more independent variables. Based on Marsh et al. (2004), multicollinearity is a well-known problem that can undermine or ruin the validity of a regression result; therefore, we chose to test for it. Multicollinear evidence can result in outcomes that are deceptive as well as incorrect interpretations, according to Marsh et al.

We decided to use a Variance Inflation Factor (VIF) approach to test, which is used to check if there is proof of multicollinearity in our explanatory variables. VIF is based on the R-squared value one obtains when regressing a predictor on all of the other predictors in the analysis (Miles, 2005).

Typically, a VIF value below 5 or 10 indicates minimal or no evidence of multicollinearity, which is considered favourable for the validity of our results. Such values suggest that the explanatory variables exhibit a sufficient level of independence and do not introduce substantial redundancy or overlap in their predictive power.

Heteroscedasticity

The next test we decided to run is the Studentized Breusch-Pagan test for heteroscedasticity. We decided to test for this because heteroscedasticity examines if the variance of the errors that occur, is constant or not. If heteroscedasticity is present, it means that the variance of errors is not constant. In a scatterplot, the observations would be further away from the average slope. If the variance of the errors is constant, we have evidence of homoscedasticity.

When there is evidence of heteroskedasticity, the standard errors that have been estimated could show evidence of biasedness. This could affect the significance of the independent variable used for our research paper. According to Kaufman (2013), if all other necessary assumptions are present, then unbiasedness could still be considered. Along with many other tests, the Breusch-Pagan test for heteroscedasticity is considered a confirmation to help increase the validity of the regression result. However, it does not conclusively mean that the result is invalid. Further tests are needed to further strengthen or weaken the methodology.

Autocorrelation

Following the Breusch-Pagan test, we wanted to test for autocorrelation, which refers to the correlation between error terms at different periods. Autocorrelation occurs when there is a systematic relationship or pattern in the residuals of a regression model that is not captured by the explanatory variables. According to Griffith (1992), it measures when certain observations in our dataset could be dependent on the previous observations in the same dataset, rather than being independent.

The fact that we are examining multiple stock data over several years on the OSE, the presence of autocorrelation can have important implications. To detect and assess autocorrelation in our regression model, we employed the Breusch-Godfrey/Woolridge test. This test extends the traditional Durbin-Watson test to account for the panel structure of the data. By conducting this approach, we can conclude if there is evidence of serial correlation in the residuals after accounting for the explanatory variables in the model.

Normality

The final test we decided to run is for normality. For our research paper, it is very important to test for normality since stock returns are usually non-normal distributions. They often present a larger, thick tail with skewness. Karoglou (2009) explains that non-normality distributions under GARCH conditions (General Autoregressive Conditional Heteroskedasticity) could increase for the stock market. The GARCH process happens when financial markets become highly volatile due to a financial crisis occurring. On the other hand, financial markets become less volatile when the general economy is experiencing steady growth. Due to this, we found it fitting to test for normality since the invasion of Ukraine may have caused financial markets to move in a more unstable fashion.

4.7 Limitations

Throughout our research and examination of the impact of the Ukraine invasion on the Norwegian financial market, we have identified certain potential limitations that may influence the interpretation of our results.

External Factors

One noteworthy limitation that occurs is the interdependence and correlation observed among financial markets. This phenomenon implies that our results may be susceptible to inflation, as the intercorrelation among markets can introduce a level of interconnectedness that affects the observed outcomes. This is a theory presented by Sandoval et al. (2012). They examine how financial markets crash at the same time during a world crisis.

Sample Size

In our study, we chose to focus on the 20 largest companies publicly traded on the Oslo Stock Exchange, as they are relevant to our economy and the circumstances surrounding the invasion. However, the sample size we used might not be substantial enough to return significant results. This is because the sample size is restricted to a single market and may not reflect how the invasion may have affected other international markets. In a research paper by McClain et al. (1996), they experienced different results when increasing or decreasing their sample size.

Additionally, several Norwegian companies are traded in many different countries and have partnerships with various international firms, which could have influenced our correlation results. Consequently, the outcomes of our study may lack generalizability since it is specific to the Norwegian market and may not apply to other countries or companies.

Search Behaviour

Further, we believe that the reliance on the Google Trends search volume may not be able to properly reflect market behaviours. People's search intentions will vary a lot and most people are only searching for Ukraine out of personal interests and not to make any financial investments that could affect stock volume and returns. Therefore, the public's search results will be affected by various other factors beyond the focus of our research.

