ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
Bachelor Thesis Economics & Business
Specialization: Financial Economics

From Corporate Ethics to Market Dynamics: The Predictive Power of ESG Factors in TARCH Models

Author: Daan Takx
Student number: 622984
Thesis supervisor: Kan Ji
Second reader: Ruben de Bliek
Finish date: July 1, 2024

ABSTRACT

This study explores the integration of ESG factors into TARCH models to improve the predictive accuracy of stock market volatility forecasts for the S&P 500 price index from 2020 to 2024. By constructing and comparing a conventional TARCH model and an ESG-enhanced TARCH model, the results indicate that the ESG-enhanced model improved forecasting accuracy, evidenced by lower RMSE and favorable likelihood-ratio and DM test results. The findings suggest that incorporating ESG metrics into financial models can enhance the prediction of market behaviors, underscoring the importance of sustainable investment practices and contributing to the evolving field of financial risk analysis.

Keywords: ESG Integration; TARCH Models; Stock Market Volatility; Predictive Accuracy; Sustainable Investment

TABLE OF CONTENTS

CHAPTER 1 Introduction

The fluctuations in price levels of assets, known as 'volatility', and the ability to forecast this phenomenon, is a cornerstone in modern financial analysis. Researchers employ different strategies to predict volatility, and several models approach this task with varying characteristics, advantages, and disadvantages. Apart from historical, realized, implied, and stochastic volatility (SV) models, examined by Koopman et al. (2005) and Andersen & Benzoni (2009), Threshold Autoregressive Conditional Heteroskedasticity (TARCH) models have shown predictive power in estimating stock return volatility. Previous research has demonstrated their effectiveness, as evidenced by studies such as Tripathi & Chaudhary (2016) and Caiado (2004). However, the Autoregressive Conditional Heteroskedasticity (ARCH) family of models has its limitations in forecasting volatility over extended periods, as evidenced by Foster & Nelson (2024). Therefore, I propose to solve this by employing a pioneering Environmental, Social, and Governance (ESG) TARCH methodology. The objective is to incorporate ESG metrics into a model that enhances the accuracy of stock price forecasts and better captures market fluctuations over time. This research not only addresses potential improvements in volatility modeling but also aims to elucidate the link between ESG factors and market volatility, thereby offering valuable insights for investors and policymakers as they navigate the evolving landscape of sustainable finance.

Recent studies by Capelli et al. (2021) find that integrating ESG risk measures with traditional financial risk measures improves volatility forecasting accuracy. Other studies have established that ESG factors reduce stock fluctuations, implying significant impact on volatility (Li & Zhang, 2021). These findings underscore the growing importance of incorporating ESG considerations into financial models, as they not only enhance predictive accuracy but also contribute to more stable and resilient investment portfolios.

Because the ARCH family of models have not incorporated ESG related variables, this study addresses a gap not yet explored by prior studies (Tripathi & Chaudhary, 2016; Caiado, 2004). While Capelli et al. (2021) have established the augmented forecasting efficacy of ESG risk measures in traditional volatility models, these considerations have not been directly applied to TARCH models. Such models have proven their precision in reflecting the asymmetric effects of news on market volatility, due to their design that allows for volatility clustering and accounting for the impact of both positive and negative news (Alberg et al., 2008). This research thereby presents a novel perspective and seeks to substantiate the efficacy of TARCH models in capturing the unique information that ESG factors capture in the context of stock volatility forecasting. Thus, as we consider the intricate dynamics between ESG factors and stock volatility, we are compelled to ask:

"How does the incorporation of ESG factors into TARCH models affect their predictive accuracy in forecasting stock market volatility?"

This study conducts a time series analysis, comparing a conventional TARCH model with an ESG-enhanced variant, focusing on the representation of ESG factors to evaluate their impact on stock market volatility. Using data from 2016 to 2024 collected from Refinitiv Eikon, we calibrate both models with historical data from 2016 to 2019 and apply them to forecast volatility from 2020 to 2024. ESG proxies from Refinitiv Eikon are transformed into categorical variables based on standardized thresholds and embedded into the TARCH model, creating the ESG TARCH model. The effectiveness of this model is tested against the standard model by comparing their Root Mean Squared Error (RMSE) and using the likelihood-ratio (LR) and Diebold-Mariano (DM) statistical test, highlighting the predictive value of ESG factors in financial market analysis.

The results found that integrating ESG factors into TARCH models significantly improved predictive accuracy for S&P 500 index volatility. The ESG-enhanced TARCH model outperformed the conventional TARCH model, evidenced by higher log-likelihood values, lower RMSE, and favorable LR and DM test results, indicating ESG factors provide valuable information for predicting market fluctuations.

The paper is structured as follows. Section 2 discusses relevant literature on volatility forecasting models and ESG factors. Section 3 introduces the dataset. Section 4 covers the empirical methodology, including model selection, specification and test introductions. Section 5 presents the main results, including a comparison of the models' predictive performance, and discusses these findings with previous literature. Section 6 provides a summary and conclusion, highlighting implications for financial investors and analysts. Additional supportive materials are provided in the Appendix.

CHAPTER 2 Theoretical Framework

2.1 Volatility Forecasting Models

It is widely understood that volatility forecasting in financial markets has come to play a fundamental role in modern finance, offering critical insights for managing risk, pricing options, and guiding strategic investment decisions. Various models have been developed to estimate and predict volatility, each with distinct strengths and limitations. Historical volatility models use the standard deviation of past daily returns to estimate future volatility, providing a simple yet sometimes lagging measure. Implied volatility models reflect market expectations, and adapt quickly to latest information (Koopman et al., 2005). Realized volatility models, which compute volatility from high-frequency intraday returns, capture actual price variability within a trading day, offering timely market sentiment.

SV models provide a more nuanced depiction by treating volatility as a latent continuous process that evolves over time. These models, such as those explored by Andersen & Benzoni (2009), capture the mean-reverting nature of volatility and accommodate sudden jumps and long-range dependencies observed in market data. According to Andersen & Benzoni, “There are two main advantages to focusing on SV models. First, much asset pricing theory is built on continuous-time models. Within this class, SV models tend to fit more naturally with a wide array of applications, including the pricing of currencies, options, and other derivatives, as well as the modeling of the term structure of interest rates" (p. 2). SV models, including the Heston model (1993) and various jump diffusion and multiscale models, offer robust fits by incorporating both continuous price components and discrete jumps. They effectively model the dynamics of asset returns, including sudden large movements in market prices, although they are more complex and need more computationally intensive methods (Raggi & Bordigon, 2006).

2.1.1 ARCH Models in Volatility Forecasting

The previously mentioned models all have their own distinctions, which also counts for ARCH models. ARCH models are widely used to forecast volatility in time series data. Initially introduced by Engle (1982), ARCH models revolutionized the understanding and modeling of time-varying volatility in that time, a feature that often characterizes financial returns. The fundamental premise of ARCH models is that current volatility is conditional on past squared returns. Specifically, these models assume that volatility at a given time point is a function of past error terms, thereby capturing the clustering of volatility often observed in financial markets. This characteristic enables ARCH models to effectively model periods of high and low volatility, which are often sequentially correlated in financial time series data.

The uniqueness of ARCH models lies in their simplicity and practicality for short-term volatility forecasting. Compared to more complex models, such as multifactor loglinear or affine-jump models, ARCH models are easier to implement and less demanding for the need for advanced estimation methods (Chernov et al., 2003). This practicality makes them highly suitable for environments where computational resources or expertise may be limited. ARCH models are particularly adept at estimating high-frequency conditional variances, capturing short-term volatility accurately. Their straightforward structure allows practitioners to quickly estimate and forecast volatility, facilitating timely financial decision-making.

However, despite their advantages, ARCH models have notable limitations compared to other volatility models. One significant drawback is their performance over medium and long-term forecasting horizons. As highlighted by Foster & Nelson (1994), ARCH models tend to perform poorly for medium and long-term forecasts due to their inherent assumption of volatility persistence based solely on past errors. This limitation is crucial as it indicates that while ARCH models are useful for short-term predictions, they may not reliably capture the evolving nature of volatility over longer periods.

One of the common perspectives states that capturing volatility accurately over extended periods requires the necessary assumption that volatility is not only dependent on its past errors and variance, but also needs the integration of unpredictable jumps. In this context, SV models offer a more nuanced depiction of volatility by treating it as a latent random process that evolves continuously over time. "Unlike ARCH-type models, SV models specify volatility as a separate random process, which provides certain advantages over the ARCH-type models for modeling the dynamics of asset returns" (Yu, p. 474). This approach provides potentially a higher forecasting accuracy by better capturing the continuous randomness of volatility, making SV models appropriate for longer horizons. Nonetheless, SV models also face several disadvantages. The latent nature of SV models adds complexity to both specification and estimation, requiring advanced statistical techniques (Raggi & Bordigon, 2006). Marginalizing latent volatilities to evaluate the likelihood complicates model estimation, especially when analytic expressions for transition densities or moments are unavailable (Andersen & Benzoni, 2008).

2.1.2 Extending ARCH Models

Extending the basic ARCH models by incorporating additional features can address some of their limitations and enhance their ability to capture volatility more dynamically. One of the most notable extensions is the Generalized ARCH (GARCH) model, introduced by Bollerslev (1986). The GARCH model extends the basic ARCH model by including lagged conditional variances in addition to lagged squared residuals. By doing this, it can better capture the persistence of volatility over time, providing a more accurate fit to financial time series data. Bollerslev states that “In this light it seems that not only does the GARCH (1,1) model provide a slightly better fit than the ARCH (8) model in Engle & Kraft (1983), but it also exhibits a more reasonable lag structure" (p. 322). Despite the improved accuracy of GARCH models, they are limited in addressing asymmetry in financial markets, known as the leverage effect. Zakoian (1991) introduced the TARCH model to account for this asymmetry, by differentiating between positive and negative shocks to volatility. It does this by applying an additional threshold variable that differs on impact on volatility depending on whether the shock is positive or negative, thus capturing the leverage effect more effectively. Zakoian found significant results for the threshold coefficients, capturing the real-world behavior of volatility better than symmetric models.

Furthermore, Nelson (1991) proposed the Exponential GARCH (EGARCH) to address asymmetry more directly. Unlike TARCH models, EGARCH models use logarithmic transformation. This ensures that the model always predicts a positive variance and allows it to handle the asymmetry of shocks in a more continuous way. Nelson found that the EGARCH model provided superior predictive accuracy for financial markets sensitive to news.

More recent approaches state that incorporating exogenous variables into GARCH models can enhance forecasting accuracy. The GARCH with exogenous variables (GARCH-X) framework integrates exogenous variables, such as economic indicators, into the volatility equation. This offers a more holistic understanding of the market dynamics of stocks. For instance, Kambouroudis & Mcmillan (2016) introduced a GARCH-type model incorporating exogenous variables, which improved volatility forecasting accuracy compared to the model excluding these variables. Considering this research, exogenous variables indicating long-term financial performance could potentially address ARCH-type models' limitations in predicting medium to long-term volatility (Nelson & Foster, 1994). Traditional ARCH models focus on short-term volatility patterns, struggling with long-term forecasts. Therefore, integrating long-term exogenous variables enables GARCH-X models to capture volatility persistence over extended periods. Research indicates that including variables such as housing starts, default spread, and realized volatility in GARCH-X models, like GARCH-MIDAS, significantly improves their ability to predict long-term stock market volatility (Fang et al., 2020).

