The Impact of ESG Performance on Idiosyncratic Volatility

A GARCH-MIDAS approach with AI-based ESG ratings

Sebastian Pickering Pedersen

Master's Essay *Data Analytics & Business Economics*

20-05-2024

Supervisor: Hossein Asgharian

Abstract

Environment, social and governance (ESG) ratings are of interest for both investors and researchers for its effect on stock performance. Previous literature has found mixed results with regard to this relationship and few papers are concerned with the volatility of ESG stocks. This paper explores the relationship between ESG scores and idiosyncratic risk. Using monthly news-based, AI generated ESG scores, I apply portfolio sorting of stocks on the S&P1500 according to ESG and perform firm-level factor regressions using six common market factors to adjust for systemic risk, followed by implementing a GARCH-MIDAS equation for modelling the conditional variance of the idiosyncratic risk. I compare high and low aggregate ESG, as well as portfolios sorted on individual environmental, social and governance scores. The findings show no major differences in conditional variance between high and low ESG portfolios, though small differences in short-term volatility persistence and asymmetric effects are observed. I expand the model to include ESG (and individual E, S, G) scores in the long-term component of GARCH-MIDAS and find ESG scores to be weakly significant for the portfolios sorted on environmental score. Inclusion of ESG is shown to have an overall model improving effect. Further research may benefit from using (AI-based) ESG scores in modelling idiosyncratic volatility in other types of models.

Keywords: ESG, idiosyncratic volatility, GARCH-MIDAS

Acknowledgements

A special thank you goes out to Hossein Asgharian for supervision throughout the whole process from the research proposal, to model suggestions, and the writing of this paper. Additional thanks go to Erik Dahlberg of Sanctify for provision of ESG data and help in this regard, and to Emily Pickering for helpful comments on academic writing and structure.

TABLE OF CONTENTS

1. Introduction

The *2024 Institutional Investor Survey on Sustainability* (Larcker et al., 2024) reveals that 67% of investors consider the environmental, social and governance (ESG) performance of companies when making investment decisions. ESG encompasses a wide number of parameters across the three categories and is often bulked together to obtain a score or rating meant to represent the overall sustainability of a company. With increased popularity in the recent 10 years, a good ESG performance has been of interest both for investors and in the academic finance field for the source of potential excess returns, as well as the potential ability to mitigate financial risk on various levels. While the former has not yielded the expected results (O'Connor, 2023), the latter is still in question. Larcker et al.'s (2024) survey states that the majority of investors believe a good ESG performance reduces risk to some extent: 78% believe it reduces tail risk, while 61% believe it decreases overall return volatility. Returns of ESG have been researched rather extensively, however, the literature solely focussed on risk – especially volatility – is rather sparse. In this paper, I attempt to fill part of this gap by exploring the relationship between the ESG performance of companies and return volatility. I focus on *idiosyncratic* volatility, the motivation being to avoid overestimating the effect of ESG by ignoring other potential (systemic) factors which will affect overall volatility such as market movement, firm size and firm value. Few studies exist which explore this relationship, and those that do (see Dayong, Kaiyuan & Wenhua, 2023; Reber, Gold & Gold, 2021) only apply a linear instrumental variables approach in the empirical analysis. However, ESG rating has potential explanatory power when included directly as a covariate in the model (Ghani, Zhu & Ghani, 2024; Bouri, Iqbal & Klein, 2022). To extend both of these analyses, I apply a *general autoregressive conditional heteroskedasticity – mixed data sampling* (GARCH-MIDAS) model, introduced by Engle, Ghysels & Sohn (2013), which has gained interest in recent years for its ability to model long-term volatility. This model, previously applied to ESG data by Ghani, Zhu & Ghani (2024), allows for modelling idiosyncratic volatility as the combination of a short-term GARCH component, and a long-term MIDAS component, in which external variables (here: ESG rating) can be included in the latter, despite being sampled at a lower frequency. I fit models both with and without the inclusion of an ESG covariate to the idiosyncratic risk of portfolios of US stocks (S&P1500) sorted according to ESG score. By doing so, I compare parameters of conditional variance between high-ESG and low-ESG, and explore the explanatory power of the ESG parameter. I further extend the empirical analysis by repeating the process using individual 'pillars' of ESG, that is, ratings of environment, governance and social performance, respectively. I implement AI-based ESG ratings by *Sanctify*, which provides several potential benefits. Due to Sanctify's rating being based on online-available news, rather than company-provided data, they have the potential for being more neutral than other scoring systems. Secondly, it provides more frequent data across all categories of scoring, allowing for the use of monthly aggregate ESG scores, as well as individual environment, social and governance scores.

I find no major differences between high and low ESG, although in some cases minor differences are observed particularly in the short-term volatility persistence and asymmetric effects. The ESG scores are only found weakly significant for portfolios sorted on the 'environment' scores, however, inclusion of ESG scores generally improves the model across all pillars, suggesting potential explanatory power, or at least model-improving effects. Additionally, coefficients (though not strongly significant) suggest higher ESG reduces conditional variance of idiosyncratic risk, and that the presence of ESG scores may increase informational transparency.

The contribution of this paper to the literature is three-fold: Firstly, to the extent of my knowledge, this is the first paper to explicitly explore the relationship between idiosyncratic volatility and ESG performance specifically in the US market. Secondly, by combining multifactor Fama/French models with the GARCH-MIDAS approach, I present a method for including ESG ratings as in-model predictor of idiosyncratic volatility. Finally, this paper explores the usage and usefulness of AI-generated, news-based ESG scores, which may be of interest for both academic researchers and investors. The paper is structured in 5 sections: Section 2 provides a theoretical background and literature overview of previous research on ESG and volatility modelling. Section 3 introduces the data and explains the methods and models applied, before presenting the empirical results and a discussion thereof in Section 4. Conclusions are presented in Section 5.

2. Theoretical background and literature review

In this section I summarize the main theoretical ideas and provide an overview of literature for the topics of sustainable investment, ESG ratings and volatility modelling.

2.1 ESG and stock performance

The relationship between ESG performance and risk has been debated over the years. Theory suggests that ESG can reduce risk both by reducing its reaction to macroeconomic and company specific risks, as well as by increasing transparency. Specifically, high ESG can mitigate risk by reacting less to shocks related to environmental disasters and climate change, human rights issues or bad governance (Fabozzi, Ma & Oliphant, 2008). Furthermore, good ESG implies higher transparency, allowing for lower fluctuation of prices, further enhanced by increased analyst attention, effectively increasing market efficiency (Dayong, Kaiyuan & Wenhua, 2023). Counter arguments, on the other hand, stem from traditional portfolio theory, suggesting that active ESG-focussed investors exclude so-called 'sin-stocks', reducing the investment universe and hence possibility for diversification (Pastor, Stambaugh & Taylor, 2021; Ott & Zincke, 2020). The effect of ESG is partially decided by the overall awareness and preferences with regard to ESG among investors. If investors are ESG-motivated, they are more likely to pay extra for high ESG stocks, which can help lower companies' cost of capital (Pastor, Stambaugh & Taylor, 2021). The general awareness of ESG also affects whether ESG risk is 'priced', and thereby whether it can be the source of excess returns (Pedersen, Fitzgibbons & Pomorski, 2021).

The empirical evidence on risk and return of ESG is mixed. Friede, Busch & Bassen (2015) performed a meta-study and found an overall positive link between ESG performance and returns. Other studies, however, find that there are no abnormal returns, or in some cases even negative returns on ESG (Halbritter & Dorfleitner, 2015; Ott & Zincke, 2020; Pedersen, Fitzgibbons & Pomorski, 2021). Although often bulked together as an aggregate measure, the three 'pillars' of ESG differ in effects. The governance ('G') factor seems to generally have the largest impact, due to being related to traditional company-specific risk-averting and stability behaviour (Pedersen, Fitzgibbons & Pomorski, 2021). For instance, the governance factor has been shown to have a positive effect on firm value, while the social factor can have a negative effect, explained by the idea that good governance is a sign of stability, while social investments may lead to additional risk (Ionescu et al., 2019). Similarly, Pedersen, Fitzgibbons & Pomorski (2021) found that their portfolio sorted on 'G' had significant returns compared to no returns of general ESG and 'E', and negative returns of 'S'. The latter was also shown by Revelli & Riviani (2015, cited in Zarafat, Liebhardt & Eratalay, 2022) who found that socially responsible investments did not outperform benchmarks. This general notion is confirmed by Larcker et al.'s (2024) investor survey, which shows that a majority of investors rate governance as being the most important category, with social being largely unimportant, or seen as having negative impacts on performance. The survey suggests the general belief that the environmental category is becoming increasingly important, with climate risk being in the forefront of decision making especially looking towards the next decade, but that it is currently not priced.

