Journal of Asset Management
https://doi.org/10.1057/s41260-025-00404-4

ORIGINAL ARTICLE

Financial returns of going green: evidence from MSCI indices

Jan Heldmann[1] · Thomas Brückner[1] · Huong Dieu Dang[2]

Revised: 5 March 2025 / Accepted: 17 April 2025
© The Author(s) 2025

Abstract

This study examines how the green criteria (GCE) used by MSCI to create green equity indices influence their financial performance. We analyze the Climate Change (CC), Paris-Aligned Benchmark (PAB), Socially Responsible Investment (SRI), and SRI Filtered PAB (SRI PAB) index variants in comparison with their standard non-green counterpart in each of the four regions: the World, the USA, Europe, and Emerging Markets (EM). Overall, the green indices often matched or exceeded the returns of their standard index without adding significant risk. With few exceptions in the EM, the green indices exhibited better long-term financial performance than their standard index. Over 2015–2023, the CC, PAB, SRI, and SRI PAB indices respectively delivered cumulative excess returns of 4.7%, 5.8%, 13.7%, and 7.5% relative to the standard index. Their returns co-moved closely with the market and the standard index's returns. The GCEs statistically and significantly contributed to green indices' relative financial outperformance.

Keywords MSCI · Green index · Cumulative return differential · Wealth relative

Introduction

The growing need for investment solutions addressing climate, social, and governance concerns has drawn great investors' interest in green products.[1] At the end of 2023, global sustainable funds managed about US$2.2 trillion and continued to outgrow conventional funds in terms of inflows (Cortina and Phyu 2024). Many of these products use green indices replicated through Exchange Traded Funds (ETFs); however, it remains uncertain whether such green indices/ETFs consistently outperform their conventional non-green counterparts. Some studies report positive relative financial performance (Fiordelisi et al. 2023; Rompotis 2023; Pavlova and de Boyrie 2022), greater resilience during financial crises (Ortas et al. 2014) and superior performance during the COVID-19 pandemic (Lin and Swain 2024). Conversely, other research finds green indices/ETFs perform comparably to traditional benchmarks, vary across regions, or show no significant difference in risk-adjusted returns (Bolognesi et al. 2024; Jonwall et al. 2024; Dumitrescu et al. 2023; Cunha et al. 2020; Jain et al. 2019; Benson et al. 2010). Some studies even suggest underperformance of green indices/ETFs relative to traditional financial products (Lean and Nguyen 2014; Ortas et al. 2012). Evidence on return correlation is also mixed, with some authors observing decoupling (Ang 2015; Lean and Nguyen 2014) and others documenting high correlations or co-movements (Rompotis 2023; Managi et al. 2012). Thus, it remains unclear whether green indices/ETFs exhibit better financial performance than their non-green counterparts. Furthermore, very little effort has been made to *directly* examine if and how criteria employed to construct green indices relate to their financial performance. If green criteria do not affect a green index's performance, any observed outperformance (relative to its standard non-green index) is likely due to luck. The existing evidence,

✉ Huong Dieu Dang
huong.dang@canterbury.ac.nz

1 University of Bayreuth, Bayreuth, Germany

2 University of Canterbury, Christchurch, New Zealand

[1] In this study, we address various sustainable investments, for example, Environmental, Social & Governance (ESG), Social Responsible Investment (SRI), sustainable, ethical, green as green investments.

Published online: 07 June 2025

mostly for ETFs, implying a mixed relationship (Abate et al. 2021; Papathanasiou and Koutsokostas 2024).

Our study is motivated by the mixed reported results on the financial performance of green indices/ETFs and the unexplored (mixed) relationship between green criteria and the financial performance of green indices (ETFs). In this study, we choose to explore Morgan Stanley Capital International (MSCI) indices as they have been widely used as reference indices for green equity ETFs.[2] We address two main research questions: (1) How did MSCI green indices perform financially compared with their respective standard (parent) index? and (2) How did the green criteria employed by MSCI to create each green index affect its financial performance?

We focus on indices rather than ETFs for several reasons. *First,* indices are more suitable for performance analysis, as the impact of total expense ratio, taxes,[3] distribution policy,[4] and replication strategy[5] varies across providers and can obscure return analyses. *Second,* the exact benchmark an ETF tracks can change over time, while the ETF keeps the same International Securities Identification Numbers (ISIN), which can distort benchmark-adjusted return analyses.[6] *Third,* ETF providers use slightly modified benchmarks regarding exclusion thresholds, compared with our benchmark being the standard non-green index. That is, providers aim to achieve a performance advantage in their respective peer group by offering ETFs that do not track exactly the competitor's benchmark but the same green exposure.[7]

We examine a set of 20 MSCI indices spanning four major regions: the World, the USA, Europe, and emerging markets (EM). Within each region, we investigate the performance of each of the four green indices—the Climate Change (CC), the Paris-Aligned Benchmark (PAB), the Socially Responsible Investment (SRI), and the SRI filtered PAB (SRI PAB) index—relative to the standard (parent) MSCI index. Our primary dataset on index returns extends from January 2015 to December 2023, providing 2160 monthly index observations. Considering the availability of firm-level data and Tobin's Q measure, our sample of 1,076,340 firm-month observations is limited to December 2022, and includes 5267 unique constituents of 50 countries and 11 sectors. The rich historical data of MSCI index constituents enables us to carry out more thorough and reliable analyses than using comparable ETFs.[8]

Our study differs from recent research on ETFs' financial performance, such as ElBannan (2024). *First*, we focus on indices instead of ETFs. By doing so, we avoid the added complexity of taxes, distribution policies, and replication methods that can influence ETF returns. *Second*, our index constituent data allows us to control for differences in sectors and levels of development across countries. *Third*, we analyze four distinct types of green indices rather than treating all ESG or SRI products as one category. This includes giving special attention to the Paris-Aligned Benchmark (PAB) and SRI PAB indices which both have not been studied in the literature.

Our study is related to Kossentini et al. (2024) and Jacob and Wilkens (2021). The former focus on four MSCI ESG Leader indices covering different parts of the European region: Europe, the Economic and Monetary Union (EMU), Emerging Markets Europe, and the Middle East and Europe, for the years from 2007 to 2020. The latter analyze four ESG and four carbon-focused MSCI World Indices from 2011 to 2019. Out of our 16 green indices, only the World SRI overlaps with Jacob and Wilken's (2021) sample. We

[2] As of October 2024, assets under management in ETFs linked to MSCI equity indexes reached approximately $1.72 trillion. See: https://ir.msci.com/aum-etfs-linked-msci-indexes.

[3] Withholding tax benefits on US dividends, depending on the fund domicile of an ETF, can have a substantial impact on fund performance. Physically replicating ETFs with US exposure domiciled in Ireland have a reduced withholding tax burden of 15% compared to, for example, 30% for ETFs domiciled in Luxembourg.

[4] For example, accumulating ETFs reinvest their dividends while distributing ETFs pay out dividends.

[5] There are two main replication strategies, physical and synthetic replication. They split into full or physically optimized replication and swap replication (which can be unfunded, funded and fully funded). The replication strategy directly influences the performance of an ETF. Physical replication benefits from securities lending while those using swaps are, depending on the qualification of the index under Section 871(m) of the US Internal Revenue Code, not burdened by tax regulations. Generally, ESG ETFs use physical replication.

[6] For example, the Amundi Index MSCI Emerging Markets (EM) SRI PAB UCITS ETF (ISIN: LU1861138961) was launched in January 2019 and tracked the MSCI EM SRI Index. In October 2019, the reference index was switched to MSCI EM SRI 5% Issuer Capped. From December 2020 onward, the ETF tracked the MSCI EM SRI Filtered PAB. Another example is the iShares MSCI World SRI UCITS ETF (ISIN: IE00BYX2JD69). It was launched in October 2017 with MSCI World SRI Select Index being the benchmark. In November 2019, it started tracking the MSCI World SRI Select Reduced Fossil Fuel Index.

[7] Please see the green product range (SFDR Article 8 and 9) from Amundi, www.amundietf.nl/en/professional/etf-products/search, iShares, www.ishares.com/us/products/etf-investments and Xtrackers, www.etf.dws.com/en-us/etf-products.

[8] MSCI provides *current* data on its indices in the factsheets and via the constituent download function at https://www.msci.com/constituents. Historical data is not freely available, but our data set extends from the beginning of 2015 to the end of 2023. The index performance of 12 indices launched after 2015 has been reconstructed (by MSCI) using the identical methods and principles of MSCI. This historical data availability makes our study unique, as we are not aware of any other study that analyses MSCI PAB and SRI PAB indices over such a long period of time, including the Corona crisis and the Ukraine war.

extend these two studies by covering a more recent period that includes significant global events such as the onset of the war in Ukraine. Our study encompasses two additional major regions, the USA, and Emerging Markets. Importantly, we place particular emphasis on indices aligned with the Paris Climate Agreement, including the PAB and SRI PAB indices, which are of particular interest to climate-conscious investors. A recent study conducted by Bolognesi et al. (2024) investigates the MSCI SRI and MSCI PAB indices, but its sample is limited to the USA for a short period, August 2021–May 2024. We employ a richer dataset that covers four regions over a longer period (2015–2023). Most importantly, we carry out index-level regression analyses that control for sector- and country-specific parameters. This allows us to directly evaluate the effectiveness of the green criteria (GCE) employed by MSCI to create each green index type.

To assess the financial performance of a green index, we use cumulative return differential (CRD) as our primary measure.[9] For robustness tests, we utilize Wealth Relative (WR), Tobin's Q,[10] the Sharpe ratio, and the Treynor ratio. CRD and WR capture the performance of a green index, relative to its standard index, over an investment horizon; each is computed directly at the index level. Other metrics are not relative measures; each is estimated using an index's monthly constituent weights and constituents' metrics. To examine an index's return co-movement with the market index, we estimate the market factor beta coefficient using Fama-French models. We also assess the dynamic conditional correlations between green and standard index returns over time using a GARCH(1,1) model with a Gaussian distribution on daily return data. To explore the relationship between green criteria and a green index's financial performance, we conduct OLS regressions and employ four dummy variables, each represents the impact of MSCI's green criteria (GCE) used to create one green index type.

Our analysis documents strong evidence supporting the attractiveness of green indices. On average, over the period January 2015–December 2023, an investor would have respectively earned a 4.7%, 5.8%, 13.7%, and 7.5% greater cumulative return if s/he had invested in the green CC, PAB, SRI, and SRI PAB index instead of the standard index. Taking into account the expense ratios of the largest ETFs which replicate the green and standard indices, an investor would have respectively earned a *net* cumulative return differential of 4%, 5.1%, 13%, and 6.8% during 2015–2023. Across the four regions, most green indices demonstrated long-term outperformance (except in EM, where only the SRI outperformed the standard index). On average, green indices across regions achieved greater Sharpe and Treynor ratios and delivered better risk-adjusted returns than their respective standard index. Furthermore, green indices' returns moved closely with those of the market index and their respective standard index. We find strong evidence that MSCI's GCEs statistically and significantly contributed to green indices' better financial performance, relative to the standard index. The positive effects of GCEs were more pronounced on the outperformance of the SRI and SRI PAB indices; both are aligned with the Paris Climate Agreement.

Our main results are consistent across robustness tests, including analyses with one-, two-, and three-month lagged control variables, alternative return measures, and various subsamples. Our key finding indicates that green indices can deliver competitive returns and, in most cases, outperform their standard non-green counterparts over a relatively long time period. By emphasizing the role of green criteria in index construction, our study contributes to ongoing discussions on the financial attractiveness of green equity investments.

