ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna

DEPARTMENT OF ECONOMICS

The evaluation of the effects of ESG scores on financial markets

Michele Costa

Quaderni - Working Paper DSE N°1189

The evaluation of the effects of ESG scores on financial markets

Michele Costa

Abstract We aim to explore the interplay between ESG scores and assets characteristics, specifically focusing on volatility. We classify stocks on the basis of both high/low ESG and high/low ESG momentum and we evaluate ESG effects by measuring the distance between the 2 group distributions. The analysis of stocks within the STOXX Europe 600 Index from 2017 to 2022 suggests that companies with higher ESG tend to exhibit lower volatility. However, we haven't observed a similar trend when examining ESG momentum. Furthermore, our findings enable us to highlight and compare the effects associated with the COVID pandemic and the conflict in Ukraine.

Key words: Sustainable finance, ESG, Stock market risk, Volatility

JEL classification: G11; C40; Q56

Michele Costa
Department of Economics, University of Bologna, e-mail: michele.costa@unibo.it

Non-technical summary

The exploration into the relationship between Environmental, Social, and Governance (ESG) scores and stock attributes unfolds within the context of market dynamics. ESG, a linchpin in assessing sustainability, has spurred a surge in empirical scrutiny. Our study delves into this discourse, aiming to decipher the interplay between ESG scores and risk in stocks. The relationship between ESG and asset characteristics is multi-faceted and not straightforward. When examining the connection between ESG scores and financial performance, it's important to note that high ESG scores may indeed contribute to better performance. Nonetheless, high-performing companies might also possess more resources, allowing them to improve their ESG scores further.

Employing a data-based approach, our analysis circumvents investor preferences, homing in directly on the nexus between ESG and asset traits. Notably, we introduce inequality decomposition methods, integrating information provided by traditional mean-based evaluations to encompass distribution-based insights. Examining companies in the STOXX Europe 600 Index from 2017 to 2022, our study offers powerful insight on the interplay between ESG and assets volatility.

Results Overview:

Analyzing ESG-scored asset groups, we discern a compelling trend: higher ESG correlates with lower volatility. Peaks in volatility during crises (e.g., COVID pandemic, Ukraine war) impact lower ESG-scored assets more significantly, unveiling nuanced market dynamics.

Inequality Decomposition Insights:

Our methodology involves meticulous scrutiny of inequality ratios, unveiling shifts in the relevance of ESG over time. Notably, the COVID crisis dampened ESG effects in 2020, but a resilient rebound ensued in 2021, hinting at ESG's rebounding prominence.

ESG Momentum Insights:

In contrast to ESG scores, the analysis of ESG momentum doesn't yield significant volatility effects. The interplay between high ESG momentum and greater variability isn't consistent throughout, contrasting starkly with ESG scores' impact on market behavior.

Our non-parametric, data-driven approach suggests ESG's influence on stock dynamics, underlining its potential as a driver of volatility. The COVID pandemic weakened ESG effects momentarily, but a robust recovery ensued. Notably, the war in Ukraine exhibited a contrasting impact, strengthening ESG's role. The method's adaptability paves the way for future expansions, potentially incorporating more asset traits and groups, offering promising avenues for integrating sustainability into market risk assessments.

1 Introduction

Environmental, Social and Governance (ESG) score represents the most widespread and used indicator to evaluate the sustainability dimension of a company and it is the key factor to assess the sustainability impact of an investment in a company. The ESG score belongs to a broader array of multidimensional indicators aimed at assessing the progress towards achieving the 17 United Nations Sustainable Development Goals, upon which the UN's 2030 Agenda is built (Ricciolini et al. [21]). Over the last years we are facing a significant shift in capital markets perception toward corporate sustainability, which has fueled a growing empirical and theoretical literature.

In this paper we aim at investigating the interplay between the ESG score and the risk of a stock, with the purpose to contribute to the debate on the relevance and the effectiveness of ESG indicators and to explore and assess the role of ESG. We develop a data-based approach that focuses exclusively and directly on the interplay between ESG and the characteristics of assets. In this way, we can avoid taking into consideration aspects such as investors' propensity and preferences for sustainability-related themes (Albuquerque et al. [1]).

In our approach we add inequality decomposition methods to the tools used to evaluate the effects of ESG scores, so as to combine distribution-based assessments with traditional assessments based on synthetic indicators such as averages. An analysis of the companies included in the STOXX Europe 600 Index points to a link between higher ESG and lower volatility, with relevant, yet contrasting effects associated with the crises of 2020 and 2022.

2 Literature

The relationship between ESG and assets characteristics is multi-faceted and not straightforward (see, e.g. Baker et al. [4], Berg et al. [6], Friede et al. [11], Hartzmark and Sussmann [13]). One of the most controversial asset's features is their returns.The high demand for assets with a high ESG score is expected to result in a relevant increase in their prices. On the other hand, assets with a low ESG score may be rewarded by a premium, known as the "sin premium", which compensates the investor for holding assets that are unsustainable (Hens and Trutwin [15], Hong and Kacperczyk [16]. Consistent with these scenarios, Authors analyzing the relationship between ESG score and returns point out to no conclusive findings, with some, as e.g. Lius et.al. [18] and Khan [17] reporting higher returns, and others, as e.g. El Ghoul [10], reporting lower returns. An enlightening review is provided by Assael et al. [2], who support relevant but opposite ESG effects on the returns of small and large capitalization companies. Furthermore, they highlight the importance to move from an aggregate approach to an analysis by sectors and size, and also with reference to specific ESG indicators such as controversies.

Moving on to the risk analysis, it is certainly true that, on one hand, companies with high ESG scores could have stronger risk management practices, which could contribute to lower volatility. However, it is also important to remember that, on the other hand, the ESG score represents only one element of the overall risk profile, which certainly depends on a variety of factors.

