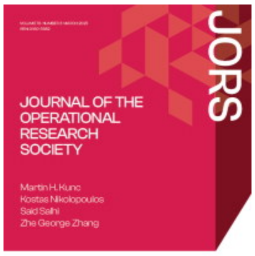

Journal of the Operational Research Society

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tjor20

Testing green finance portfolio performance

Enoch Quaye, Radu Tunaru & Diana Tunaru

JOURNAL OF THE OPERATIONAL RESEARCH SOCIETY
https://doi.org/10.1080/01605682.2025.2465895

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Testing green finance portfolio performance

Enoch Quaye[a], Radu Tunaru[b] and Diana Tunaru[c]

[a]University of Bristol, Bristol, UK; [b]University of Reading, Reading, UK; [c]University of Kent, Canterbury, UK

ABSTRACT
Green activities are measured with a green revenue adjustment factor that can be used to adjust observed market stock prices. We examine the green revenue factors for all companies that are part of the stock indexes representing the main five economies. Using multivariate correlation coefficients, we detect higher-order groupings of green indexes that may highly or lowly correlate. We employ the green revenues factor to construct portfolios that may benefit from the wedge between high green companies and low green companies, for all five economies. The quintile portfolios are compared across mean return, the CAPM beta, and realised beta. We also statistically test their comparative dollar performance using high-order stochastic dominance tests. The US portfolio has better dollar performance than the corresponding portfolios for the other economies, while the similar portfolio for Japan has the least dollar performance out of portfolios of all the other economies.

ARTICLE HISTORY
Received 29 July 2024
Accepted 4 February 2025

KEYWORDS
Green revenues factors; investment appraisal; portfolio theory; stochastic dominance test; sustainability

JEL CLASSIFICATIONS
G12; G14; G17

1. Introduction

Financial markets play an important role in facilitating the climate change agenda through significant financial innovation aiming to capture the firms' degree of green activities in the prices and the returns of market securities, see Zerbib (2019), Pástor et al. (2021), Pedersen et al. (2021), Pástor et al. (2022), Bolton and Kacperczyk (2022), Zerbib (2022). Edmans (2023) points out that at the end of 2021, the assets under management for investors who signed the Principles for Responsible Investment reached 121 trillion dollars, an increase of 50% since 2019 (Matos, 2020) and almost 20 fold since 2006. A report on Bloomberg on February 8, 2019, stated that Europe committed about 12 trillion dollars to sustainable investing whilst a report from Bank of America reported that the ratio of ESG versus non-ESG bond funds for Western Europe is almost 10%, \$2.6 billion inflows in ESG bond funds compared with \$29 billion non-ESG bond funds. For the US, the respective ratio is only 0.5% (\$0.82 billion in ESG bond funds compared with \$150 billion in non-ESG funds), see Temple-West (2023). Japan experiences an impetus in the ESG area, with a package of \$144 billion in decarbonisation bonds being announced over the next decade (Temple-West, 2023).

Portfolio optimization has been the cornerstone of financial markets for many years. Kolm et al. (2014) review the main techniques that evolved over the last several decades for portfolio construction. Their focus was mainly in building the portfolio. Another comprehensive review of the main financial applications of computational and data analytics approaches, is provided by Andriosopoulos et al. (2019) who discussed applications in portfolio management, credit analysis, banking, and insurance. Somehow surprisingly, none of the above two reviews included issues related to climate finance and non-pecuniary utility preferences.

The literature combining portfolio construction with sustainable finance investor preferences has been sparse for a long time, a notable exception being Hallerbach et al. (2004) who showed how to change the usual portfolio construction paradigm to incorporate sustainable characteristics of the firms in the investment portfolio universe. Recently, there has been much interest in portfolio construction and climate finance. Pedersen et al. (2021) generalize the Markowitz mean-variance optimization paradigm by including an additional constraint for the portfolio ESG score. The revised methodology results in a bi-criteria Sharpe ratio-ESG optimization method. Steuer and Utz (2023) took a step further and introduced a tri-criteria mean-variance ESG optimization that considers the portfolio ESG score as an objective function to be optimized. They claim that their multi-objective portfolio optimization problem always provides mean-variance – ESG-efficient solutions because it belongs to the class of

CONTACT Enoch Quaye enoch.quaye@bristol.ac.uk University of Bristol Business School, University of Bristol, Bristol BS8 1TU, UK

ε-constraint problems. However, Marohn and Auer (2024) advocate that this idea is problematic and in fact, the approach in Steuer and Utz (2023) cannot guarantee efficient portfolios. Other recent research expanded the knowledge frontier on the socially responsible multi-objective problem focusing on constructing optimal portfolios *via* the usual reward/risk maximization and incorporating a dependence structure among asset returns utilising vine copulas, see Sahamkhadam and Stephan (2024). They also employ the cumulative zero-order stochastic dominance and applied their technique to EuroStoxx 50 constituents, the results indicating that including social responsibility leads to reduced portfolio returns but also lower portfolio risk.

While the literature on portfolio construction based on optimising various objective functions is well developed and growing year on year (see Alexander & Baptista, 2004; Ban et al., 2018; Brodie et al., 2009; Capponi & Rubtsov, 2022; Cornuéjols et al., 2018; DeMiguel et al., 2009a; 2009b; 2013; DeMiguel & Nogales, 2009; Goldfarb & Iyengar, 2003; Lassance et al., 2022; Popescu, 2007), there is less research on testing the comparative portfolios that are routinely proposed using various methods. Since the seminal paper by Lo (2002), portfolio performance comparison was done using the Sharpe ratio as the main yardstick and looking at its distribution for statistical testing. Several subsequent improvements in this direction were proposed by Ledoit and Wolf (2008), Bao (2009), Liu et al. (2012), Ardia and Boudt (2015), Ledoit and Wolf (2017), Qi et al. (2018). However, it is not clear that the Sharpe ratios are the best measure for comparing portfolio performance, and furthermore, their distribution may be highly uncertain. Thus, an improved theoretical development has been employing stochastic order dominance for comparing portfolios performance using the entire distribution of returns. We advocate in this paper using stochastic dominance tests of the latest generation (Lee et al., 2023) to compare the portfolios constructed using long-short strategies based on the green revenues factor for all five major economies. These tests are of order higher than one and they are also constructed to test time consistent dominance.

In our study, we employ a novel granular dataset consisting of FTSE Russell Green Revenues Indexes for major economies and the daily green revenues factor values (GRF) for each constituent company of these indexes. The central element of the unique data we use is the GRF measure, which is calculated by FTSE Russell (part of LSEG now) and employed in the computation of Green Revenue indexes. We use the GRF information on all companies that are constituents of the major stock indexes in the top five economies to construct portfolios that reflect the level of green revenues as marketed by FTSE Russell (LSEG now). We build high minus low green revenue factor portfolios for all countries and compare their performance. We show that portfolios with US companies have a better dollar performance than portfolios using Japanese companies. Our analysis uses advanced testing for high-order stochastic dominance.

The main aim of this paper is to compare the green revenues within the major economies of the US, the UK, Europe, China and Japan, as well as of a general All-World. Furthermore, we compare the betas obtained for the green revenues adjusted share prices and show that the tilting towards green valuations impacts the distribution of the stock returns, the first two moments in particular. Our green CAPM models employ also the green revenues adjusted stock indexes, as calculated independently by the FTSE Russell (LSEG). This is the first time in the literature when a green CAPM model uses green adjusted stock prices and green adjusted stock indexes. We show that the tilting towards quasi green share prices reverts the relationship between the respective estimated betas and their corresponding realised counterparts.