The rest of our paper is structured as follows. The "Brief overview of MSCI's methods for creating green equity indices" section provides a brief overview of the process by which MSCI selects constituents and constructs the examined green indices. The "Literature review and research questions" section reviews the literature and proposes research questions. The "Data and methods" section describes data and presents the method. The "Empirical results" section discusses the empirical results. The "Conclusion" section concludes the key findings.

Brief overview of MSCI's methods for creating green equity indices

The MSCI Standard (parent) Index serves as the starting universe for determining the eligible universe of a green index.[11] The parent indices in our sample were launched in March 1986, except for EM in January 2001. The green indices are constructed by applying exclusion criteria to

[9] As we compare MSCI green indices to their corresponding MSCI standard index, we do not need to consider index provider' characteristics that affect an index's financial performance.

[10] We avoid using accounting figures from individual companies to compute index-level return on equity (ROE) and return on assets (ROA) because several factors can distort these accounting return measures. Differences in accounting standards and legal frameworks across countries can lead to inconsistent financial reporting. In addition, managers often have significant flexibility, which may allow them to adjust or smooth earnings over time. Furthermore, during periods of high inflation, depreciation expenses may be understated since they do not reflect the true replacement costs of equipment, potentially leading to artificially inflated earnings.

[11] See https://www.msci.com/index-methodology for an overview of the construction methodology of all MSCI Indices.

the parent index, leading to a smaller eligible universe. The degree of exclusion varies across the green indices, with stricter criteria resulting in fewer components compared to the parent index. Among the indices analyzed, the MSCI Climate Change (CC) Index applies the mildest exclusions, while the MSCI Socially Responsible Investing (SRI) filtered PAB Index enforces the strictest exclusions. Both the PAB and SRI PAB indices are consistent with the Paris Agreement's objectives, with the SRI PAB excluding the most securities.

The MSCI CC index, and the MSCI climate PAB index

To form the MSCI CC Index (launched in June 2019) and the MSCI Climate PAB Index (launched in October 2020), a negative screening is applied to the parent index in the first step. Companies are excluded from the parent index if they violate either the UN Global Compact Principles or sector policies, such as generating revenue above set thresholds from controversial weapons, tobacco, or fossil fuels. Next, the sustainability of the remaining companies is evaluated using the Low Carbon Transition (LCT) score, which considers three components: a company's carbon footprint, its climate-related risks, and its ability to manage those risks. Companies are categorized into five groups based on their LCT scores: Solutions, Neutral, Operational Transition, Product Transition, and Asset Stranding. Companies classified as Solutions (with the highest LCT score of three) receive a higher weighting in the final index. Weight adjustments are applied across all five categories relative to the security weights in the parent index, forming the final eligible universe for the CC Index. For instance, in the "World" region, this process transforms the MSCI World parent index into the MSCI World CC Index.

The MSCI Climate PAB Index, however, applies stricter criteria with a PAB overlay aligned with the goals of the Paris Agreement. Specifically, companies must reduce their Weighted Average Carbon Intensity (WACI) by 50% and lower their annual carbon footprint by 7% compared to the parent index. Additionally, the Climate PAB Index excludes companies deriving revenue from fossil fuel activities, such as coal, oil, natural gas exploration, processing, or power generation with excessive greenhouse gas intensity. Finally, the weighting of each company in the MSCI Climate PAB Index is determined using its LCT score.

The MSCI SRI index, and the MSCI SRI filtered PAB (SRI PAB) index

The MSCI SRI Index (launched in June 2011, except for EM in March 2014) and the MSCI SRI filtered PAB Index (launched in June 2020) apply stricter exclusion criteria than the Climate Indexes discussed above. Both indices implement more sector-specific exclusions with low thresholds, resulting in a significantly smaller eligible universe compared to the parent index. A detailed comparison of the exclusion criteria and thresholds can be found in "Appendix A."

The MSCI SRI filtered PAB Index is more exclusionary than the MSCI SRI Index due to its alignment with the Paris Agreement goals. It applies additional thresholds to companies generating revenue from oil and gas activities. Furthermore, companies with serious violations of sustainable investment objectives are flagged as "Red," receiving an ESG Controversies Score of 0, and are excluded from the index. The remaining companies are rated using MSCI ESG Ratings, which range from AAA to CCC. For inclusion, existing index constituents must have a minimum rating of BB and an ESG Controversies Score of at least 1, while new additions require a minimum rating of A and an ESG Controversies Score of 4.

Both indices employ a best-in-class approach, selecting only the top 25% of companies within each sector. The SRI filtered PAB Index also incorporates the PAB overlay, requiring companies to reduce their WACI by 50% and their carbon footprint by 7% annually relative to the parent index. To ensure diversification, the weight of any single company in the index is capped at 5%.

Literature review and research questions

From a portfolio theory perspective, both negative screening and best-in-class approaches used to construct green portfolios limit the investment universe, potentially reducing diversification and increasing risk compared to broader portfolios (Barnett and Salomon 2006; Renneboog et al. 2008). Nevertheless, several studies indicate that green stocks or those with high ESG scores exhibit favorable characteristics, such as lower unsystematic risk (Giese et al. 2019; Hong and Kacperczyk 2009) and reduced capital costs (Gregory et al. 2021; Unruh et al. 2016; Ng and Rezaee 2015). Most research on green indices or ETFs finds minimal differences in returns compared to benchmarks, and any observed differences are typically aligned with risk-adjusted returns.

Studies on SRI and ESG indices report either better or similar performance compared to conventional benchmarks. Those finding better performance include Statman (2005), Cunha et al. (2020), Jain et al., (2023), and Kossentini et al. (2024). In contrast, studies reporting similar performance include Schröder (2007), Collison et al. (2008), Consolandi et al. (2009), Benson et al. (2010), Jain et al. (2019), Jacob and Wilkens (2021), and Bolognesi et al. (2024). Several studies also document that green indices/ETFs demonstrate greater resilience during crises (Ortas et al. 2014, 2013; Omura et al. 2021; Lin and Swain 2024; Huang 2024; ElBannan 2024). Anti-ESG ETFs, however, tend to underperform (Rompotis 2024).

In terms of returns co-movement, Jain et al. (2019) find that green indices closely track their benchmarks across the USA, E.U., EM, and global markets during 2013–2017. Similarly, Managi et al. (2012) observe no distinct characteristics between green indices and their benchmarks, noting high co-movement across various market conditions. These findings *alone* suggest that green indices are not markedly distinct investment vehicles compared with their traditional benchmarks.

Despite the growing popularity of green investments, very few empirical studies *directly* investigate how green indices' construction criteria affect their financial performance. This is likely due to the challenge in obtaining indices' constituent weighting data over a reasonably long period. Related studies mainly examine funds categorized by sustainability/ESG ratings and report mixed findings. Using 634 European mutual funds during 2014–2019, Abate et al. (2021) report better performance for funds with high Morningstar Sustainability Ratings. In contrast, Papathanasiou and Koutsokostas (2024) find that ESG funds with low Morningstar ratings outperform among 235 ESG mutual funds during 2010–2022. Folger-Laronde et al. (2022) observe that higher-rated ETFs with an Eco-Fund Label have lower returns during the COVID-19 market crash. Rompotis (2022a, b) find no significant risk-adjusted return differences for ESG ETFs and a negative relationship between returns and ESG metrics, respectively. Given the above mixed findings, it remains unclear whether sustainability/ESG ratings are linked to ETFs' better financial performance. In reality, the criteria used to construct a green index are complex and extend far beyond ESG ratings, as we briefly discussed in "Brief overview of MSCI's methods for creating green equity indices" section. The unexplored relationship between green criteria and green index's financial performance motivates us to investigate how the green criteria used by MSCI affect its green indices' financial performance.

Our literature review highlights several gaps that have not been well addressed in prior studies. Most studies focus on generic ESG/SRI ETFs/indices without analyzing those aligned with the Paris Agreement's emissions reduction goals. Return comparisons often involve multiple providers against a single benchmark, ignoring provider-specific characteristics. Furthermore, the impact of green criteria used to create an index and its financial performance remains unexplored. Our study addresses these gaps by conducting a comprehensive analysis of green MSCI indices, utilizing constituent weighting data across four major regions, assessing their alignment with the Paris Agreement, and benchmarking them against their respective MSCI standard indices. We aim to address the research gaps by answering the following two research questions:

1. *How did green MSCI indices perform financially compared with its standard non-green index?*
2. *How did an index's Green Criteria Effectiveness (GCE) affect its financial performance?*

Data and methods

Data

Our study makes use of both index-level and firm-level data. Our monthly index dataset includes 20 MSCI equity indices from January 2015 to December 2023. The five examined indices (one standard and four green-oriented variants) are analyzed in each of the four regions: the World, the USA, EU, and EM. We collect daily Net Total Returns (in USD) from Bloomberg, which include dividend reinvestments and consider withholding taxes. To capture an index's exposure to the market index, we use monthly Fama-French factor returns and risk-free rates from French's (2024) online data library.[12]

For the firm-level analysis, MSCI generously provided us with monthly data on each index's constituents and their weights from January 2015 to December 2022. This dataset includes 5267 unique firms from 11 sectors across 50 countries. We are able to meticulously map 5034 of these firms (covering more than 99.9% of the total indices' weights) to their ISINs. Other firm-specific data is obtained from Refinitiv Eikon's Datastream, which covers nearly our entire sample. To our knowledge, this is one of the first studies that employ such a rich MSCI index data, including two Paris Agreement-aligned indices, over a relatively long period.

Methods

Index's financial performance measures

We employ cumulative return differential (CRD) as our primary measure that captures the financial performance of a green index gi, relative to its parent index pi. The return differential $rd_{gi,t}$ of green index gi relative to its parent index pi at time t is computed as:

$$rd_{gi,t} = r_{gi,t} - r_{pi,t} \tag{1}$$

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

The $\text{CRD}_{gi,T}$ of green index gi over the time horizon T is given by:

$$\text{CRD}_{gi,T} = \sum_{t=1}^{T} rd_{gi,t} \tag{2}$$

The average CRD, ACRD_t, for a green index gi across four regions k at time t, is given by:

$$\text{ACRD}_{gi,t} = \frac{1}{4}\sum_{k=1}^{4} \text{CRD}_{gi,t,k} \tag{3}$$

In addition, we calculate the net CRD and net ACRD by deducting one twelfth of the corresponding ETFs' average annual costs[13] from the monthly returns of an index in our sample.

In robustness analyses, we use two other performance measures namely the wealth relative (WR) and Tobin's Q. The WR is the ratio of the compounded returns between a green index gi and its parent index pi at time t, estimated as follows:

$$\text{WR}_{gi,t} = \frac{\prod_{t=1}^{T}(1 + r_{gi,t})}{\prod_{t=1}^{T}(1 + r_{pi,t})} \tag{4}$$

A WR value greater than one indicates that an investor would be better off investing in the green index gi instead of its standard index pi over the time horizon T.

We estimate an index's Tobin's Q as the weighted sum of firm j's weight ω in index i and firm j's Tobin's Q at time t, where J is the total number of firms in the index:[14]

$$\text{Tobin's } Q_{i,t} = \sum_{j=1}^{J} \omega_{i,j,t} * \text{Tobin's } Q_{j,t} \tag{5}$$

In a similar approach, we also estimate the Sharpe ratio and Treynor ratio of each index.

Returns co-movement

We estimate the Fama–French (FF) five-factor model to examine the co-movements of a green index's returns relative to the market index's returns in each region. In this analysis, we focus on the coefficient estimate of the market factor (β_1). If β_1 is close to one, it indicates that the green index i's returns move closely with the market index's returns. Our baseline FF five-factor model is specified as follows:

$$\begin{aligned} r_{i,t} - r_{f,t} = \alpha + \beta_1\left(r_{m,t} - r_{f,t}\right) + \beta_2\text{SMB}_t + \beta_3\text{HML}_t \\ + \beta_4\text{RMW}_t + \beta_5\text{CMA}_t + \epsilon_t \end{aligned} \tag{6}$$

where $r_{i,t}$ represents an index i's return at time t, and $r_{f,t}$ is the risk-free rate at time t.