Correlation and Causation

Lastly, we cannot directly assume that correlation drives causation for our research. Since we are examining the relationship between search volume and stock returns, we have only two variables that are being tested against each other. Regardless, there are still several factors not included in our research that very much affects both the search volume and stock returns. This means that a correlation might not be a result of the invasion of Ukraine, but other factors could have affected our results.

5 Results

This section contains our regression and test results. The first section (5.1) incorporates our regression results after conducting the test for the 20 largest companies on the Oslo Stock Exchange. The second section (5.2) implements lagged variable of ASVI in the regression model to investigate the eventual changes in the result. The third section (5.3) investigates how the result of the regression is affected when changing the geographics in the Google Trends search data. In the fourth section (5.4) we are narrowing down the analysis to only look into energy companies with extra explanatory variables to strengthen the regression. Lastly in the fifth section (5.5), we change the weekly data to daily data to present any differences in the results.

5.1 Panel Data Regression

Our analysis contains time series of 20 companies with 195 weekly observations. We ran two panel data regressions, one with pooled OLS and one with FE.

We will now present our base regression model. The dependent variable is the stock returns (R) which are presented in section (4.2) formula (4). The explanatory variables are ASVI, the individual stock identifier (S) which allows for differentiating the effects of different stocks and the energy dummy (E) which indicates if a stock is an energy stock or not. All these are explained further in the methodology. The model aims to capture the relationship between the return of a stock, the ASVI, and stock-specific effects, and is specified as followed:

$$R_{i,t} = \alpha + \beta_1 \cdot ASVI_t + \beta_2 \cdot S_{i,t} + \beta_3 \cdot E_{i,t} \cdot ASVI_t + \varepsilon_{i,t} \tag{6}$$

Where:

- $R_{i,t}$ represents the stock return "i" at time "t".
- $ASVI_t$ represents the abnormal search volume index at time "t".
- $S_{i,t}$ is a categorical variable representing the identifier for stock "i" at time "t".
- $E_{i,t}$ represents the energy dummy at time "t".
- β_1, β_2, and β_3 are the respective coefficients for the ASVI, S, and E variables.
- $\varepsilon_{i,t}$ represents the error term or residuals.

When examining the results from the regression in Table (1) we can see that both applied effects give a significant result towards the ASVI, where $p < 0.05$. Both models yield similar coefficients, but lower standard errors with FE. Dallal (2001) implies that the idea behind pooling is to eliminate effects that are not statistically significant from the model so the model can be refitted. Baltagi & Griffin (1984) suggests that the pooled OLS estimator is robust when analysing long-run time series, and in contrast, will FE be a better estimator for short-run time series where they offer poor long-run estimates due to large percentage bias. With our time frame being under four years, we would define it as a short time series, which loses robustness of results and the identification of long-term trends compared to utilizing a longer time frame.

Regression Results

	Dependent variable:	
	Return	
	Pooled	Fixed
	(1)	(2)
ASVI	0.158^{**}	0.157^{**}
	(0.071)	(0.062)
ASVI:Energy	-0.119	-0.107
	(0.226)	(0.196)
Constant	3.865^{***}	
	(0.041)	
Observations	3,900	3,900
R^2	0.001	0.002
Adjusted R^2	0.001	-0.004
F Statistic	2.452^{*} (df = 2; 3897)	3.239^{**} (df = 2; 3878)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 1: Regression results from the base model.

After examining both models, our results suggest that there is a significant correlation between the Norwegian search volume of "Ukraina" and the stock returns on the OSE. The coefficients suggest that on days when Google searches are higher, the stocks receive positive returns. We consider FE to be a better-suited model for our approach based on the findings of cited researchers (Clark & Linzer, 2014, Collischon & Andreas, 2020, Basumatary & Devi, 2022, Baltagi & Griffin, 1984). It is interesting to compare our results with the findings of Preis et al.

(2010), where they found no significant correlation between stock returns and the search for stock names in a given period.

