



MATE Hungarian University of Agriculture and Life Sciences

**Modeling and Forecasting Market Dynamics around Geopolitical
Shocks using Temporal Fusion Transformers**

DOI: 10.54598/005820

Steffen Robus, Campus of Kaposvár, 2025

The PhD School

Name: Doctoral School of Economic and Regional Sciences

Discipline: Management and Organizational Sciences

Head: Prof. Dr. Zoltán Bujdosó, PhD, Head of Doctoral School of Economic and Regional Sciences, H-7400 Kaposvár, Guba Sándor str. 40.

Supervisor(s): Dr. Virág Walter, PhD, Associate Professor, MATE Hungarian University of Agriculture and Life Sciences, Institute of Agricultural and Food Economics, Department of Agricultural Management and Leadership Science, H-7400 Kaposvár, Guba Sándor str. 40.

Dr. Zsolt Kőműves, Dr. habil., Associate Professor, MATE Hungarian University of Agriculture and Life Sciences, Institute of Agricultural and Food Economics, Department of Agricultural Management and Leadership Science, H-7400 Kaposvár, Guba Sándor str. 40.

.....
Approval of the
Head of Doctoral School

.....
Approval of the Supervisor(s)

.....
Approval of the Supervisor(s)

(NB: The doctoral dissertation has to be submitted in four copies bound with the original signatures and the ten theses with the photocopied signatures to the Office of the PhD School concerned.)

CONTENT

1. INTRODUCTION.....	5
1.1. Geopolitical shocks and their impact on financial markets.....	5
1.2. Volatility as a measure of uncertainty in financial markets.....	7
1.3. Research scope	8
1.4. Robustness of the results	12
1.5. Contribution to research and outline	13
2. LITERATURE OVERVIEW	14
2.1. Geopolitical Shocks and their impact on financial markets	14
2.2. Modeling and forecasting volatility in times of geopolitical shocks	20
2.2.1. Modeling and forecasting volatility in times of geopolitical shocks using GARCH-class models.....	22
2.2.2. Modeling and forecasting market dynamics using the Temporal Fusion Transformer	24
3. MATERIAL AND METHODS	31
3.1. Definitions of volatility	33
3.2. Abnormal Returns and Abnormal Volatility	34
3.3. GARCH-Class Models	35
3.4. The Temporal Fusion Transformer for volatility forecasting.	37
3.5. Forecasting Evaluation	42
4. RESULTS.....	59
4.1. Abnormal Returns for MSCI World Sector Indices	59
4.2. Identification of Abnormal Volatility	64
4.2.1. Abnormal Volatility	64
4.2.2. Cumulated Abnormal Volatility	68
4.2.3. Statistically significant differences of Abnormal Volatility	70
4.3. Univariate Volatility Modeling and Forecasting of Financial Time Series during Geopolitical Shocks	72
4.3.1. In-Sample Analysis	72
4.3.2. Univariate Out-of-Sample Volatility Forecasting	76
4.4. Multivariate Volatility Modeling and Forecasting using the Temporal Fusion Transformer	92
4.5. Multivariate direct multi-step out-of-sample Volatility Forecasting using the Temporal Fusion Transformer.....	98
4.5.1. Analysis of multi-step out-of-sample forecasting performance	99
4.5.2. Regime switching multi-step out-of-sample forecasting performance.....	111
4.5.3. Influence of hyperparameter optimization on the multi-step out-of-sample forecasting performance	118
4.5.4. The added value of a multivariate structured volatility forecasting approach.....	127
4.5.5. Model comparison between GARCH and Temporal Fusion Transformer	134
5. CONCLUSION AND RECOMMENDATIONS.....	136
6. NEW SCIENTIFIC RESULTS	143
7. SUMMARY.....	144
APPENDIX 1: BIBLIOGRAPHY	147
APPENDIX 2: FURTHER MATERIAL.....	176
ACKNOWLEDGMENTS	194

1. INTRODUCTION

1.1. Geopolitical shocks and their impact on financial markets

Over time, events with global repercussions occur time and again. These can be described as geopolitical events or geopolitical shocks. Geopolitical shocks can be defined as significant events that can influence large parts of the world, affecting economic stability, political relations, and social dynamics (Baker et al., 2016; Bloom, 2009; Caldara & Iacoviello, 2022). These shocks often arise from military conflicts, political unrest, or sudden changes in international relations, leading to widespread uncertainty and volatility across multiple sectors. One of the primary characteristics of geopolitical shocks is their ability to induce immediate and substantial economic impacts (Caldara & Iacoviello, 2022). Furthermore, geopolitical shocks impact the volatility of all assets, asset classes, sectors, and countries worldwide (Engle & Campos-Martins, 2020). Concerning the impact on the financial markets, they are accompanied by increasing broad uncertainty, which in turn leads to changes in the behavior of international market players (Baker et al., 2016; Baur & Smales, 2020; Caldara and Iacoviello, 2022). Uncertainty in the financial markets is reflected here by the fact that opinions on the valuation of assets are more unclear, resulting in stronger selling and buying (Fama & French, 1993; Fama & MacBeth, 1973). However, it can be observed that the volatility is lower for good news than bad news, which is called the asymmetry effect, e.g., Alberg et al. (2008), Awartani and Corradi (2005), Byun and Cho (2013), Pan and Liu (2018), and Wang et al. (2020).

The question of whether a shock has a global impact can only be answered ex-post in review, since a cause-and-effect relationship has to be considered. One must distinguish here between the geopolitical risk and the possible geopolitical shock Caldara and Iacoviello (2022). The geopolitical risk can be quantified using various methods and thus indicates the likelihood of certain geopolitical events occurring. Since the expectations of future developments in the financial markets are based on the thesis of efficient capital markets, the occurrence of a geopolitical shock with a certain probability is already included (Fama, 1970). The last major geopolitical events that caused turbulence in the financial markets were the COVID-19 pandemic (e.g., Albuлесcu, 2021; Liu et al., 2020; Zhang et al., 2020), the Russian attack on Ukraine (e.g., Izzeldin et al., 2021; Izzeldin et al., 2023; Umar, 2022), and the attack of Hamas on Israel (e.g., Altemur et al., 2024; Ugli, 2024).

COVID-19 has been regarded as the most significant global threat since World War II and the greatest health crisis of the 21st century (Boccaletti et al., 2020; Kickbusch & Leung, 2020). The emergence of COVID-19 was first identified in Wuhan, China, in December 2019, following a notable rise in cases of pneumonia of unknown origin (WHO, 2020). The illness was subsequently named Coronavirus Disease 2019 (COVID-19) (Huang et al., 2020). The virus soon transcended national borders, with the initial reported death occurring on January 11, 2020. To limit transmission, authorities imposed a lockdown on Wuhan by January 23, 2020. Nevertheless, the situation soon escalated, causing the World Health Organization (WHO) to proclaim a public health emergency on January 30, 2020, of International Concern (PHEIC) on January 30, 2020 (Heymann & Shindo, 2020). Despite global containment strategies, the virus continued to spread

at an alarming rate, leading to its classification as a pandemic by the WHO on March 11, 2020 (Cucinotta & Vanelli, 2020). The pandemic has since resulted in widespread health crises and significant economic and societal disruptions across the globe (Lupton & Willis, 2021). The pandemic has triggered a global economic downturn, with many countries enforcing full or partial lockdowns to curb the virus's transmission. These restrictions have drastically slowed global economic activity, forcing many businesses to scale back operations or shut down entirely, resulting in widespread job losses (Fernandes, 2020). Not only have manufacturers and agricultural sectors been impacted, but service industries, education, sports, entertainment, and the food sector have also experienced significant declines (Hale et al., 2021). However, financial markets were also heavily affected. COVID-19 impacted the stock markets like no other infection, not even the Spanish Flu (Baker et al., 2020). With rising uncertainty, stock market volatility rose significantly (Albulescu, 2021). The most important volatility index worldwide (VIX) showed a spike on the March 16, 2020 of 85.11 points (CBOE, 2024). The simple average value of the index from 2014 to 2024 is 19.20 points (CBOE, 2024).

The second geopolitical shock I wanted to examine is the Russian attack on Ukraine in 2022. This Russian invasion on February 24, 2022, marks a significant geopolitical shift in Europe's trajectory since the 1990s (Ahmed et al., 2023; Umar et al., 2022). Following the dissolution of the Soviet Union, Europe experienced increasing market integration, among other developments (Mbah & Wasum, 2022). After a period of liberalization, Russia showed increasing autocratic tendencies (Boubaker et al., 2022). Although, Europe maintained strong ties with Russia particularly through trade relations like the import of natural gas and other resources (Grätz, 2022). As part of a maneuver, Russia is already taking up position on the borders of Ukraine in 2021. The invasion that followed on February 24, 2022, caused a global shock (Boubaker et al., 2022; Grossi & Vakulenko, 2022). As seen by sell-offs across major stock indices, uncertainty surrounding Russia's actions had already been partially factored into the markets in the days leading up to the invasion. However, the invasion itself was largely unexpected. This can be seen by various significant market reactions. All major stock indices experienced losses on the day (FRED, 2024a, 2024b, 2024c; Reuters, 2024; Robus et al., 2024). Meanwhile, crude oil prices surged above USD 100 per barrel, and natural gas prices spiked (FRED, 2024e). At the same time, volatility increased (e.g., Izzeldin et al., 2023; Robus et al., 2024; Umar et al., 2022). The VIX rose to roughly 39.00 points at the end of January 2022 (CBOE, 2024). Despite manifold sanctions and political pressure from the European Union, the United States of America and other Western states, Russia's war has been going on since 2022 (Chen et al., 2023; Van Bergeijk, 2022).

The third geopolitical event I was focusing on in this article is the Hamas attack on Israel in 2023. On October 7, 2023, Hamas, the Palestinian political and military organization, launched a significant and coordinated attack against Israel, resulting in a profound escalation of hostilities and marking a pivotal moment in the ongoing Israel-Palestine conflict. This attack resulted in a tragic death toll of approximately 1,400 Israelis, including civilians, soldiers, and police forces, while around 240 individuals were taken hostage (Saar-Ashkenazy et al., 2024). The surprise nature of this assault has drawn extensive scrutiny regarding Hamas's operational capabilities and

the failures in Israel's intelligence and preparedness, raising questions about the factors contributing to Hamas's sudden success (Bordás, 2024; Paché, 2024). The context of the attack is deeply rooted in decades of conflict marked by cycles of violence, historical grievances, and territorial disputes. Notably, the ongoing tensions and violence have led to significant humanitarian crises on both sides, exacerbated by historical narratives of displacement and violence dating back to pivotal events like the Nakba in 1948 (Alsous et al., 2024). The repercussions of the October 7 attack were immediate; Israel declared a state of war and initiated extensive aerial and ground offensives in Gaza, resulting in high civilian casualties and widespread destruction of infrastructure, reflective of the broader humanitarian impact of the escalation (Levy et al., 2024; Gómez et al., 2024). The international community has responded with mixed reactions. Legal claims regarding war crimes have surfaced, particularly concerning Israel's military actions in Gaza, leading to calls for investigations under international law, such as South Africa's recent legal proceedings against Israel at the International Court of Justice (Akbar & Genovés, 2024; Soraya et al., 2024). This escalation has contributed to significant instability in the markets, which remains uncertainty (Ugli, 2024). Around the day of the Hamas terrorist attack on the 7th of October 2023, in Israel, the VIX rose only moderately to around 23 points and then flattened out again to a normal level (CBOE, 2024). It is not yet possible to foresee the global impact of the conflict between Israel and Hamas. However, there is a clear risk of significant effects on the commodity markets, as well as the entry of additional parties supporting the respective sides (e.g., the US for Israel or Iran for Hamas) and thus an escalation of the conflict in the Middle East.

1.2. Volatility as a measure of uncertainty in financial markets

All these events have in common that they have led to increased uncertainty among market participants about the future. This is expressed in an increase in market volatility. Volatility is a crucial factor influencing investor decisions in financial markets and is fundamental for asset pricing and regulatory frameworks, particularly in risk assessment and management (Black & Scholes, 1973; Merton, 1976). It serves as a key measure of risk, helping investors and financial institutions understand the degree of price fluctuations in stocks or indices and enabling more effective risk management. For instance, higher volatility may deter risk-averse investors or prompt portfolio adjustments (Black, 1986). Volatility also plays a role in shaping hedging strategies used by investors and can influence decisions about the holding period of an investment. In addition, volatility is vital in pricing derivatives, particularly options. Increased volatility leads to a higher option premium since the likelihood of the option finishing in the money rises with greater price variability (Black & Scholes, 1973). Moreover, in economic indicators, volatility often signals economic uncertainty or instability. A sharp increase in volatility may reflect investor concerns over economic policies, geopolitical events, or financial crises. Analysts and policymakers closely monitor such fluctuations to assess economic conditions and formulate appropriate responses. In this context, regulators examine volatility and market behavior to ensure stability and develop regulations to protect investors while maintaining fair, orderly, and efficient markets (Robus et al., 2024). The volatility in the financial markets influences the behavior of market participants in these markets. However, increased volatility in the financial markets can

also affect other capital flows, such as direct investment. This is because volatility influences investors' risk assessment. High volatility tends to indicate a rather unstable economic or political situation, which increases uncertainty about future returns. As a result, direct investments, which represent long-term commitments, become less attractive. Dobrota et al. (2021) examined FDI net flows and used an ARIMA procedure to model the time series. Ghosh et al. (2014) showed that interest rate volatility has a negative effect on direct capital inflows. A negative impact of increased exchange rate volatility on receiving direct investment was shown, for example, by Christiano (2014), Hoque & Yakob (2017), Latief & Lefen (2018). Conversely, many studies examine the influence of direct investment on stock markets, e.g., Demir (2019), Elimam (2017), Gupta & Ahmed (2020).

There are various methods for measuring asset price volatility, and it is useful to summarize them here briefly. A well-known approach involves calculating the standard deviation, which can be estimated as the square root of the sample variance (Granger & Poon, 2001). The standard deviation is particularly suitable for normally distributed data, serving as an appropriate measure of dispersion. Other approaches aimed at mitigating bias from extreme values have been proposed by researchers such as Ball & Torous (1984), Garman & Klass (1980), and Parkinson (1980). For market participants, volatility is closely tied to risk assessment. In financial markets, downside risks are often considered more significant than upside risks. As a result, semi-variance has emerged as a risk measurement tool that focuses specifically on the downside volatility of an asset or investment portfolio. Unlike standard deviation, which accounts for all deviations from the mean, semi-variance only considers negative deviations, or losses, that fall below a defined threshold (Choobineh & Branting, 1986). This measure offers investors a more targeted and conservative view of risk, emphasizing the potential for losses rather than treating all fluctuations equally. Another way to measure volatility, indirectly but market-related, is using option prices. Implied volatility represents the market's forecast of the future volatility of an asset, as reflected in option pricing (Day & Lewis, 1992). Higher implied volatility signals a higher anticipated level of future price fluctuations (Beckers, 1981; Black & Scholes, 1973; Canina & Figlewski, 1993).

In addition to different modeling and estimation approaches, the behavior of volatility has also been studied since at least the 1960s. Mandelbrot (1963), for example, examined changes in cotton prices and found that strong fluctuations often follow one another and are then replaced by calm phases. This bundling is generally known as volatility clustering. Another phenomenon of volatility is the so-called leverage effect. This effect describes the behavior of increasing volatility with negative asset returns (Christie, 1982; Schwert, 1989).

1.3. Research scope

In this study, I have taken a detailed analysis of the effect of geopolitical shocks on market dynamics, as well as the possibilities for modeling and forecasting volatility in times of geopolitical shocks. To carry out the most comprehensive analysis possible, I examined many different aspects. The first part of my investigation is concerned with identifying abnormal returns and abnormal volatility in times of geopolitical shocks. I was conducting a detailed event analysis

for the period around the Russian attack on Ukraine and examined the reaction of the financial market to this shock in terms of returns and volatility. I analyzed the reactions of the various MSCI World Sector indices to this shock. This detailed analysis allows us to gain further insights compared to the analysis of sector-aggregated indices, for example, a better understanding of sector dynamics. Every industry - such as technology, healthcare, finance or energy - has unique characteristics and influencing factors (e.g., Elyasiani et al., 2020). Conducting sector-specific analysis provides valuable insight into the distinct behaviors and dynamics across industries. This approach allows investors to make more informed decisions by identifying sectors likely to outperform during particular phases of the economic cycle (He et al., 2020). Economic expansions and contractions affect sectors differently. Cyclical industries, such as consumer discretionary, typically exhibit greater responsiveness to macroeconomic fluctuations, whereas defensive or non-cyclical sectors like healthcare tend to remain relatively stable regardless of the economic environment (Long & Plosser, 1987). Investors can therefore increase their returns by rotating their investments between sectors (Sasseti & Tani, 2006). It is also easier to distinguish between the expected returns and the risk taken. By understanding the specifics of a sector, investors can better assess the risks associated with investing in that sector and thus better hedge against risk. Sector analysis can also point to emerging trends, such as the increasing importance of renewable energy in the utilities sector or the impact of e-commerce on the retail sector. In addition, a sector-specific view enables better benchmarking of individual stocks and thus a better classification of performance.

I used the sector indices of the MSCI World Index (Morgan Stanley Capital International Index) to include the most relevant companies in the respective sectors in my analysis. The MSCI World Index, introduced in 1986, serves as a prominent benchmark for tracking the performance of large and mid-cap equities across 23 developed markets (MSCI, 2024a). MSCI also provides sector-specific indices designed to offer systematic coverage of market segments, enabling the analysis of sectoral performance and characteristics at a global scale (MSCI, 2024b). Representing approximately 85% of the free float-adjusted market capitalization in each constituent country, the MSCI World Index offers a comprehensive gauge of equity performance within developed economies (Fama & French, 2012). Furthermore, the index is widely used by institutional investors for index funds and as a benchmark for global equity portfolios. Its components are weighted by market capitalization and the index is reviewed quarterly to include or exclude stocks based on market capitalization and other criteria. It includes companies from North America, Europe and the Asia-Pacific region, making it a diversified option for investors seeking exposure to global equity markets. There are advantages and disadvantages by using the MSCI World Sector Indices. The advantage is that they include the largest companies from 23 industrialized countries with a large number of different business models. In addition, the methodology of the index is very transparent, and it is frequently adjusted. It is widely accepted by investors and analysts and is often used as a benchmark. Disadvantages are the focus on industrialized countries, the weighting by market capitalization (small but important companies for an industry can be underrepresented) and there is a certain risk of currency fluctuations for companies that are not listed in U.S. dollars.

In addition, the weighting process results in a high proportion of U.S. companies, which is associated with an increased country risk. My analysis aims to shed light on the impact of the Russian attack on Ukraine on the world's largest companies and how uncertainty is caused by the major geopolitical shifts arose. One must understand volatility here as a measure of risk, or as an indicator of the strength of changes, since large price fluctuations, especially if they occur frequently, are a sign of investor dissent. They disagree on expectations and on the fair pricing of assets, which results in strong upward and downward movements on the markets. However, strong volatility is also caused by large price movements in one direction. Here, investors adjust their valuation based on new information on the markets.

I wanted to find out what excessive returns and volatility the Russian attack has caused on the biggest companies and most important business models worldwide and whether returns and volatility are persistent. Moreover, I wanted to analyze if there are statistically significant differences of the volatility between different sectors. By doing an event study, I therefore used the concept of abnormal returns (Campbell & Lo, 1996; Fama, 1970) and abnormal volatility (Beaver, 1968; Brown & Warner, 1985) and analyzed the behavior of eleven different MSCI World sector indices close around the Russian attack on February 24, 2022. My concrete research questions are as follows:

- **Research Question 1:** Are there significant abnormal returns on the February 24, 2022, in the sectors of the MSCI World index, respectively?
- **Research Question 2:** Are there significant cumulative abnormal returns in the MSCI World index sectors up to 25 days after the February 24, 2022?
- **Research question 3:** Can significant abnormal volatility (-5,+5 days) around the Russian attack on Ukraine at February 24, 2022 in the MSCI market sectors be observed?
- **Research question 4:** Can a persistence in the abnormal volatility after February 24, 2022, in the MSCI market sectors be observed?
- **Research question 5:** Are there significant differences of the abnormal volatility between the MSCI market sectors?

The identification of abnormal returns and volatility is a fundamentally ex-post analysis and important for a retrospective assessment of the impact and duration of geopolitical shocks. From a practical point of view, for example that of an active market participant or a risk manager, the forward-looking aspect is more relevant, i.e. whether the evolution of market dynamics can be predicted. So, the second part of my study covers the modeling and forecasting of financial time series around geopolitical shocks. Particularly in turbulent times, action is often needed in the form of increasing hedges or adjusting the portfolios of market participants to adapt to the new situation and manage the risk of adverse effects. This always involves the best possible assessment of volatility. With my study, I wanted to provide insights for research and practice into which forecasting methods are best suited for modeling volatility on the one hand and out-of-sample forecasting on the other in phases of geopolitical shocks. I focused on the periods in which I observed a shock to the financial markets and analyze how different forecasting methods can

predict periods of high volatility. To do this, I evaluated different approaches for prediction, univariate and multivariate, as well as parametric and non-parametric approaches. The core model of my analysis is the state-of-the-art Temporal Fusion Transformer (TFT) for modeling and predicting time series. For my univariate parametric approach, and as a benchmark, I examined the frequently applied GARCH, EGARCH, and GJR-GARCH models in terms of their ability to explain and predict volatility for the most important financial instruments. From a practitioner's point of view, there is always a trade-off between model complexity and performance. Therefore, I also provided statistically significant differences in forecasting performance between model classes in my study.

In this context, I wanted to answer the following research questions:

- **Research Question 6:** Which GARCH-type model provides the best in-sample fit regarding the volatility of the analyzed financial instruments during the recent geopolitical shocks?
- **Research Question 7:** Which GARCH-type model provides the best out-of-sample forecasts regarding the volatility of the analyzed financial instruments during the recent geopolitical shocks?
- **Research Question 8:** Can the Temporal Fusion Transformer achieve better out-of-sample performance than the best performing GARCH-type model in times of geopolitical shocks?
- **Research Question 9:** Can the implementation of a regime-switching feature in the TFT increase the forecasting performance in times of geopolitical shocks?
- **Research Question 10:** Can a significantly better forecasting performance be achieved by optimizing the hyperparameters of the TFT than the performance averaged by a hyperparameter grid in times of geopolitical shocks?
- **Research Question 11:** Can a multivariate forecasting approach of the Temporal Fusion Transformer model achieve better results than a univariate approach of the TFT in times of geopolitical shocks?

To answer these questions, I have used various modifications of the Temporal Fusion Transformer to predict the volatility of some of the most important financial instruments: the highly traded stock indices S&P 500, NASDAQ 100, Nikkei 225 and Hang Seng; Gold as the essential instrument for hedging against inflation and a long-term store of value; Brent Crude Oil as the most important source of energy; the 10 years Treasury Bond as the essential reference for fixed income, a benchmark for returns and the price of time; the exchange rate of Euro vs. US-Dollar as an indicator of the balance between the two most important currencies; and Bitcoin as an instrument with increasing market capitalization, high potential of growth and a hedge against inflation. For selected periods during the geopolitical shocks of COVID-19, the Russian attack on Ukraine and the Hamas attack on Israel, I examined how well the different methods could predict these phases of increased market volatility. Using a rolling window approach, I trained the model with data from the last 250 trading days and generate direct multi-step ahead forecasts for the next 10 trading days. Then I moved the window forward by one step. For the three periods of geopolitical shocks under consideration, I analyze around 60 steps of the window each.

1.4. Robustness of the results

The robustness of the results is critical at empirical studies, so I ensured it by the following methodical approach. First, I analyze a total of nine different financial assets from different markets: stock markets, commodity markets, bond markets, currency markets, and finally, the crypto market. This allows us to observe a wide range of different markets and their dynamics. By analyzing several different time series, one can reduce sampling bias, i.e., specific behaviors of a single time series, and increase the generalizability of the results (e.g., Baltagi, 2008, Bergmeir et al., 2018, Pesaran, 2015). Furthermore, this enables us to identify similarities or structural differences. It should be noted, however, that a strong correlation between the variables can reduce robustness (Stock & Watson, 2012).

Next, I examined three different periods of geopolitical shocks to the financial markets: COVID-19, the Russian attack on Ukraine and the Hamas attack on Israel. These shocks were very different in their causes and effects and therefore offer us a good opportunity to evaluate the reactions to the various financial instruments. The robustness of the results is also supported by the inclusion of several periods with geopolitical shocks. Models that are stable over several periods are less subject to risk from specific events in one period and the fact that the results are distorted by them. Furthermore, the different periods usually also include different phases of macroeconomic influences (e.g. business cycle). In addition, possible limitations of the prediction model and sensitivities can be identified (Giacomini & Rossi, 2010; Pesaran & Timmermann, 2007; Stock & Watson, 1996). One can achieve a significant contribution to robustness by generating out-of-sample forecasts. This provides us with a realistic assessment of the forecasting performance of the models under consideration, which is of course particularly important for practical applications. In addition, the model selection made based on the out-of-sample performance carries less risk of overfitting, since a model that is too strongly adapted to the training data or the estimation period usually shows a poorer performance (e.g., Kutner, 2005; Tashman, 2000; Welch & Goyal, 2008).

Furthermore, I used a rolling window approach to estimate and train the forecasting models, as well as for the subsequent volatility forecasts. The rolling window approach is a variant of cross-validation, specifically for time series models. In this approach, the model is trained on a fixed training window and then tested on the immediately following observations. The window is then shifted in time (“rolled”) and the process is repeated step by step until the end of the time series. This ensures that forecasts are based exclusively on historical data, that the temporal structure is considered and that a realistic assessment of the out-of-sample forecasting ability is made possible (Bergmeir & Benitez, 2012; Hyndman & Athanasopoulos, 2018; Tashman, 2000). In my work, I generated multi-step forecasts of volatility for each step for 1-10 trading days into the future. These multi-step forecasts allow to evaluate the forecasting performance of the considered models more validly. Furthermore, I evaluated the forecasting performance over several time horizons. This allows to produce a more valid selection of the optimal forecasting model. Furthermore, by evaluating the longer forecasting horizons, one can identify possible effects of overfitting the model (Marcellino et al., 2006; Pesaran et al., 2011, Tashman, 2000).

1.5. Contribution to research and outline

With my dissertation, I'm contributing to the steadily growing research literature on predicting volatility in financial markets. The development of the unipolar world order into a multipolar world order suggests that the addition of political actors that can trigger geopolitical shocks in the financial markets will increase, and thus the frequency of such events is more likely to increase. These shocks carry a high risk of portfolio losses and require intensive management. A tailored hedging strategy is based first on an accurate forecast of expected developments. Otherwise, the portfolio can be both over- and under-hedged, leading to an unsatisfactory result for the market participant. I started with an ex-post analysis of the effect of a geopolitical shock on the financial markets, using the example of the MSCI World sector indices with an analysis of abnormal returns and abnormal volatility. Here I showed how strongly and how differently shocks can affect sectors and thus the most important business models. I also examine how long shocks can persist in the individual sectors. I also observed here that the Russian attack on Ukraine has exerted a shock on the financial markets, from which individual sectors have benefited. After considering how a geopolitical shock works, I turned to the task of predicting how markets will behave in turbulent times.

To do this, I examined how well the Temporal Fusion Transformer (TFT), a state-of-the-art transformer-based neural network designed specifically for time series analysis, is able to predict the volatility of key financial stocks during geopolitical shocks. I examined various modifications: univariate forecasting, multivariate forecasting and forecasting with a regime-switch feature, and compare the performance for significant differences. I used the GARCH models, which have been used in practice for years, as a benchmark for the results. In particular, there is little literature on the forecasting performance of the Temporal Fusion Transformer for describing and predicting volatility in the face of geopolitical shocks, and so I aimed to close a research gap here. By considering the most important financial instruments for the last three most recent geopolitical events, I provided systematic and comprehensive insights into the performance of the TFT in predicting volatility during phases of geopolitical shocks. With my work, I gained insights into the advantages of using the Temporal Fusion Transformer to predict the Temporal Fusion Transformer over the GARCH model approach, by providing a multivariate forecasting approach. In doing so, I provided practitioners with results that can be used to improve portfolio management based on the implementation of the TFT. Furthermore, which modifications appear useful for the TFT. For researchers, I provided a research status on which further improvement steps can be taken. The dissertation is organized as follows: Section 2 contains the literature review, which provides an overview of the relevant research on my topic. Section 3 describes the methods and procedures I used and discusses the data sets used. Section 4 discusses the results of my research. First, the identification of abnormal returns and abnormal volatility during the Russian attack for the MSCI World Sector indices. Second, the results on predicting volatility using the Temporal Fusion Transformer and benchmark GARCH class models. In Chapter 5, I draw a conclusion and present my scientific contribution to current research in Chapter 6. Finally, I summarize the whole work in chapter 7.

2. LITERATURE OVERVIEW

2.1. Geopolitical Shocks and their impact on financial markets

Geopolitical shocks are abrupt and significant events that alter the strategic landscape, disrupt established power relations, or trigger systemic uncertainty in international relations (Brütsch & Papa, 2013). Bremmer & Roubini (2011) defined them as disruptive event rooted in political instability, war, or regime change, with potential to impact global markets, security structures, or alliances. For OECD, geopolitical shocks refer to sudden, unexpected developments in international affairs, such as armed conflicts, sanctions, or diplomatic breakdowns, that affect global supply chains, energy markets, and international cooperation (OECD, 2022). These definitions have in common, that these kinds of shocks are suddenly and exogenous, with a significant multidimensional impact on the global order, economies and security architecture that can result in disturbances of political systems and society.

Famous publications regarding geopolitical shocks are the following. Brütsch & Papa (2013) examined geopolitical shocks as risks to the stability of the international system and evaluate the capacity of multilateral institutions to manage them. They identify patterns of institutional resilience and argue that global governance mechanisms often fall short in responding to sudden geopolitical crises. The study advocates for strengthening preventive and adaptive capacities within international organizations. Bremmer & Roubini (2011) analyzed the ongoing transition to a multipolar order and described a global order without a clear hegemon, in which cooperation is limited and geopolitical shocks become more frequent and less manageable. They argue that neither the United States, China, nor Europe is willing or able to provide global public goods. This power vacuum increases systemic volatility and the likelihood of economic and security-related disruptions. Mead (2024) argues that the era of liberal internationalism is giving way to a resurgence of great power competition, with revisionist states like Russia and China challenging the U.S.-led order. This return of geopolitics raises the likelihood of shocks as global institutions lose their integrative capacity. Dalby (2013) connects classical geopolitical thinking with global environmental crises, arguing that geopolitical shocks increasingly stem from ecological disruptions and planetary boundaries. He advocates for a fundamental rethinking of geopolitics in the context of the Anthropocene. The work is innovative and interdisciplinary, expanding the notion of shocks to include climate-related and ecological dimensions. Schweller (2018) analyzed U.S.–China relations through a neoclassical realist lens, emphasizing how ideological and identity-based divergence can lead to geopolitical instability. He contends that national narratives may exacerbate systemic tensions and trigger shocks, even in the absence of direct material conflicts. The study adds conceptual depth by highlighting the ideational dimension of geopolitical risk.

Consequences of recent geopolitical shocks, particularly the war in Ukraine and global supply chain disruptions, was investigated by OECD (2022). They analyzed both short-term effects (e.g., energy prices, inflation) and long-term structural shifts in trade and investment flows. The report calls for enhanced national and supranational resilience strategies. The World Bank

(2022) analyzed how geopolitical shocks, such as conflict, sanctions, and resource volatility, affect long-term growth, development, and poverty. Special emphasis is placed on emerging economies, which are disproportionately vulnerable to global disruptions. Forecast models suggest that adverse effects are likely to deepen in the absence of coordinated international responses. Blackwill & Harris (2016) introduced the concept of geoeconomics as the strategic use of economic instruments to achieve foreign policy objectives. They argue that geopolitical shocks are increasingly triggered or intensified through economic means, such as sanctions, trade wars, or investment policies. The book provides a comprehensive analytical framework for understanding contemporary power politics beyond traditional military tools.

To quantify these shocks, Caldara & Iacoviello (2022) developed the news-based Geopolitical Risk (GPR) index to measure the risk, which spikes around major conflicts (e.g. World Wars, the Cuban Missile Crisis, 9/11). Heightened geopolitical risk is associated with deteriorating economic outlooks. Caldara & Iacoviello (2022) find that surges in their GPR index precede lower investment and employment and increase downside tail risks to growth.

In financial markets, geopolitical events inject uncertainty that can alter asset valuations and volatility. In recent history it was observed that geopolitical shocks, including wars, terrorist attacks, and major political crises roil financial markets worldwide. Multiple studies underscore that geopolitical tensions engender adverse economic outcomes such as elevated commodity prices, inflationary pressures, and diminished social welfare (Aizenman et al., 2024; Bouri et al., 2024; Mahjoubi & Henchiri, 2024; Wang & Zhou, 2024). The IMF (2023) investigated how rising geopolitical tensions and strategic decoupling—particularly between major powers could fragment global trade and financial systems. Simulation models show significant long-term output losses, especially for low-income countries, if global cooperation continues to deteriorate. Campos et al. (2024) used macroeconomic models to estimate the costs of geopolitical divisions in trade, technology, and capital flows and found, that persistent fragmentation could lead to a substantial decline in global GDP and welfare, especially if technological cooperation collapses. From a perspective of international specialization, trade and connectedness, Baldwin & Freeman (2022) argued that geopolitical shocks have transformed global supply chain strategies from “just-in-time” to “just-in-case.” They document a shift toward regionalization and redundancy in production, which increases costs but enhances economic security.

A growing body of research highlights that geopolitical shocks significantly influence financial markets through complex and multifaceted channels. Geopolitical events such as regional conflicts, wars, trade wars, and political crises have been documented to induce heightened uncertainty that propagates rapidly into asset prices and market volatility. Rigobon & Sack (2005) analyze the Iraq War and demonstrate that war risk can immediately affect U.S. stock prices, interest rates, and exchange rates. Pástor & Veronesi (2012) developed a general equilibrium model indicating that political uncertainty commands a risk premium, an effect that is magnified during periods of economic weakness and market instability. Bekaert et al. (2013) distinguished between risk and uncertainty shocks and illustrated how geopolitical uncertainty contributes to financial volatility and rising risk premiums. Bouras et al. (2019) found that the GPR significantly

improves the predictability of stock returns in G7 economies, especially during periods of heightened uncertainty. Ha et al. (2021) demonstrated that uncertainty shocks stemming from events on the Korean Peninsula triggered substantial market reactions. Agyei et al. (2023) explored the relationship between geopolitical risk and stock market behavior, revealing a strong association between geopolitical developments and heightened market volatility. Similarly, Kononova & Abuzov (2023) examined how trade disputes, exemplified by the US-China trade war, result in significant declines in global share prices, emphasizing the transnational spillover effects of such shocks. Gupta et al. (2023) analyzed the role of geopolitical uncertainty in relation to the realized volatility of WTI oil-price returns. Wang et al. (2022) showed that geopolitical risk causes volatility spillovers between oil and equity markets, emphasizing the role of commodities in transmitting shocks.

The transmission mechanisms of geopolitical shocks into financial variables have been widely analyzed. Ali et al. (2023) analyzed the influence of stock markets and sectoral performance in the context of geopolitical threats over the period from 1987 to 2021. Their findings indicate that U.S. equities generally benefited from geopolitical tensions, with the financial and information technology sectors exhibiting particularly strong positive responses. Naifar & Aljarba (2023) argued that geopolitical risk elevates sovereign credit risk and contracts capital flows, suggesting that investors are highly sensitive to rising uncertainty. Complementing this perspective, Hoque et al. (2019) found that while the direct effects of geopolitical risks on local stock markets may be muted, significant indirect effects occur through channels such as oil price shocks and global economic policy uncertainty. Such findings reinforce the idea that geopolitical shocks not only affect asset prices directly through market sentiment but also indirectly by altering macroeconomic fundamentals and investor risk assessments. Będowska-Sójka et al. (2022) examined how various asset classes respond to geopolitical risk, identifying green bonds, precious metals such as gold and silver, the Swiss franc, and real estate as among the most resilient in the face of geopolitical uncertainty. Research by Balcılar et al. (2018) reveals that markets, especially in emerging economies like the BRICS, exhibit pronounced volatility in response to geopolitical tensions regardless of the nature of the shock. In addition, Apergis et al. (2018) note that investors adjust their risk management strategies by seeking industries and assets that promise relative stability as a hedge against geopolitical uncertainty. Moreover, studies focusing on investment behavior, such as the one by Tanyeri et al. (2021) in the context of the Arab Spring, underscore the broader market's sensitivity to geopolitical shocks, with abnormal stock returns and credit spread adjustments observed during periods of heightened political unrest. A broad theme in the literature also is, that geopolitical shocks prompt a "flight to safety": investors sell risky assets like equities and seek safe havens in assets such as government bonds, gold, or safe-haven currencies (e.g., Jäggi et al., 2019; Rigobon & Sack, 2005). This documents that during major conflicts, such as the Russia-Ukraine war, investors increasingly gravitate towards safe-haven investments, which underscores the role of geopolitical shocks in reshaping investment portfolios (Rania, 2023). In a related vein, Rasheed et al. (2021) find that assets such as gold serve as effective hedges against political uncertainty, especially in economies prone to political

instability. This safe-haven behavior reflects a strategic shift by market participants to mitigate the risks associated with unpredictable geopolitical events.

Collectively, these studies underscore that financial markets are highly sensitive to geopolitical developments, with impacts that are both immediate and far-reaching across asset classes and regions. Now I want to provide a closer look at the geopolitical shocks of COVID-19 and the Russian attack on Ukraine. These two events are still a growing research field. Goodell (2020), laid the groundwork by identifying that COVID-19 represents a novel shock with multifaceted economic consequences. In particular, this study highlighted that while previous pandemics and infectious disease outbreaks provide limited analogues, the scale and nature of COVID-19's impact necessitate new research avenues to more comprehensively understand its diverse ramifications on financial institutions and market dynamics (Goodell, 2020). A substantial body of literature has focused on the alteration of investor sentiment and the subsequent volatility observed in equity, bond, and commodity markets. For example, Ali et al. (2020) documented that the outbreak resulted in markedly negative returns and heightened volatility across most asset classes, with the notable exception of US Treasury bonds. Such findings are consistent with earlier research linking uncertainty with prolonged market disturbances. In a similar vein, Liu et al. (2020) employed event study methodologies across 21 major stock market indices and found that the reaction to the pandemic's onset was both rapid and profound, indicating that COVID-19 induced an immediate revaluation of risk and market fundamentals across different national contexts. The research further underscores that the disruption was not homogeneous across sectors and asset types. Baek et al. (2020) provided an industry-level analysis within the US, showing that certain sectors experienced heightened volatility compared to others, underscoring the sector-specific sensitivity to exogenous shocks such as a global pandemic.

Complementing these findings, Albulescu (2021) examines volatility at a global level, suggesting that government interventions and market responses to pandemic uncertainty play a critical role in shaping the volatility landscape of financial markets. These studies collectively emphasize that COVID-19 has not only increased market risk but also altered the structure of risk connectedness among assets (Cheng et al., 2022). Another important stream of literature has investigated how sustainability factors, such as environmental, social, and governance (ESG) criteria, can mitigate the negative effects of crisis periods. Broadstock et al. (2021) demonstrated that firms with superior ESG performance tend to exhibit lower risk during turbulent times like those induced by COVID-19, implying that non-financial performance indicators may serve as risk buffers in times of systemic stress. Such insights provide critical evidence that shifting investor focus on sustainability can modify the typical risk-return trade-offs observed during market distress. Moreover, market efficiency and contagion effects have been key themes in the emerging literature. Khan et al. (2023) applied GARCH-model methodologies to show that COVID-19 drastically increased market volatility, while Phan & Narayan (2020) relay that governmental actions taken to curb the pandemic, such as lockdowns and fiscal stimuli have had a measurable yet heterogeneous effect on stock prices across nations. Similarly, Cheng et al. (2022) investigated volatility connectedness networks across global markets, concluding that the

pandemic induced stronger spillover effects in an era characterized by intensified cross-market linkages. Aslam et al. (2020) extended this discussion to the foreign exchange markets, finding that the efficiency of forex markets was significantly impaired during the initial crisis period due to abrupt changes in market sentiment.

Lastly, innovative methods have emerged in the literature to capture the dynamics of the COVID-19 shock. Gao et al. (2021) adopt a wavelet-based quantile-on-quantile approach to compare the impacts on stock market volatility in the United States and China, revealing that COVID-19 exerted a more pronounced effect in the U.S. market, thereby highlighting regional differences that may be associated with divergent monetary policies and public health responses. Complementing these quantitative analyses, Costola et al. (2021) utilized Google search volume data as a proxy for investor sentiment, demonstrating that digital information flows played a significant role in the rapid market adjustments observed during the pandemic's outbreak.

The Russian invasion of Ukraine in 2022 constituted a profound disruption to the post-Cold War security order and triggered a major geopolitical shock, with immediate and far-reaching effects on global financial markets, including increased volatility, risk repricing, and commodity price surges (Caldara & Iacoviello, 2022; ECB, 2022; IMF, 2022; Mariotti, 2022). One body of research highlights the immediate market disruption resulting from the conflict's geopolitical shock. The invasion led to an immediate surge in commodity prices, particularly in oil and natural gas, as Russia and Ukraine collectively contribute around 12% of global oil production and 17% of natural gas supply (Martins et al., 2024; Mbah & Wasum, 2022). This surge in energy prices created a ripple effect across various asset classes, increasing volatility in equity and derivative markets as market participants reassessed risk premia amid growing uncertainty (Alam et al., 2022; Theiri et al., 2022). In addition, the implementation of fiscal and monetary interventions by leading economies—particularly sanctions targeting Russian financial operations—intensified global uncertainty and contributed to heightened market volatility (Hao, 2023). Izzeldin et al. (2023) assessed how market participants responded to the Russian invasion, focusing on both the immediacy and persistence of market reactions. Their comparative analysis, which included the 2008 global financial crisis and the onset of the COVID-19 pandemic in 2020, revealed that market responses to the Russian attack were significantly more rapid. However, the volatility period following the invasion was notably shorter than that observed during the other two crises. Among the most heavily impacted commodities were wheat and nickel. Stock market responses to the invasion have been particularly well documented. Several studies report that the day of the invasion marked a strong negative reaction among global equity markets, especially within Russia itself, where abnormal trading volumes and heightened selling pressures were detected (Alam et al., 2022; Davydov, 2022). The divergence between companies that chose to maintain operations in Russia and those that divested underscores how political risk and uncertainty can lead to heterogeneous market responses (Davydov, 2022; Kiesel et al., 2022). Ahmed et al. (2023) investigated the impact of the Russian invasion of Ukraine on the STOXX Europe 600 Index and its constituent sectors. Their findings indicate that the conflict exerted a negative influence on seven out of eleven sectors, specifically materials, consumer staples, financials, healthcare,

industrials, telecommunications, and utilities. Among these, the consumer staples sector recorded the most pronounced negative average abnormal return on the event day (February 24, 2022), whereas the energy sector exhibited a marginally positive, albeit statistically insignificant, average abnormal return.

In addition to market reactions, the imposition of sanctions led to a substantial exodus of foreign investors from the Russian financial markets, thereby reducing market liquidity and exacerbating volatility in affected equity sectors (Hoffmann & Neuenkirch, 2015; Nimani & Spahija, 2023). Complementary evidence is provided by Boubaker et al. (2022), who assessed the war's impact on global equity markets across 23 developed and 24 emerging economies, as defined by the MSCI classification. Their analysis revealed that nearly all markets experienced negative abnormal returns on the event day, with emerging markets displaying a more pronounced sensitivity. Furthermore, their study documented persistent negative cumulative returns in the aftermath of the invasion, with the exception of markets in Asia and the Middle East. In a broader cross-country study, Boungou & Yatié (2022) analyzed stock market performance in 94 countries between January 22, 2022 and March 24, 2022. Their findings offer robust empirical support for the hypothesis that the Russian invasion had a statistically significant negative impact on global equity returns. Similarly, Pandey et al. (2023) examined the immediate consequences of the invasion on major stock market indices within the European Union, further reinforcing the consensus that geopolitical conflict precipitated widespread market disruptions across regional and global financial systems. Their analysis identified a negative impact on index performance on the day of the event.

However, in the post-event period, they observed positive cumulative abnormal returns in select countries, notably Poland, Denmark, and Portugal. The financial repercussions extend into the realm of foreign exchange markets as well. Empirical evidence shows that the Russian-Ukrainian conflict led to significant fluctuations in euro and other European currencies, highlighting the fragility of the European financial system to such external shocks (Aliu et al., 2023). These sudden movements have not only impacted international trade dynamics but have also complicated monetary policy implementation for central banks dealing with dual shocks from the pandemic's residual effects and the wartime uncertainty (Aliu et al., 2023; Hao, 2023). Another critical area of study concerns the effects on cryptocurrency markets. Research indicates that market liquidity for leading digital currencies such as Bitcoin and Ethereum was notably affected during the conflict (Theiri et al., 2022). Short-term dynamics in cryptocurrency returns were observed, with studies identifying an increased correlation between cryptocurrency yields and commodity price movements, particularly in crude oil futures (e.g., Yousaf & Ali, 2023). Additionally, the cryptocurrency market has emerged as a conduit for risk transmission, showing robust connectedness with conventional financial assets and thereby serving both as a refuge and a risk amplifier during times of geopolitical turmoil (Ullah, 2024; Volosovych et al., 2024). Methodologically, various approaches have been employed to capture these dynamics. Event studies, vector autoregressive models, and network spillover estimations have all been used to determine the degree and persistence of financial market responses to the conflict (Alam et al.,

2022; Nittayakamolphon et al., 2024). These methods underscore that while some market disruptions are transient, others have longer-lasting implications, particularly in the areas of investor sentiment and the strategic allocation of assets globally (Davydov, 2022; Kiesel & Kolaric, 2023).

In sum, the existing body of literature consistently highlights that geopolitical shocks induce significant, albeit temporary, disturbances across a wide range of financial markets. Empirical evidence robustly associates such events with elevated stock market volatility, destabilization of foreign exchange rates, sharp fluctuations in energy prices, and shifts in liquidity conditions—affecting both traditional financial instruments and newer asset classes, including cryptocurrencies (Robus et al., 2024).

2.2. Modeling and forecasting volatility in times of geopolitical shocks

The modeling and forecasting of financial market volatility represent a foundational area within financial econometrics, owing to their pivotal role in risk assessment, asset valuation, portfolio optimization, and the pricing of derivative instruments (Andersen et al., 2005; Bollerslev, 1986; Engle, 1982). Volatility forecasts have widespread applications across financial institutions, including banks, hedge funds, and regulatory bodies, where they inform decisions related to capital adequacy and portfolio risk management. Accurate volatility estimates serve as essential inputs for market risk models and are central to conducting stress-testing exercises (Christoffersen et al., 2012). In the domain of derivatives, volatility modeling underpins pricing frameworks, with both implied and forecasted volatility playing key roles in dynamic hedging strategies and the calibration of models incorporating stochastic volatility (Bakshi et al., 1997).

Furthermore, in portfolio and asset allocation contexts, volatility forecasts are integral to mean-variance optimization and other risk-return trade-off models (Engle & Colacito, 2006). Further, central banks and regulators monitor market volatility as an indicator of financial instability or systemic risk. So, forecasting volatility helps detecting early signs of financial stress and inform macroprudential policy tools (Diebold & Yilmaz, 2014). Estimation and forecasting Volatility during crisis are most important, because the risk for adverse results is highly increased. So, the heightened uncertainty requires more accurate risk assessment: turbulent periods are characterized by heightened and persistent volatility, necessitating dynamic models that can capture rapid shifts in risk because standard risk measures such as VaR can significantly underestimate downside risk when volatility spikes (Christoffersen, 2011; Engle, 2001). Further, investors typically rebalance portfolios, shift into safe-haven assets, or employ volatility-targeting strategies. Accurate forecasts help avoid overreaction or mispricing, which can further destabilize markets (Anderson et al., 2003).

In crisis, volatility spikes can indicate systemic fragility. Forecasting volatility allows central banks and financial regulators to conduct stress tests and implement macroprudential policies aimed at stabilizing markets and mitigating the risk of cascading failures (Diebold & Yilmaz, 2014; Tobias & Brunnermeier, 2016). Obviously, there is a particular need to estimate and

predict volatility and their special behavior in times of crisis. The models suitable for this purpose will be presented below.

Before I'm discussing how volatility can be modeled in times of geopolitical shocks, I want to give a short summary of the so-called stylized facts, one can observe in time series of financial markets. Empirical research in financial econometrics has consistently documented a set of robust statistical regularities across different asset classes, markets, and time periods. These regularities, known as stylized facts, are not strict laws but rather empirically observed phenomena that hold with remarkable consistency. Their identification has had a profound impact on the development of theoretical and empirical models in asset pricing, volatility modeling, and market microstructure (Cont, 2001). One of the most fundamental stylized facts is the presence of heavy tails in asset return distributions. Empirical returns exhibit excess kurtosis, indicating a significantly higher probability of extreme events than would be implied by a normal distribution (Mandelbrot, 1963). This observation has motivated the widespread use of fat-tailed distributions such as the Student's *t*-distribution or generalized hyperbolic distributions in financial modeling.

Another prominent characteristic is volatility clustering, a phenomenon in which substantial price fluctuations are often succeeded by further large movements, regardless of direction, while periods of minimal change are typically followed by similarly small variations. This phenomenon results in significant autocorrelation in the second moments of return series, even when first moments (returns themselves) are uncorrelated. The presence of volatility clustering provided the impetus for the development of autoregressive conditional heteroskedasticity (ARCH) models (Engle, 1982) and their generalizations such as GARCH (Bollerslev, 1986), which explicitly model time-varying conditional volatility. A closely related empirical regularity is the absence of linear autocorrelation in raw asset returns, especially at daily or higher frequencies. This is consistent with the weak-form Efficient Market Hypothesis (Fama, 1970), which posits that prices fully reflect past information. While raw returns are often indistinguishable from white noise, measures of volatility such as squared or absolute returns exhibit strong and persistent autocorrelation, implying long memory in volatility dynamics (Ding et al., 1993).

The leverage effect constitutes another asymmetric stylized fact in financial markets, where negative returns tend to be followed by higher future volatility than positive returns of equal magnitude. This asymmetry is typically attributed to the capital structure of firms: a decline in equity value increases financial leverage, thereby raising equity risk (Black, 1976). Econometric models such as EGARCH (Nelson, 1991) and GJR-GARCH (Glosten et al., 1993) have been developed to accommodate this asymmetry in volatility responses. Additionally, a strong positive correlation between trading volume and volatility is frequently observed. This relationship has been interpreted within the framework of heterogeneous agent models and information arrival hypotheses (Karpoff, 1987), which posit that both volume and volatility respond to the flow of new information in the market.

Another notable stylized fact is aggregational Gaussianity, which refers to the observation that as one increases the time horizon over which returns are measured—from minutes to hours,

or days to weeks—the distribution of returns tends to converge toward the normal distribution. This is broadly consistent with the Central Limit Theorem under the assumption of weak dependence and finite variance, though the convergence is often slower than predicted due to the presence of long-memory effects and volatility persistence (Dacorogna et al., 2001). Finally, at very high frequencies (e.g., tick-by-tick or intraday data), market microstructure effects become prominent. Features such as bid-ask bounce, discrete pricing, and asynchronous trading introduce serial correlation and measurement error, necessitating the use of specialized econometric tools to extract meaningful signals from noisy high-frequency data (Andersen et al., 2001; Hasbrouck, 2007). Taken together, these stylized facts provide a foundational empirical reference for constructing and evaluating models of financial markets. Any realistic econometric model must either replicate or account for these phenomena in order to be considered empirically adequate.

2.2.1. Modeling and forecasting volatility in times of geopolitical shocks using GARCH-class models

When modeling and forecasting the movements of financial markets, stochastic processes (Doob, 1942; Pavliotis, 2014) are frequently used, e.g., ARIMA models (Box & Jenkins, 1970; Box et al., 2015; De Gooijer & Hyndman, 2006). These models lead to good forecasts of the time-varying expected value. A common assumption of these models was, that the variance of the stochastic error term is constant over time, which is referred to as homoscedasticity. The ARIMA models are therefore only suitable to a limited extent for modeling and predicting strong fluctuations on the markets. However, two phenomena can often be observed with financial market data. The first is that volatility is not constant over time, which is known as heteroscedasticity (Engle, 1982). Secondly, the high level of volatility occurs in clusters, i.e. there is a serial correlation (Cont, 2001; Engle, 1982; Taylor, 2007). To describe the time-varying volatility, Engle (1982) designed the model of autoregressive conditional heteroscedasticity (ARCH). Later, Bollerslev (1986) introduced the generalized autoregressive conditional heteroscedasticity model (GARCH) to better describe volatility persistence, which often occurs in financial time series. The most common extensions are the EGARCH (Exponential GARCH), IGARCH (Integrated GARCH) and TGARCH (Threshold GARCH) (Engle & Bollerslev, 1986; Glosten et al., 1993; Nelson, 1991). These models are now frequently used to predict volatility on the financial markets. As with the ARIMA models, a test-based specification of the model is also important for the GARCH models to achieve adequate forecast results (Lundbergh & Teräsvirta, 2002). They are also combined with the ARMA models to model and forecast the conditional mean and the conditional variance simultaneously. An often-investigated question is which type of model predicts volatility better. Brailsford & Faff (1996) found evidence that more complex prediction models, like ARCH class models) outperform the simpler volatility models, like the class of moving average models. Savickas (2003) analyzed statistic tests in the presence of event-induced variance and compares them according to their performance. He found that the GARCH-based approach performed best due to the highest test power. In contrast to other studies, my analysis of volatility dynamics in connection with a geopolitical shock focuses on a sector-specific view.

Volatility in financial markets is known to exhibit shifts corresponding to different market conditions – for example, tranquil periods versus crisis periods. So, regime switching volatility models are pivotal in modeling the dynamics of financial assets, as they allow for the capture of abrupt changes in market conditions. These models are primarily based on the premise that financial markets do not operate in a static environment but rather switch between different states or regimes that reflect varying levels of volatility and returns. The introduction of regime-switching models by Hamilton laid the groundwork for understanding these market phenomena (Edwards, 2005; Kim, 2014). This approach builds on early work by Hamilton (1988) and Hamilton & Susmel (1994) for regime-switching in means and by Gray (1996) who first embedded regime changes in GARCH. Traditional GARCH models assume a single set of volatility dynamics for an entire sample, which can lead to mis-estimation (e.g. overstated persistence) if structural breaks occur (Ardia et al., 2019). Regime-switching GARCH models generalize the standard GARCH framework by permitting the volatility dynamics to differ across a finite number of states. In a typical two-state MSGARCH (Markov-Switch GARCH), for instance, one state might correspond to a low-volatility regime (with lower GARCH intercept and quicker mean reversion) and another to a high-volatility regime (with higher shock impact and persistence). The switching is governed by transition probabilities, often assumed constant (Markovian) or sometimes time-varying. This allows the model to capture phenomena like sudden jumps in volatility or prolonged high-volatility episodes that a single-regime GARCH would struggle with (Ardia et al., 2019).

Over the last decade, regime-switching GARCH models have become popular for capturing nonlinear volatility patterns and they have been extended to various GARCH variants (e.g. EGARCH, GJR-GARCH) and multi-regime setups. One of the significant contributions of regime-switching models is their ability to explain the behaviors of asset returns more accurately than traditional models. Ang & Timmermann (2012) and Guidolin & Timmermann (2008) demonstrated that returns on stocks can be characterized by different regimes, such as bear markets with high volatility and low returns, and bull markets with high volatility and high returns. They argue that these regimes account for fat tails and skewness in returns, which are often observed in empirical data. By modeling these shifts, the regime-switching approach provides a robust framework for portfolio optimization and risk management (Oliveira & Pereira, 2018; Tu, 2010). Moreover, empirical studies have revealed that regime-switching models can effectively capture the long-term behaviors of financial data. For instance, Delong & Pelsser (2015) articulate that these models adequately reflect the market's transition between low and high volatility states, which align with observed economic realities. The capacity of these models to accommodate various macroeconomic variables, as discussed by Chang (2017), further supports their relevance in contemporary asset management strategies. This adaptability allows investment strategies to be tailored according to expected changes in market conditions, significantly impacting overall portfolio performance (Chang, 2017). As financial markets increasingly exhibit volatility due to external shocks—such as during the COVID-19 pandemic—researchers have turned to regime-switching frameworks to better understand and predict market dynamics. Kanamura (2022) exemplifies this, where a model was proposed to examine how COVID-19 influenced volatility

across various assets, showcasing the adaptability and relevance of regime-switching models in current financial contexts. The resilience of these models allows practitioners not only to anticipate market shifts but also to implement more effective hedging strategies during turbulent times (Lee, 2010). Shen et al. (2020) have examined the optimal investment strategies based on regime-switching pricing models, highlighting the increasing interest among financial scholars to utilize these models for complex asset management challenges.

2.2.2. Modeling and forecasting market dynamics using the Temporal Fusion Transformer

Accurate forecasting is fundamental to effective decision-making across a wide range of fields, including finance, healthcare, energy, and climate science. Time series forecasting, in particular, plays a central role in these disciplines by predicting future values based on historical observations and, in some cases, incorporating additional explanatory variables. However, the forecasting of temporal data presents substantial challenges due to its complex structure, which often involves trends, seasonality, and nonlinear dependencies. Traditional statistical models such as the autoregressive integrated moving average (ARIMA) model (Box & Jenkins, 1970; Box et al., 2015) and exponential smoothing methods (Holt, 1957; Winters, 1960) have long been employed due to their interpretability and computational simplicity.

While effective for linear and univariate data, these models often fall short in capturing intricate patterns within high-dimensional or nonlinear time series. In response to these limitations, artificial neural networks (ANNs) have gained widespread attention for their capacity to model complex, nonlinear relationships. Drawing inspiration from the structure and function of the human brain, ANNs consist of interconnected units—artificial neurons—that process and transmit information collectively. The conceptual foundation was established by McCulloch & Pitts (1943), who introduced the first formal model of a neuron. This was further developed by Rosenblatt (1958), who proposed the perceptron, a simple neural architecture consisting of a single neuron with modifiable weights and a threshold function. However, the field experienced a temporary decline following the critical analysis by Minsky & Papert (1969), who highlighted the perceptron's limitations in solving non-linearly separable problems. Interest in neural networks was revitalized in the early 1980s when Hopfield (1982) demonstrated their potential to model associative memory. The architecture of modern ANNs typically involves multiple layers: an input layer through which data enter the model, one or more hidden layers where processing and transformation occur, and an output layer that produces the final prediction. A major advancement was the introduction of the backpropagation algorithm by Rumelhart et al. (1986), which enabled the efficient training of multilayer networks and marked the beginning of deep learning. This innovation was soon applied in practice by LeCun et al. (1989), who employed convolutional neural networks (CNNs) for digit recognition in postal codes, demonstrating their utility in pattern recognition tasks. Since then, the field has expanded rapidly, both in theoretical development and practical applications. Comprehensive overviews of neural network evolution and applications can be found in Goodfellow et al. (2016) and Schmidhuber (2015).

Neural networks now encompass a broad spectrum of architectures. Shallow neural networks, as discussed by Aggarwal (2018), represent simpler forms with limited layers. Deep neural networks, or multilayer perceptrons, have been explored extensively in studies such as Schmidhuber (2015) and Yosinski et al. (2014), highlighting their capacity to learn hierarchical representations. CNNs, advanced by LeCun & Bengio (1995), LeCun et al. (1998), and Simard et al. (2003), have proven highly effective in image and spatial data processing, with further refinements presented by Gu et al. (2018). Recurrent neural networks (RNNs), introduced by Elman (1990) and extended by Hochreiter et al. (2001), were specifically designed to handle sequential data by retaining temporal dependencies. Long short-term memory networks (LSTMs), proposed by Hochreiter and Schmidhuber (1997) and subsequently improved by Gers et al. (2000), Mikolov et al. (2010), and Yu et al. (2019), address the limitations of traditional RNNs in capturing long-term dependencies. Attention-based neural networks have recently emerged as efficient tools for sequence modeling and prediction. With key contributions from Vaswani et al. (2017) they demonstrated superiority in capturing contextual information across time steps. Generative adversarial networks (GANs), first introduced by Goodfellow et al. (2014) and later refined by Creswell et al. (2018), Goodfellow et al. (2020) and Karras et al. (2018) represent another transformative innovation, enabling the generation of synthetic data and enhancing forecasting through data augmentation and simulation techniques.

Altogether, these advancements underscore the versatility and growing significance of neural networks in modeling complex time series, and their continuing evolution promises to further improve forecasting accuracy and applicability across domains. Neural networks are widely used today for estimating volatility, because they can capture intricate temporal dynamics and nonlinear relationships. Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) networks, and Gated Recurrent Units (GRU) have shown promise in modeling sequential data (Schmidhuber, 2015). They can flexibly approximate complex nonlinear relationships and potentially mimic regime-switching behavior (Zolfaghari & Hoseinzade, 2020). Neural networks offer a very different approach to modeling volatility, one that can be seen as complementary to or even a substitute for explicit regime-switching parametric model approach. A sufficiently complex neural network can, in theory, approximate the kind of piecewise dynamics that regime-switching models capture – for instance, by using nonlinear activation functions to mimic different volatility response regimes (Zolfaghari & Hoseinzade, 2020).

In practice, two strands of research have emerged: Hybrid models that integrate neural networks with regime-switching GARCH and pure machine-learning models that implicitly capture regime dynamics. For instance, the integration of neural network structures with autoregressive conditional heteroskedasticity (ARCH) frameworks, as discussed by Ramos-Pérez et al. (2021) in their study, shows substantial promise in predicting volatility in complex financial environments like the S&P market. The synergy between these approaches leverages the strengths of both statistical modeling and neural networks, providing robust predictions that outperform conventional models. Furthermore, Bildirici & Ersin (2015) explore the efficacy of combining radial basis function (RBF) and multilayer perceptron (MLP) neural networks with GARCH

models specifically for volatility forecasting in oil markets. Their findings suggest that these hybrid models significantly enhance forecasting accuracy by effectively capturing the nonlinearities inherent in financial time series.

Effective hybrid modeling is not limited to GARCH frameworks; it also extends to stochastic volatility models, as demonstrated by Cao et al. (2018) who incorporated regime switching in a hybrid model for pricing variance swaps. Their work emphasizes the significant impact of incorporating regime switching dynamics on the pricing of financial derivatives. Further enriching this discourse, Yang et al. (2023) illustrate the application of hybrid long short-term memory (LSTM) models to volatility forecasting during the COVID-19 period, reinforcing the notion that combining traditional econometric techniques with contemporary machine learning approaches leads to improved prediction outcomes. This adaptability of hybrid models allows them to capture shifts in market sentiment and structural breaks more effectively than either traditional or machine learning models alone.

But there is increasing research regarding pure machine learning approaches to model regime switching in volatility. Recent advancements in deep learning have significantly impacted the prediction of stock market volatility, as traditional forecasting methods often struggle with the complexities and non-linearity of financial data. Deep learning models, particularly variations of Recurrent Neural Networks (RNNs) and Long Short-Term Memory networks (LSTMs), have emerged as powerful tools for capturing the temporal dependencies inherent in stock price movements and volatility (Jia & Yang, 2021). They have been utilized effectively to process historical price data, uncovering complex relationships that simpler models fail to detect (Zhao et al., 2024). Velásquez et al. (2013) proposed using a dynamic neural network to forecast volatility across various international indices, implying significant advancements in predictive performance over conventional GARCH models. The integration of macroeconomic variables into these models has further enhanced their predictive capabilities, allowing for a comprehensive analysis that factors in broader market influences along with granular price patterns (Song et al., 2022). Moreover, the ability of deep learning models to adapt to high-dimensional data underscores their effectiveness in predicting volatility. For instance, studies have shown that deep neural networks excel at modeling intricate, non-linear relationships within stock price movements, thereby offering a distinct advantage over traditional linear models (Fan & Zhang, 2024; Liao et al., 2024). The introduction of advanced features, such as sentiment-driven indicators derived from news articles and social media, has also shown promise in reflecting the real-time dynamics of market sentiment, thus refining prediction accuracy (Jain et al., 2021).

Furthermore, in recent years, the Temporal Fusion Transformer (TFT) has emerged as a powerful model for multi-horizon time series forecasting, combining high accuracy with interpretability. The TFT was proposed by Lim et al. (2021) as an interpretable attention-based model combining recurrent layers and transformer-style self-attention for multi-horizon forecasting, and gained traction in numerous domains, notably in finance where forecasting tasks are often multi-scale, non-linear, and dependent on heterogeneous inputs. Its architecture includes several innovative components: gating mechanisms, static covariate encoders, variable selection

networks, and interpretable attention layers, allowing it to adaptively select relevant inputs at each time step (Lim et al., 2021). In contrast to pure Transformers (Vaswani et al., 2017), the TFT introduces mechanisms for handling static metadata and known future inputs, enabling realistic multi-step-ahead forecasting without recursive errors. The model's direct multi-horizon forecast structure and attention-based interpretability have made it particularly suitable for complex financial time series. Wu et al. (2022) and Zhou et al. (2021) compared TFT to other attention-based architectures like Autoformer and Informer, confirming TFT's superior performance in medium-range forecasting tasks. Several studies (e.g., Ji et al., 2023; Mazen et al., 2023) also introduced enhancements such as multi-task learning and hybrid TFT structures tailored for specific domains.

Volatility forecasting remains a core challenge in financial econometrics. Traditional models like GARCH (Bollerslev, 1986) and its extensions (e.g., EGARCH, FIGARCH, GJR-GARCH) are widely used but limited in capturing non-linear dependencies and high-dimensional covariates. TFT offers a promising alternative, due to their ability to utilize self-attention mechanisms, enabling them to model intricate relationships across sequential data effectively, which is crucial in capturing the volatility patterns in financial markets (Hartanto & Gunawan, 2024; Ji et al., 2023; Lim et al., 2021). The TFT model's structure allows for the incorporation of both historical price data and external indicators, such as macroeconomic variables, which have been shown to significantly impact volatility predictions (e.g. Kanarachos et al., 2023; Song et al., 2022). For example, the TFT, through its gating mechanisms, can dynamically adjust to the changes in the financial landscape, thus enhancing the predictability of complex market movements when compared to traditional models like GARCH (Ardia, 2019; Song et al., 2022). Furthermore, the adaptability of TFTs suggests their potential in analyzing the asymmetric nature of market volatility, considering factors such as leverage effects and time-varying conditions (Bhowmik & Wang, 2020; Catania & Proietti, 2019). Frank (2023) applied TFT to forecast weekly and monthly realized volatility of S&P 500 constituents, finding it outperformed GARCH, LSTM, and Random Forest models across multiple metrics. The inclusion of static and temporal covariates—such as VIX, macroeconomic indicators, and sectoral identifiers—significantly improved forecast accuracy. Similarly, Lin & Sun (2021) integrated sparse attention with TFT in forecasting asset volatility, noting improved performance over ARIMA, LSTM, and GRU models. In the cryptocurrency domain, Farooq et al. (2024) used a modified TFT model to forecast Bitcoin volatility, incorporating geopolitical risk indicators and on-chain metrics, with results showing superior accuracy to baseline deep learning models. In their original work, Lim et al. (2021) included a stock market volatility dataset as one of the case studies to demonstrate TFT's utility. Specifically, they modeled daily realized volatility of stock returns (for S&P 500 index constituents) as a multi-horizon forecasting problem. The TFT not only delivered accurate volatility forecasts, but its attention mechanisms yielded insights into the volatility dynamics. Notably, the model did not detect strong seasonal patterns in the volatility data, in contrast to other datasets like traffic or retail sales that showed clear daily/weekly seasonality. This finding aligns with financial theory: stock volatility is largely stochastic with little regular periodicity. The TFT's

attention weights for the volatility series were roughly uniformly distributed over past time steps (akin to a moving-average effect), suggesting the model was effectively smoothing out noise to capture any low-frequency trend in volatility.

This kind of interpretability is valuable for financial analysts, as it confirms that the model isn't erroneously perceiving cycles where none exist. The TFT has also shown strong results in forecasting stock returns and prices. Hu (2021) compared TFT with LSTM and SVR models for stock price forecasting and found that TFT yielded lower RMSE and higher directional accuracy. Hartanto & Gunawan (2024) applied TFT to the Indonesian stock market, reporting high accuracy and robust short-term forecasts using technical and macroeconomic indicators. Díaz Berenguer et al. (2024) enhanced TFT with causality-driven features, using transfer entropy from financial news to predict stock movements. Their TFT variant significantly outperformed LSTM, DeepAR, and pure Transformer baselines, suggesting TFT's suitability for integrating alternative data like sentiment and news. Multiple empirical studies benchmark TFT against traditional and deep learning models. In volatility prediction, TFT consistently outperforms GARCH, ARIMA, and LSTM models (Frank, 2023; Lin & Sun, 2021). For price forecasting, studies by Zhao & Yan (2024) found that TFT delivers better performance than Informer and Transformer encoder-decoder models across asset classes including FX and equities. In a survey of deep learning in finance, Fischer et al. (2023) highlighted TFT's unique position: it matches or surpasses the accuracy of other models while retaining interpretability. Comparative benchmarks (Wu et al., 2021; Zhou et al., 2021) showed that while Informer and Autoformer offer efficiency gains, TFT leads in accuracy for mid-range horizons. Bhardwaj et al. (2023) and Bhat et al. (2022) underlined that hybrid models (e.g., GARCH-LSTM) offer improvements over classical models, but TFT often surpasses even these when multiple covariates and time horizons are involved. Another core advantage of TFT is its built-in interpretability. Its variable selection network highlights the most influential features for each forecast, while attention scores identify key time steps contributing to predictions. Díaz Berenguer et al. (2024) leveraged attention heatmaps to interpret market regimes and driver variables in volatility prediction.