The independent variables include the market excess return ($r_{m,t} - r_{f,t}$), small minus big (SMB_t), high minus low (HML_t), robust minus weak (RMW_t), and conservative minus aggressive (CMA_t) factors at time t.

In addition to the baseline FF five-factor model (Fama and French 2015), we also estimate the FF three-factor model (Fama and French 1993), the Carhart Fama-French four-factor model (Carhart 1997), and the Carhart Fama-French six-factor model (Fama and French 2018) for robustness tests.

To examine the co-movements between the returns of a green index and its respective standard index (in the same region), we apply the GARCH (1,1) model with a Gaussian distribution to our daily returns. This enables us to understand the correlation dynamics over time between green indices and their parent index in each region.

OLS index-level regressions

We carry out ordinary least squares (OLS) regressions at the index level to analyze how the effectiveness of green criteria, denoted as GCE, affects an index's financial performance. For each performance metric outlined earlier (CRD, WR, Tobin's Q), we estimate the following regression model:

$$\begin{aligned} Y_{i,t} = \alpha + \sum_{k=1}^{4} \beta_i \times \text{GCE\_Index}_i + \sum_{g=1}^{4} \gamma_g \times \text{HDI\_Group}_{g,t-12} \\ + \sum_{s=1}^{10} \delta_s \times \text{Sector}_{s,t-1} + \psi_1 \text{Post\_Launch}_{i,t-1} + \psi_2 \log\ \text{MV}_{i,t-1} \\ + \psi_3 \text{Covid}_{t-1} + \psi_4 \text{Ukraine}_{t-1} + \epsilon_t \end{aligned} \tag{7}$$

where

$Y_{i,t}$: A performance measure (CRD, WR, Tobin's Q) for index i at time t, as discussed above.

GCE_Index_i: A dummy variable that takes the value of 1 if index i was constructed according to the criteria of a certain index. This means that GCE_PAB equals to 1 only for the PAB index, while it is zero for all others, including the standard, CC, SRI, and SRI PAB indices. This logic applies to all other green index types accordingly.

[13] For this purpose, we calculate the average total expense ratios of the largest ETFs which replicate green and standard indices. The average annual total expense ratios we have calculated for the standard (green) index for the World is 13 (21), for the USA 11 (20), for Europe 15 (19) and for EM 17 (25) basis points respectively.

[14] Although Tobin's Qs at the company level are annual values, the Tobin's Qs at the index level are computed monthly as the weights of constituents in an index change monthly and the composition of an index also changes over the course of a year.

$\text{HDI\_Group}_{g,t-12}$: Percentage of firms in a Human Development Indicator (HDI) group g at time $(t-12)$.[15]

$\text{Sector}_{s,t-1}$: Percentage of firms in sector s at time $(t-1)$.[16]

$\text{Post\_Launch}_{i,t-1}$: Dummy variable indicating if index i had already been launched prior to or was launched at time $(t-1)$.

lMV_{t-1} : Logarithm of the average weighted market value of index i at time $(t-1)$.

Covid_{t-1} : One-month lagged dummy variable for the COVID-19 period.

Ukraine_{t-1}: One-month lagged dummy variable for the Ukraine war period.

ϵ_t: Error term at time t.

Unlike ElBannan (2024), which uses a single dummy variable to represent various ESG funds, we analyze four green index types separately. Instead of grouping them in one category, we create four distinct dummy variables (GCE_Index_i). In equation (7), the standard index acts as the baseline category and is captured by the intercept (α). If the coefficient for GCE_Index_i is positive ($\beta_i > 0$) and statistically significant, it suggests that MSCI's criteria to create this specific green index type has a positive impact on its performance measure $Y_{i,t}$, even after controlling for relevant index-, sector-, country- and time-specific variables.

Our dataset includes 50 countries. Using multiple country-specific variables (for example, country dummy variables) leads to multicollinearity and would obscure the effects of our main explanatory variables, green criteria effectiveness (GCE). We avoid multicollinearity by using the Human Development Indicator (HDI) as a proxy for country-specific effects, and sector composition, which captures broad cross-sectional differences.[17]

Our primary financial performance measure is the cumulative return differential (CRD), which reflects the long-term perspective of investors. We use OLS regression as it offers clear and easily interpretable coefficient estimates which allow us to directly assess the influence of the GCEs on an index's relative return. Furthermore, we observe no significant differences in results between cumulative (CRD) and non-cumulative return measure (WR). The usage of cumulative values does not affect the main insights or conclusions of the study, suggesting that the OLS approach is suitable for our analysis.

To further validate our model, we conduct Residual versus Fitted Plots for our baseline regression of CRD. The residuals appear reasonably centered around zero across the range of fitted values, indicating no severe violations of linearity. We also examine added-variable plots, and the partial regression lines show a linear pattern for our GCE dummies. Our model further includes controls for sector composition, country characteristics (HDI), index's market capitalization, and major events (COVID-19, Ukraine war).

Most importantly, we use Newey–West standard errors in all analyses to address potential autocorrelation and heteroskedasticity.[18] This ensures that any omitted variable bias or violations of classical assumptions are minimized. We acknowledge that no model is perfect; however, the absence of strong evidence of non-linearity or severe specification errors supports the validity of our OLS analysis results.

In addition to our baseline OLS analysis of cumulative return differential, we conduct several robustness tests using alternative financial performance metrics (WR, Tobin's Q). We further analyze subsamples, including specific sub-indices, sub-regions, periods of heightened market uncertainty, and samples with two-month/three-month lags. The results remain consistent across these tests, reinforcing the reliability of our main findings.

Empirical results

Statistics

Table 1 presents the statistics of returns for the 20 MSCI indices in four regions over the period January 2015–December 2023 (columns 1–5), and two sub-periods of different market conditions: the pre-pandemic (market stability) period January 2015–January 2020 (columns 6–10), and the pandemic period February 2020–January 2022 (columns 11–15).[19] Across all periods, the green indices generally

[15] Each year, constituents' countries in our sample are grouped into five HDI categories based on their lagged numerical values: ≤0.6, 0.6–0.7, 0.7–0.8, 0.8–0.9, and ≥0.9. The highest category (≥0.9) serves as the reference group. Since firm weights within an index change monthly and index constituents also evolve over time, the HDI group values are updated each month. The HDI data is sourced from the United Nations Development Program (UNDP).

[16] The eleven sectors include Health Care, Consumer Discretionary, Materials, Industrials, Information Technology, Financials, Consumer Staples, Energy, Utilities, Communication Services, and Real Estate which we use as the reference category.

[17] Subsequent empirical analyses show that the green indices have betas close to 1 and they all maintain high correlations (generally above 0.95) with their respective standard indices. That is, the green indices have similar systematic risk as the market index, and macroeconomic factors, such as interest rates or inflation, likely affect both green and standard indices similarly. Thus, we do not employ country-specific macroeconomic controls.

[18] Additionally, we compute White standard errors for all models and obtain consistent results.

[19] We choose February 2020 as the starting point of the pandemic, as the World Health Organization (WHO) declared it a Public Health Emergency of International Concern on January 30 (WHO 2020). The period ends in January 2022, aligning with the onset of the Ukraine War in February 2022. We observe that by March/ April 2022, the strict measures previously enacted in response to the Covid-

Table 1 Descriptive statistics of index returns, January 2015–December 2023

Panel A: World	Total: 01/15–12/23					Pre-Covid-19: 01/15–01/20					Covid-19: 02/20–01/22				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
Standard	0.82*	1.4**	4.49	−13.2	12.8	0.73*	1.15**	3.35	−7.6	7.9	1.41	2.53	5.80	−13.2	12.8
Climate Change	0.91**	1.34**	4.62	−12.3	12.8	0.81*	1.09**	3.36	−7.6	7.6	1.53	2.49	5.86	−12.3	12.8
PAB	0.88**	1.54**	4.55	−12.7	12.7	0.85**	1.15**	3.30	−7.4	7.7	1.4	2.47	5.87	−12.7	12.7
SRI	0.91**	1.3**	4.52	−10.8	11.9	0.81*	1.14**	3.31	−7.6	7.7	1.59	2.78	5.70	−10.8	11.9
SRI PAB	0.9**	1.12**	4.54	−11.1	11.9	0.83*	1.02**	3.27	−7.7	7.7	1.57	2.6	5.66	−11.1	11.9
Panel B: USA	Total: 01/15–12/23					Pre-Covid-19: 01/15–01/20					Covid-19: 02/20–01/22				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
Standard	1**	1.4***	4.63	−12.7	13.1	0.92**	1.27***	3.45	−9.1	8.2	1.69	2.66	5.96	−12.7	13.1
Climate Change	1.11**	1.78***	4.79	−11.7	13.3	1.01**	1.54***	3.48	−8.9	8.1	1.85	2.85	6.06	−11.7	13.3
PAB	1.09**	1.59***	4.73	−11.9	13.2	1.1**	1.58***	3.42	−8.6	8.1	1.67	2.83	6.04	−11.9	13.2
SRI	1.13**	1.4***	4.73	−10.3	12.7	0.99**	1.36***	3.45	−8.6	7.8	2.03	3.38*	5.97	−10.3	12.7
SRI PAB	1.06**	1.21***	4.74	−11.4	12.3	0.93**	1.13***	3.45	−8.7	7.8	1.96	2.25	5.96	−11.4	12.3
Panel C: Europe	Total: 01/15–12/23					Pre-Covid-19: 01/15–01/20					Covid-19: 02/20–01/22				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
Standard	0.55	0.7	4.85	−14.6	17.0	0.43	0.59	3.67	−7.8	6.7	0.96	2.77	6.32	−14.6	17.0
Climate Change	0.55	0.44	4.92	−14.5	16.4	0.44	0.23	3.69	−8.2	6.2	0.94	2.99	6.35	−14.5	16.4
PAB	0.62	0.54	4.92	−14.6	16.1	0.52	0.38	3.64	−8.0	6.2	1.04	3.11	6.33	−14.6	16.1
SRI	0.68	0.75	4.75	−10.9	15.9	0.61	0.68	3.58	−7.8	7.2	1.04	3.27	6.09	−10.9	15.9
SRI PAB	0.71	1.02	4.81	−11.7	15.4	0.67	0.88	3.53	−7.7	6.9	1.18	3.17	6.10	−11.7	15.4
Panel D: EM	Total: 01/15–12/23					Pre-Covid-19: 01/15–01/20					Covid-19: 02/20–01/22				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
Standard	0.39	0.33	5.09	−15.4	14.8	0.47	0.24	4.57	−9.0	13.2	0.87	1.43	5.65	−15.4	9.2
Climate Change	0.37	0.23	5.15	−14.7	15.6	0.49	0.26	4.60	−8.9	13.1	0.85	1.26	5.58	−14.7	9.7
PAB	0.39	0.16	5.10	−15.8	13.3	0.47	0.13	4.56	−9.1	13.3	0.97	1.78	5.77	−15.8	9.7
SRI	0.54	0.21	5.44	−18.1	18.2	0.55	0.28	4.13	−8.7	13.6	1.46	2.05	6.86	−18.1	13.4
SRI PAB	0.36	0.18	5.12	−19.1	14.5	0.43	0.44	4.07	−8.6	13.2	1.01	2.03	6.75	−19.1	12.8

Panels A, B, C, D of this table respectively show the return statistics of the five indices examined for the World, the USA, Europe, and Emerging Markets (EM) region. We report the mean return, median return, the standard deviation (SD) of returns, the minimum and maximum return for the total study period (January 2015–December 2023), the pre-Covid-19 period (December 2015–January 2020), and the Covid-19 period (February 2020–January 2022). The five examined indices consist of the MSCI Standard (parent) Index, the MSCI Climate Change Index (CC), the MSCI Climate Paris-Aligned Index (PAB), the MSCI Socially Responsible Investment index (SRI) and the MSCI Socially Responsible Investment filtered Paris-Aligned Benchmark Index (SRI PAB). The asterixis next to the mean and median returns respectively indicate the *p* value for the t test and Wilcoxon rank sum test showing whether the return statistics of an index are statistically different from zero. ***p value ≤ 1%; **1% < p value ≤ 5%; *5% < p value ≤ 10%

demonstrate better performance than the standard index, as evidenced by greater or quite comparable mean returns. On average, the returns are highest in the USA and lowest in EM.

