Nottingham Business School

Nottingham Trent University

Analysis of ITA ETF returns by Geopolitical Stress Index

Dissertation thesis

Nottingham 2023 **Jakub Jamnický**

Analysis of ITA ETF returns by Geopolitical Stress Index

Dissertation thesis

Author:	Jakub Jamnický
NTU Number:	N1171886
Academic year:	2022/2023
Supervisors NTU:	Dr Bova Giuseppe
Supervisors EUBA:	Ing. Barbora Stanová, PhD.
Study program:	MSc. International Finance

Nottingham Business School
Nottingham Trent University

ABSTRACT

This dissertation thesis concentrates on analysing the relationship between the Google search query, which contains words referring to geopolitical stress, and the representative of the defense industry on the stock market, aiming to determine the predictive ability of particular terms in Google Trends searches on the ITA ETF. The research uses monthly data of the Geopolitical Stress Index and returns of ITA and SPY ETFs from July 2006 to February 2023 to investigate regression on the classic OLS model and to determine Granger causality thanks to the VAR model. The results indicate a statistically significant regression of the geopolitical stress index using Google Trends on the ITA ETF returns. Furthermore, the outcomes suggest a more substantial influence of ITA ETF returns on Google Trends than vice versa.

Key words: Google Trends, Geopolitical Stress Index, ITA ETF, stock returns

ACKNOWLEDGEMENT

I wish to express my sincere appreciation to Dr. Giuseppe Bova from Nottingham Trent University for his invaluable guidance and tireless efforts in conducting numerous consultations with me. These sessions have proved instrumental in enabling me to streamline and enhance my thought processes and conceptual ideas methodically. Furthermore, since we use the European version of Excel in our analysis to process the results, we distinguish decimal numbers using a comma, not a full-stop punctuation mark. Apologies for every inconvenience.

CONTENT

ABSTRACT ... 3
ACKNOWLEDGEMENT ... 3
1 LITERATURE REVIEW ... 6
1.1 Google Trends ... 6
1.2 Geopolitics and Geopolitical Risk ... 7
1.3 Investor sentiment ... 8
1.3.1 Concept of the "Beauty Contest" ... 9
1.4 Behavioral biases ... 9
1.4.1 Perceive information behavioral biases ... 9
1.4.2 Belief Perseverance Biases ... 11
1.5 Efficiently Market Theory ... 13
1.6 Ordinary least-squares (OLS) model ... 15
1.7 Granger causality ... 15
1.7.1 Granger-Sims causality ... 15
1.7.2 Vector Autoregression model ... 16
2 DATA & METHODOLOGY ... 18
2.1 Data ... 18
2.1.1 Geopolitical Stress Index (GSI) ... 18
2.1.2 iShares US Aerospace & Defense ETF (ITA) ... 20
2.1.3 SPDR S&P 500 ETF Trust (SPY) ... 21
2.2 Methodology ... 21
2.2.1 Regression analysis ... 21
2.2.2 Granger causality analysis ... 22
3 RESULTS ... 24
3.1 Results of OLS ... 24
3.2 Results of VAR ... 30
4 DISCUSSION & CONCLUSION ... 35
REFERENCIES ... 36

INTRODUCTION

With the rise of internet search engines, such as Google, there is now an opportunity to analyse the relationship between geopolitical stress and stock market returns using search engine data. This dissertation aspires to explore this relationship, specifically analysing the predictive power of Google Trends search data for geopolitical stress in predicting the iShares US Aerospace & Defense ETF (ITA) returns, representing the defense sector on the stock market.

In the first part of our work, we deal with Google Trends data as a measure of the general population's curiousness, and then we present geopolitical risk and events that are a possible trigger of this interest. Later in this part of the work, we will also address the sentiment of investors and related concepts, which can bring insight into the informational value or influence that the Google Trends tool can bring to investors. We also show various behavioural biases that can affect the rationality of investors and the resulting Efficient Market Hypothesis, which due to the nature of the work, which tries to prove the predictive ability of stock markets using Google Trends, is its primary basis. Further in this part of the work, we also clarify statistical models such as OLS and VAR, which we used in the practical part of the work, as well as Granger causality.

Secondly, we detail the data sources used for our analysis in the data and methodology section, including a detailed description of our Geopolitical Stress Index construction and a detailed look at the returns data from the iShares US Aerospace & Defense ETF and SPDR S&P 500 ETF Trust. In this part of the work, we also observe the close construction of models we use for data analysis, such as OLS and VAR.

Subsequently, we will present the results of our models, which will help to understand the relationship between Google Trends search data for geopolitical stress and ITA ETF returns. Finally, in the discussion and conclusion part, we will summarize the findings and discuss the implications of this research for investors, policymakers, and future research in this area.

1 LITERATURE REVIEW

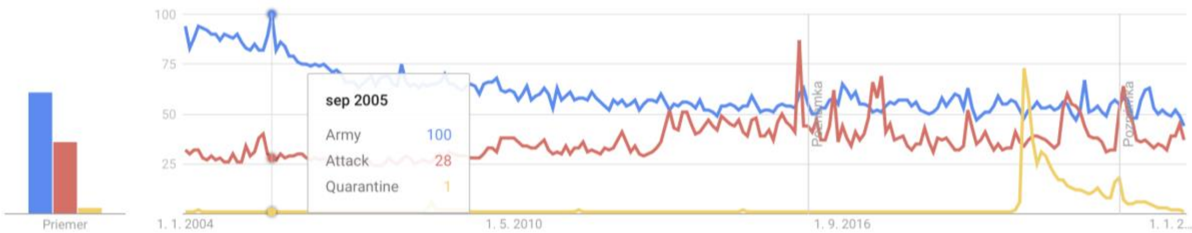
In this part of the work, we will try to present the relevant literature and an academic explanation of the models we will use in the practical part. We will also look at the theoretical views that became the basis of this work.

1.1 Google Trends

Google Trends is a tool that helps us explore how popular particular terms are in different parts of the world, languages, or regions. As the name suggests, this website by Google uses a real-time algorithm that allows us to observe terms from Google News and searches, ranking them by the relative increase in volume and absolute search volume. In addition, the site lets us search for trends and search queries of terms in all various categories simultaneously or separately. Google Trends observations include web search, image search, news search, Google shopping, and even on YouTube (Google, 2023).

Based on official public information from Google, Google Trends generates a time-series index of how frequently individuals search for a specific term over time in a geographic location or globally. The data are in relative numbers to the maximal search query of a given period. They have been available for users in CSV file format since January 2004. The Google Search Volume (GSV) index is the number of searches of investigated words entered into search engines relative to a top search query in a given observation. Therefore, as mentioned, the top search query term is set to 100, while the search query for the word at the beginning of the period is zero (Choi and Varian, 2012)

Figure 1. Search Index for terms Army, Attack, and Quarantine



Source: Google Trends, 2023

Research by Huang et al. (2020) found that Google search data can help forecast movements in the S&P 500 index. Their simple linear models performed well and generated significant excess returns in backtesting. They concluded that Google search data could be potential signals for the stock market, but the signal depends on the sentiment of the initial search term. Google search data represent not only investor attention but also the sentiment of retail investors. It indicates that Google search data could be a reasonable choice as input for a model which forecasts directional movements in the stock market.

Preis et al. (2013) suggest that Google Trends and stock market data could reflect different stages of investors' decision-making. They propose that people may gather more information during periods of concern before selling stocks at lower prices. Their study quantifies the connection between changes in the GSV index and stock market prices. They believe combining financial trading data with search query volumes could provide new insights into collective decision-making. These findings demonstrate the potential of Google Trends data to advance our understanding of complex collective behavior in society.

1.2 Geopolitics and Geopolitical Risk

Human geography concentrates on understanding the associations between people, societies, and the environment. According to Flint (2006), we can split human geography into four significant sub-disciplines: cultural, economic, political, and social geography. Each concentrates on a specific aspect of human activity. Flint (2006) claims that political geography is exceptional among these sub-disciplines. It has been a part of geography since its emergence in the late 1800s, during an era of imperial rivalry. Consequently, political geography boasts a more extensive and distinct history than the other sub-disciplines, setting it apart in its evolution.

Geopolitics was formed in 1899 when Swedish political scientist Rudolf Kjellén coined it in his book "Staternas Makt och Plats i Tillvaron", he defined the term as "the theory of the state as a geographical organism or phenomenon in space" (Cohen, 2015). This thinking by formal geopolitical reasoning about global politics helped states consider their potential to act globally based on their essentials. Therefore, geopolitics initiated to be a form of analyzing how geography and politics interact. It explains how different geographical settings and perspectives influence political power and determines spatial frameworks through which this power flows (Agnew, 2003; Cohen, 2015). In this view, Geopolitics considers international and domestic forces affecting global behavior.

Over the last few decades, the study of geopolitics has broadened to encompass a range of events and power struggles that involve different groups and organizations. These groups may include governments, corporations, rebel groups, and political parties (Caldara and Iacoviello, 2018). Because of this, the term "geopolitics" is now used to describe various events with varying causes and outcomes, such as cyber-attacks, international trade conflicts, and pandemics like COVID-19.

Following the end of the Cold War, the world experienced a phase of globalization and transnational liberalism, which presented new possibilities while simultaneously exposing unexpected geopolitical dangers, commonly called geopolitical risks (Agnew, 2005; Kurecic, 2015). Similarly, Behrendt and Khanna (2003) claim that the combination of other geopolitical events causing stress in the general public as terrorist attacks in the United States on September 11, 2001, or the conflict in Iraq, has forced organizations to take geopolitical risks in an element of decision-making.