Studies such as Alonso & Carbo (2022) and Hajek & Novotny (2024) emphasized that explainable AI models are crucial for finance professionals. Another interpretability use-case demonstrated by Lim et al. (2021) was regime identification in financial markets. By analyzing attention weight shifts, they showed that TFT could help identify high-volatility regimes or significant market events. For instance, during the 2008 financial crisis period, the TFT model's attention patterns deviated notably from baseline, highlighting those periods as out-of-distribution regimes. These attention spikes corresponded to volatility surges, effectively flagging the crisis as a distinct regime. This is an important capability: in risk management, being able to detect regime changes (such as a shift from a low-volatility market to a turmoil period) is crucial. TFT's architecture addresses this demand, offering transparency not available in many deep learning models.

Research has expanded TFT applications beyond US equities. Yang et al. (2025) applied a multi-sensor TFT to predict Sharpe ratios using multi-task learning, showing robustness in both

US and European markets. Hajek & Novotny (2024) incorporated Google Trends data and FinBERT sentiment scores into TFT for European stock forecasting. In commodities, Zha et al. (2024) used TFT for risk estimation of crude oil future prices forecasting. Zhao & Yan (2024) applied TFT to FX prediction (NZD pairs), achieving a high explanation of the data by the model for short-term exchange rate forecasts. Cross-country studies (Ang & Lim, 2022; Diaz Berenguer, 2024) demonstrate TFT's applicability to global financial markets, especially when extended to include relational data such as supply chain networks or macroeconomic dependencies. In summary, while research is still evolving, TFT has shown great promise in financial time series forecasting, especially for modeling stock return volatility. Its ability to incorporate a wide range of predictors (financial, economic, or technical) and to provide interpretable outputs makes it attractive for analysts seeking to predict not just point values but also understand the uncertainty and drivers of financial volatility.

Early studies consistently report that TFT achieves higher predictive accuracy than traditional econometric models in volatility forecasting, and its interpretability features (like attention-based regime detection) offer additional value in financial decision-making contexts. They elucidate the utility of Temporal Fusion Transformers (TFT) in time series forecasting, specifically highlighting that financial returns exhibit regime-switching characteristics, including volatility changes. Lim et al. (2021) described, how identifying these temporary shifts can be essential for modeling financial time series accurately, asserting that significant regimes or events can lead to abrupt changes in volatility. This aligns with the broader literature that considers varying market conditions, emphasizing the need for models that can adapt to sudden changes. Moreover, the research presented by Diaz Berenguer et al. (2024) integrates causality-driven dynamics into stock movement predictions, utilizing TFT to account for different market conditions that may signify regime shifts. The dynamic transfer entropy approach they adopt for measuring interactions between stock prices and external news sentiments provides insights into how volatility regimes may evolve in response to new information. This reflects the importance of understanding external factors driving regime changes in financial markets. Hartanto & Gunawan (2024) further expands on multivariate forecasting within specific regimes by employing TFT models. Their investigation into short-term stock prices illustrates how feature engineering and technical data can enhance predictive accuracy during volatile periods, thus emphasizing the role of regime-switching behavior in shaping market predictions. Additionally, the comparative study by Lin & Sun (2021) highlights the strengths of transformer models over traditional LSTM approaches in capturing complex temporal dependencies, which are valuable in scenarios characterized by regime-switching. Their findings suggest that the self-attention mechanism in transformers can provide critical insights and adaptability to changing market conditions, thereby allowing for more robust forecasting models. Tang et al. (2024) introduce a period-aggregated transformer model (LPAST) aimed at learning latent seasonality in financial time series. Their work addresses how unobservable temporal patterns contribute to market volatility and the potential for abrupt regime shifts, advocating for models that incorporate these dynamic aspects to improve long-horizon predictions in financial markets.

The literature also discusses the implications of using sentiment analysis in forecasting, as highlighted by Yao & Li (2020), who examines the information flow between economic policy uncertainty and various financial time series. This type of analysis can be crucial for recognizing regime changes by correlating shifts in sentiment with market movements. Despite the apparent advantages, several challenges remain in the application of pure neural networks for stock market volatility prediction. Issues like overfitting are particularly prevalent due to the often-limited availability of historical data for model training in financial contexts (Chen et al., 2023). Researchers have proposed various strategies to mitigate this problem, including data augmentation techniques and the careful tuning of model hyperparameters to enhance generalization (Yu & Li, 2021).

3. MATERIAL AND METHODS

In my research I focused on two major issues. First, I measured volatility around a geopolitical shock and analyzed the existence of abnormal market dynamics in the form of abnormal returns and abnormal volatility for the MSCI World sector indices. Therefore, I used the event study approach of Campbell & Lo (1996), Fama et al. (1969), and MacKinley (1997) to analyze the stock sector's reaction to the Russian attack. For the evaluation of abnormal volatility, I used the approach of Ahmed et al. (2020), Beaver (1968), Brown & Warner (1985), Landsman & Maydew (2002), and Prasad et al. (2021). In terms of formal notation, I followed Robus et al. (2024).

Second, I analyzed the out-of-sample forecasting performance of the Temporal Fusion Transformer (TFT) to predict daily volatility of high important financial instruments in three periods of geopolitical events: COVID-19, the Russian attack on Ukraine, and the attack by Hamas terrorists on Israel. Thus, I trained a multivariate TFT network and generated 1 to 10 step ahead out-of-sample forecasts in a rolling window for nine financial instruments. I analyzed the volatility of the four most significant stock indices: S&P 500 stock index, NASDAQ stock index, Nikkei 225 stock index, Hang Seng; Gold as the essential instrument for hedging against inflation and a long-term store of value; Brent Crude Oil as the most important source of energy; the 10 years Treasury Bond as the essential reference for fixed income and the price of time; the exchange rate of Euro vs. US-Dollar as an indicator of the balance between the two most essential currencies; and Bitcoin as an instrument with increasing market capitalization, high potential of growth and a hedge against inflation. In addition, I used the VIX index future and the Geopolitical Risk Index (GPR; Caldara & Iacoviello, 2022) as explanatory variables in the multivariate forecasting approach. In particular, I examined whether the TFT outperformed different GARCH-class models (GARCH, EGARCH and GJR-GARCH). Therefore, I used 250 trading days of the financial instruments under consideration as training for the TFT. I used 200 trading days only for training the neural network and 50 trading days for validating the weights of the neural network. After training procedure, I generated predictions within the time periods for COVID-19 (02/03/20 - 04/30/2020), the Russian attack on Ukraine (02/01/2022 - 29/04/2022) and for the period around the Hamas attack on Israel (10/01/2023 - 12/29/2023) for the volatility of financial instruments. The forecasts were each 1 to 10 trading days into the future. I used the rolling window approach, whereby for each step of the rolling window in the above-mentioned period, the Temporal Fusion Transformer was first trained with a combination of different hyperparameters (attention head = 1, 4, 8 and hidden size = 16, 32, 64) and then the predictions for the next 1 to 10 trading days were generated. The rolling window was then shifted by one trading day. The TFT is known for its ability as a multivariate forecasting model. For this analysis I generated and analyzed multivariate forecasts.

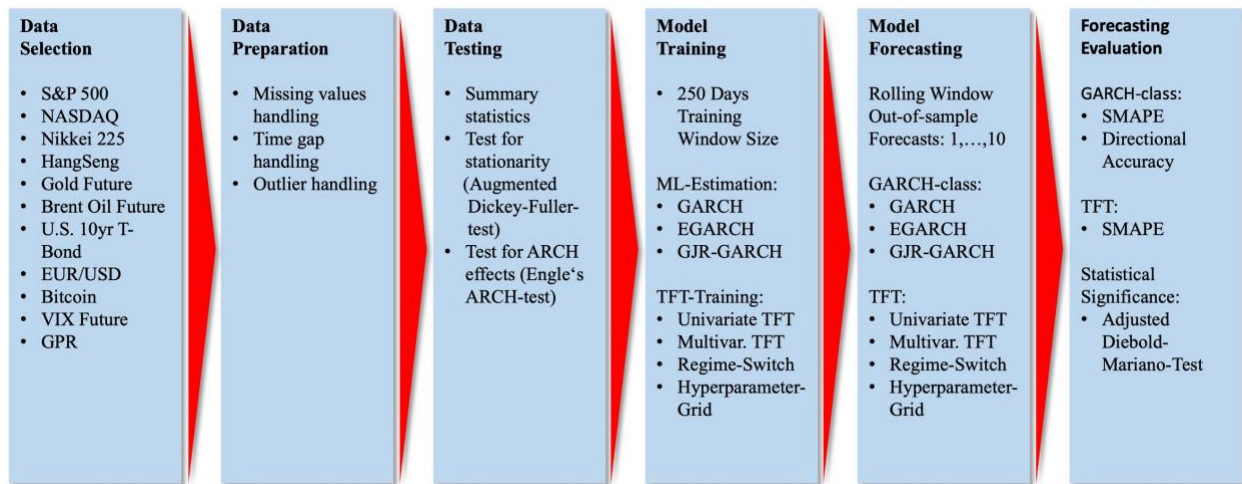


Figure 1: Research procedure for out-of-sample forecasting.

Source: Own illustration using MS PowerPoint (2025)

I predicted the volatility of all financial instruments simultaneously and included them in the training accordingly, so all the financial instruments mentioned are used as explanatory variables to forecast their volatility. In addition to the variables to be predicted, which in the method are both exogenous (as an explanatory variable for another variable to be predicted) and endogenous (as a variable to be predicted), I included the volatility index (VIX) future and the geopolitical risk index (GPR) as explanatory variables. I then tested whether the multivariate prediction approach achieves significantly better performance in predicting the volatility of the above financial instruments than the univariate approach, in which only the historical values of the financial instrument to be predicted itself serves as an explanatory variable. I evaluated the forecasting performance using the Symmetric Mean Average Percentage Error method (SMAPE).

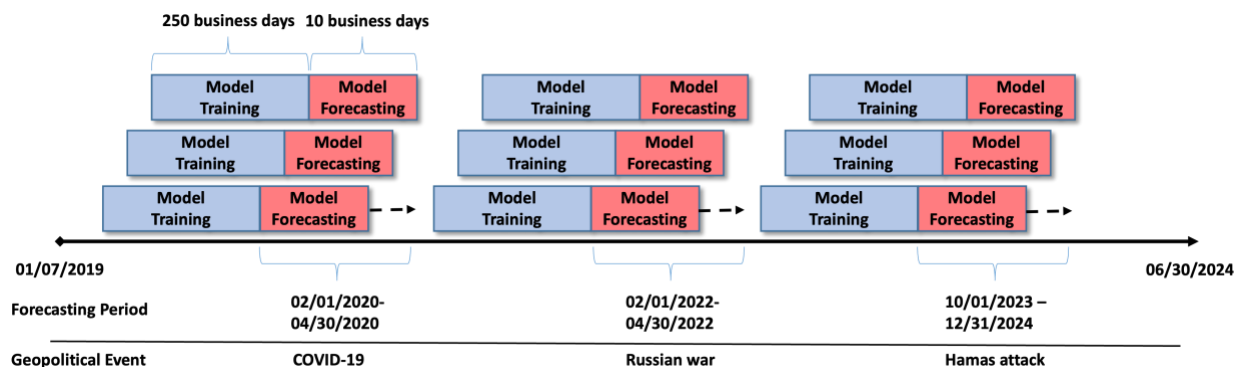


Figure 2: Rolling window approach for TFT out-of-sample forecasting.

Source: Own illustration using MS PowerPoint (2025).

My research focuses on how well the TFT can predict volatility behavior during periods of crisis. In this context, I also separately examine whether a TFT with an additional regime-switching feature performs better than the pure TFT model. As mentioned above, all forecasts were generated multiple times, on a grid of different hyperparameters of the TFT. On the one hand, this

is intended to increase the robustness of the results, and, on the other hand, I wanted to use it to examine whether variations in the hyperparameters can achieve significant improvements in forecasting performance.

3.1. Definitions of volatility

Here, I want to give a brief overview of the definitions of volatility. The statistical definition of volatility is the standard deviation of returns over a specific time horizon (Hull, 2022). This is the classical definition of volatility used in risk management and time series analysis. It assumes that returns r_t are random variables, and volatility captures their dispersion:

$$\sigma = \sqrt{\text{Var}(r_t)} = \sqrt{E[(r_t - \mu)^2]}$$

where r_t are log returns of a time series of a financial asset p_t and $\mu = E[r_t]$ is the expected value of r_t , for $t \in \{1, \dots, T\}$. The definition of conditional volatility arises from the observations, that volatility is not constant over time (Engle, 1982; Mandelbrot, 1963). Conditional volatility refers to the time-varying variance of returns, conditioned on past information (Engle, 1982). Based on this definition, ARCH (Engle, 1982) and GARCH models (Bollerslev, 1986) were developed. In the context of valuation of derivatives, the definition of implied volatility was created. Implied volatility is the market's expectation of the future volatility of an asset, as inferred from the price of options (Black & Scholes, 1973). It is derived by inverting the Black-Scholes or another option pricing model and reflects market sentiment and forward-looking expectations of risk. The definition of realized volatility approximates the true variance of returns over a day (or another interval) using summations of squared intraday returns and is e.g. applied in high-frequency finance, risk modeling and forecast accuracy tests. So, realized volatility (RV) is a non-parametric, ex-post measure of actual volatility, calculated using high-frequency intraday returns (Andersen et al., 2003):

$$RV_t = \sum_{i=1}^n r_{t,i}^2$$

where $r_{t,i}^2$ are intraday returns and interval $i = 1, \dots, n$. Frequently used in technical analysis, backtesting and VaR models the historical volatility (HV) is used. Historical volatility is the standard deviation of asset returns over a specified past window, without assuming time-variation (Poon & Granger, 2003). HV is often computed using daily, weekly or monthly returns and is a simple backward-looking definition. In the context of Bayesian econometrics, state-space modeling and asset pricing, the definition of stochastic volatility is used. In stochastic volatility models, volatility itself follows a latent stochastic process, rather than being deterministically linked to past returns (Taylor, 1986). So, volatility is modeled as a separate, unobserved process, often following a log-normal distribution, which can explain features like volatility clustering and leverage effects.

3.2. Abnormal Returns and Abnormal Volatility

To identify Abnormal Returns and Abnormal Volatility, I used the event study approach of Campbell & Lo (1996), Fama et al., (1969) and MacKinley (1997) to analyze global stock market sector's reaction to the Russian attack. For the evaluation of abnormal volatility, I use the approach of Ahmed et al., (2020), Beaver (1968), Brown & Warner (1985), Landsman & Maydew (2002) and Prasad et al., (2021). In terms of formal notation, I followed Robus et al. (2024). First, I calculated daily MSCI sector index returns by:

$$R_{s,t} = \ln\left(\frac{P_{s,t}}{P_{s,t-1}}\right)$$

where the closing price of the sector index s at day t is denoted by $P_{s,t}$. Further, I constructed the OLS market model for every sector s :

$$R_{s,t} = \alpha_s + \beta_s R_{m,t} + \varepsilon_{s,t}$$

where $\varepsilon_{s,t} \sim (0, \sigma_{s,t}^2)$ is white noise process, $R_{s,t}$ is the daily return of sector s at day t and $R_{m,t}$ is the daily market return at day t . α_s, β_s are the sector individual parameters of the market model. I used the MSCI as a benchmark for market returns $R_{m,t}$ I use the MSCI World Index. The OLS market model assumes that the conditional expectation of $R_{s,t}$ is given by:

$$E(R_{s,t} | I_{t-1}) = \alpha_s + \beta_s R_{m,t}$$

where I_{t-1} are the information which was realized up to day $t-1$.

I calculated abnormal returns by:

$$AR_{s,t} = R_{s,t} - E(R_{s,t}) = R_{s,t} - (\hat{\alpha}_s + \hat{\beta}_s R_{m,t})$$

where the estimated OLS market model parameters are denoted by $\hat{\alpha}_s$ and $\hat{\beta}_s$ for each sector s within the period of $t-250$ to $t-6$. Cumulated abnormal returns can then be calculated by:

$$CAR_{s,\tau_1,\tau_2} = \sum_{t=\tau_1}^{\tau_2} AR_{s,t}$$

Using a GARCH(1,1) approach to model the sector volatility, $\sigma_{s,t}^2$ can be described by the following:

$$\hat{\sigma}_{s,t}^2 = \hat{\omega}_s + \hat{\theta}_s \varepsilon_{s,t-1}^2 + \hat{\phi}_s \sigma_{s,t-1}^2$$

where $\hat{\omega}_s$, $\hat{\theta}_s$ and $\hat{\phi}_s$ are the estimated model parameters which must estimated within the period of $t-250$ to $t-6$. I computed abnormal volatility for each sector s at time t ($AVOLA_{s,t}$) by:

$$AVOLA_{s,t} = AR_{s,t}^2 / \hat{\sigma}_{s,t}^2$$

The procedure of an event study using Abnormal volatility is based on estimating the model parameters within the pre-event phase and then using this model to make predictions for the event

and post-event phase. Bialkowski et al., (2008) gave the notice that a one-step forecast does not produce an event-independent forecast. This problem can be solved by making the volatility forecast depends only on the information available before the event. For this reason, a h -th step ahead forecast must be used instead of h separate one step ahead forecast. I defined $T = t - 6$ and compute h -step out-of-sample Abnormal volatility forecasts for $h = 1, \dots, 31$ by:

$$AVOLA_{s,T+h} = AR_{s,T+h}^2 / \hat{\sigma}_{s,T+h}^2$$

$$AR_{s,T+h}^2 = (R_{s,T+h} - (\hat{\alpha}_s + \hat{\beta}_s R_{m,T+h}))^2$$

$$\hat{\sigma}_{s,T+h}^2 = \hat{\omega}_s \sum_{i=1}^{h-1} (\hat{\theta}_s + \hat{\phi}_s)^i + (\hat{\theta}_s + \hat{\phi}_s)^{h-1} + (\hat{\theta}_s \hat{\varepsilon}_{s,T}^2 + \hat{\phi}_s \hat{\sigma}_{s,T}^2)$$

For the analysis of whether volatility increases around the event day are persistent or not, I used the cumulative abnormal volatility ($CAVOLA_{s,t}$) with:

$$CAVOLA_{s,t+h} = \sum_{i=t}^{t+h} AR_{s,i}^2 / \sum_{j=t}^{t+h} \hat{\sigma}_{s,j}^2$$

3.3. GARCH-Class Models

Let P_t with a time series of a financial asset. Then

$$\ln \left(\frac{P_t}{P_{t-1}} \right) = r_t \quad t = 1, \dots, T$$

is called the time series of returns. The series of returns can now be described as the sum of a constant mean ($E[r_t] = \mu$) and a time dependent error term ($\sigma_t \varepsilon_t$) (Engle, 1982), where σ_t is the conditional variance and ε_t a white noise process:

$$r_t = \mu + \varepsilon_t \sigma_t \quad \text{with } \varepsilon_t \sim N(0,1) \text{ i.i.d.}$$

By demeaning the return series, it can be described as a time series of residual returns:

$$r_t - \mu = x_t = \sigma_t \varepsilon_t$$

The Generalized Autoregressive Conditional Heteroscedasticity GARCH(p,q) process was introduced by Bollerslev (1986) to model the time-varying behavior of the variance. In addition to the ARCH model developed by Engle (1982), in which only the squared historical error terms are included to model the current variance, the GARCH model also consists of the historical variance. This was due to empirical observations that volatility shows persistence. The GARCH model can be described by:

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i x_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2$$

where $\omega, \alpha_i, \beta_j, \gamma_i$ are the GARCH parameters with the following conditions:

$$\omega, \alpha_i, \beta_j, \gamma_i \geq 0 \quad i = 1, \dots, p \quad j = 1, \dots, q$$

The GARCH parameters $\omega, \alpha_i, \beta_j, \gamma_i$ determine the strength of the influence of the respective components. The EGARCH process was introduced by Nelson (1991) to model the leverage effect that occurs in many financial time series. He developed the EGARCH model for this purpose, which is good at capturing this asymmetry because it allows for the effect of the signs of the shocks on volatility. Another advantage, as pointed out by Nelson & Cao (1992), is that there are no restrictions on the model parameters:

$$\ln(\sigma_t^2) = \omega + \sum_{i=1}^p \alpha_i \left(\frac{|x_{t-i}^2|}{\sigma_{t-i}} - E \left[\frac{|x_{t-i}^2|}{\sigma_{t-i}} \right] \right) + \sum_{i=1}^p \gamma_{t-i} \frac{x_{t-i}^2}{\sigma_{t-i}} + \sum_{j=1}^q \beta_{t-j} \ln(\sigma_{t-j}^2)$$

The GJR-GARCH model of Glosten, Jagannathan, and Runkle (Glosten et al., 1993), which is also known as the threshold GARCH (T-GARCH) model, is proposed to capture the asymmetric behavior of volatility regarding good and bad news, and by allowing the current conditional variance has a different response to the past positive and negative returns, captured by the dummy variable D_t . The GJR-GARCH model is expressed as follows:

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \gamma_{t-i} D_{t-i} \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_{t-j} \sigma_{t-j}^2$$

$$\text{with } \omega, \alpha_i, \beta_j \geq 0 \quad \text{and } D_{t-i} = \begin{cases} 1, & \varepsilon_{t-i}^2 < 0 \\ 0, & \varepsilon_{t-i}^2 \geq 0 \end{cases}$$

The correct choice of the order of the GARCH model, i.e., the lags of past squared errors and variances included, is an essential question in modeling. Regarding Brooks & Burke (2003), French et al. (1987), Franses & Van Dijk (1996), Gokcan (2000), Pagan & Schwert (1990), and Visković et al. (2014) modeling σ_t as GARCH(1,1) process is sufficient to capture the volatility framework of financial data. In the analysis, I restricted the parameters to $p = 1, \dots, 5$ and $q = 1, \dots, 5$. Thus, I evaluated the best-fitting GARCH type model from a set of 75 competing models.

To estimate the GARCH(p,q) parameters, I used the approach of the maximum likelihood, regarding Bollerslev (1986), Bollerslev & Woodbridge (1992), and Nelson (1991). With the assumption of normal conditional returns, the estimation of the exemplary GARCH(1,1) process by the Maximum Likelihood function can be described as follows:

$$L(\omega, \alpha_1, \beta_1, \mu | r_1, \dots, r_T) = \prod_{i=1}^T \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(r_i - \mu)^2}{2\sigma_i^2}\right)$$

The L function is monotonically increasing, so one can maximize the log-likelihood function:

$$\max \ln L(\omega, \alpha_1, \beta_1, \mu | r_1, \dots, r_T) = -\frac{T}{2} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^T \ln \sigma_i^2 - \frac{1}{2} \sum_{i=1}^T \left(\frac{(r_i - \mu)^2}{\sigma_i^2} \right)$$

Suppose it is necessary to relax the assumption of normal distributed conditional returns, for example, to model the excess kurtosis of asset prices. In that case, one can assume that the conditional returns follow a generalized error distribution $f(x)$, e.g., Bollerslev et al. (1994) and McDonald & Newey (1988), which can be described by:

$$f(x) = \frac{\nu \exp(-\frac{1}{2}|\frac{x}{\lambda\sigma}|^{\frac{\nu}{\nu-2}})}{\lambda 2^{(1+1/\nu)} \Gamma(1/\nu)} \quad \text{where } \nu > 0 \quad \text{and } \lambda = \left[2^{-(2/\nu)} \frac{\Gamma(1/\nu)}{\Gamma(3/\nu)} \right]^{\frac{1}{2}}$$

With estimated GARCH parameters $\hat{\omega}, \hat{\alpha}_i, \hat{\beta}_j$ the estimated GARCH model can be forecasted.

For the GARCH(p,q) model and a forecast horizon h , the values of $\hat{\sigma}_{T+h}^2$ can be computed by (Campbell & Lo, 1996; Figlewski, 1997; Granger & Poon, 2001):

$$\begin{aligned} E[\hat{\sigma}_{T+1}^2 | I_T] &= \hat{\omega} + \sum_{i=1}^p \hat{\alpha}_i E[\hat{x}_{T-i+1}^2 | I_T] + \sum_{j=1}^q \hat{\beta}_j \hat{\sigma}_{T-j+1}^2 \\ E[\hat{\sigma}_{T+h}^2 | I_T] &= \hat{\omega} + \sum_{i=1}^p \hat{\alpha}_i E[\hat{x}_{T-i+1}^2 | I_T] + \sum_{j=1}^q \hat{\beta}_j \hat{\sigma}_{T-j+1}^2 \\ &= \hat{\omega} \sum_{s=1}^{h-1} \left(\sum_{i=1}^p \hat{\alpha}_i + \sum_{j=1}^q \hat{\beta}_j \right)^s + \left(\sum_{i=1}^p \hat{\alpha}_i + \sum_{j=1}^q \hat{\beta}_j \right)^{h-1} \left(\sum_{i=1}^p \hat{\alpha}_i \hat{x}_{T-i+1}^2 + \sum_{j=1}^q \hat{\beta}_j \hat{\sigma}_{T-j+1}^2 \right) \end{aligned}$$

For $h \rightarrow \infty$ it follows:

$$\lim_{h \rightarrow \infty} E[\hat{\sigma}_{T+h}^2 | I_T] = \frac{\hat{\omega}}{1 - \sum_{i=1}^p \hat{\alpha}_i - \sum_{j=1}^q \hat{\beta}_j}$$

3.4. The Temporal Fusion Transformer for volatility forecasting.

The Temporal Fusion Transformer (TFT) is an advanced deep learning model, specifically designed for forecasting multivariate time series with multiple prediction horizons. Introduced by Lim et al. (2021), it combines the strengths of various modeling approaches, including recurrent neural networks (RNNs; Hopfield, 1982; Rumelhart et al., 1986) particularly Long Short-Term Memory (LSTM) networks (Hochreiter & Schmidhuber, 1997), self-attention mechanisms (Vaswani et al., 2017), and interpretable model components. The TFT is structured to efficiently model highly complex temporal dependencies while simultaneously providing robust interpretability of the results. A central feature of the TFT is its multi-horizon forecasting capability, which allows for simultaneous predictions across multiple future time steps. The model supports probabilistic forecasting, which is especially valuable in fields such as financial market prediction. The probabilistic nature of the forecasts is typically achieved through the use of quantile regression, enabling the estimation of prediction intervals and the explicit representation of uncertainties. A distinguishing feature of the TFT is its interpretability. Beyond variable

selection, the attention mechanisms provide insights into the temporal significance of individual observations. By analyzing feature importance scores, users can understand which variables and time steps were critical for the forecast.

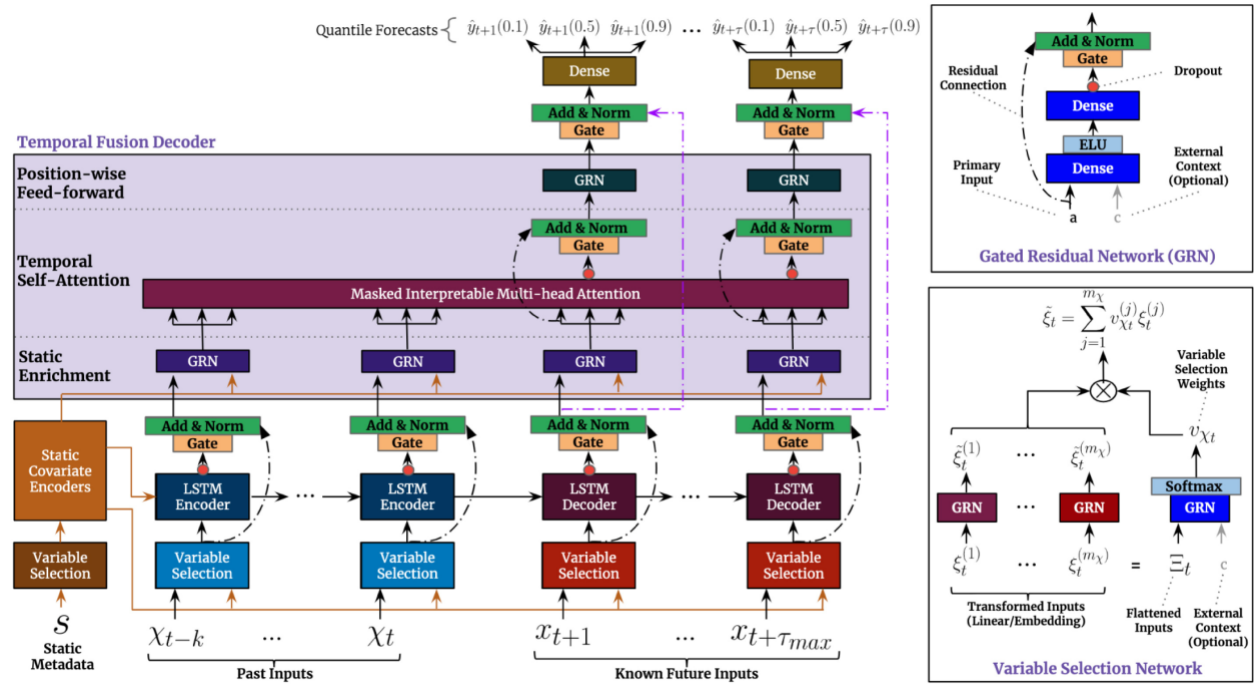


Figure 3: Temporal Fusion architecture. Source: Lim et al. (2021).

This attribute is particularly relevant in the financial sector, where model decision traceability is essential. In summary, the Temporal Fusion Transformer represents a significant advancement in time series forecasting by combining high predictive accuracy with robust interpretability. Due to its flexible architecture, the TFT is suitable for a wide range of applications, including financial market forecasting, energy consumption prediction, demand forecasting in supply chain management, and healthcare forecasting. Its ability to effectively model both short-term and long-term dependencies makes it a powerful tool for modern time series analysis tasks. Here I want to discuss just some of the TFT elements, the full description can be taken from Lim et al. (2021). The TFT forecasting framework can be described by:

$$\hat{y}_{i,t+\tau}(q, t, \tau) = f_q(\tau, y_{i,t-k:t}, \mathbf{z}_{i,t-k:t}, \mathbf{x}_{i,t-k:t+\tau}, \mathbf{s}_i)$$

where $\hat{y}_{i,t+\tau}(q, t, \tau)$ is the predicted q th sample quantile of the τ -step ahead forecast at time t , for the target variable $y_{i,t+\tau}$ and $f_q(\cdot)$ is a prediction model. The TFT incorporates different types of explanatory input features: static covariates $\mathbf{s}_i \in \mathbb{R}^{m_s}$, time-dependent observable input features, which can only be measured at each step and are unknown beforehand $\mathbf{z}_{i,t} \in \mathbb{R}^{m_z}$ and time-dependent known inputs, which can be predetermined $\mathbf{x}_{i,t} \in \mathbb{R}^{m_x}$ (e.g. the day-of-week at time t). Past information is incorporated within a finite look-back window k , where outcomes of the target variable $y_{i,t-k:t} = \{y_{i,t-k}, \dots, y_{i,t}\}$ are available up to time t . Known inputs $\mathbf{x}_{i,t-k:t+\tau} = \{\mathbf{x}_{i,t-k}, \dots, \mathbf{x}_{i,t+\tau}\}$ are available at any time across the entire time series range.

The TFT model architecture consists of the following elements (Figure 3): Variable Selection Networks (VSN), Gated Residual Networks (GRN), Static Covariate Encoders, LSTM-based local processing (Encoder and Decoder), a Multi-head Attention mechanism, the Temporal Fusion decoder and the Quantile Prediction Layer. The Variable Selection Network (VSN) dynamically weights each input feature (static, past, and future) based on its relevance to the prediction task. This is done using Gated Residual Networks (GRNs), which enable adaptive deep architectures and efficiently capture nonlinear relationships and is combined with soft attention. It allows the model to focus on the most important signals at each time step, ignore irrelevant noise, and improves interpretability and robustness, especially when many covariates are available (Lim et al. 2021):

$$\mathbf{v}_{\chi_t} = \text{softmax}(\text{GRN}_{\chi_t}(\mathbf{\Xi}_t, \mathbf{c}_s))$$

where $\mathbf{v}_{\chi_t} \in \mathbb{R}^{\chi_t}$ is a vector of variable selection weight and $\mathbf{c}_s$ is obtained from a static covariate encoder, $\xi_t^{(j)} \in \mathbb{R}^{m_\chi}$ is the transformed input of the j th variable at time t with $\mathbf{\Xi}_t = \left[\xi_t^{(1)T}, \dots, \xi_t^{(m_\chi)T} \right]^T$ is the flattened vector of all past inputs at time t .

The second element, Gated Residual Networks, transform and filter inputs via a gated mechanism, enabling the model to learn non-linear feature interactions while preserving gradient flow. They use a residual connection combined with gating and normalization. This allows the model to conditionally pass information forward or suppress it. GRNs are used extensively throughout the TFT, especially in feature selection and temporal processing. (Lim et al. 2021). These encode time-invariant inputs (e.g., asset type, geography) into embeddings that influence the entire network. Static Covariate Encoders are used to modulate various parts of the model, including variable selection, LSTM initial states, and attention scaling. This allows the model to tailor its internal dynamics depending on static properties of the input series. It provides a global conditioning mechanism for sequence modeling. Further, the TFT uses an LSTM encoder to process historical time series and a separate decoder for known future inputs. These LSTMs model short- and mid-range temporal dependencies effectively. They provide a localized understanding of temporal patterns, which is complemented by the attention mechanism for longer-range dependencies. Integrating future-known covariates such as holidays or planned interventions are done by this decoder (Lim et al. 2021).

The Multi-Head Attention layer applies parallel attention heads to focus on different time steps and types of temporal dependencies. Each head learns to attend to relevant parts of the sequence based on a different learned projection. This mechanism allows the model to capture long-term dependencies between different time steps by assigning weights that reflect each time step's relative importance for prediction. As a result, the model can effectively identify and incorporate seasonal patterns, trend shifts, abrupt events and other complex patterns that LSTMs may miss. It also supports interpretability by revealing which time steps influenced the forecast (Lim et al. 2021). The Temporal Fusion decoder combines outputs from the LSTM decoder and

attention layer, aligning and fusing them to generate final representations. It ensures the model balances both local sequential patterns (via LSTM) and global dependencies (via attention). This fusion is gated and residual in nature, maintaining efficient information flow and training stability. It prepares the context-aware features for final prediction (Lim et al. 2021). Finally, the Quantile Prediction Layer outputs multiple quantile forecasts (e.g., 10%, 50%, 90%) using fully connected layers. This allows the model to represent uncertainty in its forecasts, not just point estimates. It's particularly valuable in high-variance or risk-sensitive domains like finance or energy. The use of quantile loss (pinball loss) encourages accurate estimation of prediction intervals (Lim et al. 2021).

To train the TFT, I used the combination of hyperparameters as shown in Table 1.

Table 1: Applied hyperparameter for Temporal Fusion Transformer training and forecasting

TFT Parameters	Value
Learning Rate	0.001
Dropout Rate	0.01
Minibatch Size	32
Attention Head	1, 4, 8
Hidden Size	16, 32, 64
Max. Epochs	100
Loss Metrics	MAE, MSE

Source: Own collection (2025).

In doing so, I based the hyperparameter selection by publications such as: Frank (2023); Hartanto & Gunawan (2024) and Lim et al. (2021).

Attributed to its flexibility, TFT is well-suited for modeling financial market time series, which are the focus of this research scope. These series are characterized by complex and often non-linear temporal dynamics. These include short-term fluctuations, long-term structural dependencies, volatility clustering, regime shifts and abrupt volatility spikes (Box et al. 2015; Engle & Bollerslev, 1986; Glosten et al. 1993). Short-term behaviors, such as return autocorrelations (at intraday level), momentum effects, and reactions to market news, are captured in the TFT through a combination of Gated Residual Networks (GRNs), LSTM encoders, and decoder self-attention. GRNs dynamically weight input features at each time step, allowing the model to focus selectively on the most informative variables in response to evolving market conditions. (Farooq et al., 2024; Hartanto & Gunawan, 2024; Lim et al., 2021). For example, during a brief momentum phase, the model can increase reliance on recent return-based indicators, while deprioritizing static or lagged variables. The LSTM encoder complements this by learning sequential dependencies from recent observations, effectively capturing patterns such as price reversals, temporary shocks, or drift. In the decoding stage, the self-attention mechanism allows the model to refine its predictions by considering interdependencies across the forecast horizon, helping to model the short-term evolution of financial variables more precisely (Lim et al., 2021; Zhang et al., 2025).

Long-term structures in financial time series, such as macroeconomic cycles, seasonal patterns, or structural breaks, are addressed through temporal self-attention, static covariate encoders, and positional embeddings. The temporal self-attention mechanism allows the model to assign relevance to past time steps regardless of their distance from the present, making it ideal for capturing lagged relationships and recurring long-term dependencies. For instance, the model can learn to associate certain market behaviors with previous policy announcements or macroeconomic releases. The static covariate encoder incorporates fixed characteristics, such as asset class, sector, or region, that influence the long-run dynamics of each series. These encodings are broadcast across the model, conditioning the forecast on the broader context. Additionally, positional encodings inject information about the location of each time step within the sequence, helping the model distinguish between different phases of the calendar (e.g., month-end effects, quarterly earnings cycles) (Laborda & Zamanillo, 2023; Lim et al., 2021).

One of the defining features of financial time series is volatility clustering, the empirical tendency for periods of high volatility to be followed by further high volatility, and likewise for low-volatility periods (Bollerslev, 1986). In the TFT, volatility clustering is effectively modeled through the interplay of the LSTM encoder, GRNs, and temporal self-attention. The LSTM captures recent patterns in volatility behavior, such as surges in realized variance or rapid fluctuations in returns, by encoding them into its internal state. GRNs dynamically modulate the importance of inputs such as rolling standard deviations, implied volatility indices, or liquidity measures, enabling the model to adapt to early signals of clustering. Temporal self-attention allows the TFT to recall similar historical regimes, even if they occurred at distant time points, providing a mechanism for pattern recognition that is not limited by sequential memory constraints. Collectively, these components enable the model to learn both the persistence and evolution of volatility over time (Beck et al., 2025; Lim et al., 2021).

While volatility clustering describes persistent variance over time, volatility spikes represent sudden and often exogenous breaks in the variance structure, frequently associated with macroeconomic shocks, geopolitical events, or market crises. The TFT is well-suited to capture such discontinuities due to its architectural flexibility and contextual awareness. GRNs enable rapid re-weighting of features when market conditions change abruptly, allowing the model to respond to signals such as unexpected macroeconomic data or news sentiment indicators (Shen et al., 2025). The LSTM encoder preserves short-term temporal patterns that often precede spikes, such as increased intraday volatility or widening bid-ask spreads. Through temporal attention, the model can draw on analogs from past volatility episodes, identifying similar historical dynamics that inform its current predictions. The decoder's self-attention mechanism enables the forecast to adjust dynamically across the prediction horizon, allowing the model to capture whether a spike is expected to dissipate quickly or initiate a prolonged regime shift (Yang et al., 2025). Finally, gating mechanisms and residual connections ensure that the model can transition smoothly between stable and volatile regimes without losing predictive stability.

Quantile forecasting capabilities of the Temporal Fusion Transformer (TFT) present substantial opportunities for e.g. enhancing risk management, portfolio management, and

regulatory supervision in financial contexts. By directly modeling conditional distributions and providing accurate quantile estimates, the TFT is particularly adept at capturing complex financial dynamics, including tail risks and nonlinear market behaviors. In risk management, the TFT's quantile forecasting allows for precise estimation of key risk metrics such as Value-at-Risk and Expected Shortfall (e.g. Merlo et al., 2021; Petneház, 2021; Zha et al., 2024). Banks and financial institutions can utilize these advanced forecasts to dynamically adjust risk exposure, ensuring compliance with stringent regulatory requirements. TFT quantile forecasts further contribute to credit risk management by enabling scenario-based predictions of default probabilities and credit losses.

In portfolio management, TFT-based quantile predictions enable fund managers to better anticipate and navigate extreme market conditions. By incorporating tail-risk forecasts into portfolio construction, managers can dynamically hedge or rebalance asset allocations, effectively minimizing downside risks and optimizing returns under stressed scenarios. The TFT's ability to capture volatility clustering and spikes further strengthens the fund manager's capability to time factor exposures and sector rotations effectively (e.g. Hartanto & Gunawan, 2024; Yang et al., 2025).

Regulatory supervisory applications benefit significantly from TFT quantile forecasting, particularly through enhanced stress-testing frameworks. Supervisory authorities can deploy TFT models to generate detailed, scenario-based projections of systemic risks, improving the accuracy and interpretability of regulatory stress tests such as those conducted under CCAR and by the European Banking Authority (EBA). TFT forecasts provide supervisors with critical insights into potential vulnerabilities, facilitating proactive policy responses and macroprudential interventions aimed at maintaining financial stability. Additionally, the improved precision in modeling extreme financial events through quantile forecasting assists institutions in determining regulatory capital requirements more accurately, ensuring capital adequacy and regulatory compliance (Merlo et al., 2021; Storti & Wang, 2022; Taylor, 2019).

3.5. Forecasting Evaluation

Volatility forecasting itself has been approached by various methodologies, including GARCH-type models, neural networks, and hybrid approaches. Reviews on forecasting volatility, such as those by Gospodinov et al. (2006) and Poon & Granger (2003) emphasize that the choice of forecasting evaluation metric can be as critical as the forecasting model itself. While much of the literature has traditionally focused on metrics such as root mean squared error (RMSE) or mean absolute error (MAE), the Symmetric Mean Absolute Percentage Error (SMAPE) has gained attention as an alternative to traditional forecast accuracy metrics in financial econometrics, particularly for evaluating volatility forecasts. Unlike measures such as Mean Absolute Percentage Error (MAPE), which can become undefined or highly unstable in the presence of zero or near-zero actual values, SMAPE normalizes forecast errors by the average of the absolute predicted and actual values, yielding a bounded and symmetric error metric (Goodwin & Lawton, 1999; Makridakis, 1993). This formulation ensures that SMAPE avoids the infinite or excessively large

errors that can occur in MAPE when actual values approach zero. As such, SMAPE has been proposed as more robust and interpretable in settings where volatility varies widely or where heteroskedasticity is present (Makridakis & Hibon, 2000; Taylor, 2004). These properties make it particularly suitable for evaluating financial volatility forecasts, where both the scale and asymmetry of predicted and realized volatility can fluctuate substantially across time and assets (Bao et al., 2017; Kim & Won, 2018). The SMAPE can be described by:

$$SMAPE(\{y_t, \hat{y}_t\}_{t=1, \dots, T}) = \frac{1}{T} \sum_{t=1}^T \frac{|y_t - \hat{y}_t|}{(|y_t| + |\hat{y}_t|)/2}$$

where y_t is the observed value of the target variable and $\hat{y}_t$ a corresponding forecast in time t . SMAPE is particularly advantageous when volatility estimates are small or heterogeneous in scale, conditions under which percentage-based errors (e.g., MAPE) tend to exhibit bias or instability (Goodwin & Lawton, 1999; Makridakis, 1993). By symmetrizing the relative error, SMAPE ensures that over-predictions and under-predictions are treated with equal penalization, which is crucial when evaluating models such as GARCH-type processes, stochastic volatility models, or machine learning-based volatility forecasts (Hansen & Lunde, 2005; Taylor, 2004). This is particularly true in the evaluation of models that forecast volatility, where large shocks or outlier events can distort error measures that are not similarly bounded (Pereira et al., 2024). Furthermore, the debate regarding the optimal criteria for forecast evaluation extends beyond the use of traditional statistical loss functions. As highlighted by Degiannakis & Kafousaki (2023), the selection of an appropriate evaluation metric can lead to seemingly contradictory conclusions about model performance. In this context, SMAPE's bounded nature ensures that forecast errors are confined within fixed limits, which facilitates a more consistent comparison across models and forecast horizons. Complementary insights from Sadorsky & McKenzie (2008) further underscore the importance of choosing an error metric that accurately reflects the relative forecasting performance over different horizons, which is critical in volatility forecasting given the dynamic nature of financial markets. Several studies have employed SMAPE to assess the out-of-sample predictive performance of volatility models. For instance, Makridakis & Hibon (2000) used SMAPE in large-scale forecast competitions to compare competing models under diverse volatility regimes. In financial econometrics, Taylor (2004) applied SMAPE in comparing realized volatility forecasts generated from GARCH and stochastic volatility models. The measure has also been adopted in more recent machine learning-based volatility forecasting literature, where it provides robustness to scale variations in neural networks and ensemble models (Bao et al., 2017; Kim & Won, 2018). Despite its practical appeal, some researchers have noted potential limitations of SMAPE, including non-differentiability at zero and boundedness that may understate large relative errors when actual values are small. Nevertheless, its interpretability and symmetry make it a valuable complement to traditional metrics such as RMSE or MAE, especially when the focus is on percentage-based relative error across varying volatility scales.

The main objective of the study is to determine the usefulness of the Temporal Fusion Transformer in practical applications for predicting volatility under geopolitical shocks. With this

in mind, I addressed the question of the extent to which presenting the results of a variety of different Forecast Evaluation Criteria (FEC) favors a decision-theoretically consistent model decision. I decided to use SMAPE solitary for the following reasons. Granger & Newbold (2014) and Elliot & Timmermann (2008) noted that a consistent decision-making process is only possible if the FEC corresponds to the actual loss function. In the analysis, I wanted to predict the volatility of various financial instruments to make portfolio adjustments and hedging for a portfolio, for example. Given this goal, I'm interested in the percentage deviation of the forecast. Furthermore, Diebold & Mariano (1995) showed that models can be ranked differently under different FECs, making model selection somewhat arbitrary. In addition, statistical tests are only defined within a consistent evaluation criterion, from which it follows that the statistical significance can only be calculated consistently for one type of FEC at a time. Consequently, the simultaneous use of multiple FECs can lead to contradictory model rankings, making rational model selection more difficult (Hansen et al., 2011). Including several FECs can also lead to overfitting of the prediction models (Timmermann, 2006).

To assess the statistical significance of the difference in performance between two forecasting models, we use the Diebold-Mariano test statistics (Diebold & Mariano, 1995). The goal hereby is, to assess whether to competing forecasting models have equal predictive accuracy. Therefore, the null hypothesis can be described by:

$$H_0: E[d_t] = 0$$

where $d_t = L(y_t, \hat{y}_{1,t}) - L(y_t, \hat{y}_{2,t})$ and L is a certain loss function, y_t is the observed realization of the target variable in time t , which must be forecasted and $\hat{y}_{i,t}$ is a forecasting model to predict the value of y_t . To evaluate these hypothesis, Diebold and Mariano (1995) constructed the following test statistics:

$$DM = \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}}$$

where $\bar{d} = \frac{1}{T} \sum_{t=1}^T d_t$ is the mean loss differential and $\hat{V}(\bar{d})$ is the estimated variance of the mean loss differential. To account for possible heteroskedasticity and autocorrelation effects in the data, we adjust the Diebold-Mariano test statistics for the variance estimation, proposed by Newey & West (1987):

$$\hat{V}(\hat{d}) = \frac{1}{T} (\gamma_0 + 2 \sum_n^N w_n \gamma_n)$$

where γ_n is the autocovariance of the time series d_t with lag n and $w_n = 1 - \frac{n}{N+1}$ are the Bartlett-weights (Bartlett, 1946). For small samples, the DM test statistic can be biased, so, we adjust it by the approach of Harvey et al. (1997):

$$DM^* = \frac{DM}{\sqrt{\frac{T+1-2h+h(h-1)}{T}}}$$

where h is forecast horizon and T is sample size. The DM^* test statistics now follows a t -distribution, for which the corresponding values for the evaluation of the DM^* test must then be compared.

3.6. Applied Datasets

3.6.1. MSCI World Sector Indices for Identification of Abnormal Returns and Abnormal Volatility

In the following, I will briefly discuss the dataset and present some statistical analysis. In the analysis I used the dataset and the description of Robus et al. (2024). “The dataset contains 11 MSCI sector index time series (MSCI, 2024b) and the MSCI World (MSCI, 2024a) itself (Table 2) with closing prices (in USD) from March 04, 2021, to March 31, 2022. Each series provides a total of 281 business days. The data was collected via MSCI (2024a, 2024b). I oriented at e.g., Federle et al. (2022), Izzeldin et al. (2023) and Yousaf (2022) by defining the event date to the February 24, 2022, and name it as t .

For the event study, I have split the data set for each sector time series into three parts. First, the pre-event phase, contains 250 observations from March 4, 2021, to February 16, 2022” (Robus et al., 2024). I used this data ($t - 255$ to $t - 6$) to estimate the parameters for the OLS market model and the corresponding GARCH process. The event phase from February 17, 2022, to March 03, 2022 ($t - 5$ to $t + 5$) contains the observations which provides the market development around the Russian attack. In accordance with Robus et al. (2024) and Yousaf et al., (2022), I used this period for the analysis of the market reaction. The post event phase, period contains observations from March 04, 2022, to March 31, 2022, and was used to examine whether the abnormal volatility is persistent and how long it remains at an above-average level compared to the pre-event phase. To achieve this, I combined the event phase and the post-event phase ($t - 5$ to $t + 25$).

I also want to provide some summary statistics for MSCI sector indices shortly. Details are described at Robus et al., (2024). An illustration of the time series of log-returns is given by Figure 4. As expected, I observed close to zero average and median daily returns for the MSCI World index in the pre-event phase (Table 3), positive post-event (Table 5) and strongly negative during event phase (Table 4). Overall, it can be seen from this that the Russian attack initially led to sell-offs. However, one can see a strong recovery in the post-event phase, which is very different for the individual sectors, indicating a reallocation regarding the sectors. This is understandable insofar as the expectations of market participants regarding the development of business models within the sectors have changed and they are reacting accordingly. The highest daily losses during the event phase were recorded by the sectors: Financials (-3.41%), consumer staples (-2.82%), materials (-2.64%), telecommunications (-2.63%), consumer discretionary (-2.51%), utilities (-

1.97%), IT (-1.69%), real estate (-1.68%), industrials (-1.63%), energy (-1.46%) and health care (-1.25%).

To compare sample variances between pre-event and event phase, I applied Levene's test statistic (Table 6). The test is robust to deviations from the assumption that the samples to be compared must be normally distributed. Furthermore, the sample sizes of the two samples can be different. I found that almost all sector variances are significantly different from zero, which means that variances during the event phase are significantly higher. Data variability can also be seen, when looking at the range of the data (difference between minimum and maximum) for the individual sectors. I found that it widens for event- and post-event phase, comparing to the pre-event phase. These are all signs of an increase in the fluctuation of share prices in connection with the Russian attack.

Table 2: MSCI World Sector Indices.

Index	ISIN	Abbreviation
MSCI World Index	MIWO00000PUS	MSCI
MSCI World Consumer Discretionary Index	MIWO0CD00PUS	CD
MSCI World Consumer Staples Index	MIWO0CS00PUS	CS
MSCI World Energy Index	MIWO0EN00PUS	EN
MSCI World Financials Index	MIWO0FN00PUS	FN
MSCI World Health Care Index	MIWO0HC00PUS	HC
MSCI World Industrial Index	MIWO0IN00PUS	IN
MSCI World Information Technology Index	MIWO0IT00PUS	IT
MSCI World Materials Index	MIWO0MT00PUS	MT
MSCI World Real Estate Index	MIWO0RE00PUS	RE
MSCI World Telecommunications Index	MIWO0TC00PUS	TC
MSCI World Utilities Index	MIWO0TC00PUS	UT

Source: Morgan Stanley Capital International (MSCI, 2024b), Robus et al. (2024).

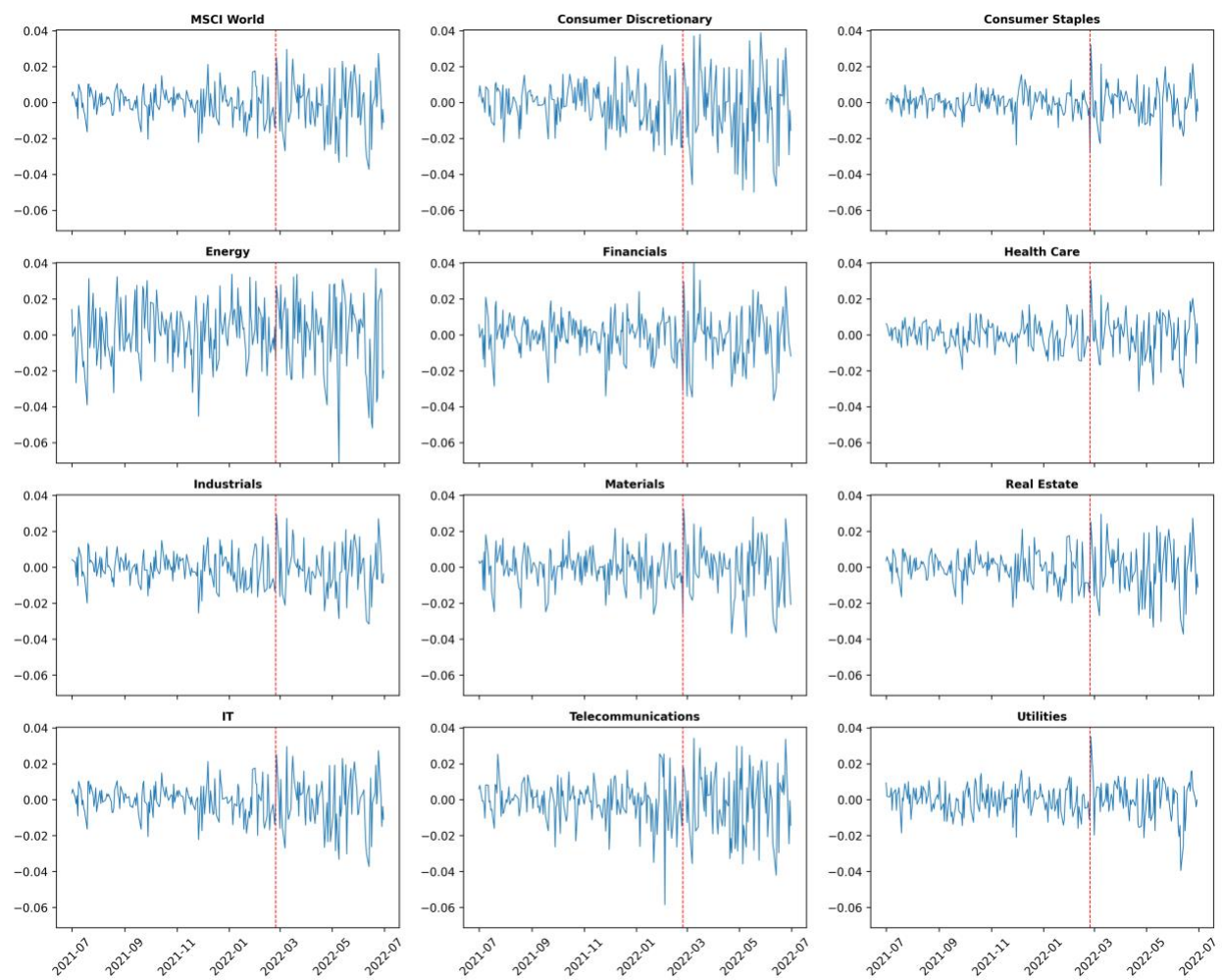


Figure 4: Log-returns for MSCI World Sector Indices (07/01/2021 – 06/30/2022). Dashed line marks the day of the Russian attack on Ukraine at February 24, 2022.

Source: Own illustration using Python Matplotlib (2025).

Table 3: Summary statistics for daily MSCI World Sector Index log-returns (in %) for pre-event phase 03/04/21 – 02/16/22

	MSCI	CD	CS	EN	FN	HC	IN	IT	MT	RE	TC	UT
Min	-2.22%	-2.92%	-2.36%	-4.54%	-3.41%	-1.92%	-2.55%	-2.22%	-2.82%	-2.22%	-5.85%	-2.20%
1.Q	-0.31%	-0.47%	-0.22%	-0.77%	-0.47%	-0.34%	-0.37%	-0.33%	-0.47%	-0.33%	-0.44%	-0.40%
Med.	0.10%	0.12%	0.07%	0.13%	0.10%	0.06%	0.02%	0.09%	0.08%	0.11%	0.01%	0.12%
Mean	0.04%	0.03%	0.05%	0.11%	0.06%	0.05%	0.02%	0.03%	0.03%	0.04%	-0.01%	0.03%
3.Q	0.50%	0.68%	0.40%	1.00%	0.70%	0.45%	0.56%	0.50%	0.61%	0.50%	0.57%	0.53%
Max	2.12%	3.20%	1.55%	3.38%	2.42%	1.68%	1.71%	2.12%	2.17%	2.12%	2.54%	1.64%
StD	0.73%	1.00%	0.54%	1.47%	0.90%	0.66%	0.77%	0.73%	0.92%	0.74%	1.01%	0.70%

Source: Morgan Stanley Capital International (MSCI, 2024b), Robus et al. (2024).

Table 4: Summary statistics for daily MSCI World Sector Index log-returns (in %) for event phase 02/17/22 – 03/03/22

	MSCI	CD	CS	EN	FN	HC	IN	IT	MT	RE	TC	UT
Min	-1.69%	-2.51%	-2.82%	-1.46%	-3.41%	-1.25%	-1.63%	-1.69%	-2.63%	-1.69%	-2.63%	-1.97%
1.Q	-1.13%	-2.15%	-0.72%	-0.79%	-1.64%	-0.42%	-1.14%	-1.13%	-0.72%	-1.13%	-1.12%	-0.45%
Med.	-0.74%	-0.91%	-0.23%	-0.17%	-0.62%	-0.29%	-0.86%	-0.74%	-0.46%	-0.74%	-0.88%	-0.11%
Mean	-0.36%	-0.77%	-0.14%	0.27%	-0.77%	-0.01%	-0.29%	-0.36%	-0.09%	-0.32%	-0.37%	0.05%
3.Q	-0.10%	0.48%	0.13%	0.88%	-0.30%	-0.10%	0.23%	-0.10%	0.50%	0.06%	0.29%	0.22%
Max	2.52%	2.25%	3.20%	2.79%	2.95%	3.07%	2.94%	2.52%	3.24%	2.52%	1.87%	3.50%
StD	1.19%	1.55%	1.36%	1.38%	1.73%	1.09%	1.29%	1.19%	1.43%	1.20%	1.30%	1.31%

Source: Morgan Stanley Capital International (MSCI, 2024b), Robus et al. (2024).

Table 5: Summary statistics for daily MSCI World Sector Index log-returns (in %) for post-event phase 03/04/22 – 03/31/22.

	MSCI	CD	CS	EN	FN	HC	IN	IT	MT	RE	TC	UT
Min	-2.69%	-4.58%	-2.28%	-2.48%	-3.46%	-1.68%	-2.11%	-2.69%	-1.84%	-2.69%	-3.56%	-0.67%
1.Q	-0.81%	-1.25%	-0.89%	-0.74%	-0.46%	-0.47%	-0.47%	-0.81%	-0.59%	-0.81%	-1.08%	-0.12%
Med.	0.12%	0.12%	0.36%	-0.04%	0.15%	0.36%	0.02%	0.12%	0.25%	0.12%	-0.03%	0.45%
Mean	0.18%	0.26%	0.03%	0.24%	0.14%	0.21%	0.14%	0.18%	0.19%	0.18%	0.12%	0.33%
3.Q	1.10%	1.87%	0.63%	1.70%	1.21%	0.85%	0.51%	1.10%	0.81%	1.10%	1.41%	0.89%
Max	2.96%	3.80%	2.14%	3.37%	4.04%	2.22%	2.72%	2.96%	2.41%	2.96%	3.43%	1.32%
StD	1.39%	2.13%	1.12%	1.80%	1.71%	1.07%	1.18%	1.39%	1.15%	1.39%	1.74%	0.61%

Source: Morgan Stanley Capital International (MSCI, 2024a, 2024b), Robus et al. (2024).

Table 6: Levene's test for equality of sample variances between pre-event and event phase (* test was rejected on a 5% level).

	MSCI	CD	CS	EN	FN	HC	IN	IT	MT	RE	TC	UT
Levene's test	2.66*	2.40*	6.34*	0.88	3.69*	2.73*	2.81*	2.66*	2.42*	2.63*	1.66	3.50*

Source: Robus et al. (2024)

3.6.2. Dataset for Temporal Fusion Transformer Forecasting

This section will present and discuss the dataset, I used to evaluate the forecasting performance of the Temporal Fusion Transformer and the different GARCH-type models. I collected daily time series data on closing prices for selected financial instruments (Table 7).

Table 7. Financial instruments in scope for volatility forecasting analysis.

Financial Instrument	ISIN
S&P 500 Index	US78378X1072
NASDAQ 100 Index	US6311011026
Nikkei 225 Index	JP9010C00002
Hang Seng Index	HK0000004322
Gold	XC0009655157
Brent Crude Oil	XC0009677409
EUR/USD	EU0009652759
10yr. Treasury-Bond	US10YT
Bitcoin	CRYPT0000BTC

Source: FRED (2024a, 2024b, 2024c, 2024d, 2024e, 2024f, 2024g, 2024h), Reuters (2024).

The used dataset contains 9 time series with closing prices from the January 07, 2019 to the June 30, 2024 (Figure 3). Overall, I collected 1,388 business days for every financial instrument from Federal Reserve Bank of St. Louis (FRED): S&P500 (FRED, 2024a), NASDAQ 100 (FRED, 2024b), Nikkei 225 (FRED, 2024c), Gold (FRED, 2024d), Brent Crude Oil (FRED, 2024e), EUR-USD exchange rate (FRED, 2024f), 10-year Treasury Bond rate (FRED, 2024g) and Bitcoin (FRED, 2024h); and data for the Hang Seng stock index from Reuters database for historical data (Reuter, 2024) and calculated log-returns for the analysis. I computed summary statistics for daily log-returns of each financial instrument for the entire dataset and selected periods: COVID-19, the Russian attack and the Hamas terrorist attack on Israel (Tables 8-11). Furthermore, I collected historical data for the CBOE Volatility Index (VIX) (FRED, 2024i) and the Geopolitical Risk Index (2024), who was constructed by Caldara & Iacoviello (2022). Looking over the entire dataset, I found heavy daily losses: Bitcoin (-46.5%), 10-year Treasury Bond (-32.4%), Brent Crude Oil (-28.0%), NASDAQ (-13.1%) and S&P 500 (-12.8%). But there were also profits. The maximum daily gains are led by the 10-year Treasury Bonds (+36.8%), followed by Brent Crude Oil (+27.4%) and Bitcoin: (+20.3%). I compared the periods of geopolitical shocks separately and found the following. The heaviest daily losses occurred during COVID-19 with average of daily losses across all financial instruments by -16.8%. It was -5.6% for the period of the Russian attack and -3.3% for the period of the Hamas terrorist attack on Israel. Bitcoin, Brent Crude Oil and Bonds suffered the most considerable daily losses in all three periods.

Comparing volatility using the standard deviation of daily returns between different periods, one can observe the following. The period of COVID-19 has the highest volatility, with an average standard deviation of daily returns of 4.6%. This is followed by the period of the

Russian attack with 2.2% and then the period of the Hamas attack with 1.4%. Bitcoin, bonds, and oil have been found to have the highest volatility in all three periods of analyzed geopolitical events. Bonds were 11.8% during COVID-19, Brent Crude oil was 7.7% during the Russian attack, and Bitcoin was 2.8% during the Hamas attack. In the next step, I checked whether I need to perform further transformations on the time series before the analysis.

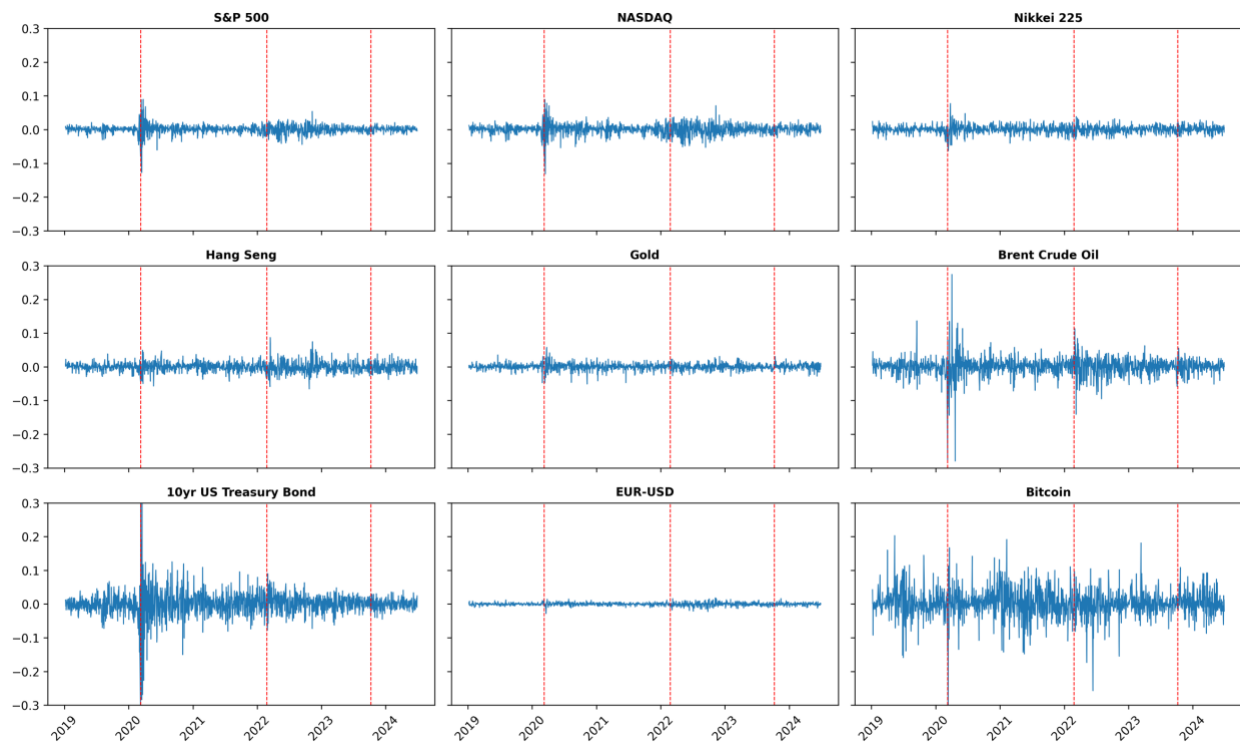


Figure 5: Daily log-returns for financial instruments in scope (01/07/2019 – 06/29/2024). Dashed lines mark geopolitical events: COVID-19 (03/09/20); Russian attack (02/24/22) and Hamas attack (10/08/23). Source: Own illustration using Python Matplotlib (2025).

This may be the case if a time series is not stationary. I tested this with the Augmented Dickey-Fuller test (Dickey & Fuller, 1979; Said & Dickey, 1984) and found that for all financial instruments, the null hypothesis, that a time series is not stationary, could be rejected (Table 12). In the last step, I wanted to test the feasibility of modeling volatility using GARCH models, i.e., whether the effects of heteroskedasticity can be observed in the time series. I use Engle's ARCH Lagrange Multiplier test (Engle, 1982). The null hypothesis assumes that there is no heteroscedasticity, whereby the parameters of an ARCH model are estimated using the data to be equal to zero. I could observe the following by applying Engle's ARCH test to the time series. The test could not be rejected for the EUR-USD exchange rate and Bitcoin but could be rejected for all other financial instruments (Table 13). This is an important indication that the current variance of the residuals of daily returns can be explained by the variance of the residuals of the past and provides the foundation for the further analysis. Furthermore, heteroscedasticity may also be present if the test is rejected. This may be because the time series under consideration has a high persistence in volatility. In this case, a GARCH model should be used for modeling. A second possibility is the presence of structural breaks that an ARCH model cannot sufficiently represent. In this case, the GJR-GARCH model shall be used.

Table 8: Summary statistics of daily log-returns for financial instruments in research scope (01/07/2019 - 06/30/2024).

	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Oil	10yr. T-Bond	EUR/USD	BTC	Average
Minimum	-0.128	-0.131	-0.063	-0.066	-0.051	-0.280	-0.324	-0.028	-0.465	-0,171
25% Quantile	-0.005	-0.006	-0.006	-0.008	-0.004	-0.010	-0.016	-0.002	-0.016	-0,008
50% Quantile	0.001	0.001	0.000	0.000	0.000	0.001	0.000	0.000	0.001	0,001
Mean	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.002	0,001
75% Quantile	0.007	0.009	0.007	0.007	0.006	0.013	0.017	0.002	0.021	0,010
Maximus	0.090	0.089	0.077	0.087	0.058	0.274	0.368	0.018	0.203	0,140
Std.	0.013	0.015	0.012	0.014	0.010	0.028	0.037	0.005	0.042	0,020
Skewness	-0.866	-0.664	0.023	0.247	-0.239	-0.789	0.052	-0.250	-1.033	-0,391
Kurtosis	15.522	7.847	3.586	2.943	4.157	23.166	22.777	2.814	13.863	10,742

Source: FRED (2024a, 2024b, 2024c, 2024d, 2024e, 2024f, 2024g, 2024h), Reuters (2024), own calculations using Matlab (2025).

Table 9: Summary statistics of daily log-returns for financial instruments in research scope in period of COVID-19 (02/01/2022 - 04/30/2022).