Additional contributions emphasize the importance, in the relationship between ESG and financial activities, of other aspects such as ESG momentum (Berk et al. [7]) and the eterogeneity and volatility of ESG scores (Avramov et al. [3]).

In analyzing the relationship between ESG scores and financial performance, it's crucial to highlight that high ESG scores may contribute to better performance. However, strong performers might also have more resources to enhance their ESG scores.

Furthermore, an important strand of the literature on the relationship between ESG and financial markets concerns the generalization of milestones, such as the mean-variance efficient frontier (Pedersen et al. [20]) and the CAPM with its extensions, like Fama and French's three-factor model (Heinkel et al. [14], Zerbib [22]), achieved by incorporating ESG information. The complexity of sustainable allocation strategies, with an interesting focus on metrics for climate change and other sustainability risks, as well as a deeper exploration of climate-related and ESG risk indicators, is well outlined by [5].

Finally, it is important to recall that there are different ESG metrics and their effect on asset characteristics may not be uniform. Khan [17] introduces new ESG metrics, focusing on the governance dimension, Berg et al. [7] specifically attribute to the presence of different ESG metrics a confounding effect on the relationship between ESG and asset's characteristics, while Gibson et al. [12] attribute a risk premium to companies with higher ESG rating disagreement. These are just a few examples of contributions that have developed and deepened the role and effects of a plurality of ESG metrics. It is important to remember how, alongside ESG scores, there are also other important indicators extremely useful for analyzing the effects on financial markets of sustainability-related issues (Monasterolo and De Angelis [19]).

Similar to all rapidly evolving situations marked by significant success, the measurement of sustainability and its effects requires the current phase of numerous, highly diverse, often conflicting contributions, with the perspective of a subsequent phase of organization, restructuring, and systematization of the various proposals.

3 Methodology

We address the interplay between ESG score and asset characteristics as a classification problem. Let Y be the characteristic of the assets under investigation, $\{y_1, ..., y_n\}$ its values in the n assets, $\bar{y}$ is its arithmetic mean, $\bar{y}_j$ is its arithmetic mean in the j-th group. With the aim to evaluate the relevance of ESG score and its effects on a specific characteristic of the assets (return, volatility, extreme risk, ...), we define different groups of assets on the basis of a high / low ESG score level, thus tracing the assessment of ESG-related effects to the comparison between different distributions.

3.1 The case of 2 *groups*

Given n assets, in order to classify them into 2 groups on the basis of their ESG score, we introduce a $n \times 1$ vector Z

$$Z = \begin{cases} z_1 \text{ with } z_1 = 1 \text{ if } ESG < p_1 \\ z_2 \text{ with } z_2 = 0 \text{ if } ESG \geq p_2 \end{cases} \tag{1}$$

where, for $p_1 = p_2$, e.g. when using the median of ESG score, all n assets are considered, while for $p_1 \neq p_2$, e.g. when the ESG quartiles are considered, it is possible to refer to a part of the assets only.

Our traditional starting point is given by the synthetic evaluation provided by the difference between the group means. A first indicator of ESG effects can be obtained by calculating the ratio between the means of the standard deviations in the two groups:

$$I_{\bar{y}} = (\bar{y}_L - \bar{y}_H)/\bar{y}_H. \tag{2}$$

where $\bar{y_L}$ and $\bar{y_H}$ are the arithmetic means of Y in the group with a low and a high ESG score, respectively.

Our purpose is to add to the information derived from the group means, the information provided by the comparison of the group distributions.

We compare the volatility (or return, or extreme risk, ...) distribution of groups High and Low: when the two distributions perfectly overlap, the ESG score does not affect the assets returns, while, for decreasing overlapping levels, the influence of ESG score on asset volatility increases. When groups High and Low do not overlap, the influence of ESG score reaches its maximum and assets are perfectly classified on the basis of their ESG score.

In order to evaluate the distance between the distributions of groups High and Low, as well as to take into account their overlapping, we resort to the decomposition of an inequality indicator, thus allowing to effectively compare different groups.

We refer to one of the most used and widespread inequality measures, the Gini index, for which many different decompositions have been proposed. Among the many contributions we use the Dagum's Gini index decomposition [9], which allows to explicitly take into account the overlap between groups.

For the case of n assets disaggregated into 2 groups of size n_1 and n_2, with $n_1 + n_2 = n$, the Gini index can be expressed as

$$G = \frac{1}{2n^2\bar{y}} \sum_{j=1}^{2} \sum_{h=1}^{2} \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} |y_{ji} - y_{hr}| \tag{3}$$

where y_{ji} is the value of Y in the i-th asset of the j-th group and, accordingly, y_{hr} is its value in the r-th asset of the h-th group. For the case $j \neq h$, let us define

$$(y_{ji} - y_{hr})^+ = max\left\{(y_{ji} - y_{hr}), 0\right\}$$

and

$$(y_{ji} - y_{hr})^- = max\left\{-(y_{ji} - y_{hr}), 0\right\}$$

such as

$$\left|y_{ji} - y_{hr}\right| = (y_{ji} - y_{hr})^+ + (y_{ji} - y_{hr})^-.$$

It is therefore possible to derive the expressions for the inequality between groups G_b and for the overlapping component G_o as

$$G_b = \frac{1}{2n^2\bar{y}} \left(\sum_{i=1}^{n_1} \sum_{r=1}^{n_2} (y_{1i} - y_{2r})^+ + \sum_{i=1}^{n_2} \sum_{r=1}^{n_1} (y_{2i} - y_{1r})^+ \right) \tag{4}$$

and

$$G_o = \frac{1}{2n^2\bar{y}} \left(\sum_{i=1}^{n_1} \sum_{r=1}^{n_2} (y_{1i} - y_{2r})^- + \sum_{i=1}^{n_2} \sum_{r=1}^{n_1} (y_{2i} - y_{1r})^- \right). \tag{5}$$

Inequality between G_b and overlapping component G_o allow to evaluate the contribution to total inequality attributable to the differences between the groups, that is, in our case, to the effect of ESG score.