2.2 ESG Scores & Their Impact

Sustainable development emphasizes aligning corporate activities with social and environmental goals, promoting long-term ecological balance and social equity. Within this framework, ESG factors have emerged as essential non-financial metrics, offering insights into a corporation's sustainability and ethical impact. ESG factors encompass a broad range of considerations that directly influence corporate behavior and performance. ESG factors have risen importance in assessing a firm's commitment to sustainable practice in recent years. Given the 3 components ESG obtains, "The environmental concerns of investors and stakeholders, for instance, are natural environment protection, climate change, and environmental impacts arising from a business operation. The social factors important to stakeholders are human rights, equality, diversity in the workplace, and contribution to society. Some of the concerned governance issues are ownership structure, board independence, equitable treatment of shareholders, minority shareholders' rights, transparency, and disclosure of corporate information" (Atan et al. 2018, p. 183). ESG scores therefore provide a multidimensional spectrum of operational components for the corporate world.

Kim & Li (2021) highlight that ESG factors are crucial in evaluating corporate profitability, particularly in firms with previously weak governance structures. This impact is particularly strong in larger companies, where sustainability issues can significantly affect operations. Friede et al. (2015) observe a positive correlation between integrating ESG into financial analysis and financial performance, particularly in markets sensitive to sustainability.

Due to its importance for corporate profitability and financial performance, investors have gained interest in publications in ESG and Corporate Social Responsibility (CSR), a subset of the ESG disclosures. Dhaliwal et al. (2011) illustrate how "Firms with CSR performance superior to that of their industry peers enjoy a reduction in the cost of equity capital after they initiate CSR reports. Further, firms initiating CSR disclosure with superior CSR performance attract dedicated institutional investors and analyst coverage"(p. 94). In this context, Lo & Kwan (2017) emphasize that firms with higher ESG scores attract more investors due to their alignment with sustainable development goals. Their work supports the idea that companies disclosing non-financial data like ESG scores are perceived to have a greater commitment to sustainability. Moreover, the transparency of ESG reporting has been found to enhance confidence in corporate governance structures. This strategy within the ESG landscape is also highlighted by Garcia-Sanchez et al. (2020), who discuss how adopting comprehensive ESG practices aligns corporate goals with sustainable development, thereby ensuring long-term value creation for stakeholders. They argue that CSR has become a crucial strategic decision, particularly in terms of environmental, social, and economic issues, and that this transparency can improve relationships with stakeholders and boost overall corporate performance.

2.2.1 *Weight of ESG Components*

The weighting of the three components in ESG scores is a complex process that significantly influences the overall assessment of a company's sustainability performance. The relative importance of each ESG component can vary depending on industry-specific risks, stakeholder priorities, and rating agency methodologies (Berg et al. 2022).

Environmental scores are recognized as arguably the most critical component within ESG ratings due to 1) the ecological impact of firms and 2) their substantial impact on financial performance and risk management. Research suggests that companies with higher environmental scores tend to exhibit better long-term growth prospects and financial stability. For instance, regulatory pressure for companies to prioritize environmental sustainability has been shown to significantly impact their ESG performance, as noted by Yan et al. (2022). Similarly, research on Indian firms identified Resource Use and Environmental Innovation scores as critical indicators of overall ESG performance (Rajesh & Rajendran, 2020).

Governance and social scores also play relevant roles. Governance indicators, reflecting event risks like fraud, are significant in the short term (Giese et al., 2020). Social factors impact corporate credit ratings, with social variables being important predictors of credit risk and financial stability. Giese et al. also state that some frameworks propose dynamic weighting schemes that adjust over time and across different market conditions to better capture the evolving nature of ESG risks and opportunities. This adaptive approach allows for more responsive and accurate ESG assessments, aligning scoring methodologies with real-world impacts and strategic goals. Given these insights, the weighting of ESG components should imply to favor environmental scores while balancing governance and social scores for the most reliable ESG score, although it is also important to take the unique industry-specific risks and stakeholder perspectives into consideration to ensure the right weights for a representative ESG score.

2.2.2 *ESG & Volatility*

Incorporating ESG factors into volatility models has garnered increased attention in financial econometrics, driven by the growing acknowledgement of their potential to moderate market fluctuations and influence investor perceptions and behaviors. Xu (2023) underscores a significant relationship where enhanced ESG performance correlates with reduced stock price volatility, indirectly indicating that these factors should be integrated as predictive variables in volatility models. This finding aligns with the observations of Lo & Kwan (2017), who report that different ESG initiatives distinctly impact stock values, highlighting how these initiatives affect investor responses and subsequently, market volatility patterns. Li & Zhang. (2021) extend this perspective by providing empirical evidence from the Chinese A-share market, where ESG factors significantly reduce both idiosyncratic and extreme stock price risks, reinforcing the notion that these factors are critical for financial analysis within TARCH models that aim to accommodate the impact of ESG on reducing negative volatility spikes. The synthesis of these insights promotes a robust theoretical foundation for employing ESG factors into volatility models, advocating for their systemic inclusion to reflect market behaviors. The integration could potentially not only enhance the predictive accuracy of financial models but also align with the global movements towards sustainability, emphasizing the strategic importance of incorporating ESG considerations, with the ongoing transition of shifting paradigms in investor priorities and regulatory landscapes.

2.2.3 Predictive Accuracy of ESG Factors in TARCH Models

In light of the previous paragraphs, integrating ESG factors into the family of ARCH models provides a novel approach to predicting stock market volatility. In addition, Akerlof (1970) introduced the concept of information asymmetry, which emphasizes how the lack of complete information between buyers and sellers can lead to market instability. Within the context of ESG, comprehensive ESG disclosures can bridge this information gap, reducing uncertainty and enhancing market stability. Dhaliwal et al. (2011) emphasized that transparent ESG reporting increases market value and mitigates risks, offering insights into sustainable development goals. By incorporating ESG factors into ARCH models, this research seeks to find a potential stabilizer in reducing information asymmetry in predicting stock volatility, as well as gaining long-term predictive insights to fill up ARCH’s limitations.

Similarly, according to the Singal Theory, Spence (1973) proposes that signals such as educational qualifications in the labor market could convey valuable information about an individual's abilities. In financial markets, ESG scores could act as a signal of management quality and corporate sustainability practices. Lo & Kwan (2017) support this theory by demonstrating that higher ESG scores attract more investors, indicating superior management and commitment to sustainability. By embedding these signals into TARCH models, this research aims to reflect a firm’s risk more accurately, allowing for improved predictions of market volatility.

Although direct relationships between ESG factors and forecasting of volatility remain limited in the current literature, recent studies indicate a growing interest in integrating these factors into volatility models. Capelli et al. (2021) proposes a hybrid ESG volatility model to capture the influence of ESG risks on volatility, providing a basis for a new ESG enhanced TARCH model. By integrating ESG factors as unknown predictors within the TARCH framework, this research aims to refine the prediction of S&P 500 market volatility. ESG factors, with their focus on long-term objectives and sustainability, could add predictive power by addressing the long-term shortcomings of regular ARCH models. This integration fulfills the need for capturing longer-term volatility dynamics, thus potentially enhancing the predictive accuracy of the models. I hypothesize that incorporating ESG scores into TARCH models will reduce predictive errors and improve forecast accuracy. Therefore, the research hypotheses are as follows:

Ho: Incorporating ESG factors into TARCH models does not significantly enhance their predictive accuracy in forecasting the volatility of the S&P 500.

H1: Incorporating ESG factors into TARCH models significantly enhances their predictive accuracy in forecasting the volatility of the S&P 500.

Drawing on the theoretical underpinnings of Information Asymmetry and Signaling Theory, this research anticipates that reducing information asymmetry through comprehensive ESG disclosures will enrich forecasting market volatility and that ESG scores will serve as effective signals of management quality, leading to improved predictive accuracy of volatility models. As we explore the intricate dynamics between ESG and stock volatility, this study aims to clarify their predictive potential and offer empirical insights for investors and financial analysts navigating the evolving landscape of sustainable finance.

CHAPTER 3 Data

3.1 S&P 500 Variable

The Daily Close Price Index represents the closing prices of the S&P 500 Composite Index on each trading day. This measure is critical for assessing the daily performance and stability of the S&P 500, an index that reflects the economic health of 500 major U.S. companies. For this analysis, the dataset spans from 2016 to 2024, containing precisely 2,131 datapoints. The values are taken directly from the market's final recorded price at the close of trading, accurately reflecting the S&P 500's valuation at the end of the day. The data for this study is sourced from Refinitiv Eikon. The dataset not only captures daily closing prices but also includes various metrics for the market analysis extending from 1963 to 2024.

Table 3.1: Summary statistics of the daily closing price index, greenhouse gas emissions, Gender Diversity Index and Asset Management Index of the S&P 500,

Variable	Mean	Standard deviation	Minimum	Maximum
Price index	3364,39	875,05	1979,26	5254,35
Greenhouse gas emissions	363.53	336.81	83.89	818.45
Gender Diversity Index	78,81	12,56	51,2	108
Asset Management Index	260,68	47,24	147,91	388,99

The price index has exhibited an upward trend from 2016 till 2024, indicating overall growth for the companies included in the S&P 500. The variance in the price index varies with periods of high volatility, followed by periods of low volatility. Because the price index faces an upward trend, the mean does not stay constant over time.

3.2 ESG Variable

The ESG variable provides a comprehensive measure of environmental, social and governance factors that can influence the performance and risk profile of companies within the S&P 500 Composite index. For this study the ESG factors are represented by the following proxies: Greenhouse Gas Emissions (environmental score), Gender Diversity Index (social score), and Asset Management Index (governance score). Each proxy has been selected due to its significant impact on the corresponding ESG pillar. The dataset plans from 2016 till 2024, containing 2,131 datapoints, will be utilized in a hybrid ESG TARCH model to examine the relationship between ESG factors and market risk, providing insights into their impact on the conditional volatility.

3.2.1 Environmental Proxy

Greenhouse gas emissions are used as a proxy for the environmental score due to their ecological impact (Eng et al., 2021). Companies involved in gas emissions tend to have a significant carbon footprint, and thus a higher score indicates a more substantial negative environmental impact. The data for this analysis is sourced from Refinitiv Eikon and provides a measure of environmental impact of gas emissions from the S&P 500 companies, encompassing 2,131 daily observations from 2016 to 2024.

The data reveals significant insights about the summary statistics, particularly the large standard deviation of 363.53, which indicates high variability. The data reveals significant insights into the summary statistics, particularly the large standard deviation of 336.81, which indicates high variability. This high standard deviation suggests that there is considerable dispersion in greenhouse gas emissions among the S&P 500 companies, reflecting a wide range of environmental impacts. This variability could be attributed to differences in industry sectors, respective environmental policies and practices, and times of economic uncertainty (Bradley & Stumper, 2021).

3.2.2 Social Proxy

Gender diversity is utilized as a proxy for the social score because it mirrors the representation of women in corporate management and board positions. This data, sourced from Refinitiv Eikon, reflects the representation of women in pivotal roles within corporations. The data is interpolated to fill gaps for more consistent time series analysis. The interpolation process involves estimating missing data points within the observed values, ensuring a smooth sequence over the period of study. The data comprises 2,131 daily observations from 2016 and 2024. The Gender Diversity Index has exhibited an upward trend from 2016 to 2024, demonstrating a gradual increase in gender diversity over time. This positive trajectory is reflected in the data, which, despite some fluctuations, shows a consistent rise in the index values. The average Gender Diversity Index over this period is approximately 78.81, with a standard deviation of 12.56, indicating low variability. This low standard deviation can be attributed to the quarterly nature of the data, which smooths out short-term fluctuations and highlights the gradual changes in gender diversity. For instance, starting from a value of 59.41 on March 8, 2016, the index steadily increases, reaching a maximum of 108 by November 2023. The lowest recorded value during this period is 51.2.