Panel A of Fig. 1 depicts the average cumulative return differentials (ACRD) for All Regions, and Panels B-E respectively show the cumulative return differentials (CRD) for the World, the USA, Europe, and EM.

As investors interested in sustainability are typically long-term oriented, our discussion below focuses on the relative performance over the *medium- and long-term* horizon. For All Regions, the World, and the USA, all the green indices exhibit positive CRDs (Panel A, B, and C). The ACRD (Panel A) highlights that the two indices with the most restrictive filters (SRI and SRI PAB) gain momentum and outperform their less green counterparts in the latter years

Footnote 19 (continued)

19 outbreak had started to ease, returning to pre-March 2020 levels, although this varied by country.

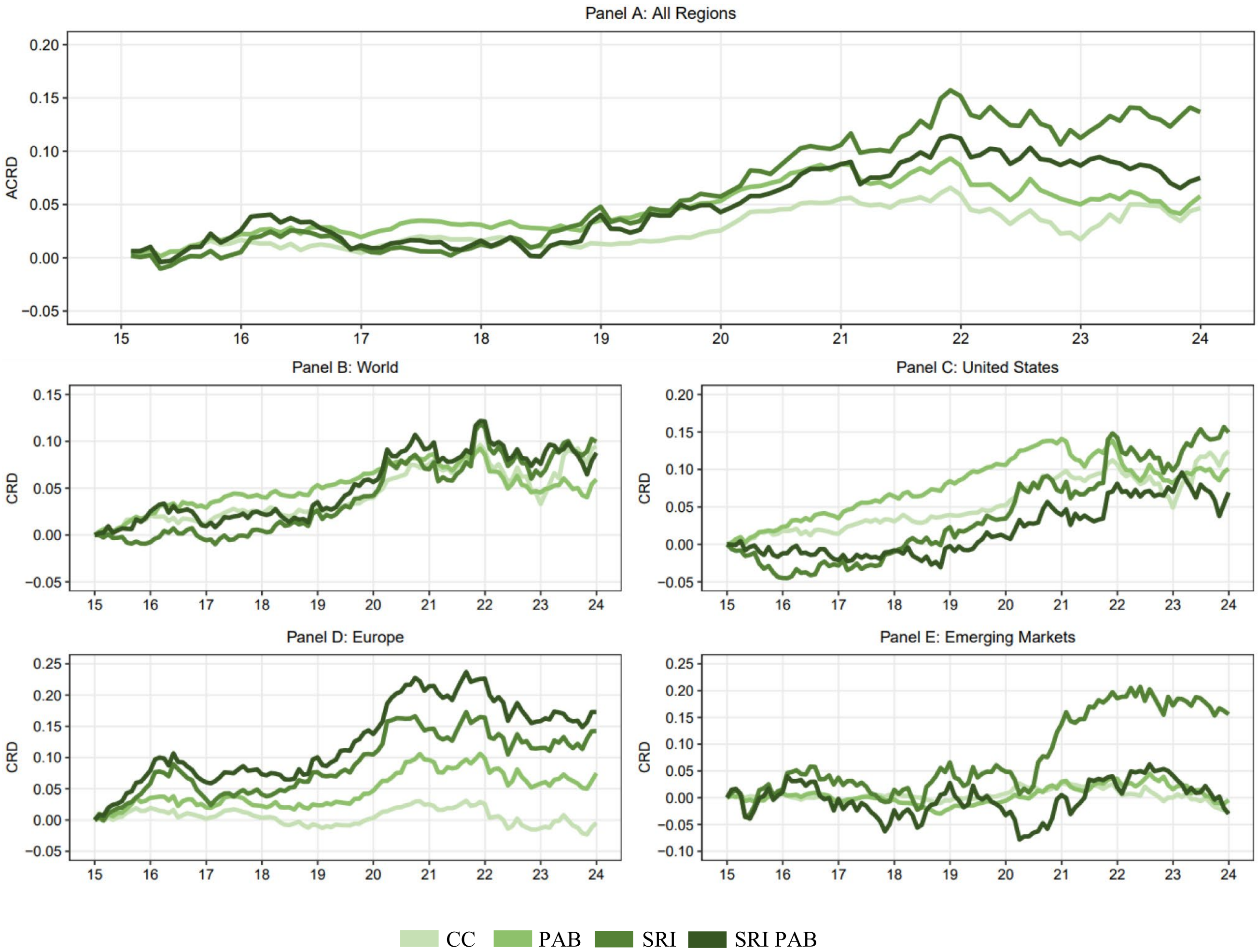

Figure 1 Cumulative return differential (CRD), January 2015–December 2023. This Figure presents the average cumulative return differential (ACRD) across regions and the cumulative return differential (CRD) of each MSCI green index relative to the MSCI standard/parent index in each region over the period January 2015-December 2023. The ACRD/ CRDs do not consider costs (ETFs' expense ratios). The four green indices consist of the MSCI Climate Change index (CC), the MSCI Climate Paris-aligned Benchmark index (PAB), the MSCI Socially Responsible Investment index (SRI) and the MSCI Socially Responsible Investment filtered Paris-aligned Benchmark Index (SRI PAB)

of the study period. For Europe, three green indices (except CC) achieve positive CRDs throughout the entire period (Panel D). However, for EM, such outperformance was not evident, with only the SRI exceeding other indices by a large margin from year five (2020) onward.

We report the ACRD for All Regions and the CRDs for each region at a five-year horizon and over the entire study period in columns 1–2, Table 2. Furthermore, we report net ACRDs and net CRDs in columns 3–4, Table 2.

Our analysis focuses on the *medium to long-term* horizons which suit investors who care about the long-term impacts of climate changes. Overall, with a few exceptions (the CC in Europe and EM, the PAB and SRI PAB in EM), the green indices in each region achieve positive net CRDs (relative to their standard index). Over our study period and across regions, the net ACRDs vary between 4% (CC) and 13% (SRI), see Panel A (column 4). The SRI is the best performer, achieving the highest net CRD in All Regions (13%, Panel A), in the World (9.2%, Panel B), in the USA (14.1%, Panel C), and in EM (14.9%, Panel E). The SRI PAB, constructed with the most restrictive criteria, delivers the best net CRD in Europe (16.9%, Panel D).

Columns 5–6 of Table 2 respectively show the wealth relative (WR) and Tobin's Q of each index averaged over the entire study period.[20] Across regions, all green indices generate WRs equivalent to or greater than one, indicating

[20] As index's constituent weighting data is unavailable beyond December 2022, our analysis of index-level Tobin's Q ends in December 2022 (instead of December 2023 as for the analysis of index's returns).

Table 2 Index's financial performance measures

Panel A: ACRD - Combined Regions						
	ACRD (without cost)		ACRD (after deducting ETF's costs)		Wealth Relative	Tobin's Q
All regions	5 Years (1)	Total (2)	5 Years (3)	Total (4)	Total (5)	Total (6)
Standard	–	–	–	–	1.00	2.08
CC	3.10%	4.70%	2.70%	4.00%	1.03	2.21
PAB	5.90%	5.80%	5.60%	5.10%	1.05	2.21
SRI	6.30%	13.70%	5.90%	13.00%	1.07	2.47
SRI PAB	4.70%	7.50%	4.30%	6.80%	1.05	2.41
Panel B: CRD - The World						
	CRD (without cost)		CRD (after deducting ETF's costs)		Wealth Relative	Tobin's Q
World	5 Years (1)	Total (2)	5 Years (3)	Total (4)	Total (5)	Total (6)
Standard	–	–	–	–	1.00	2.24
CC	4.60%	9.50%	4.20%	8.80%	1.04	2.44
PAB	7.00%	5.90%	6.50%	5.20%	1.05	2.29
SRI	4.80%	10.00%	4.40%	9.20%	1.04	2.54
SRI PAB	6.00%	8.80%	5.50%	8.10%	1.06	2.56
Panel C: CRD - The USA						
USA	5 Years (1)	Total (2)	5 Years (3)	Total (4)	Total (5)	Total (6)
Standard	–	–	–	–	1.00	2.66
CC	5.90%	12.40%	5.40%	11.60%	1.06	2.86
PAB	11.30%	10.10%	10.90%	9.30%	1.08	2.91
SRI	4.30%	14.90%	3.80%	14.10%	1.04	3.00
SRI PAB	0.70%	6.90%	0.30%	6.10%	1.02	2.94
Panel D: CRD - Europe						
Europe	5 Years (1)	Total (2)	5 Years (3)	Total (4)	Total (5)	Total (6)
Standard	–	–	–	–	1.00	1.61
CC	0.80%	−0.40%	0.60%	−0.80%	1.01	1.69
PAB	5.50%	7.50%	5.30%	7.10%	1.05	1.71
SRI	11.20%	14.30%	11.00%	13.90%	1.10	2.18
SRI PAB	14.70%	17.30%	14.50%	16.90%	1.14	2.11
Panel E: CRD - Emerging Markets (EM)						
EM	5 Years (1)	Total (2)	5 Years (3)	Total (4)	Total (5)	Total (6)
Standard	–	–	–	–	1.00	1.81
CC	1.00%	−2.80%	0.60%	−3.50%	1.00	1.86
PAB	−0.10%	−0.40%	−0.50%	−1.10%	1.00	1.93
SRI	4.80%	15.60%	4.40%	14.90%	1.08	2.14
SRI PAB	−2.60%	−3.00%	−3.00%	−3.70%	1.00	2.04

This table presents the average cumulative return differential (ACRD) in Panel A (columns 1–4), the cumulative return differential (CRD) in Panel B–E (columns 1–4), wealth relative (WR) in Panels A–E (column 5), and Tobin's Q in Panel A–E (column 6). CRD, ACRD and WR are relative to the MSCI Standard (parent) Index for each region (the World, USA, Europe, and Emerging Markets). Columns (1) and (2) of each Panel report the indices' ACRD/ CRD. Columns (3), (4) of each Panel report the net ACRD/ net CRD of corresponding synthetic Exchange Traded Funds (ETFs), which replicate our examined indices, after deducting total expense ratios. For this purpose, we calculate the average total expense ratios of the largest ETF providers for the non-sustainable and green indices, separately for each region. We then deduct 1/12 of the average annual cost from the monthly return of each index. Each performance metric (CRD, ACRD) is reported for two investment horizons starting from the study beginning in January 2015. The five-year horizon ended shortly before the spread of Covid-19 in January 2020. The analysis of index's relative returns and wealth relatives (columns 1–5) covers the period January 2015–December 2023 while the analysis of Tobin's Q (column 6) covers the period January 2015–December 2022. The calculation of the Tobin's Q of each index requires index's constituent weighting data which is only available until December 2022