5.2 Applying Lagged Variables

Our base regression model in section (5.1) yielded significant results, but we also want to implement lagged variables to see if there are changes in results. We added a lagged variable of ASVI into the regression, where the lag is equivalent to t-1. We decided to only test for one lagged variable because the result would not differ with more delay due to the period being too long. The reason behind using lagged variables is to try to capture a possible time delay, where the effects from the ASVI do not have enough time to manifest and have an impact on the OSE. We applied the lagged variables for both the standard ASVI variable and in addition the ASVI variable interacted with the Energy variable. The regression formula is based on formula (6), where we have introduced one lagged standard ASVI variable, and one lagged ASVI variable interacted with the Energy variable.

Regression Results with Lag

	Dependent variable:	
	Return	
	Pooled	Fixed
	(1)	(2)
ASVI_lag1	0.060	0.063
	(0.074)	(0.064)
ASVI_lag1:Energy	-0.131	-0.156
	(0.233)	(0.202)
Constant	3.893^{***}	
	(0.042)	
Observations	3,800	3,800
R^2	0.0002	0.0003
Adjusted R^2	-0.0003	-0.005
F Statistic	0.385 (df = 2; 3797)	0.600 (df = 2; 3778)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 2: Regression results with ASVI lagged variable.

As shown in the table (2), we can see that there is no significance in our results. Therefore, we conclude that there is no significant correlation between ASVI and returns when applying lagged variables. This means that the market reacts quickly where there are no delayed effects from the ASVI on stock returns. The results in section (5.1) support this statement.

5.3 Geographical Differences

The base regression uses Google Trends data from Norway, where the search term is “Ukraina”. We also want to test if the geographical input changes the results of the regression. Therefore, we change the geotag from Norway to worldwide and the search term from “Ukraina” to “Ukraine”. This test is only done out of curiosity, where we still believe that the search data from Norway is better suited to represent any effects on the OSE. As shown by the figures below, the search volume results are almost identical. They both share the same spikes for the given timeframe. However, the volume will be much larger for the worldwide database. The results in these graphs determine that there is no need to run regressions for both databases due to their similarity. The outcome would have little to no difference from each other since we are working with relative numbers and volume does not affect our analysis.