3.2.3 Governance Proxy

Asset Management Index is used as a proxy for the governance score. The data is also sourced from Refinitiv Eikon. The Asset Management Index is a good proxy for governance as it includes key practices such as strategic investment policy setting, accountability through meaningful reporting, and risk mitigation, all of which are core governance elements (Johnson-Calari & Strauss-Kahn, 2020). Additionally, initiatives like holistic change management and stakeholder communication strategies emphasize the governance-focused nature of asset management operations. The Asset Management Index of the S&P 500 has displayed a range of values from a minimum of 147.91 to a maximum of 388.99, with an average value of 260.68 and a standard deviation of 47.24. This variation indicates a significant fluctuation in performance, reflecting the diverse financial conditions and management efficiencies across the companies in the index. One interesting aspect is the substantial peak at 388.99, suggesting periods of exceptionally high performance.

3.3 Control Variables

In the analysis of financial market volatility, particularly when examining the impact of ESG factors, the inclusion of appropriate control variables is crucial to ensure the robustness and validity of the results. Gross domestic product (GDP), inflation rates, and interest rates are three fundamental macroeconomic variables that serve as essential control variables in volatility models with ESG as regressors. These variables capture significant aspects of the economic environment that directly and indirectly influence market volatility, thereby allowing for a more accurate assessment of the specific contributions of ESG factors (detailed summary statistics of the control variables are provided in appendix B).

GDP growth is used as control variable due to its comprehensive reflection of the overall economic health and productivity of a nation, closely linked to indices like the S&P 500. Amendola et al. (2019) extend the GARCH-MIDAS model to explore the impact of macroeconomic variables, including GDP, on S&P 500 volatility. Their findings underscore GDP growth as a critical determinant of S&P 500 volatility in GARCH models. Therefore, including GDP as a control variable allows researchers to isolate the effects of ESG factors from the broader economic trends influencing the S&P 500.

Inflation rates play a critical role in shaping market dynamics and must be controlled in volatility models. Pindyck (1983) discusses how inflation and its variance influence stock market performance by affecting the riskiness of capital investments and investor returns. Increases in the expected and actual variance of inflation contribute to higher market volatility by impacting the net real returns on both equity and bonds. Furthermore, the variance of firms' real gross marginal return on capital, which has increased over time, also plays a significant role in explaining market declines and volatility. This highlights the necessity of controlling for inflation to prevent misleading conclusions about the impact of other economic factors on market volatility.

Interest rates are another control variable due to their pervasive influence on economic activity and market conditions. Interest rates affect borrowing costs, investment decisions, and overall economic activity, all contributing to market volatility. Higher interest rates typically lead to higher borrowing costs, slowing economic growth and reducing corporate earnings, which in turn influences stock prices and market volatility. By including interest rates as a control variable, researchers can more accurately attribute changes in market volatility to ESG factors. Kang et al. (2020) also highlight the importance of interest rates in their study on oil futures volatility. Their research indicates that long-term futures price volatility is affected by interest rates, emphasizing that fluctuations in interest rates can directly influence market conditions and stability. This underscores the necessity of accounting for interest rates when examining the determinants of market volatility to ensure comprehensive and accurate analysis.

3.4 Benchmark Variable

The benchmark variable in this study is the Volatility Index (VIX), which is widely regarded as a gauge of market expectations for near-term volatility, often referred to as the "fear gauge" of the financial markets. The VIX measures the market's expectations for volatility over the coming 30 days, as derived from the prices of S&P 500 index options. This benchmark is critical for evaluating the performance of the TARCH models in forecasting volatility, as it provides an independent reference point for comparison. The VIX is commonly used as a benchmark in numerous studies to compare volatility models (Kambouroudis & McMillan, 2016; Goard and Mazur, 2013).

For this analysis, the VIX data was sourced from Yahoo Finance, spanning the period from March 8, 2016, to May 8, 2024, resulting in a total of 2,131 observations. These observations provide a comprehensive view of market sentiment and volatility over the period analysed.

The VIX data, recorded daily, captures significant fluctuations in market volatility, which can be attributed to various macroeconomic events, geopolitical developments, and market dynamics during the observed period. This benchmark variable will be used to compare the residuals of the forecasted volatilities produced by the two competing TARCH models (detailed summary statistics for the VIX are provided in Appendix A).

CHAPTER 4 Method

4.1 The TARCH Models

4.1.1 Model Transformation, Selection, and Criteria

To determine the number of lags in the TARCH model, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are used. The AIC and BIC balance model fit with model complexity, penalizing the inclusion of unnecessary parameters (Akaike, 1974; Schwarz, 1978). This methodology is justified as it ensures the selection of a model that is both parsimonious and well-fitting, avoiding overfitting and underfitting issues. Using AIC and BIC in model selection has been proven effective in numerous studies, including Hacker and Hatemi-J (2008), where the application of these criteria led to optimal lag-length choice in Value-at-risk models under various conditions, including the presence of ARCH effects. Detailed results for the selection of the number of lags using AIC and BIC can be found in Appendix A.

Furthermore, Section 3 showed that the price index, the ESG components and the control variables exhibited a changing mean over time, leading to non-stationarity. To achieve stationarity, a requirement that is crucial in time series models (Greunen et al., 2014), the price index, ESG variable, and control variables are detrended (Rupassara et al., 2023). This will ensure that the statistical properties of the series, such as the mean and variance, remain constant over time, thereby enhancing the accuracy and reliability of the time series models.

The primary focus of this study is on applying a hybrid TARCH model that incorporates ESG factors alongside a traditional TARCH model. The hybrid model captures the conditional volatility of the closing price of the S&P 500 Price Index while considering the influence of ESG proxies. Including the ESG variable into the TARCH model requires incorporating exogenous variables in the conditional volatility equation and including it in the mean equation. Ke & Hu (2011) demonstrate that incorporating multiple economic variables as exogenous predictors in a GARCH (1,1) model significantly improves volatility forecasting accuracy by providing additional information that enhances the model's performance.

4.1.2 Model Specification

The regular TARCH model consists of 3 equations: 1) the conditional mean, 2) the conditional error term, and 3) the conditional variance equation. The 3 equations in this study are specified as:

(1) $$priceindex = \alpha_0 + \beta 1 \cdot riskfreereturn + \beta 2 \cdot \inf l atinon + \beta 3 \cdot gdp + \varepsilon_t$$

(2) $$\sigma_t = \varepsilon_t z_t$$

(3) $$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma 1 \varepsilon_{t-1}^2 I_{(\varepsilon_{t-1<0})}$$

The environmental component is represented by the S&P 500 greenhouse gas emissions, with a logarithmic transformation applied to ensure that higher emissions correspond to lower scores. Specifically, the environmental component Et is defined as the logarithm of the greenhouse gas emissions. The social component is represented by the Gender Diversity Index St , which is also logarithmically transformed. The governance component Gt, denoted is represented by the Asset Management Index, with a logarithmic transformation applied to it as well. Applying logarithmic transformations to these variables stabilizes their variances, which is critical for improving the accuracy of time series forecasts and improving stationarity. Lütkepohl and Xu (2009) find that substantial forecasting improvements occur when the log transformation stabilizes the variance of the underlying series. By stabilizing the variances of the environmental, social, and governance components, the predictive accuracy for the TARCH model and ESG TARCH model will be enhanced. This approach ensures that the models effectively avoid issues of overfitting and underfitting.

As subsection 2.2.2 reveals, the environmental component of the ESG variable is often given more weight. Therefore, the three components are combined into the ESG score, where the created logarithmically transformed variables are given weights to construct the overall ESG variable:

$$(4) \qquad ESG_t = -0.4 \cdot E_t + 0.3 \cdot S_t + 0{,}3 \cdot G_t$$

Equations (1) and (2) are identical for the TARCH and ESG TARCH model. But, for the ESG TARCH model, the ESG variable is integrated into conditional variance equation, and presented by the following equation:

$$(5) \qquad \sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma 1 \varepsilon_{t-1}^2 I_{(\varepsilon_{t-1<0})} + \theta_1 \cdot ESG_t$$

4.1.3 Model Estimation

Both the TARCH and ESG TARCH model will be estimated by using observations from the 8th of March 2016 to the 31st of December 2019, containing 995 observations. The coefficients in equations (1-3) for the TARCH model and equations (1), (2), and (5) are estimated using maximum likelihood estimation (MLE). MLE is a method that estimates the parameters of a statistical model by maximizing the likelihood function, so the observed data is most probable under the model. This method is particularly suitable for time series data with non-constant variance, such as financial returns. The effectiveness of MLE in this context is heavily influenced by the iterative optimization process used to find the best-fit parameters. By employing optimization algorithms like Berndt-Hall-Hall-Hausman (BHHH) and Broyden-Fletcher-Goldfarb-Shanno (BFGS), the models can navigate the complex likelihood surface to identify the parameter values that maximize the likelihood function, ensuring that both the TARCH and ESG TARCH model reach the most accurate parameter estimates possible.

The process of optimizing the models involves using BHHH and BFGS to achieve convergence for estimating the parameters. The optimization process switches between these two methods to help the models find the right parameters more effectively. To ensure more precise estimations, the process applies stricter convergence criteria, akin to setting a higher standard for accuracy. The BHHH method is useful in this case because it simplifies the MLE process as it does not require complex third derivatives and guarantees convergence, making it suitable for complex models like GARCH (1,1) (Arnerić & Rozga, 2009). Also, the BFGS method helps by making small adjustments in a smart way to speed up the solution's finding. This method, known as gradient descent, helps the models to converge more quickly and with fewer iterations (Mokhtari & Ribeiro, 2014). The application of stringent convergence criteria ensures that the optimization process terminates correctly and provides accurate results. This method is supported by the convergence depth control strategy, which balances calculation accuracy and computation cost to avoid premature termination or excessive computation time (Wang et al., 2007).

Before comparing Both the TARCH and ESG TARCH model, the models are evaluated using the same robustness checks as presented in subsection 4.5, to ensure the reliability of the estimators. This is particularly important when conducting predictive accuracy tests and metrics to prevent invalid results and draw the wrong conclusions. If the results reveal that the estimators are not robust, the models will be estimated again using robust standard errors.