The five examined indices consist of the MSCI Standard (parent) Index, the MSCI Climate Change Index (CC), the MSCI Climate Paris-Aligned Index (PAB), the MSCI Socially Responsible Investment Index (SRI) and the MSCI Socially Responsible Investment filtered Paris-Aligned Benchmark Index (SRI PAB)

Table 3 MSCI index's Sharpe and Treynor ratios, January 2015–December 2022

Panel A: Sharpe Ratio

	World (1)	USA (2)	Europe (3)	EM (4)	Average (5)
Standard	0.77	0.91	0.55	0.51	0.68
CC	0.81	0.94	0.59	0.53	0.73
PAB	0.8	0.96	0.61	0.63	0.75
SRI	0.84	0.97	0.68	0.60	0.79
SRI_PAB	0.82	0.91	0.69	0.46	0.78

Panel B: Treynor Ratio

	World (1)	USA (2)	Europe (3)	EM (4)	Average (5)
Standard	0.251	0.295	0.163	0.225	0.234
CC	0.265	0.306	0.176	0.238	0.246
PAB	0.257	0.308	0.182	0.241	0.247
SRI	0.270	0.308	0.221	0.254	0.263
SRI_PAB	0.270	0.291	0.212	0.216	0.247

This table presents the average Sharpe Ratio (Panel A) and Treynor Ratio (Panel B) for the regions World (Column 1), the USA (Column 2), Europe (Column 3), Emerging Markets (EM) (Column 4), and the average ratios across regions (Column 5)

The five examined indices consist of the MSCI Standard (parent) Index, the MSCI Climate Change Index (CC), the MSCI Climate Paris-Aligned Index (PAB), the MSCI Socially Responsible Investment Index (SRI) and the MSCI Socially Responsible Investment filtered Paris-Aligned Benchmark Index (SRI PAB)

that investors would be better off by investing in a green index instead of the respective standard index (column 5). Consistent with CRDs reported in columns 1–4, WRs show that the most restrictive green index, the SRI PAB, in Europe was the best performer. An investor who invested in SRI PAB Europe in January 2015 would be 14% better off over the study period than if s/he invested in the corresponding standard index. In terms of valuation, all green indices in all four regions achieve greater Tobin's Q than the respective standard index (column 6). Of particular interest, across all four regions, the two most restrictive green indices, the SRI and SRI PAB, achieve the highest Tobin's Q ranging between 2.14 and 3, and 2.04 and 2.94, respectively.

Columns (1–5) of Table 3 respectively display the Sharpe ratio (Panel A) and Treynor ratio (Panel B) of each index averaged over the study period for the World, the USA, Europe, Emerging Markets (EM), and the average ratio across regions.

Except the SRI PAB in EM region, all green indices achieve better Sharpe *and* Treynor ratios than the standard index. The Sharpe ratios for the green indices range from 0.80 to 0.84 (World, column 1), 0.91–0.97 (USA, column 2), 0.59–0.69 (Europe, column 3), and 0.46–0.63 (EM, column 4), which are mostly better than the standard index's values of 0.77 (World), 0.91 (USA), 0.55 (Europe), and 0.51 (EM). On average (column 5), the green indices consistently outperform the standard index as evidenced by their greater Sharpe and Treynor ratios. The SRI is the best performer overall (with the Sharpe ratio of 0.79 and the Treynor ratio of 0.26), followed closely by the SRI PAB (0.78 and 0.247) and PAB (0.75 and 0.247).

Overall, the green indices *mostly* offer either comparable or greater CRD, relative to their respective standard index, over a *medium- to long-term* horizon. This finding holds using net CRD that considers the corresponding largest synthetic ETFs' total expense ratios which can sometimes be higher for green investment options. The evidence of the green indices outperforming the standard index is robust using alternative measures including WR, Tobin's Q, and two risk-adjusted measures namely the Sharpe ratio and the Treynor ratio.

Our findings are in line with previous literature reporting outperformance such as Fiordelisi et al. (2023) but contrast to the broader literature which does mostly not find a performance difference (Dumitrescu et al. 2023; Cunha et al. 2020; Jain et al. 2019; Kossentini et al. 2024).

Returns co-movement

A prevalent concern regarding green investment indices is the potential lack of diversification due to high exclusion rates which range between 75 and 90%.[21] To address this

[21] For example, the MSCI Emerging Markets SRI Filtered PAB had 177 constituents in March 2024 compared to the 1376 constituents of the standard index, representing an exclusion ratio of 89%. In the other regions, the exclusion rates for the SRI and SRI PAB are around 75%.

Table 4 Coefficient estimates of the market factor (r_m-r_f) in Fama–French Regressions, January 2015–December 2023

Panel A: The World					
	Obs	FF5 (1)	FF3 (2)	C4 (3)	FF6 (4)
Standard	108	1.007***	1.004***	1.009***	1.009***
Climate change	108	1.004***	1.019***	1.009***	1.000***
PAB	108	1.001***	1.007***	1.002***	0.998***
SRI	108	0.993***	0.993***	0.991***	0.991***
SRI PAB	108	0.995***	0.996***	1.001***	0.999***
Panel B: The USA					
	Obs	FF5 (1)	FF3 (2)	C4 (3)	FF6 (4)
Standard	108	0.996***	0.996***	0.995***	0.995***
Climate change	108	1.016***	1.022***	1.016***	1.013***
PAB	108	1.006***	1.011***	1.013***	1.009***
SRI	108	0.997***	0.996***	0.997***	0.998***
SRI PAB	108	0.995***	0.997***	1.005***	1.002***
Panel C: Europe					
	Obs	FF5 (1)	FF3 (2)	C4 (3)	FF6 (4)
Standard	108	0.988***	0.982***	0.982***	0.986***
Climate change	108	1.004***	0.999***	0.991***	0.995***
PAB	108	0.997***	0.997***	0.983***	0.984***
SRI	108	0.969***	0.968***	0.954***	0.956***
SRI PAB	108	0.980***	0.977***	0.973***	0.976***
Panel D: Emerging Markets (EM)					
	Obs	FF5 (1)	FF3 (2)	C4 (3)	FF6 (4)
Standard	108	1.024***	1.031***	1.025***	1.017***
Climate change	108	1.036***	1.039***	1.025***	1.021***
PAB	108	1.036***	1.034***	1.039***	1.041***
SRI	108	1.095***	1.063***	1.059***	1.092***
SRI PAB	108	1.046***	1.007***	1.000***	1.042***

This table presents the coefficient estimates of the market factor (r_m–r_f) for each index in each region, obtained from the five-factor Fama–French (FF) model (FF5, column 1, our baseline analysis), the three-factor FF model (FF3, column 2), the model with FF's three factors and Carhart's momentum factor (C4, column 3), and the model with FF's five factors and Carhart's momentum factor (FF6, column 4)

The five examined indices consist of the MSCI Standard (parent) Index, the MSCI Climate Change Index (CC), the MSCI Climate Paris-Aligned Index (PAB), the MSCI Socially Responsible Investment Index (SRI) and the MSCI Socially Responsible Investment filtered Paris-Aligned Benchmark Index (SRI PAB)

For brevity reason, the coefficient estimates of other factors are not presented. ***, ** and * represents *p* value significance at the 0.01, 0.05 and 0.10 levels respectively. The untabulated adjusted R^2s are above 93% in all models

concern, we explore the co-movements between the returns of each green index and the market index by estimating Fama-French models. We focus on the beta coefficient estimates of the market factor of the green indices. We report the market factor's beta estimates in Table 4.

We find that the market factor's beta estimates for all green indices in all four regions are close to 1 (between 0.96 and 1.095), suggesting no substantial disadvantage in terms of volatility compared with the market index (Panel A–D, column 1). The EM has all betas' estimates exceeding 1 (ranging between 1.02 and 1.095), which indicates slightly more volatility than other regions. The SRI in EM tends to be most volatile, with beta estimates varying between 1.06 and 1.095. This is not too surprising considering its outperformance relative to its green peers and the standard index

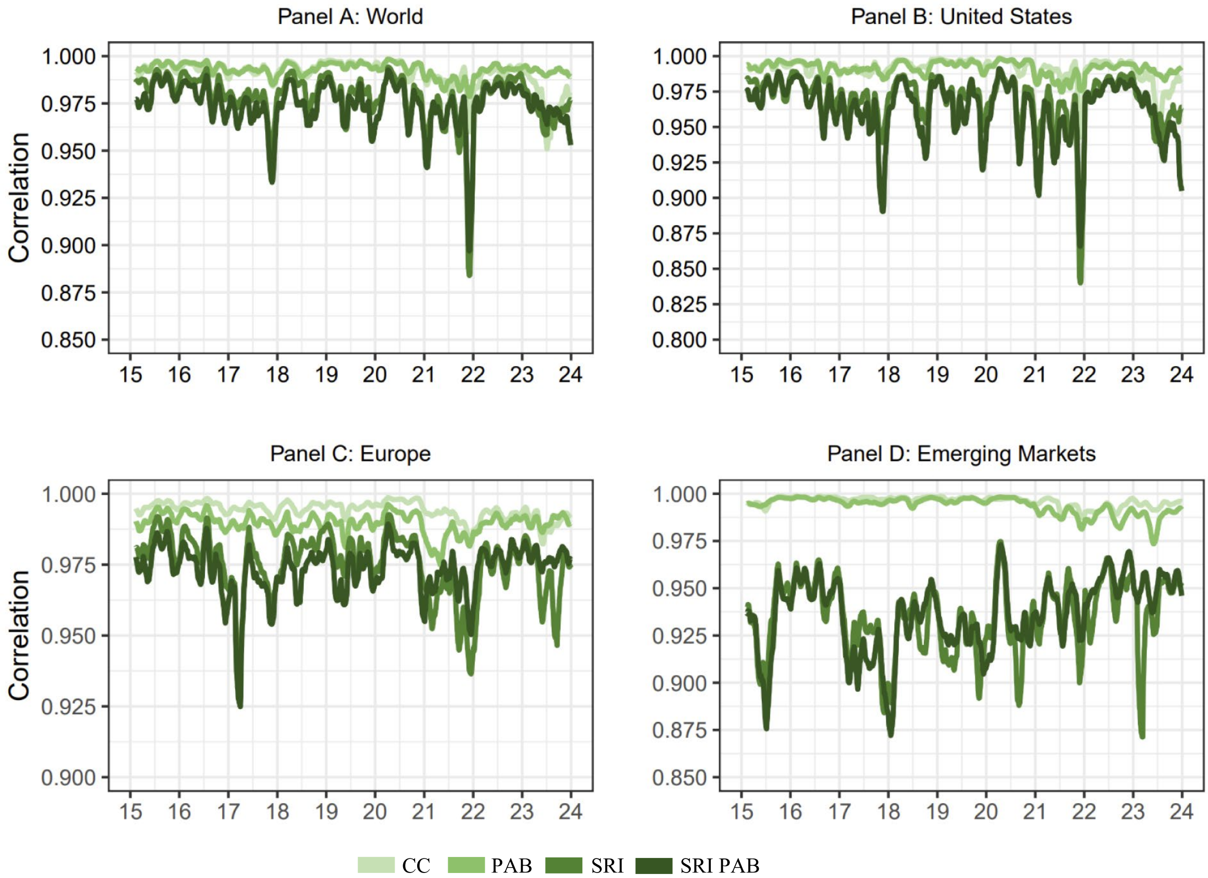

Figure 2 Dynamic conditional correlations between returns of each green index and its standard index, January 2015–December 2023. This figure features the dynamic conditional correlations between the returns of each green index and the returns of its standard parent index across four regions: World (Panel A), United States (Panel B), Europe (Panel C) and Emerging Markets (Panel B) during the period, January 2015–December 2023. The dynamic conditional correlations are obtained by estimating the GARCH (1,1) model with a Gaussian distribution using a moving average of 30 days on the obtained correlation values to smooth short-term fluctuations and reveal underlying trends. The four green indices consist of the MSCI Climate Change Index (CC), the MSCI Climate Paris-Aligned Index (PAB), the MSCI Socially Responsible Investment Index (SRI) and the MSCI Socially Responsible Investment filtered Paris-Aligned Benchmark Index (SRI PAB)

in EM region, as evidenced by a markedly higher net CRD (14.9%), a greater Sharpe ratio (0.6) and Treynor ratio (0.25).