The role of the two components is quite different. On one hand, an high (low) G_b indicates a relevant (slight) ESG effect, as total inequality is (is not) strongly influenced by inequality between. On the other hand, as pointed out above, an high (low) G_o suggests a slight (relevant) ESG effect, since complete overlapping corresponds to the absence of ESG effect, while $G_0 = 0$ (High and Low groups are perfectly separated) indicates a total stratification.

On the basis of the different meaning of G_b and G_o it is possible to derive an indicator of the relevance of ESG as

$$I_G = (Gb - G_o)/G. \tag{6}$$

When the two indicators (2) and (6) show a similar trend, the information provided by the means is supported by the indications derived by the group distributions. In the opposite case, when I_G differs from $I_{\bar{y}}$, the group means are not fully informative and it is important to also exploit information from group distributions.

Our method is extremely flexible and can be also adopted to jointly investigate different assets characteristics. We can generalize our analysis, moving from the $n \times 1$ vector Z to a $n \times m$ matrix Z, where m is the number of characteristics analyzed and each column of Z is a $0,1$ vector classifying the n asstes into two groups with respect a specific characteristic. Further possible generalizations, always within the same framework, involve the option of assigning varied weights to the different characteristics.

The joint analysis of multiple characteristics implies dividing the n assets into 2^m groups, and thus can be related to the extension to the case of k groups outlined in the next Section.

3.2 Extension to the case of k groups

The case of only two groups represents the most common and used reference, but it is not exhaustive and the extension to k groups can provide useful insight on ESG related effects.

We sort the k groups in descending order on the basis of their ESG score, thus group 1 groups assets with highest ESG score while group k contains assets with lowest ESG score. We can generalize Z to the case of more than 2 groups moving to a sequence $p_1, p_2, ..., p_k$ such as

$$Z = \begin{cases} z_1 \text{ with } z_1 = (k-1)/(k-1) = 1 \text{ if } ESG < p_1 \\ z_2 \text{ with } z_2 = (k-2)/(k-1) \text{ if } p_1 \leq ESG < p_2 \\ \dots \\ z_i \text{ with } z_i = (k-i)/(k-1) \text{ if } p_{i-1} \leq ESG < p_i \\ \dots \\ z_k \text{ with } z_k = (k-k)/(k-1) = 0 \text{ if } ESG \geq p_{k-1} \end{cases}$$

As mentioned previously, in addition to being based on $p_1, p_2, ..., p_k$, the k groups could also result from the joint analysis of multiple characteristics.

As for $I_{\bar{y}}$, the extension of expression (2) to the case of k groups can be easily obtained through a two-step procedure. First, we obtain the indexes $I_{y_{ij}}$ related to the pairwise comparison between the i-th and the j-th groups

$$I_{y_{ij}} = (\bar{y}_i - \bar{y}_j)/\bar{y}_j.$$

Second, we calculate the average of all $I_{y_{ij}}$ indexes

$$I_{\bar{y}} = \frac{2}{k(k-1)} \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} I_{y_{ij}}. \tag{7}$$

Moving to inequality decomposition based method, in ordero to generalize I_G to the case of k groups we have to rewrite accordingly expressions (3), (4) and (5). The formula of the Gini index for the case of k groups is

$$G = \frac{1}{2n^2\bar{y}} \sum_{j=1}^{k} \sum_{h=1}^{k} \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} |y_{ji} - y_{hr}| \tag{8}$$

and, for $j \neq h$, the related formulas of inequality between and overlapping are, respectively,

$$G_b = \frac{1}{2n^2\bar{y}} \sum_{j=1}^{k} \sum_{h=1}^{k} \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} (y_{ji} - y_{hr})^{+} \tag{9}$$

and

$$G_b = \frac{1}{2n^2\bar{y}} \sum_{j=1}^{k} \sum_{h=1}^{k} \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} (y_{ji} - y_{hr})^{-}. \tag{10}$$

By referring to expressions (8), (9) and (10), the indicator I_G mantains without changes its structure:

$$I_G = (Gb - G_o)/G.$$

The case of k groups is certainly appealing and potentially very interesting, but it also makes the results less immediate and more challenging to read and interpret. This is why the benefits of this extension need to be balanced with the greater effort required in analyzing the results.

4 Data

We investigate the interplay between ESG and volatility of the stocks included in the STOXX Europe 600 Index, measured by means of the standard deviation of their returns. We refer to the assets included in the index as of September 30, 2023; the ESG score is not always available for all companies. Table 1 shows the number n of assets for which the Refinitiv ESG indicator is present.

We employ monthly data ranging from 2017 to 2022, thus allowing us to also evaluate the presence of effects related to the COVID19 pandemic and the war in Ukraine. ESG score is analyzed with reference to both its levels and its variations, taking into account the notion of ESG momentum. Among the different ESG metrics available, we refer to Refinitiv methodology, fully aware that the use of a different metric could influence the results. An interesting future development could involve the joint use of multiple ESG indicators and the comparison of the deriving stock classifications, so as to analyze their effects on results.

Based on the indicator calculated by Refinitiv, from 2017 to 2022, the assetss included in the STOXX Europe 600 Index have had an average ESG score of 67.7, with a mean squared deviation of 15.3. Figure 1 illustrates the frequency distribution of the ESG score average from 2017 to 2022; it can be observed that the ESG values are concentrated within the range of 70 to 85, showing a predominantly negatively skewed distribution.

