ORIGINAL PAPER

Mixtures of generalized normal distributions and EGARCH models to analyse returns and volatility of ESG and traditional investments

Pierdomenico Duttilo[1] · **Stefano Antonio Gattone**[2] · **Barbara Iannone**[2]

Abstract

Environmental, social and governance (ESG) criteria are increasingly integrated into investment process to contribute to overcoming global sustainability challenges. Focusing on the reaction to turmoil periods, this work analyses returns and volatility of several ESG indices and makes a comparison with their traditional counterparts from 2016 to 2022. These indices comprise the following markets: Global, the US, Europe and emerging markets. Firstly, the two-component mixture of generalized normal distribution was exploited to objectively detect financial market turmoil periods with the Naïve Bayes' classifier. Secondly, the EGARCH-in-mean model with exogenous dummy variables was applied to capture the turmoil period impact. Results show that returns and volatility are both affected by turmoil periods. The return–risk performance differs by index type and market: the European ESG index is less volatile than its traditional market benchmark, while in the other markets, the estimated volatility is approximately the same. Moreover, ESG and non-ESG indices differ in terms of turmoil periods impact, risk premium and leverage effect.

Keywords Mixtures of GND · Naïve Bayes' classifier · EGARCH model · Returns distribution · Volatility · ESG investments

✉ Pierdomenico Duttilo
  pierdomenico.duttilo@unich.it

  Stefano Antonio Gattone
  gattone@unich.it

  Barbara Iannone
  biannone@unich.it

1 University "G. d'Annunzio" of Chieti-Pescara, Viale Pindaro 42, Pescara 65127, Abruzzo, Italy

2 DISFIPEQ, University "G. d'Annunzio" of Chieti-Pescara, Viale Pindaro 42, Pescara 65127, Abruzzo, Italy

🙂 Springer

1 Introduction

The Paris Agreement and the 2030 Agenda for sustainable development demonstrate how the environmental, social and governance (ESG) challenges are widely complex (Friede 2019).

In the past decades, the investors' attention on the ESG factors has been growing, becoming a recurring theme of research in different disciplines (Abate et al. 2021). While several scholars claim that ESG investments mean giving up a portion of profits for ethical reasons (Bauer et al. 2007; Bauer and Smeets 2015; Fich et al. 2015; Arouri and Pijourlet 2017), other scholars state that behind such choices, there is much more (Nilsson 2009; Krosinsky and Robins 2012; Hemerijck 2018; Kaufer and Steponaitis 2019).

Traditional investors choose their investments by adopting economic–financial criteria, such as returns, investment duration, risk aversion, risk premium, liquidity and so on. By contrast, ESG investors combine the economic performance with measurable environmental and social factors (Revelli 2017; Benlemlih and Bitar 2018; Lapanan 2018; Oikonomou et al. 2018; Chatzitheodorou et al. 2019; Rossi et al. 2019; Gomes 2020). The latter "*recognize that the generation of long-term sustainable returns is dependent on stable, well-functioning, and well-governed social, environmental, and economic systems*" (University of Cambridge 2022).

However, traditional and ESG investors assess the market risk considering a mix of economic and socio-political factors or aggregate events that may cause financial turmoil periods and deep shocks to stock markets (Zigrand 2014; Berk and DeMarzo 2018; Szczygielski et al. 2021; Duttilo et al. 2021). The dotcom bubble burst (2000–2002), the global financial crisis (2007–2008), the European sovereign debt crisis (2010–2012) and the COVID-19 pandemic are examples of turmoil periods.

This work studied the effect of turmoil periods on the performance and volatility of several Dow Jones Sustainability Indices (DJSIs) and compared them with their respective market benchmarks (traditional indices).

For this purpose, two different tools have been employed: a finite mixture of two generalized normal distributions (MGND) and the exponential GARCH-in-mean (EGARCH-M) model with exogenous dummy variables.

The MGND model is a flexible tool able to capture important stylized facts of financial returns, e.g. excess kurtosis and skewness (Wen et al. 2020). It was applied to fit the return distribution of the MSCI All-Country World Equity Index (MSCIW). This index represents the performance of large- and mid-cap stocks of 23 developed and 24 emerging markets. Financial market turmoil periods were objectively detected by applying the Naïve Bayes' classifier to the mixture model results.

EGARCH-M model is commonly used to predict financial returns and their volatility (Hoti et al. 2007). The model was applied on several ESG and non-ESG indices including an exogenous dummy variable denoting the turmoil periods as detected by the MGND model.

The entire analysis provided insights on potential differences between ESG indices and their traditional market benchmarks investigating the impact of stable

and turmoil periods on the conditional mean and volatility and other interesting aspects of financial returns such as risk premium, leverage effect and volatility persistence.

These indices comprise the following markets: Global, the US, Europe (EU) and emerging markets (EM). Hence, Global and regional DJISs are selected: Dow Jones Sustainability World Index, Dow Jones Sustainability US Composite Index, Dow Jones Sustainability Europe Index and Dow Jones Sustainability Emerging Markets. Likewise, their respective traditional market benchmarks are collected: Dow Jones Global Index, Dow Jones Industrial Average Index, Dow Jones Europe Index and Dow Jones Emerging Markets Index. The time period taken for the study is from 4 January 2016 to 30 September 2022.

The rest of the paper is organized as follows. Section 2 includes the literature review and the contribution of the present work. Section 3 describes the methodology and the data used. Section 4 illustrates the results of the analysis. Finally, Sect. 5 provides some conclusions.

2 Literature review

2.1 Previous research

There is a considerable amount of research that analyses the ESG equity indices by comparing their financial performance with traditional equity benchmarks with many different methods, techniques, dates and variables. However, there is no absolute consensus on the fact that ESG portfolios are less volatile than market benchmark portfolios (Ouchen 2022). Cunha et al. (2020) distinguished three different kinds of results about the financial performance of ESG equity indices: positive, neutral and mixed. The latter refers to studies that "*found one or more positive, neutral or negative return-risk performance findings*". In order to highlight different methods and techniques used in this research field, the existing literature has been grouped into three main methodological frameworks: portfolio performance measures, Markov-switching (MS) models and GARCH models.

The first methodological framework is featured by works that applied portfolio performance measures such as Sharpe ratios, Treynor ratios and Jensen's alpha (Schröder 2007; Collison et al. 2008; Consolandi et al. 2009; Belghitar et al. 2014; Lean and Nguyen 2014; Cunha et al. 2020).

The second methodological framework stands out for the use of MS models to analyse the performance of ESG indices (Shunsuke et al. 2012; Ouchen 2022).

In the last methodological framework, GARCH models are employed to analyse the conditional mean and volatility of daily returns on ESG indices. Usually the conditional analysis is supported by the unconditional analysis which uses different methodologies such as the portfolio performance measures.