Geopolitical risk events can significantly affect financial markets due to the fear and unpredictability of such occasions. When geopolitical risks occur, they can lead to a broad range of economic and financial consequences, including market volatility, disruptions in trade and commerce, and geopolitical tensions that can accelerate existing economic and political dangers (Balcilar et al., 2018). For example, as we mentioned, the global COVID-19 pandemic has caused reduced demand for goods and services, supply chain disruptions, and labor shortages (Kumar et al., 2020). As a result, it led to significant losses for firms, followed by their financial market investors.

In summary, geopolitical risk covers the circumstance that these events will happen and the risks that come with driving them worse. Moreover, the impact of some geopolitics dangers, such as crises, pandemics, or wars, can significantly affect a broad spectrum of sectors, including financial markets (Jung et al., 2021). Therefore, understanding geopolitical risks is essential for organizations and investors to make informed decisions. They can use Google Trends as a proxy tool to track the popularity of search terms that refers to the impact of political, economic, and social factors on global markets, signifying geopolitical risk.

1.3 Investor sentiment

Investor sentiment reflects the cognitive and affective aspects of investors' beliefs about future earnings. As a collective measure of investors' psychology and emotions, it has a considerable impact on the supply and demand for securities, which, in turn, drives market movements (Barberis et al., 1998). Moreover, it denotes investors' collective attitude, encompassing optimism, pessimism, fear, greed, or confidence (Cohen-Charash et al., 2013).

Consequently, Baker and Wurgler's (2007) empirical investigation demonstrates that investor sentiment significantly influences stock returns, even after accounting for other fundamental factors, such as dividend yields and earnings. However, their study reveals that high sentiment levels predict lower future stock returns, implying that sentiment can serve as a valuable market predictor. In summary, investor sentiment is a pivotal force in shaping the financial market's behavior and monitoring it is a critical task for investors and financial professionals.

Tetlock (2018) discovered that during times of high investor sentiment, certain stocks tend to have lower returns, such as petite, unprofitable, non-dividend-paying, high volatility, extreme growth, and distressed stocks. However, when investor sentiment is low, these patterns may disappear or reverse. Interestingly, specific company characteristics only reveal strong patterns when considering investor sentiment.

As mentioned above, geopolitical risk refers to the risk of events that occur from political tensions between countries and can significantly affect financial markets. Accordingly, investor sentiment is affected by various factors, including geopolitical risks. Consequently, geopolitical risks have the potential to influence investor sentiment, thereby triggering shifts in market behavior and producing financial consequences.

1.3.1 Concept of the "Beauty Contest"

Keynes pointed out the view of the market in two ways: investment and speculation. Investment involves predicting the yield of an asset over its whole existence. On the other hand, speculation involves forecasting the psychology of the market. This duality highlights that market psychology impacts asset prices and changes returns (Peterson, 2016).

According to Keynes' theory, considering others' thoughts is critical in investors' decision-making, as illustrated by the "Beauty Contest" concept in his book "The General Theory of Employment, Interest, and Money". In a "Beauty Contest" game, people are given a group of pictures and questioned to select the most attractive face. The winner is the one who picks the face, which most other people also choose. This game shows that people's choices may not always reflect their individual preferences but rather what they think others will choose (Keynes, 2013).

Thanks to this Keynesian concept, we can better understand the psychological pressures that cause investors, from individual thinking to high-level investors, which consider the impact of what other investors will do, regardless of their belief in the rationality of markets.

We can use this concept in the context of Google Trends data, which shows search terms' relative popularity over time. We can use it as a proxy for investors' sentiment toward a particular company or industry. For example, if the specific term is related to a particular sector, industry, or company and is being searched more frequently on Google, it could indicate increased interest in that sector or company, influencing investors' decisions and affecting returns.

1.4 Behavioral biases

Academic researchers specializing in financial decision-making have detected many different behavioral biases. These biases can overlay or be made from numerous psychological mechanisms, and these categorizations are not always clear or united. Therefore, in our research context, we will discuss two classifications of such biases: those associated with information processing and belief perseverance biases.

1.4.1 Perceive information behavioral biases

According to Krawczyk and Baxter (2020), biases are affected by current events and can affect our thoughts and behaviors. They are often developed through our experiences and force us to act in specific ways in different situations. Heuristics are related to biases, as heuristics come from the biases we develop. Our biases relate to our experiences and expertise level. It means that a highly experienced investor may hold biases based on situations they have seen play out repeatedly on the financial markets, allowing them to act with intuition. On the other hand, fewer experienced investors may need more time and resources to evaluate value and act.

To help us understand the origins and apply remedies of biases, Krawczyk and Baxter (2020) categorize behavioral biases into three groups based on cognitive processes and brain functions:

- Attention biases - arise from our brain's stimulus processing. Our brains have evolved to support foraging behavior, which includes survival circuits that help us learn about rewards and change our behavior in reaction to losses. Money activates the same courses, such as a strong motivator in everyday life.
- Memory biases - exhibit limitations of human memory, whereby the vast amount of incoming information necessitates selective retention, resulting in biases that can hinder investment outcomes.
- Knowledge biases - result from how new information interacts with our previous experiences and our brain's tendency to fit new knowledge into pre-existing structures. These structures, known as schemas, are built over time as we notice patterns and similarities in information. New experiences could modify schemas and lead to biases that negatively impact investing.

Understanding these biases to recognize them and staying objective when considering incoming information is essential. Moreover, it could lead to better narrative construction and more effective investing.

Based on these definitions, we can assume that the knowledge bias will conceivably have the most consequential impact on Google Trends of the three biases. As we know, Google Trends reflects the collective search behavior of the general public affected by their current knowledge and beliefs. As a result, people tend to search for information that confirms their existing knowledge, leading to knowledge bias. Of course, attention and memory biases may impact search queries too, but their influence is probably secondary to knowledge bias. Therefore, we should concentrate on this category of behavioral biases.

The curse of knowledge is a bias where personal knowledge interferes with accurately predicting the actions of others who lack the same expertise. As a result, it can lead to erroneous assumptions about what others know and do. Economists discovered this phenomenon by observing students trading stocks and realizing that they could not ignore their knowledge even when fully informed. This bias is relevant to investing because it can make estimating the level of one's informational advantage challenging compared to others in the market. The curse of knowledge operates at the intuition and reason levels and can misguide us by failing to discount others' estimates based on personal information (Camerer et al., 1989). When utilizing Google Trends data to guide investment decisions, the curse of knowledge may give rise to an assumption amongst investors that others are interpreting the data like themselves. Such an assumption can create a false feeling of confidence in the accuracy of the data and the decisions derived from it. Furthermore, the curse of knowledge may cause investors to ignore other important determinants of the

GSV index and stock markets. Therefore, we assume that excessive dependence on Google Trends data may result in incorrect investment decisions.

Anchoring bias is an information processing bias where an initial number greatly influences our subsequent actions. People often fail to adequately scrutinize the basis for the anchor point or its appropriateness to the situation and adequately adjust up or down from the anchor point (Krawczyk and Baxter, 2019). For example, in the study of Tversky and Kahneman (1979), a group of students was asked people to quickly estimate the product of complicated multiplication problems after only five seconds of thought. It was enough time to reproduce a few numbers, producing far too small an estimate. Students appeared to have anchored on that initial value and then inadequately modified their estimation upward. Therefore, the risk of anchoring is high in stock pricing, so investors must be attentive to avoid overly fixating on a current price, failing to predict how much the price may move simply because of the current level (Krawczyk and Baxter, 2019). If investors are anchoring on a particular stock or stock market trend, they may search for and rely heavily on information confirming their initial anchor point while failing to consider other relevant information that may lead to a more accurate assessment of the market's performance. This bias may influence the search volume index on Google and potentially lead to distorted perceptions of the market's performance.

1.4.2 Belief Perseverance Biases

In the first half of this part of the work, we discussed behavioral biases, which focus more on information processing and their connection with GSV index data. In the second half of this work, we want to present behavioral biases that are related to decision-making by investors.

Pompian (2020) states this belief perseverance biases:

- Cognitive dissonance is a psychological phenomenon that arises when newly acquired information disagrees with existing beliefs, resulting in cognitive discomfort. People may try to change their attitudes, beliefs, or values to quiet this discomfort. However, this phenomenon can exert a powerful influence on investors, compelling them to maintain their positions despite the lack of any obvious potential, even as new information emerges that undermines any rational justification for doing so. It may arise from the discomfort of acknowledging the fact that the investment was not a good choice.
- Conservatism bias is a cognitive process of individuals to uphold their existing beliefs or predictions, even when presented with new information that disagrees with them. For example, in the context of an investor receiving negative news regarding a company's earnings that contradicts a previous earnings estimate issued just a month prior, conservatism bias may lead the investors to underreact to the new information because they will maintain their prior beliefs instead of

acting on the updated data. This bias generally leads investors to remain in their pre-existing beliefs or predictions rather than comprising new information. Therefore, we can say that those investors who are more conservative in their beliefs tend to be slower to update their expectations when new information arrives, which can lead to mispricing in financial assets (Fama and French (2007).

- Confirmation bias is a type of selective perception that prioritizes ideas that align with our beliefs while disregarding anything that contradicts them. This behavior, related to belief perseverance, happens as people try to persuade themselves of what they want to believe and ignore every dissonance with over-belief. Investors overemphasize events that support our desired outcomes and minimize any opposing evidence.
- Representativeness bias explains human's tendency to classify things to understand them. When people come across something new that does not correspond to our existing categories, we try to suit it to the closest match we can find. This helps us process information quickly by drawing on our past experiences. However, sometimes investors mistakenly categorize new information as old ones when they are very different. So, depending on similarities can trigger a representativeness bias and lead to a wrong understanding of information and, therefore, unreasonable investor decision-making.
- The illusion of control bias is when people believe they have control over outcomes, even when it is impossible. For example, this bias often occurs in casinos, where individuals think they can affect random events like rolling dice. The illusion of control bias can result in excessive trading or speculative trading and under-diversification. Generally, it reaches from investor overconfidence.
- Hindsight bias is a belief perseverance bias that forces people to perceive an event as predictable after it has happened. This bias arises because actual results are more accessible for the mind to process than the infinite possibilities that could have occurred. In addition, it causes people to view events as unavoidable and predictable, which can force them to ignore past events. Finally, this bias affects future forecasting by underestimating uncertainty and underrating the outcomes that could have been but did not (Fischhoff, 2003).

