European Financial Management

WILEY

EUROPEAN FINANCIAL MANAGEMENT

ORIGINAL ARTICLE OPEN ACCESS

ESG Risk and Market Return Predictability: New Evidence From the Eurozone

Zhiyong Li[1] | Zhuoran Li[2] | Weiping Qin[3,4]

[1]School of Accounting and Finance, University of Leicester, Leicester, UK | [2]Lincoln International Business School, University of Lincoln, Lincoln, UK | [3]Southampton Business School, University of Southampton, Southampton, UK | [4]Alliance Manchester Business School, University of Manchester, Manchester, UK

Correspondence: Weiping Qin (W.Qin@soton.ac.uk)

Received: 8 October 2024 | **Revised:** 8 January 2025 | **Accepted:** 15 March 2025

Keywords: business cycle | forecast combination | incident-based ESG risk | return predictability

ABSTRACT

This study examines the predictive power of incident-based Environmental, Social and Governance (ESG) risk on the Eurozone stock market returns using a forecast combination method. We find that our constructed indicator shows significant return predictability from both a statistical and economic perspective, with an out-of-sample CER gain of 4.55% and a Sharpe ratio of 0.43, consistently outperforming the mean benchmark. Moreover, we find that the predictive power is concentrated during non-expansion periods. We attribute this mechanism to the firm's fundamentals, cash flow and discount rate channels. Our findings highlight the value of ESG information for investors.

JEL Classification: G11, G12, G17

1 | Introduction

Environmental, Social and Governance (ESG) has become an attractive topic for financial investment[1]. Academics are increasingly studying ESG's influence on financial markets, particularly its relationship with firm performance (Aouadi and Marsat 2018; Bansal et al. 2022; Bissoondoyal-Bheenick et al. 2023; Cheng et al. 2024), market reactions (Capelle-Blancard and Petit 2019; Serafeim and Yoon 2023), and market efficiency (Bofinger et al. 2022; Zhang et al. 2023). However, research exploring the predictive power of ESG information on market excess returns remains limited. Only a few studies have begun to investigate the relationship between ESG factors and market returns, hinting at the untapped value of ESG as a predictive variable (Maiti 2021).

The rationale for using ESG information in return predictability stems from the growing recognition of ESG in investment decision-making. The inclusion of ESG criteria in predictive models is gaining attention due to increasing evidence of their relevance and impact on market performance. Research results vary: some find a positive correlation between strong governance and return predictability (Gompers et al. 2003), while others argue for lower expected returns from firms committed to sustainability (Geczy et al. 2021). Additionally, some studies find higher returns for 'sin' firms or those with higher CO_2 emissions (Bolton and Kacperczyk 2021; Hong and Kacperczyk 2009; Oestreich and Tsiakas 2015). ESG-related policy changes have also been highlighted for their relevance in asset pricing (Ilhan et al. 2021; Kelly et al. 2016).

Most research focuses on stock performance based on ESG characteristics, but Chu et al. (2024) are among the few who demonstrate that aggregate ESG scores are associated with excess returns at the market-level. However, their study relies on internally disclosed ESG scores and focuses on the US market. According to Wong and Zhang (2022), investors not only consider ESG information reported by firms but also external ESG news from social media and newspapers. This paper aims to fill this gap by providing evidence from an external, incident-based

ESG risk perspective on aggregate market return predictability, focusing on the Eurozone market - a region characterised by high levels of market integration (Abad et al. 2014; Allen and Song 2005; Qin et al. 2023).

A recent report by Morningstar (2024) indicates that sustainable fund flows in Europe have consistently experienced net inflows, reflecting the potential of ESG as a factor in investment decisions. However, to the best of our knowledge, the predictability of Eurozone market returns using aggregate ESG information remains largely unexplored. Therefore, this study examines whether external ESG-related negative news has significant predictive ability for the Eurozone financial market returns, offering investment insights for investors.

We introduce an incident-based ESG risk indicator from RepRisk, a third-party specialising in negative ESG news. These news releases are from print media and various online sources, including but not limited to social media, blogs and government bodies. For this purpose, we use the RepRisk Index (*RRI*) database from over 2000 firms within the Eurozone, with more than 20,000 observations from 2007 to 2022, in response to Kelly et al. (2024)'s recommendation to consider larger information sets to improve predictability.

First, we examine the in-sample predictability of the Eurozone stock market excess returns using our ESG risk indicator. We find that the incident-based ESG risk indicator is statistically significant, with an in-sample R^2 value of 2.79% for univariate regressions, outperforming well-established predictors identified by Welch and Goyal (2008). Bivariate regression, combining the ESG risk indicator with popular predictors, further enhances the indicator's predictive power, yielding higher in-sample R^2 values.

For out-of-sample forecasting, we apply the forecast combination method by Rapach et al. (2010). Our findings indicate that our incident-based ESG risk indicator significantly predicts market excess returns, with an out-of-sample R_{os}^2 value of 1.31%. We also explore the economic significance of the ESG variable, finding that it generates an annualised certainty equivalent return (CER) gain of 4.55% and a Sharpe ratio of 0.43.

For robustness checks, we consider various conditions, including higher levels of investor risk aversion, transaction costs, different financial constraints, and alternative weighting methods on return forecasts. We also use the MSCI EMU Total Return Index and the Europe Portfolios Return Index as alternatives to the Eurozone market. All results withstand these robustness checks, outperforming popular predictors identified by Welch and Goyal (2008).

Additionally, we assess return predictability across business cycles, building on studies that highlight the heterogeneity of market excess return predictability under different economic conditions. Our results show that the predictive power of the ESG risk indicator is primarily evident during non-expansion periods. Furthermore, we find that the predictive power of our indicator extends beyond the aggregate Eurozone market to some member countries within the Eurozone.

We also explore the mechanism behind the ESG indicator's predictive power through the firm fundamentals, cash flow and discount rate channels. The evidence for the firm fundamentals channel aligns with the theoretical framework proposed by Pedersen et al. (2021), which suggests that a firm's ESG score reflects its underlying fundamentals. Besides, our analysis reveals that this predictability is also linked to the discount rate and cash flows. As noted by Inard (2023), ESG predominantly impacts a firm's market performance, and they highlight that ESG considerations can influence a firm's revenue, creating uncertainty about future cash flows. The discount rate channel has also been studied by Bolton and Kacperczyk (2021) and Pástor et al. (2021), who argue that negative ESG information influences divestment, affecting a firm's cost of capital. If such firms are avoided by investors, their cost of capital is likely to increase, resulting in diminished firm values.

This paper aligns with existing research emphasising the growing attention to ESG in stock market return predictability (see, e.g., Cao et al. 2023; Chen et al. 2023; Chu et al. 2024; Khan 2019; Serafeim and Yoon 2023). Compared to prior literature, the first contribution of this paper is that we are the first to examine the predictive power of an incident-based ESG risk indicator for market excess returns in the Eurozone. Our findings also complement Agoraki et al. (2023), who show that European firms' ESG risk information impacts their market performance. We also provide a detailed comparison with well-established predictors and find that the ESG risk indicator remains powerful and useful in the presence of the widely used predictors.

Our second contribution is that we are the first to apply purely objective ESG information in the field of return prediction. Most previous research has focused on corporate finance, particularly on the effect of scandals involving customers and suppliers (Dai et al. 2021), the impact of incidents on firms' participation in ESG actions (Li and Wu 2020) and firms' financial performance (Bansal et al. 2022). According to Cumming et al. (2024), the implications of a universally accepted sustainable risk assessment could be further explored. The application of the incident-based ESG risk in the asset pricing field, particularly in market return prediction, is explored for the first time in this article.

We also contribute to the literature on the heterogeneity of return predictability over the business cycle. While return predictability across business cycles has been widely studied (Cujean and Hasler 2017; Pettenuzzo et al. 2014; Rapach and Zhou 2013), the role of ESG information in this context has not yet been explored. This study is the first to examine potential patterns of return predictability by using our constructed ESG risk indicator. By separating the business cycle into expansion and non-expansion periods based on the Eurozone Business Cycle Clock classification, we find that return predictability over the business cycle exhibits distinct patterns. Specifically, the predictive power of the incident-based ESG risk indicator is primarily evident during non-expansion periods in the Eurozone stock market. Our findings support the argument that in periods of economic downturn, which are characterised by uncertainty and high volatility, returns become more predictable due to divergent interpretation of ESG news (Cujean and Hasler 2017).

Furthermore, this article extends the application of machine learning methods in asset pricing research. We employ the forecast combination method to construct aggregate return forecasts. Initially proposed by Bates and Granger (1969), this method is widely used in financial markets (e.g., Chen and Maung 2023; Gospodinov and Maasoumi 2021; Rapach et al. 2010; Rapach and Zhou 2022a; Xie and Hong 2016). It offers a strong shrinkage effect for high-dimensional datasets, effectively incorporating information and avoiding overfitting. However, no previous study has applied this method to return predictability using ESG information. We introduce this method to address this gap by predicting the Eurozone market returns using the incident-based ESG risk indicator.

In the next section, we describe our ESG risk indicators and other relevant variables. Section 3 outlines the methodology, and Section 4 discusses the main empirical results. Section 5 presents various additional tests, and Section 6 explains the mechanism of return predictability. Finally, we conclude this article.

2 | Data Description

2.1 | Eurozone as a Single Unit

In this paper, we treat the Eurozone as a single unit to study the predictability of the Eurozone stock market based on the ESG risks of its firms. This is attributed to the region's high degree of economic, regulatory, and financial integration.

The Eurozone represents the largest integrated economic area outside the United States, characterised by harmonised financial regulations and ESG policies. These shared frameworks, such as the EU Taxonomy and the Sustainable Finance Disclosure Regulation (SFDR), provide uniform standards for assessing ESG risks across member states. This regulatory alignment ensures consistent treatment of ESG risks, making the Eurozone a suitable candidate for examining their impact on asset returns.

Additionally, the Eurozone's economic and monetary integration further supports this unified approach. Member countries share a single currency and a common monetary policy governed by the European Central Bank (ECB), eliminating currency risks and enhancing market coherence. This integration facilitates the pooling of data across the region to analyse ESG risk collectively.

The seamless cross-border movement of goods, services, capital, and labour within the Eurozone reduces market segmentation. These factors create a cohesive financial ecosystem that mirrors the characteristics of a single market. This unique level of integration underpins our decision to treat the Eurozone as a unified entity for this study.

2.2 | Incident-Based ESG Risk Indicator

Although several agencies rate ESG performance at the firm level, the ESG scores of the same firm are often inconsistent, or even contradictory, across different agencies due to their varying definitions and criteria for specific categories. For example, Sustainalytics rates Tesla with a medium score due to its sustainable development effects but notes its poor governance. Conversely, S&P argues that Tesla is not an ESG firm and removed it from the S&P 500 ESG index in 2022, citing an unclear lower-carbon strategy and Elon Musk's negative attitude toward ESG. Recent research papers also highlight the inconsistency of ESG scores across different rating providers and the potential issues this inconsistency may cause. Berg et al. (2021) investigate the noise in ESG ratings among various raters, emphasising the potential problem of score inconsistency due to revisions, such as back-filling past values when raters change their methodology. Studies by Berg et al. (2021) and Berg et al. (2021) indicate that ESG data from Refinitiv, MSCI, and Sustainalytics may exhibit potential look-ahead bias.

To address this issue, we consider RepRisk, a media-oriented index, in this study. We select this data set for three primary reasons: First, RepRisk focuses on media-driven incidents rather than self-reported data by the firms themselves. Relevant studies have demonstrated that investors react more to ESG information from external sources than to self-disclosed data (e.g., Capelle-Blancard and Petit 2019). Second, RepRisk specifically targets negative incidents. Krüger (2015) argues that investors react more severely to negative corporate social responsibility news, while their reaction to positive corporate social responsibility news is relatively weak and neither statistically nor economically significant. Hence, RepRisk is the ideal ESG rater for this study. Third, compared to other ESG raters, RepRisk measures ratings with high frequency, providing more detailed information and allowing us to capture short-term dynamics, thereby offering crucial insights for short-term investors. Therefore, the data set has been widely applied in top journals by researchers (Bansal et al. 2022; Li and Wu 2020) and stakeholders (Dai et al. 2021; Raghunandan and Rajgopal 2022). This paper is the first to apply the data set to the research on return predictability.