	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Oil	10yr. T-Bond	EUR/USD	BTC	Average
Minimum	-0.128	-0.131	-0.063	-0.050	-0.047	-0.280	-0.324	-0.028	-0.465	-0.168
25% Quantile	-0.020	-0.017	-0.015	-0.012	-0.007	-0.041	-0.056	-0.004	-0.010	-0.020
50% Quantile	-0.001	0.001	-0.004	0.000	0.001	-0.006	-0.012	0.000	0.000	-0.002
Mean	-0.002	-0.001	-0.003	-0.002	0.001	-0.017	-0.015	0.000	-0.003	-0.005
75% Quantile	0.014	0.016	0.009	0.009	0.007	0.009	0.034	0.003	0.024	0.014
Maximum	0.090	0.089	0.077	0.049	0.058	0.274	0.368	0.014	0.167	0.132
Std.	0.039	0.038	0.025	0.020	0.020	0.077	0.118	0.006	0.074	0.046
Skewness	-0.352	-0.543	0.492	-0.116	0.301	-0.252	0.380	-0.988	-3.944	-0.558
Kurtosis	1.727	2.045	1.682	0.684	1.620	5.831	2.867	5.266	25.251	5.219

Source: FRED (2024a, 2024b, 2024c, 2024d, 2024e, 2024f, 2024g, 2024h), Reuters (2024), own calculations using Matlab (2025).

Table 10: Summary statistics of daily log-returns for financial instruments in research scope in period around Russian invasion of Ukraine (02/01/2022 – 04/29/2022).

	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Oil	10yr. T-Bond	EUR/USD	BTC	Average
Minimum	-0.030	-0.040	-0.032	-0.059	-0.027	-0.141	-0.078	-0.018	-0.081	-0.056
25% Quantile	-0.012	-0.017	-0.010	-0.014	-0.007	-0.019	-0.014	-0.005	-0.023	-0.013
50% Quantile	-0.002	-0.002	0.000	-0.002	0.002	0.000	0.003	-0.001	0.006	0.001
Mean	-0.001	-0.001	0.000	-0.003	0.001	0.002	0.006	-0.001	0.001	0.000
75% Quantile	0.011	0.016	0.010	0.006	0.008	0.025	0.028	0.002	0.020	0.014
Maximum	0.025	0.037	0.039	0.087	0.023	0.113	0.090	0.016	0.111	0.060
Std.	0.014	0.020	0.015	0.023	0.011	0.040	0.035	0.006	0.038	0.022
Skewness	-0.105	0.102	0.153	0.993	-0.300	-0.405	-0.074	0.142	0.300	0.089
Kurtosis	-0.793	-1.004	-0.147	4.372	-0.116	2.273	-0.030	1.185	0.914	0.739

Source: FRED (2024a, 2024b, 2024c, 2024d, 2024e, 2024f, 2024g, 2024h), Reuters (2024), own calculations using Matlab (2025).

Table 11: Summary statistics of daily log-returns for financial instruments in research scope in period around Hamas attack (10/01/2023 - 12/30/2023).

	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Oil	10yr. T-Bond	EUR/USD	BTC	Average
Minimum	-0.015	-0.025	-0.023	-0.027	-0.023	-0.058	-0.045	-0.009	-0.068	-0.033
25% Quantile	-0.002	-0.002	-0.006	-0.010	-0.003	-0.017	-0.016	-0.002	-0.007	-0.007
50% Quantile	0.002	0.003	0.000	-0.001	0.001	-0.001	-0.001	0.000	0.005	0.001
Mean	0.002	0.002	0.000	-0.001	0.001	-0.002	-0.002	0.001	0.008	0.001
75% Quantile	0.006	0.007	0.006	0.008	0.006	0.017	0.011	0.003	0.020	0.009
Maximum	0.019	0.023	0.025	0.038	0.031	0.055	0.031	0.016	0.109	0.039
Std.	0.008	0.010	0.012	0.014	0.009	0.024	0.018	0.005	0.028	0.014
Skewness	-0.206	-0.497	0.114	0.337	0.356	-0.187	-0.217	0.464	0.676	0.093
Kurtosis	0.194	0.366	-0.337	-0.207	1.856	-0.133	-0.449	1.339	2.588	0.580

Source: FRED (2024a, 2024b, 2024c, 2024d, 2024e, 2024f, 2024g, 2024h), Reuters (2024), own calculations using Matlab (2025).

Table 12: Augmented Dickey-Fuller test for stationarity of daily log-returns for financial instruments in research scope (01/07/2019 – 06/30/2024).

	Augmented Dickey-Fuller Test	cValue	pValue	Reject H0
S&P 500	-44.43	-1.94	0.001	yes
NASDAQ	-43.03	-1.94	0.001	yes
Nikkei 225	-37.38	-1.94	0.001	yes
Hang Seng	-36.87	-1.94	0.001	yes
Gold	-37.84	-1.94	0.001	yes
Brent Oil	-35.08	-1.94	0.001	yes
10yr T-Bond	-40.36	-1.94	0.001	yes
EUR/USD	-36.77	-1.94	0.001	yes
BTC	-38.48	-1.94	0.001	yes

Source: FRED (2024a, 2024b, 2024c, 2024d, 2024e, 2024f, 2024g, 2024h), Reuters (2024), own calculations using Matlab (2025).

Table 13: Engle's test for heteroskedasticity of daily log-returns for financial instruments in research scope (01/07/2019 – 06/30/2024).

	Engle's Test Statistic	cValue	pValue	Reject H0
S&P 500	333.14	3.84	0.000	yes
NASDAQ	274.68	3.84	0.000	yes
Nikkei 225	230.83	3.84	0.000	yes
Hang Seng	113.92	3.84	0.000	yes
Gold	34.54	3.84	0.000	yes
Brent Oil	46.37	3.84	0.000	yes
10yr T-Bond	382.06	3.84	0.000	yes
EUR/USD	2.18	3.84	0.140	no
BTC	2.63	3.84	0.110	no

Source: FRED (2024a, 2024b, 2024c, 2024d, 2024e, 2024f, 2024g, 2024h), Reuters (2024), own calculations using Matlab (2025).

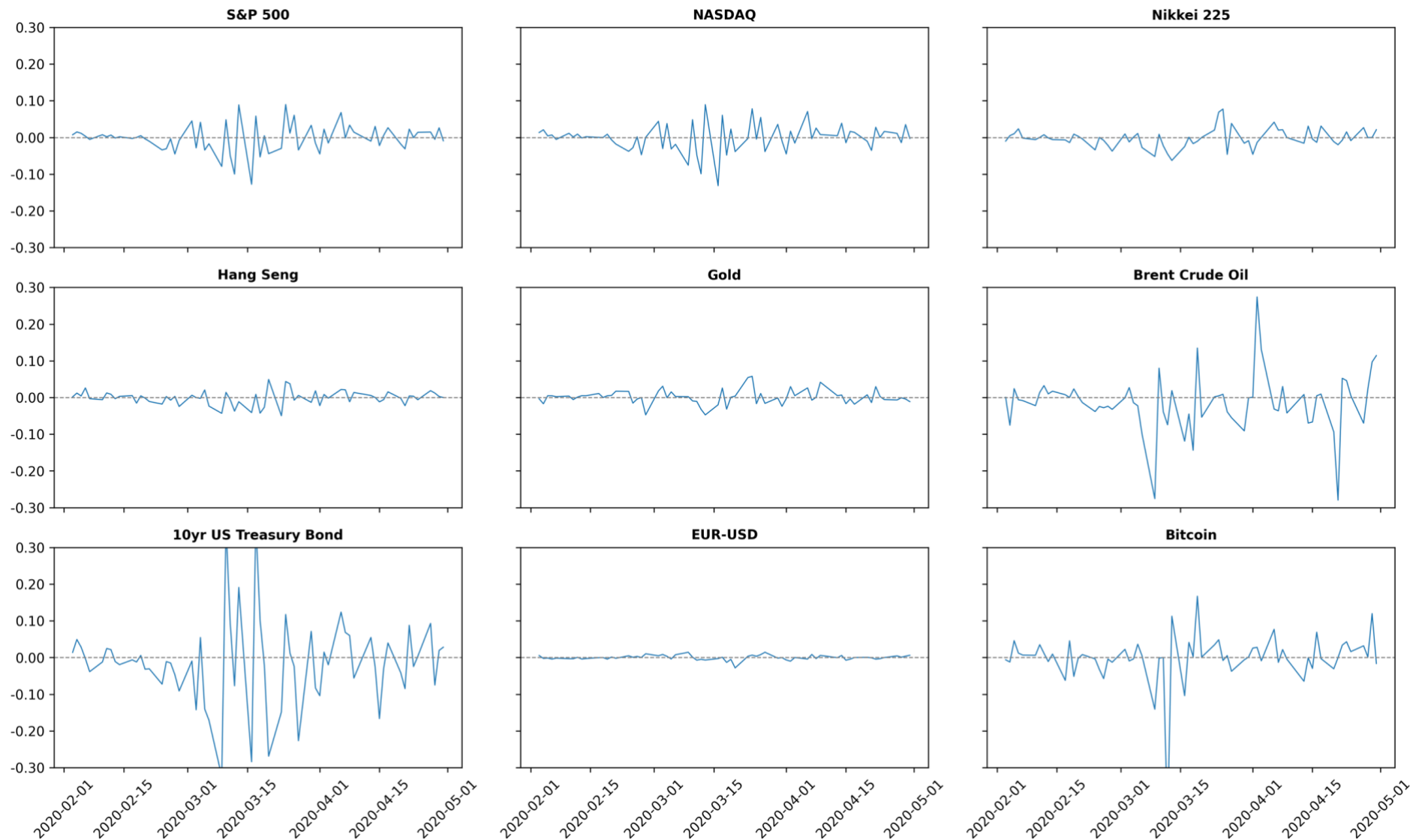


Figure 6: Daily log-returns for financial instruments in research scope in period of COVID-19 (02/03/2020 – 04/30/2020).

Source: Own illustration using Python Matplotlib (2025).

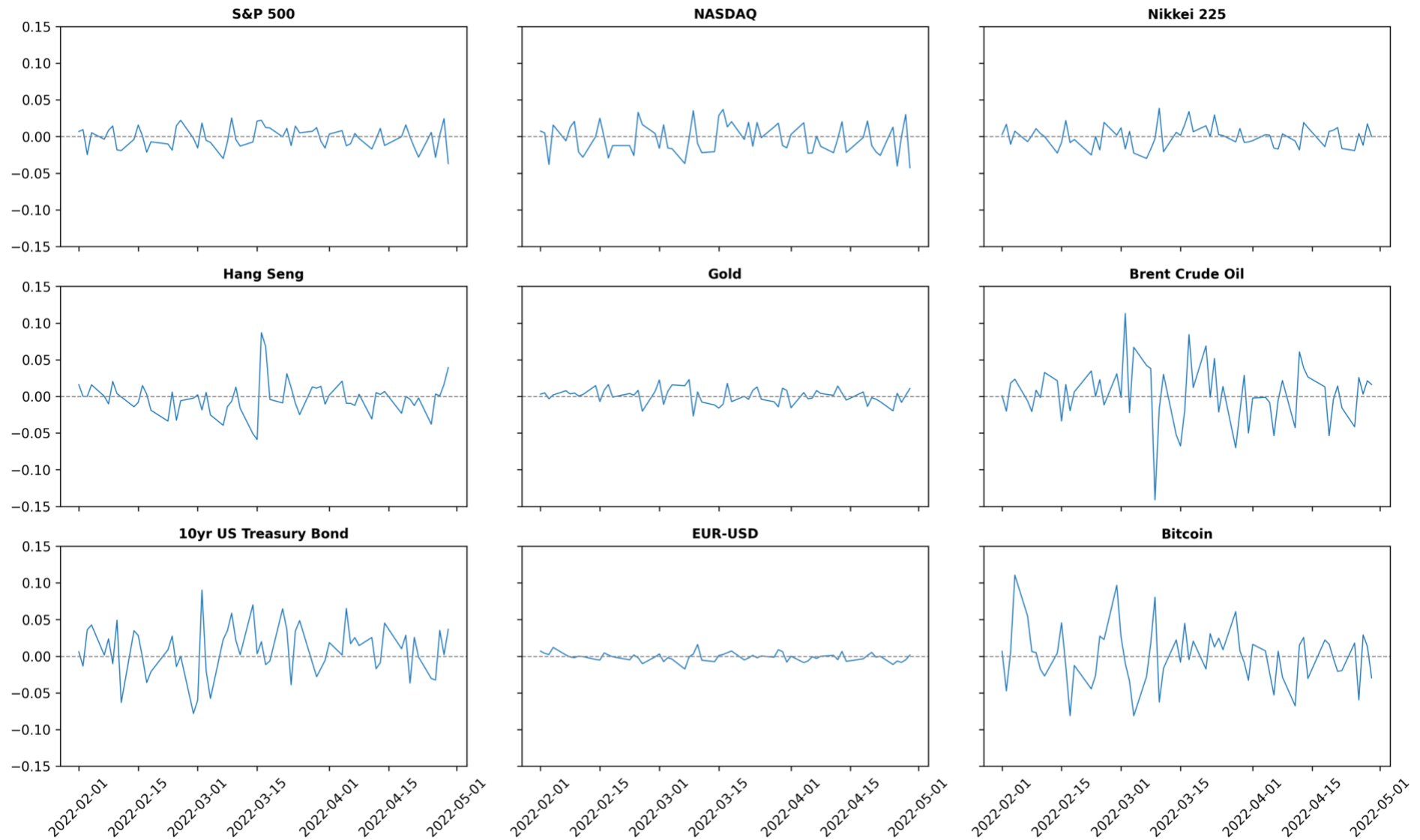


Figure 7: Daily log-returns for financial instruments in research scope around period of the Russian invasion of Ukraine (02/01/2022 – 04/29/2022).
Source: Own illustration using Python Matplotlib (2025).

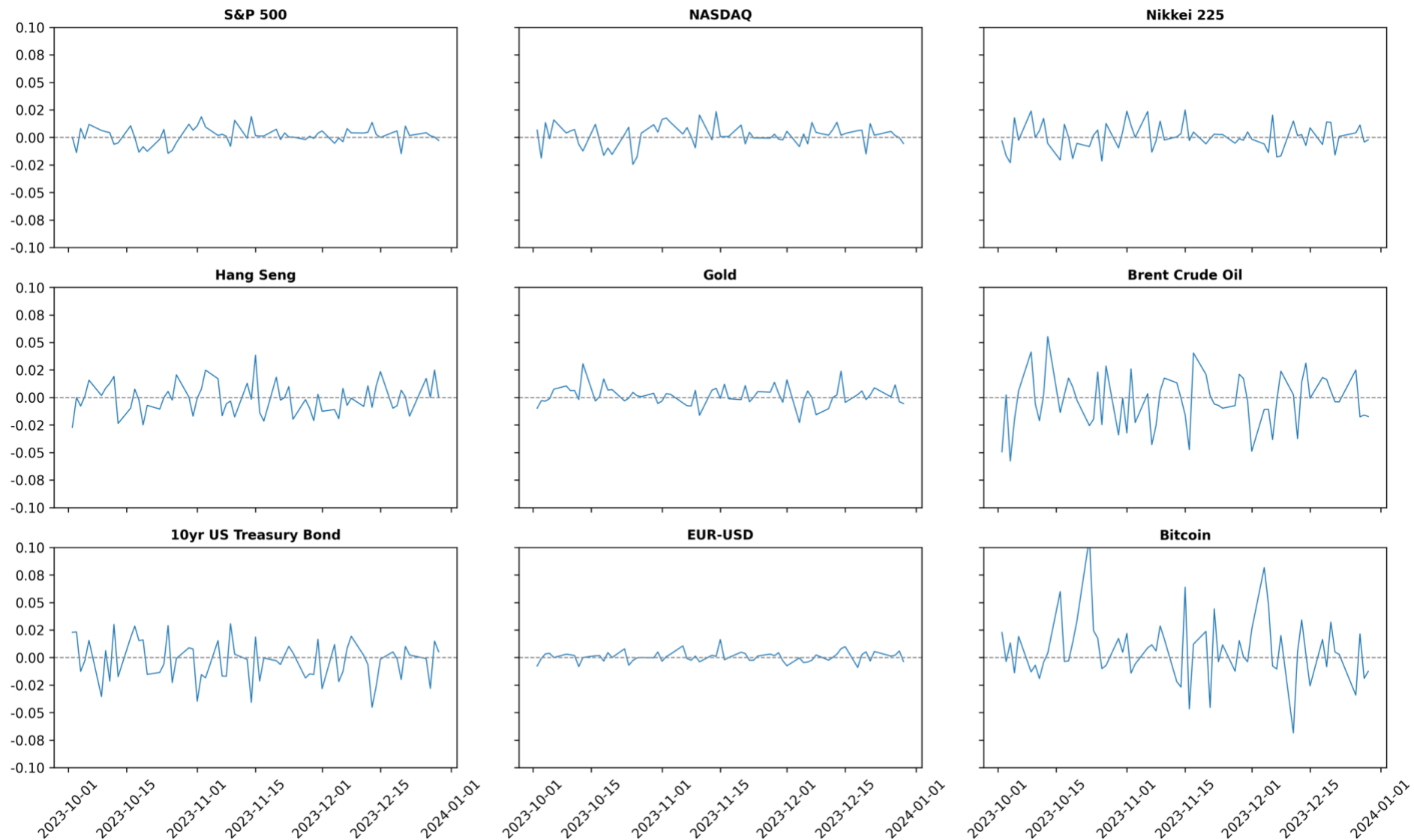


Figure 8: Daily log-returns for analyzed for financial instruments in research scope during period of the Hamas attack (10/02/2023 – 12/29/2023).
Source: Own illustration using Python Matplotlib (2025).

4. RESULTS

4.1. Abnormal Returns for MSCI World Sector Indices

Here I present the findings of abnormal returns (AR) of MSCI World sector indices surrounding the Russian invasion of Ukraine on February 24, 2022 in Table 14. The event window extends from $t-5$ to $t+5$ days to capture immediate market responses. To assess the statistical significance of these abnormal returns, I conducted t-tests, the results of which are also presented in Table 14. The data reveal a pronounced concentration of significant abnormal returns occurring on the event day and immediately thereafter. Specifically, on February 24, 2022, seven out of the eleven analyzed MSCI World sector indices exhibited statistically significant abnormal returns. The following day, February 25, 2022 ($t+1$), five indices showed significant effects, while the number of significant responses declined notably on subsequent days ($t+2$ to $t+4$). Interestingly, even prior to the invasion, a significant AR was observed in the materials sector. Beyond these focal dates, during the broader event window ($t-5$ to $t+5$), only two additional significant abnormal returns were recorded, specifically in the consumer staples and materials sectors.

Turning to the analysis of average abnormal returns (AAR), summarized in Table 15, fluctuations are evident in the days preceding the invasion, although none of these were statistically significant. On the event day itself, the market registered a significant negative AAR of -0.85% , highlighting the immediate adverse reaction to geopolitical developments. However, this loss was offset on the following day ($t+1$), which saw a significant positive AAR of $+1.32\%$. In the days that followed ($t+2$ to $t+5$), AARs remained minor and statistically insignificant, oscillating between gains and losses. Several factors may account for these short-term market dynamics. First, it is plausible that the initial sell-off reflected an overreaction to geopolitical uncertainty, a phenomenon often exacerbated by intense media coverage. Subsequently, more rational behavior prevailed, with investors taking advantage of perceived undervaluations. Moreover, given the significant U.S. weighting in the MSCI World and its sectoral indices, the post-event rally may have been driven by expectations that U.S.-based firms would benefit from a reconfiguration of global trade and investment flows, consistent with the findings of Ali et al. (2023). Comparable results have been documented in related studies. For instance, Ahmed et al. (2023) analyzed the STOXX Europe 600 Index around the same event and reported a decline of -0.41% on February 24, 2022. Although this movement mirrors the trend observed in the study, it is less pronounced, and their findings diverge more significantly in the pre- and post-event periods. Their data suggest sustained negative AARs leading up to the invasion and only modest recovery afterward. This contrast is likely due to the STOXX Europe 600's geographic proximity to the conflict, which may have heightened investor sensitivity compared to the globally diversified MSCI indices.

A sector-level analysis (Table 14) reveals heterogeneous responses across industries. The materials sector recorded the highest number of statistically significant abnormal returns, followed by the financials and utilities sectors. No significant ARs were observed in the consumer discretionary or energy sectors. On the event day, significant negative abnormal returns were found in consumer staples (-2.76%), financials (-3.01%), industrials (-0.87%), materials (-1.66%), and

information technology (−2.53%). Conversely, real estate (+1.66%) and telecommunications (+2.00%) registered significant positive abnormal returns. Among the remaining sectors, non-significant negative returns were observed in consumer discretionary (−0.14%), energy (−1.35%), healthcare (−0.36%), and utilities (−0.30%). These findings confirm the presence of sector-specific abnormal returns on the day of the invasion, thereby addressing the first research question.

A closer examination of the temporal patterns in abnormal and cumulative abnormal returns (CAR) provides further insights into how sectoral performance evolved beyond the initial shock (Table 16). For instance, the consumer discretionary sector experienced a significant decline even before the invasion ($CAR_{t-1} = -2.27\%$), followed by an insignificant increase on the event day ($AR_t = +0.63\%$) and a further decline by $t+25$ ($CAR_{t+25} = -2.94\%$). This sector comprises goods and services deemed non-essential, such as leisure, travel, and automotive products—areas likely to see demand contraction during crises due to changes in consumption priorities. The consumer staples sector, by contrast, exhibited a more resilient profile. While it experienced a sharp sell-off on the event day ($AR_t = -2.76\%$), this was immediately followed by a significant rebound ($AR_{t+1} = +2.07\%$), suggesting an overreaction had occurred. Over the longer term, the sector's CAR remained slightly negative ($CAR_{t+25} = -0.73\%$), reflecting the non-cyclical nature of demand for everyday essentials. The energy sector saw a modest positive trend in the pre-event period ($CAR_{t-1} = +1.64\%$), followed by an insignificant drop on the event day ($AR_t = -1.35\%$). Thereafter, it experienced substantial and significant cumulative gains ($CAR_{t+25} = +8.02\%$). This can be attributed to anticipated supply disruptions due to sanctions on Russian energy exports, as well as increased demand arising from military activity and reallocation toward western energy producers. The financial sector showed a neutral stance prior to the invasion ($CAR_{t-1} = +0.11\%$) but suffered a significant decline on the event day ($AR_t = -3.00\%$), continuing into the post-event period ($CAR_{t+25} = -5.50\%$). This negative trajectory likely reflects concerns over increased credit risk, loan defaults, and reduced lending activity, as well as broader economic uncertainty adversely affecting financial institutions. In contrast, the healthcare sector recorded a small decline on the event day ($AR_t = -0.36\%$) but posted significant positive CAR over the long term ($CAR_{t+25} = +4.12\%$). This trend may be linked to expectations of heightened demand for pharmaceuticals and medical services, as well as a general shift of capital toward defensive sectors. The real estate sector, which initially experienced investor uncertainty ($CAR_{t-1} = -3.19\%$), responded positively to the invasion ($AR_t = +1.65\%$) and continued to appreciate in value ($CAR_{t+25} = +4.74\%$). A breakdown of the sector reveals a strong representation of conservative sub-sectors such as industrial and telecommunications real estate, which may have benefited from capital reallocation away from riskier assets. Industrials remained relatively stable throughout the event window. While the sector exhibited minor setbacks, both before ($CAR_{t-1} = -0.53\%$) and on the day of the invasion ($AR_t = -0.87\%$), these were not significant, and long-term effects remained muted ($CAR_{t+25} = -0.29\%$). Despite some defense firms being included in this category, the sector's geographically diversified composition may have diluted potential gains. The information technology sector showed elevated pre-event volatility ($CAR_{t-1} = +0.32\%$) and suffered a significant drop on the event day ($AR_t = -1.68\%$). However, its long-term performance was

positive ($CAR_{t+25} = +2.44\%$), likely driven by expectations that technology firms—particularly those in software and semiconductors—would play a pivotal role in post-crisis economic adjustment. The materials sector gained in the days leading up to the invasion ($CAR_{t-1} = +1.64\%$) but recorded a significant event-day loss ($AR_t = -2.53\%$), which was likely part of a broader market sell-off. Nevertheless, the sector recovered strongly thereafter ($CAR_{t+25} = +3.02\%$), reflecting anticipated benefits from supply constraints and market consolidation in response to geopolitical realignments. Telecommunications, which includes large-cap firms such as Alphabet, Meta, and Netflix, also responded positively on the event day ($AR_t = +1.99\%$), despite modest pre-event losses ($CAR_{t-1} = -0.20\%$). These firms are often perceived as risky, yet their strong cash flows and low leverage likely insulated them from acute financial stress, contributing to the sector's subsequent resilience. The utilities sector displayed volatility before the invasion ($CAR_{t-1} = +0.64\%$) and experienced an insignificant decline on the event day ($AR_t = -0.30\%$). However, it emerged as one of the strongest performers over the longer term ($CAR_{t+25} = +7.29\%$). The sector includes firms involved in electricity generation and grid infrastructure, which were perceived as critical and stable in the face of potential energy shortages—paralleling the dynamics observed in the energy sector.

Overall, using the event study methodology, I have demonstrated that the Russian invasion had a clear and differentiated impact on sector-specific returns within the MSCI World Index. Notably, the energy, financials, healthcare, and utilities sectors exhibited significant cumulative abnormal returns by $t+25$, suggesting that market participants revised their expectations for these sectors in light of the geopolitical shock. These results confirm the second research question, indicating that the invasion had a sustained and heterogeneous influence across global equity markets.

Table 14: Abnormal returns around Russian invasion of Ukraine (in %) for MSCI World Sector Indices. T-test significance levels 10% (*), 5% (**) and 1% (***).

	CD	CS	EN	FN	HC	IN	IT	MT	RE	TC	UT
t-5	0.00%	1.41%*	1.04%	-0.36%	-0.15%	0.34%	0.88%	0.58%	1.14%	-0.71%	1.09%
t-4	0.12%	0.54%	-0.24%	0.26%	-0.41%	-0.33%	0.20%	-0.97%	-1.72%	0.06%	-0.07%
t-3	-0.11%	-0.11%	0.06%	0.02%	0.07%	-0.40%	-0.07%	-0.65%	-1.60%	0.23%	-0.29%
t-2	-1.44%	-0.33%	-0.04%	0.16%	0.48%	-0.06%	-0.09%	-0.43%	1.99%**	0.13%	0.32%
t-1	-0.84%	0.29%	1.64%	0.03%	0.50%	-0.09%	0.74%	-1.72%*	0.51%	0.09%	-0.41%
t	-0.14%	-2.76%***	-1.35%	-3.01%***	-0.36%	-0.87%**	-2.53%***	1.66%*	-1.69%*	2.00%***	-0.30%
t+1	3.04%	2.08%***	0.50%	0.70%	1.43%***	0.66%	1.05%	2.44%***	1.54%	-1.16%**	2.28%***
t+2	-0.09%	-0.69%	1.38%	-1.35%**	-0.24%	0.71%	0.48%	-1.55%	0.80%	0.17%	1.01%*
t+3	-1.92%	-0.14%	1.73%	-1.99%***	0.66%	-0.19%	0.29%	-0.42%	0.22%	0.55%	-1.19%**
t+4	1.37%	-0.40%	1.80%	0.32%	0.08%	0.05%	0.29%	1.88%**	-0.50%	-0.79%	-0.49%
t+5	-0.89%	0.34%	-0.13%	0.27%	0.36%	0.48%	1.22%	0.80%	0.13%	-0.12%	0.53%

Source: Robus et al. (2024).

Table 15: Average abnormal returns during Russian invasion of Ukraine (AAR %) for MSCI sector indices around event day t. T-test statistics with significance levels at 10% (*), 5% (**) and 1% (***).

	t-5	t-4	t-3	t-2	t-1	t	t+1	t+2	t+3	t+4	t+5
AAR	0.29%	-0.34%	-0.28%	0.10%	-0.01%	-0.85%**	1.32%***	0.06%	-0.22%	0.33%	0.27%
t-test	0.81	-0.94	-0.77	0.27	-0.02	-2.37	3.68	0.15	-0.61	0.91	0.76

Source: Robus et al. (2024).

Table 16: Cumulative abnormal returns (%) during Russian invasion of Ukraine for MSCI World Sector Indices around event day. T-test statistics with significance levels at 10% (*), 5% (**) and 1% (***).

	CD	CS	EN	FN	HC	IN	IT	MT	RE	TC	UT
t-5	0.00%	1.41%	1.04%	-0.36%	-0.15%	0.34%	0.88%	0.58%	1.14%	-0.71%	1.09%
t-4	0.11%	1.95%	0.80%	-0.10%	-0.56%	0.01%	1.07%	-0.39%	-0.58%	-0.65%	1.02%
t-3	0.00%	1.84%	0.86%	-0.08%	-0.49%	-0.39%	1.00%	-1.04%	-2.18%	-0.42%	0.73%
t-2	-1.43%	1.51%	0.82%	0.08%	-0.01%	-0.45%	0.91%	-1.47%	-0.19%	-0.29%	1.05%
t-1	-2.27%	1.79%	2.46%	0.11%	0.49%	-0.54%	1.64%	-3.19%	0.32%	-0.20%	0.64%
t	-1.65%	-0.97%	1.11%	-2.89%	0.13%	-1.40%	-0.89%	-1.53%	-1.37%	1.79%	0.34%
t+1	-2.44%	1.11%	1.61%	-2.19%	1.56%	-0.75%	0.16%	0.90%	0.18%	0.63%	2.62%
t+2	-1.86%	0.42%	2.99%	-3.55%	1.31%	-0.04%	0.64%	-0.65%	0.98%	0.80%	3.63%**
t+3	-1.88%	0.28%	4.72%	-5.54%***	1.97%	-0.23%	0.93%	-1.07%	1.19%	1.35%	2.44%
t+4	-2.37%	-0.12%	6.51%	-5.22%***	2.06%	-0.18%	1.21%	0.81%	0.69%	0.56%	1.94%
t+5	-3.75%	0.21%	6.39%	-4.95%***	2.42%	0.30%	2.44%	1.61%	0.82%	0.44%	2.48%
t+10	-3.81%	-2.93%**	11.79%***	-5.69%***	2.49%	0.83%	2.28%	1.25%	-2.21%	0.51%	5.26%***
t+15	-3.31%	-1.85%	5.87%	-3.16%*	4.46%***	1.91%	1.58%	2.87%	-2.21%	-1.10%	4.79%***
t+20	-3.27%	-2.36%	8.79%**	-4.22%**	3.25%**	0.46%	3.03%	1.45%	0.00%	-0.53%	3.71%**
t+25	-2.94%	-0.74%	8.03%**	-5.50%***	4.13%***	-0.29%	3.03%	4.74%	2.44%	-1.38%	7.30%***

Source: Robus et al. (2024).

4.2. Identification of Abnormal Volatility

4.2.1. Abnormal Volatility

This section presents my findings on the presence of abnormal volatility surrounding the Russian invasion of Ukraine in February 2022 (Robus et al., in press). The results are summarized in Table 17. I begin with an aggregated analysis by examining the sector-level mean abnormal volatility across the event window ($t-5$ to $t+5$). To obtain robust estimates that are not unduly influenced by extreme values, I computed the truncated mean (TMean), which excludes the minimum and maximum observations for each time point Robus et al. (in press). This approach mitigates the distortionary effect of outlier volatility spikes in individual sectors, thereby providing a more representative measure of central tendency. The abnormal volatility values (AVOLA) are expressed as multiplicative factors relative to volatility forecasts generated by a GARCH(1,1) model. For instance, an AVOLA value of 1.99 implies that realized volatility was 1.99 times—or 199%—greater than the model-based forecast derived from pre-event data. To assess the statistical significance of these deviations, I performed F-tests for each time point. The results show that, for most days within the event window, the sector-level AVOLA-TMean values exceed unity, indicating that realized volatility was systematically higher than predicted by the GARCH(1,1) baseline. Notably, elevated AVOLA values prior to the event—specifically on $t-5$, $t-2$, and $t-1$ —suggest heightened market uncertainty in anticipation of a potential military conflict. The pronounced increase in abnormal volatility on $t-2$ and $t-1$, following relatively subdued values on $t-4$ and $t-3$, likely reflects geopolitical developments that intensified investor concerns. In particular, on February 21, 2022, the Russian government officially recognized the independence of pro-Russian separatist-held territories in Donetsk and Luhansk Oblasts—territories proclaimed as the “People’s Republics”—and subsequently deployed military forces to these regions under the pretext of protection (Bocquillon et al., 2024; Sakwa, 2022).

On the event day t itself, the AVOLA-TMean reached an exceptionally high value of 8.99, indicating that realized volatility was nearly nine times greater than the GARCH(1,1) forecast. This heightened volatility persisted into the following day ($t+1$), with an AVOLA-TMean of 8.50. Both values are statistically significant, underscoring the extraordinary magnitude of market reactions across sectors in response to the invasion. It is important to note that abnormal volatility reflects the intensity of price fluctuations, irrespective of direction (Robus et al., in press). Therefore, these findings denote the scale of market turbulence, rather than uniformly negative returns, as an immediate reaction to the geopolitical shock. Robus et al., (2024), showed that there are abnormally negative returns for most sectors on the event day. Next day, however, one can see that there are many positive abnormal returns. Here one can see the corresponding fluctuations. These price fluctuations are a reaction of market participants to the changed geopolitical conditions and their expectations regarding their respective sector investment. This results in the liquidation of share portfolios as well as shifts from shares in

one sector to shares in another sector, considering the individual expectations of the respective business model. For example, in times of crisis one often see stocks in the IT sector being sold off and stocks of the consumer staples sector being bought up (e.g., Prasad et al., 2021). After the event day, at $t+2$ up to $t+5$ one see AVOLA-TMean values are less or close to one which means that the observed volatility is smaller or corresponds to the forecast volatility and that the market moves less than forecast, values less than one ($t+2$; $t+5$). Figure 1 visually underscores the pronounced spike in truncated mean abnormal volatility (AVOLA-TMean) on both the event day (t) and the following day ($t+1$). This observation reinforces my earlier findings and provides additional support for the presence of abnormal volatility around February 24, 2022, thus empirically validating Research Question 3 from an aggregated, cross-sectoral perspective (Robus et al., in press).

I now turn to the sector-specific results. On the event day, abnormal volatility (defined as $AVOLA > 1$) was observed in 8 out of the 11 MSCI World sector indices. The most pronounced deviations occurred in the financials ($AVOLA = 57.46$) and consumer staples ($AVOLA = 56.46$) sectors, followed by materials ($AVOLA = 6.28$), utilities ($AVOLA = 4.91$), telecommunications ($AVOLA = 2.70$), information technology ($AVOLA = 1.83$), industrials ($AVOLA = 1.67$), and real estate ($AVOLA = 1.45$). Among these, the financials, consumer staples, and materials sectors exhibited statistically significant deviations, indicating a robust market response to the geopolitical event. Interestingly, the reaction of market participants appears to have intensified on the day following the invasion ($t+1$), suggesting a delayed adjustment process and further underscoring the magnitude of the shock across global equity markets. Here, all sectors have an $AVOLA > 1$. These are statistically significant for 8 sectors: utilities (48.90), health care (26.72), consumer staples (19.59), industrials (15.39), IT (9.50), financials (8.15), materials (5.95) and real estate (3.12) (Robus et al., in press).

Possible explanations for the behavior of the individual sectors around the event day t are as follows. consumer discretionary, which represent companies for e.g., cars, household appliances, specialty goods, luxury goods and leisure goods, would be affected, as consumers tend to postpone major consumer spending in times of war, which has a negative impact on companies' sales. Uncertainty at consumer staples sector arose by the sanctions imposed on Russia, where the loss of the Russian market is now expected to result in a major decline in sales, particularly for consumer goods in the mass market (e.g., Coca-Cola, Nestlé and Unilever). The energy sector showed relatively small abnormal volatility over the analyzed period especially at t and $t+1$, which indicates a low level of uncertainty overall. This is possibly because investments in energy companies are less speculative and long-term oriented. Another possible explanation is, that market participants feel little uncertainty about the business model of companies in the energy sector and that no significant changes are expected as a result of the war. Also, it could be that this uncertainty had already arisen before the observation $t-5$ and any changes were already reflected in the share price. Regarding the behavior of the energy sector, I found comparable results to the analysis of abnormal returns in Robus et al. (2024). Financials sector include global operating banks and insurance companies to a large extent. During times

of war, economic losses are anticipated, often resulting in loan defaults and increased claims for insurance payouts. Additionally, the heightened uncertainty typically causes a decline in investment activity, adversely affecting banks commission-based revenue streams. Declining investment is also leading to a fall in demand for capital goods, which is having a negative impact on the industrial sector. The health care sector comprises companies involved in delivering medical services, manufacturing medical equipment or pharmaceuticals, offering medical insurance, or supporting healthcare provision to patients in various capacities. In particular, the expected increase in demand for products and services from companies in the health care sector during a realized conflict is a plausible explanation for the reaction of the sector index (Robus et al., in press).

The industrials sector covers a broad range of business models, including aerospace & defense, electronic equipment, infrastructure building, machinery and transportation. Abnormal volatility around the event day could be attributed to the relatively lower country weighting of U.S. companies in the industrials sector compared to other sectors. Potential share price declines of companies based outside the U.S., such as those in France and the UK, may also have contributed to this volatility (Robus et al., 2024). Companies in the IT sector are classically more speculative than companies in other sectors, as they are often young companies with high potential, but the return on sales is small in comparison. In addition, they often have a high debt ratio, which makes them susceptible to interest rate increases. The Russian attack increases the cost risk and profitability for these companies. As a result, investors often adjust their portfolios towards more conservative business models in times of crisis. The materials sector comprises companies involved in the production of chemical products, metal processing, industrial gases, and the mining of gold and other metals (Assefa et al., 2022). Potential opportunities may emerge from the exclusion of Russian companies, leading to supply shortages that could enhance the market position and profitability of the companies included in the index. real estates are very conservative investments but also affected by interest rate risk. Inflation in the USA, which began in 2021 after the COVID-19 pandemic had begun to subside, combined with the Russian attack, increased uncertainty in this sector. The utilities sector represents companies that supply electric power, natural gas, steam supply, water supply and sewage removal. Significant disruptions in the natural gas supply chain are to be expected, particularly as the onset of war places the long-standing energy partnership between Europe and Russia under considerable strain. This geopolitical realignment simultaneously creates strategic opportunities for alternative suppliers, most notably U.S.-based exporters of liquefied natural gas. In line with this context, the analysis revealed substantial abnormal volatility across sectors on both the event day and the subsequent trading day. However, the degree of sectoral response was heterogeneous. Notably, certain sectors—such as Energy—did not exhibit statistically significant abnormal volatility across the event window, suggesting a more muted or stabilized market reaction in light of pre-existing expectations or longer-term structural positioning within the sector. Sectors with strong abnormal volatility before the event: financials, health care and telecommunications. Sectors with abnormal volatility after the event day: e.g., consumer

discretionary and industrials. Another case of sector behavior is, that there is several abnormal volatility ($AVOLA > 1$) before the event day. These findings are consistent to the findings Prasad et al., (2021) which found evidence for stock market sector abnormal volatility during the dot-com boom, the global financial crisis and COVID-19.

Table 17: Abnormal volatility (AVOLA) for MSCI Sector Indices around Russian invasion of Ukraine (02/17/22 – 03/03/22). Statistical significance is indicated by * using an F-Test with 95% level.

	CD	CS	EN	FN	HC	IN
t-5	2.53	2.76	0.09	45.77*	3.19*	2.42
t-4	0.46	0.15	0.48	0.65	1.78	2.03
t-3	0.11	0.34	0.01	0.29	0.02	0.67
t-2	4.44*	3.80*	0.32	7.05*	0.02	1.34
t-1	3.48*	0.70	0.10	24.13*	0.37	3.30*
t	0.11	56.46*	1.04	57.46*	0.51	1.67
t+1	2.74	19.59*	3.60	8.15*	26.72*	15.39*
t+2	0.11	0.53	0.85	1.60	0.11	0.44
t+3	2.10	0.80	0.06	18.60*	0.20	3.10**
t+4	0.41	0.01	3.83*	1.24	1.04	1.27
t+5	3.00*	0.00	0.29	0.23	0.02	0.04
	IT	MT	RE	TC	UT	TMean
t-5	0.31	0.29	0.26	4.18*	0.22	1.99
t-4	4.85*	0.26	0.78	0.40	0.77	0.86
t-3	1.51	0.08	0.31	0.00	0.62	0.23
t-2	0.41	0.68	0.15	0.52	0.05	1.46
t-1	0.49	0.18	2.56	1.69	4.91*	1.60
t	1.83	6.28*	1.45	2.70	0.47	8.99*
t+1	9.50*	5.95*	3.12*	2.06	48.90*	8.50*
t+2	0.11	0.08	0.88	0.00	1.13	0.39
t+3	0.63	1.03	0.10	1.34	4.82	1.16
t+4	0.17	1.34	2.91	0.22	0.00	1.07
t+5	0.17	0.27	0.31	0.94	0.04	0.28

Source: Robus et al. (in press).

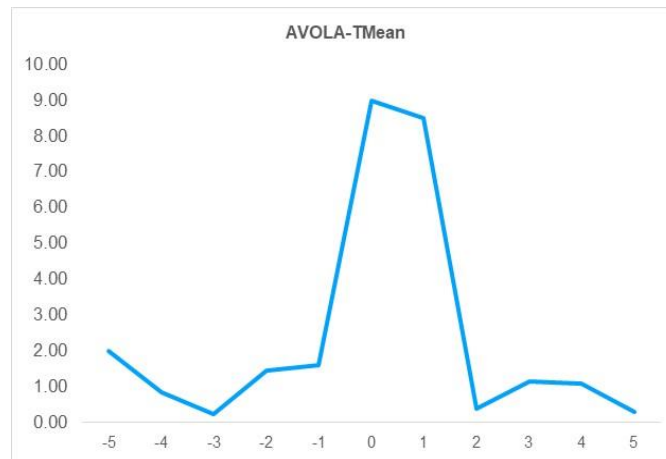


Figure 9: Trimmed mean for abnormal volatility (%) of MSCI Sector Indices around Russian invasion of Ukraine (02/17/22 – 03/03/22).
Source: Robus et al. (in press).

4.2.2. Cumulated Abnormal Volatility

I now want to discuss the second question of this paper and examine whether the abnormal volatility remains persistent over a longer period after the event day. I do this by analyzing the cumulated abnormal volatility (CAVOLA), which is defined by equation (10). The idea here is that if the ratio of the sum of actual volatility and the sum of volatility predicted by GARCH(1,1) increases, abnormal volatility > 1 has occurred over the period. If this ratio does not change, the actual volatility corresponds to the forecast volatility. If the ratio is even lower, this means that the actual volatility over the period under review is lower than the forecast volatility. I start again with the analysis of the mean value. For reasons of consistency, I also used the TMean here. The results are presented in Table 18 and illustrated in Figure 10. I have calculated cumulative abnormal volatility for $t-1$ (sum of AVOLA of the observations $t-5$ to $t-1$) as an initial value. This gives us also a benchmark for the abnormal volatility before the event day (Robus et al., in press).

I was able to observe the following for the CAVOLA-TMean. Starting from a value of 2.42 in $t-1$, it increases to 3.04 up to $t+5$. This is understandable, as although there are some sectors with increased abnormal volatility up to $t-1$, the majority can be observed in t and $t+1$, as shown in the previous analysis. What is surprising now is that the CAVOLA-TMean increases further to 3.24 by $t+10$ compared to $t+5$. This means that further abnormal volatility occurs in the time periods $t+6$ to $t+10$ and that there continue to be strong price fluctuations even days after the event day. In other words, the abnormal volatility is 20% higher than that around the event day. This also implies that there was a longer-term uncertainty regarding further developments among investors. From $t+11$ to $t+20$ one now see a decreasing values for CAVOLA-TMean ($t+15$: 3.01 and $t+20$: $t+20$). This suggests a calming of market participants, whose actions on the markets produce less volatility than the GARCH(1,1) model predicts. In the further observation period, one see that the CAVOLA-TMean of $t+20$ roughly corresponds to the CAVOLA of $t+25$. This means that the CAVOLA-TMean added from $t+21$ to $t+25$ is

approximately one, which is the value where the actual volatility corresponds to the estimated volatility and has therefore returned to a long-term level. From a mean value perspective, I was thus able to show that the abnormal volatility was persistent up to the period $t+10$ and thus lasted longer than shortly after the event day. I will now look at the behavior of individual sectors. I observed the CAVOLA-TMean behavior just described for the following sectors: consumer staples, energy, health care, industrials, IT, materials and telecommunications (Robus et al., in press).

However, I also found different behavior in individual sectors. The persistence was particularly strong in the sectors: IT, telecommunications, health care, industrials and materials. In Financials, an already high CAVOLA value was observed at time $t-1$, followed by a sharp decline up to time $t+5$, after which CAVOLA rose again, as in other sectors. The zig-zag movement here up to $t+25$, suggests that the uncertainty in this sector will persist for longer or increase again after a certain time.

Table 18: Cumulated Abnormal volatility (CAVOLA) for MSCI sector indices from 02/17/22 – 03/31/22.

	CD	CS	EN	FN	HC	IN
t-1	2.14	1.50	0.20	7.63	1.11	1.93
t+5	1.67	4.16	0.97	6.66	2.24	2.56
t+10	2.12	3.96	1.08	6.90	2.61	2.79
t+15	1.95	2.97	1.37	6.72	2.52	2.50
t+20	1.62	2.45	1.39	6.43	2.13	2.04
t+25	1.58	2.27	1.39	6.62	1.97	2.00
	IT	MT	RE	TC	UT	TMean
t-1	1.30	0.30	0.77	1.29	1.17	2.42
t+5	1.37	1.57	1.13	1.25	3.56	3.04
t+10	1.75	1.71	1.07	1.93	2.46	3.24
t+15	1.44	1.67	1.01	1.85	2.11	3.01
t+20	1.27	1.44	0.90	1.65	1.94	2.68
t+25	1.23	1.28	0.99	1.60	2.01	2.64

Source: Robus et al. (in press).

One can see little persistence of abnormal volatility due to declining CAVOLA after $t+5$ at the real estate and utilities sector. My analysis showed that abnormal volatility was persistent in many sectors until $t+10$. With these findings, I can confirm research question 4.

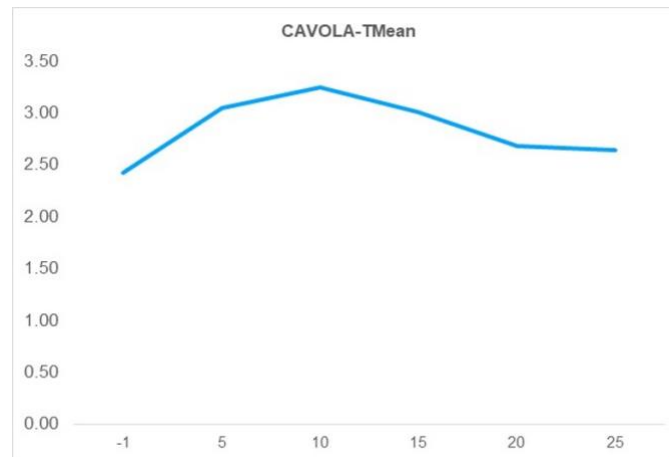


Figure 10: Trimmed mean cumulated abnormal volatility (CAVOLA-TMean) for MSCI Sector Indices around Russian invasion of Ukraine (02/17/22 – 03/31/22). Source: Robus et al. (in press).

4.2.3. Statistically significant differences of Abnormal Volatility

For the fifth research question, whether statistically significant differences in abnormal volatility between the sectors were observed. I carried out this analysis for the most prominent days, the event day t and the following day $t+1$. To evaluate statistical significance, I performed an F-test for all pairs of sectors, for t (Table 19) and $t+1$ (Table 21) respectively. For example, the F-test regarding the equality of variances between consumer discretionary and consumer staples is equal to 505.71, was rejected at a 95% level and is thus statistically significant. As one can see, there are more significant differences in abnormal volatility between the sectors on the event day than on the following day. This was already seen in the previous analysis, where I observed various sector behavior on the day of the event. Now, I want to quantify a little more. Therefore, I counted the number of statistically significant differences in abnormal volatility between the sectors. and ordered them (Table 20 and Table 22) (Robus et al., in press).

Further, I defined sectors with a frequency of up to 5 significant differences in abnormal volatility as similar regarding abnormal volatility to other sectors. Sectors with a higher frequency are dissimilar to other sectors. As one can see here, at event day t , 8 out of 11 sectors are dissimilar in terms of abnormal volatility according to the definition. This is also reflected in the average value of 6.82 (Table 20). In contrast, only the three sectors real estate, telecommunications and energy are similar. A different picture emerges for the following day of the event. I only observed two dissimilar sectors here: utilities and telecommunications. The remaining sectors are similar in terms of abnormal volatility on the day after the event. This is shown by the average value of 4.55 (Table 22). Using this approach, I demonstrated that statistically significant differences in abnormal volatility existed between sectors on the event day and the day after (Robus et al., in press).

My results provide new insights into the volatility dynamics of the global financial markets. For researchers, this study is insightful as there is little evidence on the volatility dynamics of current geopolitical shocks at the sector level of financial markets. In particular, the short-term behavior of volatility around the geopolitical event of the Russian attack has

received little attention in the literature. The specific detailed view on the sectors and thus on different business models is particularly interesting. The shown results can also be compared with those of past events, e.g., the COVID-19 pandemic. This can generate new insights into the behavior of capital markets in crises. It is also much more feasible to compare the impact of geopolitical events. For practitioners, such as fund managers, investment advisors or investors, I provided important insights into how equities behave around geopolitical shocks. In particular, my insights into the behavior of individual sectors can be used to assess the risk of potential market price fluctuations. This can be used to categorize individual stocks into different risk classes with regard to geopolitical shocks and accordingly serve as a basis for the weighting of the portfolio or as an indication of the scale of hedging positions (Robus et al., in press).

Table 19: F-Test statistics for differences between MSCI sector indices abnormal volatility at event day February 24, 2022 (* indicates statistical significance on a 95% level).

	CD	CS	EN	FN	HC	IN	IT	MT	RE	TC	UT
CD	1.00	505.71*	9.34*	514.69*	4.57*	14.93*	16.37*	56.09*	12.98*	24.14*	4.22*
CS		1.00	54.13*	1.02	110.67*	33.88*	30.90*	9.02*	38.97*	20.95*	119.81*
EN			1.00	55.09*	2.04	1.60	1.75	6.00*	1.39	2.58	2.21
FN				1.00	112.64*	34.48*	31.45*	9.18*	39.66*	21.32*	121.93*
HC					1.00	3.27*	3.58*	12.27*	2.84	5.28*	1.08
IN						1.00	1.10	3.76*	1.15	1.62	3.54*
IT							1.00	3.43*	1.26	1.47	3.88*
MT								1.00	4.32*	2.32	13.29*
RE									1.00	1.86	3.07*
TC										1.00	5.72*
UT											1.00

Source: Robus et al. (in press).

Table 20: Frequencies (f) of rejected F-Tests regarding similar abnormal volatility at event day (February 24, 2022) between MSCI sector indices.

	CD	FN	MT	CS	UT	HC	IN	IT	RE	TC	EN	Avg.
f	10	9	9	8	8	7	6	6	5	5	2	6.82

Source: Robus et al. (in press).

The results also serve as a possible benchmark for the volatility dynamics of other sector indices, e.g., for emerging markets or other regional markets. This in turn provides information about possible excess volatility, which must then be classified in relation to the resulting returns (Robus et al., in press).

Table 21: F-Test statistics for differences between MSCI sector indices Abnormal volatility on February 25, 2022 (* indicates statistical significance on a 95% level).

	CD	CS	EN	FN	HC	IN	IT	MT	RE	TC	UT
CD	1.00	7.16*	1.31	2.98	9.76*	5.62*	3.47*	2.17	1.14	1.33	17.87*
CS		1.00	5.45*	2.40	1.36	1.27	2.06	3.29*	6.28*	9.53*	2.50
EN			1.00	2.27	7.43	4.28	2.64	1.66	1.15	1.75	13.60*
FN				1.00	3.28*	1.89	1.17	1.37	2.61	3.96*	6.00*
HC					1.00	1.74	2.81	4.49*	8.57*	13.00*	1.83
IN						1.00	1.62	2.59	4.93*	7.48*	3.18*
IT							1.00	1.60	3.04*	4.62*	5.15*
MT								1.00	1.91	2.89	8.22*
RE									1.00	1.52	15.68*
TC										1.00	23.78*
UT											1.00

Source: Robus et al. (in press).

Table 22: Frequencies (f) of rejected F-Tests regarding similar abnormal volatility on February 25, 2022 between MSCI sector indices.

	UT	TC	CD	CS	HC	RE	IN	IT	FN	MT	EN	Avg.
f	8	6	5	5	5	5	4	4	3	3	2	4.55

Source: Robus et al. (in press).

4.3. Univariate Volatility Modeling and Forecasting of Financial Time Series during Geopolitical Shocks

4.3.1. In-Sample Analysis

This study aimed to assess the effectiveness of various GARCH models in fitting financial data during periods of significant geopolitical events. The analysis was divided into two main parts. In the first part, I focused on how well these models captured the volatility of

financial instruments during three specific periods: the COVID-19 pandemic (63 business days), the period around the Russian attack on Ukraine (62 business days), and the time of the Hamas attack (64 business days). From a total of 75 different GARCH model configurations, I estimated 25 GARCH models, 25 EGARCH models, and 25 GJR-GARCH models over these three periods. I evaluated the performance of each model by comparing the goodness of fit for a single financial instrument. The selection of the best-fitting model in a class was based on the Akaike Information Criterion (AIC).

Here, I present the best model from the three GARCH classes for every financial instrument and period. In addition to comparing the AIC value, I also calculated the AIC weights, e.g., Akaike (1978), Akaike (1979), and Bozdogan (1987). The AIC weights weigh the differences in the AICs between the models to be compared and thus provide information on the likelihood of a model being the best (Burnham et al., 2011; Wagenmakers & Farrell, 2004). I present the results in Table 23. These results show that the EGARCH model consistently outperforms the standard GARCH and the GJR-GARCH models across all financial instruments and periods in terms of the lowest AIC value. Specifically, the Akaike weights showed that the EGARCH model was, in almost all cases, more than 90% likely to provide the best fit than the second-best model in most cases. One explanation for this superiority of the EGARCH model can be found in the design purpose of the model. It was developed to reflect the so-called leverage effect better than the standard GARCH model. As previously mentioned, the leverage effect is when volatility increases when the price falls, i.e., negative returns are realized, so there is an asymmetric return-volatility relationship (Black, 1976; Campbell & Hentschel, 1992; French et al., 1987).

The three geopolitical events I analyzed are characterized by multiple sharp daily losses across all financial instruments, especially during COVID-19 (Table 3 and Table 4). In the COVID-19 phase analyzed, one can see this effect particularly well in the S&P 500, NASDAQ, Brent Crude Oil, and 10-year Treasury bonds. Here, a sharp daily loss is followed by a phase of higher volatility than previously. The EGARCH model can also describe high levels of volatility very well (Nelson, 1991). One can recognize this volatility level increase visually in the time series plots for the different periods and as a standard deviation in the summary statistics. The S&P 500 index's standard deviation for the whole dataset is 0.013. During the COVID-19 period, it was 0.039 and thus significantly increased. One can observe this for almost all financial instruments I analyzed. The effect is most vital for COVID-19, followed by the Russian attack. For the period of the Hamas attack, the effect is hardly recognizable. Moreover, the best-fitting EGARCH models tended to have a higher number of lags. This suggests that the capability of the higher-order EGARCH model to capture asymmetric effects, incorporation of historic volatility shocks, and volatility persistence is better suited to the complexities observed during periods of heightened geopolitical instability.

By showing that the EGARCH model achieved the best in-sample fit for the financial instruments in scope and periods of geopolitical shocks, I was able to answer my research question (Research Question 6). The literature supports my findings. Gharaibeh & Kharabsheh

(2023) examined the impact of geopolitical shocks, especially the repercussions during the Arab Spring and the Russian attack on Ukraine, on the volatility of the leading stock indices of the Mena region. They found that EGARCH outperforms the simple GARCH model regarding the goodness-of-fit in these periods. Mitsas et al. (2022) analyzed the impact of geopolitical risk indices on volatility using different GARCH models during geopolitical turbulences. At this, the EGARCH model outperformed the GJR-GARCH and the simple GARCH model regarding the lowest value of the AIC. The behavior of stock market volatility and structural breaks for fragile emerging markets was investigated by Yildirim & Celik (2020). They found evidence for the outperformance of the EGARCH model over the simple GARCH model in times with sharp volatility shifts, which is the case for geopolitical shocks. Khan et al. (2023) analyzed several important financial instruments during the period of COVID-19 and found high persistence in volatility. Also, they showed that the EGARCH model outperformed the GJR-GARCH and simple GARCH model for Bitcoin and EUR-USD exchange rate volatility.

Table 23: Best in-sample fit GARCH-type models regarding AIC and AIC weights (in %) for periods in COVID-19, during the Russian invasion and the Hamas attack on Israel (own calculations using Matlab).

	S&P 500	NASDAQ	Nikkei 225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
COVID-19	GARCH(2,1)	GARCH(2,1)	GARCH(1,1)	GARCH(1,1)	GARCH(1,2)	GARCH(1,1)	GARCH(1,1)	GARCH(1,1)	GARCH(3,1)
	-242.11 (0.00%)	-252.42 (0.01%)	-290.52 (0.03%)	-328.91 (0.00%)	-316.59 (0.00%)	-137.34 (0.00%)	-119.88 (0.01%)	-453.24 (0.00%)	-181.41 (0.00%)
	EGARCH(3,4)	EGARCH(3,4)	EGARCH(3,3)	EGARCH(3,3)	EGARCH(5,5)	EGARCH(3,3)	EGARCH(4,2)	EGARCH(5,2)	EGARCH(3,3)
	-279.78 (99.95%)	-270.55 (99.60%)	-307.01 (99.94%)	-348.85 (99.98%)	-342.25 (99.99%)	-165.97 (99.98%)	-137.85 (99.97%)	-475.21 (99.99%)	-230.16 (99.99%)
	GJR(2,1)	GJR(1,1)	GJR(3,1)	GJR(1,1)	GJR(1,1)	GJR(1,3)	GJR(1,1)	GJR(1,1)	GJR(3,1)
	-264.24 (0.04%)	-259.47 (0.39%)	-290.74 (0.03%)	-331.67 (0.02%)	-317.73 (0.00%)	-148.96 (0.02%)	-120.18 (0.01%)	-451.24 (0.00%)	-192.40 (0.00%)
Russian Invasion	GARCH(2,1)	GARCH(1,1)	GARCH(1,1)	GARCH(1,2)	GARCH(1,1)	GARCH(2,1)	GARCH(1,1)	GARCH(1,1)	GARCH(1,1)
	-242.11 (0.00%)	-300.39 (16,19%)	-342.53 (0.62%)	-300.83 (0.01%)	-385.92 (0.31%)	-223.09 (0.10%)	-234.23 (0.99%)	-454.81 (0.00%)	-223.55 (0.00%)
	EGARCH(1,2)	EGARCH(2,1)	EGARCH(1,1)	EGARCH(5,5)	EGARCH(3,4)	EGARCH(3,5)	EGARCH(1,1)	EGARCH(3,2)	EGARCH(3,2)
	-346.40 (92.13%)	-303.40 (72,91%)	-352.63 (97.30%)	-318.63 (99.96%)	-397.47 (99.53%)	-236.89 (99.87%)	-243.41 (97.70%)	-477.72 (99.99%)	-236.12 (99.99%)
Hamas Attack	GJR(1,1)	GJR(1,1)	GJR(1,1)	GJR(1,1)	GJR(1,1)	GJR(2,1)	GJR(1,1)	GJR(1,1)	GJR(1,1)
	-341.48 (7.87%)	-299.60 (10.90%)	-344.94 (2.08%)	-302.06 (0.03%)	-384.62 (0.16%)	-220.82 (0.03%)	-234.79 (1.31%)	-452.81 (0.00%)	-221.61 (0.00%)
	GARCH(1,1)	GARCH(1,1)	GARCH(1,1)	GARCH(1,1)	GARCH(4,1)	GARCH(1,1)	GARCH(1,1)	GARCH(1,1)	GARCH(1,1)
	435.71 (0.01%)	-405.58 (0.00%)	-381.56 (0.05%)	-360.15 (1.09%)	-420.19 (0.00%)	-291.43 (0.01%)	-322.95 (1.48%)	-496.06 (0.01%)	-267.25 (0.07%)
	EGARCH(1,1)	EGARCH(4,2)	EGARCH(1,1)	EGARCH(1,1)	EGARCH(4,3)	EGARCH(4,3)	EGARCH(1,1)	EGARCH(2,2)	EGARCH(3,3)
	-455.71 (99.94%)	-426.67 (99.93%)	-396.86 (99.83%)	-369.15 (98.51%)	-444.65 (99.99%)	-309.29 (99.99%)	-331.33 (97.92%)	-514.96 (99.99%)	-281.86 (99.88%)
	GJR(1,2)	GJR(1,2)	GJR(1,1)	GJR(1,1)	GJR(1,1)	GJR(1,1)	GJR(1,1)	GJR(1,1)	GJR(1,1)
	-440.65 (0.05%)	-412.06 (0.07%)	-383.45 (0.12%)	-358.14 (0.40%)	-415.67 (0.00%)	-289.48 (0.00%)	-321.13 (0.60%)	-494.06 (0.00%)	-266.71 (0.05%)

Source: Own calculations using Matlab (2025).

4.3.2. Univariate Out-of-Sample Volatility Forecasting

Time series models are widely utilized in financial research and practice, primarily for forecasting. While accurately modeling known historical data is crucial for achieving the best in-sample fit, this step is merely an intermediate phase in the broader context of predictive modeling. The ultimate objective is the model's performance in forecasting future, unseen data. Simkins (1995) highlights that models demonstrating the best in-sample fit do not necessarily yield the best out-of-sample forecasting performance. Therefore, evaluating a model's accuracy and effectiveness must be contextualized within its ability to predict out-of-sample data. In this study, I aimed to test the predictive performance of GARCH models selected during my initial analysis, particularly their ability to forecast volatility under conditions of geopolitical uncertainty. I employed a rolling window approach with 250 trading days of historical data for the financial instruments under analysis to achieve this. For each iteration of the rolling window, I estimated the GARCH models, which provided the best in-sample fit for a given financial instrument and period and generated point forecasts for volatility across a forecast horizon of 1- to 10-step-ahead. This iterative process is repeated for each step within the rolling window, ensuring a robust evaluation of the model's performance across different periods.

To assess the accuracy of the forecasts, I applied two evaluation measures. The first is the Symmetric Mean Absolute Percentage Error (SMAPE), which quantifies how close the point forecast is to the true observed value. SMAPE is selected for its ease of interpretation, as it is scale-independent and offers an intuitive measure of forecast accuracy. Additionally, unlike metrics such as the Mean Squared Error (MSE), SMAPE does not penalize large forecast errors disproportionately, offering a more balanced assessment of forecast deviations. Given the focus on point forecasts, I intentionally limit the performance evaluation to SMAPE alone. Including multiple evaluation criteria can lead to inconsistencies and complications in decision-making, mainly when different measures yield contradictory conclusions. Second, I wanted to assess the accuracy of the GARCH models in forecasting the movement of volatility. Regarding reallocating or hedging portfolios, knowing whether volatility is increasing or decreasing is essential. Lai (1990) remarks that there are many cases, where, from an investor's perspective, it is sufficient, to be frequently on the "right" side of the movement. To assess the usefulness of economic forecasts, the economic return must be evaluated rather than a statistical value (Blaskowitz & Herwartz, 2011; Granger & Pesaran, 2000). The economic return can be determined using directional accuracy measures. This measure is also easy to interpret, as the measure is to be considered as a percentage frequency for a correct directional forecast. I start by evaluating the forecasting performance using SMAPE, by considering the SMAPE of the 1-, 5- and 10-step-ahead forecasts. These are relevant values in that they correspond to the forecast for one trading day ($T+1$), one trading week ($T+5$), and two trading weeks ($T+10$). I also considered the mean value over the 1- to 10-step-ahead forecasts. This is the summary of a model's performance across all forecasting horizons. Additionally, I calculated the Diebold-

Mariano test for a statistically significant difference in forecasting performance between the model with the lowest SMAPE and the second-lowest SMAPE. I summarized the results of the SMAPE values for the analyzed financial instruments concerning the best GARCH model in a frequency table (Table 24). This shows which GARCH model class best predicted the financial instruments by the h-step ahead forecast in which frequency occurred.

The detailed results, I provide in Table 25-27. I found that in 51.85% (Table 24: Avg. 1-10-step-ahead Forecast) of all analyzed financial instruments, the EGARCH model class had the smallest percentage deviation from the observed value on average across all 1- to 10-step-ahead forecasts and thus the smallest SMAPE on average. This was followed by the GJR-GARCH model class, which had the lowest average percentage deviation from the true value in 33.33% of all cases. The simple GARCH model ranked last with 14.81%. Furthermore, I found that EGARCH best provided the 1-step-ahead forecast in 44.44% of cases, which the GJR-GARCH model followed with 33.33%. The simple GARCH model performed worst in predicting the 1-step-ahead volatility, with 22.22%. For the 5-step-ahead forecast, EGARCH and GJR-GARCH have the same performance at 40.74%; simple GARCH forecasting performance was poor again. Looking at the 10-step-ahead forecast EGARCH performed best with a frequency of 51.85%. So, one can see that the EGARCH model produces the most accurate volatility point forecast on average across all three geopolitical events.

In this section, I examined the performance of various GARCH models in forecasting financial markets during the COVID-19 pandemic, a period marked by heightened uncertainty and increased volatility due to geopolitical disturbances. The analysis covers forecast horizons ranging from 1- to 10- step-ahead, providing valuable insights into how these models perform under conditions of geopolitical shocks. For the average 1- to 10-step-ahead forecasts, the results indicate that the Exponential GARCH (EGARCH) model produced the most accurate point forecasts in 55.56% of cases across all forecast horizons. This was closely followed by the GJR-GARCH model, which was the top performer for 44.44% of the instruments. Interestingly, the simple GARCH model did not outperform either EGARCH or GJR-GARCH for any of the financial instruments considered. These results can be attributed to the underlying characteristics of the models. EGARCH, known for its ability to capture the leverage effect, performed exceptionally well during this period of regime shifts in volatility, which were largely driven by the pandemic's impact. The GJR-GARCH model, designed to handle regime switches, also fared well as volatility escalated from relatively low to high levels due to the ongoing geopolitical turmoil. Specifically, EGARCH significantly outperformed the other models for the NASDAQ index and Bitcoin, while GJR-GARCH emerged as the top performer for the Hang Seng index and the 10-year Treasury Bond (Table 25). When focusing on the 1-step-ahead forecasts, the EGARCH model showed strong performance by delivering the lowest SMAPE for 44.44% of the financial instruments, including the Nikkei 225, Brent Crude Oil, the 10-year Treasury Bond, and Bitcoin (Table 25). Meanwhile, the GJR-GARCH model provided the lowest SMAPE for 33.33% of the instruments, (Table 24), particularly the S&P 500, NASDAQ, and Gold (Table 25). The simple GARCH model showed an advantage in one-

step-ahead forecasts for the Hang Seng index and the EUR/USD exchange rate. In this context, EGARCH showed significant improvements over other models, especially for Brent Crude Oil and Bitcoin, whereas the GARCH model outperformed the others for the Hang Seng index (Table 25). For the five-step-ahead forecasts, the GJR-GARCH model emerged as the best performer for 55.56% of the instruments (Table 24), including the S&P 500, NASDAQ, Hang Seng, Gold, and the EUR/USD exchange rate. However, the EGARCH model still outperformed the competition for the Nikkei 225, the 10-year Treasury Bond, and Bitcoin, reflecting its strength in periods of volatility shifts. Notably, the simple GARCH model did not deliver superior results for any financial instrument at this horizon, reinforcing the dominance of the more advanced GARCH variants. Once again, GJR-GARCH showed its strength in forecasting for the Hang Seng index (Table 25). At the 10-step-ahead forecast horizon, the EGARCH model performed well, providing the lowest SMAPE for 44.44% of the instruments (Table 24), including the S&P 500, NASDAQ, 10-year Treasury Bond, and Bitcoin. The GJR-GARCH model proved effective for 33.33% of the instruments, such as the Nikkei 225, Hang Seng index, and Brent Crude Oil. Meanwhile, the simple GARCH model outperformed the others for Gold and the EUR/USD exchange rate, suggesting that its simplicity may still have value in specific contexts (Table 25).

Next, I analyzed the performance of GARCH models during the period of the Russian attack, another period marked by elevated market volatility and geopolitical uncertainty. I evaluated the models' forecast performance on average across horizons from 1- to 10- step-ahead. The results indicated that both the EGARCH and GJR-GARCH models performed equally well, each providing the most accurate point forecasts on average for 44.44% of the financial instruments (Table 24). However, the simple GARCH model only outperformed the others for Gold, accounting for 11.11% of the instruments (Table 24 and Table 26). Upon closer inspection, EGARCH significantly outperformed the competing models for the Nikkei 225 index, while the GJR-GARCH model showed the best performance for the Hang Seng index (Table 26). When examining the one-step-ahead forecast results in more detail, I observed that the GJR-GARCH model produced the lowest SMAPE for 44.44% of the instruments (Table 24), including the S&P 500, Nikkei 225, Hang Seng index, and the 10-year Treasury Bond. Meanwhile, EGARCH exhibited superior forecasting accuracy for 33.33% of the instruments, such as the NASDAQ index, Oil, and the EUR/USD exchange rate. Interestingly, the simple GARCH model performed best for Gold and Bitcoin (Table 26). Looking at the 5-step-ahead forecast, the results were more evenly distributed. No model consistently outperformed the others across the majority of financial instruments (Table 24). The simple GARCH model performed best for 33.33% of the instruments, specifically the NASDAQ, Gold, and Bitcoin. Similarly, EGARCH provided the lowest SMAPE for the Nikkei 225, Oil, and the 10-year Treasury Bond, also covering 33.33% of the instruments. The GJR-GARCH model performed best for the remaining 33.33%, including the S&P 500, Hang Seng index, and the EUR/USD exchange rate (Table 26). For the 10-step-ahead forecast, however, EGARCH clearly emerged as the top performer. It provided the most accurate forecasts for 55.56% of the financial

instruments (Table 24), including the S&P 500, NASDAQ, Nikkei 225, Oil, and the 10-year Treasury Bond. The simple GARCH model performed best for 22.22% of the instruments, namely Gold and the EUR/USD exchange rate, while GJR-GARCH provided the lowest SMAPE for the remaining 22.22%, particularly for the Hang Seng index and Bitcoin (Table 26). Notably, GJR-GARCH significantly outperformed the other models when forecasting the Hang Seng index at this forecast horizon, emphasizing its strength in capturing volatility patterns specific to this market.