Next, we examine the co-movements between a green index's returns and its respective standard index's returns. We apply the GARCH (1,1) model with a Gaussian distribution to daily returns, and present the correlation dynamics for the four regions in Fig. 2.

We observe a high degree of correlations between the green indices' returns and their respective standard index's returns in all four regions, particularly for the CC and PAB indices, which often mirror the movements of their parent indices very closely at a correlation coefficient of nearly 0.98. For EU, the USA, and the World, the SRI and SRI PAB display a high correlation (around 0.97) with their corresponding parent index, suggesting a strong linkage in market behaviors. For the EM, the SRI and SRI PAB consistently maintain a correlation gap of approximately 0.05–0.1 point from their parent index. This may partly be explained by the fact that constituents in the SRI and SRI PAB in EM are, on average, of smaller size[22] and invest conservatively[23] compared with those in other regions. Within the EM, the SRI and SRI PAB have 11–12.4% constituents from South Africa, greater than other indices (5.6–6.1%). Notably, the SRI PAB EM is more concentrated in the financial sector (34%) and less in the IT sector (13%) than other indices (23–26% in the Financial sector, and 21–25% in the IT sector).

[22] The weighted average market cap of SRI PAB (SRI) is USD 37 billion (79.9 billion) for EM and 61.3–137 billion (73.5-198 billion) for other regions (untabulated).

[23] The *Conservative minus Aggressive* (CMA) factor in the Fama-French (FF)5, FF6 regressions is statistically and significantly positive for SRI and SRI PAB indices in EM region only (untabulated).

A similar pairing is observed between SRI and SRI PAB indices. This clustering can be attributed to their compositions. These two indices are relatively similar regarding the number of excluded constituents, although the SRI PAB has an issuer cap of 5% and a more restrictive exclusion related to revenues realized from controversial activities.[24] Notably, in contrast to the CC and PAB, the SRI and SRI PAB exhibit several distinct phases where a deviation from the standard index occurred. This is not a surprise considering these indices' restrictive filters and the resulting constituents' re-weightings.

While the SRI and SRI PAB drift apart from the standard index at times, all green indices move quite closely with the standard index, as well as the market index. The finding of this analysis mitigates a concern of the green indices being undiversified, and is similar to the results of Jain et al. (2019) and Rompotis (2023).

OLS analysis: green criteria effectiveness (GCE) and index financial performance

We conduct index-level OLS analyses of CRD, WR, and Tobin's Q,[25] and report the results in columns 1–3, Table 5.[26] Each model is estimated using Newey–West standard errors. We lag all control variables by one month to mitigate concern of reverse causality. The GVIF for the group of four GCE dummy variables, estimated using DescTools in R, is less than 1.7, suggesting no serious concern of multicollinearity.

Our key variables of interest are the GCE dummies, each represents the criteria MSCI uses to construct a green index type. Except the CC index, which has no significant effect on Tobin's Q (column 3), the GCE coefficient estimates for all green indices are positive and statistically significant at the 1% level. The two most restrictive indices, GCE_SRI PAB (GCE_SRI), exert the most pronounced effects, respectively raising CRDs (column 1) by 7.7% (5.9%), WR (column 2) by 9.1% (6.4%), and Tobin's Q (column 3) by 0.635 (0.3).

To ensure that our results are robust, we repeat our baseline analyses using different subsamples and report the results in "Appendix B." Considering the relatively similar characteristics shared by two pairs of green indices, we first carry out separate OLS regressions of two subsamples of indices (Panel A): one with the standard, CC and PAB indices (Column 1–3), and one with the standard, SRI and SRI

Table 5 OLS financial performance analysis, January 2015–December 2022

Variables	CRD (1)	Wealth relative (2)	Tobin's Q (3)
GCE_CC	0.02***	0.037***	−0.041
	(0.007)	(0.008)	(0.035)
GCE_PAB	0.053***	0.065***	0.128***
	(0.009)	(0.01)	(0.045)
GCE_SRI	0.059***	0.064***	0.297***
	(0.004)	(0.005)	(0.019)
GCE_SRI PAB	0.077***	0.091***	0.635***
	(0.009)	(0.009)	(0.042)
Sector composite	YES	YES	YES
HDI composite	YES	YES	YES
Log of market value	YES	YES	YES
Post Launch dummy	YES	YES	YES
Ukraine dummy	YES	YES	YES
Covid dummy	YES	YES	YES
Observations	1900	1900	1900
Adj. R^2	0.528	0.547	0.894

This table reports the coefficient estimates of the Green Criteria Effectiveness (GCE) of each green index obtained from the OLS index-level analysis of cumulative return differential (CRD, column 1), wealth relative (WR, column 2), and Tobin's Q (column 3). Our analysis includes all five examined indices namely the MSCI Standard (parent) Index, the MSCI Climate Change Index (CC), the MSCI Climate Paris-Aligned Index (PAB), the MSCI Socially Responsible Investment Index (SRI) and the MSCI Socially Responsible Investment filtered Paris-Aligned Benchmark Index (SRI PAB)

Each model, estimated with the Newey–West robust standard errors, includes GCE index-type dummy values, lagged firm-weighted percentages for each Human Development Indicator (HDI) group, lagged firm-weighted percentages for each sector, logarithm of lagged firms-weighted market value, lagged dummy variables that capture an index's post-launch period, the Ukraine war and the Covid-19 pandemic. Index values (CRD, WR, Tobin's Q, and control variables except dummies) are estimated using constituents' values and their weights in the index in each month

The coefficient estimates of GCEs are reported first, followed by the standard errors (in parentheses). For brevity reason, the coefficient estimates of the intercept (the standard parent index) and control variables are not reported. ***, ** and * represents p value significance at the 0.01, 0.05 and 0.10 levels respectively. The untabulated F-statistics for the tests of the joint significance of the coefficients of variables employed are significant at the 1% level in all model specifications

[24] Please see a detailed discussion in "Brief overview of MSCI's methods for creating green equity indices" section and "Appendix A."

[25] Firm-level data for Tobin's Q was winsorized at the 1% and 99% level to remove outliers before being aggregated at the index level (Equation 5). Index's values of relative return measures (CRD, WR) are *not* affected by firm-related outliers. Untabulated additional analysis of CRD, WR, and Tobin's Q index values winsorized at the 1% and 99% level generates similarly qualitative result as the result reported in Table 5.

[26] In all regressions, the GCE values of the standard index are included in the intercept; the results do not change if the standard index observations are omitted from each regression. However, if we take out the standard index observations, one green index would be included in the intercept. Our analysis utilizes index-month observations of 20 indices over 8 years (2015–2022). Excluding 20 observations in Jan 2015 due to the use of lagged control variables, the final sample includes 1900 observations.

PAB indices (Column 4–6). Next, considering the dominance of US constituents, we analyze two subsamples of regions (Panel B): one excludes the World (Column 1–3), and one excluding the USA (Column 4–6). In panel C, we divide our sample into two sub-periods and conduct separate analyses for the market uncertainty (Column 1–3), and the market stability period (Column 4–6). Finally, to minimize concerns of reverse causality, we estimate two OLS regressions using two samples (Panel D): one with two-month lagged control variables (Column 1–3), and one with three-month lagged control variables (Column 4–6). We observe a quite consistent result that the GCEs statistically and significantly contribute to green indices' relative outperformance (CRD, WR) and, except for the GCE of CC, greater valuation (Tobin's Q). The most restrictive index, the SRI PAB, exerts the strongest impact on CRD, WR and Tobin's Q.

While the GCE of the CC index does not show a significant positive effect on Tobin's Q, this does not impact the overall implications of our study. Relative return measures CRD and WR are derived purely from market data, whereas Tobin's Q is derived based on both market and accounting data which, to some extent, is subject to earnings management. Tobin's Q shows whether an index is over-valued or under-valued (relative to the replacement cost of its assets). Tobin's Q is a valuation metric and does not capture the returns of green indices relative to their standard indices. Thus, it is *not* surprising that the result of the Tobin Q's analysis slightly differs from those of CRD and WR' analyses. The difference is mainly for the CC index (the least restrictive of the four examined green indices in each region). What matters most is that the significant effects of GCEs on Tobin's Q for the other three green indices indicate that market participants do recognize and reward the stricter green criteria employed—especially in the case of the SRI and SRI PAB indices, which are closely aligned with the Paris Agreement. Overall, the robust results obtained from CRD and WR analyses, and the significant effects of GCEs for SRI/SRI PAB (the two restrictive green indices that aligned with the Paris Climate Agreement) across all three measures (CRD, WR, Tobin's Q) remain the key drivers of our conclusions.

Conclusion

Our study analyses the financial performance 20 MSCI equity indices from 2015 to 2023 and aims at examining how the green criteria (GCE) used by MSCI to create green indices influence their financial performance. We compare the MSCI standard (parent) non-green index with four MSCI green indices in each of the four examined regions, including two that align with the Paris Agreement's objectives.

Our analyses reveal several findings. *First,* averaged across the four regions, the four green indices (CC, PAB, SRI, and SRI PAB) achieve better financial performance than the standard index, as demonstrated by long-term positive cumulative return differentials (CRD) and the wealth relatives (WR) being greater than 1 over the study period, particularly for the SRI indices. While we do see short-term underperformance, this is not unexpected considering the well-known trade-off between investing in sustainability and achieving immediate financial returns. Most importantly, all, except one, green indices deliver greater risk-adjusted returns (Sharpe and Treynor ratios), and higher Tobin's Qs indicating greater market valuation than the standard index. We are, however, cautious for the EM region as only the SRI perform substantially better than the standard index. *Second*, the returns of the four green indices are closely related to the returns of the market index and the returns of the standard index. The systematic risks of these green indices, as reflected by their beta coefficient estimates derived from Fama-French models, are similar to those of standard indices, suggesting that they do not compromise on diversification. *Third*, MSCI's GCEs used to construct green indices are significantly associated with their relative financial outperformance.

Our research adds valuable insights into the sustainable finance literature in several ways. *First,* we focus on the PAB and SRI PAB indices, making this one of the first studies to explore the financial attractiveness of Paris Agreement-Aligned Benchmark investments. *Second*, we analyze the financial performance of MSCI green indices compared to their standard MSCI non-green counterparts, using a rich MSCI dataset with monthly weighting data from 2015 to 2022; thereby eliminating concerns of potential effects of an index provider's characteristics. *Third*, we contribute to the discussion on green equity investments, showing their potential as tools for diversification and improved returns.

Our findings offer important practical insights for investors, index providers, fund managers, and policymakers. Given the observed long-term outperformance, investors seeking to align their portfolios with the Paris Agreement goals could consider ETFs tracking the MSCI PAB and SRI PAB indices (SRI PAB only for Emerging Markets). Policymakers can also use these Paris Agreement-aligned indices as reference standards for sustainable finance initiatives, for example, under the Sustainable Finance Disclosure Regulation (SFDR) Article 9 or the EU Taxonomy. They can help to verify that financial products meet defined sustainability criteria and thus increase transparency and reduce greenwashing. Using these indices as benchmarks for government pension funds or other public investment vehicles can further help to direct capital toward projects which aim to mitigate climate change. This is increasingly important considering that New Zealand Superannuation Fund and Norges Bank Investment Management, which manages Norway's sovereign wealth fund, have divested from oil companies,[27] and other sovereigns will follow suit. Our established evidence of the outperformance of the two Paris Agreement-aligned indices offers sovereign wealth funds important tools to pursue sustainable investment strategies and accelerate governments' efforts to support the transition to a low-carbon economy.