In this article, we use the 'current *RRI*' from RepRisk to measure incident-based ESG risk. *RRI* is a short-term incident-based score reflecting a firm's short-term exposure to media and stakeholder attention regarding ESG issues. The score ranges from 0 (lowest) to 100 (highest), with a higher value indicating a higher rate of ESG incidents. The extent of this increase is influenced by factors such as the severity and novelty of the incident, as well as the extent and intensity of the news coverage surrounding it. The sample is from January 2007 to December 2022.

We consider public firms headquartered in the Eurozone as our sample, since this region is the third-largest economy in the world[2], and the impact of ESG risk information on its market return has not been investigated yet. We excluded firms without any record of ESG incidents during our sample period, resulting in a total of 2132 public firms. Since the raw *RRI* might have different scales depending on factors such as firm size, industry, or other characteristics as measured by RepRisk, we follow Chu et al. (2024) to transform the raw *RRI* into its percentage level to reduce these bias. We use *IESGRI* (Incident-based ESG Risk Indicator) to represent this variable. The *IESGRI* is calculated as follows:

$$IESGRI_{i,t+1} = \frac{RRI_{i,t+1} - RRI_{i,t}}{RRI_{i,t}} \times 100\%, \quad (1)$$

where, $IESGRI_{i,t+1}$ is the percentage change in RRI for firm i from month t to $t+1$, and $RRI_{i,t+1}$ and $RRI_{i,t}$ denote the RRI for the ith firm at months $t+1$ and t, respectively. The monthly RRI is calculated as the average of the daily RRI in that month. For a firm with a missing observation in a given month, the value is filled with the cross-sectional average of the available RRI for that month due to the requirement for non-missing data in our method.

Table 1 reports the descriptive statistics of the *IESGRI* for individual firms headquartered in specific countries. When a firm experiences an unforeseen negative event, its *IESGRI* may initially spike but then gradually decline over subsequent months unless another significant event or relevant policy emerges. The number of observations for each country depends on the count of publicly traded firms and the quantitative assessment of ESG risk incidents related to firms headquartered in that country. Germany and France have more than 20,000 observations, as large firms are mostly headquartered in developed countries within the Eurozone. In contrast, countries like Croatia, Estonia, Latvia, Lithuania, Malta, Slovakia and Slovenia only have hundreds of observations since they are still developing.

2.3 | Market Returns and Other Variables

We collect the data of the Morningstar Eurozone Net Return Index from Bloomberg, which measures the performance of over 97% of stocks in the Eurozone's board regional markets based on their market capitalisation. The monthly data spans from January 2007 to December 2022. The market excess returns are calculated by subtracting the risk-free return from the monthly market returns in the Morningstar Eurozone Index. We use the EURIBOR 3-month interest rate as the representative risk-free return, considering its stability[3].

We also include a set of well-established predictors identified in Welch and Goyal (2008) as control variables. These predictors encompass Eurozone stock market ratios, including dividend-price ratio (log), dividend yield (log), earnings-price ratio (log), dividend-payout ratio (log), and book-to-market ratio, as well as key Eurozone macroeconomic variables, including short-term returns, long-term returns, yield spread, stock variance and inflation. These control variables are obtained from Bloomberg, Deutsche Bank, and Eurostat. Moreover, given the documented influence of US stock market returns on European markets (e.g., Rapach et al. 2013), we incorporate S&P 500 excess returns as an additional control variable. The S&P 500 data are from Amit Goyal's website[4]. The details of the variables employed in this article are provided in Table A1 of Appendix.

TABLE 1 | Descriptive statistics of $IESGRI_i$ across countries in the Eurozone.

Country	Num.	Mean	Min.	Max.	Std.	Obs.
Austria	86	−0.30	−0.84	0.49	0.28	4957
Belgium	97	−0.29	−0.65	1.06	0.34	5782
Croatia	20	−0.43	−0.66	0.30	0.21	668
Cyprus	34	−0.37	−0.69	0.82	0.29	1422
Estonia	15	−0.27	−0.56	0.77	0.34	604
Finland	117	−0.26	−0.61	0.41	0.23	7058
France	401	−0.27	−0.85	1.47	0.31	25,599
Germany	358	−0.26	−0.74	1.50	0.32	23,374
Greece	52	−0.42	−0.71	0.66	0.27	2198
Ireland	86	−0.14	−0.85	1.48	0.43	6003
Italy	232	−0.23	−0.72	1.77	0.33	13,801
Latvia	12	−0.47	−0.56	−0.19	0.09	473
Lithuania	13	−0.36	−0.60	0.14	0.20	608
Luxembourg	88	−0.32	−0.84	0.90	0.31	4468
Malta	11	−0.47	−0.70	−0.09	0.16	300
Netherlands	219	−0.27	−0.84	1.84	0.36	12,955
Portugal	53	−0.23	−0.64	0.61	0.31	3517
Slovakia	13	−0.42	−0.56	−0.02	0.15	489
Slovenia	9	−0.45	−0.59	−0.10	0.16	298
Spain	216	−0.27	−0.70	1.72	0.33	14,180
All Eurozone	2132	—	—	—	—	—

Note: This table reports the number of firms and their mean, minimum, maximum, standard deviation and observations across countries in the Eurozone. Mean, minimum, maximum and standard deviation are from *IESGRI* series calculated by Equation (1). Observations are the number of ESG risk incidents recorded by RepRisk. The sample spans the period from January 2007 to December 2022.

3 | Methodology

3.1 | Predictive Regression Model

We use a standard predictive regression model to assess the time-series predictive power of ESG risk for market excess returns. The regression model is as follows:

$$r_{t+1} = \alpha + \beta \times \chi_t + \epsilon_{t+1}, \text{ for } t = 1, \ldots, T-1 \tag{2}$$

where r_{t+1} is the market excess return in month $t+1, \chi_t$ is a stock return predictor, and ϵ_{t+1} is an error term.

We will conduct both in-sample and out-of-sample analyses based on Equation (2). The in-sample analysis focuses on establishing the statistical significance of the predictor. A successful predictor's coefficient should be statistically different from zero. Moreover, the in-sample analysis provides insights into the economic intuition and the underlying mechanisms driving this predictability.

However, Welch and Goyal (2008) demonstrate that many well-known in-sample predictors fail to outperform the prevailing mean benchmark predictor (i.e., a random walk with drift) in out-of-sample tests. These shortcomings often result from look-ahead bias, which arises when the entire sample is analysed at once rather than incrementally, failing to capture real-world conditions. In actual practice, investors are more interested in forward-looking forecasting performance. Additionally, overfitting commonly plagues in-sample forecasting. To address these issues, we also perform out-of-sample analyses that restrict the model's available information to 1 month ahead, thereby mitigating look-ahead bias and providing a more realistic evaluation of predictability.

3.2 | Combined Predictor and Forecast Combination

To predict market-level stock returns, we need to aggregate the firm-level IESGRIs to the market level, as individual IESGRIs cannot fully capture overall market movements. We adopt two approaches for this aggregation: (1) For the in-sample analysis, we use a simple cross-sectional mean of $IESGRI_{i,t}$ as a market-level ESG risk indicator. Rigorously, the market-level ESG risk indicator is given as follows:

$$IESGRI_t^A = \frac{1}{N} \sum_{i=1}^{N} IESGRI_{i,t}, \tag{3}$$

where, $IESGRI_t^A$ represents the market-level ESG risk indicator at month t, $IESGRI_{i,t}$ represents the ESG risk of firm i at month t, and N is the number of firms. This aggregated measure retains the intuition underlying individual firms' IESGRIs while mitigating their firm-specific idiosyncratic risk, making it a robust proxy for the overall market's ESG risk. (2) For the out-of-sample analysis, we apply a forecast combination method, which is widely recognised for improving forecast accuracy (e.g., Wang et al. 2023; Rapach et al. 2010). The main advantages of this approach are its ability to integrate data from various economic indicators while notably reducing the volatility of forecasts and maintaining relevance to the real economy. Also, this method enables diversified forecasting by balancing the performance of various predictors, similar to how portfolio diversification mitigates risk (e.g., Gospodinov and Maasoumi 2021; Rapach and Zhou 2022a)[5].

To conduct the forecast combination for the out-of-sample analysis, we first run the predictive model in Equation (2) to predict r_t, using each $IESGRI_{i,t-1}$ as the explanatory variable, restricting the data available to the model up to period t. This process yields the OLS estimates $\hat{\alpha}^{i,t}$ and $\hat{\beta}^{i,t}$, representing the intercept and the coefficient of $IESGRI_{i,t-1}$, respectively. Using these estimates along with $IESGRI_{i,t}$, we forecast the future market excess return as follows:

$$\hat{r}^{(i)}_{t+1|t} = \hat{\alpha}^{i,t} + \hat{\beta}^{i,t} \times IESGRI_{i,t}, \text{ for } i = 1, \ldots, N \tag{4}$$

where $\hat{r}^{(i)}_{t+1|t}$ is the forecasted market excess return for month $t+1$ based on firm i's ESG risk measures at month t. We apply an expanding-window approach, setting the initial period as the first 6 years, with the subsequent sample serving as the out-of-sample period.

We then calculate the market return forecast by taking a weighted average of the market return forecasts based on each firm i's ESG risk information (i.e., $\hat{r}^{(i)}_{t+1|t}$). The equation is as follows:

$$\hat{r}_{t+1|t} = \sum_{i=1}^{N} \omega_{i,t+1|t} \hat{r}^{(i)}_{t+1|t}, \tag{5}$$

where, $\omega_{i,t+1|t}$ for $i = 1, \ldots, N$ are the weights allocated to $\hat{r}^{(i)}_{t+1|t}$, and $\sum_{i=1}^{N} \omega_{i,t+1|t} = 1$.

According to Rapach et al. (2010), a simple forecast combination often outperforms the benchmark models. Consistent with this, we use the equal-weighted average of a large number of individual forecasts as our market return forecast, as it is difficult to beat in practice (Yuan and Zhou 2023). The equation can be expressed as

$$\hat{r}^{\text{Mean}}_{t+1|t} = \frac{1}{N} \sum_{i=1}^{N} \hat{r}^{(i)}_{t+1|t}, \tag{6}$$

where, the market return forecast ($\hat{r}^{\text{Mean}}_{t+1|t}$) is formed by taking the arithmetic mean of the univariate forecasts in Equation (4). We also consider alternative weighting schemes, such as the discounted mean squared prediction error (DMSPE) method used by Rapach et al. (2010) and a nonlinear weighting method applied by Yang (2004). Details of those methods are discussed in Section 5.

3.3 | Forecast Evaluation

3.3.1 | Statistical Performance Evaluation

To evaluate the out-of-sample statistical performance, we apply the widely recognised and practical method proposed by

Campbell and Thompson (2008), using the out-of-sample R^2 statistic. This statistic is used to measure the proportional reduction in mean squared forecast error (MSFE), comparing the performance of the forecast combination with that of the prevailing mean benchmark forecast:

$$R_{OS}^2 = 1 - \frac{\sum_{t=1}^{T}(r_t - \hat{r}_t)^2}{\sum_{t=1}^{T}(r_t - \bar{r}_t)^2}, \tag{7}$$

where, $\hat{r}_t$ denotes the fitted value from the forecast combination over the out-of-sample period, and $\bar{r}_t$ is the forecast from the prevailing mean benchmark, both estimated from the first observation to the last observation at month T.

We test the null hypothesis $H_0 : R_{OS}^2 \leq 0$ against the alternative hypothesis $H_A : R_{OS}^2 > 0$ via the Clark and West (2007) test (CW-test). The objective is to assess whether the forecast combination, which incorporates the incident-based ESG risk information, yields a lower MSFE than the prevailing mean forecast. If the forecast combination produces a lower MSFE, there is evidence of out-of-sample return predictability, reflected in a positive R_{OS}^2.

It is important to note that the R_{OS}^2 statistic in Equation (7) tends to be small due to the substantial unpredictable component in returns. Conversely, a large R_{OS}^2 may indicate overfitting and should be interpreted cautiously. As suggested by Campbell and Thompson (2008) and Dong et al. (2022), an R_{OS}^2 of 0.5% is considered statistically significant for out-of-sample forecasting.