In the final period with a geopolitical shock under consideration, which corresponds to the time of the Hamas attack on Israel, I again evaluated the performance of various GARCH models in forecasting financial instruments. The results for the average 1- to 10-step-ahead forecasts reveal that the EGARCH model provided the most accurate point forecasts for 55.56% of the financial instruments (Table 24), including the S&P 500, Gold, Oil, EUR/USD, and Bitcoin. In comparison, the simple GARCH model performed best for the NASDAQ, Nikkei 225, and the 10-year Treasury Bond, while the GJR-GARCH model only outperformed the others for the Hang Seng index. Notably, EGARCH significantly outperformed its counterparts for Gold, EUR/USD, and Bitcoin, indicating its strength in capturing the dynamics of these assets during this volatile period (Table 27). When examining the performance for individual forecast horizons, I begin with the one-step-ahead forecasts. EGARCH delivered the lowest SMAPE for 44.44% of the financial instruments (Table 24), including the NASDAQ, Hang Seng, Gold, and EUR/USD. The simple GARCH model performed best for the Nikkei 225 and the 10-year Treasury Bond, accounting for 22.22% of the cases, while the GJR-GARCH model showed superiority for the S&P 500 and Bitcoin, covering another 22.22%. In this horizon, EGARCH significantly outperformed the other models for the NASDAQ, Gold, and EUR/USD, while GJR-GARCH provided the significantly best forecasts for Bitcoin (Table 27). Moving to the 5-step-ahead forecasts, the EGARCH model again demonstrated strong performance, delivering the most accurate forecasts for 44.44% of the financial instruments (Table 24), specifically for the S&P 500, Nikkei 225, Gold, and Bitcoin. The GJR-GARCH model performed best for 33.33% of the instruments, particularly the Hang Seng index, Oil, and EUR/USD, while the simple GARCH model performed best for the NASDAQ and the 10-year Treasury Bond, covering 22.22% of the cases. In this scenario, EGARCH significantly outperformed the other models for the S&P 500, Gold, and Bitcoin, while the simple GARCH model remained the best option for the NASDAQ (Table 27). Finally, for the 10-step-ahead forecasts, EGARCH maintained its superior forecasting accuracy for 55.56% of the financial instruments (Table 24), including the Nikkei 225, Gold, Oil, EUR/USD, and Bitcoin. The simple GARCH model performed best for the S&P 500, Hang Seng, and 10-year Treasury Bond, accounting for 33.33% of the forecasts, while the GJR-GARCH model only outperformed the others for the NASDAQ, covering 11.11% of the instruments. EGARCH once again stood out for its performance on Gold and Bitcoin (Table 27), clearly outperforming the other models during this extended forecast horizon.

In summary, I found that on average over analyzed financial instruments and periods if geopolitical shocks, the EGARCH-type models provided the smallest SMAPE across 1- to 10 step ahead forecasts and thus the smallest percentage deviation from the observed value. This was followed by the GJR-GARCH model class. The simple GARCH model came in last. However, the results for the individual financial instruments and prediction horizons vary over financial instruments, forecast horizon and period. However, we found that the model class for a selected financial instrument is often the same for predicting the various periods of geopolitical shocks. This is also in accordance with my results of the best model for the in-sample fit. I found support for my findings in e.g. Chong et al. (1999), Cumby et al. (1993), Lama et al. (2015), Najand (2002), Pagan & Schwert (1990) and Thenmozhi & Radha (2008). This allowed me to answer the first part of my research question (Research Question 7).

Table 24: Frequency of lowest SMAPE of a GARCH type model for examined periods over all financial instruments (own calculations).

	1-step ahead Forecast			5-step ahead Forecast			10-step ahead Forecast			Avg. 1-10-step ahead Forecast		
	GARCH	EGARCH	GJR	GARCH	EGARCH	GJR	GARCH	EGARCH	GJR	GARCH	EGARCH	GJR
COVID-19	22.22%	44.44%	33.33%	0.00%	44.44%	55.56%	22.22%	44.44%	33.33%	0.00%	55.56%	44.44%
Russian War	22.22%	33.33%	44.44%	33.33%	33.33%	33.33%	22.22%	55.56%	22.22%	11.11%	44.44%	44.44%
Hamas Attack	22.22%	55.56%	22.22%	22.22%	44.44%	33.33%	33.33%	55.56%	11.11%	33.33%	55.56%	11.11%
Average	22.22%	44.44%	33.33%	18.52%	40.74%	40.74%	25.93%	51.85%	22.22%	14.81%	51.85%	33.33%

Source: Own calculations using Matlab (2025).

Table 25: SMAPE for GARCH-type multi-step ahead out-of-sample forecasting during the period of COVID-19 (02/03/2020 – 04/30/2020). Diebold-Mariano test for significant difference between best and second-best performing GARCH model for given financial instrument and forecast horizon (h) on 5% level (*) and 1% level (**).

	Modell / h	1	2	3	4	5	6	7	8	9	10	Average
S&P 500	GARCH(2,1)	0.839	0.963	1.131	1.203	1.001	1.281	1.183	1.656	1.701	1.502	1.246
	EGARCH(3,4)	0.901	0.907	0.900	0.924	0.948	0.917	0.932	0.931	0.936	0.952	0.924*
	GJR(2,1)	0.782*	0.849	0.918	0.964	0.913	0.980	1.028	1.207	1.266	0.970	0.987
NASDAQ	GARCH(2,1)	0.905	0.959	0.992	0.944	0.948	0.970	1.306	1.606	1.540	2.366	1.253
	EGARCH(3,4)	0.900	0.916	0.899	0.924	0.955	0.962	0.963	0.952	0.971	0.968**	0.941*
	GJR(1,1)	0.817*	0.889	0.923	0.875	0.869*	0.939	1.215	1.621	1.703	1.505	1.135
Nikkei 225	GARCH(1,1)	0.938	1.420	1.032	1.004	1.004	1.073	0.959	1.021	0.992	0.946	1.038
	EGARCH(3,3)	0.877*	0.955	0.921	0.925	0.867*	0.895	0.927	0.944	0.940	0.942	0.919*
	GJR((3,1)	1.137	2.086	1.100	1.618	1.232	1.043	1.354	1.056	0.933	0.919*	1.247
Hang Seng	GARCH(1,1)	1.595*	1.253	1.270	1.797	2.450	2.423	1.839	2.030	2.502	2.558	1.971
	EGARCH(3,3)	2.210	1.222	1.612	1.664	1.992	1.566	1.448	1.587	1.767	1.878	1.694
	GJR(1,1)	2.412	0.872	1.144	1.396	1.646**	1.397	1.216	1.115	1.296	1.406**	1.390**
Gold	GARCH(1,2)	0.858	0.868	0.866	0.874	0.878	0.878	0.886	0.886	0.876	0.883	0.875
	EGARCH(5,5)	0.887	0.903	0.905	0.925	0.931	0.920	0.940	0.932	0.943	0.959	0.924
	GJR(1,1)	0.824*	0.874	0.872	0.876	0.874	0.880	0.887	0.888	0.877	0.884	0.873
Brent Oil	GARCH(1,1)	1.352	0.953	1.205	1.071	1.556	1.225	1.245	1.248	1.170	0.966	1.199
	EGARCH(3,3)	0.961*	0.914	0.934	0.930	0.967*	0.945	0.911	0.889	0.948	0.935	0.933*
	GJR(1,3)	1.505	1.596	1.002	1.344	1.141	0.988	0.947	0.916	0.903	0.903	1.124
10yr. T-Bond	GARCH(1,1)	1.211	0.876	1.146	1.168	1.592	2.205	2.160	1.924	1.957	2.402	1.664
	EGARCH(4,2)	0.977*	0.991	0.987	10.974	0.960*	0.972	0.955	0.969	0.977	0.971**	1.973
	GJR(1,1)	1.065	0.834	0.921	1.216	1.390	1.873	1.654	1.602	1.416	1.946	1.391**
EUR/USD	GARCH(1,1)	0.755*	0.798	0.797	0.789	0.785	0.800	0.841	0.879	0.859	0.859	0.816
	EGARCH(5,2)	0.949	0.882	0.879	0.889	0.780	0.820	0.833	0.774	0.836	0.882	0.852
	GJR(1,1)	0.823	0.835	0.794	0.706	0.725*	0.781	0.874	0.871	0.864	0.863	0.813
BTC	GARCH(1,1)	5.603	6.107	5.922	7.482	7.353	8.376	7.406	7.355	8.073	10.475	7.415
	EGARCH(5,2)	1.849**	3.185	3.184	2.551	1.399**	1.801	2.133	1.567	1.876	1.960**	2.150**
	GJR(1,1)	4.769	5.037	4.523	5.490	9.360	10.916	12.218	11.932	16.373	12.123	9.274

Source: Own calculations using Matlab (2025).

Table 26: SMAPE for GARCH-type multi-step ahead out-of-sample forecasting during the period of the Russian invasion on Ukraine (02/01/2022 – 04/29/2022). Diebold-Mariano test for significant difference between best and second-best performing GARCH model for given financial instrument and forecast horizon (h) on 5% level (*) and 1% level (**).

	Modell / h	1	2	3	4	5	6	7	8	9	10	Average
S&P 500	GARCH(2,1)	0.805	0.778	0.739	0.778	0.773	0.805	0.786	0.836	0.866	0.871	0.804
	EGARCH(1,2)	0.676	0.640	0.727	0.688	0.766	0.744	0.784	0.804	0.826	0.826*	0.748*
	GJR(1,1)	0.642*	0.774	0.731	0.757	0.729*	0.830	0.797	0.801	0.769	0.856	0.769
NASDAQ	GARCH(1,1)	0.663	0.651	0.660	0.711	0.656*	0.690	0.662	0.680	0.716	0.720	0.681
	EGARCH(2,1)	0.575*	0.607	0.559	0.639	0.682	0.672	0.685	0.658	0.730	0.704	0.651
	GJR(1,1)	0.601	0.581	0.582	0.658	0.691	0.701	0.699	0.660	0.689	0.718	0.658
Nikkei 225	GARCH(1,1)	0.995	0.883	0.978	1.097	1.056	1.027	1.040	1.166	1.057	1.046	1.035
	EGARCH(1,1)	0.900	0.875	0.879	0.869	0.901*	0.894	0.897	0.970	0.917	0.934*	0.904*
	GJR(1,1)	0.883	0.865	0.948	1.024	1.053	1.007	1.026	1.167	1.022	0.987	0.998
Hang Seng	GARCH(1,2)	2.069	2.204	1.642	2.474	2.898	2.645	2.648	2.982	1.973	1.723	2.326
	EGARCH(5,5)	1.138	1.980	2.018	0.972	1.567	0.937	2.576	1.533	1.040	1.129	1.489
	GJR(1,1)	0.899*	0.902	0.936	0.934	0.954**	0.952	0.969	0.959	0.942	0.952*	0.940*
Gold	GARCH(1,1)	0.801*	0.799	0.754	0.754	0.755	0.799	0.798	0.798	0.764	0.731*	0.775*
	EGARCH(3,4)	4.298	3.662	2.733	1.775	1.722	2.490	3.819	3.909	2.703	1.546	2.866
	GJR(1,1)	0.873	0.844	0.834	0.833	0.834	0.852	0.851	0.853	0.785	0.785	0.831
Brent Oil	GARCH(2,1)	0.899	0.850	0.809	0.815	0.901	0.822	0.893	0.902	0.908	0.861	0.866
	EGARCH(3,5)	0.804*	0.797	0.793	0.809	0.821*	0.816	0.794	0.807	0.791	0.812*	0.804*
	GJR(2,1)	0.884	0.861	0.853	0.848	0.907	0.838	0.860	0.907	0.890	0.867	0.872
10yr. T-Bond	GARCH(1,1)	0.827	0.827	0.773	0.823	0.875	0.820	0.845	0.816	0.818	0.853	0.828
	EGARCH(1,1)	0.841	0.826	0.825	0.809	0.821*	0.804	0.806	0.793	0.784	0.808	0.812
	GJR(1,1)	0.819	0.837	0.799	0.829	0.875	0.795	0.783	0.781	0.774	0.819	0.811
EUR/USD	GARCH(1,1)	0.930	0.930	0.952	0.954	1.094	0.954	0.888	0.873	0.874	0.873	0.932
	EGARCH(3,2)	0.906*	0.898	0.953	0.946	1.129	0.917	0.869	0.853	0.882	0.877	0.923
	GJR(1,1)	0.936	0.936	0.936	0.933	0.954*	0.914	0.873	0.874	0.876	0.874	0.911
BTC	GARCH(1,1)	2.084*	2.167	2.242	2.359	2.359**	2.323	2.245	2.341	2.335	2.440	2.289
	EGARCH(3,2)	2.449	2.617	2.915	2.950	3.089	2.530	2.457	2.368	2.354	2.454	2.618
	GJR(1,1)	2.172	2.252	2.319	2.430	2.518	2.116	2.157	2.198	2.195	2.268**	2.263

Source: Own calculations using Matlab (2025).

Table 27: SMAPE for GARCH-type multi-step ahead out-of-sample forecasting during the period around the Hamas attack on Israel (10/01/2023 – 12/29/2023). Diebold-Mariano test for significant difference between best and second-best performing GARCH model for given financial instrument and forecast horizon (h) on 5% level (*) and 1% level (**).

	Modell / h	1	2	3	4	5	6	7	8	9	10	Average
S&P 500	GARCH(1,1)	1.503	1.595	1.710	1.812	1.867	1.838	1.955	2.042	1.926	1.974*	1.822
	EGARCH(1,1)	1.257	1.574	1.730	1.836	1.599**	1.735	2.167	1.559	1.858	2.041	1.736*
	GJR(1,2)	1.142*	1.453	1.837	1.921	1.772	1.910	2.016	2.125	2.257	2.305	1.874
NASDAQ	GARCH(1,1)	1.549	1.870	2.053	2.241	2.102**	2.071	2.034	1.954	2.064	2.118	2.006*
	EGARCH(4,2)	1.012**	1.534	1.500	2.415	2.593	2.519	2.181	2.050	2.362	2.362	2.067
	GJR(1,2)	1.705	2.093	2.461	2.637	2.335	2.454	1.949	2.103	2.299	1.954*	2.195
Nikkei 225	GARCH(1,1)	1.217*	1.408	1.637	1.825	2.084	2.192	2.201	2.097	2.299	2.272	1.923*
	EGARCH(1,1)	1.261	1.636	1.677	1.827	2.078	2.177	2.219	2.145	2.272	2.176	1.947
	GJR((1,1)	1.288	1.545	1.746	1.921	2.247	2.252	2.225	2.1543	2.320	2.251	1.995
Hang Seng	GARCH(1,1)	1.230	1.219	1.200	1.124	1.210	1.309	1.156	1.133	1.330	1.240	1.215
	EGARCH(4,3)	1.098*	1.052	1.522	1.366	1.570	1.424	1.048	1.289	1.578	1.431	1.338
	GJR(1,1)	1.215	1.232	1.215	1.088	1.191	1.243	1.172	1.141	1.351	1.275	1.212
Gold	GARCH(4,1)	2.228	2.454	2.311	2.244	1.879	2.096	2.235	2.517	2.752	2.490	2.320
	EGARCH(4,3)	1.711**	1.252	1.438	1.567	1.012**	1.604	1.542	1.297	1.813	1.705**	1.494**
	GJR(1,1)	2.534	2.470	2.077	2.104	2.116	2.294	2.457	2.519	2.649	2.643	2.386
Brent Oil	GARCH(1,1)	0.903	0.941	0.950	0.814	0.823	0.810	0.827	0.859	0.847	0.810	0.859
	EGARCH(4,3)	0.855*	0.918	0.938	0.830	0.924	0.785	0.792	0.822	0.805	0.787	0.845
	GJR(2,1)	1.092	0.865	0.899	0.783	0.797	0.787	0.796	0.836	0.862	0.837	0.855
10yr. T-Bond	GARCH(1,1)	0.785	0.758	0.755	0.726	0.716	0.754	0.753	0.739	0.785	0.800	0.757
	EGARCH(1,1)	0.984	1.076	0.906	0.910	1.000	0.909	0.788	0.909	1.000	1.449	0.993
	GJR(1,1)	0.791	0.781	0.764	0.709	0.742	0.730	0.754	0.754	0.792	0.824	0.764
EUR/USD	GARCH(1,1)	2.180	2.443	2.061	2.130	2.174	2.140	2.286	2.237	2.242	2.354	2.225
	EGARCH(2,2)	1.771**	1.903	1.958	2.062	2.278	1.814	1.896	2.098	1.902	2.128	1.981*
	GJR(1,1)	1.922	2.451	2.081	2.035	1.954*	2.084	2.170	2.137	2.236	2.242	2.131
BTC	GARCH(1,1)	2.367	2.595	3.268	3.528	4.218	3.070	3.362	2.736	3.438	2.989	3.157
	EGARCH(3,3)	2.243	2.422	2.610	2.692	3.328**	2.855	2.485	2.186	2.764	2.217**	2.580**
	GJR(1,1)	2.003**	2.573	2.940	3.057	3.990	3.184	3.069	2.610	3.327	3.091	2.985

Source: Own calculations using Matlab (2025).

Now I turn attention to evaluating the models' accuracy in predicting the direction of volatility movement during the COVID-19 pandemic. When examining the average 1- to 10-step-ahead forecasts, the GJR-GARCH model provided the most accurate directional predictions across 66.67% of the financial instruments (Table 28), including the S&P 500, Nikkei 225, Hang Seng, Gold, Oil, and EUR/USD. In contrast, the EGARCH model showed the highest accuracy for the NASDAQ, 10-year Treasury Bond, and Bitcoin, covering 33.33% of the instruments. The simple GARCH model performed well for Gold and the 10-year Treasury Bond. For the 1-step-ahead forecasts, GJR-GARCH demonstrated strong performance, with an average directional accuracy for 77.78% of the financial instruments, which included the S&P 500, Nikkei 225, Hang Seng, Gold, Oil, 10-year Treasury Bond, and Bitcoin. On the other hand, both the simple GARCH and EGARCH models also exhibited notable results in specific instances, with simple GARCH performing well for the Hang Seng, Gold, and Bitcoin (33.33%) and EGARCH providing strong forecasts for NASDAQ, EUR/USD, and Bitcoin (33.33%) (Table 30). Looking at the 5-step-ahead forecasts, EGARCH model class provided the best directional volatility forecasts for 55.56% of the instruments (Table 28), including the Hang Seng index, Gold, Oil, EUR/USD, and Bitcoin (Table 30). The GJR-GARCH model was also effective, delivering the best directional forecasts for NASDAQ, Nikkei 225, Gold, and the 10-year Treasury Bond, accounting for 44.44% of the instruments. The simple GARCH model showed less frequent accuracy but still provided the best forecasts for the S&P 500 and Gold, covering 22.22% of the cases. Finally, for the 10-step-ahead forecasts, the GJR-GARCH model showed the highest directional forecasting accuracy, for 77.78% (Table 28) of the financial instruments, which included the S&P 500, Nikkei 225, Hang Seng, Gold, Oil, 10-year Treasury Bond, and EUR/USD (Table 30). EGARCH also performed well, delivering the best predictions for NASDAQ, Gold, the 10-year Treasury Bond, EUR/USD, and Bitcoin, covering 55.56% of the instruments. The simple GARCH model, although less dominant, still provided accurate forecasts for Hang Seng, Gold, and Oil, accounting for 33.33% of the cases.

Next, I examined the forecasting performance for the GARCH models in scope during the period of the Russian invasion on Ukraine. For the average 1- to 10-step-ahead forecasts, the GJR-GARCH model provided the most accurate directional predictions across 55.56% of the financial instruments (Table 28). The model performed well for the S&P 500, NASDAQ, Hang Seng, EUR/USD, and Bitcoin (Table 31). The EGARCH model followed by providing the highest accuracy for 33.00% of the instruments, including the Nikkei 225, Oil, and the 10-year Treasury Bond. The simple GARCH model provided the best for Gold and Bitcoin (22.00%). When examining the one-step-ahead directional forecasts, the simple GARCH model stood out by providing the highest accuracy for 66.67% of the financial instruments (Table 28). This included the S&P 500, Nikkei 225, Gold, Oil, the 10-year Treasury Bond, and Bitcoin (Table 31). The GJR-GARCH model outperformed the other models for 44.44% of the instruments, particularly the S&P 500, NASDAQ, Nikkei 225, and Bitcoin. EGARCH also contributed, offering directional solid forecasts for the Nikkei 225, Hang Seng, and EUR/USD

(33.33%). For the 5-step-ahead forecasts, both EGARCH and GJR-GARCH models delivered the highest directional accuracy, each covering 55.56% of the financial instruments (Table 28). EGARCH performed best for the S&P 500, Gold, Oil, the 10-year Treasury Bond, and EUR/USD, while GJR-GARCH excelled for NASDAQ, Hang Seng, the 10-year Treasury Bond, EUR/USD, and Bitcoin (Table 31). The simple GARCH model, although less frequent, performed well for 44.44% of the instruments, including the S&P 500, Nikkei 225, EUR/USD, and Bitcoin. Finally, for the ten-step-ahead forecasts, the GJR-GARCH model emerged as the top performer, providing the highest directional accuracy for 77.78% of the financial instruments. This included the S&P 500, Nikkei 225, Hang Seng, Gold, the 10-year Treasury Bond, EUR/USD, and Bitcoin. EGARCH was also effective, delivering the best forecasts for 55.56% of the instruments (Table 28), particularly the S&P 500, Nikkei 225, Oil, the 10-year Treasury Bond, and EUR/USD. The simple GARCH model also performed well for a subset of the instruments, including NASDAQ, Gold, the 10-year Treasury Bond, EUR/USD, and Bitcoin (Table 31).

The final period under analysis focuses on the time of the Hamas attack on Israel. During this highly volatile period, the forecasting performance of various GARCH models was also evaluated across multiple time horizons. For the average 1- to 10-step-ahead forecasts, both the GJR-GARCH and EGARCH models provided the most accurate directional predictions for 44.44% of the financial instruments (Table 28). Specifically, the EGARCH model excelled in predicting the S&P 500, NASDAQ, Nikkei 225, Gold, and Bitcoin (Table 32), while GJR-GARCH performed best for NASDAQ, Hang Seng, Oil, and EUR/USD. The simple GARCH model showed its strength primarily with the Nikkei 225, Hang Seng, and the 10-year Treasury Bond. In the case of one-step-ahead forecasts, both EGARCH and GJR-GARCH once again proved to be the most accurate, each covering 66.67% of the financial instruments (Table 28). EGARCH provided the best directional forecasts for the S&P 500, NASDAQ, Nikkei 225, Hang Seng, Gold, and Oil (Table 32). Similarly, GJR-GARCH excelled with NASDAQ, Nikkei 225, Gold, Oil, EUR/USD, and Bitcoin. The simple GARCH model was also effective in some cases, particularly for NASDAQ, Nikkei 225, and the 10-year Treasury Bond, accounting for 33.33% (Table 28) of the instruments. For the 5-step-ahead forecasts, the simple GARCH and EGARCH models both showed high directional accuracy, each performing best for 55.56% of the instruments (Table 28). The simple GARCH model succeeded with NASDAQ, Nikkei 225, Gold, the 10-year Treasury Bond, EUR/USD, and Bitcoin, while EGARCH provided the most accurate predictions for the S&P 500, NASDAQ, Nikkei 225, Hang Seng, and EUR/USD (Table 32). The GJR-GARCH model followed closely, performing best for NASDAQ, Nikkei 225, Oil, and EUR/USD, covering 44.44% of the instruments. Lastly, for the 10-step-ahead forecasts, EGARCH demonstrated the highest accuracy for 66.67% of the financial instruments, including the S&P 500, NASDAQ, Nikkei 225, Gold, the 10-year Treasury Bond, EUR/USD, and Bitcoin. GJR-GARCH also performed well, offering strong forecasts for NASDAQ, Nikkei 225, Hang Seng, Oil, the 10-year Treasury Bond, and EUR/USD, covering 55.56% of the instruments. The simple GARCH model maintained solid performance, excelling with

NASDAQ, Nikkei 225, Hang Seng, the 10-year Treasury Bond, and EUR/USD, matching the 55.56% accuracy rate seen with GJR-GARCH. This analysis of forecasting performance during the Hamas attack on Israel highlights the strengths of each GARCH model in predicting directional volatility movements, with EGARCH and GJR-GARCH consistently delivering strong results across different horizons.

Summarizing, I found that the GJR-GARCH-class models outperformed other forecasting models in scope regarding the accuracy of predicting correct directional movements of the volatility during periods of political shocks. This result was most pronounced during the COVID-19 phase. It can be assumed that the GJR-GARCH model is better able to predict volatilities due to its asymmetric structure, especially after strong negative shocks, which we observed most strongly in COVID-19, leading to higher directional accuracy. In addition, the model performed particularly well in predicting short-term volatility. Our results are in line with the other literature (e.g., OU et al., 2022; Ramasamy, 2012).

In our analysis, we examined various GARCH-type models for their ability to predict the volatility of various financial instruments and prediction horizons in different periods of geopolitical shocks. We found that the objective of the prediction (point prediction or prediction of directional accuracy) has an influence on which model should be selected for the prediction. For point forecasting, the EGARCH model turned out to be the best model. For forecasting directional accuracy, the GJR-GARCH model. This led me to answer my research question (Research Question 7).

Table 28: Frequency of highest directional accuracy of a GARCH-type model for examined periods and financial instruments.

	1-step ahead Forecast			5-step ahead Forecast			10-step ahead Forecast			Avg. 1-10-step ahead Forecast		
	GARCH	EGARCH	GJR	GARCH	EGARCH	GJR	GARCH	EGARCH	GJR	GARCH	EGARCH	GJR
COVID-19	33.33%	33.33%	77.78%	22.22%	55.56%	44.44%	33.33%	55.56%	77.78%	11.11%	33.33%	66.67%
Russian Invasion	66.67%	33.33%	44.44%	44.44%	55.56%	55.56%	55.56%	55.56%	77.76%	22.22%	33.33%	55.56%
Hamas Attack	33.33%	66.66%	66.66%	55.56%	55.56%	44.44%	55.556%	66.67%	55.56%	33.33%	44.44%	44.44%
Average	44.44%	44.44%	62.96%	40.74	55.56%	48.15%	48.15%	59.26%	70.37%	25.93%	37.04%	55.56%

Source: Own calculations using Matlab (2025).

Table 29: Best performing GARCH-type model for multi-step ahead out-of-sample forecasting in examined periods and financial instruments.

		S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Oil	10yr. T-Bond	EUR/USD	BTC
SMAPE	COVID-19	EGARCH	EGARCH	EGARCH	GJR	GJR	EGAR	GJR	GJR	EGARCH
	Russian Invasion	EGARCH	EGARCH	EGARCH	GJR	GARCH	GJR	GJR	GJR	EGARCH
	Hamas Attack	EGARCH	GARCH	GARCH	GJR	EGARCH	EGAR	GAR	EGARCH	EGARCH
Directional Accuracy	COVID-19	GJR	EGARCH	GJR	GJR	GAR/GJR	GJR	EGARCH	GJR	EGARCH
	Russian Invasion	GJR	GJR	EGARCH	GJR	EGARCH	EGARCH	EGARCH	GJR	GAR/GJR
	Hamas Attack	EGARCH	GJR	GAR/EGARCH	GARCH/GJR	EGARCH	GJR	GARCH	GJR	EGARCH

Source: Own calculations using Matlab (2025).

Table 30: Directional accuracy of GARCH-type multi-step ahead out-of-sample forecasting in period of COVID-19 (02/03/2020 – 04/30/2020). Adjusted Diebold-Mariano test (DM*) for significant difference between best and second-best performing GARCH-type model for given financial instrument and forecast horizon (h) on 5% level (*) and 1% level (**).

	Modell / h	1	2	3	4	5	6	7	8	9	10	Average
S&P 500	GARCH(2,1)	0.700	0.733	0.667	0.700	0.667	0.650	0.633	0.550	0.616	0.683	0.660
	EGARCH(3,4)	0.716	0.700	0.667	0.616	0.583	0.616	0.566	0.650**	0.783**	0.683	0.658
	GJR(2,1)	0.800**	0.750	0.716**	0.783**	0.650	0.733**	0.683**	0.600	0.700	0.700	0.711*
NASDAQ	GARCH(2,1)	0.667	0.616	0.700	0.633	0.633	0.650	0.633	0.633	0.616	0.650	0.643
	EGARCH(3,4)	0.783**	0.667	0.700	0.667	0.650	0.700	0.700**	0.683**	0.667**	0.733**	0.695
	GJR(1,1)	0.733	0.716**	0.700	0.766**	0.733**	0.733*	0.650	0.633	0.600	0.667	0.693
Nikkei 225	GARCH(1,1)	0.683	0.716	0.667	0.683*	0.750	0.766	0.733**	0.633	0.583	0.733	0.695
	EGARCH(3,3)	0.650	0.667	0.616	0.633	0.700	0.750	0.683	0.650	0.633	0.716	0.670
	GJR(3,1)	0.700	0.716	0.716**	0.650	0.766	0.766	0.683	0.683*	0.633	0.750	0.706
Hang Seng	GARCH(1,1)	0.716	0.716	0.766	0.683	0.650	0.700	0.783	0.783	0.616	0.683	0.710
	EGARCH(3,3)	0.650	0.700	0.733	0.633	0.683*	0.667	0.733	0.750	0.650	0.667	0.686
	GJR(1,1)	0.716	0.733	0.766	0.700	0.633	0.733*	0.783	0.816*	0.650	0.683	0.721*
Gold	GARCH(1,2)	0.750	0.750	0.700	0.600	0.550	0.700	0.700	0.850	0.750	0.800	0.715
	EGARCH(5,5)	0.700	0.700	0.550	0.600	0.550	0.700	0.650	0.800	0.750	0.800	0.680
	GJR(1,1)	0.750	0.750	0.700	0.600	0.550	0.700	0.700	0.850	0.750	0.800	0.715
Brent Oil	GARCH(1,1)	0.650	0.750*	0.733	0.733	0.650	0.667	0.650	0.750*	0.733	0.700	0.701
	EGARCH(3,3)	0.667	0.667	0.750	0.733	0.733	0.783**	0.683	0.700	0.733	0.667	0.711
	GJR(1,3)	0.700*	0.716	0.750	0.700	0.716	0.700	0.716*	0.716	0.766*	0.700	0.718
10yr T-Bond	GARCH(1,1)	0.616	0.716**	0.700	0.650	0.566	0.633	0.700	0.633	0.716	0.616	0.655
	EGARCH(4,2)	0.566	0.667	0.716	0.633	0.600	0.667*	0.750**	0.683**	0.716	0.650	0.665
	GJR(1,1)	0.633	0.616	0.616	0.616	0.616	0.633	0.633	0.633	0.633	0.650	0.663
EUR/USD	GARCH(1,1)	0.716	0.800	0.716	0.700	0.683	0.750	0.800	0.783	0.633	0.716	0.730
	EGARCH(5,2)	0.766**	0.750	0.700	0.716	0.700	0.733	0.766	0.766	0.650	0.733	0.728
	GJR(1,1)	0.716	0.816	0.716	0.700	0.683	0.750	0.800	0.783	0.616	0.733	0.731
BTC	GARCH(3,1)	0.633	0.600	0.650	0.583	0.616	0.600	0.566	0.516	0.616	0.616	0.600
	EGARCH(3,3)	0.633	0.600	0.683	0.600	0.700**	0.650	0.667**	0.583**	0.683**	0.683**	0.648*
	GJR(3,1)	0.633	0.583	0.667	0.600	0.650	0.616	0.583	0.533	0.633	0.633	0.613

Source: Own calculations using Matlab (2025).

Table 31: Directional accuracy of GARCH-type multi-step ahead out-of-sample forecasting for period around the Russian invasion on Ukraine (02/01/2022 – 04/29/2022). Adjusted Diebold-Mariano test (DM*) for significant difference between best and second-best performing GARCH-type model for given financial instrument and forecast horizon (h) on 5% level (*) and 1% level (**).

	Modell	1	2	3	4	5	6	7	8	9	10	Average
S&P 500	GARCH(2,1)	0.783	0.716	0.766	0.733	0.650	0.700	0.766	0.700	0.616	0.700	0.713
	EGARCH(1,2)	0.733	0.733	0.816	0.733	0.650	0.733	0.750	0.750	0.616	0.716	0.723
	GJR(1,1)	0.783	0.750	0.816	0.750	0.633	0.716	0.750	0.750	0.633	0.716	0.730
NASDAQ	GARCH(1,1)	0.733	0.700	0.750	0.700	0.633	0.716	0.700	0.733	0.667	0.716	0.705
	EGARCH(2,1)	0.750	0.766	0.733	0.783**	0.650	0.766	0.667	0.766	0.650	0.700	0.723
	GJR(1,1)	0.766	0.766	0.783*	0.733	0.667	0.750	0.700	0.766	0.633	0.683	0.725
Nikkei 225	GARCH(1,1)	0.650	0.783	0.716	0.650	0.683	0.783	0.766	0.633	0.766	0.766	0.720
	EGARCH(1,1)	0.650	0.816*	0.750	0.683	0.650	0.783	0.800	0.667	0.783	0.783	0.736
	GJR(1,1)	0.650	0.766	0.733	0.667	0.667	0.800	0.783	0.650	0.783	0.783	0.728
Hang Seng	GARCH(1,2)	0.683	0.683	0.650	0.716	0.667	0.667	0.650	0.700	0.550	0.667*	0.663
	EGARCH(5,5)	0.745*	0.727	0.600	0.781*	0.618	0.636	0.690	0.745*	0.636	0.709	0.689
	GJR(1,1)	0.716	0.733	0.667	0.750	0.700*	0.716*	0.700	0.716	0.633	0.716	0.705
Gold	GARCH(1,1)	0.750	0.650	0.683	0.750	0.683	0.733	0.683	0.616	0.733	0.700	0.698
	EGARCH(3,4)	0.607	0.549	0.725*	0.784*	0.686	0.725	0.588	0.549	0.725	0.686	0.662*
	GJR(1,1)	0.733	0.650	0.683	0.750	0.683	0.733	0.683	0.616	0.733	0.700	0.696
Brent Oil	GARCH(2,1)	0.816*	0.716	0.783	0.700	0.650	0.783	0.633	0.733	0.700	0.700	0.725
	EGARCH(3,5)	0.783	0.750	0.766	0.716	0.683	0.783	0.650	0.716	0.716	0.750	0.731
	GJR(2,1)	0.783	0.733	0.800	0.700	0.667	0.766	0.650	0.700	0.700	0.733	0.723
10yr T-Bond	GARCH(1,1)	0.683	0.800	0.650	0.800	0.733	0.733	0.700	0.716	0.783	0.750	0.735
	EGARCH(1,1)	0.667	0.866*	0.700	0.783	0.750	0.750	0.750	0.716	0.783	0.750	0.751
	GJR(1,1)	0.616	0.833	0.683	0.816	0.750	0.766	0.733	0.716	0.783	0.750	0.745
EUR/USD	GARCH(1,1)	0.766	0.750	0.750	0.816	0.783	0.750	0.800	0.716	0.683	0.716	0.753
	EGARCH(3,2)	0.783	0.733	0.733	0.833	0.783	0.750	0.800	0.716	0.683	0.716	0.733
	GJR(1,1)	0.750*	0.750	0.766	0.833	0.783	0.750	0.800	0.716	0.683	0.716	0.755
BTC	GARCH(1,1)	0.716	0.667	0.633	0.733	0.683	0.733	0.700	0.633	0.683	0.667	0.685
	EGARCH(3,2)	0.700	0.650	0.616	0.750	0.667	0.716	0.683	0.650	0.667	0.650	0.675
	GJR(1,1)	0.716	0.667	0.663*	0.733	0.683	0.733	0.700	0.633	0.683	0.667	0.685

Source: Own calculations using Matlab (2025).

Table 32: Directional accuracy of GARCH-type multi-step ahead out-of-sample forecasting for period around Hamas attack on Israel (10/01/2023 – 12/29/2023). Adjusted Diebold-Mariano test (DM*) for significant difference between best and second-best performing GARCH model for given financial instrument and forecast horizon (h) on 5% level (*) and 1% level (**).

	Modell / h	1	2	3	4	5	6	7	8	9	10	Average
S&P 500	GARCH(1,1)	0.683	0.750	0.733	0.650	0.633	0.667	0.716	0.683	0.650	0.667	0.683
	EGARCH(1,1)	0.700	0.750	0.733	0.650	0.733**	0.667	0.783*	0.667	0.650	0.700	0.703
	GJR(1,2)	0.683	0.700	0.716	0.650	0.667	0.633	0.750	0.650	0.650	0.667	0.676
NASDAQ	GARCH(1,1)	0.850	0.733	0.750	0.667	0.633	0.633	0.716	0.667	0.683	0.750	0.708
	EGARCH(4,2)	0.850	0.766	0.783	0.700	0.633	0.616	0.700	0.650	0.700	0.750	0.715
	GJR(1,2)	0.850	0.766	0.783	0.700	0.633	0.633	0.716	0.667	0.716	0.750	0.721
Nikkei 225	GARCH(1,1)	0.783	0.667	0.816	0.783	0.783	0.783	0.633	0.766	0.733	0.750	0.750
	EGARCH(1,1)	0.783	0.667	0.800	0.800	0.783	0.783	0.633	0.766	0.733	0.750	0.750
	GJR((1,1)	0.783	0.650	0.800	0.800	0.783	0.783	0.633	0.766	0.733	0.750	0.748
Hang Seng	GARCH(1,1)	0.716	0.816	0.683	0.683	0.683	0.716	0.766	0.633	0.716	0.700	0.711
	EGARCH(4,3)	0.724	0.810	0.689	0.672	0.689	0.724	0.758	0.637	0.724	0.672	0.710
	GJR(1,1)	0.716	0.816	0.683	0.683	0.683	0.716	0.766	0.633	0.716	0.700	0.711
Gold	GARCH(4,1)	0.700	0.716	0.667	0.633	0.733*	0.716	0.733	0.600	0.583	0.700	0.678
	EGARCH(4,3)	0.716	0.683	0.700	0.716**	0.667	0.700	0.783	0.667	0.600	0.750**	0.698
	GJR(1,1)	0.716	0.716	0.700	0.667	0.700	0.733	0.766	0.600	0.616	0.700	0.691
Brent Oil	GARCH(1,1)	0.700	0.750	0.733	0.783	0.750	0.733	0.733	0.733	0.783	0.750	0.745
	EGARCH(4,3)	0.716	0.733	0.716	0.733	0.766	0.750	0.716	0.700	0.800	0.750	0.738
	GJR(2,1)	0.716	0.750	0.733	0.733	0.783	0.750	0.750	0.750	0.766	0.766	0.750
10yr T-Bond	GARCH(1,1)	0.766**	0.766	0.716	0.650	0.733	0.700	0.766	0.700	0.750	0.667	0.721
	EGARCH(1,1)	0.667	0.733	0.633	0.667	0.633	0.716	0.700	0.700	0.700	0.667	0.681
	GJR(1,1)	0.716	0.716	0.683	0.650	0.716	0.716	0.750	0.700	0.750	0.667	0.706
EUR/USD	GARCH(1,1)	0.633	0.667	0.683	0.683	0.683	0.733	0.766	0.667	0.667	0.700	0.688
	EGARCH(2,2)	0.650	0.667	0.667	0.683	0.683	0.733	0.716	0.633	0.667	0.700	0.680
	GJR(1,1)	0.667	0.667	0.667	0.683	0.683	0.733	0.766	0.667	0.667	0.700	0.690
BTC	GARCH(1,1)	0.667	0.667	0.616	0.600	0.716	0.716	0.683	0.683	0.667	0.716	0.673
	EGARCH(3,3)	0.650	0.667	0.633	0.600	0.700	0.733	0.700	0.667	0.667	0.750	0.676
	GJR(1,1)	0.683	0.650	0.600	0.583	0.683	0.716	0.667	0.667	0.683	0.700	0.663

Source: Own calculations using Matlab (2025).

4.4. Multivariate Volatility Modeling and Forecasting using the Temporal Fusion Transformer

Time series forecasting, a fundamental task in econometrics, finance, and machine learning, is used to predict future values based on historical data. A time series prediction problem of a target variable y_t , $t \in \{-\infty, +\infty\}$ can be described as the conditional expectation of y_{T+h} , given the information I_T of the time points up to T , and $T + h$ is the h -th prediction step in the future (Diebold, 1998; Granger & Newbold, 2014):

$$\hat{y}_{T+h} = E[y_{T+h}|I_T]$$

The conditional expectation in time series forecasting can now be modeled in different ways. This distinction can be made in several ways. For example, a common distinction is that between linear (e.g. autoregressive, autoregressive integrated, vector autoregressive models and generalized linear models) and non-linear models (e.g. autoregressive conditional heteroskedastic model and neural networks). Another distinction is between parametric models (e.g. ARIMA), which define the forecasting function explicitly as a combination of function parameters, which are linked together in a certain way, and non-parametric models (e.g. neural networks) where the form of the forecasting function is not explicitly defined but approximated from the data sample (e.g., Sezer et al., 2020). Univariate forecasting models vs. multivariate forecasting models is also a common distinction. A univariate time series consists of observations of a single variable recorded over time. The key assumption is that the future values of the series depend solely on its past values, without considering any external or additional explanatory variables (e.g., Hamilton, 2020). Mathematically, a univariate time series model can be expressed as (Diebold, 1998; Granger & Newbold, 2014):

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}, \varepsilon_t)$$

where $f(\cdot)$ is a function to model the dependency structure of the lagged target variable $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ and a disturbance variable ε_t . The ARIMA model class (Box & Jenkins, 1970) is a typical example of a univariate times series model and widely practically applied. A multivariate time series consists of multiple variables recorded over time that may influence each other (Lütkepohl, 2005). Instead of predicting a single variable solely, were predicting a $k \times 1$ vector $\mathbf{y}_{t-i} = (y_{1,t-i}, y_{2,t-i}, \dots, y_{k,t-i})'$ based on its past values and additional time-dependent variables. This can be represented as:

$$\mathbf{y}_t = g(\mathbf{y}_{t-1}, \dots, \mathbf{y}_{t-p}, \mathbf{x}_{t-1}, \dots, \mathbf{x}_{t-q}, \boldsymbol{\varepsilon}_t)$$

where $g(\cdot)$ is a function to model the dependency structure of the lagged target variables $\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \dots, \mathbf{y}_{t-p}$, with $\mathbf{y}_{t-i} = (y_{1,t-i}, y_{2,t-i}, \dots, y_{k,t-i})'$ for the exogenous variables $\mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \dots, \mathbf{x}_{t-q}$ with $\mathbf{x}_{t-j} = (x_{1,t-j}, \dots, x_{m,t-j})'$ that influence y_t and a disturbance variable $\boldsymbol{\varepsilon}_t = (\varepsilon_{1,t}, \dots, \varepsilon_{k,t})'$. Common multivariate models are the Vector Autoregressive Models (Sims,

1980), Dynamic Factor Models (Geweke & Singleton 1981; Stock, 1989), Vector Error Correction Models (Engle & Granger, 1987).

The neural networks I'm focusing on here, in particular the Temporal Fusion Transformer, can be implemented both as a univariate and as a multivariate model. This depends on whether other features or explanatory time series are considered in addition to the actual time series of the target variables. Multivariate time series forecasting (MTSF) in financial markets demonstrates significant advantages over univariate approaches through its ability to integrate multiple data streams and uncover complex interdependencies among financial variables (Lütkepohl, 2005; Tsay, 2013). By leveraging the correlations among different financial indicators, MTSF enhances the robustness of forecasts, particularly in volatile market environments. One key advantage of MTSF is its capacity to improve prediction accuracy by utilizing interconnected financial time series. Zhao et al. (2019) discussed multivariate directed graphical models that allow for the integration of one financial series as a predictor for others, effectively capturing the causal relationships often present in financial markets. This modeling approach is particularly beneficial for portfolio management, where the interplay between various asset classes significantly determines overall performance. The methodology employed by these models demonstrates how MTSF can uncover relationships that univariate models might overlook, thereby yielding superior forecasting results (Kolm et al., 2014). Additionally, MTSF approaches can incorporate advanced techniques such as machine learning and deep learning frameworks to harness the complexities of financial data (Ahmed et al., 2010; Lim et al., 2021). Kulendran & Witt (2003) conducted a comparative analysis illustrating how various multivariate and univariate models affect forecasting performance; their work emphasized that multivariate models often yielded more reliable outcomes when predicting different types of demand. This aligns with findings from Adebiyi et al. (2014) who articulated the practical applicability of ARIMA models in forecasting stock prices, noting that multivariate techniques generally lead to more robust forecasts compared to univariate counterparts.

The inclusion of measures like realized volatility in MTSF further enhances its predictive capability. Research by Hansen et al. (2014) discusses the use of realized measures of volatility within a multivariate framework as a strategy to improve the accuracy of forecasts, subsequently elevating risk assessment in financial market applications. Their findings underscore the importance of combining multiple perspectives on volatility to influence investment decisions effectively. Moreover, combining forecasts from multiple indicators can yield improved accuracy and reliability. Studies have shown that integrating various financial metrics, including liquidity and profitability indicators through multivariate adaptive regression splines, can enhance predictive precision compared to univariate models that may focus solely on stock prices (Shah et al., 2019). This ability to incorporate diverse financial factors provides a holistic view that is invaluable for decision-makers, further establishing multivariate forecasting as a preferred method in complex financial contexts. In conclusion, multivariate time series forecasting in financial markets provides notable advantages over univariate methods by leveraging interdependencies among multiple time series, employing advanced machine learning techniques, utilizing realized

volatility measures, and fostering improved accuracy through combined forecasting strategies. These attributes solidify MTSF as an essential tool for navigating the intricacies of financial markets effectively. The exploration of whether multivariate forecast models outperform univariate models is fundamentally intertwined with the structure of the conditional covariance matrix of the variables in question. A multivariate approach leverages the joint dynamics among multiple variables, potentially leading to more accurate forecasting outcomes.

4.4.1. Dynamic correlation of the financial assets

Dynamic correlation modeling of financial time series has gained considerable traction in financial econometrics due to its ability to capture time-varying relationships between asset returns. The foundation of such modeling lies in the extension of traditional univariate approaches, particularly Autoregressive Conditional Heteroskedasticity (ARCH) models (Engle, 1982), to multivariate frameworks. Engle's development of the Dynamic Conditional Correlation (DCC) model marked a significant advancement, allowing for more flexible estimates of variance and correlations across multiple asset returns (Engle, 2001; Engle & Sheppard, 2001). The DCC model operates under the premise of estimating univariate GARCH models for each asset while simultaneously modeling the correlations through a dynamic process. This framework accommodates the complexities of multivariate asset interactions by dynamically capturing the conditional correlations associated with varying states of market volatility (Engle, 2001; Engle & Sheppard, 2001). Evidence suggests that the DCC framework is advantageous for applications in portfolio management, financial risk assessment, and understanding complex market phenomena like financial contagion during crises (Kocaarslan et al., 2018; Rehman, 2016; Serletis & Xu, 2016).

Furthermore, recent research emphasizes the application of multivariate GARCH models to understand spillover effects in markets, particularly during times of heightened volatility. For example, the interaction between the crude oil market and financial markets is often analyzed through these models to gauge how shocks in one market affect others. This capability makes DCC models an essential tool in capturing co-movements and spillover effects during turbulent periods (Aduda et al., 2018; Rehman, 2016; Serletis & Xu, 2016). Additionally, the interplay between volatility and correlation, particularly in crisis scenarios, has highlighted how correlations between markets may shift based on external shocks (Kocaarslan et al., 2018; Thao et al., 2013). Recent developments have also integrated machine learning approaches into dynamic correlation modeling. This shift is recognized for its potential to better capture nonlinear relationships in financial time series without imposing strict parametric forms on the data, thereby allowing for richer dynamic structures (Bejger & Fiszeder, 2021; Wang et al., 2017). Such methodologies hope to overcome some limitations inherent in traditional econometric approaches by leveraging data-driven insights to model volatility and correlations dynamically. In conclusion, dynamic correlation modeling provides a robust framework for analyzing financial time series, particularly through the DCC approach. It facilitates a deeper understanding of time-varying relationships among asset returns, essential for risk management and forecasting in complex financial

environments. The continued evolution of these models, including the integration of machine learning techniques, promises to enhance their predictive capabilities and applicability in real-world financial scenarios.

The correlation among variables is crucial in enhancing the explainability and performance of forecasting models across diverse disciplines, including financial econometrics. Correlation drives the model's ability to accurately predict outcomes by establishing links and dependencies among predictors, which is foundational to forecasting accuracy and interpretability. Models have successfully applied this fact: ARIMA models (e.g., Box & Jenkins, 1970; Engle & Granger, 1987) and Vector Autoregressive Models (e.g., Lütkepohl, 2005). Another concept is the cointegration theory (Engle & Granger, 1987): if two or more non-stationary time series are cointegrated (i.e., share a long-term equilibrium relationship), their correlation structure enables better long-term forecasting. But there are also challenges and limitations. High correlation does not always imply causality or reliable predictability, particularly in financial markets where non-stationarity can distort relationships (Phillips & Perron, 1988). This is called spurious correlations. Further, financial data often exhibit regime shifts where correlations change over time (time-varying correlations), reducing the reliability of forecasts (Diebold & Rudebusch, 1996). Finally, correlation captures only linear relationships, whereas financial markets often exhibit nonlinear dependencies that traditional models might overlook. Machine learning approaches (e.g., neural networks) can sometimes improve forecasting beyond correlation-based methods. Before I used more sophisticated methods for modeling and forecasting the volatility of the financial instruments, I analyzed empirical dynamic correlation patterns for the volatility over time. Therefore, I calculated the pairwise exponential weighted Pearson correlation with a halflife of 10 for every financial instrument in my research scope, and also the Geopolitical Risk Index (GPR) and the Volatility Index Future (VIX). I will only discuss the most important findings here. The complete results can be found in Appendix 2. With this analysis, I wanted to find out which instruments are high correlated and thus are able to contribute information to the respective h-step ahead forecast. I also wanted to test whether the correlation between two instruments during the periods of geopolitical shocks differed significantly from the correlation of the entire data set.

When one looks at the average correlations over the complete dataset, the following stands out. The S&P 500 and NASDAQ show the highest correlation over all financial instruments (86.73%). This is followed by S&P 500 and VIX Future (37.55%), NASDAQ and VIX Future (32.60%), Gold and 10yr US-Treasury Bond (21.74%), S&P 500 and 10yr US-Treasury Bond, Nikkei and Hang Seng (18.34%), Nikkei and VIX Future (17.79%), NASDAQ and 10yr US-Treasury Bond (17.27%). All further correlations decrease more and more, which is why I will not present a complete representation of the results here. The Geopolitical Risk Index shows no correlation with the major financial stocks over the entire data set: S&P 500 (2.61%), NASDAQ (2.72%), Nikkei (0.06%), Hang Seng (2.83%), gold (6.20%), EUR-USD (-3.14%), Brent Crude Oil (3.22%), US-Treasury Bonds (2.51%) and Bitcoin (-2.63%).

However, as one can see, there is a correlation between the volatility of the financial instruments in my research scope. That may be suitable for the advantage of a multivariate

forecasting model. I now want to look at which instruments were particularly strongly correlated during the geopolitical shocks, and also which volatility correlation increased during the analyzed geopolitical shocks. This will give us an indication of the explanatory power of the respective financial instrument in the multivariate forecasting model in the respective periods of shocks.

Overall, one can see that the average dynamic correlation between the various financial stocks is higher during periods of geopolitical shock. In 40 of 55 (72.72%) analyzed correlation pairs, the average dynamic correlation during COVID-19 was statistically significantly higher than the overall average. I found the highest correlation between volatility for S&P 500 vs. NASDAQ (95.01%), S&P 500 vs. VIX (65.63%), NASDAQ vs. VIX (61.59%), S&P 500 vs. US T-Bonds (54.76%), US T-Bonds vs. VIX (54.60%), Nikkei vs. Hang Seng (51.89%), NASDAQ vs. US T-Bonds (50.41%). The largest increases in relation to the total average were: 10yr US T-Bonds vs. VIX (+35.71%), S&P 500 vs. 10yr US T-Bonds (+34.88%), Nikkei vs. Hang Seng (+33.55%); NASDAQ vs. 10yr US T-Bond (+33.14%), NASDAQ vs. VIX (+28.99%), S&P 500 vs. VIX (+28.08%), EUR-USD vs. VIX (+26.86%) and Hang Seng vs. VIX (+23.47%).

During the period of the Russian attack, one can see that there are fewer statistically significant deviations of the average dynamic correlation from the overall average. Only 33 out of 55 (60.00%) in total. So, the contagion effect was weaker than in the period of COVID-19. The highest correlations between volatility between the financial instruments were S&P 500 vs. NASDAQ (86.41%), GPR vs. VIX (60.51%), Nikkei vs. EUR-USD (40.46%), Gold vs. VIX (35.61%), Gold vs. GPR (33.41), S&P vs. VIX (31.91%). The largest increases during this time were: GPR vs. VIX (+50.49%), Nikkei vs. EUR-USD (+29.89%), Gold vs. VIX (+28.28%), Gold vs. GPR (+27.22%).

During the period of the Hamas attack, I found 25 out of 55 (45.45%) statistically significant deviations between average dynamic correlation of the period to the overall average. Therefore, in this period, the contagion effect was the smallest. The highest correlations in volatility were found at: S&P 500 vs. NASDAQ (91.39%), GPR vs. VIX (34.91%), Hang Seng vs. EUR-USD (29.86%), NASDAQ vs. VIX (29.07%), S&P 500 vs. VIX (28.47%), Gold vs. Brent Crude Oil (28.19%), S&P 500 vs. Hang Seng (-25.42%), S&P 500 vs. EUR-USD (-25.44%) and NASDAQ vs. EUR-USD (25.85%). The largest increases in absolute values were: S&P 500 vs. Hang Seng (+33.16%), NASDAQ vs. EUR-USD (30.19%), Nikkei vs. Gold (29.71%), S&P 500 vs. Gold (27.46%), NASDAQ vs. Hang Seng (26.66%), GPR vs. VIX (24.89%), Nikkei vs. Hang Seng (24.73%) and S&P 500 vs. Brent Crude Oil (22.86%).

Now, I want to interpret some of the patterns. The high correlation of the volatility between the S&P 500 and the NASDAQ reflects several underlying factors that can be attributed to their structural, behavioral, and economic characteristics. Both indices comprise large-cap stocks driven by similar market dynamics. The S&P 500 includes 500 of the largest companies in the U.S., while the NASDAQ Composite is heavily weighted toward technology and growth-oriented companies, many of which also appear in the S&P 500. This overlap contributes to their price movements being driven by comparable factors, including macroeconomic indicators and investor sentiment. Research has shown that daily return correlations between the S&P 500 and the NASDAQ are

significant, with studies indicating that their return series exhibit a consistent pattern during periods of market stress or boom. For instance, it has been observed that extreme return clusters often overlap among the S&P 500 and NASDAQ, indicating that both indices tend to respond similarly to extreme market conditions (Sankaran et al., 2012). Additionally, correlations tend to be elevated in periods characterized by high volatility, driven by common risk factors affecting the technology and consumer discretionary sectors predominantly listed on both indices (Flood & Rose, 2005).

Moreover, external economic variables, such as interest rates set by the Federal Reserve, impact both indices similarly. Research shows that increases in the Federal Funds Rate negatively influence stock prices across U.S. markets, including both the S&P 500 and the NASDAQ, indicating that shared macroeconomic influences contribute to their correlation (Zhao, 2023). Furthermore, the integration of these markets suggests that the covariance structure remains stable over time, indicating a strong interconnectedness between the two indices that affects their performance in response to economic changes (Qu & Perrón, 2013). I also found that the volatility of several assets is correlated with the VIX index future or increased strongly during the periods of geopolitical shocks. The VIX index is often interpreted as a barometer of fear in the market, reflecting significant predictive power for stock market returns, economic activity, and measures of financial instability. This characteristic comes into sharper focus during crises, such as the COVID-19 pandemic, when investor uncertainty escalates dramatically, resulting in closely correlated movements between the VIX and stock market indices (Dima et al., 2021). For instance, during the pandemic, the VIX indicated increased uncertainty surrounding future market conditions, echoing findings that noted market correlations tend to spike in response to VIX changes during crises (Badshah, 2017). During periods of heightened volatility, shocks to the VIX have been shown to induce increases in cross-country stock market correlations. For instance, one study found that a shock to the VIX led to increased correlations among major international stock markets, suggesting that global markets react similarly to perceived risk, thereby reinforcing the linkage driven by the VIX (Ceylan, 2021). Research has also indicated that volatility spillovers exist between the VIX and both developed and emerging markets, with the VIX acting as a conduit for information dissemination regarding market conditions (Badshah, 2018). This is particularly relevant when considering the variances in investor behavior and capital flow dynamics during financial stress, where VIX movements lead to significant implications for cross-market interactions (Vartanian & Neto, 2023). So, empirical evidence supports the notion that the correlation between the VIX and various financial markets intensifies during periods of stress, significantly affecting both equity and bond performance. Studies demonstrate that as volatility increases, so does the potential for interconnected market movements, which skews the implications for risk management strategies and underscores the critical role of the VIX in the global financial landscape during turbulent times (Bekaert et al., 2013; Smales, 2016). Moreover, the role of the VIX as a determinant of market behavior extends to various asset classes. Research by Tuna, 2022) investigates its interaction with commodities, specifically how fluctuations in the

VIX index correlate with oil and gold prices, impacting markets such as Turkey's BIST 100 during COVID-19 pandemic-induced economic disruptions.

The volatility relation between gold and fixed income assets, like U.S. Treasury Bonds, is multifaceted, influenced by speculative behaviors, macroeconomic conditions, and systemic shocks. Both assets are traditionally viewed as safe havens, yet their volatility dynamics can reflect different market sentiments and economic conditions. From a macroeconomic perspective, factors such as fiscal deficits and monetary policy have been shown to significantly influence the yield curve of U.S. Treasury bonds. Research indicates that increases in fiscal deficits correlate with higher Treasury yields, which may simultaneously lead to increased volatility in gold prices, creating an interconnected dynamic (Jiang et al., 2024). Longstaff's examination of liquidity premiums in Treasury bonds illustrates how investor sentiment shifts during crises, impacting Treasury bond prices relative to gold (Longstaff, 2004). The relationship between the volatility of US Treasury bonds and other asset classes is influenced by various market dynamics. In times of economic or financial crisis, investors sell risky equity investments and invest in safe government bonds. This increases demand for bonds, their yields fall, and volatility in the bond market often decreases while equity markets experience greater fluctuations (flight to safety). Rising interest rates make government bonds more attractive, which diverts capital from equities. This can lead to increased volatility in the stock market. Conversely, falling interest rates lead to lower bond yields and increased investment in equities. In addition, central banks influence markets through interest rate policy and liquidity measures. Announcements by the Federal Reserve can cause both bonds and equities to move sharply.

With this analysis, I was able to show the following. First, I empirically examined the relationships of volatility and their strength for the financial instruments I analyzed. This gives us an indication of which variables may have an explanatory power for other variables in the multivariate analysis. Thus, the NASDAQ may have a large influence on the S&P 500 forecast and vice versa. Also, the VIX index future may have explanatory power for several assets in the research scope. The same with the US Treasury Bonds, which volatility is also correlated with other asset classes. In addition, I was able to demonstrate the contagion effect for the periods of geopolitical shocks. As a result, the correlation of the volatility between financial instruments increases during periods of high volatility and volatility spillovers occur. For the subsequent analysis of forecasting performance.

4.5. Multivariate direct multi-step out-of-sample Volatility Forecasting using the Temporal Fusion Transformer

In the following, I discuss the results of the Temporal Fusion Transformer (TFT) forecasting examination. I trained three different TFT models with nine financial instruments in research scope (S&P 500, NASDAQ, Nikkei 225, Hang Seng, Gold, Brent Crude Oil, 10yr US-Treasury Bond, EUR-USD exchange rate and Bitcoin) and then generated predictions for the volatility of these financial instruments in periods of geopolitical shocks: COVID-19, the Russian invasion of Ukraine and the attack of the Hamas in Israel. Therefore, I used three different

modeling approaches for the training and forecasting: a univariate and a multivariate modeling approach, as well as a multivariate modeling approach with a volatility regime switching element. A rolling window approach is used to train the model and generate robust volatility forecasting results. In the univariate model approach, I modeled and forecasted the volatility of each financial instrument separately and only the historical data of the target variable is incorporated into the model, analogous to the GARCH forecasting. I compared the forecasting performance between the GARCH approach and the results of a TFT model. In doing so, I predicted the volatility values of all financial instruments under consideration simultaneously. As explanatory variables, I included the historical values of these financial instruments within a window of 250 days in the model training. In doing so, I included the mutual influences of the financial values (spill-over effects) on the explanatory power of the volatility forecast for the respective target variables. By comparing the forecasting performance between the univariate and multivariate TFT model approach, I wanted to analyze whether the complexity of implementing a multivariate TFT model is justified by the advantage in performance.

Then I checked whether the performance of the multivariate model can be improved by adding an additional explanatory variable for the TFT that indicates whether a geopolitical shock is present. I'll explain below how exactly this variable is defined in the setup. I generated forecasts for 1 to 10 steps ahead. And a special feature, the TFT can generate direct k -step-ahead forecasts, instead of creating them from the $(k - 1)$ -step-ahead (recursive) forecast. Direct and recursive forecasting differ in how multi-step-ahead predictions are generated: iterative methods produce forecasts recursively using previous predictions, while direct methods estimate a separate model for each forecast horizon (Chevillon, 2007). Iterative forecasts are more efficient under correct model specification but suffer from error accumulation at longer horizons, whereas direct forecasts are more robust to misspecification and tend to perform better in nonlinear or unstable environments, like volatility forecasting (Marcellino et al., 2006). Empirical studies suggest that while iterative methods are preferable for short horizons, direct methods yield superior accuracy for longer-term forecasts, particularly in macroeconomic and financial contexts (Tashman, 2000).

The evaluation of the forecasting performance is done by the Mean Absolute Percentage Error (SMAPE) for the TFT and GARCH-type models, as well as the directional accuracy for the GARCH-type models. Furthermore, I generated the forecasts for a defined hyperparameter grid of the TFT and then calculate the average SMAPE values from them. This gives the results additional robustness, as they cannot be generated with just one specific TFT configuration. In a further analysis, I examined whether a specific hyperparameter configuration of the TFT delivers significantly better forecasting performance than the average.

4.5.1. Analysis of multi-step out-of-sample forecasting performance

In the first analysis of the out-of-sample forecast performance of the Temporal Fusion Transformer (TFT), I examined the direct multi-step out-of-sample forecast performance, i.e. the ability of the neural network to generate correct forecasts several direct steps into the future. I generated and analyzed this for the financial instruments in scope. I used 250 trading days of the

financial instruments under consideration as training for the Temporal Fusion Transformer. I used 200 trading days only for training and 50 trading days for validating the weights of the neural network. After training the neural network, I generated predictions within the time periods for COVID-19 (02/03/20 - 04/30/2020), the Russian invasion of Ukraine (02/01/2022 - 29/04/2022) and for the period around the Hamas attack on Israel (10/01/2023 - 12/29/2023) for the volatility of financial instruments. The forecasts were each 1 to 10 trading days into the future. I used the rolling window approach, whereby for each step of the rolling window in the above-mentioned period, the Temporal Fusion Transformer was first trained with a combination of different hyperparameters (attention head = 1, 4, 8 and hidden size = 16, 32, 64) and then the predictions for the next 1 to trading days were generated. The rolling window was then shifted by one trading day.

For this analysis, I generated and analyzed multivariate forecasts. I predicted the volatility of all financial instruments simultaneously and included them in the training accordingly. In addition to the variables to be predicted, which are both, exogenous (as an explanatory variable for another variable to be predicted) and endogenous (as a variable to be predicted), I included the volatility index (VIX) future and the geopolitical risk index (GPR) as explanatory variables. The formal presentation of the approach can be found above. I evaluated the forecasting performance using the Mean Average Percentage Error method (SMAPE). As a benchmark, I used the results of the out-of-sample forecast using different GARCH class models (Table 25-27). I evaluated the significance of the advantage of the Temporal Fusion Transformer compared to the best performing GARCH-class model using a adjusted Diebold-Mariano test (DM*-test) and a bootstrapping procedure. Accordingly, the null hypothesis is that there is no statistically significant difference between the forecasting performance of the Temporal Fusion Transformer and the forecasting performance of the best performing GARCH model. The phases I examined during the COVID-19 pandemic, the Russian attack on Ukraine and the Hamas attack on Israel were characterized by the fact that there was a strong regime switch in volatility and volatility was at a higher level from this point onwards. This can be seen for all financial instruments analyzed (Figure 5-8 and Table 8-11).

I was able to demonstrate this in the analysis of the sectors of the MSCI World Index (see section 4.3.). The challenge for the methods used to predict volatility was therefore to ensure that the method quickly takes into account the regime switch in volatility and can adjust to the higher level of volatility. In summary, I gained the following insights from analyzing the results. The Temporal Fusion Transformer was able to outperform the GARCH class approach for predicting volatility in all three time periods of geopolitical shocks examined, for almost all financial instruments in all forecast horizons. I will now discuss the results systematically and will do so separately according to the examined time periods of geopolitical shocks. The TFT is better at forecasting small forecast horizons and the forecast performance for larger horizons decreases in comparison. Thus, the TFT is significantly better than the GARCH class models for small forecast horizons and relatively less good for larger forecast horizons.

4.5.1.1. The COVID-19 period

I start with the period within the COVID-19 pandemic (02/03/20 - 04/30/2020). This period was characterized by extreme uncertainty about the impact and consequences of the pandemic and a time of high volatility for all financial instruments considered. First, I compare the mean values for the individual financial instruments over time generated by the Temporal Fusion Transformer (Table 33) with those of the best-performing GARCH-type model (Table 25).

First, I would like to begin with a brief description of the results tables. These are structured in a similar way throughout the rest of the text. The respective columns of the tables contain the financial instruments under consideration (S&P 500, etc.). The rows contain the SMAPE results for the respective forecast horizon and financial instrument evaluated. Accordingly, the SMAPE for the forecasting performance of the TFT for the instrument S&P 500 and the forecast horizon = 1 is equal to 0.6699. The asterisk indicates that there is a statistically significant difference (via adjusted Diebold-Mariano test statistic at a 5% significance level) in the forecasting performance between the TFT model and the best GARCH-type model for the financial instrument and forecast horizon. The bottom row shows the aggregation across the individual forecast horizons using the average.

For the S&P 500 stock index, the average SMAPE over the ten forecast horizons was 0.721. The best performing GARCH model was an EGARCH (3,4) and achieved an average SMAPE of 0.924. The null hypothesis that the average performance of the TFT and the GARCH class model is the same could be rejected with a DM* statistic of 20.218 (p-value < 0.000). Moreover, one could observe this result for all prediction horizons. In addition, it is noticeable that the forecasting performance of the TFT decreases slightly with increasing forecast horizon, while it is rather constant for the EGARCH model. This suggests that the TFT is better able to reflect the variability of volatility, while the EGARCH model is smoother and tends to reflect the average volatility. For the NASDAQ stock index, the best GARCH class model is also the EGARCH(3,4) model, with an average SMAPE of 0.941. In contrast, the TFT achieved an average performance of 0.854. The DM*-test statistic was 11.384 with a p-value < 0.000. Accordingly, the null hypothesis could be rejected. However, when looking at the various prediction horizons, it is noticeable that the EGARCH(3,4) did not perform statistically significantly worse than the TFT for prediction horizon 5.

For the Nikkei 225 stock index, the average SMAPE of the TFT was 0.703. The best GARCH class model, an EGARCH(3,3), achieved an average SMAPE of 0.919. The DM*-test statistic was 6.195 with a p-value < 0.000. This indicates that the null hypothesis could be rejected. When looking at the forecast performance of the individual forecast horizons for the TFT, it is noticeable that the forecasts for small forecast horizons are better than for larger ones. In comparison, the SMAPE for the EGARCH model is rather stable. For h=1, for example, the SMAPE for the TFT is 0.578 and for the EGARCH model 0.877, a difference of 0.299. For h=10, the SMAPE for the TFT is 0.841 and for the GJR-GARCH 0.919, a difference of 0.042. This again indicates that the TFT is much better at predicting the immediate future than the EGARCH model. For larger prediction horizons, the performance of the TFT approaches that of the EGARCH

model. One possible explanation for this is, that the TFT, in contrast to EGARCH, has a stronger weighting on the immediate historical data, while EGARCH smoothes more strongly and predicts the average volatility. This is a significant difference, particularly in the case of regime-switching volatility. The average forecasting performance of the TFT for the Hang Seng stock index was 0.613. The best GARCH class model was the GJR-GARCH(1,1) model, and achieved a SMAPE of 1.390. The DM*-test statistic was 5.889 (p-value < 0.000). Thus, the null hypothesis was clearly rejected and the TFT had a significantly better forecasting performance. Moreover, the null hypothesis could be rejected for all forecast horizons. There was no such clear difference in the prediction of the volatility of the commodity gold. The TFT achieved an average SMAPE of 0.707. The best GARCH model, a GJR-GARCH(1,1), achieved a SMAPE of 0.873. However, the DM*-test statistic was 18.657 (p-value < 0.000), indicating, that the null hypothesis could be rejected. It is again noticeable that the TFT loses forecasting performance with the forecast horizon, while the SMAPE of the GJR-GARCH(1.1) remains rather stable.