Employing GARCH models, Hoti et al. (2007) analysed the conditional volatility of some sustainability and ethical indices: Ethibel, ASPI Eurozone, Calvert Social Index, Ethical Index and FTSE4Good (Global, the USA, the UK and Europe). Results showed differences in the volatility persistence in the short and long run and

in the leverage effect. Lean and Nguyen (2014) analysed the performance and volatility of DJSIs for the Global and three regional markets over the period 2004–2013. The conditional analysis performed via the EGARCH model highlighted that the 2008 global financial crisis had a substantial impact on both return and volatility of sustainable investments. Similarly, Ang (2015) explored the behaviour of the DJSI Korea. Results showed that both return and volatility of Korea ESG portfolio are less affected by the 2008 crash. Using mean–variance testing and GARCH models, Sudha (2015) compared the performance and volatility of the S&P ESG India Index with two market benchmarks, i.e. the Nifty and the S&P CNX 500. Although the volatility clustering featured all three indices, it was found that S&P ESG India Index was less volatile compared to the Nifty. Jain et al. (2019) explored whether ESG investments offer better financial returns than the market benchmarks in developed and emerging markets. Applying GARCH models, the study concluded that the US large-cap ESG index (TRESGUS) provided the highest return and a suitable level of risk. Sabbaghi (2022) analysed the impact of good and bad news on the volatility of ESG firms employing the MSCI indices and the GARCH framework. Results showed the presence of the leverage effect for ESG firms, i.e. bad news increase volatility by a larger amount than good news.

2.2 Our contribution to existing literature

This research work could be located in the third methodological framework. To the best of our knowledge, this is the first work that employs mixtures of GND to objectively detect financial market turmoil periods and subsequently use them as dummy variable in the EGARCH(1,1)-M model. Whereas common approaches to assess the impact of a given period on returns and volatility include the use of subjectively identified dummy variables (Shehzad et al. 2020; Bora and Basistha 2021; Duttilo et al. 2021), sub-periods (Lean and Nguyen 2014; Ang 2015; Han et al. 2019) or Markov-switching models (Shunsuke et al. 2012; Ouchen 2022). Particularly, this study answers to the following questions: is the ESG index less or more volatile than its respective market benchmark? and then, are there differences among markets? Do turmoil periods have a significant impact on ESG indices? and then, are there differences with respect to traditional indices? Which market has been most affected by the turmoil periods? Do the ESG indices have a significant risk premium and leverage effect? and Do ESG and traditional indices have the same volatility persistence?

3 Data and methodology

3.1 Data

The MSCIW index is selected to detect stable and turmoil periods. This global traditional index provides a good representation of the global market and regional ones by the performance of large- and mid-cap stocks of 23 developed and 24 emerging markets.

The DJSI Index Family tracks the stock performance of leading sustainability-driven companies which stand out in terms of economic, environmental and social criteria (Hoti et al. 2007; S&P Global 2022). This index family was launched in 1999 as the first global sustainability benchmark. The DJSI Index Family is a "*best-in-class*" benchmarks because only the top ranked companies in the *S&P Global ESG Scores* are selected for inclusion in the Index Family (S&P Global 2022). Its composition is reviewed annually and rebalanced quarterly.

In this study, data on daily closing prices of traditional and ESG indices have been collected as shown in Table 1. The time period taken for the study is from 4 January 2016 to 30 September 2022. To underline the COVID-19 pandemic scenario, the considered period does not include the global financial crisis (2007–2009) and the sovereign debt crisis (2010–2011) periods.

Next, the daily returns of all equity indices under study were calculated with the natural log difference approach (Wen et al. 2020; Shehzad et al. 2020; Duttilo et al. 2021; Ouchen 2022)

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right)100, \tag{1}$$

where

r_t is the daily percentage return on equity index at time t;
P_t is the daily closing price of equity index at time t;
P_{t-1} is the daily closing price of equity index at time $t-1$.

3.2 Finite mixtures of GND

Mixtures of distributions are widely used to fit the empirical distribution of daily returns in order to capture important stylized facts such as excess kurtosis and skewness. There are some works (Behr and Pötter 2009; Bellalah and Lavielle 2002; Kon 1984; Han et al. 2019) that applied a mixtures of Gaussians (with two

Table 1 Selected traditional and ESG indices

Index	Ticker	Type	Area
MSCI All-Country World Equity Index	MSCIW	Traditional	Global
Dow Jones Global Index	W1DOW	Traditional	Global
Dow Jones Sustainability World Index	W1SGI	ESG	Global
Dow Jones Industrial Average	DJUS	Traditional	US
Dow Jones Sustainability US Composite Index	AASGI	ESG	US
Dow Jones Europe	E1DOW	Traditional	EU
Dow Jones Sustainability Europe Index	DJSEUR	ESG	EU
Dow Jones Emerging Markets Index	W5DOW	Traditional	EM
Dow Jones Sustainability Emerging Markets	DJSEMUP	ESG	EM

Source: *spglobal.com*, *finance.yahoo.com* and *investing.com*, reference date 20 October 2022

or three components) to capture the excess kurtosis and positive or negative skewness of daily returns distribution of common stocks and indices. However, mixtures of Gaussians impose a priori specific constraints on the form of the returns distribution. Thanks to the flexibility provided by the additional shape parameter v, the finite mixture of GND can overcome this critical issue. The contribution of Wen et al. (2020) explored the univariate mixture of GND and proposed an expectation conditional maximization (ECM) algorithm for parameter estimation. Moreover, a two-component mixture of GND and a two-component mixture of Gaussians were estimated on the S&P 500 and Shanghai Stock Exchange Composite Index daily returns. Results showed that the mixture of GND better describes the excess kurtosis and skewness of daily returns compared to mixtures of Gaussians.

A random variable R is said to have the generalized normal distribution with parameters μ (location), δ (scale) and v (shape) if its *probability density function* (p.d.f.) is given by

$$f(r_t|\mu, \delta, v) = \frac{v}{2\delta\Gamma(1/v)} \exp\left\{-\left|\frac{r_t - \mu}{\delta}\right|^v\right\}, \tag{2}$$

with $\Gamma(1/v) = \int_0^\infty t^{1/v-1} \exp^{-t} dt, -\infty < r_t < \infty, -\infty < \mu < \infty, \delta > 0, v > 0$.

Thanks to the shape parameter who determines the tails of the distribution, the GND distribution is a flexible tool to capture a large class of statistical distributions (Nadarajah 2005; Wen et al. 2020), for example with $v = 1$ and $v = 2$, GND becomes a normal and Laplace distribution, respectively.

The p.d.f. of the univariate mixture of GND is given by:

$$f(r_t|\theta) = \sum_{k=1}^{K} \pi_k p(r_t|\mu_k, \delta_k, v_k),$$
$$= \sum_{k=1}^{K} \frac{\pi_k v_k}{2\delta_k\Gamma(1/v_k)} \exp\left\{-\left|\frac{r_t - \mu_k}{\delta_k}\right|^{v_k}\right\}, \tag{3}$$

where $\theta = (\pi_k, \mu_k, \delta_k, v_k), \mu_k \in R, \delta_k > 0, v_k > 0, 0 < \pi_k < 1$ and $\sum_{k=1}^{K} \pi_k = 1$. If the random variable R has the p.d.f. as in Eq. (3), then the variance of R is given by

$$\sigma^2 = \sum_{k=1}^{K} \pi_k[\mu_k^2 + \sigma_k^2], \tag{4}$$

where

$$\sigma_k^2 = \delta_k^2 \Gamma(3/v_k)/\Gamma(1/v_k) \tag{5}$$

is the variance of the component k.