3.3.2 | Economic Performance Evaluation

In addition to evaluating the accuracy of forecasts, it is crucial to investigate if forecasts based on ESG risk information can improve trading outcomes for investors. To determine the economic benefits of using such an indicator, this analysis considers a hypothetical mean-variance investor who reallocates her portfolio between a risky market portfolio and a risk-free asset each month. Specifically, at the end of month t, the investor seeks to maximise the following objective function:

$$\arg\max_{W_{t+1|t}} \left(W_{t+1|t}\hat{r}_{M,t+1|t} - 0.5\gamma W_{t+1|t}^2 \hat{\sigma}_{t+1|t}^2 \right), \tag{8}$$

where, γ denotes the coefficient of the investor's relative risk aversion, $W_{t+1|t}$ represents the weight allocated to the risky market portfolio. The remaining portion, $(1 - W_{t+1|t})$, is implicitly allocated to the risk-free asset. In our analysis, we set $\gamma = 3$, a value commonly used in the literature. We also consider $\gamma = 4$ and 5 as alternative levels of risk aversion for robustness checks. The investor's forecast of the market excess return is denoted by $\hat{r}_{M,t+1|t}$, and the variance of the market excess return for the same period is represented by $\hat{\sigma}_{t+1|t}^2$. The optimal weight allocation at the end of month t is expressed as:

$$W_{t+1|t}^* = \left(\frac{1}{\gamma}\right) \frac{\hat{r}_{M,t+1|t}}{\hat{\sigma}_{t+1|t}^2}, \tag{9}$$

where, the variance of the return $\hat{\sigma}_{t+1|t}^2$ is estimated using a 36-month rolling-window approach in this study. In line with Dong et al. (2022), we consider scenarios that include short selling and allow for up to 100% financial leverage, restricting $W_{t+1|t}^*$ to range between −1 and 2.

The average utility that the investor realises is expressed as:

$$\bar{U}_k = \bar{r}_k - 0.5\gamma\hat{\sigma}_k^2 \quad \text{for } k = 0, 1, \tag{10}$$

where, $\bar{r}_0$ ($\bar{r}_1$) and $\hat{\sigma}_0^2$ ($\hat{\sigma}_1^2$) denote the mean and variance of the portfolio return when the investor applies the prevailing mean benchmark forecast (forecast combination) of $r_{M,t+1|t}$ in Equation (9).

To assess the economic value of using the forecast combination method over the prevailing mean forecast, we calculate the utility gain, defined as the difference between $\bar{U}_1$ and $\bar{U}_0$. This utility gain, also referred to as the increase in certainty-equivalent return gain (CER gain), is computed as:

$$\Delta = \bar{U}_1 - \bar{U}_0 \tag{11}$$

We annualise the monthly CER gain calculated in Equation (11) by multiplying it by Equation (12). The utility gain can be interpreted as the annualised portfolio management fee that the investor would be willing to pay for the ESG risk information.

4 | Empirical Results

4.1 | In-Sample Return Predictability

One of the most commonly used approaches for examining return predictability is a simple linear regression, where market returns are regressed against one lagged predictor of interest. First, we predict market excess returns using the standard predictive model (Equation 2), where χ_t is the $IESGRI^A$ or any of the predictors described in Table A1 in appendix at month t. Our primary interest is in whether the coefficient β is significantly different from zero. A significant β would provide evidence that the market excess return is predictable within the sample period. This straightforward approach offers several insights relevant to existing research on market return predictability.

Panel A of Table 2 presents the predictive power of each predictor over a 1-month horizon. The results show that the incident-based ESG risk indicator, US market returns and risk-free rate are statistically significant at the 5% level across the entire sample period. Notably, one standard deviation increase in the incident-based ESG risk indicator in the current month is associated with a significant decrease of 6.38 in excess returns in the following month, outperforming all traditional fundamental factors considered in this study. These findings highlight the substantial predictive capability of the incident-based ESG risk indicator in forecasting market excess returns, exceeding that of conventional predictors typically explored in the literature.

To explore the extent to which the predictability of the incident-based ESG risk indicator aligns with that of other widely used

TABLE 2 | Results for in-sample univariate and bivariate predictive regressions.

	Panel A: Univariate regressions			Panel B: Bivariate regressions				
Predictor	β	*t*-Stat	R^2_{uni} (%)	β'	*t*-Stat	ψ	*t*-Stat	R^2_{bi} (%)
$IESGRI^A$	−6.38**	−2.16	1.90	—	—	—	—	—
SP500	0.12	1.34	0.70	−6.38**	−2.16	0.12	1.33	2.61
DP	0.60	0.08	−0.53	−6.41**	−2.18	−0.40	−0.05	1.38
DY	2.84	0.38	−0.39	−6.29**	−2.14	1.72	0.23	1.43
EP	−0.04	−1.27	0.74	−6.31**	−2.17	−0.04	−1.24	2.59
DE	6.47*	1.74	1.43	−5.97**	−2.02	5.95	1.63	3.03
BM	0.04	0.98	0.27	−6.09**	−2.12	0.04	0.83	1.93
RF	−7.22*	−1.96	2.08	−5.80**	−2.02	−6.61*	−1.81	3.55
LTY	−9.02	−0.99	0.21	−6.44**	−2.23	−9.27	−1.04	2.16
YS	9.10**	1.98	2.01	−5.62*	−1.94	8.07*	1.79	3.35
SVAR	−0.08	−0.08	−0.53	−6.51**	−2.23	−0.29	−0.29	1.44
INF	−0.45*	−1.72	2.24	−5.62*	−1.85	−0.40	−1.53	3.58
Kitchen Sink	—	—	—	−5.67**	−2.00	—	—	6.32
PCA	−0.003	−1.47	1.81	−6.25**	−2.17	−0.003	−1.42	4.12

Note: Panel A reports the univariate regression slope (β) and the Newey and West (1987) *t*-statistics (*t*-Stat), along with the in-sample R^2_{uni} (%). Panel B reports the results of the bivariate regression models from Equation (12), including coefficients (β' and ψ') and their corresponding Newey and West (1987) *t*-statistics (*t*-Stat), along with R^2_{bi} (%). The results using the kitchen sink method are derived from a single regression model that includes all predictors. The principle component analysis (PCA) approach synthesises information from all control variables and extracts the first principle component to use as a predictor in both univariate and bivariate predictive regressions. The asterisks *, ** and *** denote significance at 10%, 5% and 1% levels, respectively. The sample period spans from January 2007 to December 2022.

predictors, we use an alternative regression model that integrates two predictors: the incident-based ESG risk indicator and one of the popular predictors. The bivariate predictive regression model is expressed as

$$r_{t+1} = \alpha' + \beta' \times IESGRI_t^A + \psi \times Z_t + \epsilon'_{t+1}, \quad (12)$$

where, r_{t+1} represents the market excess returns in month $t+1$, $IESGRI_t^A$ is the incident-based ESG risk indicator, and Z_t denotes one of the predictors listed in Table A1 in Appendix. A statistically significant β' suggests that the incident-based ESG risk indicator continues to have predictive power for market excess returns, even when the influence of the additional predictor is accounted for.

Panel B of Table 2 reports that even after controlling for other widely-used predictors, the incident-based ESG risk indicator remains statistically significant at the 5% level in predicting market excess returns. This finding reinforces the notion that incident-based ESG risk indicator provides unique information that is not captured by traditional financial predictors. Furthermore, the inclusion of popular predictors in the bivariate regression models results in variations in the R^2 values, indicating the incremental explanatory power gained by incorporating ESG risk information into the predictive models. The observed increase in R^2 upon adding the incident-based ESG risk indicator as a predictor suggests that it offers additional explanatory value beyond what is provided by conventional financial predictors. This implies that investors and financial analysts can benefit from considering the incident-based ESG risk indicator alongside conventional financial indicators.

Additionally, we incorporate the information from all predictors in our predictive regression analyses using two distinct strategies. The first is the kitchen sink method, which includes all available predictors in a single regression model without applying variable selection or dimensionality reduction. While this approach can face challenges such as overfitting or multicollinearity when predictors are highly correlated (Rapach and Zhou 2022b), it is a straightforward and effective when the number of predictors is small. The second strategy employs principal component analysis (PCA) method to synthesise information from all control variables. Specifically, we extract the first principal component of the control variables and use it as a predictor in univariate and bivariate predictive regressions.

The results of these two approaches are presented in the last two rows of Table 2. For the kitchen sink method, the coefficient of our primary variable, $IESGRI^A$, remains significant at the 5% level, indicating its robustness in this specification. In the PCA approach, the univariate regression shows that the coefficient of the first principal component is not significant, suggesting that this aggregated control variable does not independently exhibit significant predictive power. However, in the bivariate regression, which includes both the first principal component and $IESGRI^A$, the coefficient of $IESGRI^A$ remains significant at the 5% level. This result further supports the robustness of $IESGRI^A$'s predictive power.

4.2 | Out-of-Sample Return Predictability

While in-sample return predictability has been established, it is widely recognised that strong in-sample results do not necessarily guarantee robust out-of-sample performance (Guidolin et al. 2009; Timmermann 2008; Welch and Goyal 2008). Welch and Goyal (2008) argue that many popular in-sample predictors fail to outperform the prevailing mean benchmark forecast (i.e., random walk with drift) in out-of-sample tests. This is primarily due to look-ahead bias, which occurs when the entire sample is applied at once and does not reflect real-world scenarios. In practice, investors are more interested in real-time forecasting performance. Additionally, overfitting is another issue with in-sample forecasting. To address this concern, we evaluate out-of-sample predictability using both statistical and economic metrics in this section. More specifically, we apply the Clark and West (2007) test to assess statistical significance, while economic significance is evaluated using the CER gain and the Sharpe ratio.

Table 3 reports the out-of-sample R^2_{OS} and the CW-test statistics for out-of-sample forecasting from January 2013 to December 2022, with the initial training period spanning the first 6 years. In Panel A, we measure out-of-sample predictability by applying univariate regressions with each individual predictor. We find that none of the predictors, except for the incident-based ESG risk indicator, is statistical significant at the 5% level. The incident-based ESG risk indicator achieves an R^2_{OS} of 1.31% with a CW-test statistic of 1.66. Most of the other predictors' R^2_{OS} values are negative, and none of them significantly outperform the benchmark model.

TABLE 3 | Results for out-of-sample univariate and bivariate predictive regressions.

	Panel A: Univariate regressions		**Panel B: Bivariate regressions**	
	R^2_{OS} (%)	**CW-test**	**R^2_{OS} (%)**	**CW-test**
IESGRI^A	1.31**	1.66	—	—
SP500	−5.09	−0.52	−0.91	0.09
DP	−0.63	−0.70	0.54	1.07
DY	−1.59	−0.59	0.19	0.42
EP	−1.82	1.28	1.71**	1.67
DE	−0.64	1.36	2.15**	1.72
BM	0.43	0.94	1.16**	2.13
RF	−0.45	1.14	1.66*	1.58
LTY	−7.50	0.94	−0.23	1.25
YS	0.28	0.99	1.40*	1.58
SVAR	−2.98	−0.83	−0.39	0.06
INF	−11.30	1.49	1.28**	1.69
PCA	−4.93	1.32	1.30**	1.67

Note: This table reports the out-of-sample R^2_{OS}(%) and the Clark and West (2007) test statistics applying well-established predictors and incident-based ESG indicator. Statistical significance for R^2_{OS} is based on the Clark and West (2007) test of the MSFE-adjusted statistic for the hypothesis: $H_0 : R^2_{\mathrm{OS}} \leq 0$ versus $H_A : R^2_{\mathrm{OS}} > 0$. The asterisks *, ** and *** denotes out-of-sample statistical significance at the 10%, 5% and 1% levels, respectively. The whole sample period spans from January 2007 to December 2022, with the first 6 years serving as the training sample.
Abbreviations: ESG, Environmental, Social and Governance; MSFE, mean squared forecast error; PCA, principal component analysis.

Next, we examine the out-of-sample predictability by applying bivariate regressions, which include the incident-based ESG risk indicator and one of the traditional predictors, with the results reported in Panel B of Table 3. When our incident-based ESG risk indicator is incorporated, the forecast accuracy for all cases is significantly improved compared to the univariate cases. In seven instances, the forecasts significantly outperform the benchmark model at the 5% or 10% levels. For the cases of EP, DE and YS, the forecasts are even better than the univariate forecast of our incident-based ESG risk indicator. These results suggest that, although these predictors may not be individually significant, they still provide complementary information to the incident-based ESG risk indicator. When included in the model, this additional information enhances the explanation of market return variability, leading to higher R^2_{OS} values.