Moving from average to annual data spanning from 2017 to 2022, it's intriguing to delve into the dynamics within this period, revealing an upward trend in the average ESG score and a decline in its variability, as shown in Figure 2 and Table 1.

Going into further detail, from 2017 to 2021, there's a consistent increase in the average ESG score, at a rate slightly surpassing 3% annually. However, between 2021 and 2022, there's a slowdown, with a variation that doesn't exceed 1%. The

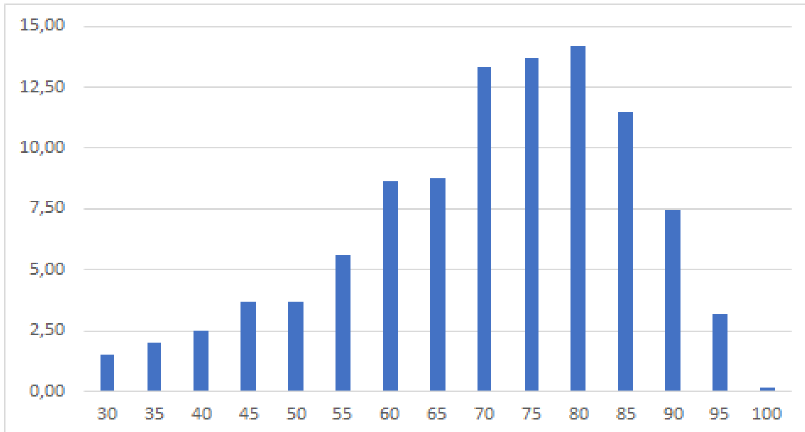
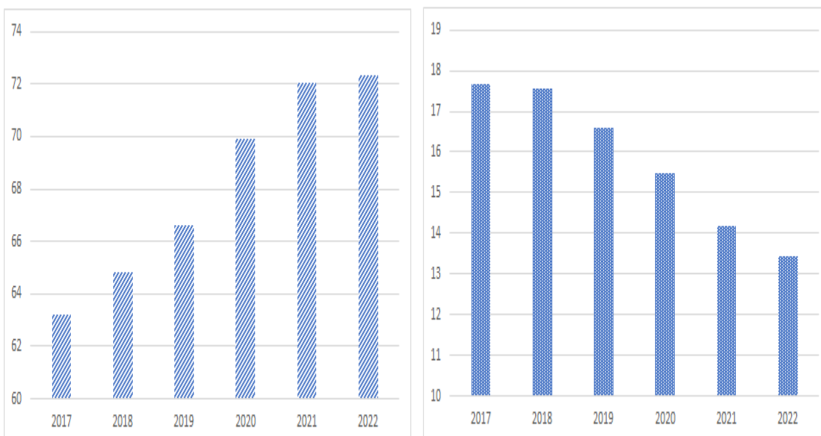


Fig. 1 Frequency distribution of ESG score average 2017-2022, STOXX Europe 600 Index assets.



(a) (b)

Fig. 2 ESG score average (a) and standard deviation (b), 2017-2022, STOXX Europe 600 Index assets.

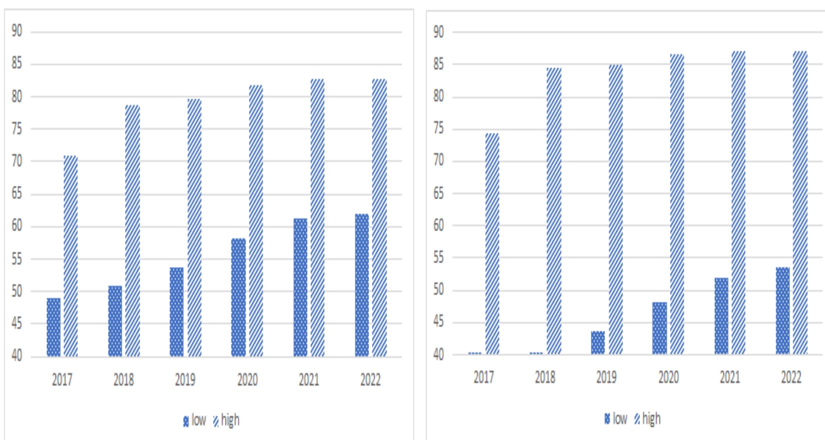
Table 1 Number *n* of assets with ESG score, mean and standard deviation of the ESG scores

	2017	2018	2019	2020	2021	2022	2017-22
n	533	562	575	584	589	468	592
mean	63.2	64.8	66.6	69.9	72.0	72.4	67.7
std.dev.	17.7	17.6	16.6	15.5	14.2	13.4	15.3

variability of the ESG scores among the stocks included in the index remains largely stable between 2017 and 2018 (decreasing by less than 1.0%). However, from 2018 to 2022, it demonstrates a significant decline, with an average annual rate exceeding -5%.The two contrasting trends observed in the average and variability of ESG scores from 2017 to 2022 reflect the increased attention and widespread sensitivity towards sustainability issues that have emerged in recent years.

Proceeding with the analysis of the annual dynamics, the stocks within the STOXX Europe 600 index are divided into $k = 2$ groups, high and low ESG. In order to classify the n assets we use both the median of ESG score with $p_1 = p_2$, and the case p_1 equal to the first quartile and p_2 equal to the third quartile.

Figure 3 depicts the averages ESG score of the two groups in both scenarios. It's interesting to note how, whether using the median or using quartiles, extremely similar results are obtained. Specifically, from 2017 to 2018, the increase in the average ESG is mainly attributable to assets with higher ESG, while starting from 2018, it's the assets with lower ESG that are driving the growth of the average ESG. Another aspect of the annual dynamics of the two high and low ESG groups under consideration concerns the variability of ESG scores.