For Brent Crude Oil, the average SMAPE was 0.622. The best GARCH class model was an EGARCH(3,3) with an average SMAPE of 0.933. The DM* statistic was 24.247 (p-value < 0.000). This clearly rejected the null hypothesis. Again, one can see a significantly better performance of the TFT for short prediction horizons than for longer ones, in contrast to the EGARCH model. The average performance of the TFT for predicting the volatility of 10-year US government bonds was 0.7900. The best GARCH model, a GJR-GARCH(1,1), achieved an average performance of 1.391. The DM* statistic was 4.627 (p-value < 0.000). This clearly rejected the null hypothesis. I was able to show that the TFT achieves a significantly better prediction performance on average than the GARCH model. With regard to the individual forecasting horizons, however, one can see a different picture compared to the results of other instruments. The forecasting performance tends to improve as the forecast horizon increases. This can be explained by the dynamics of volatility in the period under review. Compared to the other financial instruments, with the exception of Bitcoin, 10-year US government bonds recorded the strongest volatility in March 2020 (Figure 6). However, this decreased again in April. As a result, the TFT may be better able to make forecasts with a longer forecast horizon, as these are smoothed to a greater extent.

The average performance of the TFT for the euro-dollar exchange rate was 1.204. The best GARCH model, a GJR-GARCH(1,1), achieved 0.813. The DM* statistic was -13.271 (p-value = 2.000). Thus, the null hypothesis could not be rejected. However, not in the sense that the TFT outperformed the GARCH model, but vice versa. The GJR-GARCH achieved significantly better forecasts for this financial instrument. Looking at the individual forecast horizons, this picture remains the same - there is no forecast horizon for which the TFT outperformed the GJR-GARCH model. One possible explanation for this is, that the dynamics of volatility for the exchange rate are rather low compared to other financial instruments (Figure 5-6) and the prediction of smoothed volatility by the GJR-GARCH was more suitable than the TFT. The average performance of the TFT for Bitcoin was 0.929. The best GARCH class model was an EGARCH(5,2) model. Its performance was 2.150. The DM* statistic was 6.144 (p-value < 0.000) and thus the null

hypothesis that the TFT and the EGARCH model achieved the same performance could be clearly rejected - the TFT achieved a significantly better performance. There is no clear structure with regard to the performance for the prediction horizons. The best predicted horizon is $h=8$ with 0.862 and the worst is $h=5$ with 0.980. One also see no clear structure for the EGARCH model.

4.5.1.2. The period of the Russian attack on Ukraine

I continue the analysis with the period surrounding Russia's attack on Ukraine (02/01/2022 -04/29/2022). This period was also characterized by uncertainty about the effects and consequences of a European war, the possible entry of other warring parties, such as the USA, Great Britain and Poland as supporters of Ukraine, as well as China and Belarus on the Russian side and thus the development of a global conflict. Overall, it was shown that the volatility of the financial instruments under consideration remained on average below the volatility during the COVID-19 pandemic (Table 9-10) and showed less pronounced peaks in volatility. For the forecasting methods, this implies, that they had to predict less extreme events and variability. This is then also reflected in the performance, which was better during the period of the Russian attack than in the period of COVID-19 for all financial instruments considered, except for the Hang Seng index. The results can be found in (Table 34)

For the S&P 500 stock index, I found an average SMAPE of 0.519. The best GARCH model was an EGARCH (1,2) and achieved an average SMAPE of 0.748. The null hypothesis that the average performance of the TFT and the GARCH class model is the same could be rejected with the value of the DM*-test statistic of 14.329 ($p\text{-value} < 0.001$). Moreover, one can observe this result for all prediction horizons. In addition, it is noticeable that the forecasting performance of the TFT decreases slightly with increasing forecast horizon, which is similar for the EGARCH model. I observed similar results in the period of COVID-19. Furthermore, the null hypothesis could be rejected for all various forecast horizons. For the NASDAQ stock index, the best GARCH class model is an EGARCH(2,1), with an average SMAPE of 0.651, although this result is not significantly different from the result of the GJR-GARCH(1,1) model. In contrast, the TFT achieved an average performance of 0.449. The DM* statistic was 14.284 ($p\text{-value} < 0.001$). Accordingly, the null hypothesis could be rejected. The performance for all individual prediction horizons was also significantly different (better) than that of the EGARCH model. Furthermore, the performance of the EGARCH model tends to deteriorate for larger horizons, while it remained rather constant for the TFT.

For the Nikkei 225 stock index, the average SMAPE of the TFT was 0.673. The best GARCH class model, an EGARCH(1,1), achieved an average SMAPE of 0.904. The DM*-test statistic was 12.486 with a ($p\text{-value} < 0.001$). Thus, the null hypothesis could be rejected, and so for all forecast horizons. When looking at the forecast performance of the individual forecast horizons for the TFT, it is again noticeable that the forecasts for small forecast horizons are better than for medium horizons (for $h=1$ the SMAPE is 0.570 and for $h=5$ the SMAPE is 0.773) and the performance of larger horizons is also better than for medium horizons (for $h=10$ the SMAPE is 0.678). In comparison, the SMAPE for the EGARCH model is stable. The average SMAPE of the

TFT for the Hang Seng was 0.800. The best GARCH class model, a GJR-GARCH(1,1), achieved a SMAPE of 0.940 and the DM*-test statistic was 9.427 (p-value < 0.001). So, The null hypothesis could be rejected for the average SMAPE and for the SMAPE of the individual prediction horizons. Furthermore, one cannot derive a clear structure regarding the dependence of the forecast performance and the forecast horizon for both the TFT and the GJR-GARCH model. The best performance by the TFT was achieved for $h=7$ and the worst for $h=8$. For the commodity gold, I was able to recognize a clear effect between the prediction methods. The TFT achieved an average SMAPE of 0.450. The best GARCH model, a simple GARCH(1,1) model, achieved a SMAPE of 0.775. However, the DM* statistic was 30.640 (p-value < 0.001), meaning that the null hypothesis could be rejected. This also applies to all prediction horizons considered. For the performance of the TFT, depending on the forecast horizon, it is noticeable that the performance decreases for increasing h (up to $h=6$ and $h=7$) and then increases again.

For Brent Crude Oil, the average SMAPE was 0.548. The best GARCH class model was an EGARCH(3,5) model with an average SMAPE of 0.804. The DM*-test statistic was 33.912 (p-value < 0.001) and the null could be rejected clearly. I was also able to prove this for all prediction horizons. In addition, one can recognize better predictions for shorter prediction horizons. The average performance of the TFT for predicting the volatility of 10-year US government bonds was 0.5504. The best GARCH model, a GJR-GARCH(1,1), achieved an average performance of 0.811. The DM*-test statistic was 16.341 (p-value < 0.001). This clearly rejected the null hypothesis. The null hypothesis could also be rejected for the individual horizons. I was able to show that the TFT outperforms the GARCH model on average and for the individual forecast horizons. For this financial instrument, one can also see that the prediction of shorter forecast horizons achieves a better performance than for longer ones. Even during the period of the Russian attack, the volatility of the 10-year US government bonds was at a high level compared to the equity indices, with strong swings in the squared daily returns (Figure 7).

The average performance of the TFT for the euro-dollar exchange rate was 0.632. The best GARCH model, a GJR-GARCH(1,1) reached 0.811. The DM* statistic was 24.139 (p-value < 0.001). Thus, the null hypothesis could be rejected. Likewise, the null hypothesis could be rejected for each individual forecast horizon, showing that the TFT outperformed the GJR-GARCH at each forecast horizon. Furthermore, I recognized that both short and long forecast horizons could be predicted better than medium ones (e.g. $h=6, 7$). In addition to Brent Crude Oil and the 10-year government bonds, Bitcoin also recorded high volatility during this period, which, after strong spikes in February, gradually declined again during March and April (Figure 7). The average performance of the TFT for Bitcoin was 0.529. The best GARCH class model was a GJR-GARCH(1,1) model. Its performance was 2.263. The DM* statistic was 41.311 (p-value < 0.001) and thus the null hypothesis could be clearly rejected, i.e. the TFT achieved a significantly better performance. This also applies to the performance for the individual forecast horizons - the null hypothesis was rejected for all horizons. However, there is no clear structure with regard to the performance for the forecast horizons. However, long forecast horizons have a slightly better performance than short ones.

4.5.1.3. The period of the Hamas attack on Israel

I continue the analysis with the period around Hamas' terrorist attack on Israel (10/01/2023 - 12/29/2023). The terrorist attack carried out by Hamas on October 7, 2023 against Israeli civilians on Israeli soil represents one of the most serious escalations in the Israeli-Palestinian conflict. In addition to the intensification of hostilities between the two parties to the conflict, the attack also harbored the risk of global expansion, as the USA, as Israel's ally, may be drawn into a conflict with Iran and Russia. This could also lead to other Gulf states distancing themselves from Israel and the USA, which could have far-reaching effects on the oil price, which has a significant impact on energy costs and therefore also on many financial stocks. Overall, the volatility of the financial instruments under review remained on average below the volatility during the COVID-19 pandemic and the Russian attack (Table 11 and Figure 8) and showed less pronounced peaks in volatility. For example, the oil price, the financial asset with the highest volatility during this period, had a maximum daily loss of -5.8% (Table 11). In comparison, the value during COVID-19 had a maximum daily loss of -28.0% (Table 9) and the Russian attack had a value of -14.1%. For the forecasting methods, this means that they need to predict less extreme events and variability. The results can be found in Table 35.

For the S&P 500 stock index, the average SMAPE across the ten forecast horizons was 0.856. The best GARCH model was an EGARCH (1,1) and achieved an average SMAPE of 1.736. The DM*-test statistic was 19.974 (p-value < 0.001) and so, the null hypothesis could be rejected. In addition, the null hypothesis could be rejected for all individual prediction horizons. In addition, it is noticeable that the forecasting performance of the TFT decreases slightly with increasing forecast horizon, which is similar for the EGARCH model. Compared to the other periods, the performance was the lowest. For the NASDAQ index, the best GARCH class model was a GARCH(1,1), with an average SMAPE of 2.006. In contrast, the TFT achieved an average performance of 0.809. The DM*-test statistic was 13.448 (p-value < 0.001) and accordingly, the null could be rejected. The performance for all individual prediction horizons was also significantly better than that of the GARCH model. Furthermore, it could be seen that the performance of the GARCH model improves slightly for larger horizons, while it deteriorates slightly for the TFT. Overall, the performance of the TFT in this period is slightly better than for COVID-19, but worse than in the period of the Russian attack. For the Nikkei 225 stock index, the average SMAPE of the TFT was 0.777. The best GARCH class model, a GARCH(1,1), achieved an average SMAPE of 1.923. The DM*-test statistic was 11.651 (p-value < 0.001). Thus, the null could be rejected and the TFT achieved a statistically significantly better forecast performance. The null hypothesis could also be rejected for all forecast horizons. When looking at the forecast performance of the individual forecast horizons for the TFT, it is noticeable that the forecasts for small forecast horizons are better than for medium and larger horizons. This structure can also be seen in the performance of the GARCH model.

The Nikkei performed worse in this period than in previous periods. The average forecasting performance of the TFT for the Hang Seng index was 0.645. The best GARCH class model was the GJR-GARCH(1,1), and achieved an average SMAPE of 1.212. The DM*-test

statistic was 31.922 ($p\text{-value} < 0.001$). The null hypothesis was thus clearly rejected and the TFT had a significantly better forecasting performance. Moreover, the null could be rejected for all forecast horizons. Furthermore, one cannot derive a clear structure regarding the dependence of the forecast performance and the forecast horizon for both the TFT and the GJR-GARCH model. In contrast to other results, I was unable to detect a clear effect between the prediction methods for the commodity gold. The TFT achieved an average SMAPE of 1.332. After Bitcoin, gold was the worst performing financial asset during this period. The best GARCH model, an EGARCH(4,3) model, achieved a SMAPE of 1.494. The DM* statistic was 1.718 ($p\text{-value} = 0.058$), meaning that the null hypothesis could not be rejected. The null hypothesis could be rejected for the individual prediction horizons. For $h=7$ and $h=8$, however, the EGARCH model outperformed the TFT. In addition, the performance of the TFT decreased with increasing forecast horizon. The TFT thus achieved the worst average forecast for the volatility of gold in this period. For Brent crude oil, the average SMAPE of the TFT was 0.561. The best GARCH class model was an EGARCH(4,3) model with an average SMAPE of 0.845. The DM* statistic was 15.208 ($p\text{-value} < 0.001$). This clearly rejected the null hypothesis. I was also able to prove this for all prediction horizons. In addition, I was able to identify better predictions for shorter forecasting horizons.

The forecasting performance for Brent Crude Oil is slightly worse than in the period of the Russian attack and slightly better than in the period of COVID-19. Compared to other financial stocks, however, the performance is rather stable. The average performance of the TFT for predicting the volatility of 10-year US government bonds was 0.548. The best GARCH model, a GARCH(1,1), achieved an average performance of 0.757. The DM* statistic was 28.978 ($p\text{-value} < 0.001$). This clearly rejected the null hypothesis. The null hypothesis could also be rejected for the individual horizons. I was able to show that the TFT outperforms the GARCH model on average and for the individual forecast horizons. A systematic dependence of the performance on the forecast horizon cannot be recognized. Compared to the other periods, government bonds recorded less volatility in the period around the Hamas attack (Figure 8) and the performance of the TFT is best in this period.

The average performance of the TFT for the euro-dollar exchange rate was 0.503. The best GARCH model, a GJR-GARCH(1,1) reached 0.811. The DM* statistic was 30.408 ($p\text{-value} < 0.001$). Thus, the null hypothesis could be rejected. Likewise, the null hypothesis could be rejected for each individual forecast horizon, showing that the TFT outperformed the GJR-GARCH model at each forecast horizon. Furthermore, one recognize that both short and long forecast horizons could be predicted better than medium ones (e.g. $h=6, 7$). In addition to Brent Crude Oil and the 10-year government bonds, Bitcoin also recorded high volatility during this period, which, after sharp spikes in February, gradually declined again over the course of March and April (Figure 8). The average performance of the TFT for Bitcoin was 0.684. The best GARCH class model was a GJR-GARCH(1,1) model. Its performance was 2.263. The DM* statistic was 17.210 ($p\text{-value} < 0.001$) and thus the null hypothesis could be clearly rejected. So, again, the TFT outperformed the

GJR-GARCH. This also applies to the performance for the individual forecast horizons - the null hypothesis was rejected for all horizons.

This study systematically evaluates the direct multi-step out-of-sample forecasting performance of the Temporal Fusion Transformer (TFT) relative to traditional GARCH-class models across three periods of pronounced geopolitical shocks: the COVID-19 pandemic, the Russian invasion of Ukraine, and the Hamas attack on Israel. All periods were marked by structural breaks in volatility, with each regime exhibiting elevated and unstable volatility patterns. Forecasts were generated for ten forward-looking trading days using a rolling-window approach, and the models were evaluated using Symmetric Mean Absolute Percentage Error (SMAPE) and Diebold-Mariano tests with bootstrapping. Across all examined periods and asset classes, several robust patterns emerged from the analysis. First, I found a superior forecast accuracy of the TFT against the best-performing GARCH model across most assets and forecast horizons. The null hypothesis of equal predictive accuracy was rejected at the 1% level in the vast majority of cases, indicating statistically significant gains in forecast precision using the TFT. I also found a short-horizon Superiority of the TFT, as the TFT performed best at short forecast horizons ($h = 1-3$). Forecast accuracy generally declined with longer horizons, whereas GARCH models showed more stable, horizon-insensitive performance. This supports the hypothesis that the TFT's architecture places greater emphasis on recent, high-frequency patterns—especially relevant under regime-switching volatility.

During the COVID-19 period, the TFT significantly outperformed GARCH models for nearly all assets, with especially strong gains in equity indices. An exception was the EUR/USD exchange rate, where the GJR-GARCH model outperformed the TFT. For the period of the Russian war, the TFT outperformed across all instruments and horizons. And finally, during the Hamas attack, the TFT continued to outperform in most cases, although differences were less pronounced. Notably, GARCH models outperformed the TFT in forecasting the Gold volatility, suggesting that the TFT's learning-based structure may be less effective in relatively stable markets with complex or erratic signals. Moreover, I analyzed the asset class sensitivity. Here, I found, that the TFT for equities and oil consistently outperformed the GARCH-class models across all periods and horizons. For fixed income, the TFT was especially effective during periods with sharp volatility reversals (e.g., March 2020) and outperformed the GARCH by wide margins. FX and Crypto: Results were more heterogeneous. In some cases, like EUR/USD (COVID-19), GARCH delivered superior performance, likely due to its robustness in low-volatility regimes. Further, I found that the forecast horizon dynamics showed, that the Performance of the TFT generally was declining with increasing horizon length but remained statistically superior in most settings. Some assets (e.g., Nikkei 225 and EUR-USD) exhibited non-monotonic horizon-performance patterns, particularly in less volatile periods, suggesting a complex interaction between horizon length, market conditions, and model responsiveness. This allowed me to answer my research question of whether the TFT can outperform the best GARCH-type model (Research Question 8).

Table 33: SMAPE for multi-step out-of-sample multivariate forecasts during COVID-19 (02/03/2020 – 04/30/2020). SMAPE values are averages from different hyperparameter combinations (attention head = 1, 4, 8; hidden size = 16, 32, 64). Adjusted Diebold-Mariano test (DM*) for significant difference to the best performing GARCH model (Table 25) for given financial instrument and forecast horizon on 5% level (*).

Forecast Horizon	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
1	0.6699*	0.8228*	0.5786*	0.6281*	0.6576*	0.5747*	0.8043*	1.0803	0.9790*
2	0.7083*	0.8404*	0.5632*	0.6406*	0.6984*	0.6220*	0.8148	1.1819	0.9225*
3	0.7182*	0.8069*	0.6185*	0.6405*	0.6779*	0.5995*	0.8585*	1.2188	0.9168*
4	0.7222*	0.8204*	0.6392*	0.6609*	0.7328*	0.5949*	0.8230*	1.2125	0.9585*
5	0.6906*	0.8829*	0.7180*	0.5911*	0.7135*	0.6077*	0.8042*	1.2535	0.9804*
6	0.7542*	0.8765*	0.8538*	0.6162*	0.7421*	0.6461*	0.7964*	1.2236	0.8879*
7	0.6908*	0.9137	0.7376*	0.6010*	0.7536*	0.6332*	0.7532*	1.2906	0.9439*
8	0.7706*	0.8831*	0.6776*	0.6035*	0.6617*	0.6216*	0.7364*	1.2559	0.8621*
9	0.7380*	0.8643*	0.8095*	0.5547*	0.7078*	0.6530*	0.7506*	1.2067	0.9386*
10	0.7489*	0.8328*	0.8413*	0.5991*	0.7316*	0.6680*	0.7589*	1.1162	0.9047*
Average	0.7212*	0.8544*	0.7037*	0.6136*	0.7077*	0.6221*	0.7900*	1.2040	0.9294*

Source: Own calculations using Python PyTorch (2025).

Table 34: SMAPE for multi-step out-of-sample forecasts around the Russian invasion on Ukraine (02/01/2022 – 04/29/2022). SMAPE values are averages from different hyperparameter combination outcomes (attention head = 1, 4, 8; hidden size = 16, 32, 64). Adjusted Diebold-Mariano test (DM*) for significant difference to the best performing GARCH model (Table 26) for given financial instrument and forecast horizon on 5% level (*).

Forecast Horizon	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
1	0.5035*	0.4420*	0.5708*	0.8090*	0.4117*	0.5098*	0.5069*	0.5978*	0.5382*
2	0.4982*	0.4007*	0.6156*	0.8439*	0.4367*	0.5343*	0.5293*	0.6417*	0.5648*
3	0.5091*	0.4447*	0.6303*	0.7959*	0.4625*	0.5354*	0.5463*	0.6534*	0.5581*
4	0.5106*	0.4374*	0.6715*	0.7974*	0.4585*	0.5402*	0.5430*	0.6341*	0.5063*
5	0.5238*	0.4546*	0.7731*	0.8168*	0.4590*	0.5871*	0.5445*	0.6491*	0.5355*
6	0.5069*	0.4466*	0.7061*	0.7807*	0.4775*	0.5508	0.5628*	0.6512*	0.5410*
7	0.5167*	0.4694*	0.7152*	0.7485*	0.4867*	0.5850	0.6168*	0.6620*	0.5411*
8	0.5440*	0.4592*	0.6829*	0.8475*	0.4510*	0.5334*	0.5136*	0.6419*	0.5032*
9	0.5258*	0.4794*	0.6913*	0.8012*	0.4275*	0.5342*	0.5344*	0.5955*	0.5247*
10	0.5574*	0.4606*	0.6787*	0.7658*	0.4364*	0.5750*	0.6058*	0.6006*	0.4801*
Average	0.5190*	0.4495*	0.6735*	0.8007*	0.4508*	0.5485*	0.5504*	0.6327*	0.5293*

Source: Own calculations using Python PyTorch (2025).

Table 35: SMAPE for multi-step out-of-sample multivariate forecasts around the Hamas attack on Israel (10/01/2023 – 12/29/2023). SMAPE values are averages from different hyperparameter combination outcomes (attention head = 1, 4, 8; hidden size = 16, 32, 64). Adjusted Diebold-Mariano test (DM*) for significant difference to the best performing GARCH model (Table 27) for given financial instrument and forecast horizon on 5% level (*).

Forecast Horizon	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
1	0.8322*	0.7484*	0.7274*	0.6590*	1.1068*	0.5575*	0.5399*	0.4626*	0.7519*
2	0.8463*	0.7724*	0.7395*	0.6765*	0.9440*	0.5821*	0.5732*	0.4811*	0.6989*
3	0.8041*	0.7037*	0.8034*	0.6387*	0.9762*	0.5753*	0.5291*	0.4938*	0.7057*
4	0.8092*	0.7717*	0.7869*	0.6319*	1.1954*	0.5668*	0.5315*	0.4884*	0.7392*
5	0.7978*	0.7168*	0.7709*	0.5852*	1.3700	0.5776*	0.5221*	0.5031*	0.6922*
6	0.8070*	0.7711*	0.7711*	0.6626*	1.4771*	0.6060*	0.5426*	0.5321*	0.7072*
7	0.7986*	0.7727*	0.7731*	0.6404*	1.7626	0.5641*	0.5736*	0.5037*	0.6392*
8	0.9380*	0.8324*	0.7713*	0.6292*	1.3044	0.5077*	0.5361*	0.5310*	0.6597*
9	0.9343*	0.9893*	0.7729*	0.6906*	1.6226	0.5146*	0.5759*	0.5262*	0.6300*
10	1.0011*	1.0149*	0.8601*	0.6428*	1.5635*	0.5578*	0.5571*	0.5124*	0.6184*
Average	0.8569*	0.8093*	0.7776*	0.6457*	1.3323*	0.5610*	0.5481*	0.5034*	0.6842*

Source: Own calculations using Python PyTorch (2025).

Table 36: Adjusted Diebold-Mariano test (DM*) statistic (p-value) for comparing the average 1-10 step-ahead out-of-sample forecasting performance of the TFT model with the best performing GARCH-type model. The null hypothesis assumes that the average 1-10 forecast horizon SMAPE of the TFT is equal to GARCH model. * indicates, that the null hypothesis could be rejected on a 5% significance niveau.

Period	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
COVID-19	20.2181* (0.0000)	11.3840* (0.0000)	6.1960* (0.0001)	5.8894* (0.0002)	18.6579* (0.0000)	24.2472* (0.0000)	4.6274* (0.0009)	-13.2712* 2.0000	6.1444* (0.0001)
Russian Invasion	14.3296* (0.0000)	14.2845* (0.0000)	12.4864* (0.0000)	9.4278* (0.0000)	30.6403* (0.0000)	33.9125* (0.0000)	16.3416* (0.0000)	24.1394* (0.0000)	41.3117* (0.0000)
Hamas Attack	19.9748* (0.0000)	13.4482* (0.0000)	11.6515* (0.0000)	31.9221* (0.0000)	1.7181 (0.1165)	15.2089* (0.0000)	28.9784 (0.0000)	30.4088* (0.0000)	17.2105* (0.0000)

Source: Own calculations using Python PyTorch (2025).

4.5.2. Regime switching multi-step out-of-sample forecasting performance

As I discussed earlier, geopolitical shocks came along with a sudden increased uncertainty, which is expressed as a volatility jump. This means that the Volatility rises to a higher level and remains there, which is associated with stronger return-variations. In the previous section, I examined the ability of the TFT to adapt to the increased volatility caused by the geopolitical shock and to make correct predictions accordingly. But I did not give the TFT an explanatory variable that this is a phase of particularly high volatility but rather checked how the TFT managed to recognize this regime change and adapt to it on its own. In this section, I want to examine whether adding a “regime-switching” feature significantly improves the prediction quality. To do this, I examined and marked the historical volatility training values for the TFT. Values greater than the 75% quantile were given a one as a characteristic, other values were given a zero. Then I proceed as follows for the forecast. Comparable to the autoregressive logic of the switching probabilities of a Markov switch model (Hamilton, 1988), I gave the TFT the information whether the last known volatility value had a 1 as a characteristic or not. In line with the phenomenon of volatility clustering, I then assumed that if the last value had the characteristic 1, then the future value would also have this characteristic and accordingly the volatility would be above the 75% quantile. This procedure is also called past covariate.

I have done this again for all the financial instruments under consideration and for all periods of geopolitical shocks. For each financial instrument, again I used the rolling window approach, in which the TFT is first trained on the data from the last 250 trading days and then generates forecasts for the next 10 days. Each of these training sessions is carried out on the optimization grid of hyperparameters: attention size and hidden size. The rolling window goes through between 62 and 64 steps. However, the regime-switching feature is also to be implemented as a future covariate when the geopolitical shock is known. For the three geopolitical events I examined, all volatility values that occurred with and after the shock would be given a corresponding characteristic. I evaluated the forecasting performance by calculating the SMAPE. To evaluate if the regime-switching TFT (RS-TFT) outperforms the baseline multivariate TFT (BL-TFT, Table 33-35), I calculated the adjusted Diebold-Mariano test statistics and respective p-values. In terms of results, I have the following expectations. The RS-TFT should perform better in particular when the time series to be forecasted shows significant non-linear patterns and structural breaks caused by geopolitical shocks or other exogenous factors. For example, when there are clear regimes with different statistical properties. For example, if the behavior of volatility during a geopolitical shock fundamentally differs from “normal” market phases. The model can learn from the regime feature that different dynamics prevail in different market regimes, while a standard TFT tries to learn a single, continuous model structure, which can lead to distortions. Furthermore, sudden and non-linear changes in the analyzed time series lead to a significant improvement in forecasting performance when a regime switch is implemented. When a geopolitical shock does not trigger stepwise but abrupt changes in the time series, a standard model may have difficulties to adapt quickly enough to the new data structure (e.g. sharp price

drops following a shock). The regime-switching feature allows the model to immediately activate a different prediction pattern after a shock, rather than to apply the “old” pattern to new, unsuitable data. Finally, the relationships between variables in different regimes can also change. For example, in normal market phases, equity returns could depend heavily on corporate earnings, whereas in a geopolitical crisis, other external factors such as interest rate spreads or political risk dominate. I now discuss my findings (Tables 37-39).

For the S&P 500, the RS-TFT outperformed the BL-TFT in all three analyzed periods of geopolitical shock and had the third largest effect among all analyzed financial instruments. The average difference in SMAPE was 3.93% (COVID-19), 1.68% (Russian War) and 9.62% (Hamas Attack) and overall, at 5.08%. Furthermore, all three were statistically significant. I was able to observe the largest advantage of the RS model with a difference of 18.73% for the 9-step-ahead forecast during the Hamas Attack. With an overall average difference in SMAPE of 5.91%, the RS-TFT had the second largest effect for the NASDAQ and outperformed the BL-TFT in all three periods with an average difference in SMAPE at 4.60% (COVID-19), 3.91% (Russian War) and 9.22% (Hamas Attack). These differences were also statistically significant. The largest effect could be gained for the 9-step ahead forecast during the Hamas Attack with a difference of 22.65%.

The RS-TFT also achieved an advantage for the forecasting performance of the Nikkei. The average difference in SMAPE between RS-TFT and BL-TFT over all periods considered was 4.61% and 5.51% (COVID-19), 1.86% (Russian War) and 6.47% (Hamas Attack). All were statistically significant. I found the largest difference in SMAPE in the 6-step ahead forecast during COVID-19 (14.22%). For the Hang Seng, the RS-TFT was also able to achieve an advantage for the prediction performance. The average difference in SMAPE between RS-TFT and BL-TFT over all periods considered was 5.00% and 1.59% (COVID-19), 5.37% (Russian War) and 8.03% (Hamas Attack), respectively. All of them were statistically significant. Like for the S&P 500 and NASDAQ, I found the largest difference in SMAPE in the 9-step ahead forecast during the Hamas attack (12.59%). The largest difference between RS-TFT and BL-TFT was observed in the commodity gold - the average SMAPE difference across all periods was 9.04%. However, the effects are very unevenly distributed over the periods. During COVID-19, the SMAPE difference was very small at 0.38%, and the DM* test statistic could not be rejected. This was also the case during the Russian attack - the SMAPE difference here was only 0.08% and the DM* test statistic could not be rejected either. The effect originated from the period of the Hamas attack. Here, the average SMAPE difference across all forecast horizons between RS-TFT and BL-TFT was 25.85%. The largest single SMAPE difference during this period was observed for the 7-step ahead forecast (52.94%).

The usage of the RS-TFT had the smallest effect on the Brent Crude Oil commodity. The average SMAPE difference between RS-TFT and BL-TFT over all periods was 1.37%. Of this, 1.44% was accounted for by the COVID-19 period, 0.97% by the period of the Russian attack, and 1.72% by the period during the Hamas attack. Accordingly, the effect was greatest for the period. The DM* test statistic could only be rejected during the Hamas attack. Based on the negative values for the DM* test statistics, it can be deduced that the BL-TFT model is to be

preferred. The statistical significance can be found in Table 40. I observed the largest single SMAPE difference in this period of the Hamas attack for the 6-step ahead forecast (6.24%). For the 10yr US-Treasury Bond, the usage of the RS-TFT had an average SMAPE difference between RS-TFT and BL-TFT of 4.37%. In contrast to gold, the SMAPE differences between the individual periods were close: 4.56% during COVID-19, 3.95% during the Russian attack, and 4.60% during the Hamas attack. The null hypothesis of the DM* test could be rejected in all cases, thus showing that the forecasting performance of the RS-TFT model is statistically significantly different from that of the BL-TFT. I observed the largest single SMAPE difference in this period of the Russian attack for the 10-step ahead forecast (15.76%).

I observed a rather small effect for the EUR-USD exchange rate. The average SMAPE difference between RS-TFT and BL-TFT was 3.25%. Specifically, 3.12% during COVID-19, 3.02% during the Russian attack, and 3.61% during the Hamas attack. The null hypothesis of the DM* test could be rejected for all periods despite the relatively small difference. The largest single SMAPE difference was observed in this period of COVID-19 for the 8-step ahead forecast (9.67%). For Bitcoin, I found that using the RS-TFT also generated a small effect. Across all periods, the average SMAPE difference between RS-TFT and BL-TFT was 2.11%, of which - 1.53% during COVID-19, 1.78% during the Russian attack, and 6.09% during the Hamas attack. The null hypothesis of the DM* test could not be rejected for the period of COVID-19 and the BL-TFT provided better results. For both other periods, the null hypothesis of the DM* test could be rejected. The largest single SMAPE difference I observed in this period of the Hamas attack for the 4-step ahead forecast (13.98%). In summary, I found that the implementation of a regime-switching feature in the Temporal Fusion Transformer (RS-TFT) has a positive effect on all financial instruments compared to the baseline TFT (BL-TFT). Furthermore, I found that this effect was greatest for the period of the Hamas attack. For the individual financial instruments, the effect was strongest for gold (albeit due to the strong effect during the Hamas period), the NASDAQ and the S&P 500. I found almost no effect for Brent Crude Oil.

Overall, I saw, that the implementation of the regime-switching feature significantly improves model performance. On average across all financial stocks, forecast horizons and periods of geopolitical shocks, the regime-switch approach performed 4.53% better than the normal TFT approach. Furthermore, the null hypothesis that the forecasting performance of the regime-switch TFT for a selected financial instrument is identical to that of the baseline TFT could be rejected a total of 21 times out of 27 times (9 financial instruments each with three analyzed periods) over all three periods with geopolitical shocks.

Across all assets, horizons, and shock periods, the RS-TFT outperformed the BL-TFT by an average SMAPE improvement of 4.53%. In 21 out of 27 cases the null hypothesis of the DM*-test statistic could be rejected. Regarding the shock-period related results, the Hamas attack period yielded the strongest gains. This was surprising, since in this period the lowest volatility increase was observed. A possible explanation could be that the improved performance of the model incorporating the regime-switching feature during periods of low volatility response can be attributed to the features role in preventing the TFT from erroneously forecasting elevated

volatility solely due to the presence of a geopolitical shock. In contrast, during highly volatile periods, the shock is already implicitly captured through other temporal patterns, rendering the additional signal provided by the dummy less informative. This highlights the TFT's strength in modeling and leveraging non-linear contextual effects in an interpretable manner. During COVID-19 and the Ukraine war, gains were positive but more modest, reflecting the variability in volatility dynamics across crises. Regarding the financial instruments, NASDAQ and S&P 500 showed consistent and statistically significant improvements, especially at longer forecast horizons. Nikkei 225 and Hang Seng also benefited. Gold showed the largest average gain, but this effect was entirely driven by the Hamas period, with negligible gains during other crises. Brent Crude Oil exhibited minimal sensitivity to the regime-switching feature and in some cases performed better with the BL-TFT. For 10-year US Treasuries, RS-TFT achieved consistent gains across all periods. DM* tests rejected the null in every case. EUR-USD exchange rate showed modest gains and Bitcoin exhibited limited overall improvement with negative gains during COVID-19. Gains were present during later periods, but less stable than for other assets. The regime-switching feature had the strongest impact at longer forecast horizons, where the RS-TFT often achieved its maximum relative improvement. This supports the hypothesis that regime information is especially valuable for medium-to-long-term forecasts following structural breaks. This can be interpreted as relatively greater benefit of incorporating a geopolitical regime-switching feature at longer forecast horizons likely reflects the temporal nature of external shocks, whose effects on financial volatility often manifest with a delay. While short-term predictions can be effectively driven by local temporal patterns, longer-term forecasts benefit more from contextual signals that encode structural regime changes or latent risk factors. The dummy variable thus serves as a horizon-sensitive contextual anchor that helps the TFT anticipate volatility dynamics beyond the autoregressive signal.

This analysis showed that the introduction of an additional explanatory variable (regime-switching feature) to identify high and low phases of volatility significantly improves the forecasting performance and thus answers my research question (Research Question 9).

Table 37: SMAPE for regime-switching multi-step out-of-sample multivariate forecasts during COVID-19 (02/03/2020 – 04/30/2020). SMAPE values are averages from different hyperparameter combinations (attention head = 1, 4, 8; hidden size = 16, 32, 64). Adjusted Diebold-Mariano test (DM*) for significant difference between regime-switching TFT (RS-TFT) baseline TFT (BL-TFT) (Table 33) for given financial instrument and forecast horizon on 5% level (*).

Forecast Horizon	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
1	0.6429	0.7021*	0.5023*	0.6174	0.6865	0.5836	0.7531*	1.1593	0.9739
2	0.6352*	0.7848*	0.5418	0.6173	0.6623	0.5878	0.7889	1.1831	0.9173
3	0.6778	0.8224	0.5817	0.5949*	0.7123	0.5739	0.7756*	1.1395*	1.0023
4	0.6785	0.8529	0.6485	0.6010*	0.6810*	0.5589	0.7376*	1.1214*	0.9567
5	0.6645	0.8129*	0.7424	0.5672*	0.7201	0.5988	0.7389*	1.1669*	0.9528
6	0.6550*	0.8774	0.7116*	0.6148	0.7242	0.6162	0.7534	1.2736	0.8748
7	0.7372	0.7849*	0.6209*	0.5975	0.7130	0.6183	0.7286	1.2392*	0.9530
8	0.7186*	0.8148*	0.7076	0.5831	0.7101	0.6471	0.7092	1.1592*	0.9379
9	0.7040	0.8325	0.7200*	0.5965	0.7202	0.6498	0.7517	1.1595*	0.9697
10	0.7053	0.7990	0.7098*	0.5865	0.7093	0.6424	0.7073*	1.1266	0.9087
Average	0.6819*	0.8084*	0.6487*	0.5976*	0.7039	0.6077	0.7444*	1.1728*	0.9447

Source: Own calculations using Python PyTorch (2025).

Table 38: SMAPE for regime-switching multi-step out-of-sample multivariate forecasts around the Russian invasion of Ukraine (02/01/2022 – 04/29/2022). SMAPE values are averages from different hyperparameter combinations (attention head = 1, 4, 8; hidden size = 16, 32, 64). Adjusted Diebold-Mariano test (DM*) for significant difference between regime-switching TFT (RS-TFT) baseline TFT (BL-TFT) (Table 34) for given financial instrument and forecast horizon on 5% level (*).

Forecast Horizon	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
1	0.4664*	0.3785*	0.5355*	0.7970	0.4116	0.5049	0.5237	0.5806	0.5422
2	0.4833	0.3903	0.5931	0.7334*	0.4131	0.5241	0.5267	0.5802*	0.5421
3	0.4998	0.4095*	0.5631*	0.7379*	0.4265*	0.5522	0.5650	0.6104*	0.5152*
4	0.5226	0.4114	0.6780	0.7490*	0.4539	0.5249	0.5484	0.6127*	0.5562
5	0.5094	0.4078*	0.7045*	0.7644*	0.4527	0.5477*	0.4971*	0.6170*	0.4962*
6	0.4981	0.4219	0.6861	0.7409*	0.4382*	0.5624	0.5168*	0.6543	0.4918*
7	0.4868*	0.4171*	0.7091	0.7401	0.4704	0.5825	0.5106	0.6309*	0.4787*
8	0.5105*	0.4224*	0.6939	0.7759*	0.4696	0.5378	0.4937	0.6130	0.4950
9	0.5143	0.4266*	0.7103	0.7368*	0.4312	0.5136	0.4782*	0.5562*	0.4996
10	0.5365*	0.4174	0.6755	0.6945*	0.4519	0.5382*	0.4482*	0.5703*	0.4981
Average	0.5028*	0.4103*	0.6549	0.7470*	0.4419	0.5388	0.5109*	0.6026	0.5115*

Source: Own calculations using Python PyTorch (2025).

Table 39: SMAPE for regime-switching multi-step out-of-sample multivariate forecasts around the Hamas attack on Israel (10/01/2023 – 12/29/2023). SMAPE values are averages from different hyperparameter combinations (attention head = 1, 4, 8; hidden size = 16, 32, 64). Adjusted Diebold-Mariano test (DM*) for significant difference between regime-switching TFT (RS-TFT) baseline TFT (BL-TFT) (Table 35) for given financial instrument and forecast horizon on 5% level (*).

Forecast Horizon	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
1	0.7881*	0.6956*	0.6216*	0.5406*	0.9079*	0.5684	0.5090*	0.4463	0.6887*
2	0.7926*	0.7240*	0.6861*	0.6358*	0.7212*	0.5583	0.5367*	0.4624	0.6395*
3	0.7033*	0.6562*	0.6724*	0.5626*	0.8237*	0.5618	0.4882*	0.4429*	0.6653*
4	0.7153*	0.6635*	0.6850*	0.5447*	0.9131*	0.5753	0.4869*	0.4335*	0.5993*
5	0.7736	0.6675*	0.7319	0.5595	1.0146*	0.5481	0.4962	0.4496*	0.5753*
6	0.7159*	0.6920*	0.6780*	0.5663*	1.1306*	0.5435*	0.4668*	0.4698*	0.6576*
7	0.7602	0.7492	0.7421	0.5353*	1.2331*	0.5274*	0.4678*	0.4821	0.6335
8	0.7861*	0.7572*	0.7561	0.5835*	1.1695*	0.5335	0.5147	0.5203	0.6100*
9	0.7469*	0.7628*	0.7571	0.5647	1.4378*	0.5055	0.5317*	0.4932	0.5867*
10	0.8245*	0.8032*	0.7990*	0.5609*	1.3858*	0.5160*	0.5235*	0.4733**	0.5778*
Average	0.7606*	0.7171*	0.7129*	0.5654*	1.0737*	0.5438*	0.5022*	0.4673*	0.6234*

Source: Own calculations using Python PyTorch (2025).

Table 40: Adjusted Diebold-Mariano test (DM*) statistic (p-value) for comparing the out-of-sample forecasting performance of the regime-switching TFT (RS-TFT) with the baseline TFT model (BL-TFT). The null hypothesis assumes that the average 1-10 forecast horizon SMAPE of the RS-TFT is equal to the BL-TFT. * indicates, that the null hypothesis could be rejected on a 5% significance niveau.

Period	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
COVID-19	3.6888* (0.0021)	3.0809* (0.0058)	2.0781* (0.0322)	3.1201* (0.0054)	0.5910 (0.2838)	1.4154 (0.0937)	5.5399* (0.0001)	2.0345* (0.0346)	-1.3194 (0.8918)
Russian Invasion	2.6047* (0.0131)	8.1822* (0.0000)	0.9018 (0.1942)	6.1454* (0.0001)	1.3186 (0.1084)	1.3122 (0.1094)	3.6980* (0.0021)	3.1725* (0.0050)	2.1472* (0.0287)
Hamas Attack	7.6597* (0.0000)	5.8571* (0.0001)	3.8723* (0.0015)	8.9136* (0.0000)	3.4299* (0.0032)	2.3570* (0.0201)	5.8201* (0.0001)	4.3295* (0.0007)	4.8736* (0.0003)

Source: Own calculations using Python PyTorch (2025).

4.5.3. Influence of hyperparameter optimization on the multi-step out-of-sample forecasting performance

An important task in adapting a neural network for a task is optimizing the hyperparameters. Hyperparameters are configurable parameters that govern the training process of a machine learning model but are not learned from the data. They are set before training and significantly influence the model's performance, convergence speed, and generalization ability. In the context of the TFT or other deep learning models, hyperparameters include architecture-related hyperparameters (e.g. number of hidden layers, number of hidden units per layer and the number of attention heads), optimization-related hyperparameters (e.g. learning rate, batch size, weight decay and the gradient clipping threshold) and the training-specific parameters (e.g. numbers of epochs, early stopping patience and the loss function selection) (e.g., Bischl et al., 2023). In the analysis, I applied a grid search of the two important architecture hyperparameters of the TFT: attention head and hidden size.

The hyperparameter attention head is a fundamental component of the multi-head attention mechanism used in Transformer-based architectures. It represents an independent attention mechanism that learns to focus on different parts of the input sequence during training. Instead of using a single attention head, Transformer architectures use multiple heads in parallel. Each head focuses on different aspects of the input data, capturing different dependencies. So, the number of attention heads determine how well the TFT can capture different dependencies of the data (e.g., short-term vs. long-term trends in time series). Although, the number improves the model robustness by attending to different input features. The hyperparameter hidden size (also called hidden dimension) refers to the number of neurons in a hidden layer of a neural network. It determines the dimensionality of the representations learned by the model at each layer. The hidden size determines the model capacity (larger hidden size allows the model to learn more complex patterns but increases computational cost) and the representation power (a small hidden size may lead to underfitting, while too large a size can cause overfitting). What does it mean, if a model with more attention heads outperforms a model with less? If a TFT model with more attention heads outperforms a model with fewer heads in forecasting financial time series, this suggests that the model benefits from enhanced pattern extraction, better feature selection, and improved temporal dependency learning. Improved feature representation: each attention head learns to focus on different aspects of the input time series.

Financial time series are highly complex, often containing multiple interacting patterns such as: short-term fluctuations (e.g., intraday trading movements), long-term trends (e.g., market cycles), volatility clustering (e.g., periods of high or low market uncertainty) and cross-asset relationships (e.g., correlations between stocks, bonds, or commodities). So, more attention heads indicate the model can capture more diverse dependencies across different time scales and financial instruments, which enables an enhanced temporal dependencies learning, i.e. in a multi-head attention mechanism, each head learns different types of temporal dependencies. More heads allow the model to extract finer details from past observations, leading to better trend detection (e.g., identifying sustained price movements), more precise volatility estimation (e.g., capturing

GARCH-like patterns) and improved seasonality recognition (e.g., detecting cyclical market behaviors). So, with more heads, the model can track multiple independent financial signals, leading to better forecasting accuracy. Robustness to noise and market anomalies: financial markets exhibit high noise levels and non-stationarity. A model with fewer attention heads may be overfit to a dominant pattern and fail to generalize. More attention heads enable the model to distribute its focus, reducing the risk of overfitting to short-term fluctuations and missing sudden regime shifts (e.g., crashes, recoveries). The additional attention heads help create a more robust forecasting model by diversifying the learned representations. What does it mean, if a model with more attention heads outperforms a model with less? If a TFT with a higher hidden size outperforms a model with a smaller hidden size in forecasting financial time series, it suggests that the model benefits from increased representation capacity, better feature extraction, and improved temporal pattern recognition. Increased representation power: the hidden size defines the number of units (neurons) in each layer of the TFT. A higher hidden size means the model can learn more complex patterns from financial time series data.

Financial time series are highly non-linear and often exhibiting sudden regime shifts (e.g., market crashes, bull runs), volatility clustering (e.g., GARCH-like behavior) and cross-asset correlations (e.g., stock-bond inverse relationships). A larger hidden size enables the TFT to model these complex relationships more effectively, leading to better forecasting performance. Improved feature interactions and selection: The TFT uses gated residual networks (GRNs) and variable selection networks (VSNs) to determine the most important features dynamically. A higher hidden size allows more expressive feature transformations, so the model can detect interactions between macroeconomic indicators, technical signals, and order flow data. Better attention-based selection means, that the model can allocate importance to the most relevant input features. With a higher hidden size, the model makes more precise decisions about which financial indicators drive future movements. Capturing long-term and short-term dependencies: financial markets operate at multiple time scales (e.g., microsecond trades, daily price movements, monthly trends). Temporal dependencies in financial data are complex and require short-term sensitivity (e.g., response to earnings reports) and long-term memory (e.g., interest rate cycle effects). A larger hidden size helps TFT maintain and process longer time dependencies, allowing it to capture both short-term fluctuations and long-term trends. A smaller hidden size may fail to learn deep temporal structures, leading to weaker forecasts.

Now I discuss the methodology for the analysis. First, I selected the hyperparameter combination from the grid. After this, I trained the network and proceeded a multi-step out-of-sample forecast for all financial time series in scope using the rolling window approach. I have done this for all the three different periods of geopolitical shocks. Then, I evaluated the forecasting performance by using the SMAPE. To identify statistically significant differences of the forecast performance of the TFT with a different hyperparameter setup, I computed a adjusted Diebold-Mariano test statistic. Hereby I, compared the SMAPE of the best performing model (BP-TFT, model with the hyperparameter combination, that performed best in terms of the lowest SMAPE), with the average hyperparameter SMAPE, which I used the baseline TFT model for (BL-TFT,

Table 33-35). Thereby I want to analyze whether the hyperparameter selection improved the forecast performance significantly. I now discuss the results (Table 41-44). Overall, I found the following: optimization of the hyperparameters had a statistically significant effect in predicting the financial values I looked at in times of geopolitical shocks. The largest effect, I found during the COVID-19 and Hamas attack periods. In these two phases, most of the BP-TFT had a statistically significant difference from the BL-TFT. During these two phases, the null hypothesis that the BP-TFT is not statistically significantly different from the average was rejected for all but one financial value. The average effect (average difference between BP-TFT and BL-TFT over all financial instruments) was 3.94% for the COVID-19 period and 6.13% for the Hamas period. During the Russian war period, the average effect was just 2.81%. For the phase of COVID-19, it was the Hang Seng stock index and for the phase of the Hamas attack, it was Bitcoin. For the Russian attack phase, it was the NASDAQ stock index, the Nikkei stock index, gold and Bitcoin. I also evaluated the frequency with which the BP-TFT has a particular hyperparameter value. These results are not quite as clear and are close to each other. The BP-TFT used a number of 8 attention heads in 40.74% of cases. 33.33% of the BP-TFT used only one attention head. 25.92% of the best performing models used 4 attention heads. From this, one can deduce that for the observed volatility time series, it was advantageous that the TFT was able to capture different dependencies well. One can observe something similar for the hidden size hyperparameter. I found that 44.44% of the BP-TFT used 32 neurons per hidden layer. 29.63% of the BP-TFT used 64 neurons per hidden layer and 25.93% used only 16 neurons per hidden layer. The number of neurons in a hidden layer controls the complexity of the TFT and thus its ability to fit the data. When choosing the number of hidden neurons, it is important to avoid both over- and underfitting. One can see here that most of the best performing models lied exactly in the middle with 32 neurons per hidden layer. Accordingly, for the data, an attention head of 8 and a hidden size of 32 showed the best forecasting results.

Now I will discuss the results of individual financial stocks. For the S&P 500, I found, that in all periods, the SMAPE of the BP-TFT significantly outperformed the BL-TFT. The null hypothesis of the DM*-test statistic could be rejected for each period. Over the different periods, the average difference between BP-TFT and BL-TFT was 4.26%. Furthermore, I found that the optimal number of hidden size parameters was 32 for all three periods considered. In contrast, no clear structure could be derived for the number of attention heads. The NASDAQ BP-TFT significantly outperformed the BL-TFT during COVID-19 and the Hamas period. During the Russian attack the null hypothesis could not be rejected. The average difference in SMAPE over the periods was 2.87%. The BP-TFT used 1 attention head and 32 neurons per hidden layer the most frequent. However, it must be emphasized again that this is an isolated analysis and that the optimal hyperparameters cannot be chosen independently of each other and must not be re-combined. Like the NASDAQ, the BP-TFT of the Nikkei index outperformed the average model during COVID-19 and the Hamas period. The null hypothesis could not be rejected for the forecasts during the Russian attack. The average difference in SMAPE over the periods was 2.89%. The BP-TFT used 8 attention head and 32 neurons per hidden layer the most frequent. For

the Hang Seng index, the null hypothesis could not be rejected during COVID-19, but for the two other periods in scope, it could be rejected. I also found an average difference of the best performing and average model at 4.24%. Looking at the hyperparameters, it was more straight. In all cases the BP-TFT used 32 neurons per hidden layer and in 2 of 3 cases they used 8 attention heads. For the commodity gold, the BP-TFT outperformed the BL-TFT significantly during COVID-19 and the Hamas attack. For the period of the Russian attack, the null hypothesis could not be rejected. The average difference in SMAPE was 4.44%. In most cases the BP-TFT just used one attention head, but a hidden size of 64 most frequently.

Considering the commodity brent crude oil, I found that the BP-TFT outperformed the BL-TFT in every period in scope significantly. The average difference in SMAPE over the periods was at 3.00%. Moreover, I found that in most cases the architecture of the BP-TFT contains 4 attention heads and a hidden size of 64. Similarly, for the US Treasury bond time series, the best performing models significantly outperformed the average model across all periods, thus the null hypothesis of the DM* test could be rejected in any time. The average difference in SMAPE was 3.44%. From an architectural perspective, the network showed a rather inconsistent picture over the periods under review. Accordingly, the number of attention heads and hidden size for the best performing model varied. This indicates a rather changeable structure of the data in the different periods. The BP-TFT for the time series of the EUR-USD exchange rate significantly outperformed the average model in every period significantly, and the null hypothesis could always be rejected. The average difference in SMAPE was 4.77%. Furthermore, I found that most of the time the BP-TFT contain the maximal value of 8 attention heads, but the smallest number of neurons per hidden layer.

My findings showed, that hyperparameter selection affect forecasting performance significantly. In particular, for the two important architectural hyperparameters, attention head and hidden size, a suitable choice is crucial for forecasting performance. However, the rule here is not the more complex the model, the better. The advantages and disadvantages of increasing the complexity of the model must be considered. Regarding the attention head, I found that a higher number in the optimization grid increased the forecasting performance. This may be due to the fact that a higher number have an improved feature representation and thus can better capture volatility clustering patterns, which occurs in the data. Furthermore, the correlation between the financial time series, i.e. the explanatory variables, can be better captured and used for forecasting. Furthermore, more attention heads enable the model to extract finer details from past observations, leading to more precise volatility estimation (e.g. capturing GARCH-like patterns). This also refers to the data. Regarding the hidden size parameter, I found that a mean value of 32 occurred on the grid for most of the best performing models. Moreover, a higher hidden size enables the model to learn more complex patterns like exhibiting sudden regime shifts, volatility clustering (e.g., GARCH-like behavior) and cross-asset correlations more effectively, which is leading to better forecasting performance. At the same time, the possibility of modeling a high level of complexity is also associated with the potential for overfitting. Accordingly, a trade-off must be made in this regard. The results show, however, that a higher level of complexity has a positive effect on the

forecasting performance of volatility and the volatility clusters and regime switches that occur in the data. When evaluating the best performing models for Bitcoin, I found that the null hypothesis could be rejected just for the period of COVID-19. For the other two periods, there was no significant difference of the SMAPE of the best performing model and the average model. Nevertheless, the average difference was at 8.71%. However, this is due to the fact that the performance of the TFT for the period of the Hamas attack was extremely poor.

My grid-optimization of the TFT architecture revealed that both the number of attention heads and the hidden layer size significantly influence model performance in forecasting financial time series volatility during geopolitical shocks. Models with 8 attention heads and a hidden size of 32 were most frequently among the best-performing configurations, suggesting that a balance between attention diversity and model capacity yields optimal predictive accuracy. The Diebold-Mariano test confirmed that hyperparameter tuning led to statistically significant improvements in most cases, with the largest effects observed during the COVID-19 and Hamas attack periods. Asset-level analyses further highlighted that indices, Brent crude oil, and EUR-USD exchange rate consistently benefited from hyperparameter optimization across all periods, whereas other assets like Bitcoin showed mixed results due to period-specific dynamics. A higher number of attention heads allowed the model to better capture multi-scale temporal dependencies, volatility clustering, and cross-asset correlations, while a moderate hidden size provided sufficient expressiveness without overfitting. The findings emphasize that model complexity should be tailored to the data structure, as both under- and over-parameterization can reduce forecast accuracy. In conclusion, careful tuning of architectural hyperparameters enhances the TFT's ability to model complex volatility patterns, especially in environments affected by structural shocks. These results underscore the importance of adaptive model configuration in time series forecasting under non-stationary conditions (Research Question 10).

Table 41: SMAPE for multi-step out-of-sample multivariate forecasts during COVID-19 (02/03/2020 – 04/30/2020) and combinations of different hyperparameter combinations (attention head = 1, 4, 8; hidden size = 16, 32, 64). The SMAPE values are averages from forecast horizons $h=1, \dots, 10$. Adjusted Diebold-Mariano test (DM*) for significant difference of TFT (BP-TFT) and baseline TFT (BL-TFT) (Table 33). * indicates the best performing hyperparameter combination (BP-TFT). ** indicates that the BP-TFT performance is statistically significant different from BL-TFT for given financial instrument and averaged forecast horizons on 5% level.

Attention Head	Hidden Size	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
1	16	0.7763	0.8972	0.6992	0.6264	0.6848**	0.6311	0.7762	1.1615	0.9758
1	32	0.7193	0.8563	0.6705**	0.6223	0.7050	0.6282	0.7754	1.1908	0.9517
1	64	0.7239	0.8468	0.7005	0.6299	0.7186	0.6353	0.7761	1.1845	0.9433
4	16	0.7028	0.8668	0.7211	0.6197	0.7029	0.6510	0.7872	1.1845	0.9308
4	32	0.7032	0.8273	0.7061	0.5960*	0.7085	0.5961	0.7970	1.2643	0.9082
4	64	0.7493	0.8214**	0.6817	0.5975	0.7102	0.5907**	0.7877	1.2656	0.9066
8	16	0.7342	0.8588	0.7300	0.6151	0.7012	0.6429	0.7383**	1.1183**	0.9719
8	32	0.6817**	0.8255	0.7374	0.6101	0.7173	0.6178	0.8120	1.2109	0.8874**
8	64	0.6998	0.8894	0.6869	0.6051	0.7210	0.6054	0.8605	1.2347	0.8894
Average		0.7211	0.8543	0.7037	0.6135	0.7077	0.6220	0.7900	1.2016	0.9294

Source: Own calculations using Python PyTorch (2025).

Table 42: SMAPE for multi-step out-of-sample multivariate forecasts during the Russian invasion (02/01/2022 – 04/29/2022) and combinations of different hyperparameter combinations (attention head = 1, 4, 8; hidden size = 16, 32, 64). The SMAPE values are averages from forecast horizons $h=1, \dots, 10$. Adjusted Diebold-Mariano test (DM*) for significant difference of TFT (BP-TFT) and baseline TFT (BL-TFT) (Table 34). * indicates the best performing hyperparameter combination (BP-TFT). ** indicates that the BP-TFT performance is statistically significant different from BL-TFT for given financial instrument and averaged forecast horizons on 5% level.

Attention Head	Hidden Size	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
1	16	0.5003	0.4511	0.6554	0.8150	0.4683	0.5539	0.5610	0.6034**	0.5234
1	32	0.4924**	0.4332*	0.6808	0.8234	0.4462	0.5566	0.5567	0.6587	0.5240
1	64	0.5482	0.4511	0.6704	0.8126	0.4331*	0.5224**	0.5540	0.6181	0.5170*
4	16	0.5068	0.4382	0.6698	0.8111	0.4666	0.5680	0.5686	0.6681	0.5387
4	32	0.5561	0.4537	0.6947	0.8460	0.4402	0.5481	0.5460	0.6274	0.5329
4	64	0.5180	0.4621	0.6958	0.8294	0.4477	0.5477	0.5261**	0.6278	0.5399
8	16	0.5144	0.4483	0.6785	0.7728	0.4655	0.5460	0.5541	0.6489	0.5355
8	32	0.5228	0.4636	0.6521*	0.7218**	0.4505	0.5480	0.5414	0.6184	0.5352
8	64	0.5118	0.4439	0.6644	0.7740	0.4387	0.5460	0.5452	0.6238	0.5170
Average		0.5190	0.4495	0.6735	0.8007	0.4508	0.5485	0.5504	0.6327	0.5293

Source: Own calculations using Python PyTorch (2025).

Table 43: SMAPE for multi-step out-of-sample multivariate forecasts during the Hamas attack (10/01/2023 – 12/29/2023) and combinations of different hyperparameter combinations (attention head = 1, 4, 8; hidden size = 16, 32, 64). The SMAPE values are averages from forecast horizons $h=1, \dots, 10$. Adjusted Diebold-Mariano test (DM*) for significant difference of TFT (BP-TFT) and baseline TFT (BL-TFT) (Table 35). * indicates the best performing hyperparameter combination (BP-TFT). ** indicates that the BP-TFT performance is statistically significant different from BL-TFT for given financial instrument and averaged forecast horizons on 5% level.

Attention Head	Hidden Size	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Oil	10yr. T-Bond	EUR/USD	BTC
1	16	0.8551	0.7722**	0.7829	0.6500	1.2667	0.5603	0.5287	0.5117	0.6872
1	32	0.8621	0.8077	0.7782	0.6608	1.3122	0.5504	0.5690	0.4998	0.6863
1	64	0.9137	0.8400	0.7718	0.6342	1.3795	0.5558	0.5435	0.5051	0.6908
4	16	0.8717	0.7757	0.7834	0.6830	1.3619	0.5811	0.5587	0.5324	0.6866
4	32	0.7950**	0.8226	0.7937	0.6420	1.3143	0.5637	0.5421	0.4988	0.6727
4	64	0.8547	0.8170	0.7820	0.6499	1.3459	0.5282**	0.5742	0.4965	0.6635*
8	16	0.8521	0.8119	0.7459**	0.6573	1.3649	0.5723	0.5354	0.4727**	0.6856
8	32	0.8623	0.7954	0.7976	0.6147**	1.4053	0.5702	0.5207**	0.5099	0.7105
8	64	0.8449	0.8415	0.7633	0.6192	1.2397**	0.5666	0.5606	0.5050	0.6748
Average		0.8569	0.8093	0.7776	0.6457	1.3323	0.5610	0.5481	0.5034	0.6842

Source: Own calculations using Python PyTorch (2025).

Table 44: Adjusted Diebold-Mariano test (DM*) statistic (p-value) for comparing the multi-step out-of-sample forecasting performance of the best performing hyperparameter combination TFT (BP-TFT) model with the baseline TFT model (BL-TFT, Tables 33-35). The null hypothesis assumes that the average 1-10 forecast horizon SMAPE of the regime-switching multivariate TFT is equal to the baseline TFT. * indicates, that the null hypothesis could be rejected on a 5% significance niveau.

Period	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
COVID-19	3.0308* (0.0018)	2.5308* (0.0070)	2.5923* (0.0059)	1.3462 (0.0916)	1.7615* (0.0415)	2.4077* (0.0095)	3.9769* (0.0001)	6.4077* (0.0000)	3.2308* (0.0010)
Russian Invasion	2.0461* (0.0224)	1.2538 (0.1073)	1.6461 (0.0523)	6.0692* (0.0000)	1.3615 (0.0891)	2.0076* (0.0245)	1.8692* (0.0331)	2.2538* (0.0138)	0.9461 (0.1738)
Hamas Attack	4.7615* (0.0000)	2.8538* (0.0029)	2.4384* (0.0088)	2.3846* (0.0100)	7.1230* (0.0000)	2.5230* (0.0071)	2.1076* (0.0195)	2.3615* (0.0106)	1.5930 (0.0581)

Source: Own calculations using Python PyTorch (2025).

4.5.4. The added value of a multivariate structured volatility forecasting approach

The Temporal Fusion Transformer (TFT) is particularly well-suited for multivariate time series modeling and forecasting due to its architectural features. However, a multivariate forecasting approach also involves complexity, and the question therefore arises as to whether implementing a multivariate forecasting approach adds value in the form of improved forecasting performance, or whether this approach can be used to explain more relationships than a univariate forecasting approach. In this section, I will discuss the answer to this question and evaluated whether a multivariate forecasting approach outperforms a univariate approach. To this end, I compared the results presented in section 4.5.1. (Tables 33-35) with those obtained using a univariate forecasting method.

For the univariate forecast, I only used the time series itself as input, which is then also predicted. I conducted this analysis on time series data from nine financial assets: the S&P 500, NASDAQ, Nikkei 225, Hang Seng, Gold, Brent Crude Oil, the 10-year U.S. Treasury Bond, the EUR-USD exchange rate, and Bitcoin, covering three distinct periods characterized by geopolitical shocks. This is implemented using a rolling-window approach, wherein the TFT model is trained using input data from the past 250 trading days. The trained model is then used to generate forecasts for the next 1 to 10 trading days. Subsequently, the rolling window is shifted forward by one step, and the process is repeated. For each examined period of geopolitical shocks, about 60 iterations are conducted within the rolling-window framework. Additionally, I performed this procedure for a grid of nine different hyperparameter combinations. The forecasting performance is evaluated using the Symmetric Mean Absolute Percentage Error (SMAPE). I computed the average SMAPE across identical forecast horizons. To assess the statistical significance of the multivariate forecasting approach's superiority over the univariate approach, I computed an adjusted Diebold-Mariano test statistic (DM*-test). The null hypothesis, that the forecasting performance of the multivariate model (baseline TFT, BL-TFT) is equal to that of the univariate TFT (UV-TFT), is rejected if the DM*-test statistic exceeds the critical value of the t-distribution at a predefined significance level of 0.05. Before analyzing the results, I first discuss the conditions under which a multivariate model is preferable to a univariate model.

A multivariate model is likely to outperform a univariate model when the data exhibit properties that can be better captured through joint modeling of multiple time series (Engle, 2002). If the financial instruments under consideration exhibit high conditional correlation, a multivariate model can better capture the covariance structure, yielding more accurate volatility forecasts (Bollerslev, 1990). Furthermore, a multivariate model is superior when correlations among financial assets change systematically over time. Typical examples include crises, during which correlations tend to strengthen (e.g., contagion effects in financial crises)(Forbes & Rigobon, 2002). The existence of volatility spillover effects between different markets or instruments (e.g., from equities to bonds or from currencies to commodities) also favors the use of a multivariate model, as it can better capture these dynamics (Diebold & Yilmaz, 2009). Examples of such volatility spillovers include interest rate changes and their impact on stock market volatility, or fluctuations in commodity prices (e.g. Brent Crude Oil) affecting the volatility of currencies in

commodity-exporting countries (Andersen et al., 2007). Additionally, the presence of regime shifts or structural breaks suggests the use of multivariate models, as correlations between all assets often increase significantly during financial crises (Ang & Bekaert, 2002). Lastly, exogenous factors affecting multiple time series, such as macroeconomic or fundamental variables influencing several financial instruments simultaneously (e.g. interest rates, inflation, geopolitical risks), can be incorporated more consistently in a multivariate model (Chen et al., 1986). For instance, including interest rate changes in a model that jointly considers stocks and bonds can improve forecasting accuracy. However, there are also arguments in favor of a univariate forecasting model. If correlations among time series are low, the covariance structure adds little value to the forecast (Hamilton, 2020). This may occur when the data are dominated by idiosyncratic factors specific to individual financial instruments (e.g. firm-specific volatility). Additionally, if the data history is short, a multivariate model may not be estimated reliably, leading to biased parameter estimates. Finally, computational complexity may also play a decisive role, as multivariate models require significantly more resources than univariate models. As outlined in section 4.4., multiple indicators suggest that a multivariate forecasting approach is more appropriate for the data. I demonstrated that many of the financial assets analyzed are correlated over the entire dataset period (01/01/2019 to 06/30/2024). Notably, I observed correlations between the S&P 500 and NASDAQ, the S&P 500 and U.S. Treasury Bonds, Gold and U.S. Treasury Bonds, NASDAQ and U.S. Treasury Bonds, as well as between the Nikkei and Hang Seng indices. This indicates significant interdependencies in the volatility of the examined financial assets. My results showed that correlations among financial assets intensified during the geopolitical shock periods under analysis. Additionally, I observed regime shifts in the volatility of the data, particularly during the analyzed geopolitical shock periods. These findings provide strong evidence that a multivariate forecasting approach is superior to a univariate approach.

Next, I discuss the empirical results in detail. I begin with an overview and then analyze individual financial assets. For most of the analyzed financial assets, I found that the BL-TFT model outperforms the UV-TFT during specific periods of geopolitical shocks. I evaluated the average SMAPE across forecast horizons 1-10 for each financial asset and each geopolitical shock period. Out of a total of 27 performance evaluations (9 financial assets across 3 periods), the BL-TFT outperformed the UV-TFT in 23 cases (82.14%). The DM*-test statistic at a significance niveau of 0.05 further indicates that 20 of these 23 cases (86.95%) exhibit a statistically significant difference between the multivariate and univariate approaches. My results vary across the examined periods. During the COVID-19 crisis, the BL-TFT outperformed the UV-TFT for all financial assets, with statistically significant differences in all cases. During the Russian invasion of Ukraine, the BL-TFT outperformed the UV-TFT in 7 out of 9 SMAPE evaluations, with 6 of these 7 cases showing statistically significant differences. During the Hamas attack on Israel, the BL-TFT outperformed the UV-TFT for 6 out of 9 financial assets, with statistically significant differences in 5 of these cases. These findings indicate that during the COVID-19 period, which is characterized by the highest average financial market volatility (Table 9), the BL-TFT consistently outperformed the UV-TFT. During the Russian invasion, which exhibited the second-

highest average volatility, the BL-TFT outperformed the UV-TFT in fewer cases than during the COVID-19 period but more frequently than during the Hamas attack, which exhibited the lowest average volatility. This can be attributed to the increased correlations among financial assets during major crises, a phenomenon known as the contagion effect in financial markets.

The contagion effect in financial markets arises from interconnectedness, where shocks propagate across different markets, creating a ripple effect. Understanding the mechanisms driving this phenomenon has gained significant traction in empirical studies, particularly in the context of financial crises and market volatilities (citation). Several studies have examined this effect and identified comparable patterns, e.g., Matkovskyy & Jalan (2019), Muzindutsi et al. (2022), Naeem et al. (2022), and Xu & Gao (2019).