Subsequently, we evaluate the economic performance of market return forecasts that utilise the incident-based ESG risk indicator. To achieve this, we explore the CER gain and Sharpe ratio. Following the approach of Dong et al. (2022), the weight in Equation (8) is constrained between −1 and 2, allowing for short-selling while capping maximum leverage at 100%. Various investing scenarios, along with their corresponding CER gains and Sharpe ratios, are presented in Table 4. Specifically, we assess different investment strategies by considering risk

TABLE 4 | Results for asset allocation.

	No transaction cost		**10 bps transaction cost**	
	CER gain (%)	**Sharpe ratio**	**CER gain (%)**	**Sharpe ratio**
IESGRI^A	4.55	0.43	4.20	0.32
SP500	−5.09	−0.14	−6.34	−0.28
DP	−1.10	−0.17	−0.60	−0.22
DY	−0.69	−0.08	−0.19	−0.13
EP	1.56	0.38	0.37	0.30
DE	3.20	0.46	2.13	0.39
BM	1.62	0.11	2.23	0.06
RF	0.94	0.37	−0.15	0.29
LTY	2.01	0.41	0.79	0.32
YS	1.98	0.30	1.31	0.21
SVAR	−0.57	0.01	−0.88	−0.10
INF	7.26	0.60	6.75	0.53
PCA	2.35	0.45	1.07	0.37

Note: This table reports the out-of-sample economic performance (CER gain (%) and Sharpe ratio) of various predictors. On a monthly basis, the investor, guided by mean-variance principles, reallocates her wealth between the stock market and the risk-free asset. The CER gain represents the annualised certainty equivalent return differential between the utilisation of the individual predictor and the application of the prevailing mean forecast, with the weight in the risky asset constrained to be no less than −1 and no more than 2 (Dong et al. 2022). We use 36-month rolling window volatility forecasts and use $\gamma = 3$. The Sharpe ratio is the annualised average portfolio excess return over its standard deviation. Both scenarios with no transaction costs and a 10 bps transaction cost are applied. The whole sample period spans from January 2007 to December 2022, with the first 6 years serving as the training sample.
Abbreviations: CER, certainty equivalent return; PCA, principal component analysis.

aversion level of 3. In scenarios where transaction cost are excluded, the incident-based ESG risk indicator achieves an annualised CER gain of 4.55% and a Sharpe ratio of 0.43. This CER gain surpasses the 2% benchmark suggested by Dong et al. (2022) and Pástor and Stambaugh (2000), indicating substantial economic returns for investors. These results surpass those associated with most other variables, with the exception of the dividend payout ratio, which, despite generating substantial gains, does not reach statistical significance. This outcome highlights the effectiveness of the incident-based ESG risk indicator in predicting market excess returns. Additionally, when accounting for realistic transaction cost—set at 10 basis points following Farmer et al. (2023) and Shynkevich (2012)—the $IESGRI^A$ continues to deliver a positive annualised CER gain of 4.20%.

We also examine varying levels of investor risk aversion, focusing on higher risk aversion (values of 4 and 5) to assess whether conservative investors can achieve significant CER gains using the incident-based ESG risk indicator or other predictors. Table B1 in Appendix shows that the ESG risk indicator remains significant at higher risk aversion levels, with CER gains exceeding the critical 2% threshold at Level 4 but falling below it at Level 5. These findings, robust to transaction costs (Tables B2 and B3 in Appendix), highlight the indicator's predictive value.

We acknowledge the economic and cultural diversity among Eurozone countries and address this by conducting in-sample and out-of-sample analyses on country portfolio returns using individual country IESGRI data. Additionally, we use the pooled regression and the panel data model, with the latter controlling for country and time fixed effects. The in-sample analysis shows that not all country-specific coefficients of $IESGRI^A$s are significant, and the out-of-sample results indicate that forecast combination methods do not consistently outperform the benchmark model. However, both pooled regression and panel data models reveal significant predictability at the Eurozone market Level[6].

4.3 | Forecasting Performance Over Business Cycle

Numerous studies have examined market predictability across different phases of the business cycle, showing that the predictive power varies in good times and bad times (e.g., Rapach et al. 2010). In our analysis, we specifically investigate the significance of the incident-based ESG risk indicator for return predictability during both expansion and non-expansion periods. The equation is expressed as:

$$R^2_{OS,c} = 1 - \frac{\sum_{t=1}^{T} I_t^c (r_t - \hat{r}_t)^2}{\sum_{t=1}^{T} I_t^c (r_t - \bar{r}_t)^2} \quad \text{for} \quad c = \text{EXP, NON-EXP}, \tag{13}$$

where, $I_t^{EXP}\left(I_t^{NON-EXP}\right)$ is a dummy variable that equals to 1 if month t falls within an expansion (non-expansion) period, and 0 otherwise.

In this paper, we use the Business Cycle Clock provided by Eurostat to identify the phases of the business cycle. Eurostat classifies the business cycle into six phases: expansion with accelerating growth, expansion with decelerating growth, slow down, recession with decelerating growth, recession with accelerating growth and recovery. We simplify this classification into two categories: expansion periods and non-expansion periods. The expansion periods include the expansion with accelerating growth and with decelerating growth, spanning from July 2013 to August 2014, December 2014 to November 2018, and August 2020 to February 2022. The remaining periods are identified as the non-expansion periods, including slowdown, recession with decelerating growth and with accelerating growth, as well as recovery.

Table 5 reports the performance of various predictors across different phases of the business cycle. When analysing the performance of the incident-based ESG risk indicator, we find that its predictive power is stronger during non-expansion periods compared to expansion periods. Specifically, the R^2_{os} value of $IESGRI^A$ is 1.55% and is statistically significant at the 5% level during non-expansion periods. In contrast, the predictive power of the incident-based ESG risk indicator drops during expansion periods, as indicated by a lower R^2_{os} value. Although the R^2_{os} value remains positive in expansion periods, it is not statistically significant according to the CW-test, suggesting that the indicator's predictability is primarily concentrated in non-expansion periods. These findings are consistent with the established literature, which indicates that return predictability tends to be stronger during economic downturns and weaker during periods of economic growth (e.g., Rapach

TABLE 5 | Predictability in expansion and non-expansion periods.

	Expansion		**Non-expansion**	
	R^2_{os} **(%)**	**CW-test**	R^2_{os} **(%)**	**CW-test**
$IESGRI^A$	1.09	1.07	1.55**	1.74
SP500	−4.66	−0.53	−5.48	−0.3
DP	−1.41	−0.93	0.32	0.54
DY	−4.03	−0.99	1.49**	1.87
EP	−0.16*	1.61	−4.01	−0.58
DE	1.65**	1.75	−3.96	−0.25
BM	−0.39	0.05	1.59**	1.81
RF	1.48	1.46	−3.23	−0.33
LTY	−6.30	1.09	−8.95	−0.26
YS	1.39	1.18	−1.32	−0.25
SVAR	−0.31	0.20	−6.20	−0.99
INF	−2.18**	1.79	−22.41	0.45

Note: The table reports out-of-sample statistics of market return predictability during expansion and non-expansion periods. The phases of business cycle are defined by the Business Cycle Clock. $IESGRI^A$ is the incident-based ESG risk indicator. Statistical significance for R^2_{OS} is based on the Clark and West (2007) test of the MSFE-adjusted statistic for the hypothesis: $H_0: R^2_{OS} \leq 0$ versus $H_A: R^2_{OS} > 0$. The asterisks *, ** and *** denotes out-of-sample statistical significance at the 10%, 5% and 1% levels, respectively. The whole sample period spans from January 2007 to December 2022, with the first 6 years serving as the training sample.
Abbreviations: ESG, Environmental, Social and Governance; MSFE, mean squared forecast error.

et al. 2010; Rapach and Zhou 2013; Pettenuzzo et al. 2014; Cujean and Hasler 2017).

Regarding the performance of other predictors, we find that dividend yields can predict market returns during non-expansion periods, with an R^2_{os} value of 1.49%, which is also significant at the 5% level. This result extends the findings of Golez and Koudijs (2018), which show that dividend yields consistently predict market returns in the Netherlands, the UK and the United States, especially during recessions. We also find that the book-to-market ratio, dividend payout ratio, and inflation are significant predictors in certain scenarios, though none of them is consistently significant across the entire out-of-sample forecasting period.

5 | Additional Analysis

5.1 | Forecasting Country Portfolio Returns

In this section, we extend our analysis to examine whether the incident-based ESG risk indicator can effectively predict excess market returns for individual member countries within the Eurozone. Our study focuses on nine major Eurozone countries: Austria, Belgium, Finland, France, Germany, Ireland, Italy, the Netherlands and Spain. We use the Eurozone Country Portfolios Index data from the Kenneth French Data Library[7].

Table 6 presents both the in-sample and out-of-sample forecasting results for each country in the Eurozone. The in-sample analysis covers the period from January 2007 to December 2022, while the out-of-sample analysis uses the first 6 years as the training sample. The results indicate that the incident-based ESG risk indicator significantly predicts portfolio returns for most countries at the 5% or 10% significance levels. Specifically, Belgium, Italy and Spain exhibit R^2_{OS} values exceeding 1%, with statistical significance at the 5% level, indicating strong predictive power of the incident-based ESG risk indicator in these markets. In contrast, the portfolio returns for Finland and Ireland do not demonstrate statistical significance, while Austria, France, Germany, and the Netherlands show significance only at the 10% level.

These findings suggest that while the incident-based ESG risk indicator offers valuable insights into return predictability for certain regional markets within the Eurozone, its effectiveness is not consistent across all countries. This heterogeneity may be attributed to differences in market structures, economic conditions, or the varying impact of ESG incidents across these countries.

5.2 | Alternative Market Portfolios in the Eurozone

In this section, we extend our investigation to examine whether the incident-based ESG risk indicator can be applied to predict alternative market indices in the Eurozone. For this purpose, we use the MSCI EMU Total Return Index from Datastream and the Europe Portfolios Return Index from the Kenneth French Library for robustness checks.

Table 7 reports the out-of-sample forecasting results of the MSCI EMU Total Return Index and the Europe Portfolios Return Index in Panel A and Panel B, respectively. In Panel A, the R^2_{OS} value for the incident-based ESG risk indicator is 1.33%, which is statistically significant at the 5% level according to the CW-test. The annualised CER gain without transaction costs (with a 10 bps transaction cost) is 4.78% (4.22%), and the Sharpe ratio is 0.50 (0.45) for a risk aversion level of 3. Alternative risk aversion levels of 4 and 5 also support our main findings (see Table B2 in Appendix). The evidence reported in Panel B suggests that, even after accounting for a 10 bps transaction fee, the incident-based ESG risk indicator remains robust, with a CER gain of 4.33%. Overall, our findings confirm that the incident-based ESG risk indicator also predicts excess returns in other alternative markets.

TABLE 6 | In-sample and out-of-sample results of country portfolio returns.

	β	NW-test	R^2 (%)	R^2_{OS} (%)	CW-test
Austria	−7.87**	−2.27	2.67	0.97*	1.36
Belgium	−5.18*	−1.83	1.74	1.17**	1.75
Finland	−6.22*	−1.97	2.02	0.54	1.07
France	−6.11**	−2.14	2.37	1.07*	1.49
Germany	−5.61*	−1.93	1.94	0.82*	1.35
Ireland	−8.13*	−1.97	2.01	0.16	0.43
Italy	−6.66**	−1.98	2.05	1.52**	1.78
Netherlands	−6.21**	−2.12	2.34	1.02*	1.35
Spain	−7.80**	−2.48	3.17	1.87**	1.99

Note: This table reports in-sample and out-of-sample statistical performance of the incident-based ESG risk indicator on return predictability across the main countries in the Eurozone. Statistical significance for R^2 and R^2_{OS} are based on the Newey and West (1987) test, and Clark and West (2007) test of the MSFE-adjusted statistic for the hypothesis: $H_0 : R^2_{OS} \leq 0$ versus $H_A : R^2_{OS} > 0$, respectively. The asterisks *, **, and *** denote in-sample and out-of-sample statistical significance at the 10%, 5% and 1% levels, respectively. The whole sample period spans from January 2007 to December 2022, with the first 6 years serving as the training sample.
Abbreviations: ESG, Environmental, Social and Governance; MSFE, mean squared forecast error.

TABLE 7 | Out-of-sample results of alternative market portfolios.