This model nests several distributions as its sub-models (Wen et al. 2020), namely according to the shape parameter value (v_k). For example, for $K = 2$, the univariate mixture of GND reduces to:

- The mixture of Gaussians when $\nu_1 = \nu_2 = 2$;
- The mixture of Laplace distributions when $\nu_1 = \nu_2 = 1$;
- The mixture of Gaussian and Laplace distributions when $\nu_1 = 2$ and $\nu_2 = 1$;
- The mixture of Gaussian and GND distributions when $\nu_1 = 2$ and $\nu_2 > 0$;
- The mixture of Laplace and GND distributions when $\nu_1 = 1$ and $\nu_2 > 0$.

In brief, the mixture of GND does not impose a priori specific constraint on the shape of each component of the mixture.

In this work, a two-component mixture of GND (with $K = 2$) was estimated on the MSCIW index returns via the ECM algorithm (Wen et al. 2020) to objectively detect turmoil periods. Specifically, it was assumed that there are two distinct stock market periods: stable and tumultuous. The former is *predominant* and less volatile than the latter which is characterized by *extreme behaviour* (Kim and White 2004). A daily return belongs to the stability period if it belongs to the stable component, i.e. the mixture component with the highest shape parameter. Similarly, a daily return belongs to the turmoil period if it belongs to the turmoil component, i.e. the mixture component with the lowest shape parameter. Conventionally, a smaller shape parameter means a thicker tail (excess kurtosis) and then a higher standard deviation, while a higher shape parameter means a thinner tail (mild kurtosis) and then a lower standard deviation. Shifts from the stable component to the turmoil component are assumed to be due to exogenous market events, i.e. "*time-ordered shifts*" (Kon 1984).

Kon (1984) assigned daily returns to a specific mixture component through the Naïve Bayes' classifier. Originally, this classification rule was proposed for mixtures of linear models by Kon and Lau (1979) and applied by Christie (1983) and Kon (1983). As explained by Kon (1984), this procedure "*may be particularly useful for efficient markets tests when the estimated data partition can be associated with corresponding public announcements or information signals in accounting numbers released prior to the event*" (Kon 1984). In other words, it may be particularly useful to detect important market events and test hypothesis.

The two-component mixture of GND was exploited to classify the daily returns based on the Naïve Bayes' classification rule that assigns each return to the class with the highest posterior probability (Frühwirth-Schnatter 2006). The simple classification rule is defined as follows:

$$\max_k \pi_k p(r_t | \theta_k), \tag{6}$$

where the selected mixture component k generating the return r_t has the largest posterior probability. Consequently, a decoded variable which is an indicator of market turmoil was obtained by a "*soft assignment*" (Bishop 2006) of observations to classes. In this way, financial market turmoil periods are objectively detected.

3.3 Exponential GARCH-in-mean model with exogenous dummy variables

Originally, the work of Engle (1982) introduced the autoregressive conditional heteroscedastic (ARCH) model giving rise to a vast literature and variety of

models. Next, in order to capture the risk premium, Engle et al. (1987) extended the ARCH model "*to allow the conditional variance to be a determinant of the mean*", it was called ARCH-in-mean model or ARCH-M. Subsequently, Nelson (1991) introduced the exponential ARCH model to overcome some limitations of the ARCH model. The EGARCH-M model is a generalization of the exponential ARCH model. These last two asymmetric models are able to capture the leverage effect, an important stylized fact of financial time series. The leverage effect occurs when negative returns increase volatility by a larger amount than positive returns (Francq and Zakoian 2019).

In this study, the conditional mean and volatility equations have been modelled with the EGARCH-M model. In both equations, an exogenous dummy variable was included to take into account the state of the financial market at time *t*. Specifically, the dummy variable *TURMOIL$_t$* assumes the value of 1 during turmoil periods; otherwise, it is equal to 0, i.e. during stability periods. In this way, it is possible to describe the impact of stability and turmoil periods on conditional mean and volatility of equity indices. The EGARCH(1,1)-in-mean model with exogenous dummy variables is specified as follows:

Conditional mean equation

$$r_t = \mu + m_1 TURMOIL_t + \phi_1 r_{t-1} + \lambda h_t + \epsilon_t, \tag{7}$$

Conditional volatility equation

$$\ln(h_t^2) = \omega + v_1 TURMOIL_t + \alpha_1 z_{t-1} + \gamma_1(|z_{t-1}| - E[|z_{t-1}|]) + \beta_1 \ln(h_{t-1}^2)$$

$$\text{where} \quad z_t = \frac{\epsilon_t}{\sqrt{h_t^2}} \sim \text{Skewed-GND}(0, 1, v, s). \tag{8}$$

In Eq. (7), r_t and ϵ_t indicate the returns and error terms of equity index at time t, respectively. Besides, μ is the constant term. The coefficient m_1 determines the impact of turmoil periods on the conditional mean. If m_1 is negative and statistically significant, turmoil periods caused a reduction in the conditional mean. To capture the autocorrelation of returns (i.e. the linear relationship between lagged values of returns time series), the conditional mean equation includes a stationary first-order autoregressive process AR(1) like Hoti et al. (2007). The coefficient ϕ_1 measures the time link between r_t and r_{t-1}. Following Engle et al. (1987), the conditional mean equation also includes the risk premium coefficient λ. If $\lambda > 0$ and statistically significant, returns are positively related to their conditional standard deviation (h_t).

With regard to the conditional volatility in Eq. (8), $\ln(h_t^2)$ denotes the natural logarithm of the conditional variance, ω is the constant term and α_1 captures the sign effect and γ_1 the size effect, while β_1 is the GARCH effect and the volatility persistence. $E[|z_{t-1}|]$ is the expected value of the absolute standardized residual. The coefficient v_1 determines the impact of turmoil periods on the conditional volatility. If v_1 is positive and statistically significant, turmoil periods caused an increasing in the conditional volatility. In order to better describe leptokurtosis and fatter tails of returns, the standardized residuals z_t were modelled using the

skewed-GND distribution with mean 0, variance 1, v and s as shape and skewness parameters, respectively.

The EGARCH(1,1)-M model with exogenous dummy variables was estimated for all indices under study through the *R*'s *rugarch* package (Ghalanos 2022).

4 Results

4.1 Results of exploratory data analysis

Table 2 presents the basic statistics of traditional and ESG indices. The mean of traditional and ESG indices is almost the same for Global and the US markets, while the mean of ESG indices is slightly higher than the mean of traditional indices for EU and EM markets. More importantly, in the EU market, the traditional index turns out to have a higher standard deviation than its ESG counterpart. In the other markets, traditional and ESG indices have approximately the same standard deviation. All indices show negative skewness and excess kurtosis. In general, ESG indices show lower skewness and kurtosis than non-ESG indices with only two exceptions. In the EU market, the level of skewness is approximately the same, while the ESG index has a higher kurtosis and skewness than the traditional one in the EM market.

Table 3 shows the results of some preliminary statistical hypothesis tests. According to the Jarque–Bera (JB) test, the daily returns are not normally distributed. The augmented Dickey–Fuller (ADF) test and Phillips–Perron (PP) test show that the null hypothesis of unit root can be rejected, and all indices are stationary in their first difference at 1% significance. The ARCH-LM test confirms the presence of ARCH effect and heteroscedasticity because the null hypothesis of no ARCH effect is rejected at 1% significance.

4.2 Turmoil period identification

Figure 1 shows that the two-component mixture of GND well describes the heavy-tailed and leptokurtic characteristics of the daily returns of the MSCIW index. Moreover, it confirms the presence of negative skewness given by a longer tail on the left.