Our study has certain limitations as it focuses on 20 MSCI indices over a specific period, which includes the COVID-19 pandemic and the ongoing war in Ukraine, potentially influencing the results due to these unusual economic conditions. The limited availability of historical constituent weighting data across regions also restricts the scope of our analysis, making it less applicable to other indices or time periods. Future research could address these limitations by examining a wider set of indices over a longer period. It could also explore how ESG/green criteria used by different index providers influence performance under different economic conditions.

Appendix A: Comparison of exclusion thresholds in green MSCI's indices

The MSCI SRI and SRI Filtered Paris-Aligned Benchmark (PAB) indices are designed to reflect socially responsible investing principles while aligning with sustainability and climate-focused goals. These methods impose strict thresholds on revenues derived from controversial or sensitive activities to determine company (constituents) eligibility. The thresholds vary between the SRI Index and the stricter SRI Filtered PAB Index, ensuring that investments meet higher environmental, social, and governance (ESG) standards. The following summary outlines the key thresholds for these activities based on the MSCI SRI Indexes Methodology (MSCI 2024a) and the MSCI SRI Filtered PAB Indexes Methodology (MSCI 2024b). For *Controversial* Weapons, companies are excluded entirely from SRI and SRI filtered PAB indices, with a 0% threshold for all revenues. For *Conventional Weapons*, the SRI index allows up to 5% of revenues from production and 15% from components, while the SRI filtered PAB reduces this threshold to 5% for both production and components. *Civilian Firearms* are also restricted, with production revenues capped at 0% and distribution revenues capped at 5% for both the SRI and SRI filtered PAB indices. Similarly, for *Nuclear activities*, revenues from weapons production are restricted to 0%, power generation is capped at 5%, and nuclear suppliers are limited to 15% under the SRI index. However, the SRI filtered PAB reduces the supplier threshold to 5%. *Thermal Coal* is excluded entirely from the SRI and the SRI PAB index while both allow up to 5% of revenues from power generation. For *Tobacco*, production is not allowed (0% threshold) in both SRI indices, while distribution is capped at 5%. *Alcohol-related activities* have more lenient thresholds. Production revenues are capped at 5% for both SRI indices, and distribution revenues are capped at 15%. Similarly, *Gambling* is limited to 5% for ownership revenues and 15% for services in both SRI and SRI filtered PAB indices. For *Adult Entertainment*, production revenues are limited to 5%, and distribution is capped at 15% under both SRI indices. Genetically modified organisms (GMOs) are restricted to a 5% threshold for production revenues. *Oil and Gas* activities face stricter limitations. Revenues from conventional and unconventional oil and gas activities are capped at 0% for both SRI indices. However, power generation from oil and gas is allowed up to 30% under the SRI filtered PAB index.

[27] See https://www.greenpeace.org/aotearoa/press-release/super-funds-950m-fossil-fuel-divestment-an-aha-moment-for-nz-economy/ and https://www.theguardian.com/business/2017/nov/16/oil-and-gas-shares-dip-as-norways-central-bank-advises-oslo-to-divest.

Appendix B: OLS robustness analysis utilizing various subsamples

Panel A: Subsamples of indices

Variables	Sample: Standard, CC and PAB indices			Sample: Standard, SRI and SRI PAB indices		
Provider	(1)	(2)	(3)	(4)	(5)	(6)
	CRD	WR	Tobin's Q	CRD	WR	Tobin's Q
GCE_CC	0.082***	0.114***	0.23***	NA	NA	NA
	(0.005)	(0.005)	(0.053)	NA	NA	NA
GCE_PAB	0.11***	0.136***	0.322***	NA	NA	NA
	(0.007)	(0.006)	(0.059)	NA	NA	NA
GCE_SRI	NA	NA	NA	0.074***	0.08***	0.37***
	NA	NA	NA	(0.004)	(0.005)	(0.024)
GCE_SRI PAB	NA	NA	NA	0.123***	0.137***	0.618***
	NA	NA	NA	(0.007)	(0.008)	(0.051)
Sector composite	YES	YES	YES	YES	YES	YES
HDI composite	YES	YES	YES	YES	YES	YES
Log of market value	YES	YES	YES	YES	YES	YES
Post Launch dummy	YES	YES	YES	YES	YES	YES
Ukraine dummy	YES	YES	YES	YES	YES	YES
Covid dummy	YES	YES	YES	YES	YES	YES
Observations	1140	1140	1140	1140	1140	1140
Adj. R^2	0.733	0.83	0.922	0.66	0.664	0.895

Panel B: Subsamples of regions

Variables	Sample excluding "World" Region			Sample excluding "US" Region		
Provider	(1)	(2)	(3)	(4)	(5)	(6)
	CRD	WR	Tobin's Q	CRD	WR	Tobin's Q
GCE_CC	0.036***	0.038***	−0.083**	−0.013**	0.013*	−0.084**
	(0.007)	(0.008)	(0.038)	(0.006)	(0.007)	(0.037)
GCE_PAB	0.063***	0.066***	0.115**	0.012	0.033***	0.017
	(0.009)	(0.01)	(0.048)	(0.008)	(0.009)	(0.046)
GCE_SRI	0.071***	0.077***	0.332***	0.068***	0.068***	0.336***
	(0.005)	(0.006)	(0.021)	(0.004)	(0.005)	(0.021)
GCE_SRI PAB	0.094***	0.104***	0.645***	0.076***	0.092***	0.526***
	(0.009)	(0.01)	(0.043)	(0.008)	(0.009)	(0.043)
Sector composite	YES	YES	YES	YES	YES	YES
HDI composite	YES	YES	YES	YES	YES	YES
Log of market value	YES	YES	YES	YES	YES	YES
Post Launch dummy	YES	YES	YES	YES	YES	YES
Ukraine dummy	YES	YES	YES	YES	YES	YES
Covid dummy	YES	YES	YES	YES	YES	YES
Observations	1425	1425	1425	1425	1425	1425
Adj. R^2	0.56	0.564	0.901	0.603	0.598	0.888

Panel C: Subsamples of time periods

Variables	Sample market uncertainty			Sample market stability		
Provider	(1)	(2)	(3)	(4)	(5)	(6)
	CRD	WR	Tobin's Q	CRD	WR	Tobin's Q
GCE_CC	0.054***	0.009*	0.002	0.008	0.017	−0.219***
	(0.006)	(0.005)	(0.073)	(0.01)	(0.011)	(0.048)
GCE_PAB	0.136***	0.034***	0.146	0.038***	0.044***	−0.069
	(0.012)	(0.009)	(0.125)	(0.01)	(0.011)	(0.046)
GCE_SRI	0.112***	0.04***	0.377***	0.025***	0.026***	0.292***
	(0.007)	(0.005)	(0.070)	(0.004)	(0.004)	(0.02)
GCE_SRI PAB	0.152***	0.043***	0.515***	0.041***	0.047***	0.395***
	(0.008)	(0.006)	(0.081)	(0.012)	(0.014)	(0.057)
Sector composite	YES	YES	YES	YES	YES	YES
HDI composite	YES	YES	YES	YES	YES	YES
Log of market value	YES	YES	YES	YES	YES	YES
Post Launch dummy	YES	YES	YES	YES	YES	YES
Observations	700	700	700	1200	1200	1200
Adj. R^2	0.775	0.487	0.857	0.481	0.474	0.847

Panel D: Sample with two- and three-month lagged control variables

Variables	Sample with two-month lagged control variables			Sample with three-month lagged control variables		
Provider	(1)	(2)	(3)	(4)	(5)	(6)
	CRD	WR	Tobin's Q	CRD	WR	Tobin's Q
GCE_CC	0.024***	0.042***	−0.034	0.027***	0.046***	−0.058
	(0.007)	(0.008)	(0.035)	(0.007)	(0.008)	(0.036)
GCE_PAB	0.06***	0.073***	0.14***	0.064***	0.077***	0.105**
	(0.009)	(0.01)	(0.047)	(0.008)	(0.011)	(0.048)
GCE_SRI	0.063***	0.068***	0.331***	0.065***	0.071***	0.352***
	(0.005)	(0.005)	(0.023)	(0.005)	(0.005)	(0.024)
GCE_SRI PAB	0.081***	0.095***	0.653***	0.085***	0.099***	0.629***
	(0.009)	(0.01)	(0.045)	(0.008)	(0.01)	(0.045)
Sector composite	YES	YES	YES	YES	YES	YES
HDI composite	YES	YES	YES	YES	YES	YES
Log of market value	YES	YES	YES	YES	YES	YES
Post launch dummy	YES	YES	YES	YES	YES	YES
Ukraine dummy	YES	YES	YES	YES	YES	YES
Covid dummy	YES	YES	YES	YES	YES	YES
Observations	1880	1880	1880	1860	1860	1860
Adj. R^2	0.524	0.542	0.896	0.522	0.54	0.892

This "Appendix" reports the coefficient estimates of the Green Criteria Effectiveness (GCE) of each green index (the MSCI Climate Change Index-CC, the MSCI Climate Paris-Aligned Index-PAB, the MSCI Socially Responsible Investment Index-SRI, and the MSCI Socially Responsible Investment filtered Paris-Aligned Benchmark Index-SRI PAB) obtained from the OLS index-level analysis of the cumulative return differential (CRD, columns 1 and 4), wealth relative (WR, columns 2 and 5) and Tobin's Q (columns 3 and 6) for each subsample in Panels A–D

From the entire sample covering the period January 2015–December 2022, we create 8 different subsamples. In Panel A, we examine two subsamples of which one consists of the standard and two less strict CC and PAB indices (columns 1–3), and the other including the standard and two more strict SRI and SRI PAB indices (columns 4–6). In Panel B, we analyze two subsamples of which one excludes the "World" region (columns 1–3), and one exclude the "US" region (columns 4–6). In Panel C, we look at two different sub-periods: market uncertainty (columns 1–3) and market stability (columns 4–6). The market uncertainty period covers the Ukraine war and the Covid pandemic while the market stability period covers the remaining months of our study period. In Panel D, we explore a sample with two-month lagged control variables (columns 1–3) and a sample with three-month lagged control variables (columns 4–6). In both samples, the index-level HDI composite is lagged by a year

We estimate all models with the Newey–West robust standard errors. The coefficient estimates of GCEs are reported first, followed by the standard errors (in parentheses). For brevity reason, the coefficient estimates of the intercept and control variables are not reported. ***, ** and * represents p value significance at the 0.01, 0.05 and 0.10 levels respectively

Acknowledgement We would like to thank MSCI for generously providing us with historical monthly index constituent data. We are grateful for helpful comments from an anonymous reviewer, Tom Coupe, and Klaus Schäfer. All errors are ours. Huong Dang would like to thank the Erskine Grant Office and the Department of Economics and Finance at the University of Canterbury, the Bayreuth Humboldt Centre and the Chair of Finance and Banking at the University of Bayreuth (UBT) for the great support throughout her visit to UBT, during which this project was initiated.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Abate, G., I. Basile, and P. Ferrari. 2021. The level of sustainability and mutual fund performance in Europe: An empirical analysis using ESG ratings. *Corporate Social Responsibility and Environmental Management* 28 (5): 1446–1455. https://doi.org/10.1002/csr.2175.