We use the GRF adjustment factors to design investment strategies that may appeal to investors with green finance preferences. We compare the performance of these investment strategies across five major economies and we use recent statistical testing to compare the green-related investment strategies. Our results show a superior dollar return performance for portfolios using stocks from the US. This indicates that more profits can be extracted in those economies having firms that are more heterogeneous regarding ESG principles, than in those economies where perhaps due to tighter regulation, or indeed due to local cultural views towards climate change, firms are more homogeneous in their behaviour towards ESG.

The remainder of this paper is organised as follows. The next section contains a succinct literature review. This is followed by a description of the sample data, the main variables behind our analysis, and the methods employed. This section also includes a description of the process of green revenue measurement as it is independently carried out by FTSE Russell, which is part of the London Stock Exchange Group (LSEG) now. We then present in Section 4 the main empirical results regarding green portfolio comparison for the five major economies. Last section summarizes the main findings.

2. Literature review

The literature on the impact of environment on stock market prices and investors' views and behaviour has evolved over the years into a specialised strand, see Hamilton et al. (1993), Klassen and McLaughlin (1996), Konar and Cohen (2001), Heinkel et al. (2001), Kempf and Osthoff (2007). There is an intensive debate in the literature on the existence of an ESG risk premium that is directly associated with risks emerging from climate change. Cornell (2021) points out that ESG investments may be popular because of their social preferences but investors choosing this investment style should not expect high returns. From a theoretical perspective, Pástor et al. (2022) provide compelling reasons to indicate that high returns for green stocks reported in recent years should not be taken as indicative predictors of high future returns for the same stocks. They show that when investors take more green companies in their portfolios, the risk-adjusted expected returns on those firms will be lower in equilibrium. Similarly, Pedersen et al. (2021) construct an ESG adapted CAPM and show that employing a strategy based on the new efficient environment frontier does not necessarily lead to an improvement in the Sharpe ratio.

A single-period equilibrium model, built with partial segmentation and heterogeneous preferences focusing on regular investors and sustainable investors, has been developed by Zerbib (2022). This new model is a sustainable factor expanded CAPM model, which implies that sustainable investors may frequently influence the costs of capital for many firms through exclusionary screening and ESG integration.

There is an increasing strand of research combining sustainable investments and portfolio analysis. In a highly innovative paper, Ballestero et al. (2012) consider portfolio construction under an utility theory under uncertainty and also embedding an ethical goal. Their new financial-ethical bi-criteria model is derived with absolute risk aversion coefficients and targets depending on the investor's ethical preferences. Their numerical results point out that traditional efficient portfolios may outperform the strong green portfolios in terms of expected return and risk, but this is not the case with weak green investment. Gasser et al. (2017) revisit Markowitz' mean-variance methodology and suggest a way to incorporate also a social responsibility measure into the investment decision. Their method is applied in an a posteriori fashion and it allows investors to incorporate all ESG preferences. Their empirical analysis focuses on more than 6000 international companies, covering the complete universe of social responsibility-rated stocks, and it concludes that "investors opting to maximize the social impact of their investments do indeed face a statistically significant decrease in expected returns."

An excellent review of the latest green accounting and green finance literature is provided in the special issue discussed by Brooks and Schopohl (2021). Research in this area has been driven by environmental disclosures and reporting that may lead to possible future policies, but also by the impact of climate change on firm valuations. For this latter emerging strand of literature, some notable contributions are provided by Chapple et al. (2013), Clarkson et al. (2015), Johnston et al. (2008), and Matsumura et al. (2014), and the list is by no means exhaustive. Choi and Luo (2021) study the link between the size of a firm's carbon emissions and its stock market value, looking at an international sample covering 28 countries. They conclude that there is a negative relationship between carbon emissions and stock market value. However, Griffin et al. (2021) argue that the opposite may be true. Analysing panel data of 228 Canadian firms over 13 years, they find evidence that Activism, such as the Global Climate Strike on March 15, 2019, can have an impact on stock market prices. Ramelli et al. (2021) show that this event impacted negatively the market valuation of carbon-intensive firms and the financial analysts reconsidered their earnings expectations in the long run towards a lower level.

We believe that many of the criticisms of ESG empirical results have roots in methodological aspects. The different conclusions in the literature regarding the significance of carbon risk premia have been explained and reconciled in Lioui (2022) by employing an improved methodology that bypasses the problem of carbon measurement scaling. Many empirical studies rely on various ESG ratings but, as discussed in Larcker et al. (2022), there is a distinct lack of agreement of ESG ratings from different ESG ratings provides, see also Chatterji et al. (2016), Dimson et al. (2020), and Berg et al. (2022).

Yenipazarli and Vakharia (2015) argue that when introducing a new green product, it is paramount to think of consumer differentiation. The market should in general work with two prices for the green and brown versions of the product in order to maximize market development. They also show that charging green investors a premium because of their willingness-to-pay can lead to suboptimal strategies, even if the volume of trade is higher. This important idea is followed up in Quaye et al. (2024) who construct daily green-adjusted share prices for all constituent companies in the FTSE 1000 Green Revenues Index. They estimate the green-adjusted analogues to the CAPM beta and contrast the portfolio construction of a standard investor with that of a "twin" green investor who has the same dollar risk-return risk preferences but also holds high views on the necessity of green finance. Their study indicates that tilting stock returns towards climate finance could change temporarily asset pricing views, but overall the Fama-French risk factors between the two settings, standard and green, are highly correlated.

3. Data and methodology

3.1. Green revenues quantification

According to FTSE Russell, in 2022, the green economy was globally the fifth largest industry, similar in size to the fossil fuel sector. Despite the considerable diversity within the green economy, its concentration remains notable in a few key countries, particularly the United States (accounting for 54%) and China (12%). However, nations such as Japan, France, and Germany, while possessing smaller green economies in relative terms, exhibit a disproportionately high level of exposure to green activities. Using a comprehensive taxonomy that is very similar to the EU Taxonomy, the Green Revenues Data Model (GRDM) developed by FTSE Russell (part now of London Stock Exchange Group) estimates the net contribution of a company to the transition to a green economy by measuring the exposure to environmental (green) impact recorded on a company's balance sheet. FTSE Russell has applied the GRDM to an extended global dataset covering almost 99% of total global market capitalization, to estimate the net environmental impact of over 16,000 public companies across 48 developed and emerging markets.[1]

The FTSE 1000 Green Revenues Indexes reflect the green exposure that investors get by holding investments in the stock of those respective companies. Our data includes, in addition to indexes, the green revenue factors for all constituents of a given index. This is a rich database comprising cross-sectional and time series information on the Green Revenue Factors (GRF). Thus, the GRF becomes an essential yardstick that measures the level of engagement of a company over time vis-a-vis the climate change environmental agenda. The GRFs are interpreted in this paper as green/brown indicators of the net percentage of green activities. The main activities monitored by the FTSE Russsell (LSEG now) are climate change mitigation and adaptation, water, resource use, pollution, and agricultural efficiency. Table 1 presents the overview of green revenue factors for all countries covered under the green revenues database.

The GRF takes values between 0 and 2, with 2 representing a 100% green company (the ratio of green revenues to the total revenues is +1) and 0 representing the opposite, a totally brown company activity (the net ratio of green revenues to the total revenues is −1, i.e. there are no environmental benefits, but 100% damages). The mid-value of 1 is associated with a neutral level. The majority of companies in our study have GRFs close to neutral levels, indicating that the net position (green versus brown activities) of a company is close to zero.

Table 1. Summary table for green revenues factor.