	R^2_{OS} (%)	CW-Test	CER gain (%) No cost	CER gain (%) 10 bps cost	Sharpe ratio No cost	Sharpe ratio 10 bps cost
Panel A: MSCI EMU Total Return Index						
IESGRI^A	1.33**	1.67	4.78	4.22	0.50	0.45
SP500	−4.99	−0.57	−5.53	−6.92	−0.10	−0.19
DP	−0.64	−0.64	−1.21	−1.22	−0.01	−0.02
DY	−1.55	−0.54	−0.85	−1.05	0.03	0.00
EP	−1.81	1.17	2.34	2.25	0.45	0.44
DE	−0.75*	1.30	4.07	4.05	0.53	0.53
BM	0.39	0.88	1.65	1.51	0.23	0.20
RF	−0.44	1.09	1.82	1.84	0.44	0.44
LTY	−7.20	0.87	2.82	2.81	0.48	0.47
YS	0.34	1.00	2.68	2.58	0.40	0.39
SVAR	−2.74	−0.80	−0.19	−0.64	0.13	0.09
INF	−10.84*	1.43	7.01	6.80	0.62	0.61
Panel B: Europe Portfolios Return Index						
IESGRI^A	1.35**	1.69	4.98	4.33	0.56	0.51
SP500	−5.74	−0.48	−4.35	−5.84	0.01	−0.08
DP	−0.54	−0.71	−0.90	−0.90	0.11	0.11
DY	−1.49	0.61	−1.13	−1.34	0.09	0.07
EP	−1.70*	1.55	7.04	7.07	0.66	0.65
DE	−0.05*	1.59	8.41	8.43	0.71	0.71
BM	0.17	0.54	0.84	0.71	0.23	0.21
RF	0.08*	1.38	5.72	5.76	0.60	0.60
LTY	−7.91	1.21	6.61	6.65	0.64	0.64
YS	0.78	1.26	5.08	5.04	0.55	0.55
SVAR	−3.03	−0.81	0.56	0.07	0.26	0.22
INF	−10.31*	1.61	8.83	8.63	0.73	0.72

Note: This table reports the out-of-sample statistical and economic performance (CER gain (%) and Sharpe ratio) of the incident-based ESG risk indicator on return predictability for the MSCI EMU Total Return Index and the Europe Portfolios Return Index. Statistical significance for R^2_{OS} is based on the Clark and West (2007) test of the MSFE-adjusted statistic for the hypothesis: $H_0 : R^2_{OS} \leq 0$ versus $H_A : R^2_{OS} > 0$. For economic gains, we assume a risk aversion level of 3, and the portfolio weight in equities lies between −1 to 2. The CER gain represents the annualised certainty equivalent return differential between the incident-based ESG risk indicator and the prevailing mean forecast. The Sharpe ratio is the annualised average portfolio excess return over its standard deviation. Both scenarios with no transaction costs and a 10 bps transaction cost are applied. The asterisks *, ** and *** denotes out-of-sample statistical significance at the 10%, 5% and 1% levels, respectively. The whole sample period spans from January 2007 to December 2022, with the first 6 years serving as the training sample.
Abbreviations: CER, certainty equivalent return; ESG, Environmental, Social and Governance; MSFE, mean squared forecast error.

5.3 | Stricter Financial Constraints

This section explores alternative investing constraints in out-of-sample forecasting by imposing two additional financial restrictions on asset allocation weights. The first restriction, following Campbell and Thompson (2008) and Wang et al. (2019), sets the range from 0 to 1.5, prohibiting short selling and limiting maximum financial leverage to 50%. Furthermore, similar to Pettenuzzo et al. (2014), we consider a more stringent scenario where weight constraint is set between 0 and 1, disallowing both short selling and leverage.

Table 8 reports the out-of-sample economic performance of the incident-based ESG risk indicator and traditional predictors under two alternative financial restrictions. The analysis shows that the incident-based ESG risk indicator achieves an annualised CER gain of 3.34% and Sharpe ratio of 0.40 behind the constraint of 50% financial leverage. Additionally, the dividend payout ratio exhibits notable economic gains, generating a CER gain of 6.20% and a Sharpe ratio of 0.57 under the 0–1.5 portfolio weight constraint. This economic benefit remains significant, yielding CER gain of 3.72% even when both short selling and leverage are prohibited for each trade. Overall, these results suggest that the incident-based ESG risk indicator demonstrates substantial gains under various conditions. Even when the financial constraints become stricter, the ESG risk indicator still performs well, highlighting its capability to enhance portfolio performance.

TABLE 8 | Out-of-sample results of stricter financial constraints.

	Weight constraint [0, 1.5]				Weight constraint [0, 1]			
	CER gain (%)		Sharpe ratio		CER gain (%)		Sharpe ratio	
	No cost	10 bps cost	No cost	10 bps cost	No cost	10 bps cost	No cost	10 bps cost
$IESGRI^A$	3.34	2.94	0.40	0.35	3.72	3.45	0.45	0.42
SP500	1.13	0.30	0.21	0.13	1.06	0.51	0.16	0.09
DP	−0.32	−0.32	0.03	0.02	−0.21	−0.20	0.03	0.02
DY	1.86	1.75	0.25	0.23	1.43	1.36	0.20	0.18
EP	4.57	4.57	0.49	0.49	5.09	5.13	0.51	0.51
DE	6.20	6.22	0.57	0.57	6.22	6.27	0.59	0.59
BM	2.40	2.32	0.33	0.31	2.29	2.21	0.32	0.30
RF	3.86	3.89	0.46	0.46	4.91	4.97	0.50	0.50
LTY	4.19	4.20	0.47	0.47	5.00	5.05	0.51	0.50
YS	3.17	3.13	0.39	0.39	3.03	3.03	0.37	0.36
SVAR	0.86	0.55	0.18	0.15	1.73	1.52	0.24	0.21
INF	6.33	6.22	0.58	0.57	5.64	5.59	0.56	0.56

Note: This table reports the economic performance (CER gain (%) and Sharpe ratio) of the incident-based ESG risk indicator and other well-known predictors on return predictability for the Morningstar Eurozone Index Return. For economic gains, we assume risk aversion of 3, and the portfolio weight in equities lies between 0 to 1.5 and 0 to 1, respectively. The CER gain represents the annualised certainty equivalent return differential between the incident-based ESG risk indicator and the prevailing mean forecast. The Sharpe ratio is the annualised average portfolio excess return over its standard deviation. Both scenarios with no transaction costs and a 10 bps transaction cost are applied. The whole sample period spans from January 2007 to December 2022, with the first 6 years serving as the training sample.
Abbreviations: CER, certainty equivalent return; ESG, Environmental, Social and Governance.

5.4 | More Complicated Weights on Forecasts

In our main analysis, we equally allocate weights to individual forecasts when constructing aggregate forecasts. To evaluate the influence of different weighting strategies on forecasting performance, we explore forecast combination by applying discounted mean squared prediction error (DMSPE) method proposed by Stock and Watson (2004), and a non-linear weight method suggested by Yang (2004). The DMSPE method calculates combination weights based on the historical forecasting performance of each individual forecast. The weight assigned to the i-th forecast at time t, denoted by $\omega_{i,t}$, is given by:

$$\omega_{i,t} = \frac{\phi_{i,t}^{-1}}{\sum_{k=1}^{n} \phi_{k,t}^{-1}}, \tag{14}$$

where

$$\phi_{i,t} = \sum_{s=m+1}^{t-1} \theta^{t-s} (r_s - \hat{r}_{i,s})^2, \tag{15}$$

with $\phi_{i,t}$ representing the DMSPE for forecast i at time t, m is the length of the in-sample period, s indexes past time periods starting from $m+1$, which is the beginning of the out-of-sample period, k is used to sum the inverse DMSPE values of all forecasts, and θ is the discount factor. In this method, individual forecasts with lower DMSPE values (i.e., those with better recent performance) receive higher weights in the combination. When the discount factor $\theta = 1$, all past forecast errors are weighted equally without discounting. In a smaller θ, such as 0.9, higher weights are given to more accurate forecasts. We use discount factors θ of 1 and 0.9, following Rapach et al. (2010), Zhang et al. (2019) and Zhu and Zhu (2013), to denote these as DMSPE (1) and DMSPE (0.9), respectively.

Additionally, Yang (2004) argues that linear forecast combination can lead to poor performance due to their high variability in the ex-ante estimates of combination weights. To address this issue, the literature suggests using a nonlinear weighting scheme for updating the combination weights. In accordance with Wang et al. (2019) and Yang (2004), we employ a simplified exponential transformation of MSPE as a nonlinear weight approach, expressed as:

$$\omega_{i,t} = \frac{\pi_i \exp\left(-\lambda \sum_{s=1}^{t-1} (r_s - \hat{r}_{i,s})^2\right)}{\sum_{k=1}^{n} \pi_k \exp\left(-\lambda \sum_{s=1}^{t-1} (r_s - \hat{r}_{k,s})^2\right)}, \tag{16}$$

where, λ is the weighting parameter that determines the sensitivity of the exponential transformation to past forecast errors, and π is the scaling parameter for each forecast. We follow Wang et al. (2019) to set the weighting parameter λ and π to be 1 for simplicity.

Table 9 reports the results of alternative weighting approaches for forecast combinations. All methods exhibit similar out-of-sample performance compared to the simple forecast combination. The DMSPE (1) and DMSPE (0.9) methods achieve R^2_{OS} value of 1.64% and 1.84%, respectively, compared to 1.31% R^2_{OS} for the simple forecast combination. The nonlinear weight method yields an R^2_{OS} value of 1.26%. The CER gains and Sharpe ratios obtained using different estimation approaches also align closely with the main analysis. The results suggest that while different weighting approaches on forecasts can affect forecast performance, the impact is not substantial in our context. This underscores the robustness of our main analysis and the broad applicability of the the equal-weight average measurement (Claeskens et al. 2016; Rapach et al. 2010).

6 | Explanations

6.1 | Firm Fundamentals

A potential explanation for return predictability is that ESG information is linked to the future fundamentals of firms (Pedersen et al. 2021). This section explores the relationship between the incident-based ESG risk indicator and firms' future fundamentals. We consider two variables as proxies for firm performance: return on equity (ROE) (Chu et al. 2024) and return on assets (ROA) (Pedersen et al. 2021). The regression is expressed as:

$$Y_{t+1} = \alpha + \beta \times IESGRI_t^A + \epsilon_{t+1}, Y = ROE, ROA, \quad (17)$$

where, Y_{t+1} denotes aggregate firm performance at month $t+1$, and $IESGRI_t^A$ represents the incident-based ESG risk indicator at month t. Consistent with the main analysis, we use the Newey and West (1987) test to address issues arising from autocorrelation and heteroskedasticity.

Panel A in Table 10 reports the regression results on firm performance using the incident-based ESG risk indicator. The findings indicate that the incident-based ESG risk indicator significantly predicts ROE and ROA at the 1% level, with t-statistics of 7.39 and 7.51, respectively. These results are associated with the findings of Pedersen et al. (2021), who showed that positive evaluation criteria of the ESG score are related to firm's fundamentals. However, our analysis extends their understanding by demonstrating that negative evaluation criteria, such as the incident-based ESG risk indicator, are also informative. This underscores the importance of considering not only positive ESG scores but also potential risks and incidents-related indicators in assessing firms' performance and market predictability.

6.2 | Cash Flow and Discount Rate Channels

According to Cochrane (2011), if a variable can predict market returns, it must do so through the cash flow channel, the discount rate channel, or both. To explore this perspective, we test whether the market return predictability of the incident-based ESG risk indicator arises from either the cash flow channel, the discount rate channel, or both (Campbell et al. 2010). Following Wang et al. (2019), we use the earnings growth rate (EARN) as a proxy for the cash flow and dividend price ratio (DP) as a proxy for the discount rate. The regression is expressed as:

$$Y_{t+1} = \alpha + \beta \times IESGRI_t^A + \epsilon_{t+1}, Y = EARN, DP, \quad (18)$$

where, Y represents either the cash flow (EARN) or the discount rate (DP).

Panel B in Table 10 reports the predictive performance of the incident-based ESG risk indicator on the cash flow and discount rate variables. The results suggest that the ESG indicator is significantly related to the cash flow channel under the 5% significance level. This finding is consistent with the argument made by Derrien et al. (2022) that if a firm's ESG information can predict future earnings, it may affect the firm's stock market values. Additionally, Inard (2023) indicates that ESG factors can impact a firm's revenues, thereby creating uncertainty about its cash flow. On the other hand, the incident-based ESG risk indicator also predicts the discount rate channel at the 1% significance level. One explanation for this is that ESG incidents reported in the media are associated with investors' perceptions of risk about the firm's operations. This explanation aligns with Bolton and Kacperczyk (2021)'s example of carbon risk, where

TABLE 10 | Results of the firm fundamentals, cash flow and discount rate channels with the incident-based ESG risk indicator.