Table 2 Estimated basic statistics of traditional and ESG indices

	Mean		Stdev		Skew		Kur	
	Trad.	ESG	Trad.	ESG	Trad.	ESG	Trad.	ESG
Global	0.0166	0.0182	0.9718	0.9815	−1.3752	−1.3488	18.5939	17.9680
US	0.0303	0.0337	1.2115	1.2305	−1.0819	−0.9107	23.0781	17.9696
EU	−0.0098	0.0042	1.2266	1.0289	−1.3371	−1.3499	16.8315	15.6846
EM	0.0070	0.0127	0.9860	1.0262	−0.8100	−1.0156	6.1659	9.2857

Table 3 Statistical tests results

Index	JB test	ADF test	PP test	ARCH-LM test
MSCIW	24,363**	−12**	−1844**	801**
W1DOW	25,064**	−12**	−1814**	797**
W1SGI	23,420**	−12**	−1788**	1063**
DJUS	38,114**	−12**	−2130**	762**
AASGI	23,144**	−12**	−2130**	782**
E1DOW	20,606**	−12**	−1711**	3406**
DJSEUR	17,970**	−12**	−1757**	2063**
W5DOW	2885**	−12**	−1548**	816**
DJSEMUP	6412**	−11**	−1533**	768**

** denotes p value significance at the 1% level

The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were applied to compare the goodness of fit performance and complexity among the Gaussian distribution, the GND, the mixture of two GND, the mixture of two Gaussian distributions and the mixture of a Gaussian and a Laplace distribution. The latter two are nested models of the finite mixture of GND when the shape parameter v_k takes on specific values (Sect. 3.2). Table 4 shows that both the AIC and BIC support the use of the two-component mixture of GND to fit the daily return distribution of MSCIW.

As mentioned in Sect. 3.2, stable and turmoil components could be identified on the basis of the shape parameter v. Looking at the estimated coefficients of the two-component mixture of GND (Eq. 3) in Table 5, a few interesting considerations arise. Firstly, the stable component ($\pi_1 = 0.7280$) is predominant compared to the turmoil component ($\pi_2 = 0.2720$). Secondly, the estimated mixture of GND is bi-modally asymmetric $\mu_1 = 0.1244 > \mu_2 = -0.1197$. Thirdly, the tails of the stable component intermediate between the Laplace and normal distributions $1 < v_1 = 1.1019 < 2$. On the other hand,

Table 4 AIC, BIC and log-likelihood (LL) of estimated models on MSCIW index

	AIC	BIC	LL
Mixture of GND	**4113.68**	**4151.74**	−2049.84
Mixture of Gaussian	4185.32	4212.51	−2087.66
Mixture of Gaussian–Laplace	4123.13	4155.32	−2059.07
Gaussian distribution	4764.38	4775.25	−2380.19
Generalized normal distribution	4637.69	4654.00	−2315.85

Values in bold indicate the best model according to AIC and BIC

Table 5 Estimated parameters of the two-component mixture of GND on MSCIW index

Stable component				Turmoil component			
π_1	μ_1	δ_1	v_1	π_2	μ_2	δ_2	v_2
0.7280	0.1244	0.5735	1.1019	0.2720	−0.1197	0.4523	0.6977

Table 6 Estimated standard deviation and excess kurtosis of the two-component mixture of GND on MSCIW index

Stable component		Turmoil component			
σ_1	κ_1	σ_2	κ_2	σ	κ
0.6978	2.2652	1.4300	8.1379	0.9605	13.0891

$\sigma_k = \sqrt{\sigma_k^2}$, $\sigma = \sqrt{\sigma^2}$ and $\kappa_k = \Gamma(5/\nu_k)\Gamma(1/\nu_k)/\Gamma(3/\nu_k)^2 - 3$

See Wen et al. (2020) for κ
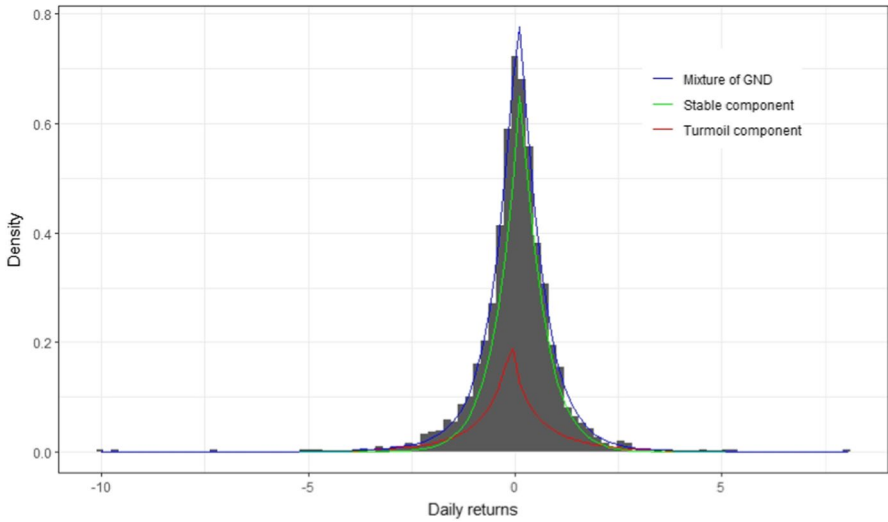
Fig. 1 Estimated density of daily returns on MSCIW index

tails of the turmoil component are more extreme than those of the Laplace distribution because $\nu_2 = 0.6977 < 1$. A smaller ν_k means a thicker tail and then a higher standard deviation (Eq. 5). Conversely, a higher ν_k means a thinner tail and then a lower standard deviation. Table 6 shows that the standard deviation of the stable component $\sigma_1 = 0.6978$ is less than the standard deviation of turmoil component $\sigma_2 = 1.4300$. The excess kurtosis of turmoil component $\kappa_2 = 8.1379$ is higher than the excess kurtosis of the stable component $\kappa_1 = 2.2652$ which have a "mild kurtosis". The combination of the excesses of kurtosis of both components fits the largest excess kurtosis $\kappa = 13.0891$ of daily returns.

Figure 2 illustrates daily returns of the MSCIW index. Grey vertical lines identify turmoil periods detected by the two-component mixture of GND and the Naïve Bayes' classifier. It can be seen that the COVID-19 pandemic constitutes the most turbulent period in terms of timing and returns fluctuation.
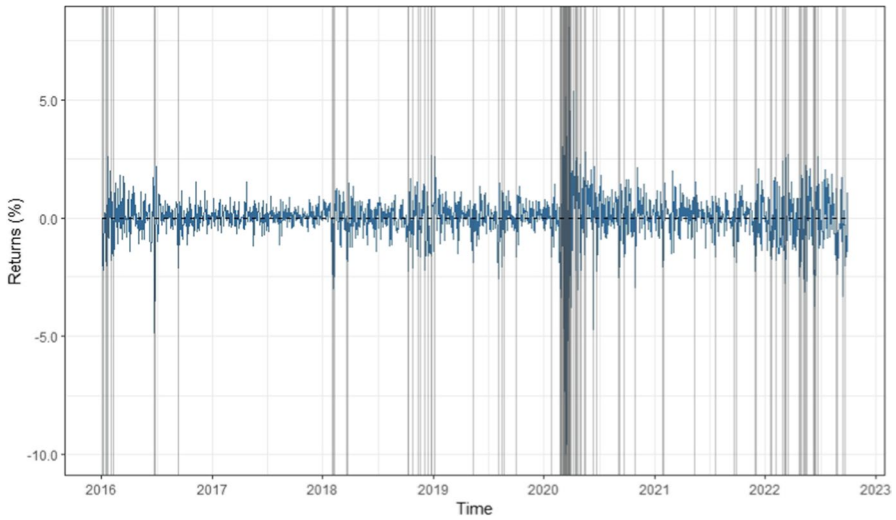
Fig. 2 Daily returns on MSCIW index and turmoil periods (grey vertical lines)