These biases can conduct in incorrect or irrational decision-making, particularly in maintaining positions lacking potential or disregarding new information that contradicts prior beliefs. In the era of the extensive information available through digital platforms like Google Trends, investors do not always remain rational in information processing and decision-making.

1.5 Efficiently Market Theory

In an efficient market, investors cannot consistently generate abnormal risk-adjusted returns that exceed their opponents, in other words, "beat the market". Instead, investors can use publicly available information to trade assets in the capital market and maximize their capital. This information includes historical price patterns, volume data, public announcements regarding company performance, and potentially intimate insider knowledge. Technical analysts are investors who use historical price patterns and volume or volatility data to identify the time and the market for allocating their assets. However, asset price movements must adhere to specific, predictable trends for technicians to outperform the market (Hodnett and Hsieh, 2012).

In the efficient markets model context, a random walk is a hypothesis that successive price changes or asset returns are independent and identically distributed. It means that since new information is unpredictable, the changes in stock prices should also be unexpected. Asset returns arise when luck causes investors' preferences and new information to combine to produce equilibria where return distributions repeat themselves over time. It is an extension of the expected return or "fair game" model, which we can use in empirical market efficiency tests. The model indicates that the current prices of assets reflect all available information and their underlying risk. Hence, it serves as a benchmark for evaluating investment decisions and testing for market efficiency. Therefore, based on the efficient market, asset prices cannot "randomly walk" change over time (Fama, 1970).

Fama (1965 and 1970) also examined the empirical literature on the random walk of asset prices and set the concept of market efficiency, categorized into weak-form, semi-strong, and strong-form market efficiency. In addition, Fama (1970) analyzes these categories on the extent to which different forms of information are quickly and accurately reflected in asset prices, thereby supporting the efficient market hypothesis. In his research, the efficient markets model is an extreme null hypothesis that assumes security prices at any point fully reflect all available information. In his research, the efficient markets model is an extreme null hypothesis that assumes security prices at any point fully reflect all available information. Furthermore, categorizing tests into weak, semi-strong, and strong forms allows for identifying the level of information at which the hypothesis breaks down.

The weak form tests pertain to the information subset of interest, including only historical prices reflected in current security prices. However, empirical research in this area is extensive, and most results support the efficiency hypothesis. Therefore, a weak-form efficient market is one in which the possibility of consistently earning positive abnormal returns is allowable. However, technical analysts who make decisions based on price charts and volume data cannot consistently make profits in this form of the efficient market (Fama 1965, 1970 and 1991; Hodnett and Hsieh, 2012).

The semi-strong form tests of efficient markets inspect if current prices include all the public information. However, according to Fama (1970), these tests look at how prices adjust to one type of information, like company earnings announcements, stock splits, financial reports, new securities, and others. Therefore, if each test only provides some evidence supporting the model, then many tests are needed to prove its validity. Even fundamental analysts who study macroeconomic forces and company performances cannot outperform the market in a semi-strong efficient market (Hodnett and Hsieh, 2012). This particular category earns our most considerable attention as Google Trends data is readily available to the public, and therefore our investigation into its influence on specific security returns makes this category particularly relevant to our research.

The strong-form tests of efficient markets examine whether investors or groups have exclusive access to information relevant to price formation. In a strong-form efficient market, insider information is absent because company insiders do not have exclusive access to information relevant to price formation (Fama 1965, 1970 and 1991; Hodnett and Hsieh, 2012). According to Fama (1970), there is limited evidence against the hypothesis that monopolistic access to information is not prevalent in the investment community.

Each form of efficiency conveys a different information processing level reflected in asset prices. Moreover, as the market becomes more efficient, this information as an investor's instrument for trading loses effectiveness as more investors learn to use it and modify their trading accordingly (Hodnett and Hsieh, 2012). Therefore, understanding the distinctions between these forms of efficiency is essential for fast evaluating the impact of various sources of information on asset prices, including the potential impact of Google Trends data on stock market returns.

Nevertheless, the efficient markets theory is imperfect and cannot be trusted to make fast and reliable profits. While market inefficiencies may exist, they are not guaranteed to result in immediate profits. However, there are other ways in which market efficiency can be wrong, such as in its interpretation of significant or geopolitical stress events (Shleifer, 2000).

Even when Fama (1998) criticized the anomalies discovered in behavioral finance, according to Shiller (2003), these criticisms are weak and do not undermine the fundamental value of the theory. Nevertheless, it is essential to consider the weaknesses of efficient markets theory and take an eclectic approach when researching financial markets. We cannot believe that financial markets always work well or that price changes always reflect genuine information. Evidence from behavioral finance shows that human behavior can cause misallocations of resources (Shiller, 2003 and 2015). Therefore, we must also include this reality in our research.

1.6 Ordinary least-squares (OLS) model

The researchers use Ordinary Least Squares (OLS) model as a statistical method for calculating the parameters of the Classical Linear Regression Model (Wooldridge, 2020). According to Brooks (2019), regression analysis is an essential tool econometricians use to describe and evaluate the relationship between one variable and one or more other variables. It helps to describe modifications in one variable by looking at changes in the other. In this case, the model only concentrates on explaining changes in one variable. In comparison to correlation, when two variables are said to be correlated, it means that there is evidence of a linear relationship between them. It does not necessarily mean that one variable causes changes in the other. In regression, the dependent variable is treated as random, while the independent variables are assumed to have fixed values. Regression is more flexible and powerful than correlation (Brooks, 2019). Greene (2020) display regression in following formula:

$$y = f(x_1, x_2, \dots, x_K) + \varepsilon = x_1\beta_1 + x_2\beta_2 + \dots + x_K\beta_K + \varepsilon$$

where:

y is the dependent or explained variable,

$x_1, x_2, \dots, x_K$ is the independent or explanatory variable,

ε is the random disturbance.

The regression model presented here is a mathematical equation explaining the relationship between an explained variable, y, and explanatory variables, $x_1, x_2, \dots, x_K$. In addition, the coefficients $\beta_1, \beta_2, \dots, \beta_K$ explain the force and direction of this relationship between variables. The model captures the net effect of missed factors in a random disturbance. Errors of measurement are significant contributors to the disturbance in an empirical model. Obtaining accurate measures of economic variables is challenging, and sometimes there may be no observable replication of the theoretical variable. Therefore, this model is widely used in econometrics to analyze and predict the behavior of convoluted systems (Brooks, 2019; Greene, 2020; Wooldridge, 2020).

1.7 Granger causality

1.7.1 Granger-Sims causality

Granger (1969) introduced a category of tests to investigate causal relationships between two variables. First, tests observe changes in one variable, called the "cause," precede and help predict changes in another variable, called the "effect," then the cause variable is said to be Granger-cause the effect variable.

Sims (1972) adjusted Granger's (1969) causality test three years later. According to Sims (1972), Granger's (1969) definitions, when limited to linear predictors and stationary covariance processes, can be assumed identical to the parameter restrictions present in the moving average or distributed lag representations of a bivariate system (Kuersteiner, 2010). In other words, according to Florens and Mouchart (1982), if we have a process called X, which has some values. We use some fields to group these values and call them X, Y, and Z. We also have another field called U with some parameters and initial conditions. Granger and Sims have different ideas about causality in this process. Granger says Y does not cause Z if Z can be predicted without using Y, while Sims says Y does not cause Z if Y does not provide any extra information to predict Z when compared to Z directly.

Brooks (2020) considers the term "causality" relatively misleading since Granger causality only indicates a correlation between the current value of one variable and the past values of another. It does not necessarily mean that one variable causes the movements of another. However, the Granger causality test provides a framework for investigating causal relationships between two variables, in the case of our research, such as GSV data and some security returns. By using this test, we can determine if changes in GSV data help forecast changes in returns and demonstrate if there is a probability of a causal relationship between the two.

1.7.2 Vector Autoregression model

Sims (1980) introduced Vector autoregressive models (VARs) to the public as a conception of univariate autoregressive models. VAR, a model with multiple dependent variables, combines univariate time series and simultaneous equations models. Generally, this model could be a reasonable substitute for large-scale simultaneous equations structural models. Hamilton (1994) supports using a VAR model to test for Gender Causality, one of the hypothesis tests that rely on a VAR estimated in levels without normal asymptotic distribution.

Brooks (2019) displayed the simplest form of a bivariate VAR model with only two variables, y_{1t}, and y_{2t}. The current values of these variables depend on different combinations of their past k values and error terms. We express Brooks' (2019) equation below:

$$y_{1t} = \beta_{10} + \beta_{11} y_{1t-1} + \ldots + \beta_{1k} y_{1t-1} + \alpha_{11} y_{2t-1} + \ldots + \alpha_{1k} y_{2t-1} + u_{1t}$$
$$y_{2t} = \beta_{20} + \beta_{21} y_{2t-1} + \ldots + \beta_{2k} y_{2t-1} + \alpha_{21} y_{1t-1} + \ldots + \alpha_{2k} y_{1t-1} + u_{2t}$$

where:

u_{it} is a white noise disturbance term with $E(u_{it}) = 0, (i = 1,2)$.

The equation supposes that the disturbances, u_{1t}, and u_{2t}, are uncorrelated across equations, meaning $\mathrm{E}(\mathrm{u}_{1\mathrm{t}}, \mathrm{u}_{2\mathrm{t}}) = 0$. Nevertheless, in practice, we can identify a coincident correlation between the disturbances as $Cov(u_{1t}, u_{2t}) = \sigma_{12}$.