	β	t-Stat	R^2 (%)
	Panel A: Firm's fundamentals		
ROE	3.86***	7.39	22.51
ROA	3.95***	7.51	23.07
	Panel B: Cash flow and discount rate channels		
EARN	−0.36**	−4.10	8.20
DP	0.49***	3.87	7.37

Note: This table reports results of regressions on return on equity (ROE), return on assets (ROA), earnings growth rate (EARN), and dividend price ratio (DP) with the incident-based Environmental, Social and Governance (ESG) risk indicator. t-Stat and R^2 are calculated by Newey and West (1987) two-sided test. The asterisks *, ** and *** denote significance at 10%, 5% and 1% levels, respectively. The sample period spans from January 2007 to December 2022.

TABLE 9 | Out-of-sample results of alternative weights on forecasts.

	R^2_{OS} (%)	CW-Test	CER gain (%)		Sharpe ratio	
			No cost	10 bps cost	No cost	10 bps cost
DMSPE (1)	1.64**	1.91	5.50	4.93	0.53	0.48
DMSPE (0.9)	1.84**	2.07	5.92	5.35	0.57	0.51
Yang (2004)	1.26**	1.69	4.47	3.97	0.45	0.41

Note: This table reports the out-of-sample statistical and economic performance of the incident-based ESG risk indicator on return predictability using various weighting methods on forecasts. Statistical significance for R^2_{OS} is based on the Clark and West (2007) test of the MSFE-adjusted statistic for the hypothesis: $H_0: R^2_{OS} \leq 0$ versus $H_A: R^2_{OS} > 0$. For economic gains, we assume a risk aversion level of 3, and the portfolio weight in equities lies between −1 to 2. The CER gain represents the annualised certainty equivalent return differential between the incident-based ESG risk indicator and the prevailing mean forecast. The Sharpe ratio is the annualised average portfolio excess return over its standard deviation. Both scenarios with no transaction costs and a 10 bps transaction cost are applied. The asterisks *, ** and *** denotes out-of-sample statistical significance at the 10%, 5% and 1% levels, respectively. The whole sample period spans from January 2007 to December 2022, with the first 6 years serving as the training sample.
Abbreviations: CER, certainty equivalent return; DMSPE, discounted mean squared prediction error; ESG, Environmental, Social and Governance.

investors may demand excess returns to compensate for perceived risks.

7 | Conclusion

In this article, we study the evidence of return predictability in the Eurozone stock market by constructing an incident-based ESG risk indicator, both in-sample and out-of-sample, covering the period from the beginning of 2007 to the end of 2022. Our motivation stems from the growing interest in sustainable investments. Europe continues to maintains a dominant position in sustainable fund, representing 84% of global sustainable fund assets (Morningstar 2024). Furthermore, in the Larry Fink's Annual Letters in 2024, BlackRock has invested 138 billion dollars to sustainable investment strategies[8]. Despite this, ESG-related variables have not yet been integrated into the established 'factor zoo', as discussed by Feng et al. (2020). Nonetheless, Rapach and Zhou (2022a) suggest that ESG factors could serve as a novel variable for return predictability. This raises our interest in exploring the relationship between market returns and an incident-based ESG risk indicator.

Our study is the first to use purely objective ESG information in the context of return prediction. By collecting a comprehensive data set from RepRisk, which covers public firms in the Eurozone, we find that the Eurozone market returns can be predicted using ESG information, both in-sample and out-of-sample. Specifically, the predictive power of the incident-based ESG risk indicator is statistically significant at the 5% level, with an out-of-sample R^2_{OS} of 1.31%. We further access its economic significance and find that investors can benefit from this indicator, achieving a CER gain of 4.55% and a Sharpe ratio of 0.43. Moreover, our analysis extends to several individual markets within the Eurozone, with notable predictive power observed in Belgium, Italy, and Spain, both in-sample and out-of-sample. This highlights the potential for broader application of the incident-based ESG risk indicator in these markets.

Additionally, our study investigates the heterogeneity of return predictability across expansion and non-expansion periods. We find the predictability is concentrated in non-expansion periods. A weaker economy tends to increase investor disagreement regarding ESG news, causing returns to react to past incidents (Cujean and Hasler 2017). This finding aligns with existing literature suggesting that returns are more predictable during recessions and periods of high volatility (e.g., Bouri et al. 2023; Rapach et al. 2010; Wang et al. 2019; Zhang et al. 2019).

Our empirical results are robust across various robustness checks. Specifically, we incorporate several popular predictors from the literature into our regression models and apply alternative representative market indices, such as MSCI EMU Index and Kenneth French's Europe Portfolios Return Index. Furthermore, the results remain robust under conditions of higher risk aversion, stricter financial constraints, the inclusion of transaction costs, and different weighting methods for return forecasts.

Moreover, we explain the predictability through the firm fundamentals, cash flow, and discount rate channels. Our findings suggest that firms are significantly impacted by the incident-based ESG risk indicator, linking their performance to ESG-related information. ESG incidents can affect the measurement of firm revenues, highlighting the presence of a cash flow channel. Simultaneously, investor perceptions of risk are influenced by the uncertainty surrounding ESG incidents, indicating the relevant of a discount rate channel.

In practical terms, our findings suggest that investors should pay close attention to the incident-based ESG risk indicator to enhance their portfolio performance. By incorporating such ESG considerations into their investment strategies, investors not only align their portfolios with their values but also gain a competitive advantage in an increasingly sustainability-focused market environment.

Acknowledgements

We would like to thank the Editor and two anonymous reviewers for their valuable comments. This paper is based on the second chapter of Zhuoran Li's PhD thesis. All authors are listed in alphabetical order.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from WRDS. Restrictions apply to the availability of these data, which were used under license for this study.

Endnotes

[1] According to the Financial Times, in 2022, exchange traded funds (ETF) focused on ESG goals constituted 65% of the total net inflows into European ETFs. European quarterly flows into ESG funds keep increasing from 2020.

[2] https://www.ecb.europa.eu/mopo/eaec/html/index.en.html.

[3] We also consider Germany 3 month government bond return as the risk-free rate, the results are similar.

[4] https://sites.google.com/view/agoyal145.

[5] We adopt different aggregation approaches for in-sample and out-of-sample analysis because (1) the forecast combination method produces only a forecast of future returns, which cannot be directly used for further economic analysis of the underlying mechanisms. In contrast, the cross-sectional mean of the predictors provides an intuitive measure that can be employed in economic interpretations. (2) For out-of-sample analysis, it is well established that forecast combination significantly enhances out-of-sample forecast accuracy as mentioned above, whereas relying solely on the cross-sectional mean may lead to a loss of complementary information contained within individual predictors.

[6] Detailed results are available upon request.

[7] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[8] https://www.blackrock.com/corporate/investor-relations/larry-fink-annual-chairmans-letter.

References

Abad, P., H. Chuliá, and M. Gómez-Puig. 2014. "Time-Varying Integration in European Government Bond Markets." *European Financial*

Management 20, no. 2: 270–290. https://doi.org/10.1111/j.1468-036X.2011.00633.x.

Agoraki, M.-E. K., M. Giaka, D. Konstantios, and V. Patsika. 2023. "Firms' Sustainability, Financial Performance, and Regulatory Dynamics: Evidence From European Firms." *Journal of International Money and Finance* 131: 102785. https://www.sciencedirect.com/science/article/pii/S0261560622001887.

Allen, F., and W.-L. Song. 2005. "Financial Integration and EMU." *European Financial Management* 11, no. 1: 7–24. https://doi.org/10.1111/j.1354-7798.2005.00273.x.

Ang, A., and G. Bekaert. 2007. "Stock Return Predictability: Is It There?" *Review of Financial Studies* 20, no. 3: 651–707. https://doi.org/10.1093/rfs/hhl021.

Aouadi, A., and S. Marsat. 2018. "Do ESG Controversies Matter for Firm Value? Evidence From International Data." *Journal of Business Ethics* 151, no. 4: 1027–1047. https://doi.org/10.1007/s10551-016-3213-8.

Baba Yara, F., M. Boons, and A. Tamoni. 2021. "Value Return Predictability Across Asset Classes and Commonalities in Risk Premia." *Review of Finance* 25, no. 2: 449–484. https://doi.org/10.1093/rof/rfaa011.

Bansal, R., D. A. Wu, and A. Yaron. 2022. "Socially Responsible Investing in Good and Bad Times." *Review of Financial Studies* 35, no. 4: 2067–2099. https://doi.org/10.1093/rfs/hhab072.

Bates, J. M., and C. W. J. Granger. 1969. "The Combination of Forecasts." *Journal of the Operational Research Society* 20, no. 4: 451–468. https://doi.org/10.1057/jors.1969.103.

Berg, F., K. Fabisik, and Z. Sautner. 2021. "Is History Repeating Itself? The (Un)Predictable Past of ESG Ratings." https://papers.ssrn.com/abstract=3722087.

Berg, F., J. F. Kölbel, A. Pavlova, and R. Rigobon. 2021. "ESG Confusion and Stock Returns: Tackling the Problem of Noise." https://papers.ssrn.com/abstract=3941514.

Bissoondoyal-Bheenick, E., R. Brooks, and H. X. Do. 2023. "ESG and Firm Performance: The Role of Size and Media Channels." *Economic Modelling* 121: 106203. https://www.sciencedirect.com/science/article/pii/S0264999323000159.

Bofinger, Y., K. J. Heyden, and B. Rock. 2022. "Corporate Social Responsibility and Market Efficiency: Evidence From ESG and Misvaluation Measures." *Journal of Banking & Finance* 134: 106322. https://www.sciencedirect.com/science/article/pii/S0378426621002739.

Bolton, P., and M. Kacperczyk. 2021. "Do Investors Care about Carbon Risk?" *Journal of Financial Economics* 142, no. 2: 517–549. https://www.sciencedirect.com/science/article/pii/S0304405X21001902.

Bouri, E., L. Rognone, A. Sokhanvar, and Z. Wang. 2023. "From Climate Risk to the Returns and Volatility of Energy Assets and Green Bonds: A Predictability Analysis Under Various Conditions." *Technological Forecasting and Social Change* 194: 122682. https://www.sciencedirect.com/science/article/pii/S0040162523003670.

Campbell, J. Y., C. Polk, and T. Vuolteenaho. 2010. "Growth or Glamour? Fundamentals and Systematic Risk in Stock Returns." *Review of Financial Studies* 23, no. 1: 305–344. https://doi.org/10.1093/rfs/hhp029.

Campbell, J. Y., and S. B. Thompson. 2008. "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?" *Review of Financial Studies* 21, no. 4: 1509–1531. https://doi.org/10.1093/rfs/hhm055.

Cao, J., S. Titman, X. Zhan, and W. Zhang. 2023. "ESG Preference, Institutional Trading, and Stock Return Patterns." *Journal of Financial and Quantitative Analysis* 58, no. 5: 1843–1877. https://www.cambridge.org/core/journals/journal-of-financial-and-quantitative-analysis/article/esg-preference-institutional-trading-and-stock-return-patterns/6FE00808FC61893DF3F9D983136BD8B3.

Capelle-Blancard, G., and A. Petit. 2019. "Every Little Helps? ESG News and Stock Market Reaction." *Journal of Business Ethics* 157, no. 2: 543–565. https://doi.org/10.1007/s10551-017-3667-3.

Cederburg, S., T. L. Johnson, and M. S. O'Doherty. 2023. "On the Economic Significance of Stock Return Predictability." *Review of Finance* 27, no. 2: 619–657. https://doi.org/10.1093/rof/rfac035.

Charles, A., O. Darné, and J. H. Kim. 2017. "International Stock Return Predictability: Evidence From New Statistical Tests." *International Review of Financial Analysis* 54: 97–113. https://www.sciencedirect.com/science/article/pii/S1057521916300904.

Chen, B., and K. Maung. 2023. "Time-Varying Forecast Combination for High-Dimensional Data." *Journal of Econometrics* 237, no. 2, Part C: 105418. https://www.sciencedirect.com/science/article/pii/S0304407623000556.

Chen, S., X. Han, Z. Zhang, and X. Zhao. 2023. "Esg Investment in China: Doing Well by Doing Good." *Pacific-Basin Finance Journal* 77: 101907. https://www.sciencedirect.com/science/article/pii/S0927538X22002025.