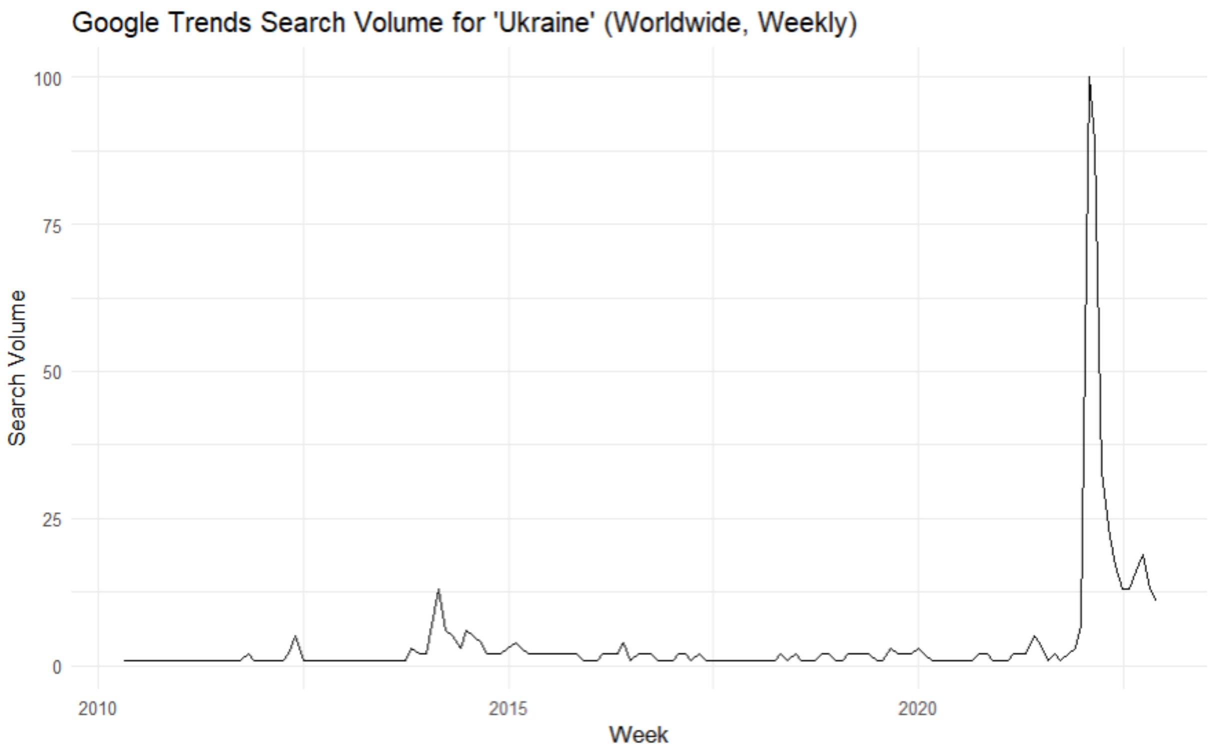

Figure 1: Search volume results for "Ukraine".

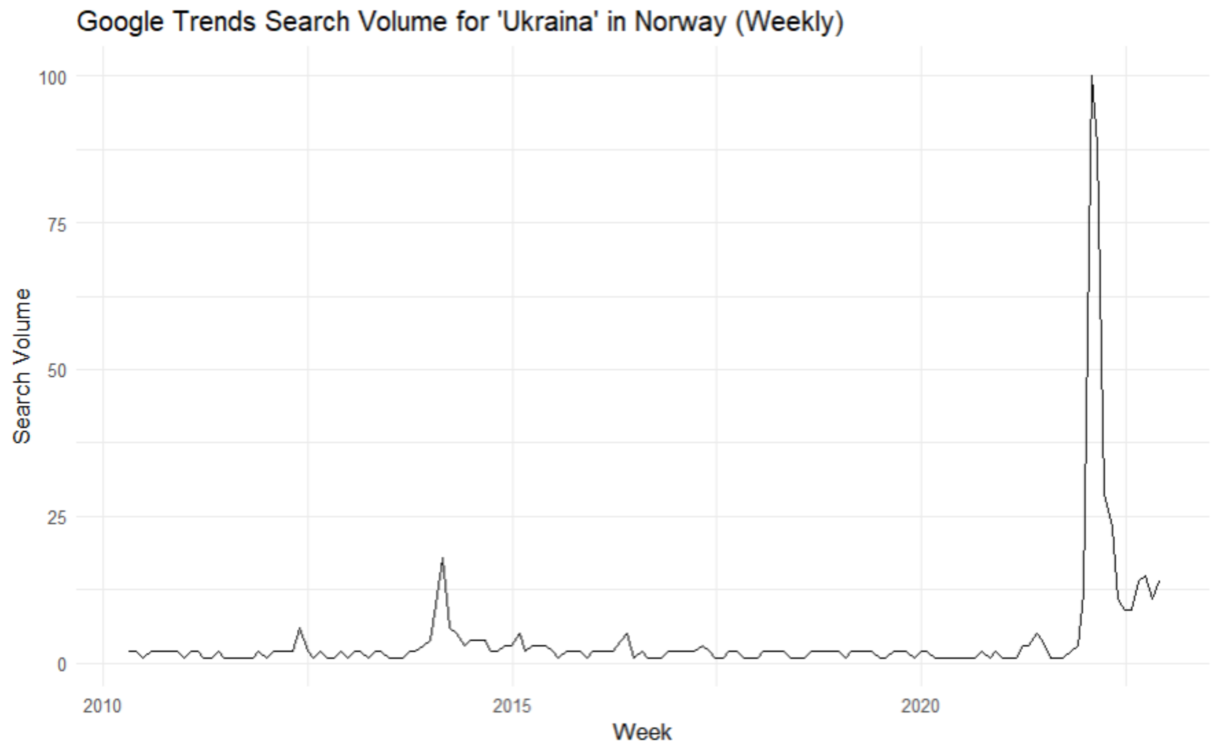

Figure 2: Search volume results for "Ukraina".

5.4 Energy-Specific Results

Our base research question is to find a correlation between stock returns on the OSE and the SVI of Google Trends for the search of Ukraine, but we are also interested in how the energy companies in our stock pool perform in the analysis on an individual basis. The invasion of Ukraine has had large effects on energy prices around the world, hence we want to run an individual panel data regression to examine if the returns of the energy sector can have any significant correlation with ASVI.