4.2 Predictive Performance

The models' performance for both the TARCH and ESG TARCH model is evaluated using an out-of-sample forecast from the models estimated in subsection 4.3. The out-of-sample is forecasted from the 1st of January 2020 to the 8th of May 2024. The models are compared by calculating their RSME and conducting LR and DM tests of the forecasted volatility to determine which model offers superior predictive accuracy. This approach is supported by Miletić & Milosevic (2019), who advocate the use of RMSE and DM tests as robust methods for evaluating and comparing the accuracy of volatility forecasting models. Additionally, an LR test is conducted to compare the models' predictive performance. This test provides a rigorous method for evaluating the relative merits of various nested ARCH models. Mckenzie & Mitchell (2002) demonstrated that the LR test is effective in identifying the best-fitting model, showing that the GARCH (1,1) model is preferred for symmetric responses and models like TARCH (1,1) with a leverage term are better for capturing asymmetries. The LR test compares the likelihoods of two models: the simpler TARCH model and the ESG TARCH model. This test evaluates whether the addition of the ESG score in the TARCH model significantly improves the model's fit. By calculating the likelihoods of both models, the test determines the probability of observing the given data under each model. The difference in these likelihoods is then assessed using a chi-squared distribution, with degrees of freedom equal to the difference in the number of parameters between the two models. If the test statistic is significant, it indicates that the ESG TARCH model provides a better fit to the data and thus predicts volatility more accurately than the TARCH model.

The RSME for each model is calculated to quantify the prediction accuracy of the models. RMSE is derived from the residuals, which are the differences between the actual observed values and the values of the conditional variance predicted by the model. For both the TARCH model and ESG TARCH model, the residuals are computed as the difference between the predicted conditional volatility and the CBOE Volatility Index (VIX), and their squared values are averaged and then square-rooted to obtain the RMSE. A lower RMSE value indicates a model with better predictive accuracy in forecasting volatility. By comparing the RMSE values of the two models, we can determine which model has a smaller average error and therefore predicts volatility more accurately.

The DM test specifically assesses whether there is a statistically significant difference in the predictive accuracy between the TARCH model and ESG TARCH model. This test evaluates Mean Absolute Error (MAE) of both models, calculated with the difference between the predicted conditional volatility and the VIX variable, and compares them to determine if one model consistently predicts volatility better than the other. In volatility modeling, the MAE is particularly relevant because it is less sensitive to outliers than RMSE, and hence, it offers a more robust measure of average forecast accuracy. The DM test considers the mean and variance of the difference in the forecasting errors, and if the result is significant, it implies that the ESG TARCH model offers improvements in prediction over the TARCH model.

4.3 Control Variables

To rigorously assess the impact of ESG factors on volatility forecasting, all control variables will be added to the TARCH model ESG TARCH model within the conditional volatility equation. This inclusion of control variables aims to evaluate whether the ESG variable is significant in the conditional volatility equation and whether the ESG TARCH model enhances predictive accuracy compared to the TARCH model, also accounting for the probability that the impact of ESG scores might be less pronounced when broader economic factors are considered. This approach ensures that the observed improvements in predictive accuracy are not merely artifacts of omitted variable bias, thereby strengthening the validity of the findings. Again, the mean equation and error term equation remain equations (1) and (2), but the control variables will be added to the conditional variance equation for both the TARCH and ESG TARCH model. Therefore, the conditional variance for the TARCH and ESG TARCH model is described by the following equations:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \varepsilon_{t-1}^2 I_{(\varepsilon_{t-1<0})} + \beta_1 \cdot riskfreereturn_t + \beta_2 \cdot \inf lation_t + \beta_3 \cdot gdp_t \tag{6}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \varepsilon_{t-1}^2 I_{(\varepsilon_{t-1<0})} + \theta \cdot ESG_t + \beta_1 \cdot riskfreereturn_t + \beta_2 \inf lation_t + \beta_3 \cdot gdp_t \tag{7}$$

4.4 Weight Variants

To explore the impact of varying the weight assigned to each ESG component on the performance of the ESG TARCH model, different weighting schemes will be tested and analysed. Specifically, three variants of the ESG score will be constructed by altering the weights assigned to the E, S, and G scores applied in equation (7).

Equation (7) is considered the baseline weight construction for the ESG TARCH model. The second variant will adjust the weights to -0.6 for Environmental, 0.2 for Social, and 0.2 for Governance components, provided in equation (11), placing a heavier emphasis on environmental factors. The third variant will increase the weight for Governance to 0.6 while reducing the Environmental weight to -0.2, keeping the social weight at 0.2 in equation (12) (detailed results of equation (11) and (12) are provided in Appendix C).

Equation (11) and (12) will be integrated into the ESG TARCH model, and their performance will be compared to the TARCH model, using AIC, BIC, LR test, RMSE and the DM test. This comparison will help determine which weighting scheme is the most effective framework for incorporating ESG factors.

4.5 Robustness Checks

For robustness checks, we will perform the Ljung-Box test, ARCH LM test, and analyse the residual plots for each model variant. The Ljung-Box test will be conducted to check for any remaining autocorrelation in the residuals, ensuring that the model adequately captures the time series dependencies. The ARCH LM test will be used to detect any remaining ARCH effects, verifying that the conditional heteroskedasticity has been appropriately modeled. Additionally, standardized residual plots for the entire dataset will be analysed to visually inspect the residuals for any patterns or clustering, further inspecting the model's effectiveness in capturing the overall volatility of the dataset. The robustness checks will first be employed to decide whether the models will need robust standard errors to secure reliable test results in subsection 5.2. Secondly the robustness checks will be analysed to evaluate the limitations of this research.

CHAPTER 5 Results & Discussion

5.1 Model Estimation & Criteria

In the interpretation of the TARCH model, the influence of the different regressors on the price index and the conditional volatility of the S&P 500 is systematically analyzed. Equation (1), as can be seen in subsection 4.2, incorporates the risk-free return, the inflation rate, and GDP as predictors of the price index from the S&P 500. All the 3 corresponding coefficients can be interpreted in the following way: A change of 1 unit of variable X, leads to a β change in the price index. Equation (2) describes the relationship between the conditional variance and the error term. Equation (2) indicates that the observed error is decomposed into the product of the conditional standard deviation and a standardized shock, which follows a standard normal distribution. In general, this means that the volatility σ at time t scales the random shock to produce the actual observed deviation from the expected mean. Equation (5) in the TARCH model incorporates past squared residuals and a threshold squared residual to model the conditional variance. The constant term for the whole conditional variance equation is accounted in the ARCH part in the TARCH model, and in the heteroskedasticity (HET) part for the ESG TARCH model. The inclusion of a threshold term in TARCH models, often represented by an indicator function, allows the models to capture the asymmetric impact of positive and negative shocks on volatility. The coefficients for ARCH and TARCH in equations (3) and (5) are interpreted in the following manner: A change of 1 unit of variable X, leads to an α change in the conditional variance. Based on the residual diagnostics presented in subsections 5.1.1 and 5.1.2, the standard errors are made robust to ensure reliable test results to compare the predictive power of both models. The estimations of the TARCH and ESG TARCH model are presented in Table 5.1

Table 5.13: Regression results for TARCH and ESG TARCH models predicting the closing price for the S&P 500 Index, with observations from 2016 till 2019. The estimators contain robust standard errors.

Variable	**TARCH**	**ESG TARCH**
priceindex		
riskfreereturn	642.44***	614.2***
	(11.90)	(13.62)
inflation	-396.22***	-357.0***
	(18.99)	(21.26)
gdp	-2.37***	10.10***
	(2.97)	(3.54)
_cons	1212.14***	1212.43***
	(23.09)	(27.60)
HET		
esg		7.468***
		(21.05)
_cons		4.22
		(0.20)
ARCH		
Larch	1.02***	0.99***
	(0.21)	(0.20)
Ltarch	-0.15	-0.43**
	(0.22)	(0.21)
_cons	258.90***	
	(37.60)	
N	995	995

Log likelihood	-6370.97	-6339.25
Walchi2	27627.13	21140.39

*Note: The numbers with parentheses represent the estimated coefficients, and the numbers between brackets the standard errors. The significance levels are indicated with the following parentheses: *** p < 0.01, ** p < 0.05, and * p < 0.10. Significant results indicate strong relationships for risk-free return, inflation, and GDP in both models. Additionally, ESG factors significantly impact the price index and conditional variance in The ESG TARCH model. Both models show significant baseline effects on volatility.*

In the context of the ESG TARCH model, the additional ESG variable is included, seen in equation (3). This inclusion allows for the assessment of how ESG factors impact conditional volatility, providing insights into how sustainable practices and governance affect financial stability. The HET term in the equation (3) is particularly important for including the exogenous ESG variable in the conditional volatility equation. The coefficient under the HET term indicates how variations in ESG scores impact the conditional variance. For instance, a 1 unit change of ESG causes the price index to change by θ.

5.1.1 The TARCH Model

To comprehensively evaluate the overall performance of the TARCH model, we consider several key metrics and diagnostic tests that highlight both its strengths and areas for improvement.

As can be seen in Table 5.1, the model's likelihood value serves as a primary indicator of how well the TARCH model fits the given data. This value provides a baseline for comparing the model's fit relative to others.

In terms of information criteria, which balance model fit with complexity, detailed results of AIC and the BIC are presented in Table 5.1. These criteria help determine the model's efficiency; lower values indicate a better model that achieves a good fit with fewer parameters.

Examining the coefficients within the equation (1) reveals significant relationships between key economic variables and the price index.. Specifically, the risk-free return coefficient and the inflation coefficient are highly significant, underscoring their strong influence on the price index. Conversely, the GDP coefficient is not significant, suggesting a more nuanced or weaker impact. The constant term also plays a significant role, indicating a baseline level effect on the price index.

Residual diagnostics point to areas where the TARCH model could have better overall performance. The Ljung-Box test, found in Table 5.2, reveals significant autocorrelation in the residuals. This indicates that the TARCH model does not fully capture all the dependencies in the data.

The examination of ARCH effects necessitates thorough analysis. The standardized residuals plot for the TARCH model, visualised in figure 5.1, offers several important observations. The residuals oscillate around zero, confirming the stationary time series. However, the residuals remain to show remaining patterns and clustered periods. This indicates that although the model captures some volatility clustering, it may not fully encompass all volatility dynamics. Additionally, the ARCH LM test results, detailed in Table 5.2, confirm that significant ARCH effects persist, underscoring that the TARCH model has not entirely encapsulated the volatility behaviors inherent in the dataset. The ARCH LM test, presented in Table 5.2 indicates that significant ARCH effects remain. This suggests that the TARCH model has not entirely captured the volatility clustering present in the data.

Figure 5.1: Standardized residuals of the mean equation for the TARCH model starting from March 8, 2016, tillMay 8, 2024.

‘

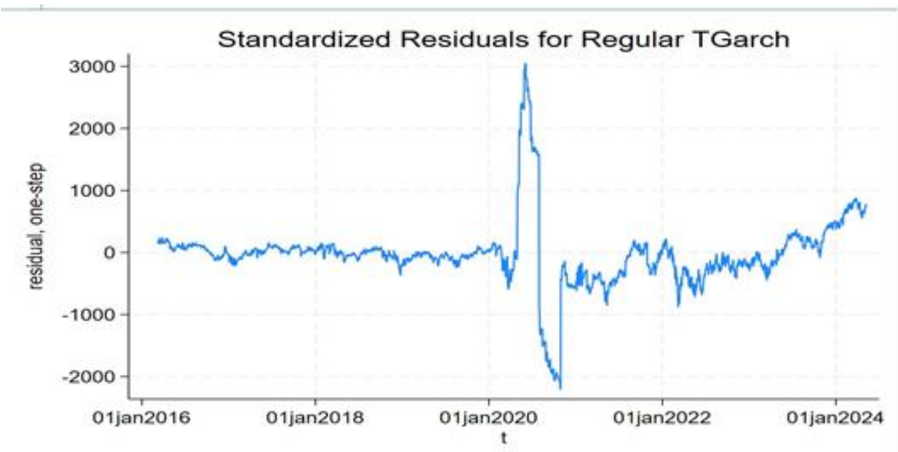

Table 5.2: Results for Ljung-Box, and ARCH-LM test for the estimated TARCH and ESG TARCH model from 2016 till 2019.