Cheng, R., H. Kim, and D. Ryu. 2024. "Esg Performance and Firm Value in the Chinese Market." *Investment Analysts Journal* 53, no. 1: 1–15. https://doi.org/10.1080/10293523.2023.2218124.

Chu, L., K. Wang, B. Zhang, and G. Zhou. 2024. " ESG and the stock market: Is ESG Exposure Systematic?" https://papers.ssrn.com/abstract=3869272.

Claeskens, G., J. R. Magnus, A. L. Vasnev, and W. Wang. 2016. "The Forecast Combination Puzzle: A Simple Theoretical Explanation." *International Journal of Forecasting* 32, no. 3: 754–762. https://www.sciencedirect.com/science/article/pii/S0169207016000327.

Clark, T. E., and K. D. West. 2007. "Approximately Normal Tests for Equal Predictive Accuracy in Nested Models." *Journal of Econometrics* 138, no. 1: 291–311. https://www.sciencedirect.com/science/article/pii/S0304407606000960.

Cochrane, J. H. 2011. "Presidential Address: Discount Rates." *Journal of Finance* 66, no. 4: 1047–1108. https://doi.org/10.1111/j.1540-6261.2011.01671.x.

Cujean, J., and M. Hasler. 2017. "Why Does Return Predictability Concentrate In Bad Times?" *The Journal of Finance* 72, no. 6: 2717–2758. https://doi.org/10.1111/jofi.12544.

Cumming, D., H. Farag, and S. Johan. 2024. "Sustainability, Climate Change and Financial Innovation: Future Research Directions." *European Financial Management* 30, no. 2: 675–679. https://doi.org/10.1111/eufm.12481.

Dai, R., H. Liang, and L. Ng. 2021. "Socially Responsible Corporate Customers." *Journal of Financial Economics* 142, no. 2: 598–626. https://www.sciencedirect.com/science/article/pii/S0304405X20300325.

Derrien, F., P. Krueger, A. Landier, and T. Yao. 2022. " ESG News, Future Cash Flows, and Firm Value." https://papers.ssrn.com/abstract=3903274.

Dong, X., Y. Li, D. E. Rapach, and G. Zhou. 2022. "Anomalies and the Expected Market Return." *Journal of Finance* 77, no. 1: 639–681. https://doi.org/10.1111/jofi.13099.

Farmer, L. E., L. Schmidt, and A. Timmermann. 2023. "Pockets of Predictability." *Journal of Finance* 78, no. 3: 1279–1341. https://doi.org/10.1111/jofi.13229.

Feng, G., S. Giglio, and D. Xiu. 2020. "Taming the Factor Zoo: A Test of New Factors." *Journal of Finance* 75, no. 3: 1327–1370. https://doi.org/10.1111/jofi.12883.

Geczy, C. C., R. F. Stambaugh, and D. Levin. 2021. "Investing in Socially Responsible Mutual Funds." *Review of Asset Pricing Studies* 11, no. 2: 309–351. https://doi.org/10.1093/rapstu/raab004.

Golez, B., and P. Koudijs. 2018. “Four Centuries of Return Predictability.” *Journal of Financial Economics* 127, no. 2: 248–263. https://www.sciencedirect.com/science/article/pii/S0304405X17303185.

Gompers, P., J. Ishii, and A. Metrick. 2003. “Corporate Governance and Equity Prices.” *Quarterly Journal of Economics* 118, no. 1: 107–156. https://doi.org/10.1162/00335530360535162.

Gospodinov, N., and E. Maasoumi. 2021. “Generalized Aggregation of Misspecified Models: With an Application to Asset Pricing.” *Journal of Econometrics* 222, no. 1, Part B: 451–467. https://www.sciencedirect.com/science/article/pii/S0304407620302001.

Guidolin, M., S. Hyde, D. McMillan, and S. Ono. 2009. “Non-Linear Predictability in Stock and Bond Returns: When and Where Is It Exploitable?” *International Journal of Forecasting* 25, no. 2: 373–399. https://www.sciencedirect.com/science/article/pii/S0169207009000065.

Haase, F., and M. Neuenkirch. 2023. “Predictability of Bull and Bear Markets: A New Look at Forecasting Stock Market Regimes (And Returns) in the Us.” *International Journal of Forecasting* 39, no. 2: 587–605. https://www.sciencedirect.com/science/article/pii/S0169207022000048.

He, M., and Y. Zhang. 2022. “Climate Policy Uncertainty and the Stock Return Predictability of the Oil Industry.” *Journal of International Financial Markets, Institutions and Money* 81: 101675. https://www.sciencedirect.com/science/article/pii/S1042443122001470.

Hong, H., and M. Kacperczyk. 2009. “The Price of Sin: The Effects of Social Norms on Markets.” *Journal of Financial Economics* 93, no. 1: 15–36. https://www.sciencedirect.com/science/article/pii/S0304405X09000634.

Ilhan, E., Z. Sautner, and G. Vilkov. 2021. “Carbon Tail Risk.” *Review of Financial Studies* 34, no. 3: 1540–1571. https://doi.org/10.1093/rfs/hhaa071.

Inard, L. 2023. “Cash Flow Valuation and ESG.” In Valuation and Sustainability: A Guide to Include Environmental, Social, and Governance Data in Business Valuation, edited by D. Glavas, 99–128. Springer International Publishing, Cham.

Jordan, S. J., A. J. Vivian, and M. E. Wohar. 2014. “Forecasting Returns: New European Evidence.” *Journal of Empirical Finance* 26: 76–95. https://www.sciencedirect.com/science/article/pii/S0927539814000085.

Kelly, B., S. Malamud, and K. Zhou. 2024. “The Virtue of Complexity in Return Prediction.” *Journal of Finance* 79, no. 1: 459–503. https://doi.org/10.1111/jofi.13298.

Kelly, B., L. Pástor, and P. Veronesi. 2016. “The Price of Political Uncertainty: Theory and Evidence From the Option Market.” *Journal of Finance* 71, no. 5: 2417–2480. https://doi.org/10.1111/jofi.2016.71.issue-5.

Khan, M. 2019. “Corporate Governance, Esg, and Stock Returns Around the World.” *Financial Analysts Journal* 75, no. 4: 103–123. https://doi.org/10.1080/0015198X.2019.1654299.

Krüger, P. 2015. “Corporate Goodness and Shareholder Wealth.” *Journal of Financial Economics* 115, no. 2: 304–329. https://www.sciencedirect.com/science/article/pii/S0304405X14001925.

Li, J., and D. A. Wu. 2020. “Do Corporate Social Responsibility Engagements Lead to Real Environmental, Social, and Governance Impact?” *Management Science* 66, no. 6: 2564–2588. https://doi.org/10.1287/mnsc.2019.3324.

Ma, F., X. Lu, J. Liu, and D. Huang. 2022. “Macroeconomic Attention and Stock Market Return Predictability.” *Journal of International Financial Markets, Institutions and Money* 79: 101603. https://www.sciencedirect.com/science/article/pii/S104244312200083X.

Maiti, M. 2021. “ Is ESG the succeeding risk factor?” Journal of Sustainable Finance & Investment 11, no. 3: 199–213. Publisher: Taylor & Francis. https://doi.org/10.1080/20430795.2020.1723380.

Morningstar. 2024. “Global Sustainable Fund Flows: Q4 2023 in Review, Quarterly Report, Morningstar.” https://www.morningstar.com/lp/global-esg-flows.

Neely, C. J., D. E. Rapach, J. Tu, and G. Zhou. 2014. “Forecasting the Equity Risk Premium: The Role of Technical Indicators.” *Management Science* 60, no. 7: 1772–1791. https://www.jstor.org/stable/42919633.

Newey, W. K., and K. D. West. 1987. “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix.” *Econometrica* 55, no. 3: 703–708. https://www.jstor.org/stable/1913610.

Oestreich, A. M., and I. Tsiakas. 2015. “Carbon Emissions and Stock Returns: Evidence From the Eu Emissions Trading Scheme.” *Journal of Banking & Finance* 58: 294–308. https://www.sciencedirect.com/science/article/pii/S0378426615001326.

Pástor, L., and R. F. Stambaugh. 2000. “Comparing Asset Pricing Models: an investment Perspective.” *Journal of Financial Economics* 56, no. 3: 335–381. https://doi.org/10.1016/S0304-405X(00)00044-1.

Pástor, L., R. F. Stambaugh, and L. A. Taylor. 2021. “Sustainable Investing in Equilibrium.” *Journal of financial economics* 142, no. 2: 550–571.

Pedersen, L. H., S. Fitzgibbons, and L. Pomorski. 2021. “Responsible Investing: The Esg-Efficient Frontier.” *Journal of Financial Economics* 142, no. 2: 572–597. https://www.sciencedirect.com/science/article/pii/S0304405X20302853.

Pettenuzzo, D., A. Timmermann, and R. Valkanov. 2014. “Forecasting Stock Returns Under Economic Constraints.” *Journal of Financial Economics* 114, no. 3: 517–553. https://www.sciencedirect.com/science/article/pii/S0304405X14001627.

Qin, W., S. Cho, and S. Hyde. 2023. “Time-Varying Bond Market Integration and the Impact of Financial Crises.” *International Review of Financial Analysis* 90: 102909. https://www.sciencedirect.com/science/article/pii/S1057521923004258.

Raghunandan, A., and S. Rajgopal. 2022. “Do Esg Funds Make Stakeholder-Friendly Investments?” *Review of Accounting Studies* 27, no. 3: 822–863. https://doi.org/10.1007/s11142-022-09693-1.

Rapach, D., J. Strauss, and G. Zhou. 2010. “Out-Of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy.” *Review of Financial Studies* 23, no. 2: 821–862. https://www.jstor.org/stable/40468327.

Rapach, D., J. Strauss, and G. Zhou. 2013. “International Stock Return Predictability: What Is the Role of the United States?” *Journal of Finance* 68, no. 4: 1633–1662. https://doi.org/10.1111/jofi.12041.

Rapach, D., M. Wohar, and J. Rangvid. 2005. “Macro Variables and International Stock Return Predictability.” *International Journal of Forecasting* 21, no. 1: 137–166. https://www.sciencedirect.com/science/article/pii/S0169207004000512.

Rapach, D., and M. Wohar. 2009. “Multi-Period Portfolio Choice and the Intertemporal Hedging Demands for Stocks and Bonds: International Evidence.” *Journal of International Money and Finance* 28, no. 3: 427–453. https://www.sciencedirect.com/science/article/pii/S0261560608001654.

Rapach, D., and G. Zhou. 2013. “Chapter 6 - Forecasting Stock Returns.” In Handbook of Economic Forecasting, edited by G. Elliott and A. Timmermann, Vol. 2 *of Handbook of Economic Forecasting*, 328–383. Elsevier. https://www.sciencedirect.com/science/article/pii/B9780444536839000062.

Rapach, D., and G. Zhou. 2022a. “Asset Pricing: Time-Series Predictability.” https://papers.ssrn.com/abstract=3941499.

Rapach, D., and G. Zhou. 2022b. “Asset Pricing: Time-series predictability.”.

Schrimpf, A. 2010. “International Stock Return Predictability Under Model Uncertainty.” *Journal of International Money and Finance* 29, no. 7: 1256–1282. https://www.sciencedirect.com/science/article/pii/S0261560610000392.

Serafeim, G., and A. Yoon. 2023. "Stock Price Reactions to Esg News: the Role of Esg Ratings and Disagreement." *Review of Accounting Studies* 28, no. 3: 1500–1530. https://doi.org/10.1007/s11142-022-09675-3.

Shynkevich, A. 2012. "Global Industry Portfolios and Short-Term Predictability of Returns: Is It There?" *Journal of Financial Markets* 15, no. 4: 438–466. https://www.sciencedirect.com/science/article/pii/S1386418112000225.

Stock, J., and M. Watson. 2004. "Combination Forecasts of Output Growth in a Seven-Country Data Set." *Journal of Forecasting* 23, no. 6: 405–430. https://doi.org/10.1002/for.928.

Timmermann, A. 2008. "Elusive Return Predictability." *International Journal of Forecasting* 24, no. 1: 1–18. https://www.sciencedirect.com/science/article/pii/S0169207007000969.

Wang, X., R. J. Hyndman, F. Li, and Y. Kang. 2023. "Forecast Combinations: An over 50-year Review." *International Journal of Forecasting* 39, no. 4: 1518–1547.