Finally, I discuss individual financial assets. The BL-TFT outperforms the UV-TFT for the S&P 500 and NASDAQ during the COVID-19 crisis and the Russian invasion. However, during the Hamas attack, the UV-TFT performed better. Although these two indices remain highly correlated during the Hamas attack, their correlation with other financial assets is lower, likely leading to overfitting in the BL-TFT and, consequently, worse performance than the UV-TFT. For the Nikkei and Hang Seng indices, the BL-TFT statistically significantly outperforms the UV-TFT during the COVID-19 crisis and the Hamas attack. Although the BL-TFT performs better during the Russian invasion, the difference is not statistically significant. For commodities such as Gold and Brent Crude Oil, the BL-TFT significantly outperforms the UV-TFT during the COVID-19 crisis and the Russian invasion. Additionally, for Brent Crude Oil, the BL-TFT outperformed the UV-TFT during the Hamas attack. In contrast, for Gold, the BL-TFT performed better during the Hamas attack, but the difference is not statistically significant. Similar patterns are observed for other financial assets, reinforcing the findings. Summarizing, the analysis reveals that the multivariate approach consistently outperforms the univariate model across all considered financial instruments and crisis periods, including the COVID-19 pandemic, the Russian invasion of Ukraine, and the Hamas attack on Israel. The improvement in forecast accuracy was statistically significant in most cases, as confirmed by adjusted Diebold-Mariano tests. The performance advantage of the multivariate approach was especially pronounced during periods characterized by abrupt and nonlinear changes in volatility. A primary interpretation for this superiority lies in the multivariate model's ability to account for interdependencies and co-movements between different financial instruments, which are particularly relevant during systemic events. Geopolitical shocks often trigger widespread reactions across markets, leading to simultaneous changes in volatility that a multivariate model can exploit. Furthermore, the inclusion of exogenous variables such as the VIX and the Geopolitical Risk Index enables the multivariate approach to integrate relevant macro-financial information, improving its responsiveness to structural breaks. In contrast, univariate models process each asset in isolation, thereby missing potential predictive signals from related markets and external indicators. This structural limitation makes them less effective in capturing the full dynamics of volatility during crises. Overall, the findings demonstrate that multivariate models offer a more robust and adaptive framework for volatility forecasting under conditions of elevated geopolitical risk (Research Question 11).

Table 45: SMAPE for univariate multi-step out-of-sample forecasts during COVID-19 (02/03/2020 – 04/30/2020). SMAPE values are averages from different hyperparameter combinations (attention head = 1, 4, 8; hidden size = 16, 32, 64). Adjusted Diebold-Mariano test (DM*) for significant difference between baseline TFT (BL-TFT) (Table 33) and univariate TFT (UV-TFT) for financial instruments in scope and forecast horizon.

* indicate that the multivariate TFT (BL-TFT) outperformed the UV-TFT significantly on 5% significance niveau.

Forecast Horizon	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
1	0.7251*	1.0039*	0.6593*	0.7134*	0.7697*	0.7045*	0.8288	1.1153	1.4945*
2	0.6950	1.0140*	0.6172*	0.7145*	0.8007*	0.7396*	0.8642*	1.2897*	1.4008*
3	0.7397	1.0244*	0.6847*	0.7247*	0.7544*	0.7204*	0.9208*	1.2935*	1.4629*
4	0.7504*	1.0297*	0.7896*	0.6716*	0.8368*	0.6772*	0.8322	1.2724*	1.4757*
5	0.7474*	1.0373*	0.8192*	0.7139*	0.8281*	0.7071*	0.8124	1.3169*	1.3675*
6	0.8001*	1.1022*	0.9892*	0.6780*	0.8324*	0.7672*	0.8223	1.2573	1.2607*
7	0.7787*	0.9916*	0.8233*	0.6963*	0.8254*	0.7195*	0.8239*	1.3419*	1.2676*
8	0.8447*	1.0387*	0.8470*	0.6774*	0.7897*	0.6957*	0.7989*	1.3592*	1.1986*
9	0.8163*	0.9571*	0.8384*	0.6335*	0.8121*	0.7469*	0.8083*	1.3169*	1.2984*
10	0.8283*	0.9863*	0.9502*	0.6670*	0.8215*	0.7984*	0.8068*	1.3181*	1.2131*
Average	0.7726*	1.0185*	0.8018*	0.6890*	0.8071*	0.7277*	0.8349*	1.2919*	1.3440*

Source: Own calculations using Python PyTorch (2025).

Table 46: SMAPE for univariate multi-step out-of-sample forecasts during the Russian invasion (02/01/2022 – 04/29/2022). SMAPE values are averages from different hyperparameter combinations (attention head = 1, 4, 8; hidden size = 16, 32, 64). Adjusted Diebold-Mariano test (DM*) for significant difference between baseline TFT (BL-TFT) (Table 33) and univariate TFT (UV-TFT) for financial instruments in scope and forecast horizon. * indicate that the multivariate TFT (BL-TFT) outperformed the UV-TFT significantly on 5% significance niveau.

Forecast Horizon	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
1	0.5459*	0.4252	0.6152*	0.8793*	0.4762*	0.5750*	0.5276*	0.8395*	0.6125*
2	0.5286*	0.4310*	0.6361*	0.9195*	0.5083*	0.6029*	0.5636*	0.8223*	0.6050*
3	0.5342*	0.4225	0.6633*	0.8615*	0.4799*	0.5991*	0.5733*	0.8673*	0.5937*
4	0.5462*	0.4277	0.7788*	0.8953*	0.5051	0.6394*	0.5421	0.8142*	0.5781*
5	0.5147	0.4113	0.7036	0.8601*	0.4738*	0.6526*	0.5474	0.7852*	0.5780*
6	0.5283*	0.4128	0.6562	0.8602*	0.4898	0.6104*	0.5250	0.7377*	0.5448
7	0.5840*	0.4242	0.6881	0.9254*	0.4705	0.5975	0.5346	0.7451*	0.5422
8	0.5876*	0.4395	0.6864	0.8169	0.4632	0.5599*	0.5057	0.7103*	0.5237*
9	0.5678*	0.4196	0.6966	0.8043	0.4637*	0.6313*	0.4995	0.7795*	0.5194
10	0.5716	0.4399	0.6772	0.8124*	0.4697*	0.5857	0.5145	0.7737*	0.5463*
Average	0.5509*	0.4254	0.6802	0.8635*	0.4800*	0.6054	0.5333	0.7875*	0.5642*

Source: Own calculations using Python PyTorch (2025).

Table 47: SMAPE for univariate multi-step out-of-sample forecasts during the Hamas attack (10/01/2023 – 12/29/2023). SMAPE values are averages from different hyperparameter combinations (attention head = 1, 4, 8; hidden size = 16, 32, 64). Adjusted Diebold-Mariano test (DM*) for significant difference between baseline TFT (BL-TFT) (Table 33) and univariate TFT (UV-TFT) for financial instruments in scope and forecast horizon.

* indicate that the multivariate TFT (BL-TFT) outperformed the UV-TFT significantly on 5% significance niveau.

Forecast Horizon	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
1	0.7932	0.7001	0.7425	0.7379*	1.1945	0.6400*	0.5967*	0.5813*	0.7506
2	0.8061	0.7411	0.7990*	0.7123*	1.1211	0.6110	0.5681	0.5559*	0.6828
3	0.7497	0.7078	0.8290	0.6436	1.1978	0.6710*	0.5669*	0.5295*	0.7099
4	0.7456	0.6819	0.8516*	0.7115*	1.2239	0.6542*	0.5301	0.5488*	0.6592
5	0.7650	0.7168	0.8525*	0.6379*	1.4343	0.6112*	0.5496*	0.5639*	0.6960
6	0.7613	0.7660	0.8297*	0.6300	1.5244	0.6458*	0.5499	0.6000*	0.7024
7	0.8025	0.7673	0.8703*	0.7302*	1.9006	0.5871	0.5928	0.6090*	0.6564
8	0.8151	0.7866	0.8894*	0.7340*	1.3971	0.5836*	0.6049*	0.6258*	0.6470
9	0.8769	0.8580	0.8927*	0.6784	1.7110	0.5850*	0.6301*	0.5546	0.6375
10	0.9336	0.9385	0.8840	0.7424*	1.6738	0.5876	0.5874*	0.5793*	0.6441
Average	0.8049	0.7664	0.8441	0.6958*	1.4378	0.6176*	0.5776*	0.5748*	0.6786

Source: Own calculations using Python PyTorch (2025).

Table 48: Adjusted Diebold-Mariano test (DM*) statistic (p-value) for comparing the multi-step out-of-sample forecasting performance of the univariate TFT (UV-TFT) model with the baseline TFT model (BL-TFT, Tables 33-35). The null hypothesis assumes that the average 1-10 forecast horizon SMAPE of the UV-TFT is equal to the BL-TFT. * indicates, that the null hypothesis could be rejected on a 5% significance niveau and thus the multivariate forecasting model of the BL-TFT outperformed the univariate forecasting model of the UV-TFT for financial instruments in scope.

Period	S&P 500	NASDAQ	Nikkei225	Hang Seng	Gold	Brent Oil	10yr. T-Bond	EUR/USD	BTC
COVID-19	5.0707* (0.0002)	14.9754* (0.0000)	2.9206* (0.0076)	7.8326* (0.0000)	9.3694* (0.0000)	11.5922* (0.0000)	3.4041* (0.0034)	4.1586* (0.0010)	34.1380* (0.0000)
Russian Invasion	5.2819* (0.0002)	-3.5481 (0.9974)	0.3682 (0.3602)	6.3669* (0.0000)	4.0320* (0.0012)	7.0214* (0.0000)	-1.4928 (0.9168)	19.4486* (0.0000)	4.2475* (0.0008)
Hamas Attack	-2,2376 (0.9754)	-1,2617 (0.8822)	5,8248* (0.0001)	5,4425* (0.0001)	1,2031 (0.1283)	6,0103* (0.0001)	4,6145* (0.0005)	9,9126* (0.0000)	-0,3936 (0.6489)

Source: Own calculations using Python PyTorch (2025).

4.5.5. Model comparison between GARCH and Temporal Fusion Transformer

Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models have long served as the methodological standard for capturing time-varying conditional variances in asset returns (Bollerslev, 1986; Engle, 1982). GARCH models are parametric econometric models predicated on a recursive structure that captures volatility clustering in financial returns by modeling the conditional variance as a function of past squared residuals and previous conditional variances. These models are typically estimated using maximum likelihood methods, and their asymptotic properties are well established (Francq & Zakoian, 2019). Moreover, extensions such as EGARCH (Nelson, 1991) and GJR-GARCH (Glosten et al., 1993) allow for asymmetric responses to positive and negative shocks, capturing the so-called leverage effect observed in equity markets. In recent years, however, the rise of deep learning has introduced novel model classes, such as the Temporal Fusion Transformer (TFT), which offer an unprecedented level of flexibility in processing complex, high-dimensional, and non-stationary time series data (Lim et al., 2021). The TFT is a non-parametric, data-driven deep learning architecture designed for multi-horizon forecasting. It integrates recurrent neural networks (GRUs), attention mechanisms, gating layers, and static covariate encoders within a unified sequence-to-sequence learning framework (Lim et al., 2021).

Unlike GARCH models, which target the second moment of a univariate series, the TFT is capable of forecasting multiple targets and can flexibly incorporate both temporal and cross-sectional covariates, rendering them suitable for high-dimensional forecasting tasks where explanatory variables may exhibit non-linear interactions. A key axis of differentiation between these two model families lies in the nature and extent of their data requirements. GARCH-type models are most effective when applied to stationary, univariate financial return series. The underlying assumptions typically include weak stationarity and finite conditional moments, and thus the model performance is highly sensitive to deviations from these assumptions (Tsay, 2010). As such, financial price series are generally transformed into logarithmic returns prior to modeling. Furthermore, GARCH models require equally spaced, regularly sampled data and perform poorly in the presence of missing values or irregular observation intervals unless explicit adjustments are made. In contrast, the Temporal Fusion Transformer is explicitly designed to handle more complex data environments. It does not require stationarity or any pre-specified data distribution, and it can incorporate missing values through learned embeddings and temporal masks. TFTs distinguish between observed historical values, known future inputs, and static features, thus allowing for rich context-dependent learning (Lim et al., 2021). However, this modeling power necessitates substantial data availability. Due to their high parameterization, TFTs are data-intensive and require large-scale datasets for effective training and generalization. Regularization strategies such as dropout and early stopping are essential to mitigate the risk of overfitting (Goodfellow et al., 2016).

There are also differences with regard to the applicability of the two models. GARCH models are designed explicitly for volatility forecasting and serve as foundational tools for Value-

at-Risk (VaR) estimation, option pricing, and stress testing in regulatory contexts (Jorion, 1997). Their parameters are economically interpretable, providing direct insights into volatility persistence, mean-reversion, and shock asymmetry. In contrast, the TFT is a general-purpose forecasting model. While it can be applied to volatility forecasting, its primary strength lies in handling complex temporal patterns across multiple time series and covariates. Recent advancements in model interpretability, particularly the use of attention scores and variable importance metrics, have enhanced the transparency of deep learning models, including the TFT (Lim et al., 2021). Nevertheless, the economic interpretability of TFT remains limited relative to GARCH models due to its black-box nature and absence of a structural economic foundation. On the other hand, GARCH-type models exhibit limited robustness to missing data, as their estimation procedures typically assume a complete and regularly spaced time series. Gaps in the data often necessitate imputation or interpolation prior to estimation, which may introduce biases or distort volatility dynamics (Tsay, 2010). In contrast, the Temporal Fusion Transformer is inherently more robust to missingness, employing temporal masking and learned embeddings that allow it to model incomplete sequences without requiring explicit imputation (Lim et al., 2021). This feature renders the TFT particularly advantageous in real-world financial environments, where data irregularities and reporting lags are common. Regarding the computational burden, the GARCH models have an advantage, due to its relatively low, owing to their parsimonious structure and closed-form likelihood functions, which facilitate efficient numerical optimization even in rolling-window or high-frequency settings. By contrast, the Temporal Fusion Transformer entails substantially higher computational complexity. Its training involves backpropagation through multiple layers of recurrent units, attention heads, and gating mechanisms, often requiring specialized hardware (e.g., GPUs) and considerable training time (Goodfellow et al., 2016; Lim et al., 2021). This increased complexity, while justifiable in large-scale forecasting tasks, may limit the practicality of TFTs in time-sensitive or resource-constrained financial applications.

Summarizing, GARCH models require stationary, univariate time series with complete data and are highly interpretable due to their economically meaningful parameters. In contrast, the Temporal Fusion Transformer (TFT) handles non-stationary, multivariate data with missing values and can incorporate rich covariate information, but at the cost of high computational complexity. While GARCH models are efficient and transparent, TFTs demand large datasets and significant resources, offering flexibility in exchange for reduced interpretability.

5. CONCLUSION AND RECOMMENDATIONS

The world is experiencing repeated shocks that do not remain local in scope but have a global impact. These geopolitical shocks often arise from military conflicts, political unrest, or sudden changes in international relations and affect the most diverse facets of the social life, including economics and thus financial markets. Furthermore, geopolitical shocks impact the volatility of all assets, asset classes, sectors, and countries worldwide which in turn leads to changes in the behavior of international market players. It is obvious that these phases must be managed very cautiously by market participants and financial market supervisors to avoid devastating crashes. This dissertation has presented a comprehensive empirical investigation into the identification and forecasting of market dynamics under the influence of major geopolitical shocks. Across four methodologically distinct but thematically interrelated chapters, this work examined the behavior of abnormal returns in global equity sectors, the emergence and persistence of abnormal volatility, the performance of GARCH-class volatility forecasting models, and, as the main part of the analysis, the predictive capacity of deep learning model Temporal Fusion Transformer (TFT), under crisis conditions. Using the COVID-19 pandemic, the Russian invasion of Ukraine, and the Hamas attack on Israel as temporal anchors for geopolitical stress, the analysis provides new insights into both, the empirical behavior of financial markets and the methodological tools suitable for capturing such behavior. To provide the most comprehensive assessment of the results, I have evaluated major capital markets financial assets. I used the important MSCI World Sector indices to analyze the abnormal returns and abnormal volatility. These indices provide a classification of the most common business models worldwide. To examine forecasting performance in times characterized by geopolitical events, I considered a wide range of the most essential financial instruments from different investment classes: major stock indices (S&P 500, NASDAQ, Nikkei 225 and Hang Seng), commodities (Gold, Brent Crude Oil), fixed income instruments (10 year US-Treasuries), foreign exchange rates (EUR-USD exchange rate) and cryptocurrency (Bitcoin).

This research has explored financial market dynamics in the context of major geopolitical shocks by analyzing the emergence of abnormal and cumulative abnormal returns, the detection and persistence of abnormal volatility, and the efficacy of econometric volatility forecasting models. Focusing primarily on the Russian invasion of Ukraine in February 2022 but also extending to other important geopolitical stress events such as the COVID-19 pandemic and the Hamas-Israel conflict, the three empirical sections contribute distinct perspectives on how global financial markets absorb and respond to sudden geopolitical disruptions. First, I employed an event study methodology, examined abnormal returns across eleven sector indices of the MSCI World index. The empirical results demonstrate that the invasion triggered an immediate and predominantly negative response among investors, with statistically significant negative abnormal returns concentrated on the event day in sectors such as energy, financials, and materials (Research Question 1). However, market reactions varied substantially across sectors. Several sectors, including healthcare, information technology, and real estate, exhibited positive cumulative abnormal returns over the subsequent 25 trading days, suggesting divergent investor expectations

regarding the long-term sectoral impact of the conflict (Research Question 2). Notably, the analysis revealed a short-lived corrective effect on the day following the invasion, during which most sectors showed positive abnormal returns, potentially reflecting rapid re-evaluation of risk or market overreaction. These findings indicate the importance of sectoral heterogeneity in return dynamics and underscore the role of investor sentiment in the immediate aftermath of geopolitical shocks.

Second, I focused on the volatility dimension of the same event, applying the Abnormal Volatility and Cumulative Abnormal Volatility frameworks to analyze excess volatility beyond GARCH(1,1)-based forecasts. The results confirmed that the event induced significant spikes in volatility across nearly all sectors, with the most pronounced reactions observed in financials, consumer staples, and materials (Research Question 3). The persistence of abnormal volatility was evident for up to ten trading days post-event in several sectors, including healthcare, IT, and industrials. These dynamics reflect both heightened investor uncertainty and the varying temporal profiles of volatility normalization across sectors (Research Question 4). Furthermore, sector pairwise comparisons of volatility behavior revealed increased convergence in volatility levels over time, indicating a gradual alignment of market participants' risk assessments in the post-shock period (Research Question 5). Third, I evaluated the forecasting performance of 75 different GARCH-type models, including standard GARCH, EGARCH, and GJR-GARCH variants, across nine major financial instruments during episodes of geopolitical turmoil. The empirical analysis found that the EGARCH model consistently outperformed its counterparts in terms of in-sample fit (Research Question 6) and out-of-sample point forecasting accuracy, particularly over medium and longer forecast horizons (Research Question 7). This superior performance is attributed to the EGARCH model's ability to capture asymmetric responses (leverage effects) and the long memory structure of volatility. In contrast, for directional forecasting accuracy, the GJR-GARCH model often performed better, especially for short-horizon forecasts and during high-volatility regimes. These findings reinforce the need to tailor model choice to the specific decision-making context, whether for risk estimation, hedging, or speculative positioning.

The fourth and most extensive empirical chapter extended the analysis into the domain of deep learning-based volatility forecasting, leveraging the Temporal Fusion Transformer (TFT) for direct multi-step, multivariate out-of-sample prediction. Across the COVID-19 pandemic, the Russian invasion, and the Hamas attack on Israel, the TFT consistently outperformed GARCH-class models in forecasting volatility for most financial instruments and forecast horizons. Statistical superiority was established using robust Diebold-Mariano tests, with the TFT showing particular strength in short-horizon forecasts, likely due to its architectural emphasis on recent high-frequency dynamics and ability to adapt to regime shifts (Research Question 8). Beyond baseline multivariate TFT, the study introduced two enhancements: a regime-switching variant (RS-TFT) incorporating structural shift information, and hyperparameter-optimized architectures. The RS-TFT model achieved statistically significant improvements in 78% of cases over the baseline TFT, with an average SMAPE reduction of 4.53%. The gains were particularly pronounced during the Hamas conflict, despite its relatively lower volatility regime, suggesting

that regime labels can serve as corrective signals that improve long-horizon performance (Research Question 9). The hyperparameter optimization analysis further demonstrated that architectural decisions (e.g., attention heads and hidden size) substantially affect forecast accuracy. The best-performing models commonly featured 8 attention heads and a hidden size of 32, striking a balance between representational power and regularization (Research Question 10).

Finally, a comparative analysis of univariate versus multivariate TFT setups demonstrated the superiority of the multivariate approach in 82% of evaluations, with statistically significant differences in most cases (Research Question 11). These findings are attributed to volatility contagion effects, co-movement in financial instruments, and the advantage of incorporating macro-financial exogenous variables such as the VIX and Geopolitical Risk Index (GPR).

The results of this research provide important implications for both academic work in financial econometrics and practical applications in risk management, asset allocation, and financial supervisory. The methodological contributions of this study lied primarily in the integration of classical econometric models with modern deep learning architectures in the context of geopolitical shocks. Empirically, the work shows that market reactions to geopolitical events are characterized by significant and heterogeneous shifts in returns and volatility, both of which demand flexible and adaptive forecasting approaches. The incorporation of regime-dependent modeling, hyperparameter optimization, and multivariate forecasting strategies into the volatility modeling framework offers several valuable insights into how financial forecasting models can be improved under conditions of structural instability.

From a methodological standpoint, the analysis demonstrated the limitations of traditional GARCH-type models in capturing the full dynamics of financial volatility during crisis periods. While EGARCH and GJR-GARCH models remain useful, particularly for capturing leverage effects and directional changes in volatility, their capacity to accommodate sudden regime changes is limited. These models tend to respond gradually to abrupt shifts in volatility, making them less suitable in environments characterized by non-linearities, contagion, and structural breaks. In contrast, the Temporal Fusion Transformer (TFT) exhibits a stronger capacity for short-term adaptability, effectively modeling complex interactions across assets, and outperforming GARCH-class models in most cases, especially during the COVID-19 pandemic and the Russian invasion of Ukraine. The integration of attention mechanisms, multivariate input structures, and flexible horizon modeling allows the TFT to internalize temporal dependencies and contextual information, which are especially important during systemic events. A particularly important methodological contribution of this dissertation is the implementation of a regime-switching feature into the TFT architecture. By incorporating an explicit binary indicator of high-volatility regimes, the RS-TFT was able to recognize and adjust to shifts in volatility dynamics that occurred during geopolitical shocks.

The empirical results indicated that this feature significantly improves the forecast accuracy of the model, particularly in less volatile crisis periods, such as the Hamas-Israel conflict, where volatility shifts are more subtle. These findings suggest that even in sophisticated learning architectures, the inclusion of domain-informed structural signals can serve as effective

mechanisms to anchor model behavior, especially for medium- and long-horizon forecasts. This supports the hypothesis that volatility forecasting is not merely a statistical challenge but also a contextual modeling problem where temporal regimes, policy responses, and investor sentiment should be explicitly encoded. The hyperparameter optimization of the TFT also yields relevant practical insights. Across all assets and shock periods, models with a higher number of attention heads and a moderate hidden layer size consistently outperformed less complex configurations. This finding highlights the importance of architectural tuning in machine learning models applied to financial time series. Specifically, configurations with eight attention heads and a hidden size of 32 offered a good trade-off between model expressiveness and generalization ability, avoiding both underfitting and overfitting. This underscores that model complexity must be aligned with the underlying data structure and volatility dynamics: overly simplistic models fail to capture key interactions, while overly complex models risk overfitting, especially when forecast horizons extend into periods of elevated uncertainty.

Beyond methodological considerations, the findings have several practical implications for financial institutions, asset managers, and policymakers. First, risk management systems should incorporate adaptive, short-horizon forecasting models such as the TFT to monitor and respond to sudden spikes in market volatility. The demonstrated superiority of the TFT in capturing regime shifts and short-term volatility dynamics suggests that such models can serve as early warning systems during geopolitical shocks. In addition, the integration of geopolitical risk indicators, such as the Geopolitical Risk Index (Caldara & Iacoviello, 2022) as exogenous inputs offers a pathway for enhancing model responsiveness to real-world events. These indicators can be updated dynamically and incorporated into forecasting engines to improve the timeliness and accuracy of risk predictions. Second, asset allocation strategies should be guided by volatility diagnostics that are sensitive to both, cross-asset correlations and sector-specific reactions. The abnormal return and volatility results demonstrated that different sectors and asset classes exhibit varied sensitivities to geopolitical events. This suggests that sector-level or asset-specific reallocation strategies, particularly those based on short-term indicators of investor uncertainty may outperform traditional static allocation models during crisis periods. Practitioners can use volatility persistence measures and sectoral heterogeneity insights to calibrate hedge ratios or adjust portfolio weights dynamically. In addition, you should also optimize your rolling volatility forecasting to be able to make a decision on adjustments that depend on volatility in order to hedge the portfolio. Third, model selection should be aligned with the specific decision-making objective. For point forecasts of volatility—important in value-at-risk calculations, options pricing, and portfolio optimization, the EGARCH and TFT models offered strong performance. However, for applications where directional changes in volatility are more critical, such as in volatility trading or regime-switching asset strategies, the GJR-GARCH model remains a valuable tool. The performance trade-offs between different model classes underscore the importance of using a multi-model forecasting framework tailored to the intended application, rather than relying on a single model across all use cases. Finally, the consistent outperformance of the multivariate

TFT over its univariate counterpart during high-volatility periods confirms the importance of modeling financial assets as a system of interrelated time series, rather than as isolated processes.

Geopolitical shocks rarely affect a single market in isolation; instead, they often propagate across asset classes through interest rate channels, currency adjustments, commodity prices, and global capital flows. Multivariate models that can capture these spillover effects offer superior forecasting capabilities, particularly when crisis dynamics are global in nature. This finding aligns with theoretical models of contagion and volatility transmission and suggests that forecast engines in practice should increasingly adopt multivariate, cross-asset frameworks. Taken together, the implications of this research are clear: as the frequency and impact of geopolitical shocks continue to rise, financial institutions must equip themselves with models that are both methodologically sophisticated and operationally adaptable. The combined use of classical econometric tools, deep learning architectures, structural regime indicators, and multivariate modeling constitutes a forward-looking strategy for forecasting market dynamics under uncertainty. Future research should continue to develop and validate hybrid models that combine interpretability, flexibility, and scalability to enhance decision-making in the presence of systemic risk.

Although, these findings carry substantial implications for several key domains of professional financial practice, particularly in light of the growing need for robust, adaptive risk modeling under conditions of structural uncertainty and exogenous shocks. In the domain of market risk management, accurate short- to medium-term volatility forecasts are critical for the calculation of Value-at-Risk (VaR) and Expected Shortfall (ES), both of which are subject to regulatory scrutiny under Basel III/IV. The demonstrated outperformance of the TFT in periods of elevated geopolitical tension, where traditional GARCH-type models often underreact to rapidly evolving risk landscapes, suggests that TFT-based volatility estimates may lead to more responsive and accurate VaR models. In practice, this enables risk managers to allocate capital more efficiently, avoid excessive procyclicality, and comply with supervisory stress-testing regimes more effectively. The superior multi-horizon forecasting accuracy of the TFT directly also benefits portfolio managers engaged in dynamic asset allocation strategies, especially under heightened geopolitical risk. Because it captures nonlinear dependencies and integrates multivariate covariates, including macroeconomic indicators and geopolitical risk indices, the TFT allows for more timely and accurate adjustments in portfolio weightings in response to anticipated changes in asset-specific risk. This is particularly advantageous for tactical allocation strategies and volatility-managed portfolios, which rely on precise forecasts of forward-looking risk to optimize performance under regime uncertainty. Options pricing and delta-hedging strategies are also highly sensitive to anticipated volatility over short horizons.

The ability of the TFT to outperform GARCH models across all tested time horizons (1–10 days) suggests that incorporating TFT-based forecasts into implied volatility surfaces or volatility modeling frameworks (e.g., for local volatility or stochastic volatility models) can improve pricing accuracy and hedging efficiency, particularly during volatile periods induced by geopolitical events. This has direct implications for trading desks managing option books, structured products, or volatility arbitrage strategies. Further, supervisory stress tests and internal

scenario analyses increasingly incorporate non-financial risk drivers, including geopolitical and macro-political shocks. The multivariate nature of the TFT, combined with its empirical robustness under geopolitical stress, makes it an ideal tool for modeling conditional volatilities under adverse scenarios. Unlike GARCH models, which rely heavily on past return dynamics, the TFT can integrate forward-looking indicators and generate stress-consistent volatility projections across multiple asset classes. This enables institutions to better quantify systemic risk, assess liquidity needs, and enhance capital planning. Another field of application are insurance companies and pension funds managing long-duration liabilities are increasingly exposed to geopolitical tail risks. Given the TFT's superior performance across longer short-term horizons (up to 10 days), integrating its forecasts into ALM frameworks may improve short-term rebalancing decisions and hedge effectiveness. In particular, insurers operating under Solvency II and similar regulatory regimes can benefit from enhanced short-horizon risk forecasting to optimize capital buffers and dynamic hedging strategies for equity-linked liabilities. Finally, for central banks and financial stability authorities, understanding volatility dynamics during geopolitical crises is essential for assessing systemic vulnerabilities. TFT-based models, with their ability to integrate high-dimensional inputs such as geopolitical risk indices, commodity prices, and cross-asset volatilities, provide a modern tool for real-time financial surveillance. The outperformance of the TFT in all instruments and horizons during geopolitical stress suggests that such models could be incorporated into early-warning systems and systemic risk dashboards to support policy decision-making.

Summarizing, the consistent outperformance of the multivariate Temporal Fusion Transformer across all tested horizons, instruments, and shock periods establishes it as a robust and highly adaptable tool for volatility forecasting in modern finance. Its superior responsiveness to geopolitical shocks, ability to process multivariate inputs, and effectiveness across time scales make it particularly well-suited for high-stakes applications in risk management, derivatives pricing, asset allocation, and regulatory oversight. Practitioners are thus encouraged to integrate TFT-based models into existing forecasting pipelines—either as a complement or as a replacement for conventional GARCH-type approaches—particularly in environments marked by elevated uncertainty and structural change. In light of the documented outperformance of the Temporal Fusion Transformer (TFT) in volatility forecasting—particularly within a multivariate setting and during periods of geopolitical shocks—a key question for practitioners is how to balance forecasting performance against model interpretability and implementation complexity when selecting an appropriate modeling framework. While the superior predictive accuracy of the TFT across all examined scenarios clearly supports its practical use, especially when additional explanatory variables are available and responsiveness to exogenous shocks is critical, other dimensions remain highly relevant. Traditional GARCH-type models continue to offer notable advantages in terms of interpretability, regulatory acceptance, and ease of implementation within established risk infrastructures.

Practitioners should adopt a use-case-specific decision framework. In highly regulated environments where model transparency is prioritized and data availability is limited, the

continued application of GARCH models remains justified. However, in data-rich and rapidly evolving contexts—particularly where geopolitical, macroeconomic, or alternative data sources play a central role—the gradual integration of TFT-based forecasting systems is recommended. A hybrid modeling strategy, wherein TFT outputs serve as complementary signal generators or are used for benchmarking alongside conventional econometric models, represents a pragmatic approach to combining predictive strength with operational feasibility.

6. NEW SCIENTIFIC RESULTS

This dissertation presents several novel scientific contributions at the intersection of financial econometrics, time series forecasting, and machine learning by analyzing, modeling, and forecasting market dynamics during geopolitical shocks.

1. The research provided a volatility modeling and forecasting framework. Systematic testing of 75 GARCH-type models showed that EGARCH models consistently provided the best in-sample fit for volatility in three different periods of geopolitical shocks (Research Question 6). It was found that EGARCH models outperformed GARCH and GJR-GARCH-type models in point forecasts, while GJR-GARCH performed better in directional forecasts (Research Question 7).
2. Key methodological innovation is the application of the Temporal Fusion Transformer (TFT), a transformer-based deep learning network, for direct, multi-step, multivariate volatility forecasting. This research found that the TFT outperformed the best-performing GARCH-type model in almost all analyzed periods of geopolitical shocks and examined financial assets in the research scope (Research Question 8).
3. Further, this research could show that hyperparameter optimization of the Temporal Fusion Transformer significantly enhanced the forecasting performance, confirming that careful configuration and hybrid econometric deep learning approaches yield better forecasting outcomes (Research Question 10).
4. Finally, the improvement of a multivariate TFT setup was compared with a univariate TFT setup. The multivariate TFT approach provided superior short-term forecast accuracy, which was crucial during high-volatility episodes like March 2020 and the Russian invasion in 2022. This makes it particularly suitable for tactical risk management applications (Research Question 11).

7. SUMMARY

This dissertation addresses a central and increasingly urgent topic in financial econometrics: the modeling and forecasting of market dynamics under conditions of geopolitical uncertainty. The financial world has seen a notable rise in the frequency and magnitude of geopolitical shocks in recent years, from the outbreak of the COVID-19 pandemic in 2020, to the Russian Federation's invasion of Ukraine in 2022, and the Hamas terrorist attacks on Israel in late 2023. Each of these events triggered abrupt volatility regime shifts in global financial markets, leading to elevated uncertainty, structural breaks in return distributions, and large-scale capital reallocations. Traditional forecasting methods, including GARCH-type models, have often struggled to adapt swiftly and effectively to these regime changes. In response, this dissertation develops a comprehensive empirical and methodological framework to better identify, interpret, and forecast financial market behavior in the presence of such shocks. The overarching aim of this work is twofold. First, it seeks to empirically quantify how financial markets, both in terms of returns and volatility, react to geopolitical shocks, with a particular emphasis on heterogeneity across sectors and asset classes. Second, it advances the methodological frontier of volatility forecasting by incorporating and systematically evaluating deep learning models, with a focus on the Temporal Fusion Transformer (TFT), a state-of-the-art deep learning architecture, designed for interpretable multivariate time series forecasting.

The literature on financial market responses to geopolitical shocks is both growing and diverse, but several gaps remain. Much of the existing research focuses on narrow market segments—such as national stock indices or commodities, and primarily employs traditional econometric models (e.g., GARCH, EGARCH, Markov-switching, wavelets). While informative, these approaches often assume stationarity or linear volatility dynamics, which may not hold during geopolitical crises. Furthermore, previous work has frequently analyzed isolated events, limiting comparative assessments across different types of shocks. In contrast, recent developments in machine learning offer new avenues for modeling high-dimensional, nonlinear, and nonstationary financial time series, yet empirical applications in financial econometrics, particularly under extreme uncertainty, remain limited.

This dissertation seeks to bridge this gap by combining deep learning methods with established econometric tools in a unified framework. It extends the literature by providing the first global sector-level event study on the Russian invasion of Ukraine, a comparative cross-event analysis of market behavior during three distinct geopolitical crises, and a comprehensive evaluation of the forecasting performance of advanced multivariate neural models across a broad set of financial instruments.

The dissertation is organized around four empirical pillars. First, the identification of abnormal returns. Using an event study methodology, this section analyzes the return behavior of eleven MSCI World sector indices in the days surrounding the Russian attack on Ukraine. The results show significant negative abnormal returns on the event day across most sectors, with some recovery the following day. The cumulative abnormal returns display strong heterogeneity, reflecting differences in business model exposure to geopolitical risk. Notably, energy, financials,

and utilities exhibit sustained negative performance, while sectors such as healthcare and information technology show signs of resilience. The second part of this dissertation deals with the identification of abnormal volatility, by applying the AVOLA framework in conjunction with a GARCH(1,1) model benchmarks. It reveals significant and persistent abnormal volatility in the aftermath of the invasion, particularly in financials, consumer staples, and industrials (Research Question 3 and Research Question 4). Moreover, statistical tests on volatility convergence across sectors suggest an evolving consensus in investor sentiment following the shock (Research Question 5). Third, in terms of point accuracy, especially during the COVID-19 pandemic and the Russian GARCH-Class volatility forecasting, systematically evaluates 75 GARCH-type model configurations, including GARCH, EGARCH, and GJR-GARCH, for their in-sample fit and out-of-sample multi-horizon volatility forecasting ability. Applied to nine major financial instruments (S&P 500, NASDAQ, Nikkei 225, Hang Seng, Gold Brent Crude Oil, U.S. Treasury Bonds, EUR-USD exchange rate and Bitcoin), the results show that EGARCH models outperform other specifications in terms of in-sample-fit and out-of-sample performance for point forecasts evaluated by SMAPE. Conversely, GJR-GARCH models demonstrate stronger directional accuracy, particularly for short-term movements.

The fourth and major section introduces the Temporal Fusion Transformer, which is applied in univariate and multivariate settings, with and without a regime-switching component. The TFT is trained using a rolling-window, direct multi-step forecasting approach for 1–10 day forecast horizons across the same nine financial instruments and three geopolitical crises. The results are striking. The TFT consistently outperforms GARCH-class models in terms of both accuracy and adaptability, particularly for short-term forecasts (1–3 days), where it is able to learn and react quickly to structural breaks. Its performance remains robust for longer horizons, though with diminishing relative gains. The regime-switching enhancement, which uses a binary volatility indicator based on quantile thresholds, further improves medium- and long-horizon forecast precision.

This work offers several original contributions to the literature. The global sector-level return and volatility analysis of the Russian invasion, revealing differentiated market reactions across the MSCI World Index and providing new insights into sector-specific risk exposures. A cross-event comparative framework analyzing COVID-19, the Russian war, and the Hamas attack, enabling generalization of model performance and volatility behavior under varying types of geopolitical stress. By applying the Temporal Fusion Transformer for a multivariate forecasting setup across heterogeneous financial instruments, I demonstrated the benefits of capturing cross-asset dependencies, volatility spillovers, and systemic dynamics. Thereby, I found evidence of a short-horizon superiority of the TFT, which excels at identifying high-frequency volatility clusters and structural regime shifts in the immediate aftermath of geopolitical shocks. When introducing a regime-switching enhancement to the TFT, it significantly improved forecasting performance in contexts with muted but persistent volatility changes, which was particularly evident during the Hamas conflict. The systematic hyperparameter optimization of the TFT, showed that

configurations with 8 attention heads and moderate hidden layer sizes (32 units) deliver the best trade-off between expressiveness and generalization across diverse market conditions. Notably, the TFT achieved particularly high accuracy for the S&P 500 and NASDAQ during the Russian invasion, for U.S. Treasury Bonds and Brent Crude Oil during COVID-19, and for EUR/USD during the Hamas conflict. In contrast, traditional GARCH models still proved competitive for assets like gold and Bitcoin in periods of moderate volatility, underscoring the importance of asset-specific modeling strategies.

The findings of this dissertation carry meaningful implications for both academic researchers and financial practitioners. On the theoretical side, the research underscores the need to move beyond stationary, linear assumptions in econometric modeling and adopt more flexible, high-capacity forecasting architectures, particularly under crisis conditions. Practically, the results suggest that risk managers, institutional investors, and policymakers can enhance their volatility forecasting accuracy by integrating multivariate deep learning models like the TFT, especially when rapid shifts in market regimes are anticipated. Moreover, the ability of the TFT to incorporate regime indicators and macro-financial inputs (e.g., the VIX, the Geopolitical Risk Index) offers promising avenues for early-warning systems and tactical portfolio adjustments. The model's interpretability, through attention weights and variable selection, also makes it compatible with regulatory expectations for model transparency and accountability.

This dissertation contributes both conceptually and methodologically to the understanding of financial market behavior under geopolitical shocks. It demonstrates that while conventional econometric models offer useful baselines, deep learning models—particularly the Temporal Fusion Transformer—set a new standard for multivariate, multi-horizon volatility forecasting in crisis environments. By offering a rigorous comparative framework and an innovative modeling approach, this work lays the foundation for more adaptive, accurate, and context-aware forecasting in the face of global uncertainty.

APPENDIX 1: BIBLIOGRAPHY

- ADEBIYI, A. A., ADEWUMI, A. O., AYO, C. K. (2014): Comparison of ARIMA and artificial neural networks models for stock price prediction. In: *Journal of Applied Mathematics*, 2014(1), 614342. p.
- ADUDA, J., MASILA, J. M., ONSONGO, E. N. (2018): The determinants of financial performance of commercial banks in Kenya. In: *Journal of Finance and Accounting*, 6(1) 1–21. p.
- AGGARWAL, C. C. (2018): *Neural Networks and Deep learning: A Textbook*. Cham: Springer. 13 p. Springer Series in the Data Sciences (10)
- AGYEI, S. K. (2023): Emerging markets equities' response to geopolitical risk: Time-frequency evidence from the Russian-Ukrainian conflict era. In: *Heliyon*, 9(2). p.
- AHMED, N. K., ATIYA, A. F., GAYAR, N. E., EL-SHISHINY, H. (2010): An empirical comparison of machine learning models for time series forecasting. In: *Econometric Reviews*, 29(5-6) 594-621. p.
- AHMED, S., HASAN, M. M., KAMAL, M. R. (2023): Russia–Ukraine crisis: The effects on the European stock market. In: *European Financial Management*, 29(4) 1078-1118. p.
- AHMED, A. S., MCMARTIN, A. S., SAFDAR, I. (2020): Earnings volatility, ambiguity, and crisis-period stock returns. In: *Accounting & Finance*, 60(3) 2939-2963. p.
- AIZENMAN, J., LINDAHL, R., STENVALL, D., UDDIN, G. S. (2024): Geopolitical shocks and commodity market dynamics: New evidence from the Russia–Ukraine conflict. In: *European Journal of Political Economy*, 85 102574. p.
- AKAIKE, H. (1978): On the likelihood of a time series model. In: *Journal of the Royal Statistical Society: Series D (The Statistician)*, 27(3-4) 217-235. p.
- AKAIKE, H. (1979): A Bayesian extension of the minimum AIC procedure of autoregressive model fitting. In: *Biometrika*, 66(2) 237-242. p.
- AKBAR, M., GENOVÉS, M. (2024): South africa sues international court over israel's palestinian genocide under international law. In: *Lampung Journal of International Law*, 6(2) 83-94. p.
- ALAM, M. K., TABASH, M. I., BILLAH, M., KUMAR, S., ANAGREH, S. (2022): The impacts of the Russia–Ukraine invasion on global markets and commodities: a dynamic connectedness among G7 and BRIC markets. In: *Journal of Risk and Financial Management*, 15(8) 352. p.
- ALBERG, D., SHALIT, H., YOSEF, R. (2008): Estimating stock market volatility using asymmetric GARCH models. In: *Applied Financial Economics*, 18(15) 1201-1208. p.
- ALBULESCU, C. T. (2021): COVID-19 and the United States financial markets' volatility. In: *Finance Research Letters* 38, 101699. p.

- ALI, M., ALAM, N., RIZVI, S. (2020): Coronavirus (Covid-19) — an epidemic or pandemic for financial markets. In: *Journal of Behavioral and Experimental Finance*, 27 100341. p.
- ALI, S. R. M., ANIK, K. I., HASAN, M. N., KAMAL, M. R. (2023): Geopolitical threats, equity returns, and optimal hedging. In: *International Review of Financial Analysis*, 90 102835. p.
- ALIU, F., HAŠKOVÁ, S., BAJRA, U. Q. (2023): Consequences of Russian invasion on Ukraine: evidence from foreign exchange rates. In: *The Journal of Risk Finance* 24(1) 40-58. p.
- ALONSO, F., & CARBO, S. (2022). Artificial intelligence and explainability in financial services. In: *Journal of Financial Regulation and Compliance*, 30(3) 358–370. p.
- ALSOUS, M., AL MUHAISSEN, B., MASSAD, T., AL BSOUL, A. (2024): Exploring depression, PTSD, insomnia, and fibromyalgia symptoms in women exposed to Gaza war news: A community-based study from Jordan. In: *International Journal of Social Psychiatry*, 70(8) 1470–1480. p.
- ALTEMUR, N., EREN, B. S., KARACA, S. S. (2024): The impact of the Hamas-Israel conflict on stock market indices in the middle east and Turkey: an event study analysis. In: *Journal of Economics and Business Issues*, 4(1) 77-85. p.
- ANDERSEN, T. G., BOLLERSLEV, T., DIEBOLD, F. X., LABYS, P. (2001): The distribution of realized exchange rate volatility. In: *Journal of the American Statistical Association*, 96(453) 42–55. p.
- ANDERSEN, T. G., BOLLERSLEV, T., DIEBOLD, F. X., LABYS, P. (2003): Modeling and forecasting realized volatility. In: *Econometrica*, 71(2) 579-625. p.
- ANDERSEN, T. G., BOLLERSLEV, T., DIEBOLD, F. X., WU, J. (2005): A framework for exploring the macroeconomic determinants of systematic risk. In: *American Economic Review*, 95(2) 398-404. p.
- ANDERSEN, T. G., BOLLERSLEV, T., DIEBOLD, F. X., VEGA, C. (2007): Real-Time Price Discovery in Global Stock, Bond and Foreign Exchange Markets. In: *Journal of International Economics*, 73(2) 251–277. p.
- ANG, A., BEKAERT, G. (2002): International Asset Allocation with Regime Shifts. In: *Review of Financial Studies*, 15(4) 1137–1187. p.
- ANG, A., TIMMERMAN, A. (2012): Regime changes and financial markets. In: *Annual Review of Financial Economics*, 4(1) 313–337. p.
- ANG, G., LIM, E.-P. (2022): Guided attention multimodal multitask financial forecasting with inter-company relationships and global and local news. 6313-6326 p. In: MURESAN, S., NAKOV, P., VILLAVICENCIO, A. Villavicencio (Eds.): *PROCEEDINGS OF THE 60th ANNUAL MEETING OF THE ASSOCIATION OF FOR COMPUTATIONAL LINGUISTICS (VOLUME 1: LONG PAPERS)*. Association for Computational Linguistics.
- APERGIS, N., BONATO, M., GUPTA, R., KYEI, C. (2018): Does geopolitical risks predict stock returns and volatility of leading defense companies? Evidence from a nonparametric approach. In: *Defence and Peace Economics*, 29(6) 684-696. p.

- ARDIA, D., BLUTEAU, K., BOUDT, K., CATANIA, L., TROTTIER, D. A. (2019): Markov-switching GARCH models in R: The MSGARCH package. In: *Journal of Statistical Software*, 91 1-38. p.
- ASLAM, F., AZIZ, S., NGUYEN, D., MUGHAL, K., KHAN, M. (2020): On the efficiency of foreign exchange markets in times of the covid-19 pandemic. In: *Technological Forecasting and Social Change*, 161 120261. p.
- ASSEFA, T., SENGUPTA, S., WILLIAMS, S. (2022, October): The Effect of COVID-19 Announcement on the US Stock Market. 33-34p. In: PROCEEDINGS OF THE 45th ANNUAL MEETING OF THE NORTHEAST BUSINESS & ECONOMICS ASSOCIATION. Northeast Business & Economics Association.
- AWARTANI, B. M., CORRADI, V. (2005): Predicting the volatility of the S&P-500 stock index via GARCH models: the role of asymmetries. In: *International Journal of Forecasting*, 21(1) 167-183. p.
- BADSHAH, I. (2017): Volatility spillover from the fear index to developed and emerging markets. In: *Emerging Markets Finance and Trade*, 54(1) 27-40. p.
- BAEK, S., MOHANTY, S., GLAMBOSKY, M. (2020): Covid-19 and stock market volatility: an industry level analysis. In: *Finance Research Letters*, 37 101748. p.
- BAKER, S. R., BLOOM, N., DAVIS, S. J. (2016): Measuring economic policy uncertainty. In: *The Quarterly Journal of Economics*, 131(4) 1593-1636. p.
- BAKER, S. R., BLOOM, N., DAVIS, S. J., KOST, K., SAMMON, M., VIRATYOSIN, T. (2020): The unprecedented stock market reaction to COVID-19. In: *The Review of Asset Pricing Studies*, 10(4) 742-758. p.
- BAKSHI, G., CAO, C., CHEN, Z. (1997): Empirical performance of alternative option pricing models. In: *The Journal of Finance*, 52(5) 2003-2049. p.
- BALCILAR, M., BONATO, M., DEMIRER, R., GUPTA, R. (2018): Geopolitical risks and stock market dynamics of the BRICS. In: *Economic Systems*, 42(2) 295-306. p.
- BALDWIN, R., FREEMAN, R. (2022): Risks and global supply chains: What we know and what we need to know. In: *Annual Review of Economics*, 14(1) 153-180. p.
- BALL, C. A., TOROUS, W. N. (1984): The maximum likelihood estimation of security price volatility: Theory, evidence, and application to option pricing. In: *The Journal of Business*, 57(1) 97-112. p.
- BALTAGI, B. H. (2008): *Econometric Analysis of Panel Data*. Chichester: John Wiley & Sons. 135-145 p.
- BAO, W., YUE, J., RAO, Y. (2017): A deep learning framework for financial time series using stacked autoencoders and long-short term memory. In: *PLOS ONE* 12(7), e0180944. p.
- BARTLETT, M. S. (1946): On the theoretical specification of sampling properties of autocorrelated time series. In: *Journal of the Royal Statistical Society*, 8(1) 27-41. p.

- BAUR, D. G., SMALES, L. A. (2020): Hedging geopolitical risk with precious metals. In: *Journal of Banking & Finance*, 117 105823. p.
- BEAVER, W. H. (1968): The information content of annual earnings announcements. In: *Journal of Accounting Research*, 6 67-92. p.
- BECK, N., DOVERN, J., VOGL, S. (2025): Mind the naive forecast! a rigorous evaluation of forecasting models for time series with low predictability. In: *Applied Intelligence*, 55(6). p.
- BECKERS, S. (1981): Standard deviations implied in option prices as predictors of future stock price variability. In: *Journal of Banking & Finance*, 5(3) 363-381. p.
- BĘDOWSKA-SÓJKA, B., DEMIR, E., ZAREMBA, A. (2022): Hedging geopolitical risks with different asset classes: A focus on the Russian invasion of Ukraine. In: *Finance Research Letters*, 50 103192. p.
- BEJGER, S., FISZEDER, P. (2021): Dynamic spillovers between stock markets during the COVID-19 crisis: Evidence from a TVP-VAR connectedness approach. In: *Finance Research Letters*, 43 101957. p.
- BEKAERT, G., HOEROVA, M., DUCA, M. L. (2013): Risk, uncertainty and monetary policy. In: *Journal of Monetary Economics*, 60(7) 771-788. p.
- BERGMEIR, C., BENÍTEZ, J. M. (2012): On the use of cross-validation for time series predictor evaluation. In: *Information Sciences*, 191 192-213. p.
- BERGMEIR, C., HYNDMAN, R. J., KOO, B. (2018): A note on the validity of cross-validation for evaluating autoregressive time series prediction. In: *Computational Statistics & Data Analysis*, 120 70-83. p.
- BHAT, S. A., SHAH, A., SHAFIQ, M. (2022): A hybrid GARCH-LSTM model for financial time series forecasting. In: *Expert Systems with Applications*, 195 116645. p.
- BHARDWAJ, R., KAKADE, K., MISHRA, A. (2023): A hybrid ensemble learning GARCH-LSTM-based approach for Value-at-Risk forecasting. In: *Resources Policy*, 78 102903. p.
- BHOWMIK, R. AND WANG, S. (2020): Stock market volatility and return analysis: a systematic literature review. In: *Entropy*, 22(5) 522. p.
- BIALKOWSKI, J., GOTTSCHALK, K., & WISNIEWSKI, T. P. (2008): Stock market volatility around national elections. In: *Journal of Banking & Finance*, 32(9) 1941-1953. p.
- BILDIRICI, M., ERSIN, Ö. (2015): Forecasting volatility in oil prices with a class of nonlinear volatility models: Smooth transition RBF and MLP neural networks augmented GARCH approach. In: *Petroleum Science*, 12(3) 534–552. p.
- BISCHL, B., BINDER, M., LANG, ET AL. (2023): Hyperparameter optimization: foundations, algorithms, best practices, and open challenges. In: *Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery*, 13(2) e1484. p.
- BLACK, F. (1986): Noise. In: *The Journal of Finance*, 41(3) 528-543. p.

- BLACK, F., SCHOLLES, M. (1973): The pricing of options and corporate liabilities. In: *Journal of political economy*, 81(3) 637-654. p.
- BLACKWILL, R.D., HARRIS, J.M. (2016): War by other means: Geoeconomics and statecraft. Cambridge, MA: Harvard University Press. 49 p.
- BLASKOWITZ, O., HERWARTZ, H. (2011): On economic evaluation of directional forecasts. In: *International Journal of Forecasting*, 27(4) 1058-1065. p.
- BLOOM, N. (2009): The impact of uncertainty shocks. In: *Econometrica*, 77(3) 623-685. p.
- BOCCALETTI, S., DITTO, W. L., MINDLIN, G. B., ATANGANA, A. (2020): Modeling and forecasting of epidemic spreading: the case of covid-19 and beyond. In: *Chaos, Solitons and Fractals*, 135 109794. p.
- BOCQUILLON, P., DOYLE, S., S. JAMES, T., MASON, R., PARK, S., ROSINA, M. (2024): The effects of wars: lessons from the war in Ukraine. In: *Policy Studies*, 45(3-4) 261-281. p.
- BOLLERSLEV, T. (1986): Generalized Autoregressive conditional heteroscedasticity. In: *Journal of Econometrics*, 31 307-327. p.
- BOLLERSLEV, T. (1990): "Modeling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model". In: *The Review of Economics and Statistics*, 72(3) 498–505. p.
- BOLLERSLEV, T., WOOLDRIDGE, J. M. (1992): Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. In: *Econometric Reviews*, 11(2) 143-172. p.
- BOLLERSLEV, T., ENGLE, R. F., NELSON, D. B. (1994): ARCH models. In: *Handbook of Econometrics*, 4 2959-3038. p.
- BORDÁS, M. (2024): Hamas-israel war: a brief analysis of first two phases of war. In: *European Scientific Journal, ESJ*, 20(11), 1-14. p.
- BOUBAKER, S., GOODELL, J. W., PANDEY, D. K., KUMARI, V. (2022): Heterogeneous impacts of wars on global equity markets: evidence from the invasion of Ukraine. In: *Finance Research Letters*, 48 102934. p.
- BOUNGOU, W., YATIÉ, A. (2022): The impact of the Ukraine–Russia war on world stock market returns. In: *Economics letters*, 215 110516. p.
- BOURAS, C., CHRISTOU, C., GUPTA, R., SULEMAN, T. (2019): Geopolitical risks, returns, and volatility in emerging stock markets: Evidence from a panel GARCH model. In: *Emerging Markets Finance and Trade*, 55(8) 1841-1856. p.
- BOURI, E., JAIN, A., ROUBAUD, D., XU, Y. (2024): Geopolitical risk and food price inflation: A GVAR analysis of the Russia–Ukraine conflict. In: *Emerging Markets Finance and Trade*, 60(1) 1–17. p.
- BOX, G. E. P., JENKINS, G. M. (1970): Time series analysis: Forecasting and Control. San Francisco: Holden Day. 73 p.

- BOX, G. E., JENKINS, G. M., REINSEL, G. C., LJUNG, G. M. (2015): Time series analysis: forecasting and control (5th ed.). Hoboken, NJ: John Wiley & Sons. 75 p.
- BOZDOGAN, H. (1987): Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. In: *Psychometrika*, 52(3) 345-370. p.
- BRAILSFORD, T. J., FAFF, R. W. (1996): An evaluation of volatility forecasting techniques. In: *Journal of Banking & Finance*, 20(3) 419-438. p.
- BREMMER, I., ROUBINI, N. (2011): A G-zero world-the new economic club will produce conflict, not cooperation. In: *Foreign Affairs*, 90(2) 2-7. p.
- BROADSTOCK, D., CHAN, K., CHENG, L., WANG, X. (2021): The role of esg performance during times of financial crisis: evidence from covid-19 in China. In: *Finance Research Letters*, 38 101716. p.
- BROOKS, C., BURKE, S. P. (2003): Information criteria for GARCH model selection. In: *The European Journal of Finance*, 9(6) 557-580. p.
- BROWN, S. J., Warner, J. B. (1985): Using daily stock returns: The case of event studies. In: *Journal of Financial Economics*, 14(1) 3-31. p.
- BRÜTSCH, C., PAPA, M. (2013): Dealing with Geopolitical Shocks: Multilateralism as a Global Risk Management Strategy. In: *Global Policy*, 4(1) 56–66. p.
- BURNHAM, K. P., ANDERSON, D. R., HUYVAERT, K. P. (2011): AIC model selection and multi-model inference in behavioral ecology: some background, observations, and comparisons. In: *Behavioral Ecology and Sociobiology*, 65 23-35. p.
- BYUN, S. J., CHO, H. (2013): Forecasting carbon futures volatility using GARCH models with energy volatilities. In: *Energy Economics*, 40 207-221. p.
- CALDARA, D., CONLISK, S., IACOVIELLO, M., PENN, M. (2022): The effect of the war in Ukraine on global activity and inflation. In: *FEDS Notes*.
- CALDARA, D., IACOVIELLO, M. (2022): Measuring geopolitical risk. In: *American Economic Review*, 112(4) 1194-1225. p.
- CAMPBELL, J. Y., LO, A. W. (1996): The Econometrics of Financial Markets. Princeton, NJ: Princeton University Press. 155 p.
- CAMPBELL, J. Y., HENTSCHEL, L. (1992): No news is good news: An asymmetric model of changing volatility in stock returns. In: *Journal of financial Economics*, 31(3) 281-318. p.
- CAMPOS, R. G., HEID, B., TIMINI, J. (2024): The economic consequences of geopolitical fragmentation: Evidence from the Cold War. In: *CESifo Working Paper Series*, 11057. p.
- CANINA, L., FIGLEWSKI, S. (1993): The informational content of implied volatility. In: *The Review of Financial Studies* 6(3) 659-681. p.

- CAO, J., ROSLAN, T., ZHANG, W. (2018): Pricing variance swaps in a hybrid model of stochastic volatility and interest rate with regime-switching. In: *Methodology and Computing in Applied Probability*, 20(4) 1359-1379. p.
- CATANIA, L. AND PROIETTI, T. (2019): Forecasting volatility with time-varying leverage and volatility of volatility effects. In: *International Journal of Forecasting*, 36(2) 389–403. p.
- CBOE (2024): Historical Data for VIX-Index.
https://www.cboe.com/tradable_products/vix/vix_historical_data/. Search engine: Google. Key word: cboe vix historical data. Date of search:2024.08.09).
- CHANG, Y., CHOI, Y., PARK, J. Y. (2017): A new approach to model regime switching. In: *Journal of Econometrics*, 196(1) 127–143. p.
- CHANG, B. Y., CHRISTOFFERSEN, P., JACOBS, K. (2013): Market skewness risk and the cross section of stock returns. In: *Journal of Financial Economics*, 107(1) 46-68. p.
- CHANG, Y.-C. (2017). Regime-switching models and their application in asset allocation. In: *International Review of Economics & Finance*, 51 537–554. p.
- CHEN, N. F., ROLL, R., ROSS, S. A. (1986): Economic forces and the stock market. In: *Journal of business*, 59(3) 383-403. p.
- CHEN, C., XUE, L., XING, W. (2023): Research on improved GRU-based stock price prediction method. In: *Applied Sciences*, 13(15) 8813. p.
- CHEN, S., BOUTESKA, A., SHARIF, T., ABEDIN, M. Z. (2023): The Russia–Ukraine war and energy market volatility: A novel application of the volatility ratio in the context of natural gas. In: *Resources Policy*, 85 103792. p.
- CHENG, T., LIU, J., YAO, W., ZHAO, A. (2022): The impact of Covid-19 pandemic on the volatility connectedness network of global stock market. In: *Pacific-Basin Finance Journal*, 71 101678. p.
- CHEVILLON, G. (2007): Direct multi-step estimation and forecasting. In: *Journal of Economic Surveys*, 21(4) 746-785. p.
- CHOOBINEH, F., BRANTING, D. (1986): A simple approximation for semivariance. In: *European Journal of Operational Research*, 27(3) 364-370. p.
- CHONG, C. W., AHMAD, M. I., ABDULLAH, M. Y. (1999): Performance of GARCH models in forecasting stock market volatility. In: *Journal of Forecasting*, 18(5) 333-343. p.
- CHRISTIANO, A. D. O. (2014): Investment and exchange rate uncertainty under different regimes. In: *Economic Studies (São Paulo)*, 44(3) 553-577. p.
- CHRISTIE, A. A. (1982): The stochastic behavior of common stock variances: Value, leverage and interest rate effects. In: *Journal of Financial Economics*, 10(4) 407-432. p.
- CHRISTOFFERSEN, P. (2011): Elements of Financial Risk Management (2nd ed.). San Diego, CA: Academic Press. 67 p.

- CHRISTOFFERSEN, P., JACOBS, K., ORNTHANALAI, C. (2012): Dynamic jump intensities and risk premiums: Evidence from S&P500 returns and options. In: *Journal of Financial Economics*, 106(3) 447-472. p.
- CEYLAN, Ö. (2021): Dynamics of global stock market correlations: the VIX and attention allocation. In: *Journal of Applied Economics*, 24(1) 392-400. p.
- CONRAD, J., DITTMAR, R. F., GHYSELS, E. (2013): Ex ante skewness and expected stock returns. In: *The Journal of Finance*, 68(1) 85-124. p.
- CONT, R. (2001): Empirical properties of asset returns: Stylized facts and statistical issues. In: *Quantitative Finance*, 1(2) 223-236. p.
- COSTOLA, M., IACOPINI, M., SANTAGIUSTINA, C. (2021): Google search volumes and the financial markets during the covid-19 outbreak. In: *Finance Research Letters*, 42, 101884. p.
- CRESWELL, A., WHITE, T., DUMOULIN, V., ARULKUMARAN, K., SENGUPTA, B., BHARATH, A. A. (2018): Generative adversarial networks: An overview. In: *IEEE Signal Processing Magazine*, 35(1) 53-65. p.
- CUCINOTTA, D., VANELLI, M. (2020): WHO declares COVID-19 a pandemic. In: *Acta Bio Medica: Atenei Parmensis*, 91(1) 157. p.
- CUMBY, R. E., FIGLEWSKI, S., HASBROUCK, J. (1993): Forecasting volatilities and correlations with egarch models. In: *The Journal of Derivatives*, 1(2) 51-63. p.
- DACOROGNA, M. M., GENCAY, R., MÜLLER, U. A., OLSEN, R. B., PICTET, O. V. (2001): An Introduction to High-Frequency Finance. San Diego, CA: Academic Press. 161 p.
- DALBY, S. (2013): The geopolitics of climate change. In: *Political Geography*, 37 38-47. p.
- DAVYDOV, A. Y. (2022): US and Russian Financial Markets: Comparative Analysis. In: *Herald of the Russian Academy of Sciences*, 92(15) 1430-1434. p.
- DAY, T. E. AND LEWIS, C. (1992): Stock market volatility and the information content of stock index options. In: *Journal of Econometrics* 52(1-2), 267-287.
- DEGIANNAKIS, S., KAFOUSAKI, E. (2023): Forecasting VIX: the illusion of forecast evaluation criteria. In: *Economics and Business Letters*, 12(3) 231-240. p.
- DELONG, L., PELSSER, A. (2015): Instantaneous mean-variance hedging and Sharpe ratio pricing in a regime-switching financial model. In: *Stochastic Models*, 31(1) 67-97. p.
- DEMIR, C. (2019): Macroeconomic determinants of stock market fluctuations: the case of bist-100. In: *Economies*, 7(1) 8. p.
- DICKEY, D. A., FULLER, W. A. (1979): Distribution of the estimators for autoregressive time series with a unit root. In: *Journal of the American Statistical Association*, 74(366a) 427-431. p.
- DE GOOIJER, J. G., HYNDMAN, R. J. (2006): 25 years of time series forecasting. In: *International Journal of Forecasting*, 22(3) 443-473. p.

- DELONG, L., PELSSER, A. (2015): Instantaneous mean-variance hedging and instantaneous Sharpe ratio pricing in a regime-switching financial model. In: *Stochastic Models*, 31(1) 67-91. p.
- DÍAZ BERENGUER, A., DA, Y., BOSSA, M. N., OVENEKE, M. C., SAHLI, H. (2024): Causality-driven multivariate stock movement forecasting. In: *PLOS ONE*, 19(4) e0302197. p.
- DIEBOLD, F. X. (1998): *Elements of Forecasting*. Cincinnati, OH: South-Western College Publishing. 45 p.
- DIEBOLD, F. X., MARIANO, R. S. (1995): Comparing predictive accuracy. In: *Journal of Business & Economic Statistics*, 13(3) 253–263. p.
- DIEBOLD, F. X., RUDEBUSCH, G. D. (1996): Measuring business cycles: A modern perspective. In: *NBER Working Paper*, 4643.
- DIEBOLD, F. X., YILMAZ, K. (2009): Measuring Financial Asset Return and Volatility Spillovers, with Application to Global Equity Markets. In: *The Economic Journal*, 119(534) 158–171. p.
- DIEBOLD, F. X., YILMAZ, K. (2014): On the network topology of variance decompositions: Measuring the connectedness of financial firms. In: *Journal of Econometrics*, 182(1) 119-134. p.
- DIMA, B., DIMA, Ș., IOAN, R. (2021): Remarks on the behaviour of financial market efficiency during the covid-19 pandemic. The case of VIX. In: *Finance Research Letters*, 43, 101967. p.
- DING, Z., GRANGER, C. W. J., ENGLE, R. F. (1993): A long memory property of stock market returns and a new model. In: *Journal of Empirical Finance* 1(1) 83–106. p.
- DOBROTA, M., ZORNIĆ, N., MARKOVIĆ, A. (2021): FDI time series forecasts: evidence from emerging markets. In: *Journal of Sustainable Business and Management Solutions in Emerging Economies*, 26(2) 77-88. p.
- DOOB, J. L. (1942): What is a stochastic process?. In: *The American Mathematical Monthly*, 49(10) 648-653. p.
- DUBEY, S. R., SINGH, S. K., CHAUDHURI, B. B. (2022): Activation functions in deep learning: A comprehensive survey and benchmark. In: *Neurocomputing*, 503 92-108. p.
- ECB (2022): European Central Bank Economic Bulletin (4/2022). <https://www.ecb.europa.eu/press/economic-bulletin/html/eb202204.en.html>. Search engine: Google. Key word: ecb economic bulletin 04/2022. Date of search: 2024.11.10.
- EDWARDS, C. (2005): Derivative pricing models with regime switching: A general approach. In: *Journal of Derivatives*, 13(1) 41. p.
- ELLIOTT, G., TIMMERMAN, A. (2008): Economic Forecasting. In: *Journal of Economic Literature*, 46(1) 3-56. p.
- ELIMAM, H. (2017): Determinants of foreign direct investment in saudi arabia: a review. In: *International Journal of Economics and Finance*, 9(7) 222. p.

- ELMAN, J. L. (1990): Finding structure in time. In: *Cognitive science*, 14(2) 179-211. p.
- ELYASIANI, E., MANSUR, I., ODUSAMI, B. (2022): Sectoral stock return sensitivity to oil price changes: A double-threshold FIGARCH model. 359-388 p. In: GEMAN, H. (1st ed.): *Commodities: Markets, performance, and strategies*. Boca Raton, FL: Chapman and Hall/CRC.
- ENGLE, R. F. (1982): Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. In: *Econometrica* 50(4) 987-1007. p.
- ENGLE, R. F., BOLLERSLEV, T. (1986): Modelling the persistence of conditional variances. In: *Econometric Reviews*, 5(1) 1-50. p.
- ENGLE, R. F. (2001): GARCH 101: The use of ARCH/GARCH models in applied econometrics. In: *Journal of economic perspectives*, 15(4) 157-168. p.
- ENGLE, R. F. (2002): Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. In: *Journal of Business & Economic Statistics*, 20(3) 339-350. p.
- ENGLE, R. F., COLACITO, R. (2006): Testing and valuing dynamic correlations for asset allocation. In: *Journal of Business & Economic Statistics*, 24(2) 238-253. p.
- ENGLE, R. F., GRANGER, C. W. J. (1987): Co-Integration and Error Correction: Representation, Estimation, and Testing. In: *Econometrica*, 55(2) 251-276. p.
- ENGLE, R. F., SHEPPARD, K. (2001): Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. In: *NBER Working Paper*, 8554.
- FAMA, E. F. (1970): Efficient capital markets: A review of theory and empirical work. In: *Journal of Finance*, 25(2) 383-417. p.
- FAMA, E., FRENCH, K. (1993): Common risk factors in the returns on stocks and bonds. In: *Journal of Financial Economics*, 33(1) 3-56. p.
- FAMA, E. F., FRENCH, K. R. (2012). Size, value, and momentum in international stock returns. In: *Journal of Financial Economics*, 105(3) 457-472. p.
- FAMA, E. F., MACBETH, J. D. (1973): Risk, return, and equilibrium: empirical tests. In: *Journal of Political Economy*, 81(3) 607-636. p.
- FAMA, E. F., FISHER, L., JENSEN, M. C., ROLL, R. (1969): The adjustment of stock prices to new information. In: *International Economic Review*, 10(1) 1-21. p.
- FAN, C., ZHANG, X. (2024): Stock price nowcasting and forecasting with deep learning. In: *Journal of Intelligent Information Systems*, 1-18. p.
- FAROOQ, A. et al. (2024): Interpretable multi-horizon time series forecasting of cryptocurrencies by leverage temporal fusion transformer. In: *Heliyon* 10(22).

Federal Reserve Bank of St. Louis (FRED)(2024a): Historical data for S&P 500. <https://fred.stlouisfed.org/series/SP500>. Search engine: Google. Keywords: fred s&p500. Date of search: 2024.08.11.

Federal Reserve Bank of St. Louis (FRED)(2024b): Historical data for NASDAQ 100. <https://fred.stlouisfed.org/series/NASDAQ100>. Search engine: Google. Keywords: fred nasdaq. Date of search: 2024.08.11.

Federal Reserve Bank of St. Louis (FRED)(2024c): Historical data for Nikkei 225. <https://fred.stlouisfed.org/series/NIKKEI225>. Search engine: Google. Keywords: fred nikkei. Date of search: 2024.08.11.