Country	Mean	SD	Min	q_{25}	q_{50}	q_{75}	Max	Firm count
Australia	1.11	0.13	0.91	1.09	1.09	1.12	2.23	89
Belgium	1.13	0.11	1.08	1.09	1.09	1.11	1.47	12
Brazil	1.14	0.19	0.81	1.08	1.09	1.13	2.02	61
Canada	1.12	0.18	0.94	1.09	1.09	1.10	2.20	48
Chile	1.22	0.27	0.98	1.09	1.09	1.27	2.11	19
China	1.14	0.22	0.59	1.09	1.09	1.12	2.31	226
Colombia	1.11	0.05	1.08	1.08	1.09	1.10	1.22	8
Czech Republic	1.11	0.15	0.97	1.02	1.09	1.20	1.29	4
Denmark	1.19	0.28	1.01	1.09	1.09	1.13	2.15	18
Egypt	1.11	0.02	1.09	1.09	1.10	1.12	1.14	5
Finland	1.16	0.11	1.08	1.09	1.12	1.19	1.42	14
France	1.16	0.21	0.92	1.09	1.09	1.12	2.23	77
Germany	1.12	0.14	0.75	1.09	1.09	1.12	1.95	70
Greece	1.10	0.03	1.07	1.09	1.09	1.10	1.17	8
Hong Kong	1.12	0.17	0.91	1.09	1.09	1.09	2.17	79
Hungary	1.11	0.03	1.08	1.09	1.11	1.13	1.14	4
India	1.11	0.11	0.77	1.09	1.09	1.12	2.03	124
Indonesia	1.11	0.08	1.06	1.09	1.09	1.09	1.52	28
Ireland	1.10	0.03	1.07	1.08	1.09	1.11	1.13	4
Iceland	0.92	0.00	0.92	0.92	0.92	0.92	0.92	10
Israel	1.09	0.03	1.01	1.09	1.09	1.09	1.14	25
Italy	1.12	0.13	0.88	1.09	1.09	1.12	1.56	29
Japan	1.13	0.11	0.62	1.09	1.10	1.12	2.08	467
South Korea	1.11	0.08	0.75	1.09	1.09	1.12	1.57	121
Kuwait	0.94	0.02	0.91	0.93	0.94	0.94	0.98	10
Malaysia	1.12	0.07	1.03	1.09	1.09	1.12	1.43	39
Mexico	1.10	0.07	0.93	1.09	1.09	1.09	1.29	31
Netherlands	1.13	0.15	0.94	1.09	1.09	1.12	1.65	21
Norway	1.21	0.34	1.07	1.09	1.09	1.14	2.24	11
New Zealand	1.22	0.33	1.03	1.09	1.09	1.20	2.16	12
Austria	1.12	0.04	1.06	1.10	1.11	1.14	1.19	7
Pakistan	1.11	0.03	1.09	1.09	1.09	1.12	1.15	4
Peru	1.17	0.02	1.16	1.16	1.17	1.18	1.18	2
Philippines	1.13	0.12	1.02	1.09	1.09	1.12	1.58	23
Poland	1.09	0.04	1.04	1.09	1.09	1.09	1.17	13
Portugal	1.45	0.55	1.07	1.10	1.24	1.81	2.25	4
Qatar	1.11	0.14	0.99	1.09	1.10	1.10	1.56	17
Romania	0.95	0.00	0.95	0.95	0.95	0.95	0.95	1
Russia	1.13	0.09	1.00	1.11	1.11	1.12	1.46	29
South Africa	1.10	0.08	0.94	1.09	1.09	1.09	1.50	54
Saudi Arabia	0.84	0.04	0.69	0.84	0.84	0.84	0.90	43
Slovenia	1.11	0.06	1.03	1.09	1.09	1.11	1.33	29
Spain	1.19	0.24	1.01	1.09	1.09	1.14	2.12	24
Sweden	1.10	0.03	1.01	1.09	1.09	1.12	1.15	33
Switzerland	1.11	0.05	0.99	1.09	1.09	1.12	1.31	44
Thailand	1.17	0.21	0.97	1.09	1.09	1.13	2.03	34
Turkey	1.09	0.05	1.00	1.09	1.09	1.09	1.28	25
Taiwan	1.13	0.16	1.04	1.09	1.09	1.10	2.24	87
United Arab Emirates	1.08	0.05	0.92	1.09	1.09	1.09	1.10	13
United Kingdom	1.14	0.20	0.87	1.09	1.09	1.12	2.24	107
United States	1.12	0.13	0.81	1.09	1.09	1.12	2.29	514

Notes: This table reports the time series mean of the cross-sectional summary statistics of green revenue factors of the respective countries. The "firm count" column indicates the time series mean of the average number of firms per cross-section. The dataset spans 50 countries from May 26, 2016, to December 21, 2022, at a daily frequency, with different starting dates. Firms with over 40% missing data are excluded for each country.

The daily green-adjusted returns based on the green-adjusted share prices are computed as:

$$R^*_{i,t} = \ln\left(\frac{S^*_{i,t}}{S^*_{i,t-1}}\right) = \ln\left(\frac{S_{i,t}}{S_{i,t-1}}\right) + \ln\left(\frac{GRF_{i,t}}{GRF_{i,t-1}}\right) = R_{i,t} + \ln\left(\frac{GRF_{i,t}}{GRF_{i,t-1}}\right) \quad (1)$$

The green-adjustment translates into an additive tilting factor applied to the standard returns which correctly captures the level of engagement with the green agenda of a company over time. More specifically, an increase in the GRF factor over a period of time (more positive environmental impact), results in a reward in term of returns, while a decrease in the GRF (more negative environmental impact) leads to a penalty, with a green-adjusted return lower than the standard return. The green-adjusted prices and returns defined above together with the Green Revenues Indexes provide the basis for a new green-adjusted investment universe that investors and policy makers could explore further to better understand how the green efforts made by companies around the world are reflected in the financial markets.

3.2. Data description

The indexes we study are: the US FTSE Russell 1000, the UK All Share index, China, Europe, Japan, and the All-World index. The daily time series of the green revenues indexes and the GRF time series for the constituents of those indexes are downloaded from the FTSE Russell database. The overall sample period spans from 26 May 2016 to 21 December 2022, for a total of 1711 trading days. As dictated by data availability, the coverage in the FTSE Russell database begins at different times for different indexes and hence, the number of observations varies across indexes. For example, data for the FTSE Russell 1000 green index prices is available from 26 May 2016, while coverage for Japan begins later on 21 March 2017.

The number of firms utilized for each index varies across different economies by design, and within the same economy because of delisted companies. Table 2 shows that the US and China work with about 1000 firms on average while the UK, Europe and Japan with roughly 500–600 firms. The All-World Index has on average 3600 firms.

We present few summary statistics of the number of firms that are the constituents of the equity indexes in the five main economies and also in the All-World index in Table 3.

Regarding the GRF data, we compare the means of the cross-sectional averages across the six indexes and observe that there is a clear ordering, with the smallest value for the UK, followed by Japan, then the US, all lower than 1. It continues with Europe, followed by China and then All-World. Since a green revenue factor larger than 1 implies more green activities, one may infer that firms from China are on average greener than firms from Europe, for example. However, a more informed view can be obtained by looking at all quantiles listed in columns 5–7 of that table. Furthermore, the

Table 2. Summary statistics for constituents counts.

Index	Avg ^{firms}	Min ^{firms}	Max ^{firms}
UK	626	599	645
Russell 1000 (US)	1001	972	1031
Japan	508	493	520
Europe	663	634	689
China	1064	420	1774
All-World	3614	3060	4177

Notes: This table presents summary statistics for the number of constituents in each index over the entire sample period, from 26 May 2016 to 21 December 2022. We denote by Avg ^{firms}, Min ^{firms}, and Max ^{firms} the average, minimum, and maximum number of firms in each index. These summaries are calculated from daily time series data on the number of listed companies in each index over the sample period.

Table 3. Summary statistics for green revenue factor around the world.