To strengthen the analysis, we have added two energy price indices to act as supportive variables in the regression. The stocks we will investigate are Equinor and Aker BP, which is the first and third biggest stock on the OSE. Therefore, we think they are representative of presenting the effects of the energy sector with a combined market capitalization of approximately $ 100 billion. The energy price indices we have utilized as a support are Brent Crude Oil and Natural Gas. The regression formula is based on formula (6), where we have included an index for both Brent Crude Oil and Natural Gas in addition to our original ASVI variable. Return is now the returns of EQNR and AKRBP.OL. We have also added interaction terms between ASVI and the supportive indices.

Regression Results of Energy Sector		
	Dependent variable:	
	Return	
	Pooled	Fixed
	(1)	(2)
ASVI	1.656	1.656
	(5.113)	(5.037)
ASVI:Crude_Oil	-0.441	-0.441
	(1.401)	(1.380)
ASVI:Natural_Gas	0.212	0.212
	(0.937)	(0.923)
Constant	4.710***	
	(0.140)	
Observations	390	390
R^2	0.0004	0.0004
Adjusted R^2	-0.007	-0.010
F Statistic	0.049 (df = 3; 386)	0.051 (df = 3; 385)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 3: Regression results within the energy sector.

The results indicate no presence of a significant correlation between energy stock returns and ASVI, nor with ASVI interacting with the added supportive indices. As mentioned, we have only used two stocks as a proxy for the whole energy sector, Equinor and Aker BP. Even though they have a large combined market capitalization, there may be a lack of diversification in our energy stock pool that may be a causal factor in the results we received.

5.5 Observation Frequency Differences

Lastly, we want to look further into the observation frequency and how the results may differ. As mentioned earlier, our base regression model contains weekly observations. As a final attempt, we will run the regression with daily data to see if the frequency yields a different result than the base regression. The daily regression has been conducted in the same way as the weekly regression. However, there are a few differences. Since the dates when using daily data do not match between the ASVI and the stock returns, we had to remove all observations for the search volume where the dates did not align with each other. This happens when the financial markets are closed, which is during the weekend and special holidays.

The daily regression also uses a different timeframe than the weekly regression. We decided to use 2022-01-01 up until 2022-08-31. This gives us 3 245 observations after removing not applicable observations, which is a sufficient amount for the regression. The model is based on formula (6) used to present the first regression model of this paper. The same variables are applied, only with daily frequency. As done earlier for the weekly data, we also wanted to test lagged variables for the daily data as well. However, since we are working with daily data, it makes sense to increase the lagged interval. In the former test, we lagged the observations with t-1, which represents 1 week. For the daily data, we decided to lag the observations with t-5, which represents a lag of 5 days for the ASVI variable.

Regression Results with Daily Observations

	Dependent variable:	
	Return	
	Pooled	Fixed
	(1)	(2)
ASVI	0.004*	0.004*
	(0.002)	(0.002)
ASVI_lag5	-0.004*	-0.004*
	(0.002)	(0.002)
ASVI:Energy	0.003	0.003
	(0.008)	(0.008)
ASVI_lag5:Energy	-0.003	-0.004
	(0.008)	(0.008)
Constant	-0.001	
	(0.0005)	
Observations	3,245	3,245
R^2	0.001	0.001
Adjusted R^2	0.00005	-0.006
F Statistic	1.039 (df = 4; 3240)	1.071 (df = 4; 3221)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 4: Regression results with daily observations.

The results in Table (4) conclude that there is little to no significant correlation between the ASVI and stock returns for the given period. The regression yielded a p-value of $p < 0.1$ towards both the regular and lagged ASVI variable, which does not give us enough evidence to reject the null hypothesis. The coefficients and standard errors of the models are similar, but this is because of the numbers being rounded up, where the differences are visible in the lower decimals. One thing to note is that the observed window of time is only approximately 9 months. This was mainly to try and capture the short-time effects of the invasion. The timeframe may be too narrow to gain any form of significance in the data, or the removal of non-matching observations can be another explanation.

6 Conclusion

In conclusion, our study explored the relationship between Google search activity, investor sentiment, and stock market volatility in the context of the ongoing Ukraine invasion. We employed a panel data regression model and utilized Google Trends data to capture investor attention and sentiment through the ASVI.

Initially, we applied regressions with both pooled OLS and FE to examine the association between ASVI and stock returns. Our results showed a significant correlation in the base regression, suggesting that investor search behaviour may have some predictive power for stock market movements. But it is important to remember that the SVI of Google Trends does not only contain searches related to investor behaviour.