Federal Reserve Bank of St. Louis (FRED)(2024d): Historical data for Gold. <https://fred.stlouisfed.org/tags/series?t=gold%3Bprice>. Search engine: Google. Keywords: fred gold price. Date of search: 2024.08.11.

Federal Reserve Bank of St. Louis (FRED)(2024e): Historical data for Brent Crude Oil. <https://fred.stlouisfed.org/series/POILBREUSDM>. Search engine: Google. Keywords: fred brent crude oil. Date of search: 2024.08.11.

Federal Reserve Bank of St. Louis (FRED)(2024f): Historical data for EUR-USD exchange rate. <https://fred.stlouisfed.org/series/CCUSMA02EZM618N>. Search engine: Google. Keywords: fred eur usd exchange rate. Date of search: 2024.08.11.

Federal Reserve Bank of St. Louis (FRED)(2024g): Historical data for 10-year U.S. Treasury Securities. <https://fred.stlouisfed.org/series/DGS10>. Search engine: Google. Keywords: fred 10-year us treasury rate. Date of search: 2024.08.11.

Federal Reserve Bank of St. Louis (FRED)(2024h): Historical data for Bitcoin. <https://fred.stlouisfed.org/series/CBBTCUSD>. Search engine: Google. Keywords: fred bitcoin. Date of search: 2024.08.11.

Federal Reserve Bank of St. Louis (FRED)(2024i): Historical data for CBOE Volatility Index (VIX). <https://fred.stlouisfed.org/series/VIXCLS>. Search engine: Google. Keywords: fred cboe vix index. Date of search: 2024.08.11.

FEDERLE, J., MEIER, A., MÜLLER, G. J., SEHN, V. (2022): Proximity to War: The stock market response to the Russian invasion of Ukraine. In: *Journal of Money, Credit and Banking*, 56(7) 1935–1970. p.

FIGLEWSKI, S. (1997): Forecasting volatility. In: *Financial Markets, Institutions & Instruments*, 6(1) 1-88. p.

FERNANDES, N. (2020): Economic effects of coronavirus outbreak (COVID-19) on the world economy. In: *Journal of Business Research*, 117 272–281. p.

FISCHER, T., KRAUSS, C., TREVINO, A. (2023): Deep Learning in Finance: A Survey of Applications and Techniques. In: *Journal of Financial Data Science*, 5(4) 10–30. p.

FLOOD, R. P., ROSE, A. K. (2005): Estimating the expected marginal rate of substitution: A systematic exploitation of idiosyncratic risk. In: *Journal of Monetary Economics*, 52(5) 951-969. p.

- FORBES, K. J., RIGOBON, R. (2002): No Contagion, Only Interdependence: Measuring Stock Market Comovements. In: *The Journal of Finance*, 57(5) 2223–2261. p.
- FRANCQ, C., ZAKOIAN, J. M. (2019): GARCH models: Structure, statistical inference and financial applications. Chichester, England: John Wiley & Sons. 175 p.
- FRANK, J. (2023): Forecasting realized volatility in turbulent times using temporal fusion transformers. In: *FAU Discussion Papers in Economics 03/2023*.
<https://hdl.handle.net/10419/268951>
- FRANSES, P. H., VAN DIJK, D. (1996): Forecasting stock market volatility using (non-linear) GARCH models. In: *Journal of Forecasting*, 15(3) 229-235. p.
- FRENCH, K. R., SCHWERT, G. W., STAMBAUGH, R. F. (1987): Expected stock returns and volatility. In: *Journal of Financial Economics*, 19(1) 3-29. p.
- GAO, X., REN, Y., UMAR, M. (2021): To what extent does covid-19 drive stock market volatility? a comparison between the U.S. and China. In: *Economic Research-Ekonomska Istraživanja*, 35(1) 1686-1706. p.
- GARMAN, M. B., KLASS, M. J. (1980): On the estimation of security price volatilities from historical data. In: *Journal of Business*, 53(1) 67-78. p.
- Geopolitical Risk Index (GPR) (2024): Historical data for the Geopolitical Risk Index.
<https://www.policyuncertainty.com/gpr.html>. Search engine: Google. Keywords: gpr geopolitical risk index. Date of search: 2024.08.11.
- GERS, F. A., SCHMIDHUBER, J., CUMMINS, F. (2000): Learning to forget: Continual prediction with LSTM. In: *Neural computation*, 12(10) 2451-2471. p.
- GEWEKE, J., SINGLETON, K. (1981): Maximum likelihood "confirmatory" factor analysis of economic time series. In: *International Economic Review*, 22(1) 37. p.
- GHARAIBEH, O., KHARABSHEH, B. (2023): Geopolitical Risks, Returns, and Volatility in the MENA Financial Markets: Evidence from GARCH and EGARCH Models. In: *Montenegrin Journal of Economics*, 19(3) 21-36. p.
- GIACOMINI, R., ROSSI, B. (2010): Forecast comparisons in unstable environments. In: *Journal of Applied Econometrics*, 25(4) 595-620. p.
- GLOSTEN, L. R., JAGANNATHAN, R., RUNKLE, D. E. (1993): On the relation between the expected value and the volatility of the nominal excess return on stocks. In: *The Journal of Finance*, 48(5) 1779-1801. p.
- GOKCAN, S. (2000): Forecasting volatility of emerging stock markets: linear versus non-linear GARCH models. In: *Journal of Forecasting*, 19(6) 499-504. p.
- GÓMEZ, R., MUSTILLO, T., VERCESI, M. (2024): Political developments and data in 2023. In: *European Journal of Political Research Political Data Yearbook*, 63(1) 3-13. p.
- GOODELL, J. (2020): Covid-19 and finance: agendas for future research. In: *Finance Research Letters*, 35, 101512. p.

- GOODFELLOW, I., BENGIO, Y., COURVILLE, A., BENGIO, Y. (2016): Deep learning. Cambridge, MA: MIT Press. 151-154 p.
- GOODFELLOW, I. et al. (2020): Generative adversarial networks. In: *Communications of the ACM*, 63(11) 139-144. p.
- GOODWIN, P., LAWTON, R. (1999): On the asymmetry of the symmetric MAPE. In: *International Journal of Forecasting*, 15(4) 405-408. p.
- GHOSH, A. R., QURESHI, M. S., KIM, J. I., ZALDUENDO, J. (2014): Surges. In: *Journal of International Economics*, 92(2) 266-285. p.
- GOSPODINOV, N., GAVALA, A., JIANG, D. (2006): Forecasting volatility. In: *Journal of Forecasting*, 25(6) 381-400. p.
- GRANGER, C. W., NEWBOLD, P. (1973): Some comments on the evaluation of economic forecasts. In: *Applied Economics*, 5(1) 35-47. p.
- GRANGER, C. W. J., NEWBOLD, P. (2014): Forecasting economic time series. San Diego, CA: Academic Press. 120 p.
- GRANGER, C. W., PESARAN, M. H. (2000): Economic and statistical measures of forecast accuracy. In: *Journal of Forecasting*, 19(7) 537-560. p.
- GRANGER, C. W., POON, S. H. (2001): Forecasting financial market volatility: A review. In: *Journal of Economic Literature*, 41(2) 478-593. p.
- GRAY, S. F. (1996): Modeling the conditional distribution of interest rates as a regime-switching process. In: *Journal of Financial Economics*, 42(1) 27-62. p.
- GRÄTZ, J. (2022): Energy sanctions and the limits of economic statecraft: EU-Russia energy relations after 2022. In: *Journal of International Affairs*, 75(2) 33-47. p.
- GROSSI, G. AND VAKULENKO, V. (2022): New development: accounting for human-made disasters—comparative analysis of the support to Ukraine in times of war. In: *Public Money & Management*, 42(6) 467-471. p.
- GU, R., YANG, Z., JI, Y. (2020): Machine learning for intelligent optical networks: a comprehensive survey. In: *Journal of Network and Computer Applications*, 157 102576. p.
- GUIDOLIN, M., TIMMERMAN, A. (2008): International asset allocation under regime switching, skew, and kurtosis preferences. In: *The Review of Financial Studies*, 21(2) 889-935. p.
- GUPTA, K., AHMED, S. (2020): Determinants of foreign portfolio flows to Indian debt market. In: *Journal of Indian Business Research*, 12(4) 459-479. p.
- GUPTA, R., KARMAKAR, S., PIERDZIOCH, C. (2023): Safe havens, machine learning, and the sources of geopolitical risk: a forecasting analysis using over a century of data. In: *Computational Economics*, 64(1) 487-513. p.

GUPTA, R., PIERDZIOCH, C. (2022): Forecasting the realized variance of oil-price returns: a disaggregated analysis of the role of uncertainty and geopolitical risk. In: *Environmental Science and Pollution Research*, 29(34) 52070-52082. p.

HA, J., LEE, S., SO, I. (2021): The impact of uncertainty shocks: evidence from geopolitical swings on the Korean peninsula. In: *Oxford Bulletin of Economics and Statistics*, 84(1) 21-56. p.

HAFNER, C. M., REZNIKOVA, O. (2012): On the estimation of dynamic conditional correlation models. In: *Computational Statistics & Data Analysis*, 56(11) 3533-3545. p.

HÁJEK, P., NOVOTNÝ, J. (2024): Beyond sentiment in stock price prediction: Integrating News Sentiment and Investor Attention with Temporal Fusion Transformer. 30-43 p. In: *IFIP International Conference on Artificial Intelligence Applications and Innovations*. Cham, Switzerland: Springer Nature

Hasanov, A. S., Burkhanov, A. U., Usmonov, B., Khajimuratov, N. S., & qizi Khurramova, M. M. (2024). The role of sudden variance shifts in predicting volatility in bioenergy crop markets under structural breaks. *Energy*, 293, 130535.

MAGLOGIANNIS, L. et al. (eds.): *Artificial Intelligence Applications and Innovations*. Cham, Switzerland: Springer Nature.

HALE, T., ANGRIST, N., GOLDSZMIDT, R., KIRA, B., PETHERICK, A., PHILLIPS, T., TATLOW, H. (2021): A global panel database of pandemic policies (oxford covid-19 government response tracker). In: *Nature Human Behaviour*, 5(4) 529-538. p.

HAMILTON, J. D. (1988): Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates. In: *Journal of Economic Dynamics and Control*, 12(2-3), 385-423. p.

HAMILTON, J. D. (2020): Time series analysis. Princeton, NJ: Princeton University Press. 72 p.

HAMILTON, J. D., SUSMEL, R. (1994): Autoregressive conditional heteroskedasticity and changes in regime. In: *Journal of econometrics*, 64(1-2) 307-333. p.

HANSEN, P. R., LUNDE, A. (2005): A forecast comparison of volatility models: does anything beat a GARCH (1, 1)?. In: *Journal of Applied Econometrics*, 20(7) 873-889. p.

HANSEN, P. R., LUNDE, A., NASON, J. M. (2011): The model confidence set. In: *Econometrica*, 79(2) 453-497. p.

HANSEN, P. R., LUNDE, A., VOEV, V. (2014): Realized beta GARCH: A multivariate GARCH model with realized measures of volatility. In: *Journal of Applied Econometrics*, 29(5) 774-799. p.

HARTANTO, S., GUNAWAN, A. (2024): Temporal fusion transformers for enhanced multivariate time series forecasting of indonesian stock prices. In: *International Journal of Advanced Computer Science and Applications*, 15(7).

HARVEY, D., LEYBOURNE, S., NEWBOLD, P. (1997). Testing the equality of prediction mean squared errors. In: *International Journal of Forecasting*, 13(2), 281–291. p.

- HASBROUCK, J. (2007): Empirical Market Microstructure: The Institutions, Economics, and Econometrics of Securities Trading. Ney York, NY: Oxford University Press. 4 p.
- HE, P., SUN, Y., ZHANG, Y., LI, T. (2020): Covid-19's impact on stock prices across different sectors—an event study based on the Chinese stock market. In: *Emerging Markets Finance and Trade*, 56(10) 2198-2212. p.
- HEYMANN, D. L., SHINDO, N. (2020): COVID-19: What is next for public health? In: *The Lancet*, 395(10224) 542–545. p.
- HOCHREITER, S., SCHMIDHUBER, J. (1997): Long Short-Term Memory. In: *Neural Computation*, 9(8) 1735-1780. p.
- HOCHREITER, S., BENGIO, Y., FRASCONI, P., SCHMIDHUBER, J. (2001): Gradient flow in recurrent nets: The difficulty of learning long term dependencies. 237-243. p. In: KREMER, S. C., KOHLEN, J. F. (Eds.): *A Field Guide to Dynamical Recurrent Neural Networks* (pp. 237–243). Piscataway, NJ: IEEE Press.
- HOFFMANN, M., NEUENKIRCH, M. (2017): The pro-Russian conflict and its impact on stock returns in Russia and the Ukraine. In: *International Economics and Economic Policy*, 14(1) 61-73. p.
- HOLT, C. C. (1957): Forecasting trends and seasonals by exponentially weighted moving averages. In: *O.N.R Memorandum* 52(52) 5-10. p.
- HOPFIELD, J. J. (1982): Neural networks and physical systems with emergent collective computational abilities. In: *Proceedings of the National Academy of Sciences*, 79(8) 2554-2558. p.
- HOQUE, M. E., YAKOB, N. A. (2017): Revisiting stock market development and economic growth nexus: The moderating role of foreign capital inflows and exchange rates. In: *Cogent Economics & Finance*, 5(1) 1329975. p.
- HOQUE, M., WAH, L., ZAIDI, M. (2019): Oil price shocks, global economic policy uncertainty, geopolitical risk, and stock price in malaysia: factor augmented var approach. In: *Economic Research-Ekonomska Istraživanja*, 32(1) 3700-3732. p.
- HUANG, C. et al. (2020): Clinical features of patients infected with 2019 novel coronavirus in wuhan, china. In: *The Lancet*, 395(10223) 497-506. p.
- HULL, J. C. (2022): Options, Futures, and Other Derivatives. Harlow, England: Pearson. 451 p.
- HYNDMAN, R. J., ATHANASOPOULOS, G. (2018): Forecasting: Principles and Practice. Melbourne, Australia: OTexts. 106-108. p.
- INAYATI, S., IRIAWAN, N., IRHAMAH, I. (2024): A markov switching autoregressive model with time-varying parameters. In: *Forecasting*, 6(3) 568-590. p.

International Monetary Fund (2022): World Economic Outlook: War Sets Back the Global Recovery (April 2022). International Monetary Fund.
<https://www.imf.org/en/Publications/WEO/Issues/2022/04/19/world-economic-outlook-april-2022>. Search engine: Google. Keywords: imf world economic outlook 04/2022. Date of search: 2024.05.11.

International Monetary Fund (2023): World Economic Outlook: Geoeconomic Fragmentation and Foreign Direct Investment (April 2023). International Monetary Fund.
<https://www.imf.org/en/Publications/WEO/Issues/2023/04/11/world-economic-outlook-april-2023>. Search engine: Google. Keywords: imf world economic outlook 04/2023. Date of search: 2024.05.11.

IZZELDIN, M. et al. (2021): The impact of Covid-19 on G7 stock markets volatility: Evidence from a ST-HAR model. In: *International Review of Financial Analysis*, 74 101671. p.

IZZELDIN, M. et al. (2023): The impact of the Russian Ukrainian war on global financial markets. In: *International Review of Financial Analysis*, 87 102598. p.

JAIN, N., MOTIANI, M., KAUR, P. (2021): Stock direction prediction using sentiment analysis of news articles. 195-206 p. In: GUPTA, H. et al. (Eds.): *Advanced computing and intelligent technologies*. Singapore: Springer.

JÄGGI, A., SCHLEGEL, M., ZANETTI, A. (2019): Macroeconomic surprises, market environment, and safe-haven currencies. In: *Swiss Journal of Economics and Statistics*, 155(1) 1-21. p.

Ji, W., CAO, Z., LI, X. (2023): Multi-Task Learning and Temporal-Fusion-Transformer-Based Forecasting of Building Power Consumption. In: *Electronics*, 12(22) 4656. p.

JIA, F. AND YANG, B. (2021): Forecasting volatility of stock index: deep learning model with likelihood-based loss function. In: *Complexity*, 2021(1) 5511802. p.

JIANG, C., LIU, X., ZHANG, T. (2024): Study of factors influencing U.S. treasury yields based on time series linear regression models. In: *Advances in Economics Management and Political Sciences*, 95(1) 19-26. p.

JONDEAU, E., ROCKINGER, M. (2003): Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements. In: *Journal of Economic Dynamics and Control*, 27(10) 1699-1737. p.

JORION, P. (1997): Value at risk: the new benchmark for managing financial risk. New York, NY: McGraw-Hill. 81 p.

JUBINSKI, D. AND LIPTON, A. (2011): Equity volatility, bond yields, and yield spreads. In: *Journal of Futures Markets*, 32(5) 480-503. p.

KANAMURA, T. (2022): Timing differences in the impact of Covid-19 on price volatility between assets. In: *Finance Research Letters*, 46(PB) 102401. p.

KANARACHOS, S., GEORGIU, G., BOUKAS, E. (2023): Deep learning for financial forecasting: A comprehensive survey. In: *Applied Soft Computing*, 138 110888. p.

KARPOFF, J. M. (1987): The relation between price changes and trading volume: A survey. In: *Journal of Financial and Quantitative Analysis*, 22(1) 109–126. p.

KICKBUSCH, I., LEUNG, G. (2020): Response to the COVID-19 pandemic: a call for global health governance reform. In: *The Lancet*, 395(10241) 1685–1688. p.

KIESEL, F., KOLARIC, S. (2023): Should I stay or should I go? Stock market reactions to companies' decisions in the wake of the Russia-Ukraine conflict. In: *Journal of International Financial Markets, Institutions and Money*, 89 101862. p.

KIM, H. Y., WON, C. H. (2018): Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. In: *Expert Systems with Applications*, 103 25-37. p.

KHAN, M. et al. (2023): Covid-19 pandemic & financial market volatility; evidence from garch models. In: *Journal of Risk and Financial Management*, 16(1) 50. p.

KOCAARSLAN, B., SOYTAS, U., SOYTAS, M. A. (2018): The asymmetric impact of oil prices, interest rates and oil price uncertainty on stock returns in Turkey: New evidence from a NARDL approach. In: *Energy Economics*, 71 343–352. p.

KOLM, P. N., TÜTÜNCÜ, R., FABOZZI, F. J. (2014): 60 years of portfolio optimization: Practical challenges and current trends. In: *European Journal of Operational Research*, 234(2) 356-371. p.

KONOVALOVA, M., ABUZOV, A. (2023): Geopolitical crises, the energy sector, and the financial capital market. In: *E3s Web of Conferences*, 381 01042. p.

KARRAS, T., LAINE, S., AILA, T. (2019): A style-based generator architecture for generative adversarial networks. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4401-4410. p.

KIM, J. (2014): Pricing of power options under the regime-switching model. In: *Journal of Applied Mathematics & Informatics*, 32(5-6) 665-673. p.

KIM, H. Y., WON, C. H. (2018): Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. In: *Expert Systems with Applications*, 103 25-37. p.

KULENDRAN, N., WITT, S. F. (2003): Forecasting the demand for international business tourism. In: *Journal of Travel Research*, 41(3) 265-271. p.

KUO, P. AND HUANG, C. (2018): A high precision artificial neural networks model for short-term energy load forecasting. In: *Energies*, 11(1) 213. p.

KUTNER, M. H. et al. (2005). *Applied Linear Statistical Models* (5th Ed.). Boston, MA: McGraw-Hill/Irwin. 183 p.

LABORDA, J. R. S., ZAMANILLO, I. (2023). Multi-country and multi-horizon GDP forecasting using Temporal Fusion Transformers. In: *Mathematics*, 11(12) 2625. p.

- LAI, K. S. (1990): An Evaluation of Survey Exchange Rate Forecasts. In: *Economics Letters*, 32 61–65. p.
- LAMA, A., JHA, G. K., PAUL, R. K., GURUNG, B. (2015): Modelling and forecasting of price volatility: an application of GARCH and EGARCH models. In: *Agricultural Economics Research Review*, 28(1) 73-82. p.
- LANDSMAN, W. R., MAYDEW, E. L. (2002): Has the information content of quarterly earnings announcements declined in the past three decades?. In: *Journal of Accounting Research*, 40(3) 797-808. p.
- LATIEF, R., LEFEN, L. (2018): The effect of exchange rate volatility on international trade and foreign direct investment (FDI) in developing countries along “one belt and one road”. In: *International Journal of Financial Studies*, 6(4), 86. p.
- LECUN, Y., BOSER, B., DENKER, J. S., HENDERSON, D., HOWARD, R. E., HUBBARD, W., JACKEL, L. D. (1989): Backpropagation applied to handwritten zip code recognition. In: *Neural computation*, 1(4) 541-551. p.
- LECUN, Y., BENGIO, Y. (1995): Convolutional networks for images, speech, and time series. 255-258 p. In: ARBIB, M. A. (Ed.) *The handbook of brain theory and neural networks*. Cambridge, MA: MIT Press.
- LECUN, Y., BENGIO, Y., HINTON, G. (2015): Deep learning. In: *Nature*, 521(7553) 436-444. p.
- LECUN, Y., BOTTOU, L., BENGIO, Y., HAFFNER, P. (1998): Gradient-based learning applied to document recognition. In: *Proceedings of the IEEE*, 86(11) 2278-2324. p.
- LEE, H. T. (2010): Regime switching correlation hedging. In: *Journal of Banking & Finance*, 34(11) 2728-2741. p.
- LEE, T. H., LONG, X. (2009): Copula-based multivariate GARCH model with uncorrelated dependent errors. In: *Journal of Econometrics*, 150(2) 207-218. p.
- LEVY, C., GIVATY, G., OVADIA, Y., SABAN, M. (2024): Treating wartime injuries amidst attack: insights from a medical facility on the edge of combat. In: *Conflict and Health*, 18(1) 1. p.
- LIAO, H., HUANG, J., TANG, Y. (2024): Leet: stock market forecast with long-term emotional change enhanced temporal model. In: *Peerj Computer Science*, 10 e1969. p.
- LIM, B., ARIK, S. Ö., LOEFF, N., PFISTER, T. (2021): Temporal fusion transformers for interpretable multi-horizon time series forecasting. In: *International Journal of Forecasting*, 37(4) 1748-1764. p.
- LIN, C., SUN, L. (2021): Financial volatility forecasting: A sparse multi-head attention neural network. In: *Information*, 12(10) 419. p.
- LIU, H., MANZOOR, A., WANG, C., ZHANG, L., MANZOOR, Z. (2020): The covid-19 outbreak and affected countries stock markets response. In: *International Journal of Environmental Research and Public Health*, 17(8) 2800. p.

- LIU, L., SHU, H., WEI, K. (2017): The impacts of political uncertainty on asset prices: evidence from the bo scandal in China. In: *Journal of Financial Economics*, 125(2) 286-310. p.
- LIU, R., LUX, T. (2017): Generalized Method of Moment estimation of multivariate multifractal models. In: *Economic Modelling*, 67 136-148. p.
- LONG, J. B., PLOSSER C. I. (1987): Sectoral vs. Aggregate Shocks in the Business Cycle. In: *American Economic Review, Papers and Proceedings*, 77(2) 333-36. p.
- LONGSTAFF, F. (2004): The flight-to-liquidity premium in U.S. treasury bond prices. *The Journal of Business*, 77(3) 511-526. p.
- LÜTKEPOHL, H. (2005): New Introduction to Multiple Time Series Analysis. Berlin, Heidelberg: Springer-Verlag. 1-3 p.
- LUNDBERGH, S., TERÄSVIRTA, T. (2002): Evaluating GARCH models. In: *Journal of Econometrics*, 110(2) 417-435. p.
- MA, F. et al. (2019): Volatility forecasting: Long memory, regime switching and heteroscedasticity. In: *Applied Economics*, 51(38) 4151–4163. p.
- MA, T., WANG, W., CHEN, Y. (2023): Attention is all you need: An interpretable transformer-based asset allocation approach. In: *International Review of Financial Analysis*, 90 102876. p.
- MACKINLAY, A. C. (1997): Event studies in economics and finance. In: *Journal of Economic Literature*, 35(1) 13-39. p.
- MAHJOUBI, M., HENCHIRI, J. E. (2024): The effect of policy uncertainty on the volatility of bitcoin. In: *Journal of Financial Economic Policy*, 16(4) 429-441. p.
- MAKRIDAKIS, S. (1993): Accuracy measures: theoretical and practical concerns. In: *International Journal of Forecasting*, 9(4) 527–529. p.
- MAKRIDAKIS, S., HIBON, M. (2000): The M3-Competition: results, conclusions and implications. In: *International journal of forecasting*, 16(4) 451-476. p.
- MANDELBROT, B. (1963): The Variation of Certain Speculative Prices. In: *Journal of Business*, 36(4) 394–419. p.
- MARCELLINO, M., STOCK, J. H., WATSON, M. W. (2006): A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. In: *Journal of Econometrics*, 135(1-2) 499-526. p.
- MARIOTTI, S. (2022): A warning from the russian–ukrainian war: avoiding a future that rhymes with the past. In: *Journal of Industrial and Business Economics*, 49(4) 761-782. p.
- MARSHALL, A., MAULANA, T., TANG, L. (2009): The estimation and determinants of emerging market country risk and the dynamic conditional correlation GARCH model. In: *International Review of Financial Analysis*, 18(5) 250-259. p.
- MARTINS, A. M., CORREIA, P., GOUVEIA, R. (2023): Russia-Ukraine conflict: The effect on European banks' stock market returns. In: *Journal of Multinational Financial Management*, 67 100786. p.

MATKOVSKYY, R., JALAN, A. (2019): From financial markets to bitcoin markets: a fresh look at the contagion effect. In: *Finance Research Letters*, 31 93-97. p.

MAZEN, F. M. A., SHAKER, Y., ABULA SEOUD, R. A. (2023): Forecasting of Solar Power Using GRU–Temporal Fusion Transformer Model and DILATE Loss Function. In: *Energies*, 16(24) 8105.

MBAH, R. E., WASUM, D. F. (2022): Russian-Ukraine 2022 War: A review of the economic impact of Russian-Ukraine crisis on the USA, UK, Canada, and Europe. In: *Advances in Social Sciences Research Journal*, 9(3) 144-153. p.

MCCULLOCH, W. S., PITTS, W. (1943): A logical calculus of the ideas immanent in nervous activity. In: *The Bulletin of Mathematical Biophysics*, 5(4) 115-133. p.

MCDONALD, J. B., NEWHEY, W. K. (1988): Partially adaptive estimation of regression models via the generalized t distribution. In: *Econometric Theory*, 4(3) 428-457. p.

MEAD, W. R. (2024): The Return of Hamiltonian Statecraft: A Grand Strategy for a Turbulent World. In: *Foreign Affairs*, 103(1) 52. p.

MERLO, L., PETRELLA, L., RAPONI, V. (2021): Forecasting VaR and ES using a joint quantile regression and its implications in portfolio allocation. In: *Journal of Banking & Finance*, 133 106248. p.

MERTON, R. C. (1976): Option pricing when underlying stock returns are discontinuous. In: *Journal of Financial Economics*, 3(1-2) 125-144. p.

MIKOLOV, T. et al. (2010, September): Recurrent neural network-based language model. In: *Interspeech*, 2(3) 1045-1048. p.

MINSKY, M., PAPERT, S. (1969): An introduction to computational geometry. Cambridge, MA: MIT Press. 16 p.

MITRAS, S., GOLITSIS, P., KHUDOYKULOV, K. (2022): Investigating the impact of geopolitical risks on the commodity futures. In: *Cogent Economics & Finance*, 10(1) 2049477. p.

MSCI (2024a): MSCI World Index. <https://www.msci.com/indexes/index/990100>. Search engine: Google. Keywords: msci world index. Date of search: 2024.07.04.

MSCI (2024b): MSCI World Sector Index. <https://www.msci.com/indexes/group/sector-indexes#featured-indexes>. Search engine: Google. Keywords: msci world sector index. Date of search: 2024.07.04.

MUZINDUTSI, P. et al. (2022): Contagion risk in equity markets during financial crises and covid-19: a comparison of developed and emerging markets. In: *Scientific Annals of Economics and Business*, 69(4) 615-629. p.

NAEEM, M. et al. (2022): Global factors and the transmission between united states and emerging stock markets. In: *International Journal of Finance & Economics*, 28(4) 3488-3510. p.

- NAIFAR, N., ALJARBA, S. (2023): Does geopolitical risk matter for sovereign credit risk? fresh evidence from nonlinear analysis. In: *Journal of Risk and Financial Management*, 16(3) 148.
- NAJAND, M. (2002): Forecasting stock index futures price volatility: Linear vs. nonlinear models. In: *nonlinear models. Financial Review*, 37(1) 93-104. p.
- NELSON, D. B. (1991): Conditional heteroskedasticity in asset returns: a new approach. In: *Econometrica*, 59(2) 347-370. p.
- NELSON, D. B., CAO, C. Q. (1992): Inequality constraints in the univariate GARCH model. In: *Journal of Business & Economic Statistics*, 10(2) 229-235. p.
- NEWAY, W. K., WEST, K. D. (1987): A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. In: *Econometrica*, 55(3) 703–708. p.
- NITTAYAKAMOLPHUN, P., BEJRANANDA, T., PHOLKERD, P. (2024): Asymmetric Effects of Uncertainty and Commodity Markets on Sustainable Stock in Seven Emerging Markets. In: *Journal of Risk and Financial Management*, 17(4) 155. p.
- NGUYEN, T., TRAN, M., KOHN, R. (2022): Recurrent conditional heteroskedasticity. In: *Journal of Applied Econometrics*, 37(5) 1031-1054. p.
- NIMANI, A., SPAHIJA, D. (2023): Financial markets and price increases in Europe after the Russian-Ukrainian War. In: *Scientific Horizons*, 26 (3) 135-145. p.
- NOVIANDY, T. et al. (2023): Deep learning-based bitcoin price forecasting using neural prophet. In: *Ekonomikalia Journal of Economics*, 1(1) 19-25. p.
- Organisation of Economic Co-operation and Development (OECD) (2022): Geopolitical Shocks and Global Interdependence. In: *Economic Outlook 2022(2)*. OECD Publishing.
<https://www.oecd.org/economic-outlook/>. Search engine: Google. Keywords: oecd economic outlook 2022 volume 2. Date of search: 2024.02.12
- OLIVEIRA, A. B., PEREIRA, P. L. V. (2018): Asset allocation with markovian regime switching: Efficient frontier and tangent portfolio with regime switching. In: *Brazilian Review of Econometrics*, 38(1) 97-127. p.
- OU, J. et al. (2022). Traffic volatility forecasting using an omnibus family GARCH modeling framework. In: *Entropy*, 24(10) 1392. p.
- PACHÉ, G. (2024): Israeli-palestinian conflict: towards a major logistical and environmental crisis?. In: *Technium Social Sciences Journal*, 53(1) 252-258. p.
- PARKINSON, M. (1980): The extreme value method for estimating the variance of the rate of return. In: *Journal of Business*, 53(1) 61-65. p.
- PAGAN, A., SCHWERT, G. W. (1990): Alternative models for conditional stock volatility. In: *Journal of Econometrics*, 45(1-2) 267-290. p.
- PAN, Z., LIU, L. (2018): Forecasting stock return volatility: A comparison between the roles of short-term and long-term leverage effects. In: *Physica A: Statistical Mechanics and its Applications*, 492 168-180.

- PANDEY, D. K., ASSAF, R., RAI, V. K. (2023): Did the Indian stock market sail the Russia-Ukraine storm safely?. In: *The Journal of Economic Asymmetries*, 28 e00319. p.
- PASTOR, L., VERONESI, P. (2012): Uncertainty about Government Policy and Stock Prices. In: *The Journal of Finance*, 67(4) 1219–1264. p.
- PAVLIOTIS, G. A. (2014): Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations. In: *Texts in Applied Mathematics*, 60.
- PEREIRA, F., FERREIRA, P., MARTINS, L. (2024): Evaluating volatility forecasts under extreme events: A comparative study of traditional and machine learning models. In: *Journal of Financial Econometrics*, 22(2) 345–367. p.
- PESARAN, M. H. (2015): Testing weak cross-sectional dependence in large panels. In: *Econometric reviews*, 34(6-10) 1089-1117. p.
- PESARAN, B., PESARAN, M. H. (2007): Modelling volatilities and conditional correlations in futures markets with a multivariate t distribution. In: *Journal of Empirical Finance*, 14(5) 622–640. p.
- PESARAN, M. H., PICK, A., TIMMERMAN, A. (2011): Variable selection, estimation and inference for multi-period forecasting problems. In: *Journal of Econometrics*, 164(1) 173-187. p.
- PESARAN, M. H., SMITH, R. (1995): Estimating long-run relationships from dynamic heterogeneous panels. In: *Journal of econometrics*, 68(1) 79-113. p.
- PESARAN, M. H., TIMMERMAN, A. (2007): Selection of estimation window in the presence of breaks. In: *Journal of Econometrics*, 137(1) 134-161. p.
- PETNEHÁZI, G. (2021): Quantile convolutional neural networks for value at risk forecasting. In: *Machine Learning with Applications*, 6 100096. p.
- PHAN, D. AND NARAYAN, P. (2020): Country responses and the reaction of the stock market to covid-19 - a preliminary exposition. In: *Emerging Markets Finance and Trade*, 56(10) 2138-2150. p.
- PHILLIPS, P. C. B. (1986): Understanding Spurious Regressions in Econometrics. In: *Journal of Econometrics*, 33(3) 311-340. p.
- PHILLIPS, P. C., PERRON, P. (1988): Testing for a unit root in time series regression. In: *Biometrika*, 75(2) 335-346. p.
- POON, S. H., GRANGER, C. W. J. (2003): Forecasting volatility in financial markets: A review. In: *Journal of economic literature*, 41(2) 478-539. p.
- PRASAD, M., BAKRY, W., VARUA, M. E. (2021): Abnormal Volatility in seasoned equity offerings during economic disruptions. In: *Journal of Behavioral and Experimental Finance*, 30 100509. p.
- RAMASAMY, R., MUNISAMY, S. (2012). Predictive accuracy of GARCH, GJR and EGARCH models select exchange rates application. In: *Global Journal of Management and Business Research*, 12(15) 89-100. p.

RAMOS-PÉREZ, E., GONZÁLEZ, P., NÚÑEZ-VELÁZQUEZ, J. (2021): Multi-transformer: a new neural network-based architecture for forecasting & volatility. In: *Mathematics*, 9(15) 1794. p.

RANIA, F. (2023): The dynamics of safe-haven assets and european stock markets before and during the russia-ukraine war. In: *European Journal of Business Management and Research*, 8(4) 46-50. p.

RASHEED, H., AHMAD, H., JAVID, A. (2021): Is gold a hedge and safe haven during political uncertainties?. In: *Business & Economic Review*, 13(2) 1-28. p.

REHMAN, M. U. (2016): Financial contagion and the global financial crisis: Evidence from Asia. In: *International Review of Economics & Finance*, 45 411–425. p.

Reuters (2024): Historical data of Hang Seng Stock Index.
<https://www.reuters.com/markets/quote/.HSI/>. Search engine: Google. Keywords: reuters historical data hangseng. Date of search: 2024.08.11.

RIGOBON, R., SACK, B. (2005): The effects of war risk on US financial markets. In: *Journal of Banking & Finance*, 29(7) 1769-1789. p.

ROBUS, S., WALTER, V., KÖMÜVES, Z. (2024): What is the impact of the Russian war in the Ukraine on global stock market sectors in 2022?. In: *Multidiszciplináris kihívások, sokszínű válaszok-Gazdálkodás-és Szervezéstudományi folyóirat*, (1) 51-81.p.

ROBUS, S., WALTER, V., KÖMÜVES, Z. (in press): Did the Russian war against Ukraine increase global stock market sector volatility in 2022?. In: *Közgazdász Fórum / Forum on Economics and Business*.

ROSENBLATT, F. (1958): The perceptron: a probabilistic model for information storage and organization in the brain. In: *Psychological Review*, 65(6) 386. p.

RUMELHART, D., HINTON, G., WILLIAMS, R. (1986): Learning representations by back-propagating errors. In: *Nature*, 323 533–536. p.

QU, Z., PERRON, P. (2013): A stochastic volatility model with random level shifts and its applications to S&P 500 and NASDAQ return indices. In: *The Econometrics Journal*, 16(3) 309-339. p.

SAID, S. E., DICKEY, D. A. (1984): Testing for unit roots in autoregressive-moving average models of unknown order. In: *Biometrika*, 71(3) 599-607. p.

SADORSKY, P., MCKENZIE, M. D. (2008): Power transformation models and volatility forecasting. In: *Journal of Forecasting*, 27(7) 587-606. p.

SAAR-ASHKENAZY, R., BERGMAN, Y., ASHKENAZY, O., GUEZ, J. (2024): Traumatic stress, active engagement and resilience in first responders and civilians in the outbreak of war. In: *European Journal of Psychotraumatology*, 15(1).

SAKWA, R. (2022): The Ukraine crisis and the return of geopolitics. In: *International Affairs*, 98(3) 721–739. p.

- SAMEK, W., MONTAVON, G., LAPUSCHKIN, S., ANDERS, CJ, MÜLLER, KR (2021): Explaining deep neural networks and beyond: A review of methods and applications. In: *Proceedings of the IEEE*, 109(3) 247-278. p.
- SANKARAN, H., NGUYEN, A., HARIKUMAR, J. (2012): Extreme return correlation and volatility: a two-threshold approach. In: *American Journal of Business*, 27(2) 154-173. p.
- SARWAR, G. AND KHAN, W. (2018): Interrelations of U.S. market fears and emerging markets returns: global evidence. In: *International Journal of Finance & Economics*, 24(1) 527-539. p.
- SASSETTI, P., TANI, M. (2006): Dynamic asset allocation using systematic sector rotation. In: *The Journal of Wealth Management*, 8(4) 59-70. p.
- SAULNIER, G., CASTRO, J., COOK, C. (2020): Forecasting inpatient glycemic control: extension of damped trend methods to subpopulations. In: *Future Science OA*, 6(10). p.
- SAVICKAS, R. (2003): Event-induced volatility and tests for abnormal performance. In: *Journal of Financial Research*, 26(2) 165-178. p.
- SCHMIDHUBER, J. (2015): Deep learning in neural networks: An overview. In: *Neural networks*, 61 85-117. p.
- SCHWELLER, R. (2018): Opposite but compatible nationalisms: a neoclassical realist approach to the future of US–China relations. In: *The Chinese Journal of International Politics*, 11(1) 23-48. p.
- SCHWERT, G. W. (1989): Why does stock market volatility change over time?. In: *The Journal of Finance*, 44(5) 1115-1153. p.
- SERLETIS, A., XU, L. (2016): Modeling volatility and contagion in international equity markets. In: *Journal of Macroeconomics*, 48 276–291. p.
- SEZER, O. B., GUDELEK, M. U., OZBAYOGLU, A. M. (2020): Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. In: *Applied soft computing*, 90 106181.
- SHAH, A., ISAH, H., ZULKERNINE, F. (2019): Stock market analysis: A review and taxonomy of prediction techniques. In: *International Journal of Financial Studies*, 7(2) 26. p.
- SHEN, J., QIN, P., ZHONG, R., HAN, P. (2025): Optimization of temporal feature attribution and sequential dependency modeling for high-precision multi-step resource forecasting: a methodological framework and empirical evaluation. In: *Mathematics*, 13(8) 1339. p.
- SIAW, R., OFOSU-HENE, E., TEE, E. (2017): Investment portfolio optimization with GARCH models. In: *ELK Asia Pacific Journal of Finance and Risk Management*, 8(2).
- SIMARD, P. Y., STEINKRAUS, D., PLATT, J. C. (2003, August): Best practices for convolutional neural networks applied to visual document analysis. In: *Proceedings of the Seventh International Conference on Document Analysis and Recognition*, 3 958–963. p.
- SIMKINS, S. (1995): Forecasting with vector autoregressive (var) models subject to business cycle restrictions. In: *International Journal of Forecasting*, 11(4) 569-583. p.

- SIMS, C. A. (1980): Macroeconomics and reality. In: *Econometrica: Journal of the Econometric Society*, 48(1) 1-48.
- SHEN, Y., WEI, J., ZHAO, Q. (2020): Mean–variance asset–liability management problem under non-Markovian regime-switching models. In: *Applied Mathematics & Optimization*, 81 859-897. p.
- SMALES, L. A. (2021): Geopolitical risk and volatility spillovers in oil and stock markets. In: *The Quarterly Review of Economics and Finance* 80 358-366. p.
- SONG, Y., TANG, X., WANG, H., MA, Z. (2022): Volatility forecasting for stock market incorporating macroeconomic variables based on garch-midas and deep learning models. In: *Journal of Forecasting*, 42(1) 51-59. p.
- SORAYA, N., MUHAMMAD, A., LADIQI, S. (2024): Icc jurisdiction: against israeli war and humanitarian crimes targeting palestinian civilians 2023. In: *Journal Media Hukum*, 31(1) 59-77. p.
- STOCK, J. H. (1989): New Indexes of Coincident and Leading Economic Indicators. In: *NBER Macroeconomics Annual*, 4 351–394. p. MIT Press.
- STOCK, J. H., WATSON, M. W. (1996): Evidence on Structural Instability in Macroeconomic Time Series Relations. In: *Journal of Business & Economic Statistics* 14(1) 11–30. p.
- STOCK, J. H., WATSON, M. W. (2007): Why has US inflation become harder to forecast?. In: *Journal of Money, Credit and Banking*, 39 3-33. p.
- STOCK, J. H., WATSON, M. W. (2012): Generalized shrinkage methods for forecasting using many predictors. In: *Journal of Business & Economic Statistics*, 30(4) 481-493. p.
- SU, L., ZUO, X., LI, R., WANG, X., ZHAO, H., HUANG, B. (2025): A systematic review for transformer-based long-term series forecasting. In: *Artificial Intelligence Review*, 58(3) 80. p.
- STORTI, G., WANG, C. (2022): Nonparametric expected shortfall forecasting incorporating weighted quantiles. In: *International Journal of Forecasting*, 38(1) 224-239. p.
- TANG, Z., HUANG, J., RINPRASERTMEECHAI, D. (2024): Period-aggregated transformer for learning latent seasonalities in long-horizon financial time series. In: *PLOS ONE*, 19(8) e0308488. p.
- TANYERI, B., SAVAŞER, T., USUL, N. (2021): The stock and cds market consequences of political uncertainty: the arab spring. In: *Emerging Markets Finance and Trade*, 58(7) 1821-1837. p.
- TASHMAN, L. J. (2000): Out-of-sample tests of forecasting accuracy: an analysis and review. In: *International Journal of Forecasting*, 16(4) 437-450. p.
- TAYLOR, J. W. (2004): Volatility forecasting with smooth transition exponential smoothing. In: *International Journal of Forecasting*, 20(2) 273-286. p.

- TAYLOR, J. W. (2019): Forecasting value at risk and expected shortfall using a semiparametric approach based on the asymmetric Laplace distribution. In: *Journal of Business & Economic Statistics*, 37(1) 121-133. p.
- TAYLOR, S. J. (1987): Forecasting the volatility of currency exchange rates. In: *International Journal of Forecasting*, 3(1) 159-170. p.
- TAYLOR, S. J. (2007): *Modelling Financial Time Series*. Singapore: World Scientific Publishing. 35p.
- THAO, P. T. P., DALY, K., ELLIS, C. (2013): Dynamic linkages among ASEAN stock markets: Evidence from rolling estimations and GARCH models. In: *International Review of Financial Analysis*, 29 1–10. p.
- THEIRI, S., NEKHILI, R., SULTAN, J. (2022): Cryptocurrency liquidity during the Russia–Ukraine war: The case of Bitcoin and Ethereum. In: *The Journal of Risk Finance*, 24(1) 59–71. p.
- THENMOZHI, M. AND RADHA, S. (2008): Dynamic interaction among stock returns, short-term interest rates, spread and expected inflation. In: *International Journal of Indian Culture and Business Management*, 1(3) 296. p.
- TIMMERMAN, A. (2006): Forecast combinations. In: *Handbook of Economic Forecasting*, 1 135-196. p.
- TOBIAS, A., BRUNNERMEIER, M. K. (2016): CoVaR. In: *The American Economic Review*, 106(7) 1705. p.
- TSAY, R. S. (2010): *Analysis of Financial Time Series*. Hoboken, NJ: John Wiley & Sons. 110 p.
- TSAY, R. S. (2013): *Multivariate Time Series Analysis: with R and Financial Applications*. Hoboken, NJ: John Wiley & Sons. 18 p.
- TU, J. (2010): Is regime switching in stock returns important in portfolio decisions? In: *Management Science*, 56(7) 1198–1215. p.
- TUNA, A. (2022): The effects of volatilities in oil price, gold price and Vix index on Turkish BIST 100 stock index in pandemic period. In: *İstanbul İktisat Dergisi-Istanbul Journal of Economics*, 72(1).
- UGLI, D. K. K. M. (2024): Palestine-Israel Conflict: The Influence of Arab Countries on the Oil Policy and Geopolitics of the Middle East. In: *London Journal of Research in Humanities and Social Sciences*, 24(8) 83-89. p.
- ULLAH, A., BIAO, H., ULLAH, A. (2024): Unveiling the Nexus Between Crises, Investor Sentiment, and Volatility of Tourism-Related Stocks: Empirical Findings From Pakistan. In: *SAGE Open*, 14(3) 21582440241256236. p.
- UMAR, Z., POLAT, O., CHOI, S. Y., TEPLOVA, T. (2022): The impact of the Russia-Ukraine conflict on the connectedness of financial markets. In: *Finance Research Letters*, 48, 102976. p

VAN BERGEIJK, P. A. (2022): Sanctions against the Russian war on Ukraine: Lessons from history and current prospects. In: *Journal of World Trade*, 56(4).

VARTANIAN, P. R., NETO, R. S. (2023): The VIX Index and the volatility of the Latin American and G7 stock exchanges before and during the COVID 19 pandemic. In: *International Journal of Economics and Finance*, 15(2) 25-38. p.

VASWANI, A. et al. (2017): Attention is all you need. In: *Advances in Neural Information Processing Systems*, 30.

VELÁSQUEZ HENAO, J. D., RUEDA MEJIA, V., FRANCO CARDONA, C. J. (2013): Electricity demand forecasting using a SARIMA-multiplicative single neuron hybrid model. In: *Dyna*, 80(180) 4-8. p.

VISKOVIĆ, J., ARNERIĆ, J., ROZGA, A. (2014): Volatility Switching between Two Regimes. In: *International Journal of Social, Management, Economics and Business Engineering*, 8(3) 682–686. p.

VOLOSOVYCH, S. et al. 2024): The Cryptocurrency Asset Markets in the Conditions of Military Aggression. In: *Financial & Credit Activity: Problems of Theory & Practice*, 1(54).

WAGENMAKERS, E. J., FARRELL, S. (2004): AIC model selection using Akaike weights. In: *Psychonomic Bulletin & Review*, 11 192-196. p.

WANG, G.-J., XIE, C., LIN, M., STANLEY, H. E. (2017): Stock market contagion during the global financial crisis: A multiscale approach. In: *Finance Research Letters*, 22 163–168. p.

WANG, Y., BOURI, E., FAREED, Z., DAI, Y. (2022): Geopolitical risk and the systemic risk in the commodity markets under the war in Ukraine. In: *Finance Research Letters*, 49 103066. p.

WANG, L., MA, F., LIU, J., YANG, L. (2020): Forecasting stock price volatility: New evidence from the GARCH-MIDAS model. In: *International Journal of Forecasting*, 36(2) 684-694. p.

WANG, Y., ZHOU, H. (2024): Geopolitical shocks and commodity market dynamics: New evidence from the Russia–Ukraine conflict. In: *Economic Modelling*, 124, 106334. p.

WELCH, I., GOYAL, A. (2008): A comprehensive look at the empirical performance of equity premium prediction. In: *The Review of Financial Studies*, 21(4) 1455-1508. p.

World Health Organisation (WHO) (2020): WHO Timeline – COVID-19, April 2020. World Health Organisation. <https://www.who.int/news/item/27-04-2020-who-timeline---covid-19>. Search engine: Google. Key words: who covid-19 timeline. Date of search: 2024.03.05.

WINTERS, P. R. (1960): Forecasting sales by exponentially weighted moving averages. In: *Management Science*, 6(3) 324-342. p.

World Bank (2022): Global economic prospects, June 2022. The World Bank. <https://www.worldbank.org/en/publication/global-economic-prospects>. Search engine: Google. Key words: world bank global economic prospects june 2022. Date of search: 2024.07.04.

XU, G. AND GAO, W. (2019): Financial risk contagion in stock markets: causality and measurement aspects. In: *Sustainability*, 11(5) 1402. p.

- WU, H., XU, J., WANG, J., LONG, M. (2021): Autoformer: Decomposition transformers with autocorrelation for long-term series forecasting. In: *Advances in Neural Information Processing Systems*, 34 22419-22430. p
- YANG, J., LI, P., CUI, Y., HAN, X., ZHOU, M. (2025): Multi-Sensor Temporal Fusion Transformer for Stock Performance Prediction: An Adaptive Sharpe Ratio Approach. In: *Sensors*, 25(3) 976. p.
- YANG, A., YE, Q., ZHAI, J. (2024): Volatility forecasting with Hybrid-long short-term memory models: Evidence from the COVID-19 period. In: *International Journal of Finance & Economics*, 29(3) 2766-2786. p.
- YAO, C.-Z., LI, H.-Y. (2020): Effective transfer entropy approach to information flow among EPU, investor sentiment and stock market. In: *Frontiers in Physics*, 8 206. p.
- YAZIZ, S., ZAKARIA, R., BOLAND, J. (2020): Multistep forecasting for highly volatile data using a new box-jenkins and garch procedure. In: *ASM Science Journal*, 13(1) 1-6. p.
- YILDIRIM, D., & CELIK, A. K. (2020). Stock market volatility and structural breaks: An empirical analysis of fragile five countries using GARCH and EGARCH models. In: *Journal of Applied Economics and Business Research*, 10(3) 148-163. p.
- YOUSAF, I., ALI, S. (2023): Dynamic connectedness and portfolio implications among gold, oil, US dollar and cryptocurrencies during the COVID-19 crisis. In: *Resources Policy*, 82 103473. p.
- YOUSAF, I., PATEL, R., YAROVAYA, L., 2022. The reaction of G20+ stock markets to the Russia–Ukraine conflict “black-swan” event: Evidence from event study approach. In: *Journal of Behavioral and Experimental Finance* 35, 100723. p.
- YOSINSKI, J., CLUNE, J., BENGIO, Y., LIPSON, H. (2014): How transferable are features in deep neural networks?. In: *Advances in Neural Information Processing Systems*, 27.
- YU, X. AND LI, D. (2021): Important trading point prediction using a hybrid convolutional recurrent neural network. In: *Applied Sciences*, 11(9) 3984. p.
- YU, Y., SI, X., HU, C., ZHANG, J. (2019): A review of recurrent neural networks: LSTM cells and network architectures. In: *Neural Computation*, 31(7) 1235-1270. p.
- ZHA, R., HE, K., YU, L., XI, X., Su, Y. (2024). Risk estimation of crude oil future price using temporal fusion transformer model. In: *Procedia Computer Science*, 242, 313-317. p.
- ZHAO, F. (2023): Bank Lending in Dual Narratives: Economic Policy Uncertainty and Interest-Rate Derivatives. In: *The Journal of Applied Business and Economics*, 25(6) 182-192. p.
- ZHANG, D., HU, M., JI, Q. (2020): Financial markets under the global pandemic of COVID-19. In: *Finance research letters*, 36 101528. p.
- ZHANG, Q., LI, X., GAO, P. (2025). Forecasting sales in live-streaming cross-border e-commerce in the uk using the temporal fusion transformer model. In: *Journal of Theoretical and Applied Electronic Commerce Research*, 20(2), 92. p.

ZHAO, B., WANG, Y. S., KOLAR, M. (2019): Direct estimation of differential functional graphical models. In: *Advances in Neural Information Processing Systems*, 32.

ZHAO, R., LEI, Z., ZHAO, Z. (2024): Research on the application of deep learning techniques in stock market prediction and investment decision-making in financial management. In: *Frontiers in Energy Research*, 12, 1342130. p.

ZHAO, L., YAN, W. Q. (2024): Prediction of Currency Exchange Rate Based on Transformers. In: *Journal of Risk and Financial Management*, 17(8) 332. p.

ZHOU, H., ZHANG, S., PENG, J., ZHANG, S., LI, J., XIONG, H., & ZHANG, W. (2021): Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(12) 11106–11115. p.

ZOLFAGHARI, M., HOSEINZADE, S. (2020): Impact of exchange rate on uncertainty in stock market: Evidence from Markov regime-switching GARCH family models. In: *Cogent Economics & Finance*, 8(1) 1802806. p.

APPENDIX 2: FURTHER MATERIAL

Financial Instruments	DC COVID-19	DC Complete	t-statistic	p-value
S&P500 vs NASDAQ	0.9501	0.8673	14.7885	0.0000
S&P500 vs VIX	0.6563	0.3755	17.7835	0.0000
NASDAQ vs VIX	0.6159	0.3260	17.7560	0.0000
S&P500 vs 10yr US T-Bond	0.5476	0.1988	27.0965	0.0000
10yr US T-Bond vs VIX	0.5460	0.1889	24.3247	0.0000
Nikkei 225 vs HangSeng	0.5189	0.1834	16.2950	0.0000
NASDAQ vs 10yr US T-Bond	0.5041	0.1727	23.7250	0.0000
Nikkei 225 vs VIX	0.3913	0.1779	12.2638	0.0000
S&P500 vs Nikkei 225	0.3870	0.1087	11.2058	0.0000
HangSeng vs VIX	0.3709	0.1362	7.0391	0.0000
NASDAQ vs Nikkei 225	0.3586	0.0816	11.8138	0.0000
EUR-USD vs VIX	0.3298	0.0612	13.6055	0.0000
S&P500 vs Gold	0.3295	0.1048	11.2292	0.0000
NASDAQ vs Brent Crude Oil	0.3260	0.0981	9.8653	0.0000
Brent Crude Oil vs 10yr US T-Bond	0.3151	0.1266	6.5735	0.0000
Gold vs VIX	0.2998	0.0733	9.1401	0.0000
S&P500 vs Brent Crude Oil	0.2941	0.1300	6.8764	0.0000
HangSeng vs EUR-USD	0.2935	0.1150	6.2421	0.0000
Brent Crude Oil vs VIX	0.2914	0.1377	6.8291	0.0000
HangSeng vs Gold	0.2749	0.0470	11.7219	0.0000
NASDAQ vs Gold	0.2694	0.1034	8.0475	0.0000
10yr US T-Bond vs EUR-USD	0.2645	0.0913	5.1291	0.0000
S&P500 vs HangSeng	0.2554	0.0774	5.6310	0.0000
HangSeng vs Brent Crude Oil	0.2525	0.0748	5.5603	0.0000
S&P500 vs Bitcoin	0.2397	0.1115	3.6639	0.0005
Nikkei 225 vs EUR-USD	0.2319	0.1057	4.1009	0.0001
Gold vs 10yr US T-Bond	0.2247	0.2174	0.3762	0.7079
NASDAQ vs Bitcoin	0.2212	0.1301	2.1733	0.0334
NASDAQ vs HangSeng	0.2174	0.0671	5.6433	0.0000
NASDAQ vs EUR-USD	0.1999	0.0433	9.8891	0.0000
HangSeng vs Bitcoin	0.1979	0.0060	7.8909	0.0000
Nikkei 225 vs Bitcoin	0.1841	0.0475	5.0986	0.0000
S&P500 vs EUR-USD	0.1840	0.0202	10.0960	0.0000
Gold vs EUR-USD	0.1822	0.1071	2.6731	0.0095
Bitcoin vs VIX	0.1740	0.0841	3.0135	0.0037
Gold vs GPR	0.1705	0.0620	5.1867	0.0000
Gold vs Brent Crude Oil	0.1626	0.1080	1.2964	0.1995
Nikkei 225 vs 10yr US T-Bond	0.1429	0.0797	1.9786	0.0521
S&P500 vs HangSeng	0.2554	0.0774	5.6310	0.0000
HangSeng vs Brent Crude Oil	0.2525	0.0748	5.5603	0.0000
S&P500 vs Bitcoin	0.2397	0.1115	3.6639	0.0005
Nikkei 225 vs EUR-USD	0.2319	0.1057	4.1009	0.0001
Gold vs 10yr US T-Bond	0.2247	0.2174	0.3762	0.7079
NASDAQ vs Bitcoin	0.2212	0.1301	2.1733	0.0334
NASDAQ vs HangSeng	0.2174	0.0671	5.6433	0.0000
NASDAQ vs EUR-USD	0.1999	0.0433	9.8891	0.0000
HangSeng vs Bitcoin	0.1979	0.0060	7.8909	0.0000
Nikkei 225 vs Bitcoin	0.1841	0.0475	5.0986	0.0000
S&P500 vs EUR-USD	0.1840	0.0202	10.0960	0.0000
Nikkei 225 vs Brent Crude Oil	0.0970	0.0487	1.4003	0.1662
Brent Crude Oil vs Bitcoin	0.0963	0.0156	2.9217	0.0048
Gold vs Bitcoin	0.0884	0.0567	1.6385	0.1057
HangSeng vs GPR	0.0322	0.0283	0.1723	0.8637
Bitcoin vs GPR	0.0169	-0.0263	3.1626	0.0022
Nikkei 225 vs GPR	0.0113	0.0061	0.1939	0.8468
10yr US T-Bond vs Bitcoin	-0.0060	0.0661	-2.8983	0.0051
S&P500 vs GPR	-0.0332	0.0261	-3.6359	0.0005
EUR-USD vs Bitcoin	-0.0356	0.0672	-4.7107	0.0000
NASDAQ vs GPR	-0.0420	0.0272	-4.3867	0.0000
EUR-USD vs GPR	-0.1063	-0.0314	-5.1073	0.0000
10yr US T-Bond vs GPR	-0.1354	0.0251	-9.4566	0.0000
GPR vs VIX	-0.1681	0.1002	-20.0258	0.0000
Brent Crude Oil vs GPR	-0.1822	0.0322	-12.4323	0.0000

Table 49: Dynamic correlation (DC) by exponential Pearson correlation of analyzed financial instruments, Geopolitical Risk Index (GPR) and VIX during COVID-19 in 02/03/2020 – 04/30/2020. Statistically significant differences between period of geopolitical shock and full sample (DC Complete) by t-test statistic. Source: Own calculations using Matlab (2025).

Financial Instruments	DC Russian War	DC Complete	t-statistic	p-value
S&P500 vs NASDAQ	0.8641	0.8673	-0.5821	0.5618
GPR vs VIX	0.6051	0.1002	22.1412	0.0000
Nikkei 225 vs EUR-USD	0.4046	0.1057	15.0289	0.0000
Gold vs VIX	0.3561	0.0733	16.6758	0.0000
Gold vs GPR	0.3341	0.0620	18.9583	0.0000
S&P500 vs VIX	0.3191	0.3755	-4.4356	0.0000
Nikkei 225 vs Gold	0.3041	0.0685	9.9618	0.0000
NASDAQ vs VIX	0.2690	0.3260	-5.2017	0.0000
Gold vs Brent Crude Oil	0.2576	0.1080	5.7683	0.0000
Nikkei 225 vs HangSeng	0.2558	0.1834	3.3898	0.0011
Brent Crude Oil vs 10yr US T-Bond	0.2553	0.1266	5.8151	0.0000
Nikkei 225 vs VIX	0.2514	0.1779	5.1405	0.0000
Brent Crude Oil vs VIX	0.2497	0.1377	7.3302	0.0000
10yr US T-Bond vs VIX	0.2146	0.1889	1.8770	0.0638
Brent Crude Oil vs Bitcoin	0.2035	0.0156	11.3906	0.0000
HangSeng vs VIX	0.1949	0.1362	3.7973	0.0003
Nikkei 225 vs GPR	0.1889	0.0061	19.0111	0.0000
NASDAQ vs Bitcoin	0.1865	0.1301	3.2416	0.0017
S&P500 vs GPR	0.1831	0.0261	9.3331	0.0000
Brent Crude Oil vs GPR	0.1771	0.0322	10.0209	0.0000
Bitcoin vs VIX	0.1685	0.0841	5.0529	0.0000
S&P500 vs Brent Crude Oil	0.1685	0.1300	2.3100	0.0237
NASDAQ vs HangSeng	0.1671	0.0671	3.7892	0.0003
10yr US T-Bond vs Bitcoin	0.1630	0.0661	6.7901	0.0000
Nikkei 225 vs Brent Crude Oil	0.1429	0.0487	4.7224	0.0000
HangSeng vs GPR	0.1394	0.0283	4.7560	0.0000
NASDAQ vs Brent Crude Oil	0.1271	0.0981	2.1333	36.0000
HangSeng vs EUR-USD	0.1239	0.1150	0.4646	0.6436
HangSeng vs Brent Crude Oil	0.1198	0.0748	2.1825	0.0324
10yr US T-Bond vs GPR	0.1144	0.0251	3.8911	0.0002
Gold vs EUR-USD	0.1126	0.1071	0.3261	0.7453
S&P500 vs Gold	0.1117	0.1048	0.2059	0.8375
S&P500 vs HangSeng	0.1104	0.0774	1.1791	0.2426
EUR-USD vs VIX	0.1096	0.0612	2.8703	0.0054
EUR-USD vs GPR	0.0971	-0.0314	8.9871	0.0000
S&P500 vs Bitcoin	0.0940	0.1115	-0.8833	0.3800
NASDAQ vs GPR	0.0866	0.0272	4.1956	0.0001
EUR-USD vs Bitcoin	0.0650	0.0672	-0.0904	0.9283
S&P500 vs EUR-USD	0.0574	0.0202	1.5024	0.1377
Gold vs 10yr US T-Bond	0.0553	0.2174	-10.7424	0.0000
HangSeng vs Gold	0.0491	0.0470	0.1089	0.9136
NASDAQ vs EUR-USD	0.0439	0.0433	0.0295	0.9765
Bitcoin vs GPR	0.0415	-0.0263	4.4784	0.0000
Brent Crude Oil vs EUR-USD	0.0265	0.0623	-2.2974	0.0244
NASDAQ vs 10yr US T-Bond	0.0116	0.1727	-8.9536	0.0000
S&P500 vs 10yr US T-Bond	0.0083	0.1988	-9.8089	0.0000
S&P500 vs Nikkei 225	-0.0197	0.1087	-8.0181	0.0000
NASDAQ vs Gold	-0.0198	0.1034	-5.4790	0.0000
Gold vs Bitcoin	-0.0425	0.0567	-5.4437	0.0000
10yr US T-Bond vs EUR-USD	-0.0515	0.0913	-8.9726	0.0000
Nikkei 225 vs Bitcoin	-0.0788	0.0475	-11.4608	0.0000
HangSeng vs Bitcoin	-0.0857	0.0060	-6.9121	0.0000
NASDAQ vs Nikkei 225	-0.1116	0.0816	-12.4151	0.0000
HangSeng vs 10yr US T-Bond	-0.1436	0.0331	-11.0352	0.0000
Nikkei 225 vs 10yr US T-Bond	-0.1738	0.0797	-17.6149	0.0000
S&P500 vs NASDAQ	0.8641	0.8673	-0.5821	0.5618
GPR vs VIX	0.6051	0.1002	22.1412	0.0000
Nikkei 225 vs EUR-USD	0.4046	0.1057	15.0289	0.0000
Gold vs VIX	0.3561	0.0733	16.6758	0.0000
Gold vs GPR	0.3341	0.0620	18.9583	0.0000
S&P500 vs VIX	0.3191	0.3755	-4.4356	0.0000
Nikkei 225 vs Gold	0.3041	0.0685	9.9618	0.0000
NASDAQ vs VIX	0.2690	0.3260	-5.2017	0.0000

Table 50: Dynamic correlation (DC) by exponential Pearson correlation of analyzed financial instruments, Geopolitical Risk Index (GPR) and VIX during the Russian invasion in 02/01/2022 – 04/29/2022. Statistically significant differences between period of geopolitical shock and full sample (DC Complete) by t-test statistic. Source: Own calculations using Matlab (2025).

Financial Instruments	DC Hamas Attack	DC Complete	t-statistic	p-value
S&P500 vs NASDAQ	0.9139	0.8673	11.5801	0.0000
GPR vs VIX	0.3491	0.1002	7.5812	0.0000
HangSeng vs EUR-USD	0.2986	0.1150	7.6480	0.0000
NASDAQ vs VIX	0.2907	0.3260	-2.7212	0.0078
S&P500 vs VIX	0.2847	0.3755	-6.7881	0.0000
Gold vs Brent Crude Oil	0.2819	0.1080	9.6443	0.0000
NASDAQ vs 10yr US T-Bond	0.2816	0.1727	6.7459	0.0000
S&P500 vs 10yr US T-Bond	0.2654	0.1988	3.9830	0.0001
Nikkei 225 vs 10yr US T-Bond	0.2420	0.0797	7.6194	0.0000
S&P500 vs GPR	0.2245	0.0261	14.8214	0.0000
Nikkei 225 vs EUR-USD	0.2096	0.1057	4.6531	0.0000
Nikkei 225 vs VIX	0.2029	0.1779	1.7142	0.0904
EUR-USD vs Bitcoin	0.1914	0.0672	5.5915	0.0000
NASDAQ vs GPR	0.1833	0.0272	10.6425	0.0000
10yr US T-Bond vs GPR	0.1608	0.0251	7.0622	0.0000
Brent Crude Oil vs VIX	0.1472	0.1377	0.6313	0.5297
Brent Crude Oil vs 10yr US T-Bond	0.1395	0.1266	0.8606	0.3919
Gold vs 10yr US T-Bond	0.1388	0.2174	-5.5345	0.0000
Bitcoin vs GPR	0.1318	-0.0263	9.0823	0.0000
S&P500 vs Nikkei 225	0.1201	0.1087	0.6216	536.0000
NASDAQ vs Gold	0.1186	0.1034	0.6588	0.5122
Gold vs GPR	0.1031	0.0620	2.3906	0.0194
NASDAQ vs Nikkei 225	0.0915	0.0816	0.5803	0.5634
HangSeng vs Brent Crude Oil	0.0903	0.0748	1.0265	0.3076
Nikkei 225 vs Brent Crude Oil	0.0890	0.0487	1.9712	0.0526
10yr US T-Bond vs VIX	0.0884	0.1889	-6.5645	0.0000
Nikkei 225 vs Bitcoin	0.0874	0.0475	3.2008	0.0019
10yr US T-Bond vs EUR-USD	0.0760	0.0913	-1.1689	0.2452
S&P500 vs Gold	0.0718	0.1048	-1.8908	0.0624
HangSeng vs Gold	0.0707	0.0470	0.9306	0.3553
Gold vs VIX	0.0496	0.0733	-0.9769	0.3319
HangSeng vs VIX	0.0304	0.1362	-5.7676	0.0000
10yr US T-Bond vs Bitcoin	0.0293	0.0661	-4.1213	0.0001
Gold vs Bitcoin	0.0281	0.0567	-1.1417	0.2575
HangSeng vs Bitcoin	0.0249	0.0060	1.3270	0.1881
Nikkei 225 vs GPR	0.0064	0.0061	0.0181	0.9856
Bitcoin vs VIX	0.0061	0.0841	-4.3821	0.0000
Brent Crude Oil vs Bitcoin	-0.0053	0.0156	-1.1653	0.2477
NASDAQ vs Brent Crude Oil	-0.0068	0.0981	-6.8980	0.0000
Brent Crude Oil vs EUR-USD	-0.0095	0.0623	-2.9235	0.0047
NASDAQ vs Bitcoin	-0.0105	0.1301	-9.3661	0.0000
S&P500 vs Bitcoin	-0.0197	0.1115	-9.4656	0.0000
Brent Crude Oil vs GPR	-0.0374	0.0322	-3.8828	0.0002
Nikkei 225 vs HangSeng	-0.0639	0.1834	-9.4219	0.0000
Gold vs EUR-USD	-0.0894	0.1071	-11.3407	0.0000
S&P500 vs Brent Crude Oil	-0.0968	0.1300	-16.4279	0.0000
HangSeng vs GPR	-0.1077	0.0283	-8.3285	0.0000
EUR-USD vs VIX	-0.1150	0.0612	-18.4708	0.0000
HangSeng vs 10yr US T-Bond	-0.1303	0.0331	-8.8607	0.0000
EUR-USD vs GPR	-0.1440	-0.0314	-7.9347	0.0000
NASDAQ vs HangSeng	-0.1996	0.0671	-16.0907	0.0000
Nikkei 225 vs Gold	-0.2286	0.0685	-20.1083	0.0000
S&P500 vs HangSeng	-0.2542	0.0774	-24.0926	0.0000
S&P500 vs EUR-USD	-0.2544	0.0202	-24.0646	0.0000
NASDAQ vs EUR-USD	-0.2585	0.0433	-29.4377	0.0000
S&P500 vs NASDAQ	0.9139	0.8673	11.5801	0.0000
GPR vs VIX	0.3491	0.1002	7.5812	0.0000
HangSeng vs EUR-USD	0.2986	0.1150	7.6480	0.0000
NASDAQ vs VIX	0.2907	0.3260	-2.7212	0.0078
S&P500 vs VIX	0.2847	0.3755	-6.7881	0.0000
Gold vs Brent Crude Oil	0.2819	0.1080	9.6443	0.0000
NASDAQ vs 10yr US T-Bond	0.2816	0.1727	6.7459	0.0000
S&P500 vs 10yr US T-Bond	0.2654	0.1988	3.9830	0.0001

Table 51: Dynamic correlation (DC) by exponential Pearson correlation of analyzed financial instruments, Geopolitical Risk Index (GPR) and VIX during the Hamas attack from 10/01/2023 – 12/29/2023. Statistically significant differences between period of geopolitical shock and full sample (DC Complete) by t-test statistic. Source: Own calculations using Matlab (2025).