The VAR models are less suitable for theoretical analysis and policy recommendations. There is also a higher risk of obtaining spurious relationships only by "data mining". As a result, interpreting the coefficient estimates in VAR models can be unclear (Hamilton 1994, Brooks 2019). Therefore, VAR models cannot lack theoretical guidance in their specifications. These models need to have a solid theoretical base.

2 DATA & METHODOLOGY

In this part of the work, we will present the data and analysis methods of the relationship between search queries of selected words associated with geopolitical stress on Google Trends and the price returns of the ITA index, which includes the most significant companies operating in the arms industry. The goal of our research is to find out what impact these search volumes have not only on the development of the stock market returns in the field of the arms industry but also explore the Google Trends tool as one of the potential tools for predicting market actions.

2.1 Data

2.1.1 Geopolitical Stress Index (GSI)

One of the essential data on which our research is focused comes from the Geopolitical Stress Index (furthermore referred to as GSI). The GSI is the monthly average web worldwide search query of specifically chosen words on Google Trends. We can also display it in the equation as follows:

$$GSI = \frac{\frac{\sum w_1, w_2, w_3, \dots, w_n}{n}}{100}$$

where:

w is the relative search volume of the given word in the relevant month to the benchmark word in the monitored period,
n is number of all observative words in the index.

Since the relative search volumes are in percentage, we divided their monthly average by 100 to obtain a coefficient.

The words we use in the GSI index for measuring geopolitical tension originate in the GPR index[1] by Dario Caldara and Matteo Iacoviello (2018). Observations also showed that the most statistically significant for our research were the words of the GPR index, which in web searches tracked by Google Trends reached an average search volume of more than 3% for the entire observed period. We can assume that the low average search volumes of these words were disturbing for the whole sample, mainly because these words did not attract enough attention to provoke a reaction from the general public. Therefore, we select only a few specific words used in the GPR index, such as:

arm, arms, army, attack, bomb, clash, concern, conflict, crisis, danger, enem, fear, foe, hostage, invasion, kill, military, missile, offensive, peace, quarantine, raid, revolution, risk, strike, tension, threat, weapon.

[1] The Geopolitical Risk index or GPR includes the automated text-search outcomes obtained from the electronic archives of 10 newspapers. They calculated the index by monthly counting each newspaper's articles on adverse geopolitical events.

We used data from three periods when collecting data from July 2006 to February 2023 on the volume of searches for these words from Google Trends for our index. As mentioned above, Google Trends provides query search data as a relative measure of the maximum search volume in the monitored period. Therefore, dividing the total period into three monitored windows gave us more accurate data on the interest in word query searches.

Table 1. Time windows of collecting data for GSI

	Window 1	**Window 2**	**Window 3**
Period of duration	2006/07 - 2012/06	2012/07 - 2018/06	2018/07 - 2023/02
Word of maximum search query	army	attack	quarantine
Date of maximum search query	2006/11	2015/11	2020/03

Source: Data collected from Google Trends

As we can see in Table 1., each window relates to a different word that reached the maximum search query. However, we naturally assumed that the words with the highest search query would change over time as a result of various world events that occurred in the whole observed period.

For example, in November 2006, several events originated in geopolitical stress connected to the search term "army". Firstly, the enormous losses of the American army in the continuing war in Iraq, where November 2006 was also referred to as a deadly month, with over 100 American soldiers killed (CBS, 2006). Alternatively, the Fijian army, under the leadership of Commodore Frank Bainimarama, staged a coup against the government of Prime Minister Laisenia Qarase. As a result, the army rapidly established its control over the nation (United Nations, 2006). We mentioned only two geopolitical stress events, but we should note that in November 2006, the term "army" was the most searched query, also thanks to many other events related to geopolitical stress (NPR, 2006). These events assumably contributed to the increased interest and search for the term "army" during that time.

The second window offers a more specific answer to the maximum searched term and given period. We can consider that the word attack reached the top in November 2015 due to the historical event of the terrorist attack in Paris (BBC, 2015; The Guardian, 2015; NBC, 2015).

It is more than evident that the term "quarantine" had a top search query in March 2020 due to the outbreak of the COVID-19 pandemic. We can assume that this is because quarantine was one of the critical measures of the government to control the virus's spread, leading to increased interest in the term as people seek information about it (Financial Times, 2020; The Guardian, 2020).

As we mentioned, search queries for specific terms change over time due to various world events. For example, "army" and "attack" were most searched during specific periods due to geopolitical stress events like war and coups. We also believe that the COVID-19 pandemic directed the term "quarantine" to become the most searched query as it was an essential measure in controlling the virus's spread.

These data construct our geopolitical stress index. In our regression model, this variable also appears under the abbreviation GSI.

2.1.2 iShares US Aerospace & Defense ETF (ITA)

iShares U.S. Aerospace & Defense ETF (furthermore referred to as ITA) is a fund that tracks the performance of companies in the aerospace and defense industry in the United States. Accordingly, with free available data from March 2023, The iShares U.S. Aerospace & Defense ETF (ITA) has over $5.64 billion in assets under management (Yahoo Finance, 2023; Financial Times, 2023). The index has the most significant AUM from ETFs focused on the defense industry. The ITA ETF is a financial product created by BlackRock to track a benchmark index, the Dow Jones U.S. Select Aerospace & Defense Index (BlackRock, 2022). ITA, as a product, allows investors or the general public to invest in the aerospace and defense sectors. Therefore, related to our research, it could reflect the actual sentiment impact on the industry, which we assume is most sensitive regarding worldwide geopolitical stress.

The index basis on the prices of the stocks of the companies contained in the index. Therefore, the index's performance is calculated by the changes in stock values over time. In our research, we need to observe the effect of a search query on this ETF or vice versa. Therefore, we need to work with the ITA returns. For calculating the returns of ITA, we used the following equation:

$$retITA = \frac{\frac{PriceITA_{t_0} - PriceITA_{t_1}}{PriceITA_{t_0}}}{100}$$

where:

$PriceITA_{t_0}$is previous month's price of the ITA index in *average price,*

$PriceITA_{t_1}$is current month's price of the ITA index in *average price.*

As mentioned in the GSI index part, we work with coefficients. Therefore, we are dividing the result by 100. We use the average of the open, high, low, and close prices for our period. Using the average price can provide us with a complete view of the price behaviour of an ETF over a certain period.

Because it is very problematic for us to determine whether the market will react positively or negatively to events associated with geopolitical stress, we have to work with the absolute value of returns. Therefore, in our model, we operate with absITA data, demonstrated in the following equation:

$$absITA = abs\left(\frac{\frac{PriceITA_{t_0} - PriceITA_{t_1}}{PriceITA_{t_0}}}{100}\right)$$

This equation provides us with more accurate data for our research because, with absolute numbers, the analysis focuses more on the size of change.

2.1.3 SPDR S&P 500 ETF Trust (SPY)

Considering that we are trying to track the impact of geopolitical stress using the GSI index on a specific sector represented by the ITA index, it is also appropriate to consider market conditions that may affect this index. To address these conditions, we utilized an SPDR S&P 500 ETF Trust (SPY) to consider their impact. The SPY is widely acknowledged as a trustworthy indicator of market conditions, as it tracks the performance of the S&P 500 index. ETF aspires to replicate the returns of the S&P 500 index as near as possible by investing in identical stocks with the exact proportions.

Similarly, as with ITA, we operated with absolute returns also in this case. Therefore, the equation for collecting SPY returns data can display as this:

$$absSPY = abs\left(\frac{\frac{PriceSPY_{t_0} - PriceSPY_{t_1}}{PriceSPY_{t_0}}}{100}\right)$$

where:

$PriceSPY_{t_0}$ is previous month's price of the SPY index in *average price,*

$PriceSPY_{t_1}$ is current month's price of the SPY index in *average price.*

2.2 Methodology

2.2.1 Regression analysis

In our model, we try to include two market ETFs representing the market situation. We can assume that the price value of the SPY reflects a broader group of factors than the ITA ETF. Given that ITA is an exchange-traded fund that invests in companies involved in the aerospace and defense industries in the United States, we can assume that the price value of the index will be more sensitive to geopolitical stress events and their effects. Therefore, in our model, the absolute value of the average monthly returns of the ITA (absITA) will be the dependent variable. If we want to analyse the influence of our GSI index, then we can display our OLS model in the following equation:

$$absITA_t = \beta_1 GSI_t + \beta_2 absSPY_t + \mu$$

This OLS model aims to estimate the values of the regression coefficients β_1 and β_2 that minimize the sum of the squared errors or residuals and to test whether the independent variables have a statistically significant effect on the absITA.

2.2.2 Granger causality analysis

In the case of the OLS model, we talk only about one time series, which provides us a view of our variable simultaneously. But, if we assume that it could be a delay exists between new information and reaction to them on the market, such as in the case of conservative investors that are slow to update their opinions, then we should test Granger's causality, which analyses relationships between our variables using multiple time series. For this purpose, we can use the VAR model, which enables the assessment of the lagged coefficients between two-time series and identifying whether past values of one series aid in forecasting the future values of the other series. So, unlike the OLS model, if we use the VAR model to test Granger causality, we can predict the size of the returns based on Google trends. Or vice versa, we can expect a possible volume of search queries for specific words for higher or lower returns. The vector autoregression in our model we can also express as this:

$$GSI_t = \beta_1 + a_{11}GSI_{t-1} + a_{12}absITA_{t-1} + \ldots + a_1kGSI_{t-k} + a_{1k}absITA_{t-k} + \varepsilon_{1t}$$

$$absITA_t = \beta_2 + a_{21}GSI_{t-1} + a_{22}absITA_{t-1} + \ldots + a_2kGSI_{t-k} + a_{2k}absITA_{t-k} + \varepsilon_{2t}$$

where:
GSI_t and $absITA_t$ are the values of the variables at time t,
GSI_{t-i} and $absITA_{t-i}$ are the lagged values of the variables at time $t-i$, where i = 1, 2, ..., k,
β_1 and β_2 are the intercept terms for each equation.
$a_{11}, a_{12}, a_{21}, a_{22}, a_{1k}, a_{2k}, a_{1k}, a_{2k}$ are the coefficients capture the coexisting and lagged relationships between the variables.
ε_{1t} and ε_{2t} are the error terms for each equation, which capture any unobserved elements that impact the variables.