4.3 Estimated exponential GARCH-in-mean model with exogenous dummy variables

4.3.1 Results of conditional mean equation

Table 7 shows the estimated coefficients of the conditional mean equation. The constant μ is negative and statistically significant for all indices, except for the DJSE-MUP index. Turmoil periods caused a decrease in the conditional mean because the coefficient m_1 is negative and statistically significant for all indices. Figure 3 provides an easily interpretable view of the turmoil periods impact on the conditional
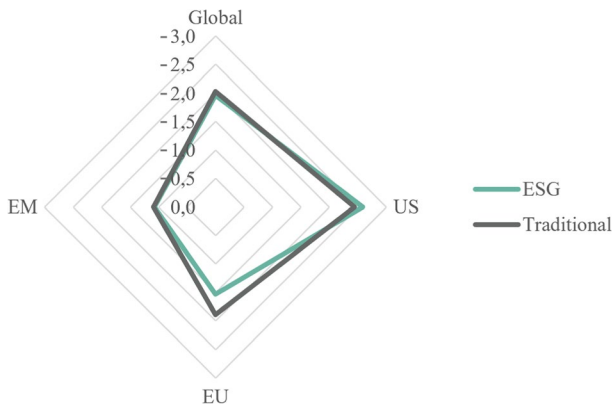
Fig. 3 Impact of turmoil periods on the conditional mean (m_1 coefficient) by index type and market

mean by index type and market. The magnitude of the impact is higher in the US market, which is followed by the Global, EU and EM markets. Results showed no significant differences between ESG and non-ESG indices except for the EU market where the conditional mean of the traditional index is more impacted than the conditional mean of the ESG index.

The AR(1) coefficient ϕ_1 is statistically significant for all indices except for the DJUS. In addition, it is positive for the Global and EM indices but negative for the US and EU indices. These different signs may be related to the spurious consequences of the non-synchronous trading among the index's component stocks (Campbell et al. 1997; Koutmos 1997; Basher et al. 2007).

Moreover, the results reveal a statistically significant coefficient of the risk premium for all indices. The ESG indices have a higher risk premium than traditional indices in the EU and EM markets.

4.3.2 Results of conditional volatility equation

Table 8 illustrates the estimated coefficients of the conditional volatility equation. The constant ω is negative and statistically significant for all indices. Turmoil periods caused an increase in the conditional volatility because the coefficient v_1 is positive and statistically significant for all indices. According to Fig. 4, the magnitude of the impact varies by index type and market. It is higher in the US market, followed by the Global, EU and EM markets. A more pronounced impact on the conditional volatility has been found for the traditional indices in the Global and EM markets.

Coefficients α_1 and β_1 are statistically significant highlighting the presence of ARCH and GARCH effects for all indices. Additionally, the conditional volatility is characterized by a negative leverage effect given by the negative sign of α_1 and the statistical significance of γ_1. Generally, the leverage effect is lower for ESG indices compared to their market benchmark, except for the US and EM markets which exhibit a reverse pattern.
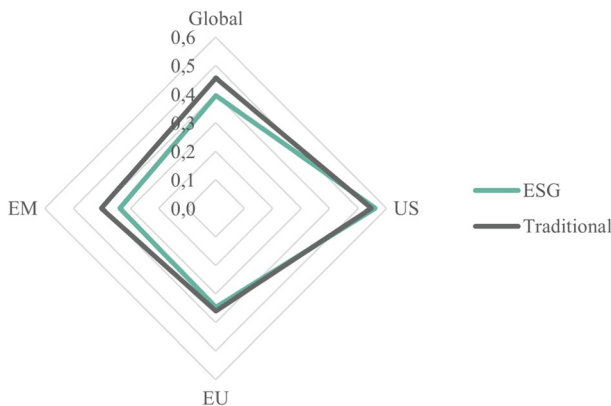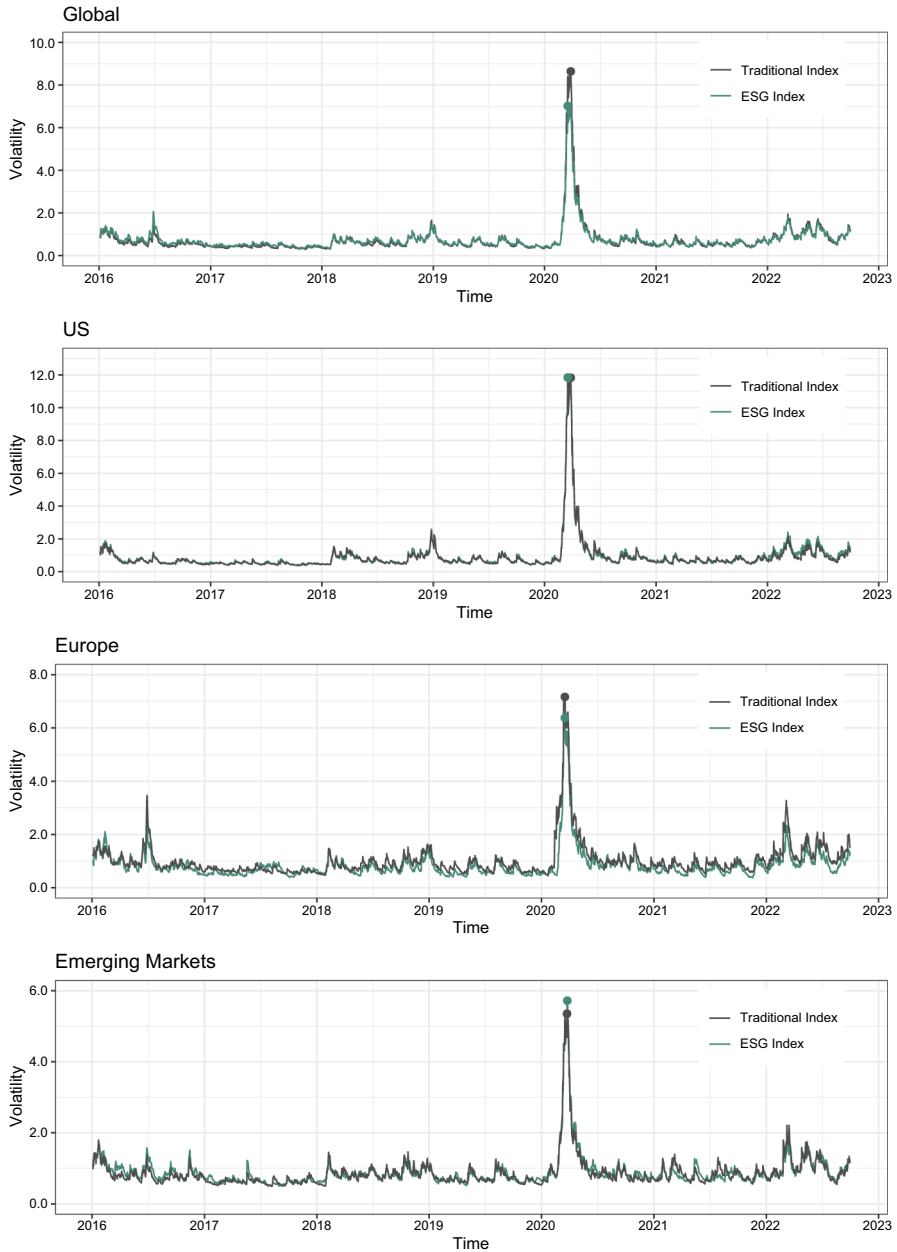