ESG has been shown to have a reducing effect on various measures of risk. Chen & Ying (2023) found that highly rated ESG stocks can have lower realized volatility compared to low rated stocks, and Ott & Zincke (2020) showed that high ESG scores (particularly the governance pillar) reduces tail-risk. Zarafat, Liebhardt & Eratalay (2022) studied the difference in asymmetry effects and found that ESG has an influence on the difference in reaction to negative and positive news but that the exact effect varies between industries as well as over time, suggesting that the effect of ESG was different before and after the COVID-19 pandemic. Furthermore, as shown for the Chinese market, a high ESG rating may reduce idiosyncratic risk further compared to a low one (Dayong, Kaiyuan & Wenhua, 2023). Idiosyncratic risk (also known as unsystematic risk) is firm specific, i.e. not explained by market factors. Often, this is assumed to be able to be diversified away, however this is not always the case and is shown to have an effect on the overall stock returns (Bechetti, Ciciretti & Hasan, 2015). If a stock is more informationally transparent, idiosyncratic risk can be reduced to systemic risk (Dayong, Kaiyuan & Wenhua, 2023); disclosing ESG information can add to this informational transparency (Feng et al., 2022; Reber, Gold & Gold, 2021).

2.2 ESG ratings

Although the concept of sustainable investment has existed for centuries (Ott and Zincke, 2020), ESG scores emerged in the 1980s and have evolved hand in hand with the interest in sustainable investment (Berg, Kölbel & Rigobon, 2022). The core idea is to give a quantitative measure of a company's performance across the three categories (Environment, Social, Governance). The exact definition of ESG, however, is unclear; it depends on a number of different and evolving values and is in many cases limited to the information companies themselves choose to disclose, something which varies from jurisdiction to jurisdiction and country to country. Often ESG disclosure is voluntary. As a result of this, ESG ratings differ substantially depending on the provider. Berg, Kölbel and Rigobon (2022) explore the divergence between ESG ratings on six different major providers and find that the correlations range between 0.38 and 0.71. There are several underlying reasons for this divergence, including the method of measuring and the exact attributes which are considered. This inconsistency makes it difficult to draw general conclusions and informed decisions for investors and academic researchers alike.

An alternative approach to ESG is basing the rating on available news rather than company disclosed information. This paper uses ratings provided by Swedish tech start-up Sanctify, who implements natural language processing (NLP, more precisely a BERT method) to scrape news data related to companies' ESG performance. By doing so, they present an alternative approach to obtaining ratings, which, rather than being based on company disclosed information, uses available news regarding company performance and controversies available online. By implementing NLP, Sanctify is able to provide frequently updated, aimed-to-be unbiased ratings for a large number of companies worldwide. Sanctify's ESG scores are available at various frequencies from daily to yearly, and offer an aggregate ESG score as well as sub-scores for the environmental, social and governance categories, respectively. A further introduction to the methods behind the creation of the Sanctify ratings can be found in Sanctify (2024).

2.3 Modelling of (idiosyncratic) volatility

The seminal papers of Engle (1982) and Bollerslev (1986) introduce the autoregressive conditional heteroskedasticity (ARCH) and general ARCH (GARCH) models, respectively. The core of these models is that returns are explained by a mean equation and a variance equation. The variance equation is described by a (G)ARCH process, where model variance depends on lagged values of the residuals and lagged residual variance. Since then, the family of GARCH models has grown to cover various inconsistencies in the baseline models, such as asymmetry and leverage effects as well as the incorporation of exogenous variables; a general overview of GARCH models, as well as other volatility approaches can be found in So et al. (2021) and Engle & Patton (2007). This paper is concerned with the GJR-GARCH model of Glosten, Jagannathan & Runkle (1993). This model adjusts for the asymmetry (the difference in reaction to positive and negative shocks) by including an indicator function for when the residual (ARCH) term is negative. Although new alternatives to GARCH models exist, including the Heterogeneous Autoregressive (HAR) model (Corsi, 2009) and machine learning methods (Kristensen, Sigaard & Veliyev., 2021; Filipovich & Khalilzadeh, 2021), the GARCH family of models remains popular due to a well-recognised place in literature and the simplicity of interpretation.

GARCH models are as a rule univariate, despite volatility likely depending on external factors as well (Engle & Patton, 2007). To overcome this, Engle, Ghysels & Sohn (2013) introduces combining GARCH with the mixed data sampling (MIDAS) approach. This allows for modelling short-term volatility as a GARCH process, while including exogenous variables sampled at a lower frequency. Conrad & Kleen (2019) test this approach using both Monte Carlo methods and real-world stock data and find that while the heterogeneous autoregressive (HAR) model of Corsi (2009) generally performs the best on short term forecasts, GARCH-MIDAS has a very good long-term forecasting ability. The performance of GARCH-MIDAS depends on the choice of exogenous variable (Conrad & Kleen, 2019), which depends on the specific study. The most common exogenous factor is realized volatility (Conrad & Kleen, 2019; Asgharian, Christiansen & Hou, 2015) but there are examples which go beyond, such as the incorporation of macroeconomic variables (Asgharian, Christiansen & Hou, 2015) and ESG / climate policy uncertainty (CPU) indexes (Ghani, Zhu & Ghani, 2024).

The GARCH infrastructure involves modelling the mean and variance of a series. In some cases, the mean may be assumed constant, but is often modelled as an autoregressive moving average (ARMA) model (Engle & Patton, 2007). This is however not exclusive. Several papers (e.g. Xiao, Huang & Newton, 2024; Yamani & Swanson, 2014; Ahmed & Alhabd, 2020) apply a Fama-French multifactor model as a mean equation and model the residuals using GARCH-type models. By implementing this approach, the idiosyncratic volatility is modelled, rather than the overall variance of returns. The factor approach of Fama & French is one of the most seminal and applied approaches to asset pricing; the literature of this, which is based on extending the capital asset pricing model (CAPM) of Sharpe (1964) and Arbitrage pricing theory of Ross (1976), spans more than half a century. Notable papers include Fama & French (1993) which introduces the (perhaps most famous) three-factor model, and Fama & French (2015) which introduces the five-factor model. The factor approach of Fama & French has frequently been applied to study the relationship between ESG and stock performance, such as in Ott & Zincke (2020), Halbritter & Dorfleitner (2015), Pedersen, Fitzgibbons & Pomorski (2021), and Bechetti, Ciciretti & Hasan (2015).

3. Methods and Data

In this section I describe the data source and sample selection, as well as the software used for analysis (3.1). Subsection 3.2 presents the general empirical approach and covers the sorting of portfolios according to ESG, calculation of residuals using Fama-French regression, and the modelling of volatility using GARCH-MIDAS with and without the ESG covariate.

3.1 Data & software

The quantitative data used in this paper consists of three main components: stock returns, factor portfolios and ESG ratings. The sample is based on the S&P1500 index which combines three American indexes, in total encompassing large, mid and small cap stocks trading on the US stock exchanges. The choice to focus on the US market is due to data availability and the fact that, to my knowledge, no studies exist which explore the relationship between idiosyncratic volatility and ESG specifically for the US market. Returns data is downloaded from Yahoo finance and is based on the difference between daily adjusted close values. Factor portfolios are based on the six factors of Fama & French (see section 3.2) and are downloaded together with the risk-free rate from Kenneth French's website. The design and updating procedure of the portfolios can be found in Fama & French (2023).

ESG ratings are downloaded using Sanctify's application programming interface (API). The ESG data provides four separate monthly scores for each company: Environmental (ENV), governance (GOV), social (SOC) and aggregated average ESG. The scores are based on text-based news collected using natural language processing and range from -100 to 100 with the sign indicating bad or good performance. A score of 0 indicates neutral performance – this is also the score allocated when there is no (news) information available for a given company in a given month. Sanctify ESG has three options for the 'term' of the score: short, mid and long, which determine the extent to which past scores affect current scores. For this paper, I will be using the mid-term scores for a balanced approach. Each of the categories ENV, GOV and SOC are based on a number of subcategories/criteria, the ESG score is an average of the ENV, GOV and SOC score together. Further details on calculations can be found in Sanctify (2024).

The time series of the data spans from January 2016 to December 2023, chosen according to the availability of ESG data (although Sanctify's data base dates back to 2010, the data prior to 2016 is fairly sparse). Stocks containing NAs and/or no available ESG information were removed for a final sample size of 1165 stocks. All data was retrieved in April 2024. Data processing, cleaning, portfolio sorts and Fama-French regressions were performed using Python v. 3.11.5. GARCH models were fit in *R* v. 4.3.2 using the `mfGARCH` library (Kleen, 2021).