As we can see in the equation above, our model captures the dynamic interrelationships only between the endogenous variables in the system. The model assumes that past values and the other variables' past values influence both the GSI and absITA variables. Therefore, we do not include any exogenous variables in our model because we want to model the behaviour of our two endogenous variables over time based on their past values and mutual interactions. From the view of Granger causality, it can be a valuable way to explore how the variables relate and predict their future behaviour.

However, Brooks (2019) states that if all the variables in VAR are stationary, it is possible to test the joint hypotheses easily using the F-test framework. Because of that, the individual set of restrictions involves parameters from only one equation. Therefore, we can estimate equations separately with OLS to get an unrestricted Residual Sum of Squares, then set restrictions and re-estimate models for a restricted Residual Sum of Squares. Hence, to evaluate the significance of variables in a VAR, we typically conduct joint tests on all lags of a variable rather than examining individual coefficients.

Therefore, given our monthly data, we examined lags in the horizon from one year to observe as large a spectrum as possible, which represents 12 OLS lags. We created one test for each lag. We could also display these models in the following equations:

$$absITA_{t_1} = \beta_1 GSI_{t_{-1}} + \beta_2 absSPY_{t_{-1}} + \mu$$
$$absITA_{t_1} = \beta_1 GSI_{t_{-2}} + \beta_2 absSPY_{t_{-2}} + \mu$$
$$absITA_{t_1} = \beta_1 GSI_{t_{-3}} + \beta_2 absSPY_{t_{-3}} + \mu$$
$$absITA_{t_1} = \beta_1 GSI_{t_{-4}} + \beta_2 absSPY_{t-4} + \mu$$
$$absITA_{t_1} = \beta_1 GSI_{t_{-5}} + \beta_2 absSPY_{t-5} + \mu$$
$$absITA_{t_1} = \beta_1 GSI_{t_{-6}} + \beta_2 absSPY_{t-6} + \mu$$
$$absITA_{t_1} = \beta_1 GSI_{t_{-7}} + \beta_2 absSPY_{t-7} + \mu$$
$$absITA_{t_1} = \beta_1 GSI_{t_{-8}} + \beta_2 absSPY_{t-8} + \mu$$
$$absITA_{t_1} = \beta_1 GSI_{t_{-9}} + \beta_2 absSPY_{t-9} + \mu$$
$$absITA_{t_1} = \beta_1 GSI_{t_{-10}} + \beta_2 absSPY_{t-10} + \mu$$
$$absITA_{t_1} = \beta_1 GSI_{t_{-11}} + \beta_2 absSPY_{t-11} + \mu$$
$$absITA_{t_1} = \beta_1 GSI_{t_{-12}} + \beta_2 absSPY_{t-12} + \mu$$

where:

$t_{-1,2,\dots,12}$ are lag of months

In this case, we subject both independent variables to delays while the dependent variable remains constant. In this way, we can compare the regressions of all lags against the dependent variable with an unrestricted Residual Sum of Squares.

3 RESULTS

This section represents the culmination of the data and methodology discussed in the primary part of this research and provides a comprehensive summary of the outcomes and their corresponding interpretations within the two models mentioned above.

3.1 Results of OLS

The OLS model is valuable for determining the relationship between independent and dependent variables. In addition, it allows us to establish a more accurate cause-and-effect relationship between the variables. Finally, it lets us recognize whether the GSI index, as an indicator of geopolitical stress, really has such an impact on the stock market related to the defense industry, as we assumed. Considering that we are using the monthly data, 200 observations are an appropriate sample for running this model.

After processing the OLS model in the Gretl program with our data, where the dependent variable is absITA and the independent variables are our absSPY and GSI index for the whole observed period with one benchmark peek in the search query, we will get the following result:

Table 2. The results of OLS model with GSI one period

Depended variable:	absITA		
Time period:	2006/07 - 2023/02		
Variable	*Coefficient*	*p-value*	
const	-0,042	0,086	*
GSI	0,247	0,053	*
absSPY	1,086	1,87E-38	***
Observation		200	
Sum squared resid		0,087	
R-Squered		0,604	
Durbin-Watson		1,851	
F-Statistic		*150,136*	
P-value (F)		2,46E-40	

Source: Own Analysis in Gretl processed in Microsoft Excel 2022

We want to use the regression results shown above to demonstrate how the results will be modified after we adapt the GSI index methodology to be more accurate. The P-value of our GSI index in this regression is more than 0.05 but for a small amount. However, it still means that the result is insignificant at the 5% level. Therefore, this result also motivated us to modify the GSI index data further. Due to the nature of the GSI index data, we used the method of several monitored periods, which we explained in more detail in the methodology part of this work. Based on this, we obtained the following regression results:

Table 3. The results of OLS model with GSI multiple period's windows

Depended variable:	absITA		
Time period:	2006/07 - 2023/02		
Variable	*Coefficient*	*p-value*	
const	-0,044	0,069	*
GSI	0,213	0,042	**
absSPY	1,086	2,57E-38	***
Observation		200	
Sum squared resid		0,087	
R-Squered		0,605	
Durbin-Watson		1,853	
F-Statistic		*150,636*	
P-value (F)		2,02E-40	

Source: Own Analysis in Gretl processed in Microsoft Excel 2022

The above table presents a regression analysis conducted on the dependent variable, ITA ETF monthly returns in absolute values, from July 2006 to February 2023. In addition, the table presents the regression coefficients for the independent variables with their associated p-values.

The coefficient for the GSI variable is 0,213 with a p-value of 0,042, indicating a statistically significant relationship between this independent variable and the dependent variable at the 95% confidence level. The results indicate a positive relationship in the given model between the GSI independent variable and the absITA dependent variable. In other words, if the Geopolitical Stress Index increase by 1 unit, and the absSPY variable is 0, then ceteris paribus the returns of ITA ETF in absolute numbers also increase by 0,213. So far, results suggest a potential conflict with Efficiency Market Hypothesis. In this case, if investors will closely monitor geopolitical developments via GSI and understand the impact of geopolitical events on the market and investors' behaviour and expectations, they should be able to generate more increased returns. Nevertheless, it could pose significant risks if the geopolitical stress event leads to more expansive market fluctuation or economic downturns. Therefore, we want to highlight the curse of knowledge behavioural bias in this particular point. We acknowledge that the results presented above can lead to false speculation and create an incorrect feeling of confidence in the accuracy of the data and its decisions.

The coefficient for the absSPY variable is 1,086, with an extremely low p-value of 2,57E-38, indicating a highly statistically significant relationship between this independent variable and the dependent variable. If the "absSPY" variable increases by 1 unit, and another independent variable is 0, than we can ceteris paribus expect the "absITA" to increase by 1,086 units. We internally assume this number is exceptionally high. It signifies that the absSPY variable should be an essential predictor of the dependent variable and that it significantly impacts the returns in absolute numbers of ITA ETF. Given this result, the returns of SPY ETF are, as the independent variable absSPY, significant on the Alpha 0,05. The variable absSPY also shows a positive relationship with the dependent variable, and in this case, we cannot argue similarly to the case of the variable GSI in this model. Regardless, in this case, we do not dare to claim that there may be a conflict with the Efficient Market Hypothesis specifically because it is a benchmark ETF sensitive to similar factors influencing the stock market as the ITA ETF. It may eventually mean only equivalent reactions on causality that simultaneously affect both ETFs based on their identical asset types. Consequently, this outcome is unsurprising for us, and we do not pay much attention to it.

The regression also includes a constant term with a coefficient of -0,044 and a p-value of 0,069. However, the constant term may not be influential in our regression models since it is improbable that the GSI index and change in SPY ETF returns would be equal to zero in real-world situations. Furthermore, the independent variable represents absolute numbers. Therefore, negative coefficient is inconceivable.

The R-squared value for the regression is 0,605, which implies that the independent variables explain 60,5% of the variation in the dependent variable. It is a relatively high number. However, since we include the market factor in the model, i.e., absSPY, not a surprise.

The Durbin-Watson statistic is 1,853, indicating no significant autocorrelation in the residuals, where the critical values for Durbin-Watson statistics are dL = 1,7483 and dU = 1,7887, we can deduce that there is no significant autocorrelation in the regression residuals at the 95% confidence level. The F-statistic for the regression is 150,636, accompanied by an extremely low p-value of 2.02E-40. This result reveals that our regression with the OLS model is significant overall.

Furthermore, the Sum Squared residuals (SSR) in the model are relatively good, namely 0,087. So, given this data, the slight difference in the residuals indicates that the model is relatively good at explaining the variance in the data, and therefore the predicted values could be closer to the actual values. However, even though lower SSR can indicate that this model may be appropriate for assembling accurate forecasts, our independent variables are not subject to lag. Therefore, it suggests that the predictive ability of GSI and absSPY variables on the ITA ETF returns in absolute numbers can be only in less than one month.

When we sum it up, results suggest that our GSI index, created from specific geopolitical stress words, and absSPY, returns of benchmark ETF in absolute numbers, have a considerable effect on the dependent variable absITA, representative ETF for the defense sector. Moreover, it indicates that the model is acceptable for explaining the relationship between these variables.