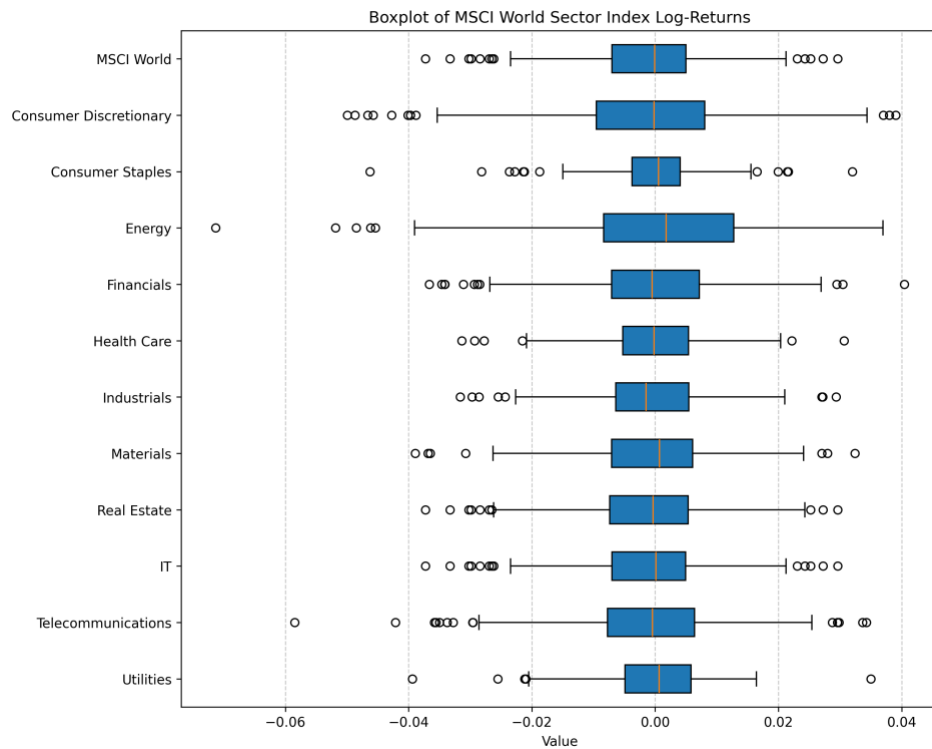


Figure 11: Boxplots of MSCI World Sector Index log-returns for 07/01/2021 – 06/30/2022.
Source: Own illustration using Python Matplotlib (2025).

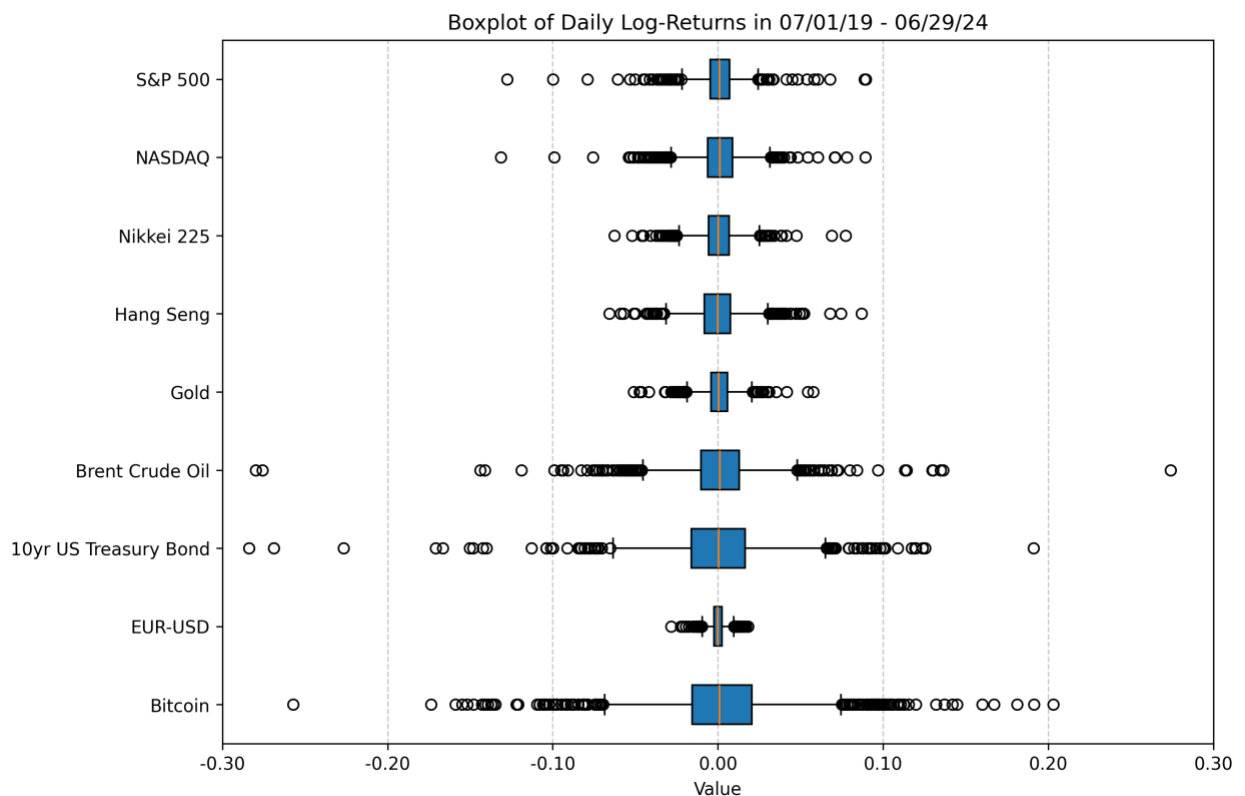


Figure 12: Boxplots of daily log-returns for for financial instrument in 07/01/2019 – 06/29/2024.
Source: Own illustration using Python Matplotlib (2025).

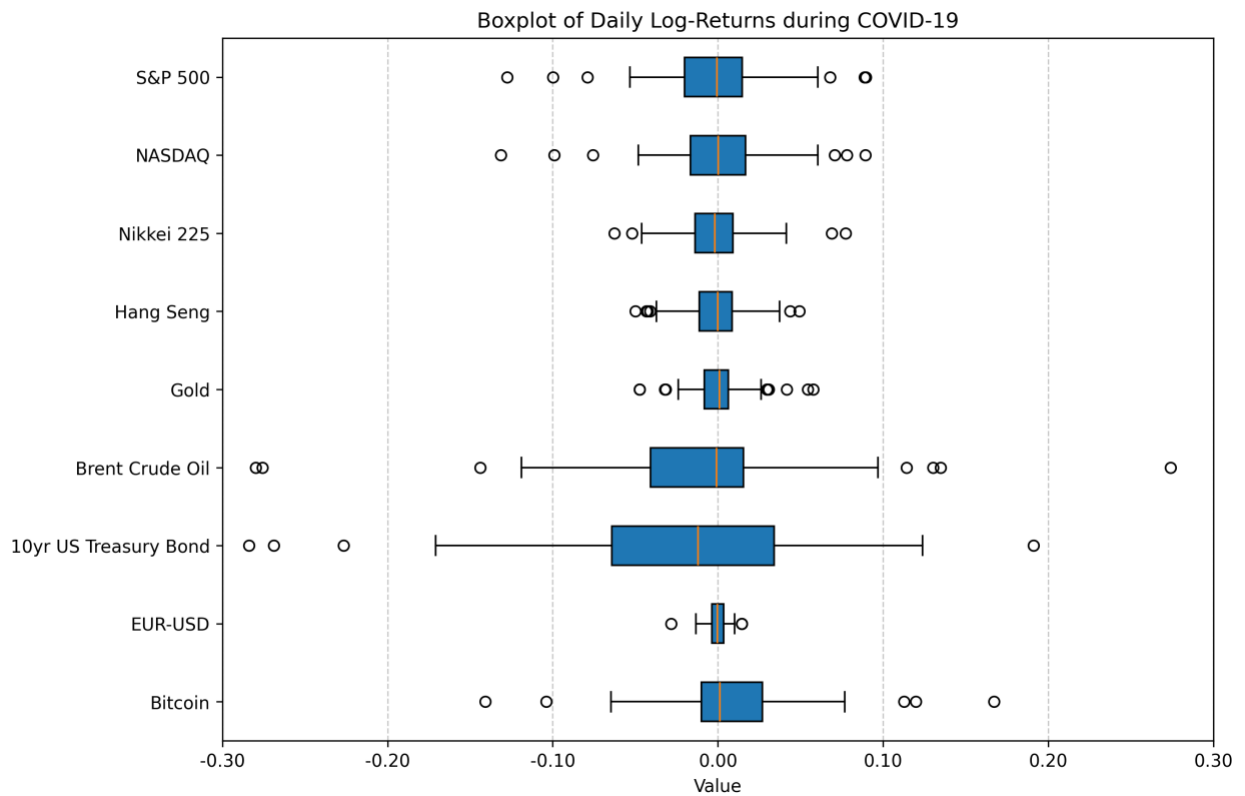


Figure 13: Boxplots of daily log-returns for selected financial instruments during COVID-19 period in 02/03/2020 – 04/30/2020. Source: Own illustration using Python Matplotlib (2025).

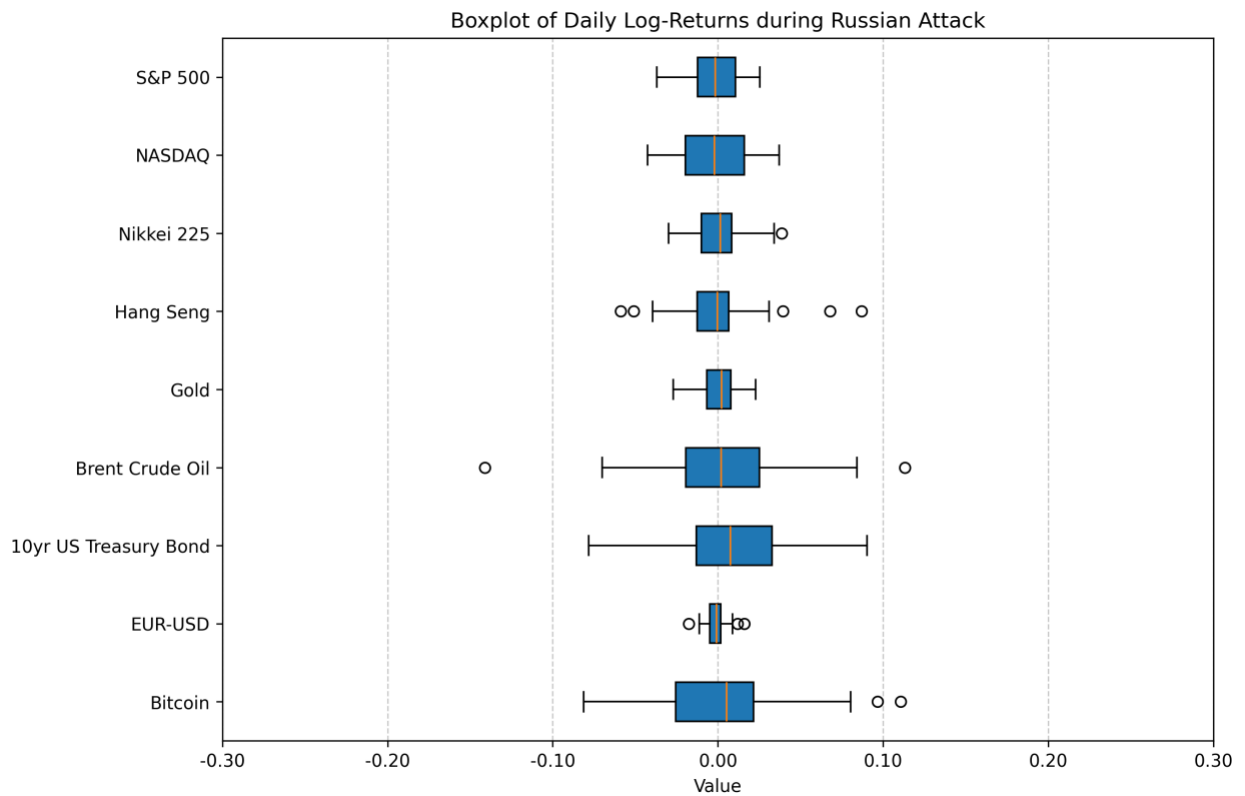


Figure 14: Boxplots of daily log-returns for selected financial instruments during the Russian attack period in 02/01/2022 – 04/29/2022.

Source: Own illustration using Python Matplotlib (2025).

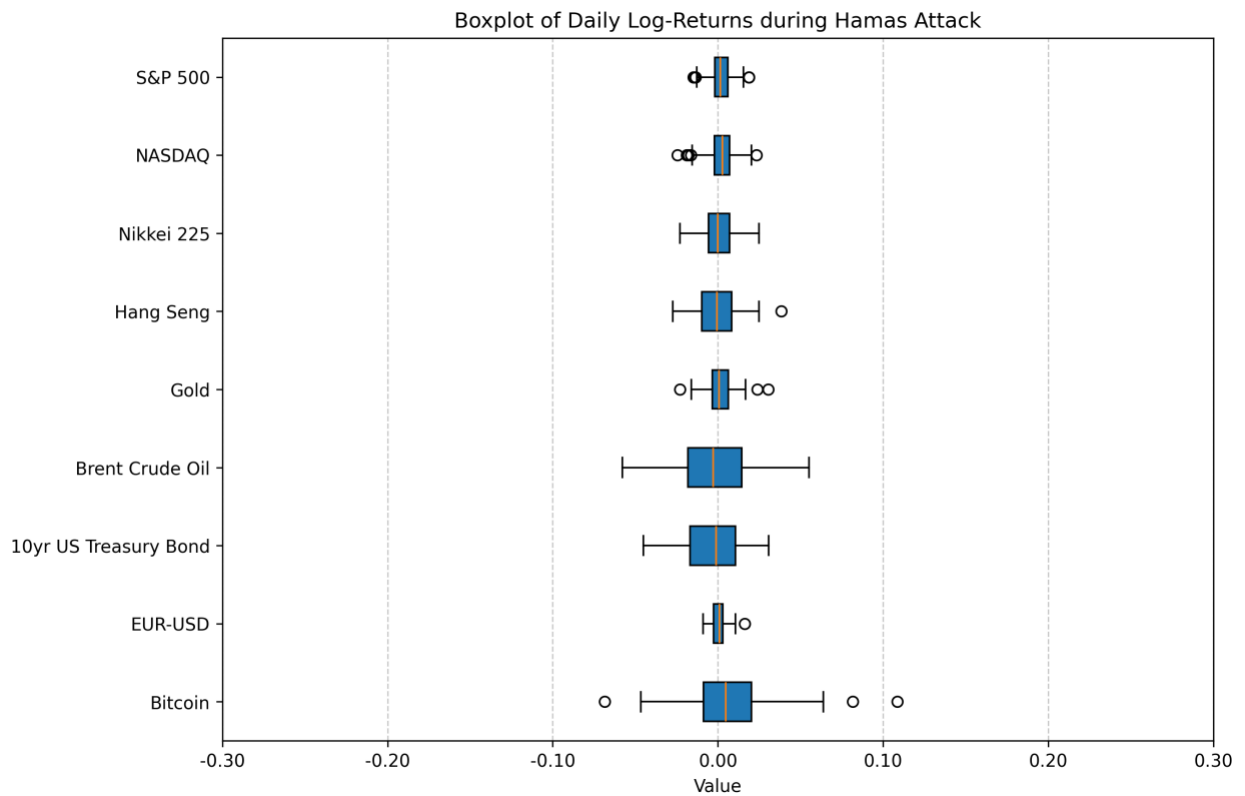


Figure 15: Boxplots of daily log-returns for selected financial instruments during the Hamas attack period in 10/02/2023 – 12/29/2023.
Source: Own illustration using Python Matplotlib (2025).

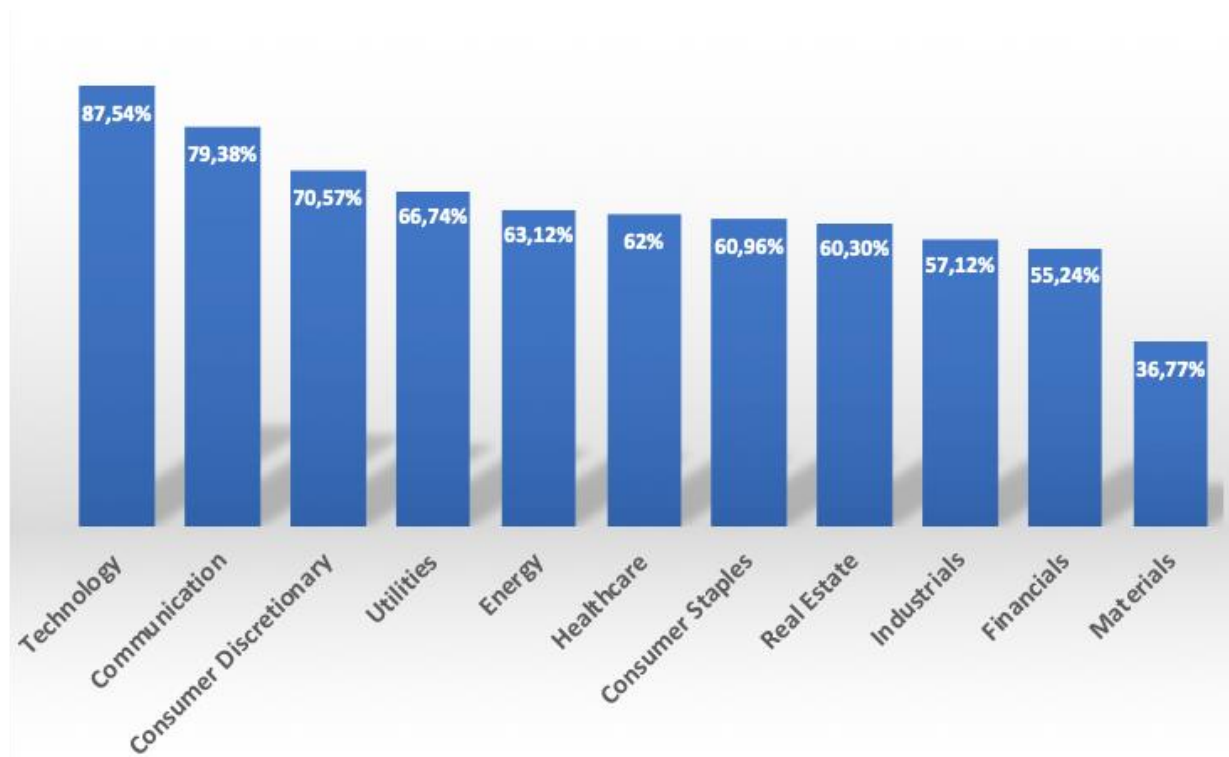


Figure 16: Weighting by market capitalization of U.S. companies in MSCI World sector indices as at 02/2022 (%). Source: Own edition using MS Excel (2024).

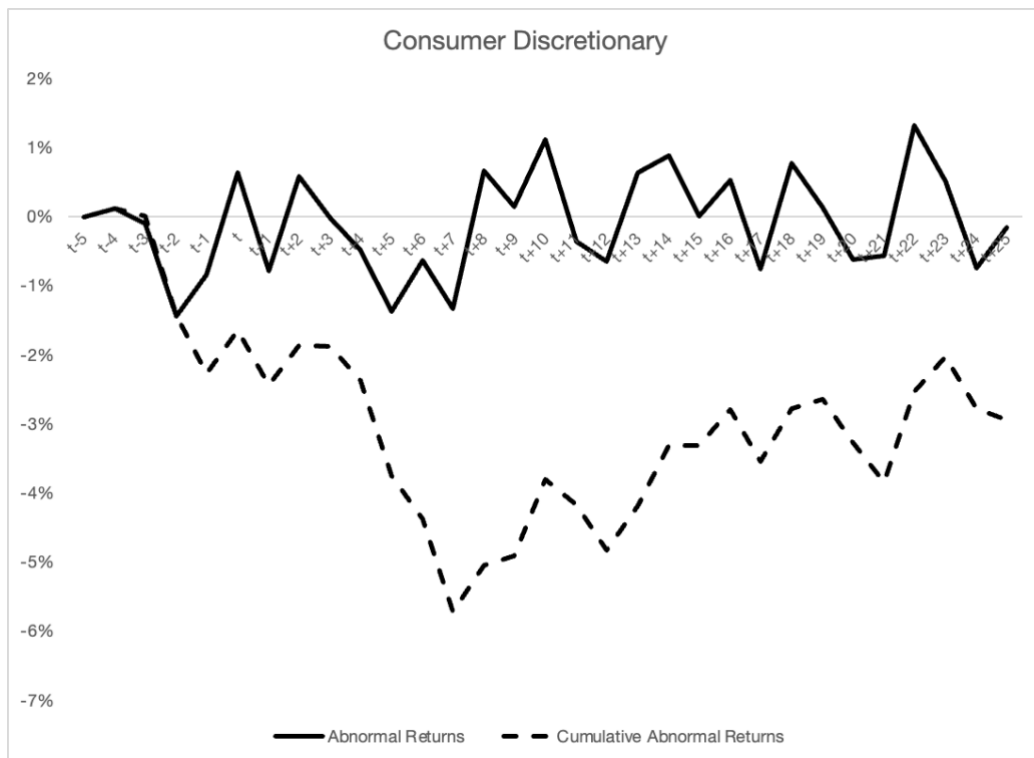


Figure 17: Abnormal and cumulative abnormal returns around event day for MSCI World Consumer Discretionary Index (%).
Source: Robus et al. (2024).



Figure 18: Abnormal and cumulative abnormal returns around event day for MSCI World Consumer Staples Index (%). Source: Robus et al. (2024).

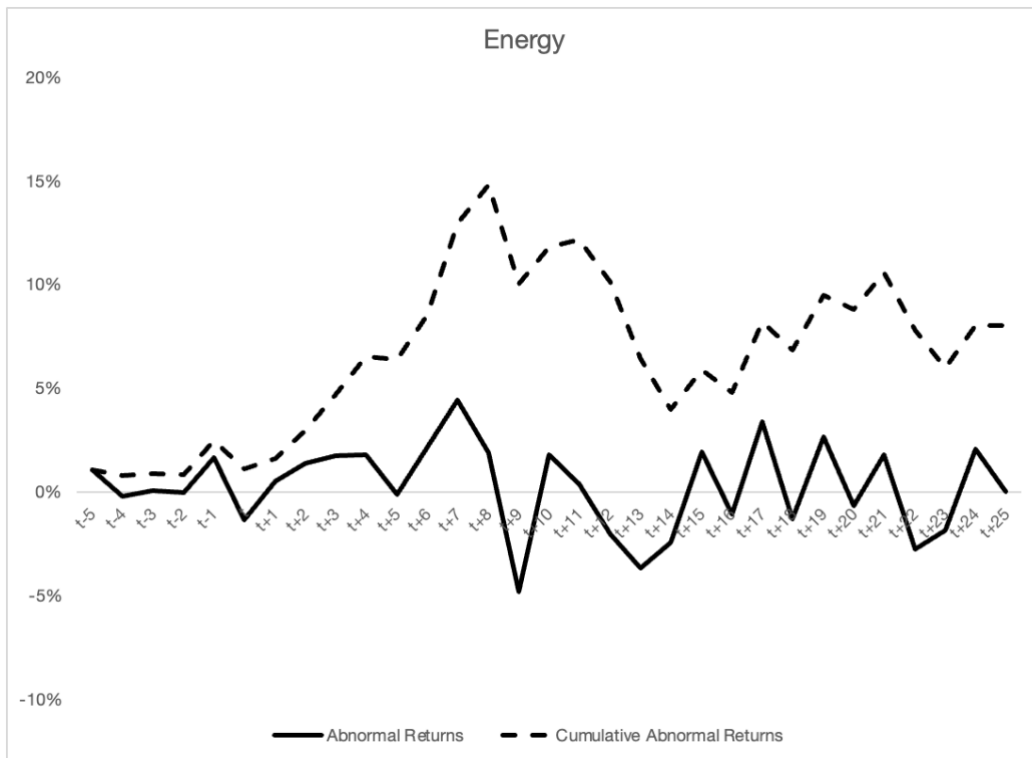


Figure 19: Abnormal and cumulative abnormal returns around event day for MSCI World Energy Index (%). Source: Robus et al. (2024).

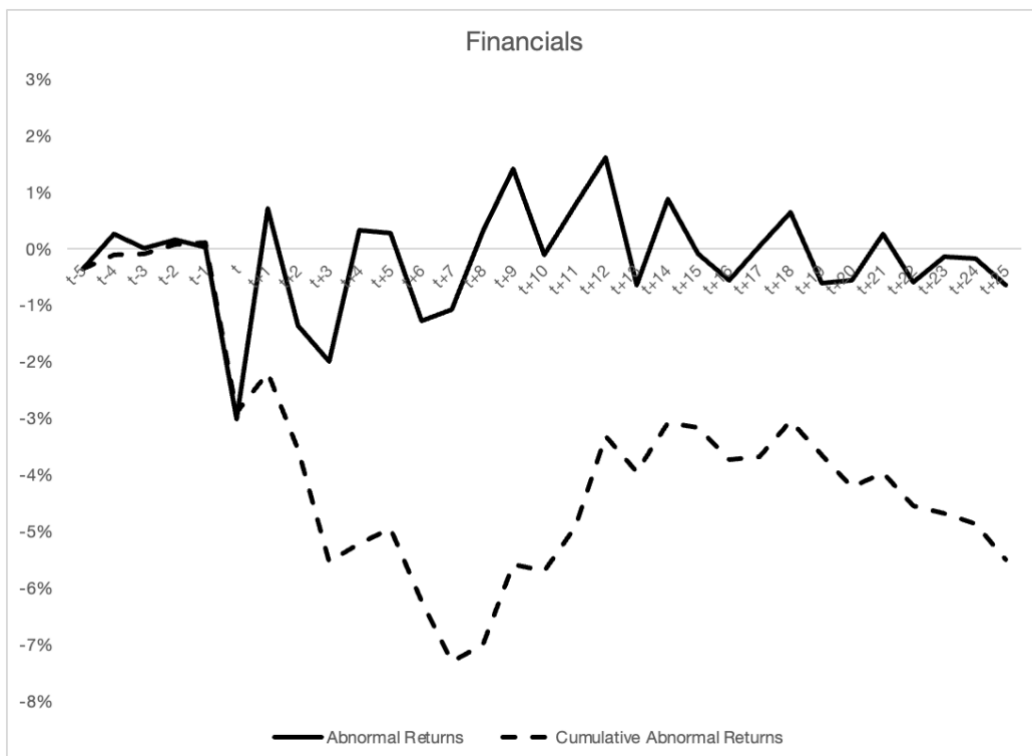


Figure 20: Abnormal and cumulative abnormal returns around event day for MSCI World Financials Index (%). Source: Robus et al. (2024).

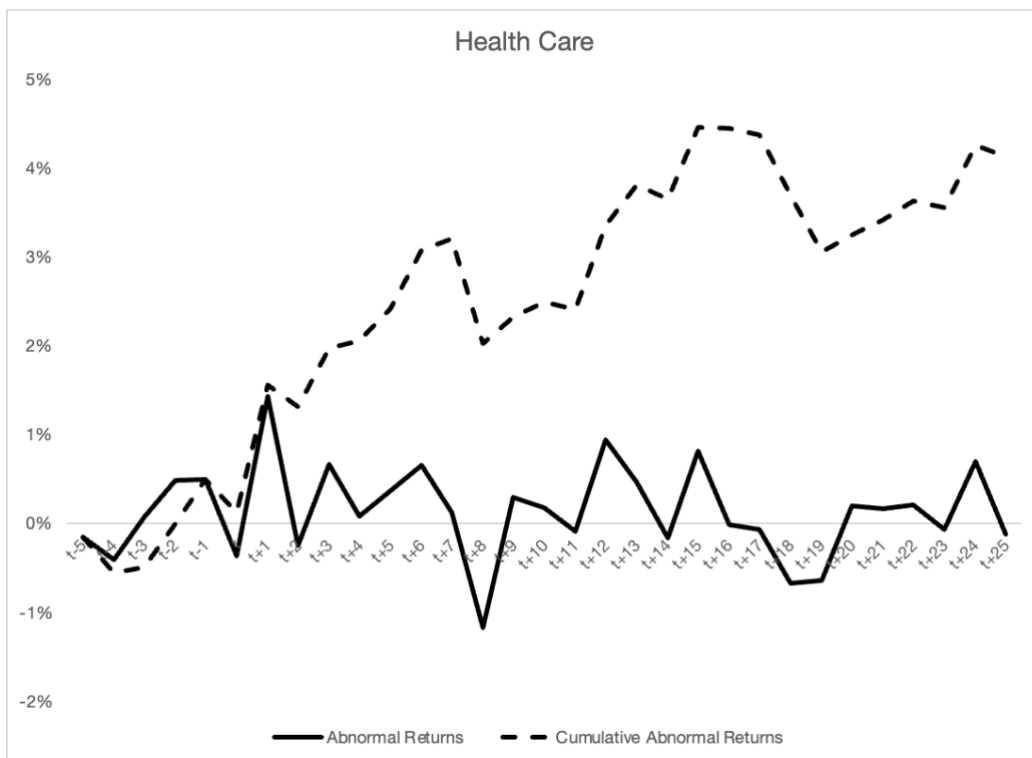


Figure 21: Abnormal and cumulative abnormal returns around event day for MSCI World Health Care Index (%). Source: Robus et al. (2024).

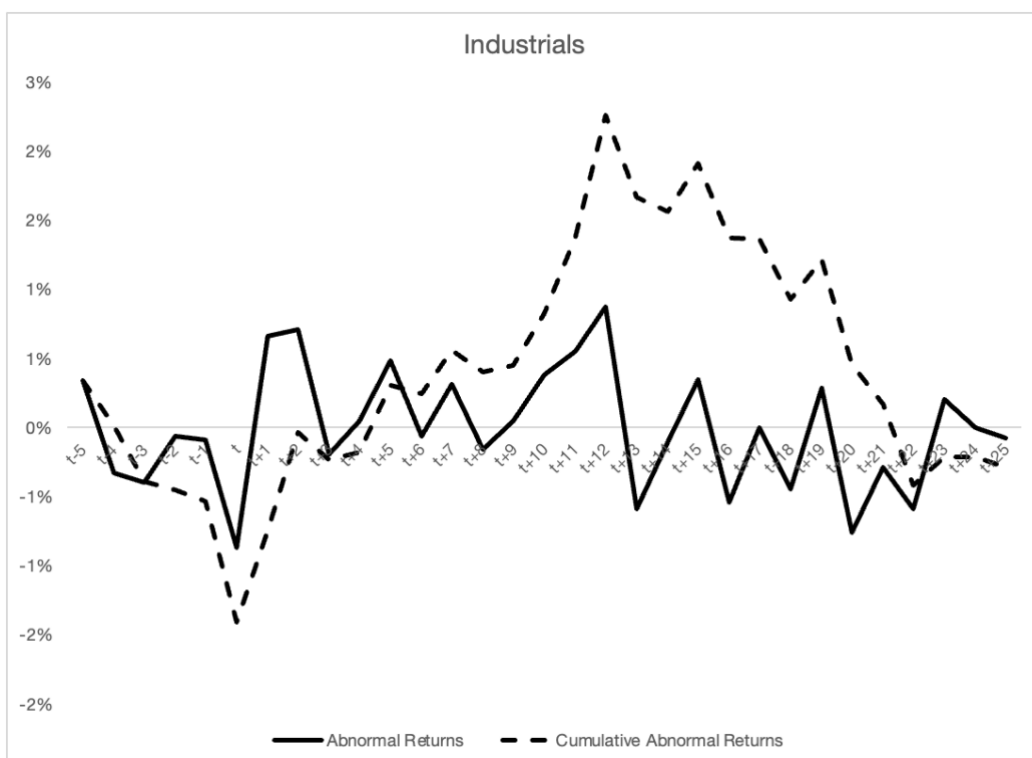


Figure 22: Abnormal and cumulative abnormal returns around event day for MSCI World Industrials Index (%). Source: Robus et al. (2024).

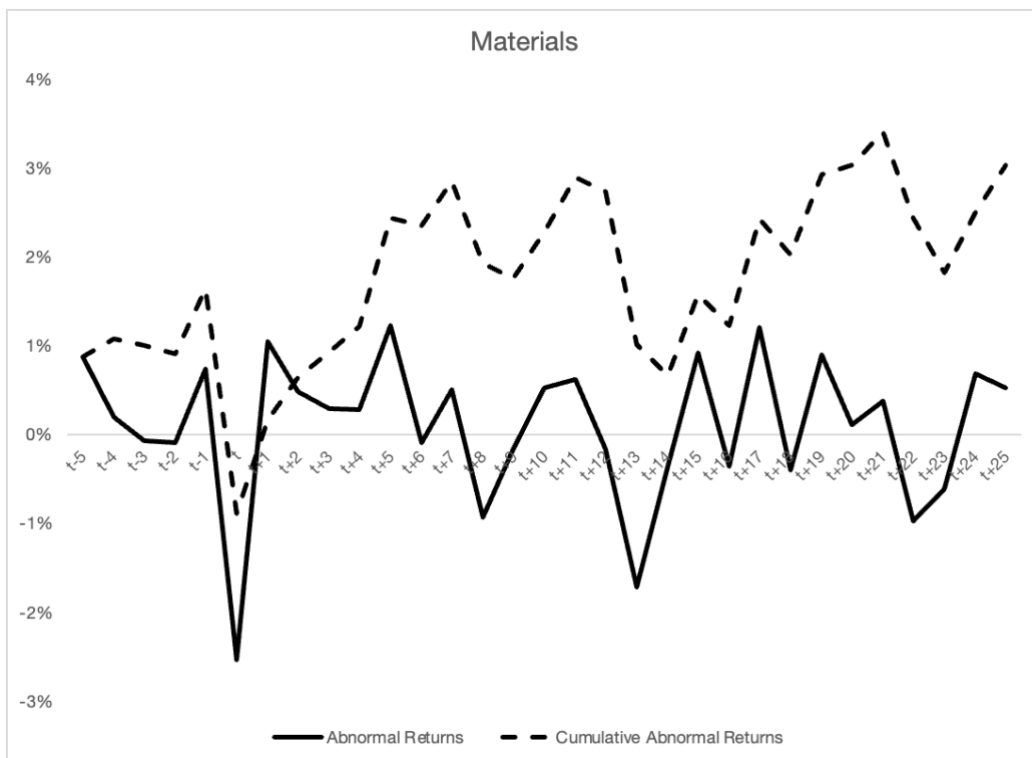


Figure 23: Abnormal and cumulative abnormal returns around event day for MSCI World Materials Index (%). Source: Robus et al. (2024).

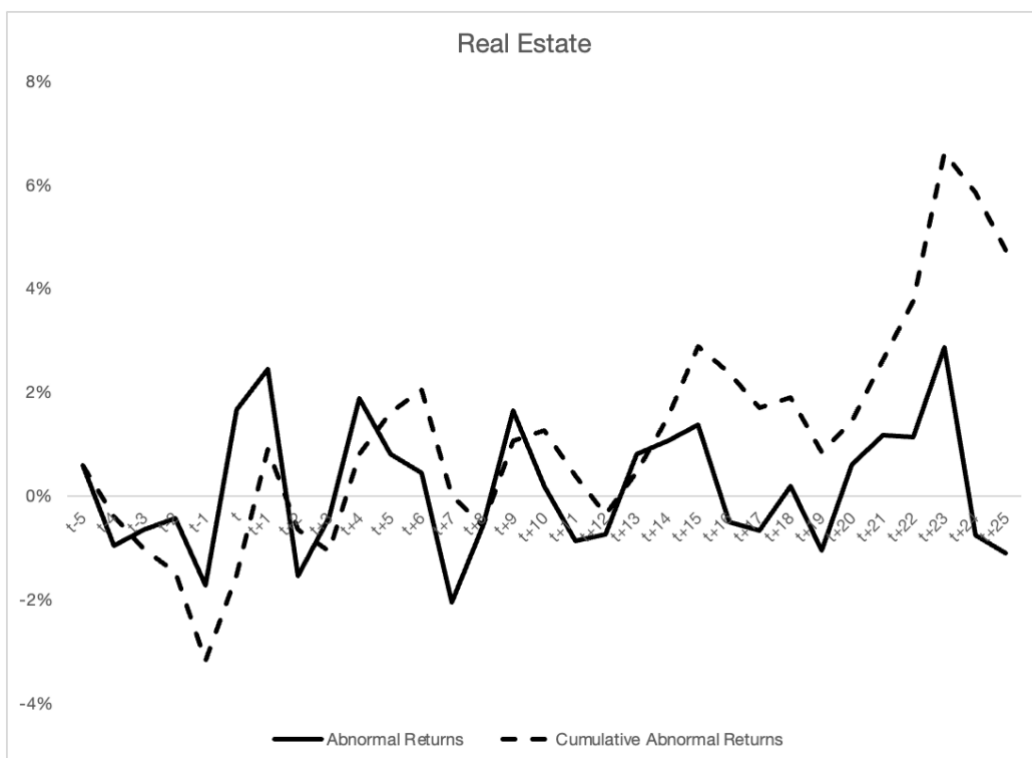


Figure 24: Abnormal and cumulative abnormal returns around event day for MSCI World Real Estate Index (%). Source: Robus et al. (2024).

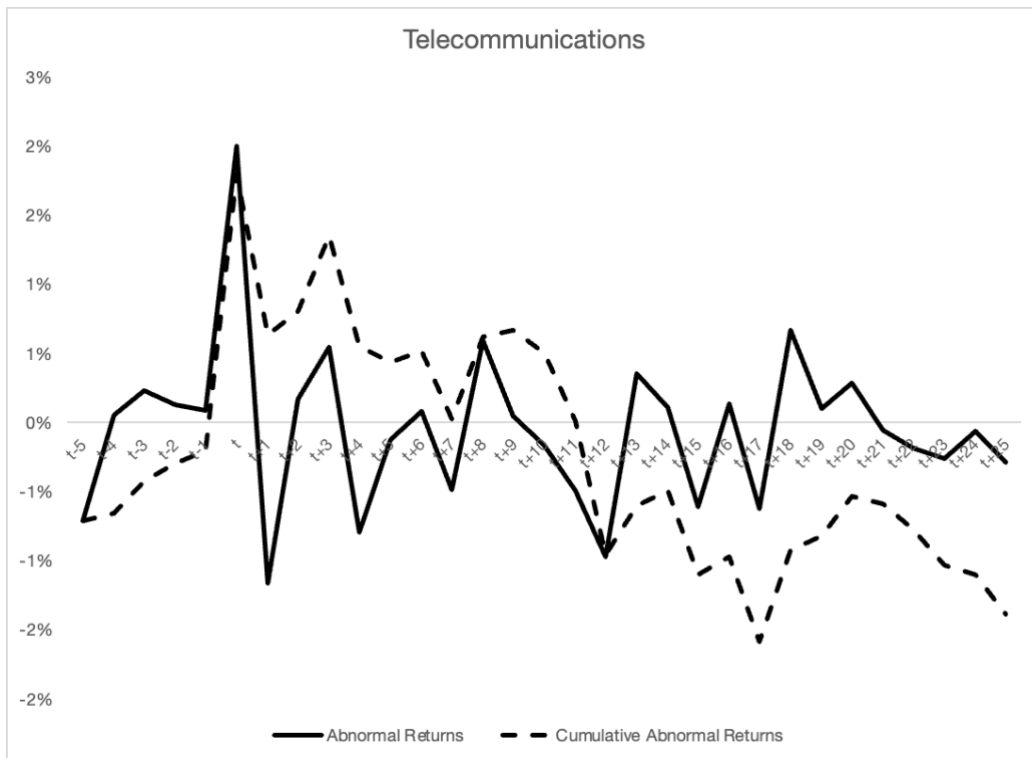


Figure 25: Abnormal and cumulative abnormal returns around event day for MSCI World Telecommunications Index (%). Source: Robus et al. (2024).

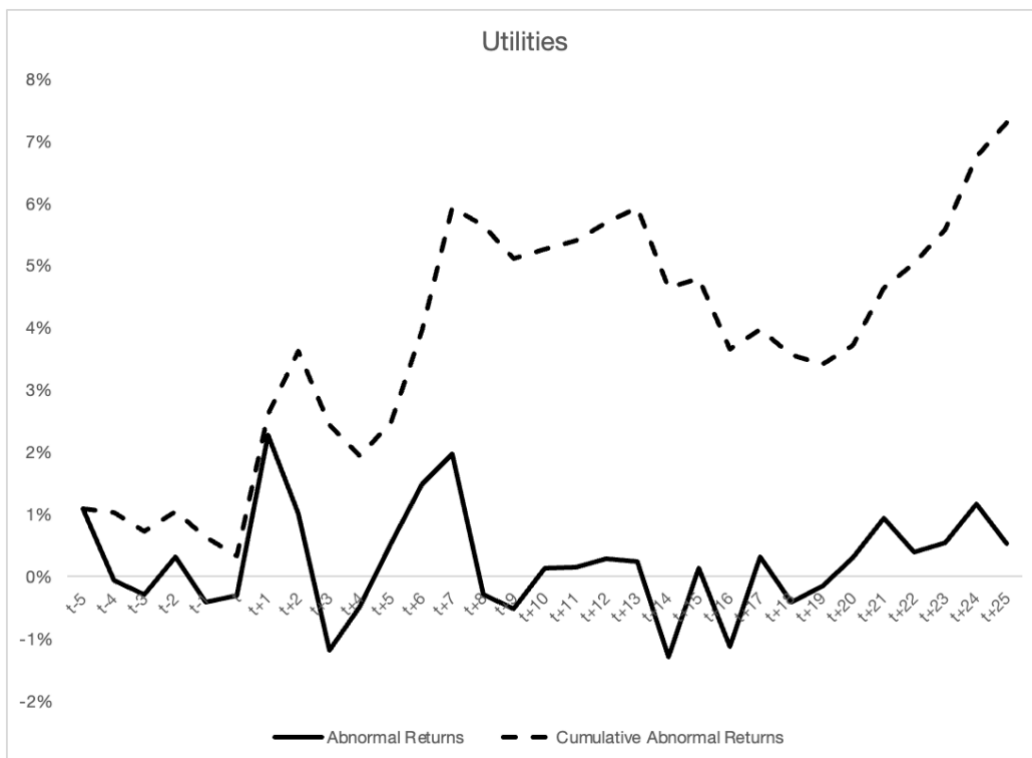


Figure 26: Abnormal and cumulative abnormal returns around event day for MSCI World Utilities Index (%). Source: Robus et al. (2024).

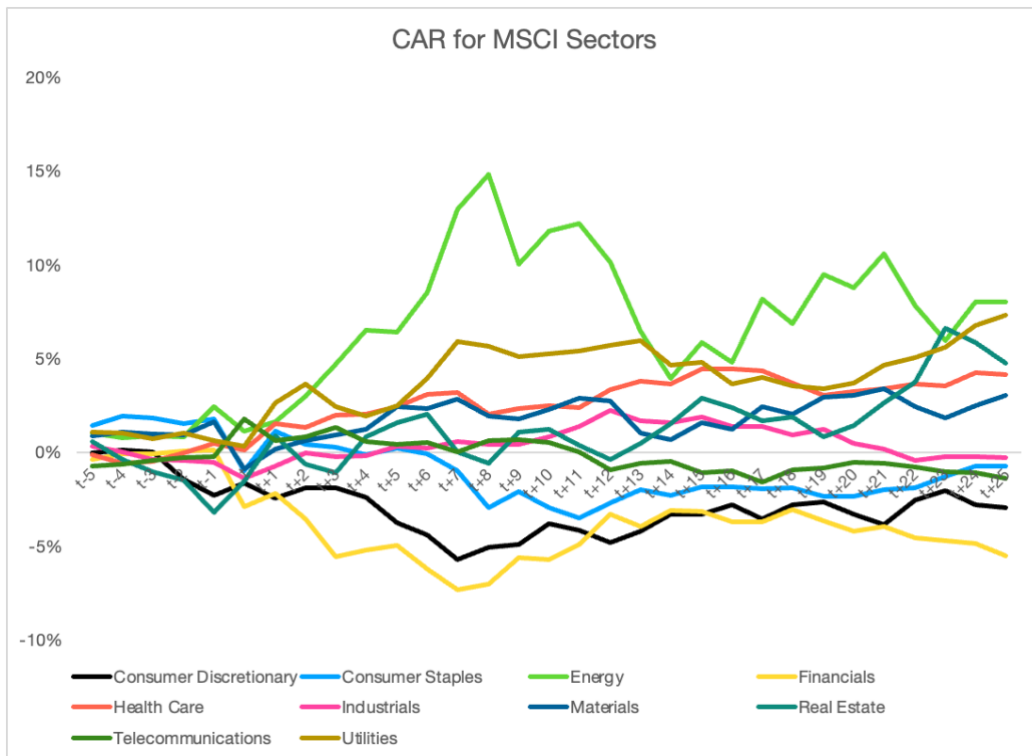


Figure 27: Cumulative abnormal returns around event day for all MSCI World Sector Indices (%). Source: Robus et al. (2024).

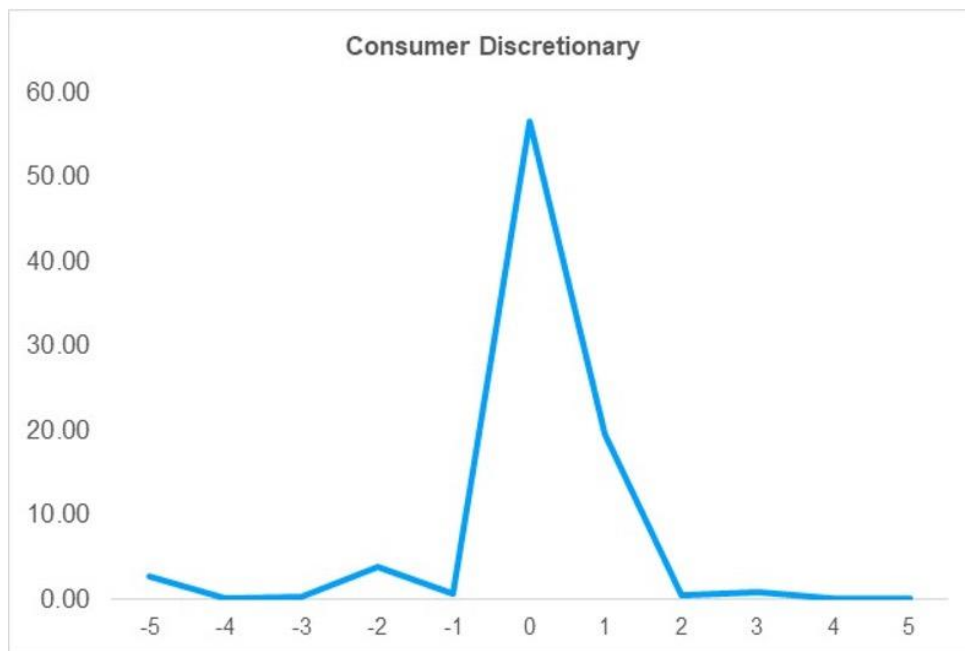


Figure 28: Abnormal Volatility (%) of Consumer Discretionary Sector Index (02/17/22 – 03/03/22). Source: Robus et al. (in press).

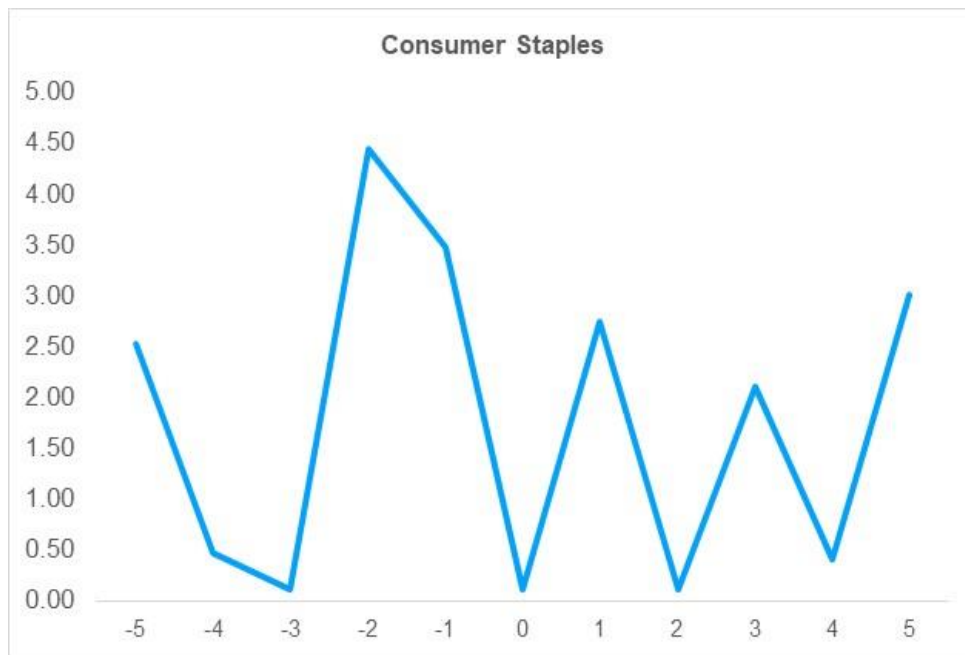


Figure 29: Abnormal Volatility (%) of Consumer Staples Sector Index (02/17/22 – 03/03/22). Source: Robus et al. (in press).

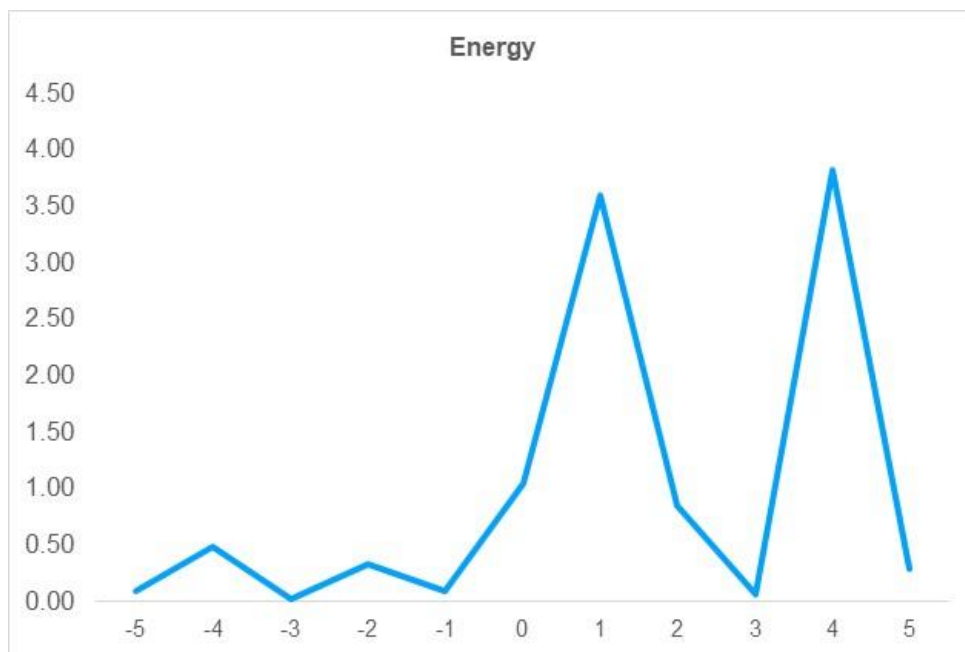


Figure 30: Abnormal Volatility (%) of Energy Sector Index (02/17/22 – 03/03/22). Source: Robus et al. (in press).

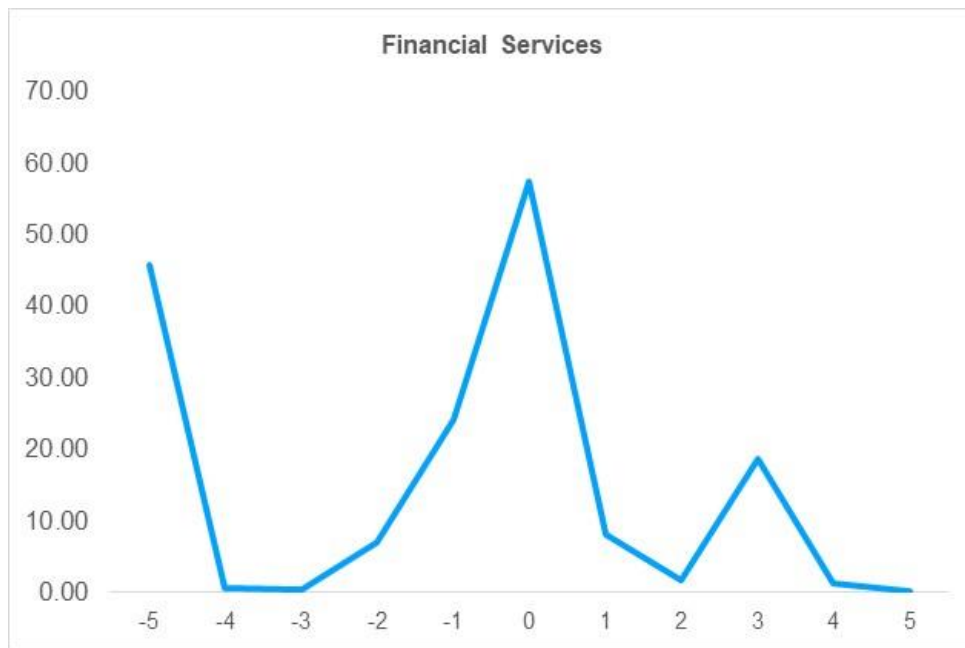


Figure 31: Abnormal Volatility (%) of Financials Sector Index (02/17/22 – 03/03/22). Source: Robus et al. (press).

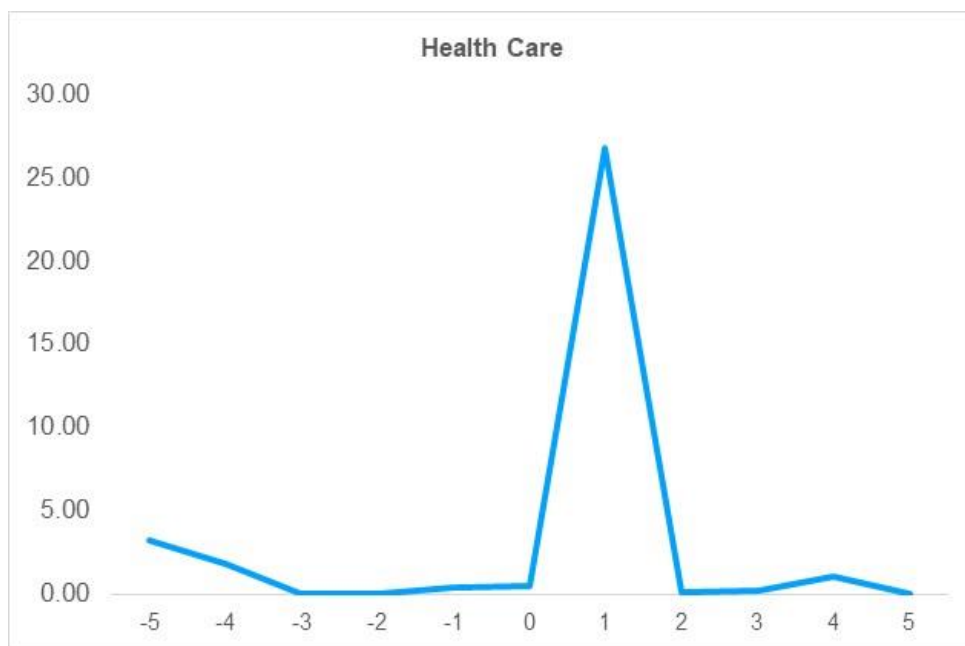


Figure 32: Abnormal Volatility (%) of Health Care Sector Index (02/17/22 – 03/03/22). Source: Robus et al. (in press).

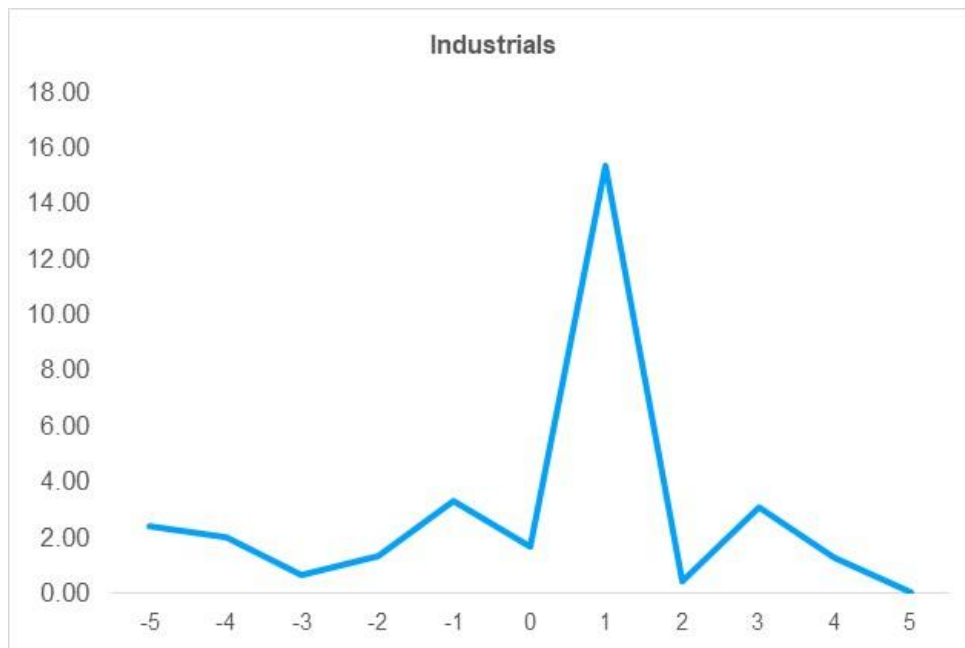


Figure 33: Abnormal Volatility (%) of Industrials Sector Index (02/17/22 – 03/03/22). Source: Robus et al. (in press).

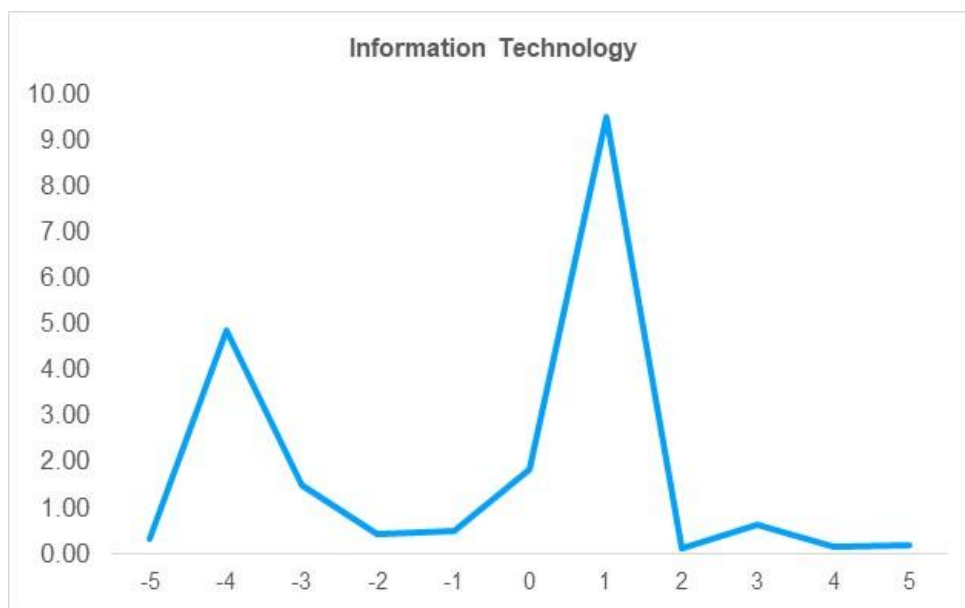


Figure 34: Abnormal Volatility (%) of IT Sector Index (02/17/22 – 03/03/22). Source: Robus et al. (in press).

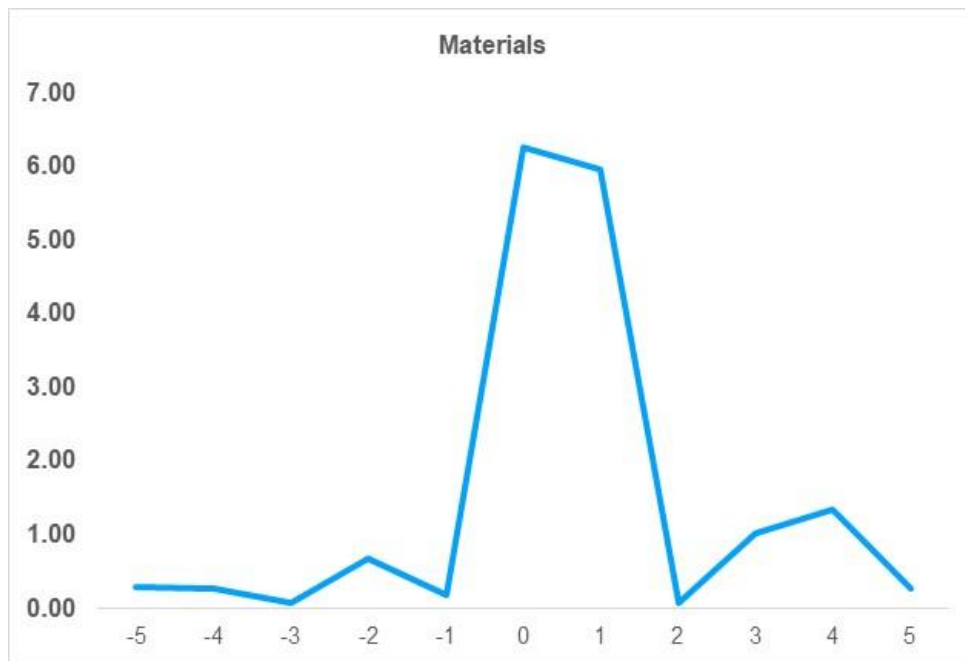


Figure 35: Abnormal Volatility (%) of Materials Sector Index (02/17/22 – 03/03/22). Source: Robus et al. (in press).

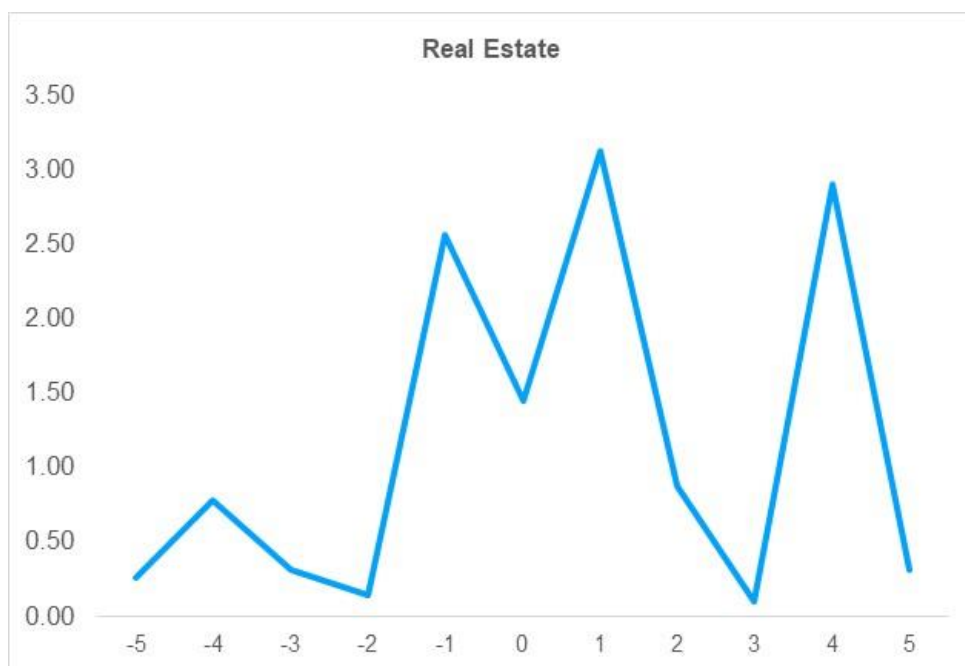


Figure 36: Abnormal Volatility (%) of Real Estate Sector Index (02/17/22 – 03/03/22). Source: Robus et al. (in press).

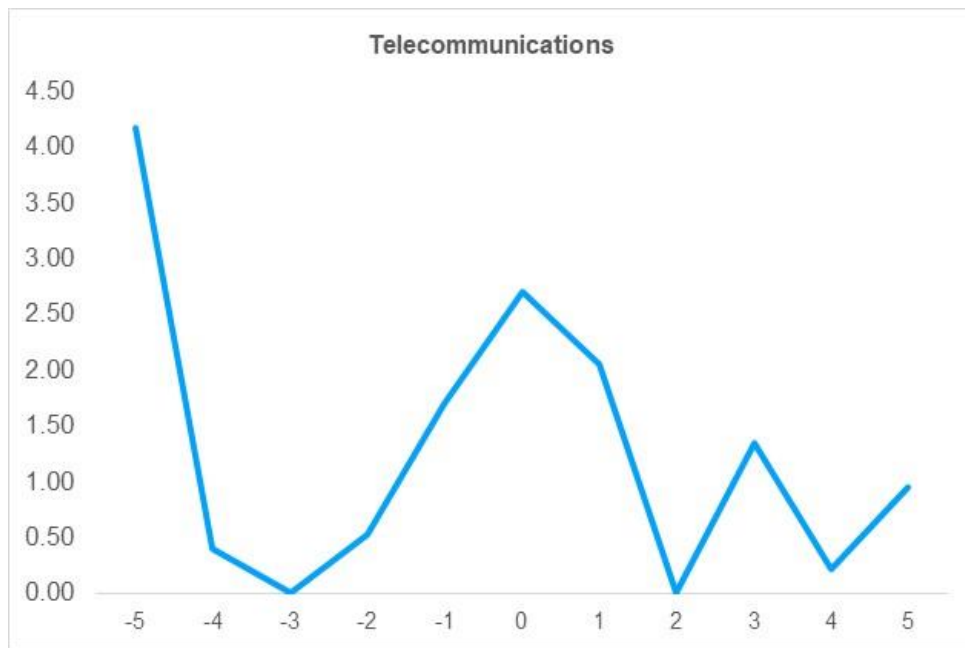


Figure 37: Abnormal Volatility (%) of Telecommunications Sector Index (02/17/22 – 03/03/22). Source: Robus et al. (in press).

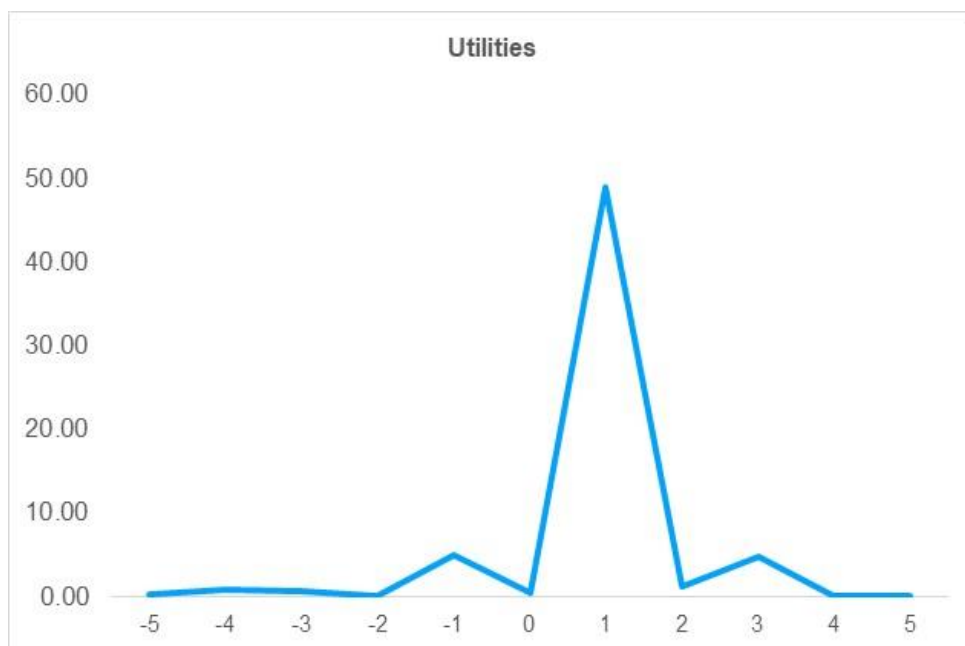


Figure 38: Abnormal Volatility (%) of Utilities Sector Index (02/17/22 – 03/03/22). Source: Robus et al. (in press).

ACKNOWLEDGMENTS

At this point, I would like to thank all the people who have accompanied and supported me during the writing of this dissertation.

My special thanks go to my doctoral supervisors, Dr. Virág Walter and Dr. habil. Zsolt Kőműves who was always at my side with their expertise, patient support and valuable suggestions. Their critical questions and comments have contributed significantly to the quality of this work.

I would also like to thank the staff of the Hungarian University of Agriculture and Life Sciences for their support and assistance during my doctoral studies.

I would also like to thank my colleague Rainer Remke for his support in getting the funding and for his encouragement during the assessment for my doctoral program.

My deepest gratitude goes to my family, especially my grandparents, whom I would like to commemorate here, and who have always inspired and supported me.

Finally, I would like to thank my beloved wife for her support, patience and encouragement. Without you, this work would not have been possible.