Fig. 4 Impact of turmoil periods on the conditional volatility (v_1 coefficient) by index type and market

Volatility shocks have "*long memory*" because the volatility persistence (β_1 coefficient) is close to 1 for all indices. However, it is necessary to consider the estimated (conditional) volatility and its distribution for risk assessment purposes. Figure 5

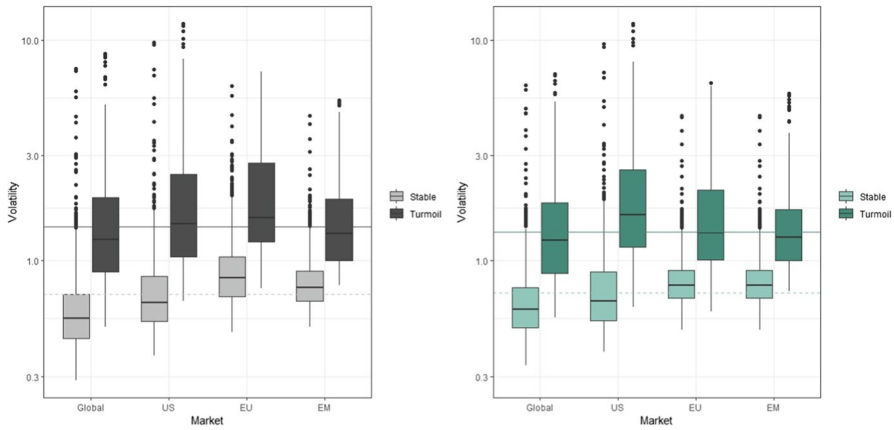**Fig. 5** Estimated conditional volatility by index type and market

Fig. 6 Distribution of the estimated conditional volatility by index type and market

Table 7 EGARCH: estimated coefficients of the conditional mean equation

Index	μ	m_1	ϕ_1	λ_1
W1DOW	−0.1047**	−2.0329**	0.1083**	0.3435**
W1SGI	−0.1159**	−1.9621**	0.0865**	0.3352**
DJUS	−0.1275**	−2.4445**	−0.0361	0.3586**
AASGI	−0.1425**	−2.5841**	−0.0756**	0.3746**
E1DOW	−0.1008**	−1.8971**	−0.0569**	0.1998**
DJSEUR	−0.1413**	−1.5315**	−0.0525*	0.2771**
W5DOW	−0.0917**	−1.0966**	0.1641**	0.1903**
DJSEMUP	−0.1844	−1.0750**	0.1675**	0.3159*

* and ** denote p value significance at the 5% and 1% levels, respectively

Table 8 EGARCH: estimated coefficients of the conditional volatility equation

Index	ω	v_1	α_1	γ_1	β_1	v	s
W1DOW	−0.0850**	0.4574**	−0.1153**	0.1404**	0.9391**	1.7114**	1.0115**
W1SGI	−0.0728**	0.3956**	−0.1185**	0.1275**	0.9398**	1.5935**	1.0005**
DJUS	−0.0764**	0.5450**	−0.1020**	0.1527**	0.9268**	1.5195**	1.0117**
AASGI	−0.0738**	0.5584**	−0.1098**	0.1640**	0.9242**	1.5978**	0.9906**
E1DOW	−0.0327**	0.3595**	−0.1023**	0.1942**	0.9348**	1.2096**	0.9230**
DJSEUR	−0.0564**	0.3459**	−0.1616**	0.1046**	0.9323**	1.3622**	0.8717**
W5DOW	−0.0601**	0.4020**	−0.1020**	0.1348**	0.9105**	1.6336**	0.9065**
DJSEMUP	−0.0489**	0.3364**	−0.0520**	0.1699**	0.9227**	1.4946**	0.9104**

** denotes p value significance at the 1% level

shows the estimated volatility by index type and market. The graphical examination suggests some interesting considerations.

Global: The difference in terms of estimated volatility between the ESG and traditional indices is slight. During the financial crisis due to the COVID-19 pandemic, the traditional index exhibits a highest peak of estimated volatility (8.64) compared to the ESG index (7.02).

The US: The estimated volatility of the ESG and traditional index is approximately the same. During the financial crisis due to the COVID-19 pandemic, the ESG index exhibits the same peak of estimated volatility (11.84) compared to the traditional index (11.82).

Europe: The estimated volatility of the ESG and traditional index is different because the latter is higher than the former. During the financial crisis due to COVID-19 pandemic, the traditional index exhibits a highest peak of estimated volatility (7.17) compared to the ESG index (6.38). In addition, the time series highlight other two pronounced volatility peaks: The first occurs during the Brexit vote (June 2016), while the second occurs during the Russia–Ukraine War (February 2022).

Emerging Markets: The estimated volatility of the ESG and traditional index is approximately the same for almost the entire study period. During the COVID-19 pandemic, the ESG index exhibits a slightly highest peak of estimated volatility (5.72) compared to the traditional index (5.35).

These findings are also confirmed by the distribution of the estimated volatility in Fig. 6. It is possible to infer that stable periods are characterized by low volatility, while turmoil periods by high volatility, for both index types (in line with the assumptions in Sect. 3.2). Specifically, the median volatility of both index types is approximately the same during stable periods (dashed lines). In contrast, the median volatility of traditional indices is higher than that of ESG indices during turmoil periods (straight lines), 1.43 against 1.34.

4.4 Diagnostic test results

The skewed-GND distribution suitably captured the leptokurtosis and the skewness of the standardized error terms (Eq. 8) because the coefficients v and s are statistically significant for all indices (Table 8).

In order to ascertain the absence of autocorrelation and heteroscedasticity in the standardized residuals of the models, the weighted Ljung–Box test and the weighted ARCH-LM test are performed. Table 9 illustrates that the standardized residuals are not affected by both the serial correlation and heteroskedasticity because all test statistics have p values> 0.05.

5 Conclusions

This study adds to the literature as it investigates the reactions of ESG and traditional indices to turmoil periods.

Table 9 EGARCH: diagnostic test results

Index	Weighted Ljung–Box test		Weighted ARCH-LM test	
	Test statistic	*p* value	Test statistic	*p* value
W1DOW	1.8554	0.7525	0.4569	0.4991
W1SGI	2.3509	0.6093	0.1420	0.7063
DJUS	0.3955	0.9971	0.0048	0.9455
AASGI	1.5563	0.8327	1.0320	0.3097
E1DOW	3.0091	0.4290	0.1228	0.7261
DJSEUR	2.6158	0.5337	6e-05	0.9934
W5DOW	2.6954	0.5116	0.1250	0.7237
DJSEMUP	1.6800	0.8006	0.7300	0.3929

Firstly, a finite mixture of two GND was estimated on the MSCIW index returns. According to the initial hypotheses (Sect. 3.2), findings show that the two-component mixture of GND well describes the heavy-tailed and leptokurtic characteristics of the daily returns. Similarly to Kim and White (2004) and Wen et al. (2020), the estimated mixture contains a "*predominant*" stable component and a "*relatively rare*" turmoil component. The former is characterized by a high shape parameter that implies a thinner tail and then a lower standard deviation, while the latter is featured by a small shape parameter that implies a thicker tail and then a higher standard deviation. Consequently, the mixture model results were exploited to objectively detect financial market turmoil periods.