In the end, these results outline some exciting highlights for us. First, our GSI index and the market portfolio are significant in this case without delay. The significance of the SPY index should be a surprise because it is a benchmark index, and we expected that it would considerably impact the ITA index. Nevertheless, regarding the GSI index, we can assume that, in this case, the interest in geopolitical events within one month is well reflected in the movement of the returns of the ITA index. We internally assumed such results since, in general, we can assume that events causing geopolitical stress will substantially impact investors' sentiment precisely at the peak of their popularization, which we can determine as a moment when they could be so-called "hot news".

As results show in this case, the impact of Google Trends data on stock market returns is particularly relevant to the research on the semi-strong form of efficient markets. Given the effect of investor sentiment on financial markets, changes in the GSI index may reflect changes in investor sentiment as investors search for information related to their beliefs about future earnings of the defense sector, with should be closely related to geopolitical stress events. As such, thanks to this, monitoring changes in the GSI index may provide helpful insights into investor sentiment and its potential influence on market behaviour. As we can assume, the GSI index is especially relevant to the research on the semi-strong form of efficient markets. Investors whose fundamental research depends excessively on Google Trends data may experience various behavioural biases, leading to incorrect investment decisions because, as we can see, the market portfolio as the SPY index has a significant real-time impact. These results prove a significant regression between the index formed from terms related to geopolitical stress and the ETF representing the defense sector on the stock market. However, without lags, we cannot analyse the predictive ability of these findings. Therefore, we further subjected the OLS data to lags that could provide better information about the relationship between variables in different periods. The results of OLS models with various lags periods for independent variables are in the following table:

Table 4. The results of OLS models with various lags

Depended variable: absITA

Time period: 2006/07 - 2023/03

Lag 1				**Lag 7**			
Variable	const	GSI	absSPY	**Variable**	const	GSI	absSPY
Coefficient	-0,040	0,275	0,390	**Coefficient**	-0,031	0,262	0,183
p-value	0,274	0,081	0,000	**p-value**	0,421	0,115	0,091
R-Squered		0,103		**R-Squered**		0,036	
P-value (F)		0,000		**P-value (F)**		0,031	
Sum squred resid		0,196		**Sum squred resid**		0,209	
Lag 2				**Lag 8**			
Variable	const	GSI	absSPY	**Variable**	const	GSI	absSPY
Coefficient	-0,066	0,422	0,092	**Coefficient**	0,015	0,072	0,113
p-value	0,080	0,010	0,379	**p-value**	0,693	0,667	0,305
R-Squered		0,045		**R-Squered**		0,008	
P-value (F)		0,012		**P-value (F)**		0,465	
Sum squred resid		0,207		**Sum squred resid**		0,215	
Lag 3				**Lag 9**			
Variable	const	GSI	absSPY	**Variable**	const	GSI	absSPY
Coefficient	-0,007	0,153	0,239	**Coefficient**	0,019	0,046	0,206
p-value	0,850	0,349	0,024	**p-value**	0,618	0,785	0,061
R-Squered		0,037		**R-Squered**		0,021	
P-value (F)		0,026		**P-value (F)**		0,132	
Sum squred resid		0,209		**Sum squred resid**		0,211	
Lag 4				**Lag 10**			
Variable	const	GSI	absSPY	**Variable**	const	GSI	absSPY
Coefficient	0,052	-0,120	0,425	**Coefficient**	0,030	-0,001	0,207
p-value	0,161	0,452	5,93E-05	**p-value**	0,436	0,995	0,062
R-Squered		0,081		**R-Squered**		0,020	
P-value (F)		0,000		**P-value (F)**		0,158	
Sum squred resid		0,199		**Sum squred resid**		0,212	
Lag 5				**Lag 11**			
Variable	const	GSI	absSPY	**Variable**	const	GSI	absSPY
Coefficient	0,014	0,066	0,226	**Coefficient**	0,048	-0,061	0,065
p-value	0,719	0,690	0,04	**p-value**	0,225	0,723	0,559
R-Squered		0,027		**R-Squered**		0,002	
P-value (F)		0,073		**P-value (F)**		0,822	
Sum squred resid		0,211		**Sum squred resid**		0,215	
Lag 6				**Lag 12**			
Variable	const	GSI	absSPY	**Variable**	const	GSI	absSPY
Coefficient	-0,044	0,319	0,167	**Coefficient**	0,046	-0,050	0,032
p-value	0,251	0,054	0,121	**p-value**	0,253	0,777	0,775
R-Squered		0,040		**R-Squered**		0,001	
P-value (F)		0,019		**P-value (F)**		0,937	
Sum squred resid		0,208		**Sum squred resid**		0,215	

Source: Own Analysis in Gretl processed in Microsoft Excel 2022

Because we worked only with variables without any delays in the regressions shown in Tables 2 and 3, we could not verify individual independent variables' predictive ability and lag effect on the ITA index. Therefore, we investigated every single lag for the independent variables in a separate OLS model within a one-year horizon representing 12 lags. Each model contains a coefficient and a p-value for all monitored variables. We also present the model's R-squared, P-value, and Sum squared residual to demonstrate its statistical significance.

The OLS model, in which we subjected the variable GSI and absSPY to a 1-month delay, shows that in this period, GSI, compared to the model where they are variable without delay, shows a decrease in statistical significance, unlike absSPY, which is still a statistically very significant independent variable even though that its coefficient shows the impact on the movement of the dependent variable is less than half that in OLS model without lag. Accordingly, the model's results with this lag suggest that the GSI index, which represents search demand for words associated with geopolitical stress, does not significantly impact the movement of the ITA index returns instead of the SPY index. However, as mentioned above, we can see the reason in the topicality, which can influence investors to the greatest extent. Therefore, we assume that influence is more remarkable, specifically when the events that cause geopolitical stress are the most topical.

What particularly interested us in the results of the OLS models, where we work with lags, is precisely the 2-month lag of the independent variables, where our GSI reached a p-value of less than 0.01, which means that it has statistical significance at the 1% level. In this case, even more than the absSPY index, which performed worse and is statistically significant only at the 5% level. We can explain these results from several behavioral biases, leading investors to underreact the new information. For instance, considering conservatism bias, we have just revealed when a conservative investor reacts to events causing geopolitical stress. As a result, it could be why this reaction on the stock market with a focus on the defense sector appears only after a specific time. The other OLS models with lags do not offer as exciting results as particularly this one.

Nevertheless, if we look at the significance of the model as a whole, the r-squared in the one-month lag is 0,103, which means that the independent variables explain only 10.30% of the variability of the dependent variable, given that this is the highest measured value of all lags and, at the same time, we can talk about six times smaller value than in the case of the OLS model, whose independent variables do not experience a lag. Therefore, in broad, but also compared to our OLS model without lags, these models' significance is too small. This statement also confirms the results of the Sum squared residual, whose values range from 0,196 to 0,215, which is approximately two and a half times more than in the case of the OLS model, where we did not use lags. These measures display a poor fit and suggest that the model may not be suitable for making accurate predictions, which means that according to these data, we cannot predict the ITA index returns with returns of the benchmark ETF index or even with our geopolitical stress index.

These results also indicate that the efficient market hypothesis is valid for the ITA index returns. According to the hypothesis in semi-strong tests, financial markets are efficient, and all available information is immediately incorporated into asset prices, making it impossible to outperform the market. Furthermore, as mentioned, the inadequate fit of the OLS model indicates that the information provided by the independent variables does not help predict the ITA index returns, supporting the idea that the market is efficient.

3.2 Results of VAR

As we mentioned, in the methodology part of this work, in our VAR model, we work only with the endogenous variables in the system. Our VAR model, with 12 lags and 24 explanatory variables in each of two equations, including lagged values of GSI and another endogenous variable, can help explain the dynamics and relationships among these variables over time and endogenous variable absITA. In General, the model should provide an understanding of the interdependencies and causal relationships among the variables included in the analysis over multiple time series. The results of our VAR model are in the following table:

Table 5. The results of VAR model

Time period:	2006/07 - 2023/03	**AIC**	-9,916
Observations	188	**BIC**	-9,055
Log-likelihood	982,102	**HQC**	-9,567
Determinant of covariance matrix	9,94E-08	**Portmanteau test**	LB(47) = 158,296, df = 140 [0,1382]

Equation 1: absITA	*Coefficient*	*p-value*		**Equation 2:** GSI	*Coefficient*	*p-value*	
const	-0,016	0,810		const	0,046	0,063	*
absITA_1	0,322	9,68E-05	***	absITA_1	0,032	0,278	
absITA_2	-0,051	0,541		absITA_2	0,074	0,017	**
absITA_3	0,084	0,327		absITA_3	-0,052	0,098	*
absITA_4	0,035	0,690		absITA_4	0,003	0,929	
absITA_5	0,046	0,601		absITA_5	-0,045	0,167	
absITA_6	0,028	0,749		absITA_6	-0,005	0,884	
absITA_7	0,023	0,794		absITA_7	-0,023	0,473	
absITA_8	0,020	0,818		absITA_8	0,027	0,396	
absITA_9	0,102	0,246		absITA_9	-0,002	0,950	
absITA_10	-0,030	0,734		absITA_10	0,018	0,563	
absITA_11	-0,090	0,301		absITA_11	-0,005	0,870	
absITA_12	0,011	0,891		absITA_12	-0,005	0,859	
GSI_1	0,044	0,839		GSI_1	0,470	2,02E-08	***
GSI_2	0,464	0,054	*	GSI_2	0,041	0,644	
GSI_3	-0,155	0,521		GSI_3	-0,004	0,964	
GSI_4	-0,325	0,180		GSI_4	0,083	0,346	
GSI_5	0,004	0,988		GSI_5	-0,007	0,933	
GSI_6	0,343	0,158		GSI_6	0,202	0,024	**
GSI_7	0,088	0,716		GSI_7	-0,097	0,276	
GSI_8	-0,341	0,161		GSI_8	-0,086	0,331	
GSI_9	0,161	0,507		GSI_9	0,012	0,894	
GSI_10	0,101	0,675		GSI_10	-0,079	0,367	
GSI_11	-0,162	0,493		GSI_11	0,103	0,237	
GSI_12	-0,078	0,722		GSI_12	0,167	0,038	**
R-squared	0,225			**R-squared**	0,457		
Durbin-Watson	2,004			**Durbin-Watson**	1,962		
F-Statistic	1,971			***F-Statistic***	5,709		
P-value (F)	0,007			**P-value (F)**	3,12E-12		
Sum squared resid	0,167			**Sum squared resid**	0,022		
F-tests of zero restrictions				*F-tests of zero restrictions*			
All lags of absITA	F(12, 163) = 2,2075 [0,0135]			All lags of absITA	F(12, 163) = 1,1208 [0,3466]		
All lags of GSI	F(12, 163) = 0,98443 [0,4660]			All lags of GSI	F(12, 163) = 8,6335 [0,0000]		