To further investigate the dynamics between search activity and stock market volatility, we introduced lagged variables into our regression model. However, the lagged variables did not yield any significant results, indicating that there are no delayed effects on the results and that the market reacts quickly enough to capture the effects from Google searches. In addition to exploring lagged effects, we examined the potential geographical variations in search behaviour by comparing Google Trends search for “Ukraina” in Norway, to the search for “Ukraine” worldwide. We found no significant differences in search patterns between Norway and worldwide, suggesting that the search queries related to Ukraine were not solely driven by specific geographic factors.

To gain a more comprehensive understanding of the impact of geopolitical events on the energy sector, we incorporated energy price indices, specifically Brent Crude Oil and Natural Gas, into our regression model. The results indicated no evidence of a significant correlation between the returns of energy stocks and ASVI. We received the same results regarding the interaction towards ASVI interacted with the supportive indices. Furthermore, we explored the impact of using daily observations instead of weekly data in our regression model. However, the daily data did not provide any significant insights beyond what we observed in the weekly analysis, indicating that the weekly frequency was sufficient for capturing the relationship between search activity and stock market dynamics.

In summary, our study contributes to the growing literature on the relationship between online search activity and financial markets. We found evidence of a significant correlation between Google search activity for Ukraine and returns of the 20 biggest stocks on the OSE. Hence, we conclude that Google Trends search volume may be used as a proxy to determine stock returns in times of geopolitical conflict. This implies that changes in search volume on Google can provide valuable information or insights regarding investor sentiment and market behaviour, which in turn can be associated with stock returns during periods of geopolitical conflict. However, it is important to note that while there is a correlation in our study, using search volume as a proxy does not guarantee a causal relationship with stock returns, and other factors should also be considered in comprehensive analysis.

Overall, our research highlights the value of incorporating alternative data sources, such as online search activity, in understanding investor sentiment and stock market dynamics during geopolitical events. The insights gained from this study can inform policymakers, investors, and researchers in their decision-making processes and contribute to a deeper understanding of the complex interplay between online information, investor behaviour, and financial markets.

References

Baltagi, B. H. & Griffin, J. M. (1984). Short and Long Run Effects in Pooled Models. *International Economic Review*, 25 (3), 631-645. https://www.jstor.org/stable/2526223

Basumatary, K. & Devi, M. (2022). Pooled OLS and Fixed Effect Estimation of Wage Structure and Differential in Handloom Sector: Choosing the Better Method. *Journal of Social Economics Research, Conscientia Beam*, 9 (2), 137-146.

Bijl, L., Kringhaug, G., Molnár, P. & Sanvik, E. (2016). Google searches and stock returns. *International Review of Financial Analysis*, 45, 150-156. https://doi.org/10.1016/j.irfa.2016.03.015

Bui, V, X. & Nguyen, H, T. (2019). Stock market activity and Google Trends: the case of a developing economy. *Journal of Economics and Development*, 21 (2). http://dx.doi.org/10.1108/JED-07-2019-0017

Clark, T. S. & Linzer, D. A. (2014). Should I Use Fixed or Random Effects? *Political Science Research and Methods*, 3 (2), 399-408. https://doi.org/10.1017/psrm.2014.32

Collischon, M. & Eberl, A. (2020). Let's Talk About Fixed Effects: Let's Talk About All the Good Things and the Bad Things. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 72 (2), 289-299. https://doi.org/10.1007/s11577-020-00699-8

Da, Z., Engelberg, J. & Gao, P. (2011). In Search of Attention. *The Journal of Finance*, 66 (5), 1461-1499. https://doi.org/10.1111/j.1540-6261.2011.01679.x

Dallal, G. E. (2001). *Pooling Effects*. http://www.jerrydallal.com/lhsp/pool.htm

Fama, E. F. (1965a). The Behaviour of Stock-Market Prices. *The Journal of Business*, 38 (1), 34-105. http://www.jstor.org/stable/2350752

Fama, E. F. (1965b). Random Walks in Stock Market Prices. *Financial Analysts Journal*, 21 (5), 55-59. https://www.jstor.org/stable/4469865

Garman, M. B & Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *The Journal of Business*, 53 (1), 67-78. https://www.jstor.org/stable/2352358