Test	TARCH		ESG TARCH	
	Test statistic	P-value	Test statistic	P-value
Ljung-Box (Q)	10610.12	0.00***	10578.54	0.00***
ARCH-LM	1696.29	0.00***	1695.35	0.00***

*Note: The Ljung-Box test statistic should be interpreted with a significant value indicating potential model inadequacies in capturing dependencies. The ARCH-LM test statistic is interpreted by p-values. In general, a p-value is interpreted as the probability of observing the test statistic under the null hypothesis, with a low p-value indicating statistical significance (*** p < 0.01, ** p < 0.05, and * p < 0.10) and suggesting that the null hypothesis can be rejected.*

Assessing the conditional variance dynamics of equation (3) in the TARCH model demonstrates persistence in volatility. The ARCH L1 coefficient is significant, indicating periods of high volatility are followed by high volatility and vice versa. However, the TARCH L1 coefficient is not significant, showing that model fails to account for leverage effects, where negative shocks impact volatility differently than positive shocks.

Interestingly, this inability to capture leverage effects persists despite the determination of the model's lag structure based on the AIC and BIC criteria. The AIC and BIC are specifically designed to identify model specification that provides the best balance between goodness of fit and model complexity. In this case, the chosen number of lags results in the lowest AIC and BIC values, suggesting an optimal model configuration. However, even with this optimal lag structure, the TARCH model fails to capture leverage effects, which are crucial for accurately modeling financial markets. This finding underscores a notable limitation: while the TARCH model is statistically optimal according to AIC and BIC, it still misses an important aspect of market dynamics. This highlights the complexity of financial data and suggests that additional factors or alternative model specifications might be necessary to fully capture the asymmetric impacts of shocks on volatility.

5.1.2 The ESG TARCH Model

Examining the ESG TARCH model can be analysed similarly to subsection 5.1.1. Starting with the likelihood value in Table 5.1, indicating its proficiency in aligning with the dataset. This metric stands as a testament to the model's fit, setting a benchmark for comparative analysis.

The model's information criteria in Appendix A— AIC and BIC—highlight its balance between complexity and performance. These values suggest a model that integrates these number of parameters without overfitting the most effectively.

Analysing the coefficients within the equation (1) for the ESG TARCH model uncovers significant relationships among various economic indicators and the price index. The coefficients for risk-free return, inflation rate, GDP, are profoundly significant. The constant term further substantiates a baseline impact on the price index, reinforcing the model's robustness.

Despite these strengths, residual diagnostics in this model also expose areas for refinement. The Ljung-Box test, detailed in Table 5.2 reveals significant autocorrelation in the residuals, suggesting model does not fully encapsulate all dependencies. The presence of ARCH effects warrants careful examination. The standardized residuals plot for the model (detailed visualisation provided in appendix D), reveals several critical insights. Notably, there are significant periods of elevated variance. These periods are marked by pronounced spikes in the residuals, indicating that the volatility is not constant over time. Therefore, the presence of discernible patterns and clusters indicates that the volatility dynamics may not be fully addressed by the model. Consequently, the ARCH LM test results, detailed in Table 5.2, confirm that significant ARCH effects persist, underscoring that the ESG TARCH model has not entirely encapsulated the volatility clustering inherent in the dataset.

In terms of conditional variance dynamics for equation (5), the ESG TARCH model excels with a significant ARCH L1 coefficient, detailed in Table 5.1. However, model stumbles with the TARCH L1 coefficient, failing to capture leverage effects where negative shocks impact volatility more than positive shocks.

Again, what is particularly intriguing is this shortcoming given the model's lag structure, determined through AIC and BIC criteria. These criteria aim to identify the optimal model configuration, balancing fit and complexity. Despite achieving the lowest AIC and BIC values, which suggest an ideal setup, the ESG TARCH model still falls short in capturing leverage effects. This anomaly highlights a significant limitation: even an "optimal" model according to statistical standards can miss crucial market dynamics. This realization invites further exploration into alternative specifications or additional variables that might bridge this gap.

Furthermore, the inclusion of control variables is examined to look for look for potential improvements in the model's explanatory power and predictive accuracy. Detailed results about the estimations are provided in Appendix A. Notably, the ESG variable remains significant despite the addition of 3 control variables, underscoring its robust impact on conditional volatility. This consistency affirms the crucial role of ESG factors in volatility modeling, regardless of the addition of other control variables. Also, while the log-likelihood improves slightly with the addition of these control variables, this enhancement is marginal. The limited significance of most control variables on conditional volatility suggests that these additional factors do not potentially influence the model's performance. This observation suggests that while the overall performance of model is not greatly enhanced by the additional control variables, the consistent significance of the ESG factor highlights its crucial role in affecting conditional volatility.

5.2 Predictive Performance

In comparing the predictive power of the TARCH model and ESG TARCH model, several statistical tests and evaluation criteria are utilized for these two models to determine which model predicts volatility more accurately. These tests are crucial in an econometric context because they offer insights into how well each model captures the inherent volatility in the data. The following figure illustrates the predicted variance by both models, providing a visual comparison to complement the statistical evaluation.

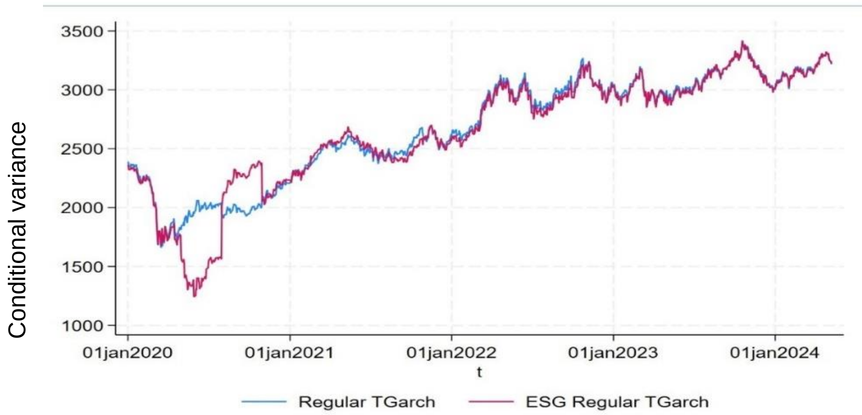

Figure 5.1: Predicted conditional variance of the TARCH and ESG TARCH model from January 1st, 2020, till the 8th of May 2024.

Firstly, the log-likelihood values for both models are considered. As detailed in Table 5.1, the log-likelihood value for the ESG TARCH model is higher than the TARCH model. A higher log-likelihood value indicates that model is a better fit for the data. This is because the log-likelihood measures the probability of observing the given data under the specified model. Therefore, a higher value suggests that model with ESG factors is more likely to capture the underlying volatility dynamics of the price index accurately.

The AIC and the BIC are also calculated for both models (detailed results provided in Appendix A). Both AIC and BIC penalize models for their complexity; hence, lower values indicate a model that achieves a good fit with fewer parameters. The lower AIC and BIC values for the ESG TARCH model indicate that it provides a more parsimonious representation of the data, which is critical in volatility modeling where overfitting can lead to poor out-of-sample predictions.

The LR test is employed to compare the goodness of fit between the two models directly. The test yielded significant results and rejects the null hypothesis that the TARCH model is as good as the ESG TARCH model. Therefore, the ESG TARCH model significantly improves the fit by capturing additional volatility patterns that the TARCH model misses. This improvement can be attributed to the inclusion of ESG factors, which likely contain relevant information for predicting volatility.

Table 5.3: Results for the LR test for the TARCH and ESG TARCH model from 2020 till 2024.

Statistic/Metric	**TARCH**	**ESG TARCH**	**Difference**
Loglikelihood	-6370.70	-6339.25	31.45
Degrees of freedom	7	8	1
LR test statistic (Q)		63.44	
P-value		0.00***	

*Note: parentheses *** p < 0.01, ** p < 0.05, and * p < 0.10*

In evaluating predictive accuracy more thoroughly, the RSME is calculated for both models. The RMSE for the ESG TARCH model is slightly lower than the TARCH model. To further address the forecasting performance and ensure strengthened proof, the DM test is conducted, focusing on the Mean Absolute Error (MAE) of the forecasts. The ESG TARCH model demonstrated a significantly lower MAE compared to the TARCH model. In combination with the RMSE and LR test, the much lower MAE for the ESG TARCH model shows that it provides more accurate volatility forecasts, as detailed in Table 5.4 and 5.5.

Table 5.4: RSME results for the residuals created by the actual volatility (VIX) and the predicted conditional volatility for the TARCH and ESG TARCH model from 2020 till 2024.

Model	**RMSE**
TARCH	1969.36
ESG TARCH	1955.48

Table 5.5: DM test for the residuals created by the actual volatility (VIX) and the predicted conditional volatility for the TARCH and ESG TARCH model from 2020 till 2024.

Statistic/Metric	TARCH	ESG TARCH	Difference
MAE	2661	2634	26.37
DM test statistic		2.681	
P-value		0.00***	

*Note: parentheses *** p < 0.01, ** p < 0.05, and * p < 0.10*

Additionally, predictive accuracy is also tested with the inclusion of control variables. As detailed in Table 5.6 and 5.7, including the ESG variable in model (6) improves its fit, with a higher log-likelihood. The LR test supports this improvement. The AIC and BIC values also show improvements, suggesting a more parsimonious model. (Details of the visualisation of the forecasted volatility from 2020 till 2024 for models (5) and (6) can be found in Appendix D). The RMSE for model (6) is lower than model (5). Similarly, the MAE is 2651 compared to 2654, with the DM test indicating significant difference in forecast accuracy. Therefore, predictive accuracy improves substantially. The inclusion of control variables indicates that the inclusion of ESG variables in the conditional variance enhances predictive power for the VIX while not suffering from omitted variable bias.

Given these findings, the hypothesis that incorporating ESG factors into TARCH models enhances the predictive accuracy in forecasting stock market volatility is supported. The ESG TARCH model showed a significantly better fit for the data, evidenced by higher log-likelihood values, lower AIC and BIC values, a lower RMSE and superior predictive accuracy as indicated by the DM test results. The null hypothesis is rejected, and the alternative hypothesis that ESG-enhanced TARCH models provide a more accurate forecast of stock market volatility is accepted. Therefore, the integration of ESG factors into TARCH models indeed improves their predictive performance in capturing stock market volatility.

Each model including the ESG variable showed significant improvements for the TARCH model with the data under consideration and utilized model specifications. However, based on these results, does this study state that the results proof that incorporating ESG factors improve volatility TARCH models universally? This assertion needs to be critically nuanced by the fact that the overall model does not perform in such a way that all-encompassing conclusions about forecasting accuracies can be made. As detailed in subsection 5.7, the overall TARCH model does not perform in a manner that the coefficients are reliable enough to draw definitive conclusions about the long-term effectiveness and general applicability of incorporating ESG factors in volatility forecasting. This implies that while the inclusion of ESG variables enhanced predictive power in the context of this study, the limitations and inconsistencies of the model necessitate further research and refinement. Nevertheless, the results of 5 different metrics (AIC & BIC, loglikelihood, LR test, RMSE and DM test) consistently suggest that the ESG TARCH model performs better than the TARCH model in terms of volatility forecasting, indicating that the incorporation is projected to provide meaningful improvement.