Index	Mean	SD	Min	q25	q50	q75	Max
UK	0.951	0.153	0.519	0.925	0.925	0.927	1.957
Russell 1000 (US)	0.991	0.125	0.395	0.965	0.968	0.997	2.000
Japan	0.986	0.111	0.616	0.936	0.939	1.000	1.909
Europe	1.025	0.159	0.620	0.982	0.983	1.013	1.999
China	1.038	0.211	0.513	0.973	0.988	0.989	1.993
All-World	1.077	0.134	0.439	1.043	1.043	1.074	1.999

Notes: This table reports the time series averages per index of cross-sectional summary statistics for the green revenue factor of all index constituents. The cross-sectional summary statistics include Mean (Mean), Standard Deviation (SD), Minimum (Min), Maximum (Max), 25th quantile (q_{25}), 50th quantile (q_{25}), and 75th quantile (q_{75}) index constituents.

standard deviations indicate that actually Japan has the most consistent green revenue factors and, together with Europe have the largest minima. The lowest minimum GRF and the largest maximum GRF are for the US.

It would be also useful to see the evolution over time of the green revenues valuations measured by GRF. The GRFs have different distributions and evolutions across the economies compared in this study. This can be observed from the graphs in Figure 1 that illustrate the monthly time series of cross-sectional averages of GRFs for companies from all five economies and All-World. The confidence intervals are computed based on the cross-sectional 2.5% and 97.5% cross-sectional quantiles of GRFs from the respective economies.

The average and most of the GRF quantities for US are below 1, indicating that between 2016 and 2022 the US was not highly geared towards climate change adjustments. This is perhaps not surprising given that this period coincided with the Trump administration, which formally withdraw from the Paris climate agreement in June 2017. Although there seems to have been a short-lived recovery towards positive green adjustments in early 2020, the COVID period pushed back the GRFs below 1 in the US.

Interestingly, the GRFs for Japan are not as high as expected perhaps with values in the neutral territory until COVID-19 eruption when there is a clear downward shift. By contrast, for the Chinese

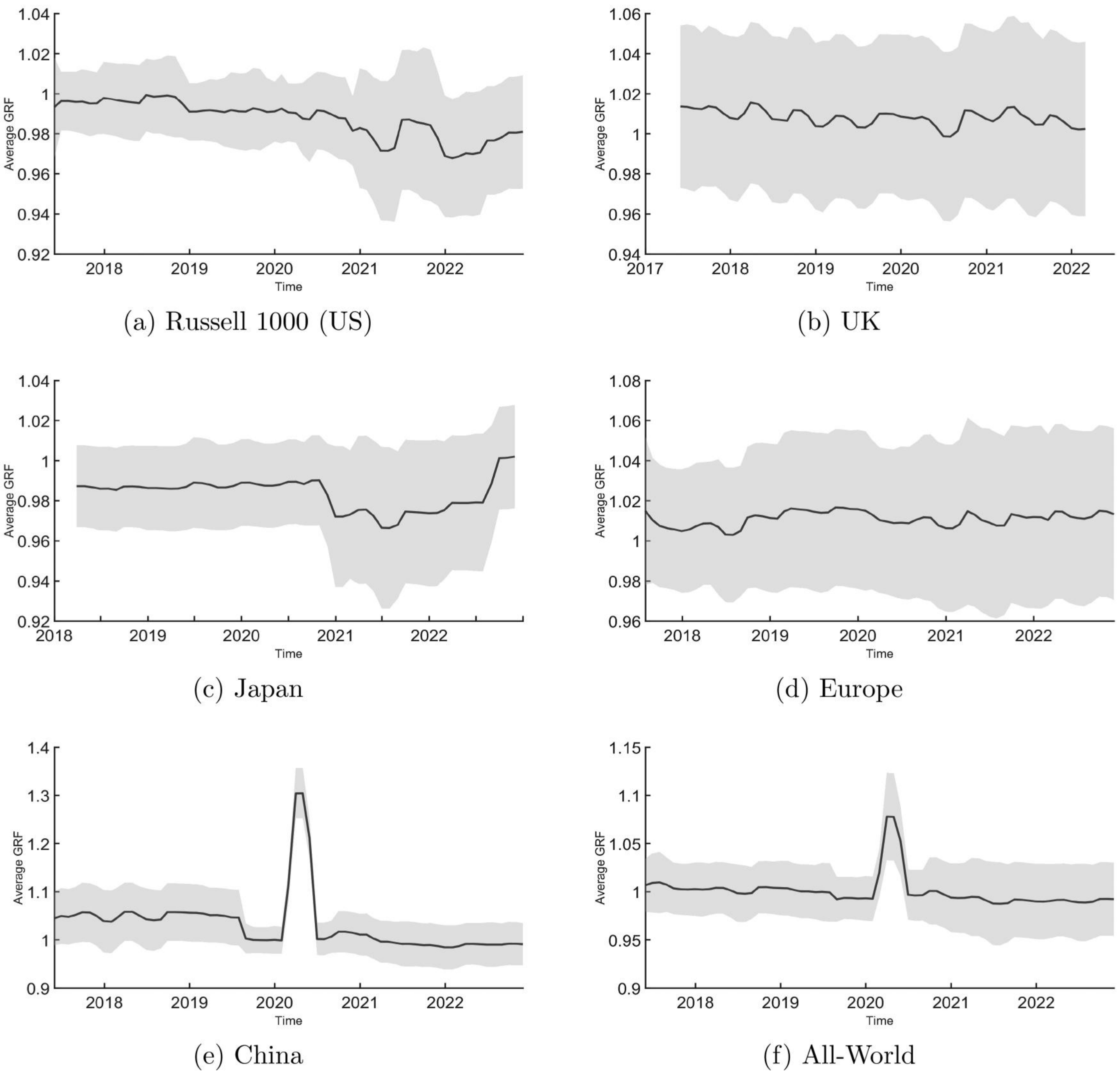

Figure 1. Time series of cross-sectional average GRF. Notes: For each index, at the end of each month, we first calculate the time series average of GRF for each constituent stock using daily data from the past 1-year including the date of calculation. We then compute the cross-sectional average of the values obtained in the first step and plot them as monthly time series along with its 95% confidence interval depicted by the shaded area.

companies that are the constituents of the FTSE Russell China Green Revenues, the GRFs are quite positive, with almost the entire distribution above 1. The average GRF scores for the UK have been consistently above 1, with a quite tight confidence interval for GRF values, roughly between 0.98 and 1.04. The GRF evolution for companies in Europe is very similar to the UK, with a cross-sectional average above 1 and the lower confidence boundary roughly at 0.98. It is worth pointing out that the COVID-19 pandemic period impacted the green revenues assessments for companies in the US and Japan, increasing in those economies the uncertainty about green activities. China experienced a positive shock in the GRF factor in 2020 followed by a more downward trend afterwards.

Perhaps surprisingly, UK and Europe were almost undeterred by the COVID shock, although for UK a slight downward trend for GRFs can also be noticed over the entire study period.

3.3. Methods

3.3.1. Multiple correlation coefficients

Crisci (2023) surveys methodologically several measures for the identification of the strength of association between a response variable and covariates and she applies generalized estimating equations to gauge the impact of governance factors on environmental policy disclosure. Her research emphasizes the important pitfalls around estimation when working with multivariate data. The proposed solution based on estimating equations requires only a mean-covariance specification and not the entire distribution.

In our paper, we take a look at the multivariate correlation coefficients that capture more than the pairwise correlation relationships. The groupings interactions or associations are particularly relevant for portfolio construction. For groups that are highly correlated the information can be used for cross-hedging exercises that are routinely executed in financial markets. Identifying groups that are not highly correlated helps with diversification.