3.2 Fama-French – GARCH-MIDAS approach

The following section describes the process of model creation. The process consists of three main parts: portfolio sorting according to ESG, Fama-French (FF) regression to estimate residuals, and GARCH-MIDAS modelling idiosyncratic volatility. A summary of the full procedure is found at the end of the section.

Following the approach of Ott & Zincke (2020), I sort the stocks in the sample according to their ESG rating, dividing them into deciles. Although much of the Fama/French literature prefers quintiles (Fama & French, 1993), I chose to bulk in a smaller bin width in order to compare extremes. Across the data, a large number of companies have a monthly ESG rating close to zero (neutral), hence in order to have a true high vs. low comparison, a smaller bin width is preferred. The portfolios are updated monthly according to the ESG scores; the 1st, 5th and 10th decile create the high, neutral and low ESG portfolio, respectively. The portfolios are weighted equally, meaning that besides the ranking, ESG only determines which stocks go into the portfolios, but not how much of each stock is being held. This choice was made to avoid effects of singular stocks to skew the results. Separate sorting and portfolio creation are performed for aggregate ESG, ENV, GOV and SOC. Summary statistics for the portfolios can be found in Appendix 1 and are discussed in section 4.1.

While in the literature, portfolios are typically formed *prior* to fitting the FF regression, I choose an alternative approach. I run FF regressions on a stock specific level, and then consolidate the residuals into portfolios according to ESG afterwards. The reason for this is that it allows to adjust returns according to size, book-to-market and the other factors on a company specific level, rather than on a portfolio level, allowing the residuals of the portfolio to be adjusted for company specific attributes. The next section describes the FF regression approach to estimate idiosyncratic risk, and the GARCH-MIDAS model used to model the idiosyncratic volatility.

Following a similar approach to Xiao, Huang & Newton (2024) and Yamani & Swanson (2014), the mean equation of excess return is modelled by the Fama/French six-factor (FF6) model, with the variance being described by a GARCH-type process. The initial regression makes use of Fama & French's (2015) five-factor model, extending it with the momentum factor introduced by Jegadeesh & Titman (1993) and applied in Carhart (1997), amongst others. The combination of all factors is a standard way of estimating idiosyncratic risk (Fink, Fink & He, 2012). The six factors are the excess market return, and 5 factor portfolios, each designed as the difference between the 'best' or highest rated portfolio and the lowest rated. These are book-to-market (high minus low), size (big minus small), level of investment (conservative minus aggressive), profitability (robust minus weak) and momentum (winners minus losers). Summary statistics can be seen in Appendix 4, which also provides specific references for each factor for further reading on design and background.

The FF6 regression serves as the mean equation and is estimated as:

$$R_{i,d} - rf_d = b_{0,i} + \sum_{k=1}^{K=6} b_k F_{k,d} + \epsilon_{i,d}$$

(1)

where $R_{i.d} - rf$ is the *excess return* over the risk-free rate for stock i on day d. F_k are the respective Fama/French factors described above. $\epsilon_{i,t}$ is the *idiosyncratic risk*, to be understood as the excess return unrelated to the market and systemic factors and estimated by the residuals of the FF regression from the section above. The volatility of excess returns is determined by this term, which I model by a GARCH-MIDAS, such that:

$$\epsilon_{d,t} = \eta_{d,t}\sqrt{\tau_t \times g_{d,t}}$$

(2)

where τ_t is the long-term component, $g_{d,t}$ is the short-term component, and $\eta_d \sim iid(0,1)$ is a white noise term. The short-term component $g_{d,t}$ for day d in month t follows a GJR-GARCH process:

$$g_{d,t} = \mu + (\alpha + \gamma S)\frac{\left(\epsilon_{d-1,t} - \bar{\epsilon_t}\right)^2}{\tau_t} + \beta g_{d-1,t}$$

(3)

where $S_{d-1,t} = \begin{cases} 1, & if\ \epsilon_{d,t} < 0 \\ 0, & otherwise \end{cases}$

is the indicator function of GJR, and $\mu = (1 - \alpha - \beta - \gamma/2) > 0$. α is the coefficient of the residual (ARCH) term, β of the short-term volatility (GARCH) term, and γ is the GJR asymmetry coefficient, indicating the relative effect of negative shocks on volatility compared to positive.

The long-term component is given by the MIDAS equation:

$$\log(\tau_t) = m + \theta_X \sum_{k=1}^{K} \phi(w_1, w_2) X_{t-k}$$

(4)

where X_{t-k} is a lagged exogenous variable, varying between models (see below). Following Asgharian, Christensen & Hou (2015), the chosen lag length for all models in this paper is $K =$

12, i.e. 1 year. $\phi(w_1, w_2) > 0$ is the weighting scheme, known as the beta-function, where $w_1 = 1$ and w_2 is adjusted according to the function. This function determines how the lags of the exogenous variable are weighted (Conrad & Kleen, 2019). I apply a 'restricted' beta function, such that the lags are weighted decreasingly.

I implement four different models for the variance, two benchmark models and two where I extend these with ESG ratings. The first bench-mark model has no long-term component and the variance equation reverts to a standard GJR-GARCH(1,1) model. The second (model 2, *GM-RV*) has monthly realized idiosyncratic volatility as the long-term component. The long-term component is given by:

$$\log(\tau_t) = m + \theta_{RV} \sum_{k=1}^{K} \phi(w_1, w_2) RV_{t-k} \tag{5}$$

where RV_{t-k} is lagged realized monthly idiosyncratic volatility calculated as the root of the sum of squared daily residuals per month:

$$RV_t = \sqrt{\sum_{d=1}^{N_t} \epsilon_{d,t}^2} \tag{6}$$

Model 3 (*GM-ESG*) is a GARCH-MIDAS where the respective score for ESG is implemented as the only long-term covariate:

$$\log(\tau_t) = m + \theta_{ESG} \sum_{k=1}^{K} \phi(w_1, w_2) ESG_{t-k} \tag{7}$$

ESG_{t-k} is the lagged ESG rating. Equivalent models are created for using the individual ENV, SOC, and GOV ratings. Finally, model 4 (*GM-RV-ESG*) combines model 3 and 4 such that both RV and the respective ESG score are included as covariates.

$$\log(\tau_t) = m + \theta_{RV} \sum_{k=1}^{K} \phi(w_1, w_2) RV_{t-k} + \theta_{ESG} \sum_{k=1}^{K} \phi(w_1, w_2) ESG_{t-k} \tag{8}$$

The overall approach is performed as follows: FF regressions are estimated on a firm level for every firm in the sample. Firms are sorted according to monthly ESG value, and residuals are then summed per portfolio, forming three portfolios (high, neutral, low) according to the respective rating. This process is repeated for each pillar (aggregate ESG, ENV, SOC, GOV). The average score and realized monthly idiosyncratic risk are calculated per portfolio. Each of the variance models 1 through 4 is then fit to each portfolio using the respective covariate as described above. The next section analyses the empirical results.

4. Empirical results and discussion

This section presents the empirical findings of the models described in the section above. I start by presenting summary statistics of the main variables, followed by an analysis of the benchmark models for differences in portfolios. I then proceed to comment on the inclusion of ESG ratings in the model and compare model performance.

4.1 Summary statistics

Table A1.1 (Appendix 1) shows the summary statistics of the three ESG portfolios. The extreme values of the residuals are due to the additive nature of how they were calculated; since these are added up after the portfolio formation, these become very large in periods of general large deviations. The augmented dickey fuller (ADF) tests reject unit root presence at 1%, so I can assume stationarity of the residuals. This is also the pattern exhibited in Figure 1, which also suggests volatility clustering, confirming that a GARCH-type model is appropriate. Note that realized volatility (RV) is stationary as well, but the ESG ratings are not. This is due to the nature of the ESG scores, and may have an effect on the model; this is discussed in section 4.4. There is no immediate difference in overall standard deviation between the portfolios, nor do the trends in residuals over time reveal any major differences in overall volatility (Figure 1). The neutral portfolio has the largest standard deviation, this is a consistent result across individual pillars as well. The average ESG rating is generally close to zero, a major difference lies in the sign: high ESG has positive ESG ratings, low ESG has negative ratings, and the neutral portfolio has a rating of ~0 for the majority of the sample. This is explained by the nature of the scoring methods, where ‘good’ and ‘bad’ ESG performance have positive and negative scores, respectively, while those that are neutral or lack sufficient information are scored 0.