Table 5.6: AIC, BIC, RSME, and DM test results for the TARCH (C) and ESG TARCH (C), both including all control variables, with forecasted observations from 2020 till 2024.

Test/Metric	**TARCH (C)**	**ESG TARCH (C)**
AIC	12755.87	12689.66
BIC	12804.9	12743.59
RMSE	1964.61	1962.60
DM-test (MAE)	2654	2651***

*Note: The test statistic for the DM test is noted with parentheses *** p < 0.01, ** p < 0.05, and * p < 0.10. This means that Model (5) would get parentheses if the MAE is significantly lower, suggesting a significantly better forecast.*

Table 5.7: LR test for the TARCH vs. the ESG TARCH, both including control variables and tested with observations from 2020 till 2024.

Statistic/Metric	**TARCH (C)**	**ESG TARCH (C)**	**Difference**
Loglikelihood	-6367.94	-6333.83	44.11
Degrees of freedom	8	11	3
LR test statistic (Q)		68.21	
P-value		0.00**	

*Note: parentheses *** p < 0.01, ** p < 0.05, and * p < 0.10*

5.3 Weight Variants

When constructing ESG scores, various weight variants can be employed to observe their impact on model performance. For the second weight variant, the ESG scores are constructed through equation (11), assigning a heavier weight to the environmental component, as proposed by subsection 2.2.2. The coefficients for risk-free return, inflation, GDP in the mean equation are significant and can be seen in Appendix A. The constant term is also significant. Residual diagnostics (detailed results can be found in Appendix B). indicate significant autocorrelation with a significant Ljung-Box Portmanteau statistic. The ARCH LM showed persistent ARCH effects. The ARCH L1 coefficient shows strong volatility persistence, although the TARCH L1 coefficient remains insignificant. The log-likelihood, AIC, BIC, RSME and DM test are presented in Table 5.9 and 5.10. Comparing the forecasting accuracy of these weight variants with the TARCH model, it is evident that varying the weights impacts the model's predictive power. Weight variant 2, emphasizing the environmental factor, results in the lowest error metrics of all weight variants, suggesting that a heavier focus on environmental data introduces better predictions for volatility forecasts. Conversely, weight variant 3, which places more weight on governance, shows decreased predictive performance compared to the TARCH model, indicating that governance factors might provide more instable and unreliable signals for volatility prediction. These findings highlight that while the weighting scheme for ESG factors can significantly affect model outcomes. This insight underscores the importance of carefully selecting weightings based on the specific characteristics and stability of the dataset under consideration, aligning with the strategic goals of the analysis. Adjusting weights offers a nuanced approach to model optimization, demonstrating that even within the same modeling framework, fine-tuning can lead to more accurate and reliable forecasts.

Table 5.9: AIC, BIC, RSME, and DM test results from the forecasted period of 2020 till 2024 for the TARCH and ESG TARCH model with the different weight variants, computed by equation (11) and (12) for the ESG TARCH model.

Statistic/Metric	**TARCH**	**ESG TARCH**	**ESG TARCH (W2)**	**ESG TARCH (W3)**
AIC	12749.87	12694.5	12755.87	12687.87
BIC	12789.11	12733.72	12804.9	12741.8
RMSE	1969.36	1955.49	1954.94	1959.95
DM-test	2661	2634***	2634***	2637*

*Note: The 3rd column is the ESG TARCH model with the weight variant from equation (11), and the 4th column is the ESG TARCH model with equation (12). The test statistic for the DM test is noted with parentheses *** p < 0.01, ** p < 0.05, and * p < 0.10. This means that The ESG TARCH model, ESG TARCH(W2) and ESG TARCH(W3) would get parentheses if the MAE is significantly lower than the regular TARCH model, suggesting a significantly better forecast.*

Table 5.10: LR test for the ESG TARCH and the different weight variants for the ESG TARCH model, compared to the regular TARCH model, from 2020 till 2024.

Statistic/Metric	**TARCH**	**ESG TARCH**	**ESG TARCH (W2)**	**ESG TARCH (W3)**	**Difference**
Log-likelihood	-6370.97	- 6339.25	- 6341.26	-6360.70	
Degrees of freedom	7	8	8	8	1
LR test statistic (Q)	63.44	10.84	59.42	20.54	
P-value	0.00***	0.00***	0.00***	0.00***	

*Note: ESG TARCH (W2) is the ESG TARCH model with the weight variant from equation (11), and ESG TARCH (W3) is the ESG TARCH model with equation (12). ESG TARCH, ESG TARCH (W2) and ESG TARCH (W3) are compared to The TARCH model. Parentheses are *** p < 0.01, ** p < 0.05, and * p < 0.10.*

5.4 Robustness Checks

Ensuring the robustness of the models examined in previous sections involves a critical evaluation of several diagnostic tests. With the following tests, the validity and efficiency of the specifications and estimations of the ESG TARCH is inspected. The robustness checks can be found in Table 5.2.

Starting with residual diagnostics, the Ljung-Box test uncovers significant autocorrelation within the residuals, as evidenced by a significant Portmanteau statistic. This result indicates that the ESG TARCH model fails to fully encapsulate all time-series dependencies, suggesting that certain patterns remain unmodeled. This statistic reinforces the need for additional refinement to address these lingering dependencies. Moreover, the plot of the standardized residuals depicts spikes in the residuals, revealing that the TARCH models has not captured all volatility dynamics of the price index (detailed in Appendix D).

The ARCH LM test provides further evidence of the model's limitations. The test yields a substantial statistic, indicating that significant ARCH effects are still present. This finding points to the model's partial capture of volatility clustering, highlighting areas where it can be enhanced to better reflect the volatility patterns observed in the data.

Examining the conditional variance dynamics offers another layer of insight. The ESG TARCH model exhibits a significant ARCH L1 coefficient, demonstrating strong volatility persistence. This indicates that periods of high volatility are likely to be followed by continued high volatility, and similarly for low volatility periods. However, the model's inability to capture leverage effects is evident from the insignificance of the TARCH L1 coefficient. This suggests the ESG TARCH model does not adequately differentiate between negative and positive shocks' impacts on volatility.

This failure to capture leverage effects is particularly noteworthy given that the model's lag structure is determined using the AIC and BIC criteria, which are designed to balance model fit and complexity. Despite these criteria suggesting an optimal model configuration, the ESG TARCH model still misses this critical aspect of market behavior. This discrepancy highlights a fundamental limitation: even statistically optimal models according to traditional criteria may overlook essential dynamics such as asymmetric volatility responses.

Therefore, the ESG TARCH model does not fully meet all the assumptions needed for the most optimal assessment of 1) being classified as an accurate model 2) being classified as the better model compared to the TARCH model universally. Although the ESG TARCH model, creates a better fit using AIC and BIC, and outperforms the TARCH model in predictive accuracy, these results need to be interpreted with caution as the standardized residuals still show discernible patterns and the ESG TARCH model misses out on capturing all the ARCH effects.

To examine if there are better alternatives than the ESG TARCH model for incorporating asymmetric information and predicting volatility better, I compare the ESG TARCH model with various other ARCH models. In this way, I can see if the ESG TARCH model is the right model to incorporate ESG factors for better predictive accuracy, or if there are other models that could enhance the forecasting power even more and be a better fit for the dataset under consideration. Firstly, the EGARCH model (9), including the ESG factors in the volatility equation, is compared to the ESG TARCH model with the same control variables as in the other tests. When looking at Table 5.11, which conducts the same tests as the ones used in the comparison of the TARCH model and the ESG TARCH model (except the LR test, as this test cannot be performed since every model has the same number of degrees of freedom), the ESG TARCH model outperforms model (9). The same results emerged when comparing the ESG TARCH model with GARCH (10) and SAARCH (11) models, all of which included the ESG factors and control variables in their model. Estimations of the coefficients for model the EGARCH, GARCH, and SAARCH are provided in Appendix A.

Table 5.11: *AIC, BIC, RSME, and DM test results from the forecasted period of 2020 till 2024 for the ESG TARCH and the different ARCH-type models*

Statistic/Metric	**ESG TARCH**	**EGARCH**	**GARCH**	**SAARCH**
AIC	12694.5	12853.87	12697.87	12726.01
BIC	12733.72	12904.9	12741.8	12765.23
RMSE	1955.49	1985.83	1955.95	1977.46
DM-test	2635	2655	2698	2689

*Note: The test statistic for the DM test is noted with parentheses *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. This means that Model (9), (10) and (11) would get parentheses if the MAE is significantly lower than the regular TARCH model, suggesting a significantly better forecast.*

The comparisons highlight several key points about the ESG TARCH model. First, it demonstrates superior predictive power compared to other ARCH-type models when incorporating ESG factors, indicating it is particularly well-suited for scenarios where ESG considerations are crucial for volatility forecasting. Second, the model's consistent performance across various ARCH model comparisons suggests a level of resilience, maintaining robustness and reliability even when benchmarked against more complex or differently structured models. Third, while the ESG TARCH model is robust, diagnostics indicate areas for potential improvement, particularly in fully capturing leverage effects and residual autocorrelation. Finally, the comparative analysis validates the choice of the TARCH framework for incorporating ESG factors, suggesting it is the most effective structure among the models tested, thereby justifying its use despite the identified limitations.

5.5 Discussion

A compelling convergence in results and theoretical frameworks across various studies examines the integration of ESG factors into volatility forecasting models. Despite differences in methodologies and contexts, the consensus underscores the robustness of ESG factors in enhancing predictive accuracy and financial stability.

The main insight in this study proves that the ESG TARCH model outperforms the TARCH model in predicting volatility. This finding aligns with the research by Lo & Kwan (2017), which illustrates that companies incorporating ESG factors exhibit lower volatility and higher returns. Although the main purposes of the research differ, both studies emphasize that ESG factors are a crucial factor in explaining financial volatility.

Examining the findings with Guo et al. (2020), who introduced a deep learning framework to predict stock volatility from ESG news, reinforces the shared understanding of ESG factors' positive impact. Despite employing a different methodology, their findings align with this study. Their ‘ESG2Risk’ model showed significant improvements in predictive accuracy, like the enhancements observed in the ESG TARCH model. Guo et al’s study demonstrated that integrating ESG news reduced prediction errors and captured volatility dynamics more accurately. This convergence underscores the versatility and robustness of ESG factors in financial modeling, and complements the statistical approaches used in my study.

The findings are also consistent with the results of the analysis of Jakobsson & Lundberg (2018), which highlighted that higher ESG scores are significantly negatively related to both realized daily volatility and daily GARCH (1,1) estimates of volatility. The use several statistical tests, along with the statistically significant negative relationship found between high ESG scores and share price volatility, supports the robustness of ESG factors in financial modeling. These results indicate that higher ESG scores are associated with lower share price volatility, reinforcing the significance of incorporating ESG factors in predictive financial models.

Furthermore, the results of the weight variants aligned with my hypothesis and the discussed literature. Weight variant 2, emphasizing weight on the environmental factor, showed the lowest RSME and produced the most significant p-value in the DM test. Combining the findings from Giese et al. (2021), and Capelli et al. (2021), it can be stated that ESG factors influence forecasting accuracy significantly, and by giving more appropriate weights to the ESG variable, the predictive accuracy will be enhanced.