Source: Own Analysis in Gretl processed in Microsoft Excel 2022

As we can see in Table 5., the model's log-likelihood is 982,102, which we can consider a high number, indicating the model's fit to the data. However, log-likelihood does not necessarily guarantee that the model is the best possible representation of the data. Nevertheless, as we can see on the top of Table no. 5, the AIC is -9,916, the BIC is -9,055, and the HQC is -9,567. These values are all negative, which is pleasing, and the AIC and BIC are particularly close, which denotes that the model is a proper fit. In addition, the Portmanteau test indicates no significant autocorrelation left in the residuals after fitting the model. The covariance determinant is equal to 9,94E-08, which is very small. Therefore, considering this number as the degree of correlation between the GSI and absITA, our two variables are highly correlated. Furthermore, in the context of Granger causality, our covariance determinant suggests that our variables can help predict each other, but for a more helpful explanation, we should look at the coefficients and p-values of explanatory variables.

Looking at the right side of Table no. 5, where we observe equation 1 with the dependent variable absITA, we can see only one coefficient whose p-value is significant at the 1% level. It is one month lag of absITA itself. As mentioned in this part of the work, it is not unusual for variables such as ours to have a statistically significant relationship with themselves without a lag, especially in a VAR model. In other words, it indicates that the past returns of the ITA ETF could be a valuable predictor also of its current returns.

From all explanatory variables in Equation 1, we find no other significant p-value at the 5% percent level or better. However, the nearest it can be the GSI with a delay of 2 lags, the second most significant result. It reaches a p-value of 0,054, almost significant at the 5% level. We could observe a similar result in the first part of our research, where we followed the lags of simple OLS models, the independent variable GSI came out as significant in the model, where our independent variables experienced two lags, and in this case, GSI variable was significant even at the 1% level. As we have already explained, we consider this a consequence of behavioural biases and, regardless of that, the late reaction to new information.

Moreover, Equation 1 with the dependent variable absITA has an r-squared equal to 0,225, a relatively adequate number. Accordingly, because we observe only two endogenous variables in our VAR model, we can consider the explanation of 22,5% of variability as a decent result. Furthermore, the result of the Durbin-Watson statistic is 2,004, which suggests no significant autocorrelation in the errors, which is also good. Furthermore, F-statistics with a corresponding p-value of 0,007 is a relatively small number. This result indicates that Equation 1 as a model is significant and confirms that at least one of the explanatory variables is significant. In our case, it is absITA with one month lag. The SSR of 0,167 is a slight disappointment for us. Although, compared to the other models that we performed, we can call this number relatively high, it indicates that the model does not fit the data precisely and should not be able to make correct predictions.

In the F-test of zero restrictions, all lags of absITA have an F-statistic value of 2,208 with a corresponding p-value of 0,014, representing a statistically significant result. As we have already mentioned, this may be the result of the influence of past returns on the future returns of the ITA ETF. On the contrary, the results of all lags of GSI show an F-statistic of 0,984 with a corresponding p-value of 0,466, which is an insignificant result, and therefore the lags of GSI are not able to explain the current value of absITA. This result conflicts with our hypothesis that GSI can predict the movements of ITA returns, and at this point, the VAR model refutes this hypothesis.

Unlike the first equation, the second explains the movement of a search query of geopolitical terms in our index by lagged versions of our endogenous variables. Equation GSI presents several statistically significant explanatory variables. The most statistically significant variable is the GSI, with a one-month lag, with a p-value of 2,02E-08, which is significant on the 1% level. In this case, we think there is a resonance in which the searching query for some event transforms into the current query. In other words, it means that the past values of the search query affect the current values. Other GSI lags, such as 6-month and 12-month, also reach statistical significance at a 5% level. We consider this a possible consequence of some commemorations and seasonal events, which are highly likely to occur in such periods. However, compared to the 1-lag GSI coefficient of 0,470, the impact of these explanatory variables is barely half. More surprising was the lag absITA results in Equation 2. These indicate a statistically solid significance within the explanatory variable with two lags. Moreover, according to this, the p-value of the relevant variable at 0,017 is statistically significant at the 5% level. We assume that this result suggests a possible impact of ITA ETF returns on search demand for terms in GSI with a 2-month delay.

The R-squared in our second Equation reaches a remarkable value of 0,457. The fact that our two explanatory variables can explain up to 45,7% of the variability of the dependent variable in this model should attract our attention. The Durbin-Watson statistical test is 1,962. These results uncover the potential of autocorrelation existence in our Equation 2 model. Therefore, we make an additional test for autocorrelation presence in our model for single lags. We display the results in the following table:

Table 6. The results of Autocorrelation test in VAR model

	Rao F	Approx dist.	p-value
lag 1	1,134	F(4, 320)	0,340
lag 2	1,617	F(8, 316)	0,119
lag 3	1,353	F(12, 312)	0,188
lag 4	1,311	F(16, 308)	0,189
lag 5	1,175	F(20, 304)	0,275
lag 6	1,003	F(24, 300)	0,462
lag 7	0,98	F(28, 296)	0,497
lag 8	0,912	F(32, 292)	0,608
lag 9	1,056	F(36, 288)	0,388
lag 10	0,981	F(40, 284)	0,509
lag 11	0,948	F(44, 280)	0,569
lag 12	1,1	F(48, 276)	0,314

Source: Own Analysis in Gretl processed in Microsoft Excel 2022

As shown in the table above, p-values for all lags are insignificant at 5%. It means that significant autocorrelation is non-existent, and therefore the model is not contrary to the assumption of independently and identically distributed error. Therefore, we can use the model to draw valid statistical inferences and make a possible prediction on that. The F-statistic of 5,709 with the associated very low p-value 3,12E-12 confirms the overall significance of the model.

Regarding looking at the Sum squared residual result of the VAR model in the second equation, we can see that the model where the explanatory variable is the GSI has an SSR of 0,022, which is a relatively small number and indicates that the predicted values should be close to the actual values, even if we cannot interpret these results with an unmistakable probability, we can claim that based on the results of the VAR model of both equations, there is a greater probability of GSI prediction using the results of the ITA index than vice versa. In this case, the VAR model indicates that the ITA index, primarily with a delay of 2 months, can have a significant impact at the 5% level for the GSI search, but it is necessary to mention that its delayed values also have a significant influence on the GSI, namely a delay of 1 month, which is significant at the 1% level.

If we sum it up, in the context of Granger causality, the results of the VAR model indicate that the ITA returns in absolute numbers significantly impact the GSI search, which implies that the ITA index Granger causes the GSI. Furthermore, our VAR analysis results also suggest that the delayed values of the ITA index significantly influence the GSI, which implies that the past values of the ITA index can predict the current and future values of the GSI. Nevertheless, as we can see in Equation 1, the results do not provide relevant evidence of reverse causality, which means that the GSI does not Granger cause the ITA index, as we initially assumed. Therefore, based on the results of the VAR model, it is reasonable to conclude that the ITA index can be a valuable predictor of the GSI search, and the delayed values of the ITA index can significantly influence the GSI search and not vice versa, as we expected.

4 DISCUSSION & CONCLUSION

The primary purpose of our work was the analysis of geopolitical stress converted into Google Trends search data on the development of the returns of the defense sector, whose representative in our work was the iShares US Aerospace & Defense ETF. To smooth out other conditions that can primarily affect the ITA ETF, we used the benchmark market portfolio SPDR S&P 500 ETF, whose returns represent the market's overall performance. Subsequently, we analysed these variables in detail in several OLS and VAR models. Moreover, we interpreted the results in light of existing literature.

Our analysis found a statistically significant relationship between Google search volumes and ITA ETF returns in the OLS model without lags. This result indicated that our Geopolitical Stress Index could be a potentially valuable tool for predicting stock returns, especially in the defense sector. However, this model requires further research to explore the full possibilities of the smaller period sample, mainly because our data was monthly. Therefore, at this point, we will point out the biggest weakness of our research, mainly because our research suggests that search queries related to geopolitical stress, recast in our GSI index, which is the most recent, have the most significant impact.

As for the VAR model, our tool for determining Granger causality, the result did not agree with our inner belief and pointed out that the ITA index most conceivably causes the GSI. However, this result is a submission of the demand for closer academic research examining the potential of ITA index returns as a proxy indicator of geopolitical stress for the general public, whose reaction reflects in search engines after two months.

This research implies that the GSI index could help predict ITA ETF returns. Nevertheless, exploring the full potential of the GSI index, particularly with a weekly or even daily sample period, is needed. In terms of avenues for further research, our study suggests a more extended investigation focus on the causes of the ITA index on the GSI index, which could be a valid predictor of future search queries.