Wang and Zheng (2020) generalized the Pearson coefficient of linear correlation to multiple variables. First, let us consider the case of correlation (and uncorrelation) for one extra variable, so that correlation measures are defined for a joint triplet of variables. If X, Y, Z are three random variables and ρ_{xy}, ρ_{xz} and ρ_{yz} are the Pearson linear correlation coefficients for the respective pairs of variables then the *triple correlation coefficient* among X, Y and Z is defined by

$$\rho^2_{XYZ} = \rho^2_{xy} + \rho^2_{xz} + \rho^2_{yz} - 2\rho_{xy}\rho_{xz}\rho_{yz} \tag{2}$$

This correlation measure has properties that are very similar to the Pearson correlation coefficient. The following properties are proved by Wang and Zheng (2020). For any three random variables X, Y, Z with non-vanishing variance, the triple correlation coefficient satisfies that $0 \le \rho^2_{XYZ} \le 1$; $\rho^2_{XYZ} = 1$ if and only if the sample values are linearly dependent; and $\rho^2_{XYZ} = 0$ if and only if the variables X, Y, Z are mutually uncorrelated.

For the multiple correlation coefficient defined in (2), the multiple uncorrelation coefficient is in a sense a dual measure to the multiple correlation coefficient that can be defined as

$$\psi^2_{XYZ} = 1 - \rho^2_{XYZ}. \tag{3}$$

The generalization presented above can be continued to an arbitrary dimension of random vector. Consider now a set of variables $X_1, X_2, \dots X_d$ and let us denote by R the correlation matrix constructed from pairwise Pearson correlation coefficients $\rho_{x_i x_j}$ for the countries i and j.

$$R = \begin{pmatrix} 1 & \rho_{x_1x_2} & \cdots & \rho_{x_1x_d} \\ \rho_{x_1x_2} & 1 & \cdots & \rho_{x_2x_d} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{x_dx_1} & \rho_{x_dx_2} & \cdots & 1 \end{pmatrix} \tag{4}$$

The d-multiple uncorrelation coefficient (MUC) is defined by Wang and Zheng (2020) as

$$\psi^2_{x_1x_2\dots x_d} = det(R) \tag{5}$$

and then it follows by complementarity that the multiple correlation coefficient (MCC) is

$$\rho^2_{x_1x_2\dots x_d} = 1 - \psi^2_{x_1x_2\dots x_d} \tag{6}$$

Similar to the 3-dimensional case, there are similar properties for the more general d-dimensional case, which are discussed in detail in Wang and Zheng (2020). In this paper, we compute multiple correlation and uncorrelation squared coefficients up to order 5.

3.3.2. Beta estimation

For a better understanding of the impact caused by the green revenues adjustment or tilting of the market share price for companies in main economies, we also consider any changes that may appear in conventional asset pricing exercises. The unconditional beta is defined directly from the capital asset pricing model (CAPM) as

$$\beta_i = cov(R_i, R_m)/var(R_m) \tag{7}$$

where R_i, R_m are the share price return of company i and the return of market portfolio, respectively. The standard CAPM beta is estimated using the historical series of returns with the regression:

$$r_{i,t} = \alpha_i + \beta_i r_{m,t} + \varepsilon_{i,t} \tag{8}$$

where $r_{i,t}, r_{m,t}$ are the excess return of company i and of the market portfolio over the risk-free rate, respectively, at time t. This leads to the most common beta estimate $\beta_i^{HIST} = \hat{\beta}_i$. Typically, the regression model is estimated based on a one-year rolling window of daily excess return data. That is, at the end of each month t, the regression model is estimated based on the previous 12-month period of daily return data, covering months $t-11$ through t, inclusively. Andersen et al. (2006) define the realised beta as

$$\beta^R_{i,\tau} = \frac{\sum_{t=1}^{t=N} r_{i,t} r_{m,t}}{\sum_{t=1}^{t=N} r^2_{m,t}} \tag{9}$$

where N is the number of observations during the estimation window τ. It is known, see Andersen et al. (2006), that under weak regularity conditions this is the only consistent measure for the true beta.

The green revenue factor allows a direct transformation of the market share price of a firm into a green-revenue adjusted dollar price. We take advantage of being able to generate the tilted green revenues adjusted share prices and compute the corresponding CAPM green beta and the realised green beta for all firms that are the constituents of our green revenues indexes.

3.4. Time stochastic dominance test for green minus brown strategy

The GRFs permit us a quintile tranching based on time series average GRFs for the respective countries. Those portfolios can be compared using static measures such as mean returns or betas. It is of great academic and practical interest to compare also the high minus low portfolios for each economy. At the end of each month t, we sort stocks by their green revenue factor (GRF) and divide them into five quintiles: $1(Brown), 2, 3, 4, 5(Green)$, from lowest to highest. Returns in each quintile are weighted by market capitalisation at time t. We form a zero-cost long-short strategy with a long position in the highest GRF portfolio $5(Green)$ and a short position in the lowest GRF portfolio $1(Brown)$. This method is applied to UK, US, Japan, Europe, and China. If these strategies pass the statistical test then we can conjecture that there is a possible green revenues adjustment factor that can be computed and utilized based on the GRF.

We use the nonparametric time stochastic dominance (TSD) test discussed in Lee et al. (2023) to evaluate the dynamic performance of the green minus brown (long-short) strategy in country pairs. Testing long-short portfolio returns helps investors assess environmental sustainability and green tilting benefits by measuring and managing systematic differences across countries. The TSD test partially orders long-short portfolio strategies over time based on expected net present value criteria for general utility and discount functions of investors in these countries. We let $\mathcal{Y}_i := \{Y_{i,t} : t \in \tau\}$ and $\mathcal{Y}_j := \{Y_{j,t} : t \in \tau\}$ represent the value-weighted [2] returns from a high minus low quintile strategy for country pairs i and j, where $(i,j) \in \{\text{UK}, \text{US}, \text{Japan}, \text{Europe}, \text{China}\}$.

The time-path of value-weighted returns have a common support $Y = [\underline{y}, \overline{y}]$ for all $t \in \tau$ and $f_k(\cdot, t)$ and $F_k(\cdot, t) := \int_{\underline{y}}^{\cdot} f_k(z, t) dz$ denote the density and distribution functions of $\mathcal{Y}_i$ and $\mathcal{Y}_j$ for $t \in \tau$. The utility and time-discount functions are respectively denoted by $u : Y \mapsto \mathbb{R}$ and $v : \tau \mapsto \mathbb{R}$ both of which are continuously differentiable. The expected discounted utility of prospect $Y_{(i)}$ and $Y_{(j)}$ at time $t = 0$ is given by

$$NPV_{v,u}(\mathcal{Y}_i) = \sum_{t=0}^{T} v(t) \mathbf{E}_{F_i(\cdot,t)} u(Y_{i,t}) = \sum_{t=0}^{T} v(t) \left(\int_Y u(y) f_i(y,t) dy \right)$$

$$NPV_{v,u}(\mathcal{Y}_j) = \sum_{t=0}^{T} v(t) \mathbf{E}_{F_j(\cdot,t)} u(Y_{j,t}) = \sum_{t=0}^{T} v(t) \left(\int_Y u(y) f_j(y,t) dy \right)$$

which depends on the utility function. Investors are assumed to rank outcomes rationally according to the values of $NPV_{v,u}(Y_i)$ and $NPV_{v,u}(Y_j)$.

Utility functions are from nested classes: $\mathcal{U}_1$, a class of monotonically increasing functions, and $\mathcal{U}_2$, a concave CRRA class for risk aversion. They are defined as follows: $\mathcal{U}_1 = u : u^{(1)}(y) \geq 0$ and $\mathcal{U}_2 = u : u \in \mathcal{U}_1, u^{(2)}(y) \leq 0$, where $u^{(1)}$ and $u^{(2)}$ represent the first and second derivatives of the utility function, respectively. A recursive definition is employed for higher-order utility function classes $\mathcal{U}_m = u : u \in \mathcal{U}_{m-1}, (-1)^m u^{(m)} \leq 0$ for $m \geq 2$.