Despite that ESG factors strengthened the predictive power of volatility, the robustness checks in my study reveal limitations in both the TARCH and ESG TARCH models, particularly in their inability to fully capture volatility dynamics across the dataset under consideration. Linking these findings, the uniqueness of ARCH models lies in their simplicity and practicality for short-term volatility forecasting (2.1.2). However, as highlighted, these models face significant limitations in medium and long-term forecasting (Nelson & Foster, 1994).The results in 5.4 and 5.6 point out that the limitations could be nestled within the choice of using TARCH models for this dataset and timespan, underscoring the need for further refinement, alternative modeling approaches, or selection of the data and variables used to fully encapsulate the volatility patterns observed in financial markets.

Notably, other SV models are known for their ability to better capture long-term volatility dynamics (2.1). This could imply that incorporating such SV models might provide a more accurate and comprehensive understanding of the volatility behavior in datasets that exhibit long-term dependencies and structural breaks, applicable for the dataset used in this study. Adopting these advanced models could therefore address the limitations observed in the TARCH and ESG TARCH models, leading a more reliable presentation of the impact of ESG factors on volatility forecasting, and potentially improving forecasting accuracy in financial modeling.

Despite the evidence that ARCH-type models stay limited in performing well in forecasting long-term volatility, the overarching theme across the findings from both this study as other studies are the consistent with the findings from this study that suggest that ESG integration enhances forecasting performance, independent of what methodology is used. These empirical examples, such as Lo and Kwan's (2017) demonstration of reduced volatility and higher returns, or Guo et al.'s (2020) improved accuracy with ESG news, illustrate the multifaceted benefits of ESG integration. These converging results and theoretical alignments underscore the importance of incorporating ESG factors into financial models and provide a profound framework for improving predictive accuracy and enhancing financial stability. The dynamic interplay of methodologies and perspectives across these studies implies the multifaceted benefits of ESG integration, reinforcing the findings of the ESG TARCH model and highlighting the broader implications for financial modeling and investment strategies.

CHAPTER 6 Conclusion

In this thesis, I have examined the integration of ESG factors into TARCH models to enhance the predictive accuracy of stock market volatility forecasts. Previous research has demonstrated the potential of TARCH models in estimating stock return volatility, highlighting their utility in various financial markets. However, the inclusion of ESG factors as an unknown predictor in these models represents a new approach that aims to address both the limits of TARCH and the increasing demand for sustainable investment practices. This study's significance lies in its potential to fulfill the limitations of TARCH models compared to more advanced and complex volatility models and contribute to the evolving understanding of the role of ESG factors in financial risk analysis. Specifically, the study focuses on the S&P 500 price index, utilizing data from 2016 to 2024 to calibrate and compare the performance of a traditional TARCH model against those enhanced with ESG factors. This work not only responds to innovating financial econometric models but also seeks to clarify the ESG-volatility link, thus providing valuable insights for investors and financial analysts in the complexities of sustainable finance. Therefore, the research question that was studied in this dissertation was:

"How does the incorporation of ESG factors into TARCH models affect their predictive accuracy in forecasting stock market volatility?"

To answer this research question, I conducted a time series analysis using data from the S&P 500 index, spanning from 2016 to 2024. The analysis involved constructing a conventional TARCH model and an ESG-enhanced TARCH model, incorporating ESG factors as unknown predictors. ESG proxies such as greenhouse gas emissions, the Gender Diversity Index, and Asset Management Index were transformed into categorical variables based on standardized thresholds and embedded into the TARCH model. The models were calibrated using historical data from 2016 to 2019 and then applied to forecast volatility from 2020 to 2024. The performance of the models was evaluated by comparing their RMSE and conducting the LR and DM statistical test. The results indicated that the ESG-enhanced model demonstrated a significant improvement in predictive accuracy compared to the conventional TARCH model. The inclusion of ESG factors enhanced the model's ability to capture market fluctuations, thereby providing more reliable volatility forecasts.

However, the robustness checks revealed the need to acknowledge the limitations of this research. The Ljung-Box test, ARCH LM test, and analysis of the standardized residuals implied that the estimates could be biased, underscoring the need for further refinement of TARCH models in capturing volatility. Despite these limitations, the discussed literature and the results strongly indicated that the models (4 and 6) including ESG factors outperformed the conventional TARCH models (1 and 5) consistently across all tests, converging to a conclusion that favors the inclusion of ESG factors in TARCH models to predict stock price volatility more accurately.

Therefore, under the given dataset and model specifications, the integration of ESG factors into TARCH models enhances the predictive accuracy of stock market volatility forecasts, while considering the research's limitations. This finding is particularly relevant for investors and quantitative financial analysts who are increasingly focused on improving financial metrics and have gaining interest in sustainable finance. The results show that, in this research's context, ESG factors provide additional information that improves the ability of volatility models to capture market dynamics and investor behavior.

By incorporating ESG metrics such as greenhouse gas emissions, the Gender Diversity Index, and Asset Management Index into financial models, researchers and practitioners can achieve more accurate and reliable forecasts of market volatility. Nonetheless, future research should address the identified limitations and further validate these findings by ensuring that all relevant assumptions are optimally met.

Other researchers studying the intersection of ESG factors and financial volatility can learn from this thesis that the inclusion of ESG variables not only aligns with increasing interest in sustainibility but also offers practical benefits in terms of risk management and investment decision-making. The enhanced predictive power of ESG-enhanced TARCH models underscores the importance of considering sustainability metrics in financial analysis. This research contributes to the broader understanding of how sustainable practices impact financial markets, providing a foundation for further exploration into the role of ESG factors in financial risk assessment and modeling.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716-723.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics, 84*(3), 488-500.

Alberg, D., Shalit, H., & Yosef, R. (2008). Estimating stock market volatility using asymmetric GARCH models. *Applied Financial Economics, 18*(15), 1201-1208.

Amendola, A., Candila, V., & Gallo, G. (2019). On the asymmetric impact of macro–variables on volatility. *Economic Modelling*, *76*, 135-152.

Andersen, T. G., & Benzoni, L. (2008). *Stochastic volatility* (NBER Working Paper No. 14297). *National Bureau of Economic Research*.

Arnerić, J., & Rozga, A. (2009). Numerical Optimization within Vector of Parameters Estimation in Volatility Models. *Journal of Information and Communication Convergence Engineering, 3*(1), 112-116.

Atan, R., Alam, M. M., Said, J., & Zamri, M. (2018). The Impacts of Environmental, Social, and Governance Factors on Firm Performance: Panel Study of Malaysian Companies. *Management of Environmental Quality: An International Journal, 29*(2), 182-194.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307-327.

Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate Confusion: The Divergence of ESG Ratings. *Review of Finance*, *26*(6), 1315-1344.

Bradley & Stumpner. (2021). The impact of COVID-19 on capital markets, one year in. *McKinsey & Company.*

Caiado, J. (2004). Modelling and forecasting the volatility of the Portuguese stock Index PSI-20. *Portuguese Journal of Management Studies 9*(1), 67-87.

Campbell, J., & Vuolteenaho, T. (2004). Inflation illusion and stock prices. *NBER Working Paper Series, 94*(2), 19-23.

Capelli, P., Ielasi, F., & Russo, A. (2021). Forecasting volatility by integrating financial risk with environmental, social and governance risk. *Corporate Social Responsibility and Environmental Management*, *28*(5), 1483-1495.

Chernov, M., Gallant, A. R., Ghysels, E., & Tauchen, G. (2003). Alternative models for stock price dynamics. *Journal of Econometrics, 116*(1-2), 225-257

Dhaliwal, D.S., Li, O.Z. Tsang, A., & Jang Y.G. (2011). Voluntary Nonfinancial Disclosure and the Cos t of Equity Capital: The Initiation of Corporate Social Responsibility Reporting. *The Accounting Review, 86*(1), 59-100.

Eng, L., Fikru, M. G., & Vichitsarawong, T. (2021). Comparing the informativeness of sustainability disclosures versus ESG disclosure ratings. *Sustainability Accounting, Management and Policy Journal*, *13*(2), 494-518.

Engle, R.F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, *50*(4), 987-1007.

Engle, R. F., & Kraft, D. F. (1983). Multiperiod forecast error variances of inflation estimated from ARCH models. *Applied Time Series Analysis of Economic Data*, 293-307.

Fang, T., Lee, T.-H., & Su, Z. (2020). Predicting the long-term stock market volatility: A GARCH-MIDAS model with variable selection. *Journal of Empirical Finance, 58,* 36-49.

Foster, D. P., & Nelson, D. B. (1994). Continuous record asymptotics for rolling sample variance estimators. *Econometrica, 64*(1), 139-174.

Friede, G., Busch, T., & Bassen, A. (2015). ESG and Financial Performance: Aggregated Evidence From More than 2002 Empirical Studies. *Journal of Sustainable Finance & Investment, 5*(4), 210-233.

Garcia-Sanchez, I. M., Rodrigues-Dominguez, L., & Gallego-Alvarez, I. (2020). The role of the board of directors in the adoption of GDP guidelines for the disclosure of CSR information. *Journal of Cleaner Production, 141*, 737-750.

Giese, G., Lee, L.-E., Melas, D., Nagy, Z., & Nishikawa, L. (2020). Foundations of ESG Investing: How ESG Affects Equity Valuation, Risk, and Performance. *Journal of Portfolio Management, 45*(5), 69-83.

Goard, J., & Mazur, M. (2012). Stochastic volatility models and the pricing of VIX options. *Mathematical Finance*, *23*(3), 439-458

Greunen, J. van, Heymans, A., van Heerden, C., & Van Vuuren, G. V. (2014). The Prominence of Stationarity in Time Series Forecasting. *Studies in Economics and Econometrics*, *38*(1), 1-16.

Guo, T., Jamet, N., Betrix, V., Piquet, L.-A., & Hauptmann, E. (2020). ESG2Risk: A Deep Learning Framework from ESG News to Stock Volatility Prediction. *ERN: Stock Market Risk.*

Hacker, R. S., & Hatemi-J, A. (2008). Optimal lag-length choice in stable and unstable VAR models under situations of homoscedasticity and ARCH. *Journal of Applied Statistics, 35*(6), 601-615.

Heston, S. L. (1993). "A closed-form solution for options with stochastic volatility with applications to bond and currency options." *The Review of Financial Studies*, *6*(2), 327-343.

Jakobsson, R., & Lundberg, L. (2018). The Effect of ESG Performance on Share Price Volatility. *Environmental Science, Business, Economics.*

Johnson-Calari, J., & Strauss-Kahn, I. (2020). Good governance: Principles, pitfalls, and best practice. In J. Bjorheim (Ed.), *Asset management at central banks and monetary authorities,* 305-321.

Kambouroudis, D. S., & McMillan, D. (2016). Does VIX or volume improve GARCH volatility forecasts? *Applied Economics, 48*(13), 1210-1228.

Kang, B., Sklibosios Nikitopoulos, C., & Prokopczuk, M. (2020). Economic Determinants of Oil Futures Volatility: A Term Structure Perspective. *Energy Economics, 88,* 3-27.

Ke, T.-H., & Hu, Y.-P. (2011). Volatility forecasting with many predictors. *Journal of Forecasting, 32*(8), 743-754.

Kim, S., & Li, Z. (2021). Understanding the Impact of ESG Practices in Corporate Finance. *Sustainability, 13*(3746), 1-15.