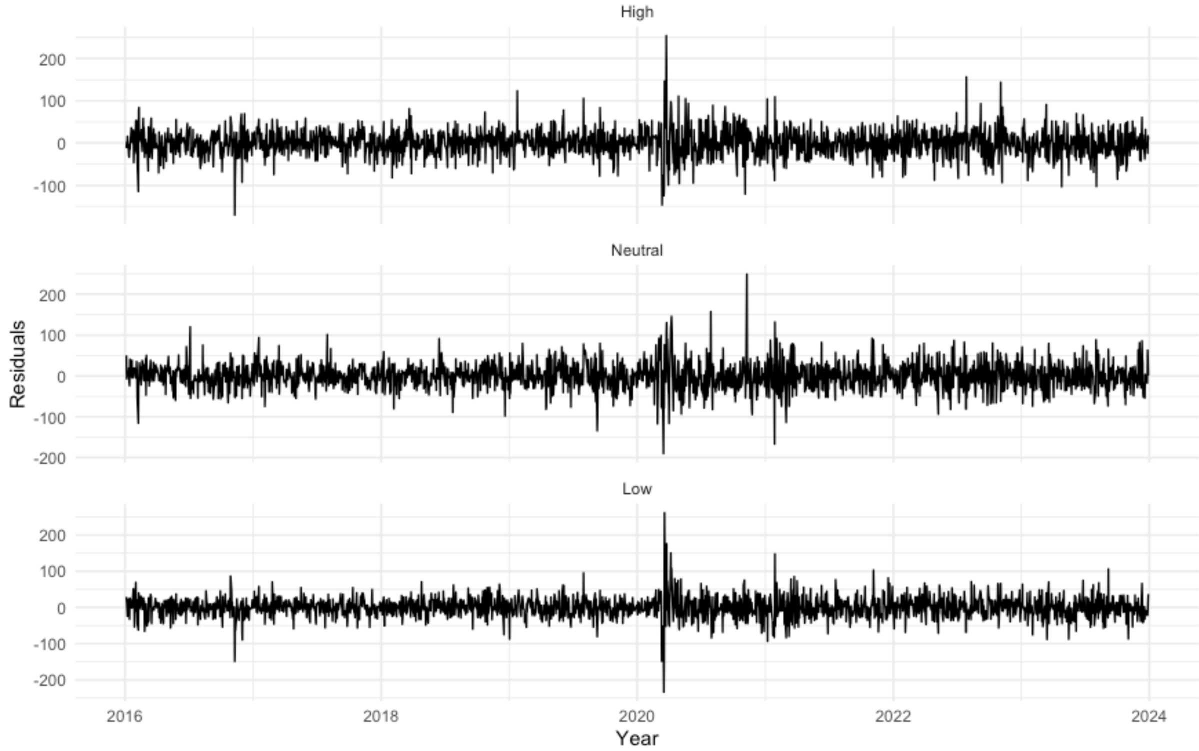

Figure 1: Daily Idiosyncratic risk for different portfolios sorted according to aggregate ESG rating.

4.2 Benchmark models

To identify differences in conditional variance between high and low ESG portfolios, I compare the parameters of the benchmark models. The main focus is on the aggregate ESG portfolio, although comparisons are made across all pillars. Tables representing the numeric results of ENV, SOC and GOV portfolios can be found in Appendix 2 (A2). Tables for aggregate ESG are repeated in text for convenience. For the coefficients, * , **, and *** represent significance at the 10%, 5%, and 1% level, respectively, calculated using robust standard errors for volatility models of Bollerslev & Wooldridge (2007).

Table 1 shows the results of the first benchmark model, GJR-GARCH(1,1). α and β are highly significant showing strong evidence that one-day lagged values of residuals and volatility impact the conditional variance. β is close to 1, indicating very strong volatility persistence and confirming the presence of clustering patterns. α is much smaller, suggesting only a small effect of short-term shocks on the conditional variance. The asymmetry term γ is small and insignificant, indicating no difference in reaction between negative and positive news on a daily basis. The low-ESG portfolio has the largest α, and the lowest β term, indicating relatively stronger reaction to short-term shocks and a slightly less volatility persistence compared to the other two portfolios. The neutral portfolio exhibits the opposite pattern, indicating the largest volatility persistence, but the lowest impact of short-term shocks. Note that some trade-off between α and β is natural do to the model restrictions presented in section 3.2, equation (3).

Table 1: Results of GJR-GARCH(1,1) for different portfolios sorted on aggregate ESG rating.

Model 1	μ	α	β	γ
GJR-GARCH				
High	-0.323	0.0348***	0.9615***	0.0074
Neutral	0.2633	0.0151***	0.9807***	0.0084
Low	-0.1728	0.0577***	0.9425***	-4.00E-04

Portfolio sorts on ENV and SOC exhibit a similar pattern as that of aggregate ESG with small and significant α, close-to-unity and significant β, and insignificant intercepts and γ (Table A2.2, A2.3). SOC has identical α and β to two decimal places across the three portfolios, suggesting that for a simple GARCH, there is no difference between high and low social rating in the effect on the conditional variance. GOV sorted portfolio shows γ as significant on 1% level for all three portfolios, indicating strong evidence for asymmetry effects (Table A2.3). High and neutral GOV portfolio has a moderate asymmetry effect, while the low-GOV the coefficient is notably lower. This suggests the low governance rated firms exhibits less asymmetry; negative and positive shocks have a more equal effect, whereas higher rated GOV companies may experience a larger impact from negative news compared to positive.

In model 2 (GARCH-MIDAS with realized (idiosyncratic) volatility (RV) as the long-term component), the ARCH and GARCH components have a similar interpretation as in the simple GJR-GARCH, however we do see a slightly reduced β value, indicating some of the lagged daily volatility can be explained by monthly volatility (Table 2). The neutral portfolio has the smallest effect of short-term shocks and the largest persistence of short-term volatility similar to model 1. The coefficient of RV is small, negative (~ -0.01) and insignificant. This is in contrast to previous literature, which finds θ_{RV} significant and close to 1 (Conrad & Kleen, 2019); however, it should be noted that most previous literature focuses on modelling return volatility directly, rather than idiosyncratic volatility, and the difference may stem from this.

Table 2: Results of GARCH-MIDAS-RV for portfolios sorted according to ESG.

Model 2	μ	α	β	γ	m	θ_{RV}	$w2_{RV}$
GM-RV							
High	0.1546	0.0377^{***}	0.959^{***}	0.0066	1.0674	-0.0101	4.0023
Neutral	0.153	0.0276^{***}	0.9663^{***}	0.0122	1.5376	-0.0113	3.4284
Low	0.4248	0.0468^{***}	0.9474^{***}	0.0116	0.7967	-0.0086	3.6008

The asymmetry effect is very small for high-ENV, and largest for low-ENV (Table A2.2). This pattern is also present in the simple GJR-GARCH, but more pronounced in the GM-RV model. This implies that an increase in environmental performance decreases the relative effect of negative news on short-term volatility; however, as the parameter is insignificant this should be viewed with caution. The SOC portfolio shows both γ and θ_{RV} significant at 10% level (Table A2.3). Here, the neutral portfolio is the least asymmetric, indicating that both high and low social stocks have higher asymmetry. Assuming investors are aware of the social impact, this could indicate that companies involved in news-worthy stories about social issues react more strongly to negative news on both sides of the spectrum, but less so for those not who are neutral (scores are very close to zero for the whole spectrum). A potential explanation lies in the design of the scoring system: close-to-zero values of the neutral portfolio stem from lack of news/information, rather than a true neutral social performance, and hence is less reactive. The RV coefficient is small and negative, a negative effect suggesting RV has a decreasing effect on the long-term conditional variance.

Overall, the benchmark models do not suggest any major differences in short or long-term idiosyncratic volatility, particularly in the ARCH and GARCH terms. There is weak evidence for differences in asymmetry; the significance and effect vary between pillars. RV does not seem to carry information, except for the SOC pillar in which it shows a small, decreasing effect. The next section interprets the results of the models where ESG (and the respective pillars) is included in the long-term MIDAS component, and compares in-sample performance of the models for each portfolio.

4.3 Models with ESG as covariate

For the aggregate ESG model, α and β have the same relative relationship as in the benchmark models, with a similar reduction in β compared to model 1 as RV (Table 3). γ is still insignificant and shows no difference between portfolios. Although the coefficient of ESG rating θ_{ESG} is not significant on the 10% level, we see some interesting differences between the portfolios. For the high portfolio, the coefficient is -0.41 suggesting a moderate, decreasing effect of ESG rating on long-term volatility. The low portfolio has a large positive effect, showing increasing effect. The magnitude here should not be given too much attention as it is likely to simply reflect the difference in magnitude between the average scores of the portfolios. Furthermore, the signs of the coefficient will naturally be reversed, as the low portfolio has negative ESG ratings. Hence, the interpretation is that higher ESG ratings will decrease the conditional variance in the long term, while negative ESG scores indicates an increase. The coefficient, however, is not significant and interpretations are therefore not conclusive. None of the individual pillars show θ_{rating} as significant. ENV (A2.2) and SOC (A2.3) show a similar pattern to aggregate ESG, with a negative coefficient for high rated portfolios, and larger positive coefficient for low-rated portfolios. For GOV, θ_{GOV} is small and negative for all portfolios (Table A2.3).