Discount functions can be classified into three categories: strictly positive $\mathcal{V}_0$ (indicating a positive degree of time preference), strictly decreasing $\mathcal{V}_1$ (indicating increasing impatience over time), or strictly decreasing and convex $\mathcal{V}_2$ (indicating decreasing impatience over time). The definitions are as follows: $\mathcal{V}_0 = v : v(t) > 0$, $\mathcal{V}_1 = v : v \in \mathcal{V}_0, v^{(1)}(t) < 0$, and $\mathcal{V}_2 = v : v \in \mathcal{V}_1, v^{(2)} > 0$. To apply the hypothesis testing method, a definition is required to relate unobservable utility-based comparisons to observable distribution comparisons. Accordingly, an investment decision $\mathcal{Y}_i$ is said to first-order time and first-order stochastically dominate $\mathcal{Y}_j$, denoted $\mathcal{Y}_i \succcurlyeq_{1T1SD} \mathcal{Y}_j$, if and only if

a. $NPV_{v,u}(\mathcal{Y}_i) - NPV_{v,u}(\mathcal{Y}_j) \geq 0, \forall (v,u) \in \mathcal{V}_1 \times \mathcal{U}_2$, or

b. $D^{(1,1)}(y,t) \leq 0, \forall (y,t) \in Y \times \tau$

where $D^{(1,1)}(y,t) = F_i^{(1,1)}(y,t) - F_j^{(1,1)}(y,t)$ and $F_k^{(1,1)}(y,t) = \sum_{s=0}^{t} F_k(y,s) = \sum_{s=0}^{t} \int_{\underline{y}}^{\overline{y}} f_k(z,s) dz, k = 1,2$ (see Dietz and Matei, 2016).

A generalisation of this definition states that, investment decision $\mathcal{Y}_i$ n-th order time and m-th order stochastic dominates $\mathcal{Y}_j$, if and only if

a. $NPV_{v,u}(\mathcal{Y}_i) - NPV_{v,u}(\mathcal{Y}_j), \forall (v,u) \in \mathcal{V}_n \times \mathcal{U}_m$, *or*
b. For $a = 0, \ldots, n-1$, and $b = 1, \ldots, m-1$, (i) $D^{(a+1,b+1)}(\overline{y}, t) \leq 0$, (ii) $D^{(n,b+1)}(\overline{x}, t) \leq 0$, $\forall t \in \tau$, (iii) $D^{(a+1,m)}(y, T) \leq 0$, $\forall y \in Y$, and (iv) $D^{(n,m)}(y,t) \in Y \times \tau$

$n, m = 1, 2$ and u and v, respectively, denote the utility and time-discount functions both assumed to be continuously differentiable. $i \neq j$. The null hypothesis of the n-order time and m-order stochastic dominance is given by

$$H_0^{(n,m)}: NPV_{v,u}(\mathcal{Y}_i) \succcurlyeq NPV_{v,u}(\mathcal{Y}_j), \ \forall (v,u) \in \mathcal{V}_n \times \mathcal{U}_m \tag{10}$$

equivalent to F_i and F_j satisfying first order TSD or higher order TSD.

For example, if we fail to reject $H_0^{(2,2)}$ then all risk averse investors who discount with decreasing and convex discount functions would prefer investment $\mathcal{Y}_i$ to investment $\mathcal{Y}_j$. If we reject $H_0^{(n,m)}$ it means that there exists at least one investor with $(v,u) \in \mathcal{V}_n \times \mathcal{U}_m$ who ranks the prospects as equivalent. More specifically, a rejection of the null, suggest that the NPV of the long-short strategy in country i is more or less the same as that in country j. Likewise, if we fail to reject $H_0^{(1,1)}: \mathcal{Y}_i \succcurlyeq \mathcal{Y}_j$, we conclude that a given investor or economic agent having increasing utility and decreasing discount function would assign higher NPV to long-short strategy in country i than long-short strategy in country j.

According to Lee et al. (2023), the test statistic is a one-sided L_p-type test statistic written as

$$T_N = r_N^p \int_{\mathcal{X}} \Lambda_p(\hat{v}_1(x), ..., \hat{v}_L(x)) \qquad (11)$$

where $r_N := \sqrt{\frac{N_1 \cdot N_2}{N_1+N_2}}$. For critical value calculations, Lee et al. (2023) suggested two alternatives: the contact set method and the numerical delta method. The contact set method improves over the conservative least favorable case (LFC), directly imitating the limiting distribution under the null hypothesis without computing the LFC-based critical value when the asymptotic distribution degenerates to zero. The numerical delta method is based on Fang and Santos (2019), Hong and Li (2018), and Dümbgen (1993).

The green revenues factors can be used as a yardstick to classify firms into green, neutral and brown. The difference between green firms and brown firms represents a wedge that may be used to construct stock portfolios that would help to greenify economies over time. If such a portfolio is profitable in dollars then that implies the respective economy still needs to do better to improve their climate agenda credentials. In a green perfect world, all firms would gave maximum green revenues factors and there will not be a wedge. We will construct this green minus brown type of portfolios for all economies investigated and we will pairwise test their portfolio dollar performance.

4. Empirical results

Returns on individual equities are winsorised at 1% and 99%. We require equities to have at least 200 non-missing returns when estimating the 1-year horizon beta and at least 15 non-missing returns in the 1-month horizon beta.

4.1. Multidimensional correlations of green adjusted equity indexes returns

We calculate the daily logarithmic returns for all firms time series, denoted $R_{i,t+1} = \ln \frac{X_{i,t+1}}{X_{i,t}}$ for each i-th index X. The summary statistics of the green index return series reported in Table 4 indicate a predominantly positive mean daily return for the sampled indexes, with the only exception of UK and China. The distribution of returns as depicted by the 25th-, 50th-, and 75th-quantiles does not vary substantially across different indexes, generally being within the same order of magnitude. China reported the highest standard deviation, which is not surprising given the severe adverse impact of the pandemic.

The results in Panel B of the Table 4 suggest that the revenues from the green revenues adjusted equity indexes in some economies are much more interdependent than other pairings. Despite its universally positive sign, sample correlation $\rho(\cdot,\cdot)$ varies substantially across the different pairs of green index returns, ranging from a minimum of 19.7% for $\rho(Japan, US)$, to a maximum of 94.9% $\rho(All\ World, US)$. Respectively, the correlation coefficient between Japan and other indexes is consistently below 39% and that between China and other indexes is 50%. The largest pairwise Pearson correlation coefficients are for US and All-World, at almost 95% and between UK and Europe, at 94%.

The former relationship is perhaps not surprising given the economic dominance of the US economy on economies in other parts of the world. However, the latter strong connection upon green tilting of firms' in the UK and Europe was perhaps expected

Table 4. Summary statistics of green Revenues Adjusted Equity Index returns.