Wang, Y., Z. Pan, L. Liu, and C. Wu. 2019. "Oil Price Increases and the Predictability of Equity Premium." *Journal of Banking & Finance* 102: 43–58. https://www.sciencedirect.com/science/article/pii/S0378426619300640.

Welch, I., and A. Goyal. 2008. "A Comprehensive Look At the Empirical Performance of Equity Premium Prediction." *Review of Financial Studies* 21, no. 4: 1455–1508. https://www.jstor.org/stable/40056859.

Wong, J. B., and Q. Zhang. 2022. "Stock Market Reactions to Adverse Esg Disclosure via Media Channels. The British." *Accounting Review* 54, no. 1: 101045. https://www.sciencedirect.com/science/article/pii/S0890838921000718.

Xie, J., and T. Hong. 2016. "GEFCom2014 Probabilistic Electric Load Forecasting: An Integrated Solution With Forecast Combination and Residual Simulation." *International Journal of Forecasting* 32, no. 3: 1012–1016. https://www.sciencedirect.com/science/article/pii/S0169207015001405.

Yang, Y. 2004. "Combining Forecasting Procedures: Some Theoretical Results." *Econometric Theory* 20, no. 1: 176–222. https://www.cambridge.org/core/journals/econometric-theory/article/combining-forecasting-procedures-some-theoretical-results/42A92E2137102421030994FEC382040C.

Yu, D., D. Huang, and L. Chen. 2023. "Stock Return Predictability and Cyclical Movements in Valuation Ratios." *Journal of Empirical Finance* 72: 36–53. https://www.sciencedirect.com/science/article/pii/S0927539823000221.

Yuan, M., and G. Zhou. 2023. "Why Naive 1/N Diversification is Not so Naive, and How to Beat It?" https://papers.ssrn.com/abstract=4281138.

Zhang, Q., R. Ding, D. Chen, and X. Zhang. 2023. "The Effects of Mandatory Esg Disclosure on Price Discovery Efficiency Around the World." *International Review of Financial Analysis* 89: 102811. https://www.sciencedirect.com/science/article/pii/S1057521923003277.

Zhang, Y., F. Ma, C. Liang, and Y. Zhang. 2021. "Good Variance, Bad Variance, and Stock Return Predictability." *International Journal of Finance & Economics* 26, no. 3: 4410–4423. https://doi.org/10.1002/ijfe.2022.

Zhang, Y., Y. Wei, F. Ma, and Y. Yi. 2019. "Economic Constraints and Stock Return Predictability: A New Approach." *International Review of Financial Analysis* 63: 1–9. https://www.sciencedirect.com/science/article/pii/S1057521918307658.

Zhu, X., and J. Zhu. 2013. "Predicting Stock Returns: A Regime-Switching Combination Approach and Economic Links." *Journal of Banking & Finance* 37, no. 11: 4120–4133. https://www.sciencedirect.com/science/article/pii/S0378426613002860.

Appendix A

Variable Description

Table A1 shows the descriptions and data sources for the control variables used in this article.

TABLE A1 | Variable description.

Name	Abbr.	Brief Description	Source	Cited by
US Market Excess Return	SP 500	Percent change of S&P 500 return index.	Amit Goyal's website	Rapach et al. (2013)
Dividend Price Ratio	DP	The difference between the log of dividends and the log of market prices.	Bloomberg	Charles et al. (2017); Jordan et al. (2014)
Dividend Yield	DY	The difference between the log of dividends and the log of lagged prices.	Bloomberg	Ang and Bekaert (2007); Charles et al. (2017); Jordan et al. (2014)
Earnings Price Ratio	EP	The difference between the log of earnings and the log of market prices.	Bloomberg	Charles et al. (2017); Jordan et al. (2014)
Dividend Payout Ratio	DE	The difference between the log of dividends and the log of earnings.	Bloomberg	Charles et al. (2017); Jordan et al. (2014)
Book-to-Market Ratio	BM	The ratio of book value to market value.	Bloomberg	Baba Yara et al. (2021)
Short-term Return	RF	EURIBOR 3-month interest rates.	Eurostat	Rapach and Wohar (2009)
Long-term Return	LTY	EURIBOR 10-year interest rates.	Eurostat	Rapach and Wohar (2009)
Yield Spread	YS	The difference between short-term return and long-term return.	Author's calculation	Schrimpf (2010)
Stock Variance	SVAR	Sum of squared monthly returns on the Morningstar Eurozone index.	Bloomberg	Jordan et al. (2014)
Inflation	INF	Harmonised Index of Consumer Prices in Euro area (changing composition). We follow Welch and Goyal (2008) to lag one-month, as the inflation information in current month only releases data in last month.	Eurostat	Rapach et al. (2005); Schrimpf (2010)

Appendix B

Out-of-sample Results of Alternative Risk Aversions and Market Portfolios

We consider various levels of risk aversion in our study, recognising that the risk tolerance of investors varies in financial markets. Lower risk aversion suggests a higher tolerance for risk, with investors being more willing to accept greater uncertainty in pursuit of higher returns. Conversely, higher risk aversion corresponds to a lower tolerance for risk, leading investors to adopt more conservative investment strategies. In this section, we consider higher risk aversions to examine whether conservative investors can still generate high CER gain by applying the incident-based ESG risk indicator or other well-established predictors.

In Table B1, we present results for investors with risk aversion values of 4 (Zhang et al. 2021) and 5 (Cederburg et al. 2023; He and Zhang 2022; Neely et al. 2014), respectively. These values are higher than widely used risk aversion value of 3 (e.g., Dong et al. 2022; Haase and Neuenkirch 2023; Ma et al. 2022; Yu et al. 2023). The incident-based ESG risk indicator remains consistently significant with the R_{os}^2 value of 1.35% and annual economic gains of the 3.61% and 2.77% without transaction costs for risk aversion levels of 4 and 5, respectively. The sizable gains exceed the critical threshold of 2% annual gains suggested by Pástor and Stambaugh (2000) and Dong et al. (2022). Inflation is another predictor that is economically significant at a risk aversion level of 4. However, when the risk aversion level increases to 5, the CER gains drop to 1.33%, falling below the critical threshold.

We also consider transaction costs across different risk aversion values, and the results remain similar in Table B2. These robustness checks confirm our findings that the incident-based ESG risk indicator is a valuable tool for predicting market returns. Table B2 reports the out-of-sample results of MSCI EMU Total Return Index and Europe Portfolios Return Index in the risk aversions of 4 and 5.

Table B3 reports the out-of-sample results of the Morningstar Eurozone Net Return Index by identifying the market excess return as the difference between the index and Germany 3-month interest rate. The results are similar with the main results which set EURIBOR 3-month interest rate as the risk-free rate.

TABLE B1 | Out-of-sample results of alternative risk aversions.

	No transaction cost		**10 bps transaction cost**	
	CER gain (%)	**Sharpe ratio**	**CER gain (%)**	**Sharpe ratio**
Panel A: $\gamma = 4$				
$IESGRI^A$	3.61	0.47	3.14	0.42
SP500	−5.42	−0.13	−6.63	−0.22
DP	−0.89	−0.08	−0.93	−0.13
DY	−1.17	−0.05	−1.33	−0.08
EP	−1.72	0.34	−1.86	0.33
DE	−1.17	0.41	−1.24	0.41
BM	1.35	0.19	1.23	0.17
RF	−2.44	0.35	−2.47	0.34
LTY	−2.83	0.33	−2.90	0.33
YS	0.26	0.28	0.13	0.27
SVAR	−0.76	0.03	−1.11	−0.01
INF	4.19	0.60	3.96	0.59
Panel B: $\gamma = 5$				
$IESGRI^A$	2.77	0.45	2.37	0.40
SP500	−4.83	−0.10	−5.88	−0.18
DP	−0.72	−0.08	−0.74	−0.13
DY	−1.27	−0.07	−1.40	−0.10
EP	−2.68	0.33	−2.83	0.32
DE	−4.26	0.36	−4.35	0.35
BM	0.99	0.18	0.89	0.15
RF	−3.99	0.31	−4.06	0.30
LTY	−5.65	0.27	−5.74	0.26
YS	−0.17	0.26	−0.27	0.25
SVAR	−0.82	0.01	−1.11	−0.03
INF	1.33	0.57	1.12	0.56

Note: This table reports the economic performance (CER gain (%) and Sharpe ratio) of the incident-based ESG risk indicator and other well-known predictors on return predictability for the Morningstar Eurozone Index Return. For economic gains, we assume risk aversion of 4 and 5, and the portfolio weight in equities lies between −1.5 to 2. The CER gain represents the annualised certainty equivalent return differential between the incident-based ESG risk indicator and the prevailing mean forecast. The Sharpe ratio is the annualised average portfolio excess return over its standard deviation. Both scenarios with no transaction costs and a 10 bps transaction cost are applied. The whole sample period spans from January 2007 to December 2022, with the first 6 years serving as the training sample.
Abbreviations: CER, certainty equivalent return; ESG, Environmental, Social and Governance.

TABLE B2 | Out-of-sample results of alternative risk aversions: MSCI EMU total return index and Europe portfolios return index.

	R^2_{OS} (%)	CW-test	CER gain (%)		Sharpe ratio	
			No cost	10 bps cost	No cost	10 bps cost
Panel A: MSCI EMU Total Return Index						
$\gamma = 4$	1.33**	1.67	3.76	3.30	0.51	0.46
$\gamma = 5$	1.33**	1.67	2.90	2.51	0.50	0.45
Panel B: Europe Portfolios Return Index						
$\gamma = 4$	1.35**	1.69	3.87	3.33	0.56	0.51
$\gamma = 5$	1.35**	1.69	3.16	2.69	0.56	0.51

Note: This table reports the out-of-sample statistical and economic performance (CER gain (%) and Sharpe ratio) of the incident-based ESG risk indicator on return predictability for the MSCI EMU Total Return Index and the Europe Portfolios Return Index. Statistical significance for R^2_{OS} is based on the Clark and West (2007) test of the MSFE-adjusted statistic for the hypothesis: $H_0 : R^2_{OS} \leq 0$ versus $H_A : R^2_{OS} > 0$. For economic gains, we assume risk aversion levels of 4 and 5, and the portfolio weight in equities lies between −1 to 2. The CER gain represents the annualised certainty equivalent return differential between the incident-based ESG risk indicator and the prevailing mean forecast. The Sharpe ratio is the annualised average portfolio excess return over its standard deviation. Both scenarios with no transaction costs and a 10 bps transaction cost are applied. The asterisks *, ** and *** denotes out-of-sample statistical significance at the 10%, 5% and 1% levels, respectively. The whole sample period spans from January 2007 to December 2022, with the first 6 years serving as the training sample.
Abbreviations: CER, certainty equivalent return; ESG, Environmental, Social and Governance; MSFE, mean squared forecast error.

TABLE B3 | Out-of-sample results of the morningstar Eurozone total return index - Using Germany 3 month interest as the risk-free rate.

Risk aversion	R^2_{OS} (%)	CW-test	CER gain (%)		Sharpe ratio	
			No cost	10 bps cost	No cost	10 bps cost
$\gamma = 3$	1.29**	1.65	4.57	4.13	0.46	0.34
$\gamma = 4$	1.29**	1.65	3.59	3.24	0.44	0.36
$\gamma = 5$	1.29**	1.65	2.74	2.44	0.42	0.34

Note: This table reports the out-of-sample statistical and economic performance (CER gain (%) and Sharpe ratio) of the incident-based ESG risk indicator on return predictability for the MSCI EMU Total Return Index by using ERIBOR as the risk-free rate. Statistical significance for R^2_{OS} is based on the Clark and West (2007) test of the MSFE-adjusted statistic for the hypothesis: $H_0 : R^2_{OS} \leq 0$ versus $H_A : R^2_{OS} > 0$. For economic gains, we assume risk aversion levels of 3, 4, and 5, and the portfolio weight in equities lies between −1 to 2. The CER gain represents the annualised certainty equivalent return differential between the incident-based ESG risk indicator and the prevailing mean forecast. The Sharpe ratio is the annualised average portfolio excess return over its standard deviation. Both scenarios with no transaction costs and a 10 bps transaction cost are applied. The asterisks *, ** and *** denotes out-of-sample statistical significance at the 10%, 5% and 1% levels, respectively. The whole sample period spans from January 2007 to December 2022, with the first 6 years serving as the training sample.
Abbreviations: CER, certainty equivalent return; ESG, Environmental, Social and Governance; MSFE, mean squared forecast error.