Ang, W.R. 2015. Sustainable investment in Korea does not catch a cold when the United States sneezes. *Journal of Sustainable Finance & Investment* 5 (1–2): 16–26. https://doi.org/10.1080/20430795.2015.1042737.

Barnett, M.L., and R.M. Salomon. 2006. Beyond dichotomy: The curvilinear relationship between social responsibility and financial performance. *Strategic Management Journal* 27 (11): 1101–1122. https://doi.org/10.1002/smj.557.

Benson, C.C., N. Gupta, and R. Mateti. 2010. Does risk reduction mitigate the costs of going green? An empirical study of sustainable investing. *Southern Journal of Business and Ethics* 2: 7–25.

Bolognesi, E., A. Dreassi, M. Migliavacca, and A. Paltrinieri. 2024. Mapping sustainable investing: Exploring ambiguities and consistencies among sustainable indices. *Journal of Environmental Management* 367: 122081. https://doi.org/10.1016/j.jenvman.2024.122081.

Carhart, M.M. 1997. On persistence in mutual fund performance. *The Journal of Finance* 52 (1): 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x.

Collison, D.J., G. Cobb, D.M. Power, and L.A. Stevenson. 2008. The financial performance of the FTSE4Good indices. *Corporate Social Responsibility and Environmental Management* 15 (1): 14–28. https://doi.org/10.1002/csr.144.

Consolandi, C., A. Jaiswal-Dale, E. Poggiani, and A. Vercelli. 2009. Global Standards and Ethical Stock Indexes: The Case of the Dow Jones Sustainability Stoxx index. *Journal of Business Ethics* 87 (S1): 185–197. https://doi.org/10.1007/s10551-008-9793-1.

Cortina, H., and M. Phyu. 2024. Sustainable funds continue to outgrow peers in 2023. *ISS Insights*. Retrieved from https://insights.issgovernance.com/posts/sustainable-funds-continue-to-outgrow-peers-in-2023/.

Cunha, F.A.F.D.S., E.M. De Oliveira, R.J. Orsato, M.C. Klotzle, F.L. Cyrino Oliveira, and R.G.G. Caiado. 2020. Can sustainable investments outperform traditional benchmarks? Evidence from global stock markets. *Business Strategy and the Environment* 29 (2): 682–697. https://doi.org/10.1002/bse.2397.

Dumitrescu, A., J. Järvinen, and M. Zakriya. 2023. Hidden Gem or Fool's Gold: Can passive ESG ETFs outperform the benchmarks?. *International Review of Financial Analysis* 86: 102540. https://doi.org/10.1016/j.irfa.2023.102540.

ElBannan, M.A. 2024. Returns behavior of ESG ETFs in the COVID-19 market crash: Are green funds more resilient?. *Journal of Corporate Accounting & Finance* 35 (2): 187–223. https://doi.org/10.1002/jcaf.22680.

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

Fama, E.F., and K.R. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116 (1): 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010.

Fama, E.F., and K.R. French. 2018. Choosing factors. *Journal of Financial Economics* 128 (2): 234–252. https://doi.org/10.1016/j.jfineco.2018.02.012.

Fiordelisi, F., G. Galloppo, G. Lattanzio, and V. Paimanova. 2023. Looking at socially responsible investment strategies through the lenses of the global ETF industry. *Journal of International Money and Finance* 137: 102917. https://doi.org/10.1016/j.jimonfin.2023.102917.

Folger-Laronde, Z., S. Pashang, L. Feor, and A. ElAlfy. 2022. ESG ratings and financial performance of exchange-traded funds during the COVID-19 pandemic. *Journal of Sustainable Finance & Investment* 12 (2): 490–496. https://doi.org/10.1080/20430795.2020.1782814.

French, K.R. 2024. *Data Library*. Retrieved from https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Giese, G., L.-E. Lee, D. Melas, Z. Nagy, and L. Nishikawa. 2019. Foundations of ESG investing: How ESG affects equity valuation, risk, and performance. *The Journal of Portfolio Management* 45 (5): 69–83. https://doi.org/10.3905/jpm.2019.45.5.069.

Gregory, R.P., J.G. Stead, and E. Stead. 2021. The global pricing of environmental, social, and governance (ESG) criteria. *Journal of Sustainable Finance & Investment* 11 (4): 310–329. https://doi.org/10.1080/20430795.2020.1731786.

Hong, H., and M. Kacperczyk. 2009. The price of sin: The effects of social norms on markets. *Journal of Financial Economics* 93 (1): 15–36. https://doi.org/10.1016/j.jfineco.2008.09.001.

Huang, Y. 2024. Do ESG ETFs provide downside risk protection during Covid-19? Evidence from forecast combination models. *International Review of Financial Analysis* 94: 103320. https://doi.org/10.1016/j.irfa.2024.103320.

Jacob, A., and M. Wilkens. 2021. What drives sustainable indices? A framework for analyzing the sustainable index landscape. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3874437.

Jain, M., G.D. Sharma, and M. Srivastava. 2019. Can sustainable investment yield better financial returns: A comparative study of ESG indices and MSCI indices. *Risks* 7 (1): 15. https://doi.org/10.3390/risks7010015.

Jain, N., V. Mehrotra, and P. Mendiratta. 2023. Performance analysis of ESG indices: A step towards fulfilling social responsibility. *Bimaquest* 23 (1): 22–37.

Jonwall, R., S. Gupta, and S. Pahuja. 2024. Do socially responsible indices outperform conventional indices? Evidence from before and after the onset of Covid-19. *Corporate Social Responsibility and Environmental Management* 31 (5): 4995–5011. https://doi.org/10.1002/csr.2839.

Kossentini, H., O. Belhassine, and A. Zenaidi. 2024. ESG index performance: European evidence. *Journal of Asset Management* 25 (7): 653–665. https://doi.org/10.1057/s41260-024-00361-4.

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

Lin, X., and R.B. Swain. 2024. Performance of negatively screened sustainable investments during crisis. *International Review of Economics & Finance* 93: 1226–1247. https://doi.org/10.1016/j.iref.2024.04.001.

Managi, S., T. Okimoto, and A. Matsuda. 2012. Do socially responsible investment indexes outperform conventional indexes?. *Applied Financial Economics* 22 (18): 1511–1527. https://doi.org/10.1080/09603107.2012.665593.

MSCI. 2024a. MSCI SRI indexes methodology. Retrieved from: https://www.msci.com/documents/10199/2f754e5a-310a-0b5f-5273-5921648b1a7c.

MSCI. 2024b. MSCI SRI filtered PAB indexes methodology. Retrieved from: https://www.msci.com/documents/10199/d21129da-4fe6-9c55-81d0-eb42d4931b51.

Ng, A.C., and Z. Rezaee. 2015. Business sustainability performance and cost of equity capital. *Journal of Corporate Finance* 34: 128–149. https://doi.org/10.1016/j.jcorpfin.2015.08.003.

Omura, A., E. Roca, and M. Nakai. 2021. Does responsible investing pay during economic downturns: Evidence from the COVID-19 pandemic. *Finance Research Letters* 42: 101914. https://doi.org/10.1016/j.frl.2020.101914.

Ortas, E., R.L. Burritt, and J.M. Moneva. 2013. Socially responsible investment and cleaner production in the Asia Pacific: Does it pay to be good?. *Journal of Cleaner Production* 52: 272–280. https://doi.org/10.1016/j.jclepro.2013.02.024.

Ortas, E., J.M. Moneva, R. Burritt, and J. Tingey-Holyoak. 2014. Does sustainability investment provide adaptive resilience to ethical investors? Evidence from Spain. *Journal of Business Ethics* 124 (2): 297–309. https://doi.org/10.1007/s10551-013-1873-1.

Ortas, E., J.M. Moneva, and M. Salvador. 2012. Does socially responsible investment equity indexes in emerging markets pay off? Evidence from Brazil. *Emerging Markets Review* 13 (4): 581–597. https://doi.org/10.1016/j.ememar.2012.09.004.

Papathanasiou, S., and D. Koutsokostas. 2024. Sustainability ratings and fund performance: New evidence from European ESG equity mutual funds. *Finance Research Letters* 62: 105095. https://doi.org/10.1016/j.frl.2024.105095.

Pavlova, I., and M.E. De Boyrie. 2022. ESG ETFs and the COVID-19 stock market crash of 2020: Did clean funds fare better?. *Finance Research Letters* 44: 102051. https://doi.org/10.1016/j.frl.2021.102051.

Renneboog, L., J. Ter Horst, and C. Zhang. 2008. The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance.* 14 (3): 302–322. https://doi.org/10.1016/j.jcorpfin.2008.03.009.

Rompotis, G.G. 2022a. The ESG ETFs in the UK. *Journal of Asset Management* 23 (2): 114–129. https://doi.org/10.1057/s41260-021-00251-z.

Rompotis, G.G. 2022b. Environmental, social and governance responsibility, financial performance and assets: A study of exchange traded funds. *Research Papers in Economics and Finance* 6 (2): 23–49. https://doi.org/10.18559/ref.2022.2.2.

Rompotis, G.G. 2023. The performance of ESG ETFs in the US. *Capital Markets Review* 31 (2): 89–101.

Rompotis, G.G. 2024. The performance of anti-ESG ETFs in the United States. *Journal of Asset Management* 25 (7): 700–713. https://doi.org/10.1057/s41260-024-00381-0.

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

Statman, M. 2005. Socially responsible indexes: Composition and performance. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.705344.

Unruh, G., D. Kiron, N. Kruschwitz, M. Reeves, H. Rubel, and Meyer zum Felde, A. 2016. Investing for a Sustainable Future. *MIT Sloan Management Review*.

World Health Organization. 2020. Archived: WHO Timeline—COVID-19. Retrieved from: https://www.who.int/news/item/27-04-2020-who-timeline---covid-19.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Jan Heldmann is a PhD candidate and Research Associate at the Chair of Finance and Banking, University of Bayreuth, Germany. He earned both his Bachelor's degree in International Economics & Development (2021) and his Master's in Business Administration (specializing in Finance, Accounting, Controlling, and Taxation) (2023) from the University of Bayreuth. His Master's thesis, "Empirical Analysis of Variations in Factor Models," received the Bayreuth Award for Financial Services in 2024. From 2023 to early 2025, he served as a Project Manager for the BMBF-funded initiative titled "Climate Reporting by SMEs" (Klimaberichterstattung bei KMUs). His research interests include ESG, risk management, and factor models.

Thomas Brückner is a Master's candidate at the Chair of Finance and Banking, University of Bayreuth, Germany. He earned his Bachelor's degree in Finance in 2022 and will complete his Master's study in 2025, with a thesis titled "Factor models on the German stock market: An empirical comparison." Since 2023, he has worked as a Client Account Manager specializing in the Exchange-Traded Fund business at Amundi Investment Solutions (Germany), a leading European asset manager and one of the top three asset managers in terms of voting performance on environmental and social issues in 2023. In 2024, he was selected for Crédit Agricole's DEJourney talent development program (2024–2025). His research interests include financial performance, ESG, and risk management.

Huong Dieu Dang CFA, FRM is Senior Lecturer above the Bar at the University of Canterbury, New Zealand, and was a Visiting Professor at the University of Bayreuth, Germany. She earned an MBA from the University of Arizona (2006) and an MSc in Management (Finance concentration) from the University of Arizona (2006), before obtaining a PhD degree in Economics (Finance Discipline) at the University of Sydney (2010). Since 2020, she has served as an Associate Editor of the *Journal of Applied Accounting Research* published by Emerald. Her research interests include credit risk, investment, social norm, and sustainability. Her co-authored research has been published in respected peer-reviewed journals such as the *Journal of Financial Economics, ABACUS*, and *Pacific-Basin Finance Journal.*