	All-World	UK All Share	China	Europe	Japan	Russell 1000 (US)
			Panel A: Descriptive statistics			
N (days)	1700	1625	1492	1702	1468	1710
Mean(%)	0.023	−0.001	−0.007	0.006	0.003	0.036
SD(%)	1.080	1.329	1.589	1.266	1.125	1.355
Min(%)	−9.963	−13.707	−7.956	−14.056	−6.539	−12.985
Max(%)	7.950	10.885	12.641	8.499	6.937	9.039
q_{25} (%)	−0.371	−0.561	−0.818	−0.500	−0.584	−0.433
q_{50} (%)	0.072	0.080	0.038	0.074	0.031	0.057
q_{75} (%)	0.507	0.595	0.833	0.595	0.622	0.672
			Panel B: Unconditional correlation			
All-World	1.000	0.767	0.497	0.793	0.383	0.949
UK All Share	0.767	1.000	0.419	0.941	0.359	0.594
China	0.497	0.419	1.000	0.431	0.342	0.345
Europe	0.793	0.941	0.431	1.000	0.367	0.612
Japan	0.383	0.359	0.342	0.367	1.000	0.197
Russell 1000 (US)	0.949	0.594	0.345	0.612	0.197	1.000

Notes: This Table reports the averages of cross-sectional summary statistics for the green index return from 26 May 2016 to 21 December 2022 in Panel A. The cross-sectional summary statistics include Mean (Mean), Standard Deviation (SD), Minimum (Min), Maximum (Max), Median (Median), 25th quantile (q_{25}), 75th quantile (q_{75}), and the available number of sample points (N) for each index return time series. The unconditional sample correlation of index returns between the All-World Index, UK All Share Index, China Index, Europe Index, Japan Index, and FTSE Russell1000 indexes are reported in Panel B.

Table 5. Multiple uncorrelation and correlation coefficients for Green Revenues Indexes around the world.

Groupings	Ψ^2	ρ^2	Groupings	Ψ^2	ρ^2
All, UK, China	0.3085	0.6914	All, UK, China, Europe	0.0003	0.9997
All, UK, Europe	0.0420	0.9579	All, UK, China, Japan	0.2514	0.7486
All, UK, Japan	0.3470	0.6529	All, UK, China, US	0.0119	0.9881
All, UK, US	0.0994	0.9006	All, UK, Europe, Japan	0.0354	0.9646
All, China, Europe	0.2781	0.7218	All, UK, Europe, US	0.0019	0.9981
All, China, Japan	0.6195	0.3804	All, UK, Japan, US	0.0107	0.9893
All, China, US	0.0588	0.9411	All, China, Europe, Japan	0.2265	0.7735
All, Europe, Japan	0.0571	0.9428	All, China, Europe, US	0.0081	0.9919
All, Europe, US	0.0171	0.9828	All, China, Japan, US	0.0334	0.9666
All, Japan, US	0.0571	0.9428	All, Europe, Japan, US	0.0069	0.9931
UK, China, Europe	0.0930	0.9069	UK, China, Europe, Japan	0.6815	0.3185
UK, China, Japan	0.6814	0.3185	UK, China, Europe, US	0.0572	0.9428
UK, China, US	0.4123	0.5876	UK, China, Japan, US	0.4322	0.5678
UK, Europe, Japan	0.0989	0.9010	UK, Europe, Japan, US	0.0615	0.9385
UK, Europe, US	0.6254	0.3745	China, Europe, Japan, US	0.7717	0.2283
UK, Japan, US	0.5634	0.4365	All, UK, China, Europe, Japan	0.0256	0.9744
China, Europe, Japan	0.6707	0.3292	All, UK, China, Europe, US	0.0009	0.9991
China, Europe, US	0.5026	0.4973	All, UK, China, Japan, US	0.0050	0.9950
China, Japan, US	0.7716	0.2283	All, China, Europe, Japan, US	0.0026	0.9974
Europe, Japan, US	0.5404	0.4595	UK, China, Europe, Japan, US	0.0469	0.9531

Note: This table presents multiple uncorrelation and correlation square coefficients for the return time series of Green Revenues Indexes of All-World, UK, China, Europe, Japan and US. The sample period is from May 2016 and ending December 2022.

in sign but not so much in magnitude. It should also be noted the relative low correlation coefficients for Japan with all the other economies and also for China with all other economies. These results point out to different roles played worldwide by different economies, after adjusting firms' share prices for green activities.

The multidimensional correlation coefficients offer a more insightful view of interactions between different groupings of economies. The classical pair-wise correlation coefficients of return time series of green revenues adjusted equity indexes of the five major economies and All-World as well, were presented in Table 4 in Panel B. This normalised correlation matrix is also all that is needed to compute the multidimensional uncorrelation and correlation coefficients in formulae (5) and (6).

In Table 5, we report the multidimensional correlation coefficients for all possible 3-dimensional, 4-dimensional and 5-dimensional groupings. At the 3-dimensional level the strongest squared correlation coefficient is observed for ρ^2 (All-World, Europe, US) at 98.28% followed closely by ρ^2 (All-World, UK, Europe) at 95.79%. The lowest squared 3-dimensional correlation coefficient is for ρ^2 (China, Japan, US) at 22.83% followed by ρ^2 (UK, China, Japan) at 31.85%.

It should be also noted that higher correlation coefficients do not imply higher green values for share prices of firms from those respective economies. It rather depicts the perception of investors of those companies as very similar, even after adjusting for green revenues. The 4-dimensional and 5-dimensional squared correlation coefficients reveal an interesting view. There are low coefficients for ρ^2 (China, Europe, Japan, US) at 22.83% and ρ^2 (UK, China, Europe, Japan) at 31.85% but combining those two groups into ρ^2 (UK, China, Europe, Japan, US) gives a coefficient of 95.31%. This points out that an investor that has strong green risk preferences may need to have a portfolio diversified in such a way to include firms from all five economies, in order to capture high order of interaction post green revenues adjustments. One possible explanation for these clear discrepancies may be related to cultural differences. It has been observed in the literature recently, see Auzepy et al. (2023), that sustainability-linked loan borrowers are perceived quite differently in the EU and the US.

4.2. Portfolio sorts

We test whether GRF has an effect on realised stock returns. At the end of each month, we sort stocks in ascending order using their GRFs. We form quintile portfolios so that stocks with the lowest GRFs are assigned to quintile 1 and those with the highest GRFs are assigned to quintile 5. Based on the sorting outcome, we implement a trading strategy that takes a goes long on stocks with the highest GRF (quintile 5) and shorts stocks with the lowest GRFs (quintile 1). On this basis, we can attribute differences in average returns to differences inherited from the spread in the GRF variable.

In Table 6, we report the results for the quintiles portfolios formed on the basis of the GRF for each economy. The first two panels present the time series average of each quintile portfolio. Upon the sorting, Japan has the largest betas compared like for like with the other economies (except for the lowest quintile for which US is the highest), whilst China has the lowest betas. One can also note the quintiles portfolios for the US and UK give similar betas. Panel C of the same table shows the average values of the GRF for the respective quintile portfolios and economies. For the lowest quintile, the lowest

Table 6. Quintiles portfolios sorted on the green revenue factor.

Portfolio	UK	Russell 1000 (US)	Japan	Europe	China	All-World
		Panel A: CAPM Beta				
1(Low)	0.746	0.856	0.797	0.613	0.536	0.628
2	0.821	0.868	0.879	0.698	0.478	0.532
3	0.853	0.821	0.889	0.646	0.441	0.570
4	0.783	0.894	0.986	0.735	0.396	0.707
5(High)	0.851	0.866	0.942	0.717	0.439	0.610
		Panel B: Realised Beta				
1(Low)	0.743	0.851	0.795	0.610	0.533	0.623
2	0.819	0.865	0.878	0.697	0.475	0.529
3	0.853	0.819	0.888	0.645	0.439	0.569
4	0.782	0.893	0.985	0.733	0.393	0.705
5(High)	0.849	0.864	0.941	0.715	0.437	0.607
		Panel C: GRF				
1(Low)	0.955	0.948	0.919	0.953	0.946	0.947
2	0.977	0.965	0.930	0.966	0.977	0.962
3	0.978	0.967	0.948	0.967	0.986	0.966
4	0.981	0.981	1.001	0.991	0.990	0.993
5(High)	1.139	1.083	1.126	1.174	1.280	1.138
		Panel D: Mean Return				
1(Low)	−0.018	−0.049	−0.099	−0.048	−0.064	−0.057
2	−0.032	−0.007	−0.130	−0.025	0.007	−0.026
3	−0.019	0.014	−0.089	−0.005	0.019	0.008
4	0.014	0.014	−0.108	−0.043	0.034	−0.018
5(High)	−0.013	0.033	−0.127	−0.033	−0.010	−0.029