Table 3: Results of GARCH-MIDAS with ESG as covariate for portfolios sorted according to aggregate ESG.

Model 3	$\boldsymbol{\mu}$	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$	**m**	$\boldsymbol{\theta_{ESG}}$	$\boldsymbol{w2_{ESG}}$
GM-ESG							
High	0.2894	0.0288***	0.964***	0.0144	0.0132	-0.4065	3.5897
Neutral	0.4081	0.0166***	0.9757***	0.0154	-0.0202	0.2275	3.3703
Low	0.4083	0.0482***	0.9438***	0.016	0.962	9.6899	1.0001

For the model with two covariates (*GM-RV-ESG*), ESG, SOC and GOV show no major change in the effect of the parameters when including both RV and ESG/SOC/GOV rating (Table 4, Table A2.4). RV remains small and insignificant, and the patterns of ESG/SOC/GOV exhibited in the GM-ESG models remain the same. For the environmental-focused portfolios (A2.2), θ_{ENV} becomes significant at 10% level. The effect is decreasing, suggesting evidence that better environmental performance reduces long-term idiosyncratic volatility, while 'bad' environmental performance leads to an increase. The effect is notably larger for the high and low ENV portfolios compared to the neutral one, suggesting the effect is more pronounced for the either end of the spectrum.

Table 4: Results of GARCH-MIDAS with ESG score and realized idiosyncratic volatility as covariates for portfolios sorted on aggregate ESG.

Model 4	μ	α	β	γ	m	θ_{RV}	$w2_{RV}$	θ_{ESG}	$w2_{ESG}$
GM-RV-ESG									
High	0.141	0.0394***	0.9579***	0.005	1.305	-0.008	4.547	-0.675	2.744
Neutral	0.223	0.0319***	0.9578***	0.021	2.043	-0.014	3.328	-1.842	3.151
Low	0.465	0.044***	0.9497***	0.013	1.373	-0.010	3.301	3.846	1.820

Despite only weak significance, measures of model performance suggest that inclusion of ESG improves the overall model. Table 5 compares the models for each portfolio using Bayesian Information Criterion (BIC) and log-likelihood (LLH). Equivalent tables for other pillars are shown in Appendix 3.

Table 5: In-sample model performance for portfolios sorted according to aggregate ESG.

	High		**Neutral**		**Low**	
	BIC	LLH	BIC	LLH	BIC	LLH
GJR	19540	-9751	19671	-9817	19104	-9533
GM-RV	**17090**	**-8519**	**17283**	-8615	**16748**	-8348
GM-ESG	17108	-8528	17303	-8626	16766	-8357
GM-RV-ESG	17106	**-8519**	17295	**-8614**	16761	**-8347**

For aggregate ESG, both models implementing ESG outperforms basic GJR-GARCH on both measures, suggesting that including ESG as a covariate improves the overall model. Using BIC, the best performing model is GM-RV, while LLH selects GM-RV-ESG. This confirms previous literature where it has been shown that largely RV is the stronger indicator, but that by including both RV and ESG, there is a possibility to improve the model. A similar pattern is shown for ENV, GOV and SOC. In some cases, the log-likelihood is larger for the GM-RV-ESG models, indicating ESG may have model-improving effects overall. Opposite of Bouri, Iqbal & Klein (2022), the results are consistent across the different portfolios, suggesting there is no difference in explanatory power for a low or high ESG portfolio. The results are similar across the four pillars as well. In some cases, BIC selects the GM-ESG or GM-RV above GM-RV-ESG, due to the parsimony-preferring nature of the criterion.

In line with the findings of Ghani, Zhu & Ghani (2024), including ESG rating as a parameter in GARCH-MIDAS improves the overall model especially compared to a simple GJR-GARCH. The coefficients of ESG parameters suggests that higher ESG ratings reduces long-term conditional variance of idiosyncratic risk, although significance of the parameter is weak at best, suggesting that conclusions with regard to this should be interpreted with caution. We generally see a pattern that the neutral portfolio behaves differently than the others: since neutral may indicate lack of information, rather than true neutral, this favours the argument of Reber, Gold & Gold (2021) that the presence of ESG information leads to higher informational transparency, but not consistently that of Dayong, Kaiyuan & Wenhua (2023) who finds a significant inverse linear correlation ESG score and idiosyncratic volatility (i.e. higher ESG results in lower volatility).

4.4 Points of improvement and suggestions for further research

The use of Sanctify ESG scores may lead to slight misspecifications in the model. Across the full sample, many observations (especially for the aggregate ESG ratings) have value 0 or close to 0. As described in Sanctify (2024), this indicates a period for which no news were present, hence it is treated as neutral. To overcome this, Sanctify has their data divided into 3 terms, as discussed in section 3.1, allowing previous scores to have an effect on later scores, which decreases the risk of many neutral values and is perhaps more reflective of how investors see ratings (very negative news in one month are likely to be reflected in the following months). However, this automatically results in non-stationarity of the ESG scores, which is an assumption when using them as a MIDAS variable (Conrad & Kleen, 2019). Possibly this can be mitigated by the use of short-term scores, but there will naturally be a trade-off between having non-zero scores and stationarity. This misspecification may be affecting the results presented in this paper, hence additional research with less restricted models may be beneficial for more robust conclusions. An alternative model which may be of interest could be HAR with external variables, which takes advantage of a linear, lagged-volatility based approach (Corsi, 2009). Furthermore, machine learning methods could outperform traditional methods (Kristensen, Sigaard & Veliyev, 2021; Filipovitc & Khalilzadeh, 2021). Particularly, as highlighted by Kristensen, Sigaard & Veliyev (2021), a long short-term memory (LSTM) neural network model can work very well for predicting volatility, with a similar structure to that of GARCH, but fewer limitations in terms of assumptions. Applying these approaches can expand the use of ESG, particularly for its forecasting abilities. The downside of this type of model is that it loses some of the ease of interpretation which the GARCH-type models have, which may be of continued academic interest.

Additional to model specifications, different time periods also have an impact on the effect of ESG. The effect of ESG may vary between different time periods, especially in ‘crisis’ periods, such as the COVID-19 pandemic (2020-21) (Zarafat, Liebhardt, & Eratalay, 2022). Furthermore, investor preferences of ESG change over time (Pedersen, Fitzgibbons & Pomorski, 2021), which may also have an effect on returns and hence volatility. Thus, future research should account for changes in periods, either by modelling each window separately, or by including dummy variables for specific time periods. A constraint here is that the sample period is already rather short due to the availability of consistent ESG ratings, and hence the quality of analysis will naturally improve as time passes and more ESG information is made available. Lastly, Larcker et al. (2024) shows ESG preferences and legal frameworks to vary between the US and EU, so for a complete picture, multiple universes should be investigated.

5. Conclusion

ESG investing has been widely researched and is of increasing interest for both investors and academics. Previous literature has studied both the effect on returns and risk; results vary due to differences between each pillar (environment, social, governance) and divergence between ESG scores between different providers. To mitigate this, I implement monthly news-based ESG and individual E, S, & G scores. There are multiple ways of modelling stock volatility, including those based on GARCH. This paper studies the relationship between idiosyncratic (unsystematic) risk and ESG by applying a GARCH-MIDAS model to the residuals of Fama-French 6-factor model, allowing the inclusion of ESG rating in the long-term component. I find only slight differences in the conditional variance between high and low ESG, but including ESG rating as a long-term component generally has an effect on the coefficients and results in better performance measured by BIC and log-likelihood. In multiple models, I find that the neutral portfolio exhibits stronger volatility clustering than high and low ESG, and for portfolios sorted on 'social' rating a stronger asymmetry effect. This suggests that presence of ESG ratings improve transparency, whether or not ESG performance is 'good' or 'bad'. The portfolios sorted on 'governance' show strong evidence of asymmetry, with low-governance having the smallest effect. For the portfolios sorted on 'environmental' score, the environment parameter is found significant at $p < 0.1$ when including realized idiosyncratic volatility as a covariate, and suggests a decreasing effect on the long-term volatility. A similar pattern is observed for other portfolios; however, the parameter is not shown to be significant. The use of news-based ESG scores offers the possibility of frequently sampled data across four pillars, but in this model may carry more information as to whether news exists or not, rather than whether it is positive or negative. There are some underlying model issues which require attention, such as non-stationarity of ESG scores and potential varying effects between time periods and investment universes. Further research should take this into consideration, along with applying other types of models such as those based on machine learning. In conclusion, this paper shows the potential for ESG ratings to influence both transparency and long-term idiosyncratic volatility and showcases the potential for future development of models working with ESG performance to evaluate the effects on investment strategies and outcomes.