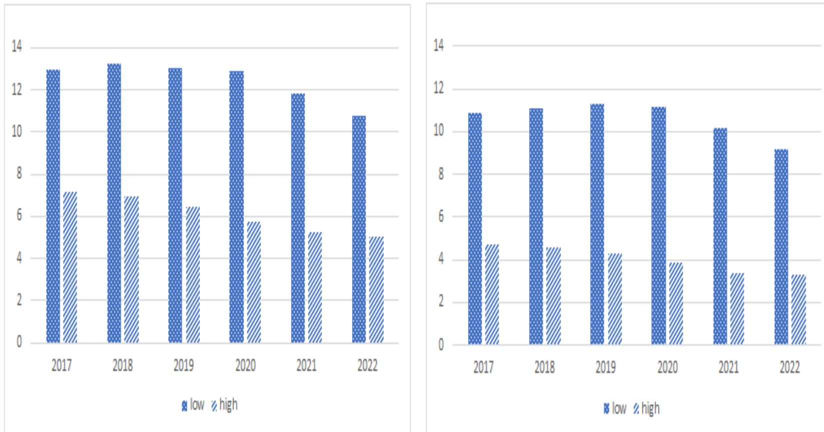


(a) (b)

Fig. 3 Average ESG scores in low and high ESG score groups based on the median (a) and on the first and last quartiles (b), 2017-2022, STOXX Europe 600 Index assets.

Figure 4 illustrates the standard deviation of ESG scores in the two high and low ESG groups for the two cases analyzed: it can be observed that the standard deviation is consistently higher in the lower ESG group. Over the considered period, there is a general decrease in ESG variability across all groups, with more pronounced declines between 2020 and 2021.

Alongside the levels of ESG scores, aiming to assess the effects related to dynamics and revisiting the concept of momentum, we evaluate the difference between the



(a) (b)

Fig. 4 Standard deviation ESG scores in low and high ESG score groups based on the median (a) and on the first and last quartiles (b), 2017-2022, STOXX Europe 600 Index assets.

ESG score at time t and the ESG score at the previous time $t-1$:

$$m_t = ESG_t - ESG_{t-1}$$

In Figure 5, the average ESG momentum is shown, with reference to the two high and low ESG momentum groups, always obtained by classifying the stocks based on either the median ESG momentum or ESG momentum quartiles.

Consistent with the observed growth in the ESG score during the period under consideration, the ESG momentum also appears positive, with higher values in 2018 and 2020, and a notable slowdown in 2022.

Regarding the variability of ESG momentum, Figure 6 shows the standard deviation of ESG momentum within high and low ESG momentum groups: it can be observed that, during the analyzed period, the variability decreases, and the low ESG momentum group consistently exhibits a lower standard deviation compared to that of the high ESG momentum group.

Figure 7 shows the frequency distribution of the average percentage returns of STOXX Europe 600 index assets. From 2017 to 2022, on average, the assets of the STOXX 600 Europe index yielded 0.77 per month, with a standard deviation of 0.97. The median of the distribution of the means is 0.65, quite close to the mean. The tenth percentile is -0.2 and the ninetieth is 1.96. These data, along with a skewness index of 0.66 (where symmetry = 0) and a kurtosis index of 3.21 (normal distribution = 0), indicate a slight departure from normal distribution.

Among the various characteristics of interest concerning the assets included in the STOXX index, the one we aim to analyze is volatility. Figure 8 displays the frequency distribution of the standard deviations of returns for the securities within the index. The mean of the standard deviations of monthly returns during the period

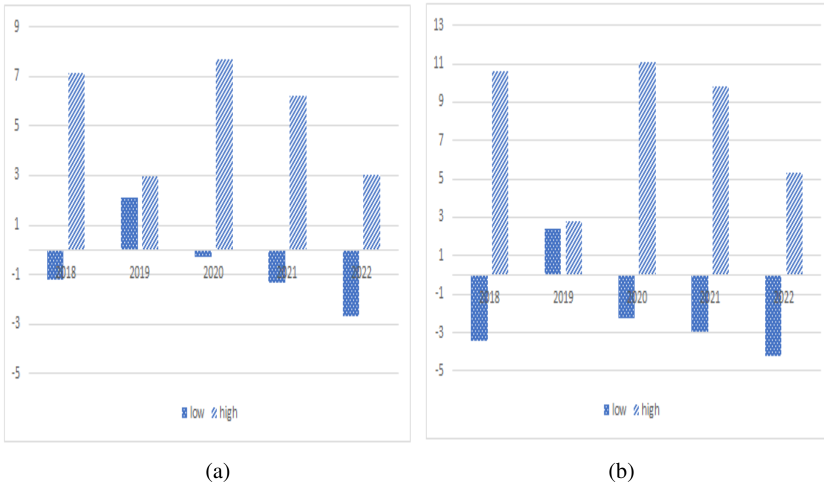


Fig. 5 Average ESG momentum in low and high ESG momentum groups based on the median (a) and on the first and last quartiles (b), 2017-2022, STOXX Europe 600 Index assets.