Additionally, we want to highlight the need for a closer examination of the relationship between search engine data and stock returns as a research topic with considerable potential to contradict with Efficiency Market Hypothesis.

REFERENCIES

Agnew, J., 2005. Hegemony: The New Shape Of Global Power. Namur: Temple University Press.

Agnew, J.A., 2003. Geopolitics: re-visioning world politics. London: Routledge. 10.4324/9780203015612.

Baker, M., Wurgler, J., 2006. Investor Sentiment and the Cross-Section of Stock Returns. The Journal of finance (New York), 61(4), pp.1645–1680. 10.1111/j.1540-6261.2006.00885.x.

Balcilar, M. et al., 2018. Geopolitical risks and stock market dynamics of the BRICS. Economic systems, 42(2), pp.295–306. 10.1016/j.ecosys.2017.05.008.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. Journal of financial economics, 49(3), pp.307–343. 10.1016/S0304-405X(98)00027-0.

BBC News. (2015, November 14). Paris attacks: What happened on the night. BBC News. Retrieved from https://www.bbc.com/news/world-europe-34818994. Accessed 02. Apr. 2023.

Behrendt, S., Khanna, P. 2003. Geopolitics and the Global Corporation. Strategy & Business.

BlackRock. (2022). iShares U.S. Aerospace & Defense ETF (ITA) Fund Fact Sheet. Retrieved from https://www.ishares.com/us/literature/fact-sheet/ita-ishares-u-s-aerospace-defense-etf-fund-fact-sheet-en-us.pdf. Accessed 02. Apr. 2023.

Brooks, C., 2019. Introductory Econometrics for Finance. Cambridge University Press. 10.1017/9781108524872.

Caldara, D., Iacoviello, M., 2018. Measuring Geopolitical Risk. IDEAS Working Paper Series from RePEc. 10.17016/IFDP.2018.1221.

Camerer, C., Loewenstein, G., Weber, M., 1989. The Curse of Knowledge in Economic Settings: An Experimental Analysis. The Journal of political economy, 97(5), pp.1232–1254. 10.1086/261651.

CBS News. (2006, November 15). Violence, Confusion Reign In Baghdad. Retrieved from https://www.cbsnews.com/news/violence-confusion-reign-in-baghdad/. Accessed 11. Apr. 2023.

Choi, H., Varian, H., 2012. Predicting the Present with Google Trends. The Economic record, 88(s1), pp.2–9. 10.1111/j.1475-4932.2012.00809.x.

Cohen-Charash, Y. et al., 2013. Mood and the market: can press reports of investors' mood predict stock prices? PloS one, 8(8), pp.e72031–e72031. 10.1371/journal.pone.0072031.

Cohen, S.B., 2015. Geopolitics : the geography of international relations Third edition. Lanham, [Maryland: Rowman & Littlefield.

Fama, E.F., 1965. The Behavior of Stock-Market Prices. The Journal of business (Chicago, Ill.), 38(1), pp.34–105. 10.1086/294743.

Fama, E.F., 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of finance (New York), 25(2), p.383–. 10.2307/2325486.

Fama, E.F., 1991. Efficient Capital Markets: II. The Journal of finance (New York), 46(5), pp.1575–1617. 10.1111/j.1540-6261.1991.tb04636.x.

Fama, E.F., 1998. Market efficiency, long-term returns, and behavioral finance. Journal of financial economics, 49(3), pp.283–306. 10.1016/S0304-405X(98)00026-9.

Fama, E.F., French, K.R., 2007. Disagreement, tastes, and asset prices. Journal of financial economics, 83(3), pp.667–689. 10.1016/j.jfineco.2006.01.003.

Financial Times. (2020, March 29). Trump decides against coronavirus quarantine of states. Retrieved from https://www.ft.com/content/f1bb743c-5235-4056-b6bf-8f68a0a38878. Accessed 11. Apr. 2023.

Financial Times. (2023). iShares U.S. Aerospace & Defense ETF (ITA:BTQ:USD) - ETF. Retrieved from https://markets.ft.com/data/etfs/tearsheet/summary?s=ITA:BTQ:USD. Accessed 02. Apr. 2023.

Fischhoff, B., 2003. Hindsight ≠ foresight: the effect of outcome knowledge on judgment under uncertainty. Quality & Safety in Health Care, 12(4), pp.304–311. 10.1136/qhc.12.4.304.

Flint, C., Taylor, P.J., 2018. Political Geography: World-Economy, Nation-State and Locality 7th ed. Milton: Routledge. 10.4324/9781315164380.

Google. About Google Trends data. Retrieved from https://support.google.com/trends/answer/4365533?hl=sk&sjid=14852739257502267561-EU. Accessed 06. Apr. 2023.

Granger, C.W.J., 1969. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. Econometrica, 37(3), pp.424–438. 10.2307/1912791.

Greene, W.H., 2020. Econometric analysis Eighth, Global edition. Harlow, England: Pearson.

Hamilton, J.D., 1994. Time series analysis. Princeton, N.J: Princeton University Press.

Hodnett, K., Hsieh, H.-H., 2012. Capital Market Theories: Market Efficiency Versus Investor Prospects. The international business & economics research journal, 11(8), p.849–. 10.19030/iber.v11i8.7163.

Huang, M.Y., Rojas, R.R., Convery, P.D., 2020. Forecasting stock market movements using Google Trend searches. Empirical economics, 59(6), pp.2821–2839. 10.1007/s00181-019-01725-1.

Jung, S., Lee, J., Lee, S., 2021. The impact of geopolitical risk on stock returns: Evidence from inter-Korea geopolitics. IDEAS Working Paper Series from RePEc.

Keynes, J.M., 2013. The collected writings of John Maynard Keynes. Volume 7, The general theory of employment, interest and money. [New edition]. Cambridge: Cambridge University Press for the Royal Economic Society.

Krawczyk, D.C., Baxter, G.H., 2020. Understanding behavioral BIAS : a guide to improving financial decision-making. New York, New York: Business Expert Press.

Kuersteiner, G.M. 2010. Granger-Sims causality. in Durlauf, S.N. and Blume, L.E. (eds) Macroeconometrics and Time Series Analysis. The New Palgrave Economics Collection. London: Palgrave Macmillan, pp. 273-287. 10.1057/9780230280830_14.

Kumar, A. et al., 2020. COVID-19 impact on sustainable production and operations management. Sustainable Operations and Computers, 1, pp.1–7. 10.1016/j.susoc.2020.06.001.

Kurecic, P., 2015. Geoeconomic and Geopolitical Conflicts: Outcomes of the Geopolitical Economy in a Contemporary World. World review of political economy, 6(4), pp.522–543. 10.13169/worlrevipoliecon.6.4.0522.

NBC News. (2015, November 14). Paris Attacks: ISIS Claims Responsibility for 129 Dead in Multiple Terrorist Strikes. NBC News. Retrieved from https://www.nbcnews.com/storyline/paris-terror-attacks/isis-claims-responsibility-deadly-paris-terror-attacks-n463446. Accessed 02. Apr. 2023.

NPR. (2006, November 17). Afghan Army Making Progress, Still Reliant on NATO. Retrieved from https://www.npr.org/2006/11/17/6499377/afghan-army-making-progress-still-reliant-on-nato. Accessed 11. Apr. 2023.

Peterson, R.L., 2016. Trading on sentiment : the power of minds over markets. Hoboken, New Jersey: Wiley.

Pompian, M.M., 2012. Behavioral finance and wealth management : how to build investment strategies that account for investor biases 2nd ed. Hoboken, N.J: Wiley.

Preis, T., Moat, H.S., Stanley, H.E., 2013. Quantifying trading behavior in financial markets using Google Trends. Scientific reports, 3(1), pp.1684–1684. 10.1038/srep01684.

Shiller, R. 2016. Irrational Exuberance. Princeton: Princeton University Press. 10.1515/9781400865536.

Shiller, R.J., 2003. From Efficient Markets Theory to Behavioral Finance. Journal of Economic Perspectives, 17(1), pp.83–104. 10.1257/089533003321164967.

Shleifer, A., 2000. Inefficient markets : an introduction to behavioral finance. Oxford: Oxford University Press.

Sims, C.A., 1972. Money, Income, and Causality. The American economic review, 62(4), pp.540–552.

Sims, C.A., 1980. Macroeconomics and Reality. Econometrica, 48(1), pp.1–48. 10.2307/1912017.

Tetlock, P.C., 2007. Giving Content to Investor Sentiment: The Role of Media in the Stock Market. The Journal of finance (New York), 62(3), pp.1139–1168. 10.1111/j.1540-6261.2007.01232.x.

The Guardian. (2015, November 14). Paris attacks: What we know so far. The Guardian. Retrieved from https://www.theguardian.com/world/2015/nov/14/paris-attacks-what-we-know-so-far. Accessed 02. Apr. 2023.

The Guardian. (2020, March 30). 'Treated worse than criminals': Australian arrivals put into quarantine lament conditions. Retrieved from https://www.theguardian.com/world/2020/mar/30/treated-worse-than-criminals-australian-arrivals-put-into-quarantine-lament-conditions. Accessed 11. Apr. 2023.

Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: heuristics and biases. Science, 185(4157), pp. 1124-1131.

United Nations. (2006, December 6). Fiji must abide by global obligations on rights, fundamental freedoms: UN rights chief. Retrieved from: https://news.un.org/en/story/2006/12/202182-fiji-must-abide-global-obligations-rights-fundamental-freedoms-un-rights-chief. Accessed 11. Apr. 2023.

Wooldridge, J.M., 2020. Introductory econometrics a modern approach Seventh edition. Australia: Cengage.

Yahoo Finance (2023). iShares U.S. Aerospace & Defense ETF. Retrieved from https://finance.yahoo.com/quote/ITA?p=ITA. Accessed 02. Apr. 2023.