References

Ahmed, M., & Alhadab, M. (2020). Momentum, asymmetric volatility and idiosyncratic risk-momentum relation: Does technology-sector matter? *The Quarterly Review of Economics and Finance*, Volume 78, 2020, pp. 355-371, https://doi.org/10.1016/j.qref.2020.05.005

Asgharian, H., Christiansen, C., & Hou, A. (2015). Effects of macroeconomic uncertainty on the stock and bond markets. *Finance Research Letters*, Volume 13, 2015, Pages 10-16, ISSN 1544-6123. https://www.sciencedirect.com/science/article/pii/S1544612315000367

Becchetti, L., Ciciretti, R. & Hasan, I. (2015). Corporate social responsibility, stakeholder risk, and idiosyncratic volatility. *Journal of Corporate Finance*, Volume 35, 2015, pp. 297-309, ISSN 0929-1199. https://doi.org/10.1016/j.jcorpfin.2015.09.007

Berg, F., Kölbel, J., & Rigobon, R. (2022). Aggregate Confusion: The divergence of ESG ratings. *Review of Finance*, Volume 26, Issue 6, November 2022, Pages 1315–1344. https://doi.org/10.1093/rof/rfac033

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31(3), pp. 307-327.

Bollerslev, T., & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews*, 11(2), 143–172. https://doi.org/10.1080/07474939208800229

Bouri, E., Iqbal, N., & Klein, T. (2022). Climate policy uncertainty and the price dynamics of green and brown energy stocks. *Finance Research Letters*, Volume 47, Part B, 2022. https://doi.org/10.1016/j.frl.2022.102740

Carhart, M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, vol. 52, issue 1, pp. 57-82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Chen, J., & Ying, Y. (2023). ESG Performance and Stock Price Volatility. *Academic Journal of Business & Management*, 2023, 5(22). https://francis-press.com/papers/13181

Conrad, C., & Kleen, O. (2018). Two Are Better Than One: Volatility forecasting using multiplicative component GARCH models. *Journal of Applied Econometrics*, Vol. 35, issue 1 (2020), pp. 19-45. https://doi.org/10.2139/ssrn.2752354

Corsi, F. (2009). A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 2009, vol. 7, issue 2, 174-196.

Dayong, L., Kaiyuan, G., & Wenhua, H. (2023). ESG performance and stock idiosyncratic volatility. *Finance Research Letters*, Volume 58, Part B, 2023, 104393. ISSN 1544-6123. https://doi.org/10.1016/j.frl.2023.104393

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

Engle, R., & Patton, A. (2007). What good is a volatility model? *Forecasting volatility in the financial markets*. Butterworth-Heinemann, 2007. 47-63. https://www.sciencedirect.com/science/article/pii/B9780750669429500042

Engle, R. F., Ghysels, E., & Sohn, B. (2013). Stock Market Volatility and Macroeconomic Fundamentals. *The Review of Economics and Statistics*, 95(3), 776–797. http://www.jstor.org/stable/43554794

Fama, E., & French, K. (1993). Common risk factors in the returns of stocks and bonds. *Journal of Financial Economics*, volume 33, Issue 1, February 1993, pp. 3-56. https://doi.org/10.1016/0304-405X(93)90023-5

Fama, E., & French, K. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, Volume 116, Issue 1, April 2015, Pages 1-22. https://doi.org/10.1016/j.jfineco.2014.10.010

Fama, E., & French, K. (2023). Production of U.S. Rm-Rf, SMB, and HML in the Fama-French Data Library. *Chicago Booth Research Paper No. 23-22*. http://dx.doi.org/10.2139/ssrn.4629613

Fabozzi, F., Ma, K., & Oliphant, B. (2008). Sin Stock Returns. *The Journal of Portfolio Management*, Fall 2008, 35 (1), pp. 82-94. https://www.pm-research.com/content/iijpormgmt/35/1/82

Feng, H., Shuqi Q., Yuanyuan L., Ji W. (2022): CSR and idiosyncratic risk: Evidence from ESG information disclosure, *Finance Research Letters*, Volume 49, 2022. https://doi.org/10.1016/j.frl.2022.102936

Filipovic, D., & Khalilzadeh, A. (2021). Machine Learning for Predicting Stock Return Volatility. *Swiss Finance Institute Research Paper*, No. 21-95. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3995529

Fink, J., Fink, K., & Hui, H. (2012). Expected Idiosyncratic Volatility Measures and Expected Returns. *Financial Management*, Vol. 41, No. 3 (2012), pp. 519-553. https://www.jstor.org/stable/23325742

Friede, G., Busch, T., & Bassen, A. (2015). ESG and Financial Performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, Volume 5, Issue 4, p. 210-233, 2015. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2699610

French, K. (2024). Variable definitions. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/variable_definitions.html

Ghani, U., Zhu, B., Qin, Q., & Ghani, M. (2024). Forecasting US Stock Market Volatility: Evidence from ESG and CPU indices. *Finance Research Letters*, Volume 59, 2024. https://doi.org/10.1016/j.frl.2023.104811

Glosten, L.R., Jagannathan, R., & Runkle, D.E. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance*, 48(5), pp. 1779-1801. https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1993.tb05128.x

Halbritter, G., & Dorfleitner, G. (2015). The wages of social responsibility — where are they? A critical review of ESG investing. *Review of Financial Economics*, Volume 26, 2015, pp. 25-35. https://doi.org/10.1016/j.rfe.2015.03.004

Ionescu, G. H., Firoiu, D., Pirvu, R. and Vilag, R. D. (2019): The impact of ESG factors on market value of companies from travel and tourism industry, *Technological and Economic Development of Economy*, 25(5), pp. 820-849. https://doi.org/10.3846/tede.2019.10294

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance*, vol. 48 issue 1, pp. 65-91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x

Kleen, O. (2021). mfGARCH: Mixed-Frequency GARCH Models, R package version 0.2.1. https://github.com/onnokleen/mfGARCH/

Kristensen, K., Sigaard, M., & Veliyev, B. (2021). A machine learning approach to volatility forecasting. *CREATES Department of Economics and Business Economics Aarhus University*. https://econ.au.dk/fileadmin/site_files/filer_oekonomi/Working_Papers/CREATES/2021/rp21_03.pdf

Larcker, D., Seru, A., Tayan, B., & Lee, L-E. (2024). 2024 Institutional Investor Survey on Sustainability. *Stanford Graduate School of Business*. https://www.gsb.stanford.edu/faculty-research/publications/2024-institutional-investor-survey-sustainability

Lioui, A., & Tarelli, A. (2022). Chasing the ESG factor. *Journal of Banking & Finance*, Volume 139, 2022, 106498, ISSN 0378-4266. https://doi.org/10.1016/j.jbankfin.2022.106498

Naffa, H., & Fain, M. (2021): A factor approach to the performance of ESG leaders and laggards. *Finance Research Letters*, Volume 44, 2022, 102073, ISSN 1544-6123. https://doi.org/10.1016/j.frl.2021.102073

O'Connor, T. (2023): Fears ESG interest 'running out of steam' as record outflows recorded. *FTAdvisor,* Oct. 5th, 2023. https://www.ftadviser.com/investments/2023/10/05/fears-esg-interest-running-out-of-steam-as-record-outflows-recorded/

Ott, N. & Zincke, R. (2020): The Impact of ESG Ratings on Risk & Return of Firms: An empirical analysis, *Copenhagen Business School,* https://research.cbs.dk/en/studentProjects/the-impact-of-esg-ratings-on-risk-amp-return-of-firms-an-empirica

Pástor, L., Stambaugh, R., & Taylor, L. (2021). Sustainable investing in equilibrium. *Journal of Financial Economics*, Volume 142, Issue 2, 2021, pp. 550-571. https://doi.org/10.1016/j.jfineco.2020.12.011

Pedersen, L., Fitzgibbons, S., & Pomorski, L. (2021). Responsible Investing: The ESG-efficient frontier. *Journal of Financial Economics*, Volume 142, Issue 2, 2021, pp. 572-597. https://doi.org/10.1016/j.jfineco.2020.11.001

Reber, B., Gold, A. & Gold, S. (2021): ESG Disclosure and Idiosyncratic Risk in Initial Public Offerings. *J Bus Ethics* 179, 867–886 (2022). https://doi.org/10.1007/s10551-021-04847-8

Ross, S.A. (1976). The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory*, 13(3), pp. 341-360.

Sanctify ESG (2024). Scoring Documentation 1.1. *Sanctify.ai*. https://sanctify.ai/methodology/

Sharpe, W.F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, 19(3), pp. 425-442.