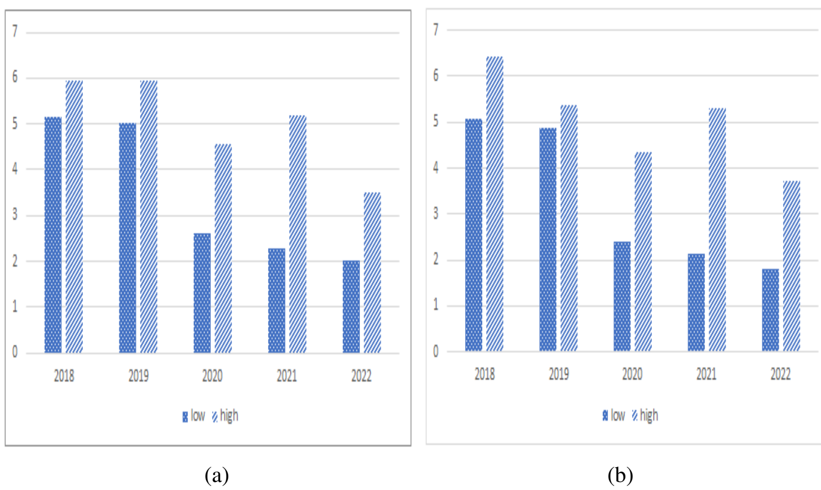


Fig. 6 Standard deviation ESG momentum in low and high ESG momentum groups based on the median (a) and on the first and last quartiles (b), 2017-2022, STOXX Europe 600 Index assets.

2017-2022 is 8.96, the median is 8.55, and the tenth and ninetieth percentiles are 5.86 and 12.56, respectively. The distribution is essentially symmetric, characterized by a right tail primarily due to the standard deviation of a single asset.

In the next section, we will analyze, referring to the methods outlined in Section 2, the relationship between the standard deviation of returns and ESG, consistently referencing both the ESG score and ESG momentum.

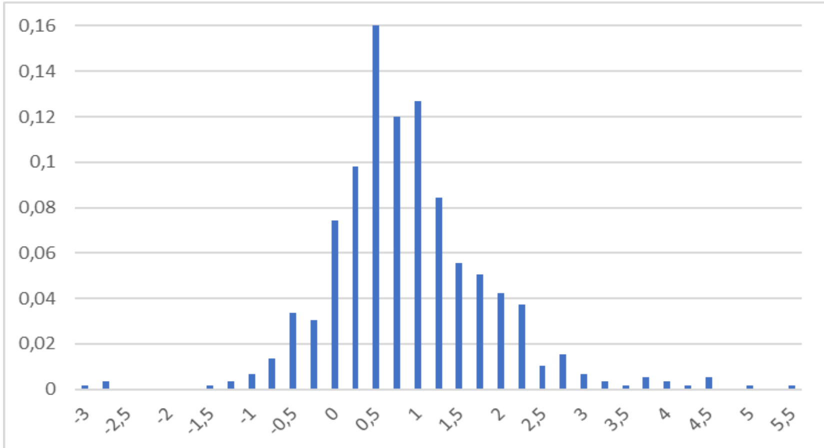


Fig. 7 Frequency distribution of assets average return 2017-2022, STOXX Europe 600 Index assets.

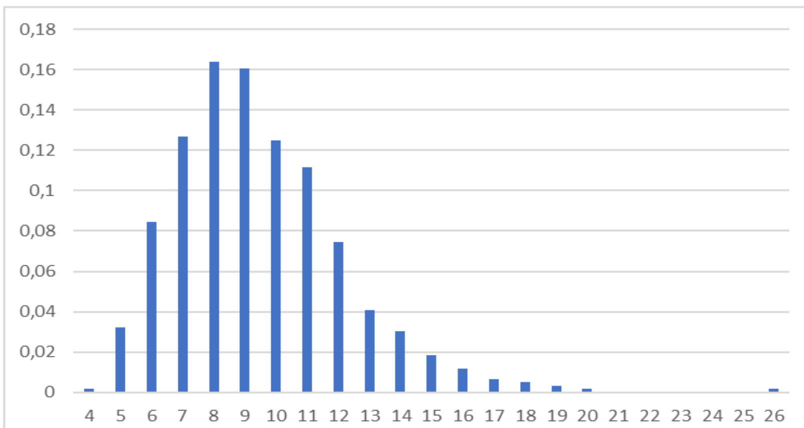


Fig. 8 Frequency distribution of assets standard deviation return 2017-2022, STOXX Europe 600 Index assets.

5 Results

The following results provide a comprehensive depiction of outcomes derived from asset classifications based on their ESG scores and ESG momentum, focusing on return volatility.

Table 2 illustrates the results related to the two groups obtained by classifying the assets on the basis of the median of their ESG score (the results for the first and last quartile are quite similar). First we report the mean of the standard deviations of the assets in the two groups. We can observe how, from 2017 to 2022, the average of the standard deviations increases, with significant peaks in 2020 and in 2022 corresponding to the COVID pandemic and the war in Ukraine. The values of volatility in the group with lower ESG are steadily higher.

We assess the relevance of our results by developing a bootstrap procedure resampling assets from the original list according to an i.i.d. sampling with replacement. Table 2 includes the mean and the standard deviation of 100,000 replicates of the bootstrap. The bootstrap mean is always equal or very close to the observed values; this element, together with the low values of the standard deviation, strongly supports our results.

To assess the significance of the difference between the volatility of stocks with low and high ESG scores, from the bootstrap procedure we also implement a test for the null hypothesis

$$H_0 : \bar{y}_L = \bar{y}_H.$$

Table 2 contains the bootstrap p-values, indicating that the hypothesis of equality of means between the two groups is always rejected except for 2020.

The second part of Table 2 shows the ratios G_b/G and G_o/G which, by evaluating the weight of inequality between groups and of overlapping component on total inequality, allow us to obtain further information on the role of ESG scores.

From 2017 to 2022 the relevance of G_b increases slightly, with the exception of the negative peak of 2020, suggesting a stronger ESG effect and a deep impact of COVID pandemic.

The overlapping component also indicates, by decreasing, a stronger ESG effect, still except in 2020. It is interesting to observe the difference between the two crises of 2020 and 2022. In both cases, volatility increases significantly. However, in the first case, the differences between the two groups decrease (G_b decreases and G_o increases), while in the second case, the opposite occurs. The trends of G_b and G_o suggest a stronger effect of the ESG score during the 2022 crisis but not in 2020.