Koopman, S. J., Jungbacker, B., & Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance, 12*(3), 445-475.

Li, Y., & Zhang, Y. (2021). Investor sentiment, idiosyncratic risk, and stock price premium: Evidence from Chinese cross-listed companies. *SAGE Open, 11*(2).

Lo, K. Y., & Kwan, C. L. (2017). The Effect of Environmental, Social, Governance and Sustainability Initiatives on Stock Value – Examining Market Response to Initiatives Undertaken by Listed Companies. *Corporate Social Responsibility and Environmental Management, 24*(6), 606-619.

Lütkepohl, H., & Xu, F. (2009). The role of the log transformation in forecasting economic variables. *Empirical Economics, 42*, 619-638.

McKenzie, M., & Mitchell, H. (2002). Generalized asymmetric power ARCH modelling of exchange rate volatility. *Applied Financial Economics, 12*(8), 555-564.

Miletić, S., & Milosevic, D. (2019). Modeling and forecasting exchange rate volatility in EEC countries. *Anali Poslovne Ekonomije, 6*(11), 1-17.

Mokhtari, A., & Ribeiro, A. (2014). Regularized stochastic BFGS algorithm. *IEEE Transactions on Signal Processing, 62*(23), 6089-6104.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. E*conometrica, 59*(2), 347-370.

Rajesh, R., & Rajendran, C. (2020). Relating Environmental, Social, and Governance scores and sustainability performances of firms: An empirical analysis. *Business Strategy and the Environment, 29*, 1247-1267.

Raggi, D., & Bordignon, S. (2006). Comparing stochastic volatility models through Monte Carlo simulations. *Computational Statistics & Data Analysis, 50*(7), 1678-1699.

Rupassara, U., Frantsvog, S., Holen, A., & Robinson, K. (2023). Analysis of the stationarity and correlation of the global temperature and carbon dioxide time series [version 1; peer review: awaiting peer review]. *F1000Research, 1*, 1-16

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*(2), 461-464.

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics, 87*(4), 355-374.

Tripathi, V., & Chaudhary, P. (2016). Estimating stock return volatility in Indian and Chinese stock market., *International Journal of Banking, Risk and Insurance, 4*(2), 36-49.

Wang, K., Shao, Z., Zhang, Z., Chen, Z., Fang, X., Zhou, Z., & Qian, J. (2007). Convergence depth control for process system optimization. *Industrial & Engineering Chemistry Research, 46*(23), 7729-7738.

Xu, X. (2023). Has ESG performance reduced stock price volatility. *Journal of Innovation and Development, 3*(1), 59-66.

Yan, Y., Cheng, Q., Huang, M., Lin, Q., & Lin, W. (2023). Government environmental regulation and corporate ESG performance: Evidence from natural resource accountability audits in China. *International Journal of Environmental Research and Public Health*, *20*(1), 1-16

Yu, J. (2012). A semiparametric stochastic volatility model. *Journal of Econometrics, 167*(2), 473-482.

Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control, 18*(5), 931-955.

APPENDIX

A

Table 5.12: AIC and BIC of the regular TARCH model, with different number of lags.

Model	AIC	BIC
ARCH(1) TARCH (1)	12755.94	12790.26
ARCH(2) TARCH (1)	13126.09	13160.41
ARCH(3) TARCH (1)	13180.80	13215.12
ARCH(1) TARCH (2)	12763.26	12797.58
ARCH(2) TARCH (2)	13228.39	13262.71
ARCH(3) TARCH (2)	13236.80	13291.12
ARCH(1) TARCH (3)	12763.20	12797.52
ARCH(2) TARCH (3)	13220.07	13254.39
ARCH(3) TARCH (3)	13237.66	13271.98

Note: The numbers between brackets present the number of lags the ARCH and TARCH components consist of in the given model.

Table 5.13: Results for AIC and BIC regarding the TARCH and ESG TARCH model estimations from March 8, 2016 till December 31, 2019.

Model	AIC	BIC
TARCH	12755.94	12790.26
ESG TARCH	12694.52	12733.72

Table 5.14: Estimation of the TARCH model and ESG TARCH model with both models including the variables the risk-free return, inflation rates, and GDP in the conditional volatility equation from March 8, 2016, till December 31, 2019. The estimators contain robust standard errors.

Variable	**TARCH (C)**	**ESG TARCH (C)**
priceindex		
riskfreereturn	627.0***	623.75***
	(15.40)	(13.25)
inflation	-370.11***	-366.05***
	(16.79)	(21.00)
gdp	5.75.***	-5.41
	(3.01)	(3.02)
_cons	1245.21***	1248.67***
	(43.92)	(25.92)
HET		
esg		8.00***
		(0.63)
riskfreereturn	1.80	0.67
	(1.22)	(0.97)
inflation	-1.94	-2.12
	(1.99)	(1.49)
gdp	0.59	0.70**
	(0.41)	(0.32)
_cons	0.46	1.33
	(21.08)	(2.12)
ARCH		
L.arch	1.01***	1.03***
	(4.92)	(0.20)
L.tarch	-0.02*	-0.480*
	(2.01)	(2.30)
N	995	995
Log likelihood	-6367.94	-6333.83
Waldchi2	24473.29	22778.96

*Note: The (C) stands for the inclusion of all the control variables into the volatility equation. The numbers with parentheses represent the estimated coefficients, and the numbers between brackets the standard errors. The significance levels are indicated with the following parentheses: *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$. Significant results indicate strong relationships for risk-free return, inflation, and GDP in both models.*

Table 5.15: Estimation results for TARCH, GARCH, EGARCH and SAARCH models, all including the ESG variable and control variables. The estimators contain robust standard errors.

Variable	**ESG TARCH (C)**	**ESG EGARCH**	**ESG GARCH**	**ESG SAARCH**
Priceindex				
Risk Free Return	623.7***	631.4***	667.1***	627.1***
	(47.08)	(18.39)	(23.37)	(40.92)
Inflation	-366.00***	-286.80***	-390.40***	-370.10***
	(17.42)	(5.54)	(18.19)	(15.06)
GDP	-5.41	27.54***	54.38***	-5.71
	(1.79)	(3.34)	(5.50)	(1.90)
Constant	1248.70***	1025.60***	942.90***	1245.10***
	(48.17)	(16.47)	(15.66)	(43.69)
HET				
Risk Free Return	0.67	1.55***	-0.53	1.79
	(0.07)	(0.36)	(0.08)	(0.47)
Inflation	-2.12	-3.51***	0.06**	-1.92
	(0.42)	(0.52)	(0.00)	(-0.17)
GDP	0.70*	0.33*	0.05	0.59
	(2.19)	(2.54)	(1.89)	(1.42)
esg	8.04***	3.33***	4.45***	4.26
	(1.69)	(0.47)	(1.01)	(2.22)
Constant	1.32	6.78***	9.07***	0.50
	(0.62)	(6.43)	(3.54)	(0.19)
ARCH				
L.arch	1.03***	0.00*	0.57***	1.00***
	(3.21)	(1.09)	(0.10)	(0.94)
L.saarch				-0.00
				(0.22)
L.egarch		-0.09		
		(.0.28)		
L.garch			-0.523	
			(-1.22)	
L.tarch	-0.48*			
	(-2.30)			

N	995	995	995	995
Log likelihood	-6333.828	-6561.659	-6439.7	-6297.509
Waldchi2	22778.96	1753.91	6411.22	21628.12

*Note: The numbers with parentheses represent the estimated coefficients, and the numbers between brackets the standard errors. The significance levels are indicated with the following parentheses: *** p < 0.01, ** p < 0.05, and * p < 0.10. Significant results indicate strong relationships for risk-free return, inflation, and GDP in both models.*

B

Table 5.16: Results for Ljung-Box, and ARCH-LM test for the estimated TARCH and ESG TARCH model from 2016 till 2019, with equations (8) and (9) applied for each model.

Test	**ESG TARCH (W2)**		**ESG TARCH (W3)**	
	Test statistic	P-value	Test statistic	P-value
Ljung-Box (Q)	10583.14	0.00***	10539.97	0.00***
ARCH-LM	1695.50	0.00***	1694.02	0.00***

*Note: The Ljung-Box test statistic should be interpreted with a significant value indicating potential model inadequacies in capturing dependencies. The ARCH-LM test statistic is interpreted by p-values. In general, a p-value is interpreted as the probability of observing the test statistic under the null hypothesis, with a low p-value indicating statistical significance (*** p < 0.01, ** p < 0.05, and * p < 0.10) and suggesting that the null hypothesis can be rejected.*

Table 5.17: Summary statistics for risk-free return, inflation rates, and GDP.

Variable	**Mean**	**Standard deviation**	**Minimum**	**Maximum**
Risk-free return	2,40	1,04	0,50	4,98
Inflation	0,38	0,90	-1,20	2,53
GDP	2,73	9,02	-55,00	33,90

C

(8) $ESGt = -0.6 \cdot Et + 0.2 \cdot St + 0.2 \cdot Gt$

(9) $ESGt = -0.2 \cdot Et + 0.2 \cdot St + 0.6 \cdot Gt$

D

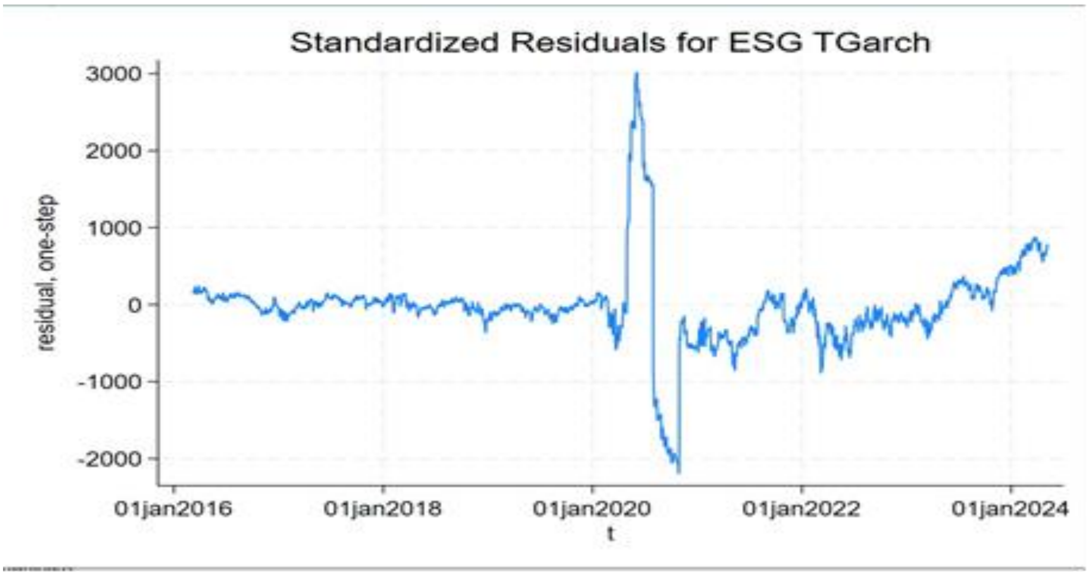

Figure 5.2: Standardized residuals of the mean equation for the ESG TARCH model starting from March 8, 2016, till May 8, 2024.

Figure 5.3: Predicted conditional variance of the TARCH and ESG TARCH model from January 1st 2020 till the 8th of May 2024, including both risk-free return, inflation, and GDP as control variables in the conditional volatility equation.