Secondly, turmoil periods were included as exogenous dummy variables in the EGARCH-in-mean model to capture their impact on the conditional mean and conditional volatility of several ESG and non-ESG Dow Jones indices. Results show that the return–risk performance as well as the impact of turmoil periods on return and volatility vary among ESG and traditional indices and across markets. Like Hoti et al. (2007), Lean and Nguyen (2014) and Ang (2015) mixed return–risk findings (Cunha et al. 2020) have been obtained.

Whereas the European ESG index results to be less volatile than its traditional market benchmark, the estimated conditional volatility is approximately the same in the other markets.

The turmoil periods impact on the conditional mean and volatility of both index types shows a similar pattern: The most affected market is the US followed by the Global, EU and EM markets. However, the conditional mean of ESG indices is less affected in the EU market, while they are more resilient in the Global and EM markets in terms of conditional volatility.

Other interesting key aspects concern the risk premium and leverage effect. ESG indices have a higher risk premium than traditional indices in the EU and EM markets. As identified by Sabbaghi (2022), results confirm the presence of the leverage effect for ESG indices. The leverage effect is lower for ESG indices compared to their market benchmark, except for the US market and EMs which exhibit a reverse pattern.

Findings are important to equity portfolio managers, investors, policy-makers, academics and anyone who decide to encourage ESG investments. In order to contain the risk during turmoil periods, portfolio managers may apply rebalancing strategies placing higher weight on the less volatile asset. Besides, they may obtain more investment information about decisions, in terms of diversifying the risks considering also the turmoil period impact, risk premium and leverage effect (Markowitz 1952). Lastly, corporate executives yet shall use it to benchmark their own results against peers and track news as well.

It is important to note that the results obtained depend on the definition of the turmoil period (Sect. 4.2). For example, considering market specific turmoil periods, larger differences between ESG and non-ESG indices might be observed. Additionally, the Naïve Bayes' classification rule makes the assignment without considering the temporal dependence of the returns. It would be interesting to explore a dynamic time-varying estimation of the mixture model.

To conclude, another extension of this work could be the study of the evolution of correlations among ESG and traditional indices focusing on the impact of turmoil periods and spillover effects among markets.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Abate, G., Basile, I., Ferrari, P.: The level of sustainability and mutual fund performance in Europe: an empirical analysis using esg ratings. Corp. Soc. Responsib. Environ. Manag. **28**(5), 1446–1455 (2021). https://doi.org/10.1002/csr.2175

Ang, W.R.: Sustainable investment in Korea does not catch a cold when the united states sneezes. J. Sustain. Finance Invest. **5**(1–2), 16–26 (2015). https://doi.org/10.1080/20430795.2015.1042737

Arouri, M., Pijourlet, G.: Csr performance and the value of cash holdings: international evidence. J. Bus. Ethics **140**(2), 263–284 (2017). https://doi.org/10.1007/s10551-015-2658-5

Basher, S.A., Hassan, M.K., Islam, A.M.: Time-varying volatility and equity returns in Bangladesh stock market. Appl. Fin. Econ. **17**(17), 1393–1407 (2007). https://doi.org/10.1080/09603100600771034

Bauer, R., Smeets, P.: Social identification and investment decisions. J. Econ. Behav. Organ. **117**(C), 121–134 (2015)

Bauer, R., Derwall, J., Otten, R.: The ethical mutual fund performance debate: new evidence from Canada. J. Econ. Behav. Organ. **70**(2), 111–124 (2007). https://doi.org/10.1007/s10551-006-9099-0

Behr, A., Pötter, U.: Alternatives to the normal model of stock returns: Gaussian mixture, generalised logf and generalised hyperbolic models. Ann. Finance **5**, 49–68 (2009). https://doi.org/10.1007/s10436-007-0089-8

Belghitar, Y., Clark, E., Deshmukh, N.: Does it pay to be ethical? Evidence from the ftse4good. J. Bank. Finance **47**, 54–62 (2014). https://doi.org/10.1016/j.jbankfin.2014.06.027

Bellalah, M., Lavielle, M.: A decomposition of empirical distributions with applications to the valuation of derivative assets. Multinatl. Finance J. **6**(2), 99–130 (2002)

Benlemlih, M., Bitar, M.: Corporate social responsibility and investment efficiency. J. Bus. Ethics **148**, 647–671 (2018). https://doi.org/10.1007/s10551-016-3020-2

Berk, J., DeMarzo, P.: Finanza aziendale 1 Fondamenti, 4th edn (2018). Pearson, https://books.google.it/books?id=a3ViwkwYCKQC

Bishop, C.M.: Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag, Berlin, Heidelberg (2006)

Bora, D., Basistha, D.: The outbreak of covid-19 pandemic and its impact on stock market volatility: evidence from a worst-affected economy. J. Publ. Affairs **n/a**(n/a), e2623 (2021). https://doi.org/10.1002/pa.2623

Campbell, J.Y., Lo, A.W., MacKinlay, A.: The Econometrics of Financial Markets. Princeton University Press (1997), http://www.jstor.org/stable/j.ctt7skm5

Chatzitheodorou, K., Skouloudis, A., Evangelinos, K., et al.: Exploring socially responsible investment perspectives: a literature mapping and an investor classification. Sustain. Prod. Consump. **19**, 117–129 (2019). https://doi.org/10.1016/j.spc.2019.03.006

Christie, A.: On Information Arrival and Hypothesis Testing in Event Studies. Tech. rep., University of Rochester, working paper (1983)

Collison, D.J., Cobb, G., Power, D.M., et al.: The financial performance of the FTSE4Good indices. Corp. Soc. Responsib. Environ. Manag. **15**(1), 14–28 (2008). https://doi.org/10.1002/csr.144

Consolandi, C., Jaiswal-Dale, A., Poggiani, E., et al.: Global standards and ethical stock indexes: the case of the Dow Jones sustainability Stoxx index. J. Bus. Ethics **87**(1573–0697), 185–197 (2009). https://doi.org/10.1007/s10551-008-9793-1

Cunha, F.AFd.S., de Oliveira, E.M., Orsato, R.J., et al.: Can sustainable investments outperform traditional benchmarks? Evidence from global stock markets. Bus. Strategy Environ. **29**(2), 682–697 (2020). https://doi.org/10.1002/bse.2397

Duttilo, P., Gattone, S.A., Di Battista, T.: Volatility modeling: an overview of equity markets in the Euro area during covid-19 pandemic. Mathematics (2021). https://doi.org/10.3390/math9111212

Engle, R.F.: Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica **50**(4), 987–1007 (1982)

Engle, R.F., Lilien, D.M., Robins, R.P.: Estimating time varying risk premia in the term structure: the arch-m model. Econometrica **55**(2), 391–407 (1987)