Also for G_b and G_o the bootstrap results support our findings: the means of 10000 bootstrap replicates are always equal or very close to the observed values, with an almost negligible standard deviation.

The last part of Table 2 shows indicators $I_{\bar{y}}$ and I_G which agree in suggesting an increase in the ESG effect over the last 6 years, together with detecting a strong impact from both the COVID pandemic and the war in Ukraine.

An interesting element arising from the comparison between $I_{\bar{y}}$ and I_G highlights the dynamic after the pandemic: from 2019 to 2022, I_G shows a stronger increase with respect to $I_{\bar{y}}$, suggesting that relying solely on averages may underestimate the importance of the ESG score, which, on the contrary, would be fully captured by distribution-based indicators.

Unlike the case with ESG scores, the analysis of ESG momentum does not indicate significant effects on stock volatility. Table 3 presents the results related to ESG momentum, following the same structure as Table 2, and thus presenting observed values alongside the mean and standard deviation of 100,000 replicates of the bootstrap developed by resampling assets from the original list according to an i.i.d. method.

Table 2 Mean of the observed standard deviations of the assets in group with High / Low ESG score; observed ratios G_b/G and G_o/G; mean and standard deviation of 100,000 bootstrap replicates

		2017	2018	2019	2020	2021	2022	2017-22
	$\bar{y}$	5.26	6.31	6.54	11.71	6.65	9.18	8.46
High ESG	mean boot	5.26	6.31	6.54	11.71	6.65	9.18	8.46
	std.dev.boot	0.12	0.13	0.17	0.30	0.14	0.17	0.13
	$\bar{y}$	5.63	7.01	7.17	12.09	7.23	10.59	9.22
Low ESG	mean boot	5.64	7.02	7.18	12.07	7.21	10.60	9.21
	std.dev.boot	0.19	0.18	0.18	0.32	0.17	0.23	0.18
$H_0 : \bar{y}_L = \bar{y}_H$		< 0.001	< 0.001	< 0.001	0.20	< 0.001	< 0.001	< 0.001
	G_b/G	0.29	0.32	0.30	0.27	0.31	0.36	0.32
Ineq. between	mean boot	0.29	0.32	0.30	0.27	0.31	0.36	0.32
	std.dev.boot	0.02	0.02	0.02	0.02	0.02	0.02	0.02
	G_o/G	0.21	0.19	0.20	0.23	0.19	0.16	0.19
Overlapping	mean boot	0.21	0.19	0.20	0.23	0.20	0.16	0.19
	std.dev.boot	0.02	0.02	0.02	0.01	0.02	0.02	0.02
	$I_{\bar{y}}$	0.07	0.11	0.10	0.03	0.09	0.15	0.09
	I_G	0.08	0.13	0.10	0.04	0.12	0.20	0.13

From the first part of the Table 3, it can be observed that the averages of the standard deviations of stocks in the two groups, high and low ESG momentum, do not significantly differ. Not only that, but the sign of the difference is not consistent over the considered period: on average, stocks with high ESG momentum also exhibit greater variability, $\bar{y}_L < \bar{y}_H$, but in the years 2020 and 2021, stocks with low ESG momentum, instead, show higher standard deviation in returns. The hypothesis of equality between the means of the two groups $H_0 : \bar{y}_L = \bar{y}_H$ is never rejected for any of the years considered, whereas it is rejected when analyzing the period 2018-22 as a whole.

In the second part of Table 3, consistent with the smaller difference between the means, the inequality between the groups G_b is also less pronounced, while the overlap component G_o appears more consistent. Both proposed summary indicators $I_{\bar{y}}$ and I_G show, in the case of ESG momentum, more contained and stable values during the analyzed period, highlighting a notable difference compared to the analysis of ESG scores and not revealing significant effects of ESG momentum on the variability of STOXX Europe 600 index assets.

The analysis carried out can be extended in several directions, always maintaining the methodological framework outlined in Section 3, which proves to be extremely flexible and adaptable to numerous generalizations. Among possible future developments, one can consider stock returns computed at different horizons, the inclusion of additional features of stocks, to be evaluated individually or collectively, and, naturally, the classification of stocks into k groups, as specified in Section 3.

Table 3 Mean of the observed standard deviations of the assets in group with High / Low ESG momentum; observed ratios G_b/G and G_o/G; mean and standard deviation of 100,000 bootstrap replicates

		2018	2019	2020	2021	2022	2018-22
	$\bar{y}$	6.82	6.90	11.89	6.94	9.93	9.20
High Momentum	mean boot	6.82	6.91	11.90	6.94	9.93	9.20
	std.dev.boot	0.16	0.17	0.31	0.14	0.21	0.16
	$\bar{y}$	6.60	7.07	12.14	7.00	9.83	8.72
Low Momentum	mean boot	6.60	7.07	12.12	7.01	9.81	8.70
	std.dev.boot	0.16	0.17	0.31	0.16	0.21	0.16
$H_0 : \bar{y}_L = \bar{y}_H$		0.16	0.31	0.44	0.66	0.64	0.01
	G_b/G	0.27	0.26	0.26	0.26	0.26	0.29
Ineq. between	mean boot	0.27	0.27	0.27	0.27	0.27	0.29
	std.dev.boot	0.02	0.01	0.01	0.01	0.01	0.02
	G_o/G	0.23	0.24	0.24	0.24	0.24	0.21
Overlapping	mean boot	0.23	0.23	0.23	0.23	0.23	0.21
	std.dev.boot	0.02	0.01	0.01	0.01	0.01	0.02
	$I_{\bar{y}}$	0.03	0.02	0.02	0.01	0.01	0.08
	I_G	0.04	0.03	0.02	0.01	0.01	0.06

6 Conclusions

A data-based, simple yet effective method is proposed to assess and evaluate the role of the ESG score. This is a non-parametric approach that doesn't require specific preliminary assumptions and lacks the elegance, depth, and implications, particularly in predictive terms, of many econometric models. However, it compensates for these gaps with extreme simplicity and strong effectiveness.