Google LLC. (2023). Google Trends. https://trends.google.com/trends/

Griffith, D. A. (1992). What is spatial autocorrelation? Reflections on the past 25 years of spatial statistics. *L'Espace géographique*, 21 (3), 265-280. https://www.jstor.org/stable/44381737

Heiden, M. & Hamid, A. (2015). Forecasting volatility with empirical similarity and Google Trends. *Journal of Economic Behaviour & Organization*, 117, 62-81. https://doi.org/10.1016/j.jebo.2015.06.005

Hudson, R. S. & Gregoriou, A. (2015). Calculating and comparing security returns is harder than you think: A comparison between logarithmic and simple returns. *International Review of Financial Analysis*, 38, 151-162. https://doi.org/10.1016/j.irfa.2014.10.008

Jara-Días, S. & Rosales-Salas, J. (2015). Understanding time use: Daily or weekly data? *Transportation Research Part A: Policy and Practice*, 76, 38-57. https://doi.org/10.1016/j.tra.2014.07.009

Joseph, K., Wintoki, M. B. & Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. *International Journal of Forecasting*, 27 (4), 1116-1127. https://doi.org/10.1016/j.ijforecast.2010.11.001

Karoglou, M. (2009). Breaking down the non-normality of stock returns. *The European Journal of Finance*, 16 (1), 79-95. https://doi.org/10.1080/13518470902872343

Kaufman, R. L. (2013). *Heteroskedasticity in Regression: Detection and Correction.* Sage Publications Inc, California, US. https://doi.org/10.4135/9781452270128

Liu, Y., Peng, G., Hu, L., Dong, J. & Zhang, Q. (2019). Using Google Trends and Baidu Index to analyze the impacts of disaster events on company stock prices. *Industrial Management & Data Systems*, 120 (2). http://dx.doi.org/10.1108/IMDS-03-2019-0190

Marsh, H. W., Dowson, M., Pietsch, J., & Walker, R. (2004). Why Multicollinearity Matters: A Reexamination of Relations Between Self-Efficacy, Self-Concept, and Achievement. *Journal of Educational Psychology*, 96 (3), 518–522. https://doi.org/10.1037/0022 0663.96.3.518

McClain, K, T., Humphreys, H, B., Boscan, A. (1996). Measuring risk in the mining sector with ARCH models with important observations on sample size. *Journal of Empirical Finance*, 3 (4), 369-391. https://doi.org/10.1016/S0927-5398(96)00006-0

Miles, J. (2005). Tolerance and Variance Inflation Factor. *Encyclopaedia of Statistics in Behavioural Science*. http://dx.doi.org/10.1002/0470013192.bsa683

Molnár, P. (2012). Properties of range-based volatility estimators. *International Review of Financial Analysis*. 23, 20-29. https://doi.org/10.1016/j.irfa.2011.06.012

Preis, T., Reith, D. & Stanley, H. E. (2010). Complex dynamics of our economic life on different scales: insights from search engine query data. *Philosophical Transactions of the Royal Society A*, 368, 5705-5719. https://doi.org/10.1098/rsta.2010.0284

Preis, T., Moat, H. S. & Stanley, H. E. (2013). Quantifying Trading Behaviour in Financial Markets Using Google Trends. *Scientific Reports*, 3 (1684). https://doi.org/10.1038/srep01684

Salisu, A. A, Ogbonna, A. E. & Adediran, I. (2020). Stock-induced Google trends and the predictability of sectoral stock returns. *Journal of Forecasting*, 40 (2), 327-345. https://doi.org/10.1002/for.2722

Sandoval, L, J., Franca, I, D, P. (2012). Correlation of financial markets in times of crisis. *Physica A: Statistical Mechanics and its Applications*, 391 (1-2), 187-208. https://doi.org/10.1016/j.physa.2011.07.023

Swamy, V., Dharani, M. & Takeda, F. (2019). Investor attention and Google Search Volume Index: Evidence from an emerging market using quantile regression analysis. *Research in International Business and Finance*, 50, 1-17. http://dx.doi.org/10.1016/j.ribaf.2019.04.010

Yahoo Inc. (2023). Yahoo Finance. https://finance.yahoo.com/

Zhang, Y., He, J., He, M. & Li, S. (2023). Geopolitical risk and stock market volatility: A global perspective. *Finance Research Letters*, 53. https://doi.org/10.1016/j.frl.2022.103620