Fich, E.M., Harford, J., Tran, A.L.: Motivated monitors: the importance of institutional investors' portfolio weights. J. Financ. Econ. **118**(1), 21–48 (2015). https://doi.org/10.1016/j.jfineco.2015.06.014

Francq, C., Zakoian, J.M.: GARCH Models: Structure, Statistical Inference and Financial Applications, 2nd edn. John Wiley & Sons Ltd (2019)

Frühwirth-Schnatter, S.: Finite Mixture and Markov Switching Models. Springer Science, New York, USA (2006)

Friede, G.: Why don't we see more action? A metasynthesis of the investor impediments to integrate environmental, social, and governance factors. Bus. Strat. Environ. **28**(6), 1260–1282 (2019). https://doi.org/10.1002/bse.2346

Ghalanos, A.: Rugarch: Univariate GARCH models. R package version 1.4-7 (2022)

Gomes, O.: Optimal growth under socially responsible investment: a dynamic theoretical model of the trade-off between financial gains and emotional rewards. Int. J. Corp. Soc. Respons. **5**(1), 5 (2020). https://doi.org/10.1186/s40991-020-00049-z

Han, L., Yan, H., Zheng, C.: Normal mixture method for stock daily returns over different sub-periods. Commun. Stat. Simul. Comput. **48**(2), 447–457 (2019). https://doi.org/10.1080/03610918.2017.1383423

Hemerijck, A.: Social investment as a policy paradigm. J. Eur. Publ. Pol. **25**(6), 810–827 (2018). https://doi.org/10.1080/13501763.2017.1401111

Hoti, S., McAleer, M., Pauwels, L.L.: Measuring risk in environmental finance. J. Econ. Surv. **21**(5), 970–998 (2007). https://doi.org/10.1111/j.1467-6419.2007.00526.x

Jain, M., Sharma, G.D., Srivastava, M.: Can sustainable investment yield better financial returns: a comparative study of ESG indices and MSCI indices. Risks (2019). https://doi.org/10.3390/risks7010015

Kaufer, K., Steponaitis, L.: Just money: From ego-system to ecosystem finance. In: Just Money: Mission-Driven Banks and the Future of Finance. The MIT Press (2019), https://doi.org/10.7551/mitpress/12582.003.0006

Kim, T.H., White, H.: On more robust estimation of skewness and kurtosis. Finance Res. Lett. **1**(1), 56–73 (2004). https://doi.org/10.1016/S1544-6123(03)00003-5

Kon, S.J.: The market-timing performance of mutual fund managers. The J. Bus. **56**(3), 323–347 (1983)

Kon, S.J.: Models of stock returns-a comparison. The J. Finance **39**(1), 147–165 (1984)

Kon, S.J., Lau, W.P.: Specification tests for portfolio regression parameter stationarity and the implications for empirical research. The J. Finance **34**(2), 451–465 (1979). https://doi.org/10.1111/j.1540-6261.1979.tb02108.x

Koutmos, G.: Feedback trading and the autocorrelation pattern of stock returns: further empirical evidence. J. Int. Money Finance **16**(4), 625–636 (1997). https://doi.org/10.1016/S0261-5606(97)00021-1

Krosinsky, C., Robins, N.: Sustainable Investing: The Art of Long-Term Performance. Environmental Market Insights, Taylor & Francis Group (2012) , https://books.google.it/books?id=a3ViwkwYCKQC

Lapanan, N.: The investment behavior of socially responsible individual investors. The Quart. Rev. Econ. Finance **70**, 214–226 (2018). https://doi.org/10.1016/j.qref.2018.05.014

Lean, H.H., Nguyen, D.K.: Policy uncertainty and performance characteristics of sustainable investments across regions around the global financial crisis. Appl. Fin. Econ. **24**(21), 1367–1373 (2014). https://doi.org/10.1080/09603107.2014.925063

Markowitz, H.: Portfolio selection. The J. Finance **7**(1), 77–91 (1952)

Nadarajah, S.: A generalized normal distribution. J. Appl. Stat. **32**(7), 685–694 (2005). https://doi.org/10.1080/02664760500079464

Nelson, D.B.: Conditional heteroskedasticity in asset returns: a new approach. Econometrica **59**(2), 347–70 (1991)

Nilsson, J.: Segmenting socially responsible mutual fund investors: the influence of financial return and social responsibility. Int. J. Bank Market. **27**, 5–31 (2009). https://doi.org/10.1108/02652320910928218

Oikonomou, I., Platanakis, E., Sutcliffe, C.: Socially responsible investment portfolios: Does the optimization process matter? The Br. Account. Rev. **50**(4), 379–401 (2018)

Ouchen, A.: Is the ESG portfolio less turbulent than a market benchmark portfolio? Risk Manag. **24**(1), 1–33 (2022). https://doi.org/10.1057/s41283-021-00077-4

Revelli, C.: Socially responsible investing (sri): From mainstream to margin? Res. Int. Bus. Finance **39**, 711–717 (2017). https://doi.org/10.1016/j.ribaf.2015.11.003

Rossi, M., Sansone, D., van Soest, A., et al.: Household preferences for socially responsible investments. J. Bank. Finance **105**, 107–120 (2019). https://doi.org/10.1016/j.jbankfin.2019.05.018

Sabbaghi, O.: The impact of news on the volatility of ESG firms. Glob. Finance J. **51**(100), 570 (2022). https://doi.org/10.1016/j.gfj.2020.100570

Schröder, M.: Is there a difference? The performance characteristics of SRI equity indices. J. Bus. Finance Account. **34**(1–2), 331–348 (2007). https://doi.org/10.1111/j.1468-5957.2006.00647.x

Shehzad, K., Xiaoxing, L., Kazouz, H.: Covid-19's disasters are perilous than global financial crisis: A rumor or fact? Finance Res. Lett. **36**(101), 669 (2020). https://doi.org/10.1016/j.frl.2020.101669

Shunsuke, M., Tatsuyoshi, O., Akimi, M.: Do socially responsible investment indexes outperform conventional indexes? Appl. Fin. Econ. **22**(18), 1511–1527 (2012). https://doi.org/10.1080/09603107.2012.665593

S&P Global, DJSI Index Family, S&P Global. https://www.spglobal.com/esg/performance/indices/djsi-index-family. Accessed 10 Jan 2022

Sudha, S.: Risk-return and volatility analysis of sustainability index in India. Environ. Dev. Sustain. **17**(6), 1329–1342 (2015). https://doi.org/10.1007/s10668-014-9608-8

Szczygielski, J.J., Bwanya, P.R., Charteris, A., et al.: The only certainty is uncertainty: an analysis of the impact of covid-19 uncertainty on regional stock markets. Finance Res. Lett. **43**(101), 945 (2021). https://doi.org/10.1016/j.frl.2021.101945

University of Cambridge, What is responsible investment? https://www.cisl.cam.ac.uk/business-action/sustainable-finance/investment-leaders-group/what-is-responsible-investment, accessed: March 16, 2022

Wen, L., Qiu, Y., Wang, M., et al.: Numerical characteristics and parameter estimation of finite mixed generalized normal distribution. Commun. Stat. Simul. Comput. (2020). https://doi.org/10.1080/03610918.2020.1720733

Zigrand, J.P.: Systems and systemic risk in finance and economics (2014). http://eprints.lse.ac.uk/id/eprint/61220