So, Mike K. P., Chu, Amanda M. Y., Lo, Cliff C., & Ip, Chun Yin (2021). Volatility and dynamic dependence modeling: Review, applications, and financial risk management. *Wires Computational Statistics*, vol. 14, issue 5, 2022. https://doi.org/10.1002/wics.1567

Xiao, C., Huang, W., & Newton, D.P. (2024). Predicting expected idiosyncratic volatility: Empirical evidence from ARFIMA, HAR, and EGARCH models. *Rev Quant Finan Acc* (2024). https://doi.org/10.1007/s11156-024-01279-z

Yamani, E., & Swanson, P. (2014). Financial crises and the global value premium: Revisiting Fama and French. *Journal of International Financial Markets, Institutions and Money*, Volume 33, 2014, pp. 115-136. https://doi.org/10.1016/j.intfin.2014.07.012

Zarafat, H., Liebhardt, S., & Eratalay, M. H. (2022). Do ESG Ratings Reduce the Asymmetry Behavior in Volatility? *Journal of Risk and Financial Management*, 15(8): 320. https://doi.org/10.3390/jrfm15080320

Appendix 1: Summary statistics

Res represents the summed residuals of the FF regressions, the daily idiosyncratic risk.
RV represents the monthly realized idiosyncratic volatility.
ESG, ENV, SOC, GOV represents the respective ratings.
Asterixes indicate significance at 10% (*), 5% (**) or 1% (***), respectively.

Table A1.1: Aggregate ESG

	Mean	Median	Min.	Max.	1st Quant.	3rd Quant.	St. Dev.	Skew	Kurtosis	ADF
High										
Res	0.04	-0.33	-260.12	211.83	-16.41	15.77	29.70	0.16	10.01	-12.69***
RV	124.89	115.03	63.23	490.37	95.81	134.99	53.91	3.88	24.64	-3.61**
ESG	0.40	0.33	0.05	1.15	0.27	0.49	0.24	1.19	4.11	-1.73
Neutral										
Res	-0.04	-0.78	-234.33	229.75	-20.71	20.31	35.93	0.05	6.21	-14.19***
RV	152.47	140.50	72.20	522.29	113.07	177.39	61.86	2.67	15.09	-3.14
ESG	0.00	0.00000	-0.00057	0.01104	0.00	0.01706	0.00220	3.45504	14.32	-2.47
Low										
Res	0.71	0.22	-230.31	196.50	-15.44	16.02	28.21	0.16	9.31	-11.42***
RV	119.68	114.57	46.22	456.60	92.27	131.91	48.70	3.63	25.28	-3.52**
ESG	-2.59	-2.9378	-3.6596	-0.3345	-3.2837	-2.3921	0.9713	1.2422	3.2457	-1.50

Table A1.2: Environmental (ENV)

	Mean	Median	Min.	Max.	1st Quant.	3rd Quant	St. Dev.	Skew	Kurtosis	ADF
High										
Res	-0.30	0.02	-174.38	173.47	-17.66	18.18	31.30	-0.05	5.48	-12.21***
RV	134.00	123.54	68.26	401.22	99.19	155.34	50.90	2.36	11.92	-3.45*
ENV	0.62	0.55	0.09	1.57	0.49	0.66	0.33	1.17	4.27	-2.70
Neutral										
Res	0.28	-0.94	-229.31	369.19	-21.66	21.17	37.16	0.70	11.10	-13.86***
RV	156.53	139.24	79.59	617.97	127.42	169.47	66.78	4.09	26.42	-3.55**
ENV	0.00114	0.0	0.0	0.00993	0.0	0.00053	0.00240	2.31	7.11	-3.07
Low										
Res	0.00	-1.01	-285.19	232.22	-16.55	15.61	29.25	0.04	11.35	-11.29***
RV	123.04	114.74	59.12	499.20	92.27	136.34	52.97	4.00	27.93	-3.28*
ENV	-1.33	-1.54	-2.04	-0.16	-1.79	-1.09	0.56	0.84	2.49	-2.07

Table A1.3: Social (SOC)

	Mean	Median	Min.	Max.	1st Quant.	3rd Quant	St. Dev.	Skew	Kurtosis	ADF
High										
Res	-0.60	0.34	-142.49	201.01	-16.90	16.47	29.74	0.04	6.57	-12.23^{***}
RV	126.60	117.32	61.93	434.02	97.41	145.54	50.23	3.29	19.28	-3.85^{**}
SOC	1.20	1.14	0.19	2.80	0.96	1.32	0.59	1.04	4.26	-2.85
Neutral										
Res	-0.04	-0.78	-234.33	229.75	-20.71	20.31	35.93	0.05	6.21	-14.19^{***}
RV	152.47	140.50	72.20	522.29	113.07	177.39	61.86	2.67	15.09	-3.14
SOC	0.0007	0.0	-0.0006	0.0110	0.0	0.0017	0.0022	3.46	14.32	-2.47
Low										
Res	0.71	0.22	-230.31	196.50	-15.44	16.02	28.21	0.16	9.31	-11.42^{***}
RV	119.68	114.57	46.22	456.60	92.27	131.91	48.70	3.63	25.28	-3.52^{**}
SOC	-2.59	-2.94	-3.66	-0.33	-3.28	-2.39	0.97	1.24	3.25	-1.50

Table A1.4: Governance (GOV)

	Mean	Median	Min.	Max.	1st Quant.	3rd Quant	St. Dev.	Skew	Kurtosis	ADF
High										
Res	0.04	-0.33	-260.12	211.83	-16.41	15.77	29.70	0.16	10.01	-12.69^{***}
RV	124.89	115.03	63.23	490.37	95.81	134.99	53.91	3.88	24.64	-3.61^{**}
GOV	0.40	0.33	0.05	1.15	0.27	0.49	0.24	1.19	4.11	-1.73
Neutral										
Res	-0.04	-0.78	-234.33	229.75	-20.71	20.31	35.93	0.05	6.21	-14.19^{***}
RV	152.47	140.50	72.20	522.29	113.07	177.39	61.86	2.67	15.09	-3.14
GOV	0.00069	0	-0.00057	0.01104	0	0.00017	0.00220	3.46	14.32	-2.47
Low										
Res	0.71	0.22	-230.31	196.50	-15.44	16.02	28.21	0.16	9.31	-11.42^{***}
RV	119.68	114.57	46.22	456.60	92.27	131.91	48.70	3.63	25.28	-3.52^{**}
GOV	-2.59	-2.94	-3.66	-0.33	-3.28	-2.39	0.97	1.24	3.25	-1.50

Appendix 2: GARCH-MIDAS results

This appendix presents results of various GARCH models. Results are divided according to pillars (ESG (A2.1), ENV (A2.2), SOC (A2.3), GOV (A2.4). For each pillar the models are noted as follows: GJR = benchmark GJR-GARCH(1,1), GM-RV = GARCH-MIDAS w. realized idiosyncratic volatility as covariate, GM-ESG/GM-ENV/GM-SOC/GM-GOV = GARCH-MIDAS w. respective ESG score as covariate, GM-RV-ESG = GARCH-MIDAS with both realized idiosyncratic volatility and respective ESG score as covariate.
Asterixes indicate significance at 10% (*), 5% (**) or 1% (***), respectively.

Table A2.1: Aggregate ESG

Model	Portfolio	μ	α	β	γ	m	θ_{RV}	$w2_{RV}$	θ_{ESG}	$w2_{ESG}$
GJR	*High*	-0.323	0.0348***	0.9615***	0.007	0.168	-	-	-	-
	Neutral	0.263	0.0151***	0.9807***	0.008	-0.232	-	-	-	-
	Low	-0.173	0.0577***	0.9425***	0.000	-0.035	-	-	-	-
GM-RV	*High*	0.155	0.0377***	0.959***	0.007	1.067	-0.010	4.002	-	-
	Neutral	0.153	0.0276***	0.9663***	0.012	1.538	-0.011	3.428	-	-
	Low	0.425	0.0468***	0.9474***	0.012	0.797	-0.009	3.601	-	-
GM-ESG	*High*	0.289	0.0288***	0.964***	0.014	0.013	-	-	-0.407	3.590
	Neutral	0.408	0.0166***	0.9757***	0.015	-0.020	-	-	0.228	3.370
	Low	0.408	0.0482***	0.9438***	0.016	0.962	-	-	9.690	1.000
GM-RV-ESG	*High*	0.141	0.0394***	0.9579***	0.005	1.305	-0.008	4.547	-0.675	2.744
	Neutral	0.223	0.0319***	0.9578***	0.021	2.043	-0.014	3.328	-1.842	3.151
	Low	0.465	0.044***	0.9497***	0.013	1.373	-0.010	3.301	3.846	1.820