The analysis of the interplay between ESG score and assets characteristics can be addressed in the framework of classification and can greatly benefit from employing inequality decomposition methods, able to also take into account the overlapping between group distributions. On the basis of these methods it is possible to exploit the information related to the entire distribution of the groups and not rely only on the group means. With this objective in mind, we introduce an indicator I_G that effectively supplements and integrates information provided by the more traditional comparison of means $I_{\bar{y}}$.

We analyze the stocks included in the STOXX Europe 600 Index and refer to Refinitiv ESG score from 2017 to 2022. Our findinds support ESG as a positive driver of assets returns, with a relationship between higher ESG score and lower volatility. Covid pandemic strongly affected ESG effect, which was greatly weakeened in 2020, the only year in which the average volatility of low ESG stocks is not significantly different from the average volatility of high ESG stocks. A strong recovery of ESG effect can already be observed in 2021. The size of the recovery is stronger on the basis of I_G, that is when evaluating the entire group distributions.

The effects of the shock caused by the war in Ukraine are profoundly different from those of the COVID crisis: the role of the ESG score is not weakened but rather strengthened. This indication, already evident when looking at the averages, becomes even stronger when evaluating the entire group distributions.

The extreme flexibility of our proposal makes it possible to add other asset characteristics to the analysis, also considering them jointly, with numerous further developments, such as the possibility of more than two groups, capable of providing promising results and effectively including sustainability issues in stock market risk evaluation.

References

1. Albuquerque R., Koskinen Y., Santioni R.: Mutual fund trading and ESG stock resilience during the Covid-19 stock market crash. Temi di discussione, Bank of Italy n. 1371 (2022)
2. Assael J., Carlier L., Challet D.: Dissecting the explanatory power of ESG features on equity returns by sector, capitalization, and year with interpretable machine learning. J. Risk Financial Manag. **16**,3, 1-22 (2023)
3. Avramov D., Cheng S., Lioui A., Tarelli A.: Sustainable investing with ESG rating uncertainty. J. Fin. Eco. **145**, 642–664 (2022)
4. Baker M., Egan M.L., Sarkar S.K.: How do investors value ESG? NBER WP n. 30708 (2022)
5. Banca Italia: Rapporto annuale sugli investimenti sostenibili e sui rischi climatici. **2**, 1–57 (2023)
6. Berg F., Koelbel J.F., Pavlova A., Rigobon R.: ESG confusion and stock returns: tackling the problem of noise. NBER WP n. 30562 (2022)
7. Berk I., Guidolin M., Magnani M.: Strong vs stable: the impact of ESG ratings momentum and their volatility on the cost of equity capital. Bocconi WP n. 202 (2023)
8. Cao M., Duan K., Ibrahim H.: Green investments and their impact on ESG ratings: an evidence from China. Eco. Lett. **232** (2023)
9. Dagum C.: A new approach to the decomposition of the Gini income inequality ratio. Emp. Eco. **22**, 515–531 (1997)
10. El Ghoul S., Guedhami O., Kwok C.C.Y:, Mishra D.R.: Does corporate social responsibility affect the cost of capital?. J. Bank. Fin. **35**, 2388–2406 (2011)
11. Friede G., Busch T., Bassen A.: ESG and financial performance: aggregate evidence from more than 2000 empirical studies. J. Sust. Fin. & Inv. **5**, 210–233 (2015)
12. Gibson R., Krueger P., Schmidt P.S.: ESG rating disagreement and stock returns. Swiss Finance Institute WP n. 19-67 (2021)
13. Hartzmark S.M., Sussmann A.B.: Do investors value sustainability? A natural experiment examining ranking and fund flows. J. Fin. **74**, 2789–2837 (2019)
14. Heinkel R., Kraus A., Zechner J.: The effect of green investment on corporate behavior. J. Fin. Quant. An. **36**, 431–449 (2001)
15. Hens T., Trutwin E.: Modelling sustainable investing in the CAPM. Swiss Finance Institute WP n. 23-56 (2023)
16. Hong H., Kacperczyk M.: The price of sin: the effects of social norms on markets. J. Fin. Eco. **93**, 15–36 (2009)
17. Khan M.: Corporate governance, ESG and stock returns around the world. Fin. Anal. J. **75**, 103–123 (2019)
18. Liu M., Guo T., Ping W.: Sustainability and stability: Will ESG investment reduce the return and volatility spillover effects across the Chinese financial market?,. Ene. Eco. **121**, 1-16 (2023)
19. Monasterolo I., De Angelis L.: Blind to carbon risk? An analysis of stock market reaction to the Paris Agreement. Ecol. Eco. **170**, 515–531 (2020)
20. Pedersen L.H., Fitzgibbons S., Pomorski L.: Responsible investing: the esg-efficient frontier. J. Fin. Eco. **142**, 572–597 (2021)

21. Ricciolini E., Tiralti A., Paolotti L., Rocchi L., Boggia A: Sustainable development according to 2030 agenda in European Union countries: Evidence of the enlargement policy. Sust. Dev., 1-19 (2023)
22. Zerbib O.D.: A sustainable asset pricing model: evidence from environmental integration and sin stock exclusion. Rev. Fin. **26**, 1345–1388 (2022)

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Strada Maggiore 45
40125 Bologna - Italy
Tel. +39 051 2092604
Fax +39 051 2092664
http://www.dse.unibo.it