Table A2.2: ENV

Model	Portfolio	μ	α	β	γ	m	θ_{RV}	$w2_{RV}$	θ_{ESG}	$w2_{ESG}$
GJR	*High*	0.3002	0.0203***	0.9722***	0.0148	0.0031	-	-	-	-
	Neutral	0.2654	0.0368***	0.9502***	0.0259	-0.1287	-	-	-	-
	Low	0.4471	0.0262***	0.9489***	0.0498	0.0596	-	-	-	-
GM-RV	*High*	0.5333	0.0503***	0.9462***	0.007	0.6868	-0.0107	3.3493	-	-
	Neutral	0.1332	0.0534***	0.9367***	0.0199	0.6416	-0.0064	3.1214	-	-
	Low	0.8041	0.0274***	0.948***	0.0494	0.9128	-0.0092	4.7211	-	-
GM-ESG	*High*	0.5999	0.0438***	0.9499***	0.0127	-0.1208	-	-	-0.7732	43.913
	Neutral	0.3004	0.041***	0.9432***	0.0314	-0.2994	-	-	0.1722	3.4373
	Low	0.7942	0.0225***	0.9478***	0.0594	3.2135	-	-	1.9093	2.598
GM-RV-ESG	*High*	0.537	0.0487***	0.9473***	0.008	1.0576	-0.0097	3.6149	-1.0635*	24.4565
	Neutral	0.2145	0.0423***	0.9441***	0.0272	0.5502	-0.0058	2.9865	-0.1596*	3.1009
	Low	0.6602	0.022***	0.9517***	0.0525	3.7474	-0.0132	3.3943	1.401*	4.1528

Table A2.3: SOC

Model	Portfolio	μ	α	β	γ	m	θ_{RV}	$w2_{RV}$	θ_{ESG}	$w2_{ESG}$
GJR	*High*	-0.3558	0.0368***	0.9573***	0.0118	0.0927	-	-	-	-
	Neutral	-0.108	0.036***	0.9553***	0.0173	0.2625	-	-	-	-
	Low	0.6778	0.0371***	0.9545***	0.0168	0.2576	-		-	-
GM-RV	*High*	0.1114	0.0355***	0.9533***	0.0223*	0.6581	-0.0118*	2.7713*	-	-
	Neutral	-0.4715	0.0458***	0.9526***	0.0033*	-0.495	-0.0061*	2.6739*	-	-
	Low	1.2635	0.0288***	0.9592***	0.0241*	1.3038	-0.0152*	3.7154*	-	
GM-ESG	*High*	0.1112	0.0355***	0.9543***	0.0204	-0.4439	-	-	-0.1087	1.0155
	Neutral	-0.3754	0.0363***	0.9592***	0.0089	-1.2338	-	-	1.4067	4.3454
	Low	1.384	0.0279***	0.9578***	0.0286	0.6609	-	-	0.2916	2.1598
GM-RV-ESG	*High*	0.0466	0.042***	0.9505***	0.0149	0.3446	-0.0073	3.5212	-0.1637	2.9188
	Neutral	-0.4546	0.0449***	0.9524***	0.0053	-0.0598	-0.0096	2.2618	0.6008	3.5543
	Low	1.298	0.0267***	0.9592***	0.0282	1.3065	-0.0176	3.4193	-0.0784	1.0352

Table A2.4: GOV

Model	Portfolio	μ	α	β	γ	m	θ_{RV}	$w2_{RV}$	θ_{ESG}	$w2_{ESG}$
GJR	*High*	0.367	0.0261^{***}	0.952^{***}	0.0438^{***}	0.492				
	Neutral	0.666	0.0261^{***}	0.9524^{***}	0.0429^{***}	0.020				
	Low	-0.279	0.0505^{***}	0.9448^{***}	0.0094^{***}	0.264				
GM-RV	*High*	0.766	0.0334^{***}	0.9474^{***}	0.0385^{***}	0.624	-0.008	3.374		
	Neutral	0.778	0.0278^{***}	0.9508^{***}	0.0429^{***}	0.725	-0.007	3.215		
	Low	0.109	0.0476^{***}	0.945^{***}	0.0149^{***}	0.575	-0.010	3.312		
GM-ESG	*High*	0.854	0.0258^{***}	0.9509^{***}	0.0465^{***}	0.158			-1.127	12.851
	Neutral	0.893	0.0305^{***}	0.9456^{***}	0.0478^{***}	-0.083			-0.039	3.111
	Low	0.104	0.05^{***}	0.9415^{***}	0.0171^{***}	1.557			0.924	2.202
GM-RV-ESG	*High*	0.754	0.0307^{***}	0.9489^{***}	0.0407^{***}	1.385	-0.012	3.035	-1.490	14.287
	Neutral	0.762	0.0298^{***}	0.9491^{***}	0.0423^{***}	0.856	-0.008	3.086	1.109	2.012
	Low	0.067	0.0476^{***}	0.9454^{***}	0.0142^{***}	1.796	-0.011	2.887	0.504	3.490

Appendix 3: Model comparisons

Comparisons between the performance of each model per pillar and per portfolio measured by Bayesian Information Criterion (BIC) and log-likelihood (LLH). Models marked in **bold** indicate the best performance for the respective measure (lowest value for BIC, highest value for LLH).

Table A3.1: Aggregate ESG

	High		**Neutral**		**Low**	
	BIC	LLH	BIC	LLH	BIC	LLH
GJR	19540	-9751	19671	-9817	19104	-9533
GM-RV	**17090**	**-8519**	**17283**	-8615	**16748**	-8348
GM-ESG	17108	-8528	17303	-8626	16766	-8357
GM-RV-ESG	17106	**-8519**	17295	**-8614**	16761	**-8347**

Table A3.2: ENV

	High		**Neutral**		**Low**	
	BIC	LLH	BIC	LLH	BIC	LLH
GJR	19398	-9680	19945	-9953	19004	-9483
GM-RV	**17002**	-8475	**17495**	**-8721**	**16666**	-8307
GM-ENV	17012	-8480	17503	-8725	16682	-8315
GM-RV-ENV	17014	**-8473**	17509	**-8721**	16673	**-8303**

Table A3.3: SOC

	High		Neutral		Low	
	BIC	LLH	BIC	LLH	BIC	LLH
GJR	19137	-9549	19805	-9883	18957	-9459
GM-RV	**16801**	**-8374**	**17397**	-8672	**16611**	-8280
GM-SOC	16815	-8381	17403	-8675	16642	-8295
GM-RV-SOC	16819	-8376	17410	**-8671**	16625	**-8279**

Table A3.4: GOV

	High		Neutral		Low	
	BIC	LLH	BIC	LLH	BIC	LLH
GJR	19055	-9508	19705	-9834	18807	-9385
GM-RV	**16732**	-8340	**17329**	**-8638**	**16491**	**-8219**
GM-GOV	16741	-8344	17341	-8644	16506	-8227
GM-RV-GOV	16742	**-8337**	17343	**-8638**	16505	**-8219**

Appendix 4: Fama-French factor portfolios

Table A4.1: Factors for Fama/French regression.
Each factor (besides Mkt-Rf) is created as diversified portfolios, and seeks to explain excess return from market factors. Mkt-Rf is the excess return on the market, as presented by CAPM. Please see dedicated references, as well as Kenneth French's website (French, 2024) for in-depth description and discussion of the portfolios and their creation.

Symbol	**Description**	**Reference**	**Mean**	**St. dev.**	**Min**	**1st quartile**	**Median**	**3rd quartile**	**Max**
Mkt-Rf	Market factor. Return on market minus the risk-free rate.	Sharpe (1964), Fama & French (1993, 2015)	0.0502	1.2049	-12.0	-0.4	0.06	0.6225	9.34
HML	Value. High minus low *book-to-market.*	Fama & French (1993, 2015)	0.0185	0.5445	-2.16	-0.2925	0.01	0.32	4.2
SMB	Size. Small minus big.	Fama & French (1993, 2015)	-0.0024	0.6910	-4.55	-0.4	-0.02	0.38	5.7
CMA	Investment. Conservative minus aggressive.	Fama & French (2015)	0.0031	0.4871	-2.73	-0.26	-0.01	0.26	2.46

RMW	Profitability. Robust minus weak.	Fama & French (2015)	-0.0030	0.9561	-5.02	-0.48	-0.04	0.46	6.73
MOM	Momentum. 'Winners minus losers'.	Jegadeesh & Titman (1993), Carhart (1997)	-0.0059	1.1616	-14.37	-0.52	0.06	0.6	5.93
Rf	Risk-free rate.	Fama & French (1993, 2015)	0.005742	0.006284	0	0	0.004	0.009	0.022