Note: At the end of each month, we sort stocks in each index into 5 annualized value-weighted portfolios according to their green revenue factor (GRF). We report the time series average of each quintile portfolio's value-weighted average CAPM beta, Realised beta, and the GRF used for sorting. Mean return denotes the annualized average portfolio excess return. The period of analysis starts from May 2016 and ends December 2022.

average GRF is for Japan and the highest is for UK, followed closely by Europe. For the highest quintile portfolios, the lowest average GRF is for US at 1.083 while the largest is for China at 1.280. Europe is second largest at 1.174, followed by UK with 1.139 and Japan with 1.126. Therefore, the greenest portfolio can be constructed with companies from China, while the less green would be for Japan. US quintile portfolios have average GRFs varying between 0.948 and 1.083, indicating that most companies from the US basket are classified more or less as less green or net green neutral. Japan is the only economy with two quintiles (4th and 5th) with average GRF larger than 1.

Based on the above, it is very interesting to see the realised returns performance for those quintiles portfolios. For Japan and Europe, all quintiles have negative average returns, hence a high minus low strategy would generate a positive performance for Europe and negative performance for Japan. China has negative average returns for both low and high quintiles and positive average returns for all the other three middle quintiles. A high-minus-low strategy will also give positive returns for China. UK has negative average returns for all quintiles except the 4th and a high minus low strategy would also generate positive returns. US is perhaps the best-structured tranche of quintile portfolios, with negative average returns for the first two quintiles and positive average returns for the last three quintiles.

Table 7. Time stochastic dominance (TSD) test for high minus low strategy.

Test	LFC (1,1)	Contact (1,1)	NDM (1,1)
$\text{NPV}_{u,v}(\mathcal{Y}_{UK}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{US})$	0.010	0.000	0.005
$\text{NPV}_{u,v}(\mathcal{Y}_{US}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{UK})$	0.350	0.265	0.220
$\text{NPV}_{u,v}(\mathcal{Y}_{UK}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{JPN})$	0.575	0.575	0.465
$\text{NPV}_{u,v}(\mathcal{Y}_{JPN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{UK})$	0.120	0.120	0.105
$\text{NPV}_{u,v}(\mathcal{Y}_{UK}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{EUR})$	0.360	0.360	0.330
$\text{NPV}_{u,v}(\mathcal{Y}_{EUR}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{UK})$	0.64	0.640	0.59
$\text{NPV}_{u,v}(\mathcal{Y}_{UK}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{CHN})$	0.010	0.000	0.005
$\text{NPV}_{u,v}(\mathcal{Y}_{CHN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{UK})$	0.115	0.045	0.065
$\text{NPV}_{u,v}(\mathcal{Y}_{US}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{JPN})$	0.835	0.825	0.775
$\text{NPV}_{u,v}(\mathcal{Y}_{JPN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{US})$	0.030	0.010	0.030
$\text{NPV}_{u,v}(\mathcal{Y}_{US}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{EUR})$	0.295	0.285	0.195
$\text{NPV}_{u,v}(\mathcal{Y}_{EUR}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{US})$	0.010	0.010	0.010
$\text{NPV}_{u,v}(\mathcal{Y}_{US}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{CHN})$	0.680	0.680	0.625
$\text{NPV}_{u,v}(\mathcal{Y}_{CHN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{US})$	0.340	0.340	0.315
$\text{NPV}_{u,v}(\mathcal{Y}_{JPN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{EUR})$	0.155	0.155	0.140
$\text{NPV}_{u,v}(\mathcal{Y}_{EUR}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{JPN})$	0.840	0.840	0.795
$\text{NPV}_{u,v}(\mathcal{Y}_{JPN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{CHN})$	0.030	0.020	0.030
$\text{NPV}_{u,v}(\mathcal{Y}_{CHN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{JPN})$	0.375	0.365	0.280
$\text{NPV}_{u,v}(\mathcal{Y}_{EUR}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{CHN})$	0.015	0.015	0.015
$\text{NPV}_{u,v}(\mathcal{Y}_{CHN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{EUR})$	0.075	0.045	0.040

Note: This Table reports the *p*-values of the Lee et al. (2023) time stochastic dominance (TSD) test using the null hypothesis specified in Equation (10). The null hypothesis states that the expected discounted utility of the strategy in country *i*, $\text{NPV}_{u,v}(\mathcal{Y}_i)$, first-order time and first-order stochastic dominates that of *j*, $\text{NPV}_{u,v}(\mathcal{Y}_j)$; thus, $n = 1$ and $m = 1$. The reported *p*-values are those obtained from the LFC algorithm, Contact-set approach, and numerical delta method (NDM).

A high minus-low strategy would give the largest positive returns out of all economies.

4.3. Time stochastic dominance test results

Table 7 shows the results for time stochastic dominance for all possible pairings of economies. The All-World economy was left out of this analysis, since there is no clear regulatory, legal, and economic jurisdiction. For robustness, we present three different methods for calculation of *p*-values, i.e. the least favorable case (LFC), contact set (Contact) algorithm, and the numerical delta method (NDM). Rejecting the null hypotheses in the first-time and first-order TSD test means that the NPV of long-short strategies implemented for countries on the right-hand side of ($\succcurlyeq$) is more or less the same as that of countries indicated on its left-hand side. For example, rejecting $H_0^{(1,1)}$: $\text{NPV}_{u,v}(\mathcal{Y}_{UK}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{EUR})$ means that the NPV of the long-short strategy implemented in the EUR is more or less the same as that of the UK. When the null hypotheses are not rejected, it implies that the results show that NPV of long-short strategies in the first country, first-time and first-order stochastic dominates that of the second country. Take for example, $H_0^{(1,1)}$: $\text{NPV}_{u,v}(\mathcal{Y}_{UK}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{US})$ and $H_0^{(1,1)}$: $\text{NPV}_{u,v}(\mathcal{Y}_{US}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{UK})$ with strong evidence against the null (see *p*-values in the first and second row of Table 7). This indicates that an investor with green-preference, having increasing utility and decreasing discounting function would assign high NPV to long-short strategies in the US

Table 8. Time stochastic dominance (TSD) test for high minus low strategy.

Test	LFC	Contact	NDM
$\text{NPV}_{u,v}(\mathcal{Y}_{UK}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{US})$	0.135	0.030	0.130
$\text{NPV}_{u,v}(\mathcal{Y}_{US}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{UK})$	0.400	0.330	0.345
$\text{NPV}_{u,v}(\mathcal{Y}_{UK}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{JPN})$	1.000	1.000	1.000
$\text{NPV}_{u,v}(\mathcal{Y}_{JPN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{UK})$	0.095	0.095	0.095
$\text{NPV}_{u,v}(\mathcal{Y}_{UK}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{EUR})$	0.405	0.405	0.395
$\text{NPV}_{u,v}(\mathcal{Y}_{EUR}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{UK})$	0.660	0.660	0.645
$\text{NPV}_{u,v}(\mathcal{Y}_{UK}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{CHN})$	0.285	0.280	0.270
$\text{NPV}_{u,v}(\mathcal{Y}_{CHN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{UK})$	0.355	0.355	0.340
$\text{NPV}_{u,v}(\mathcal{Y}_{US}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{JPN})$	0.625	0.560	0.555
$\text{NPV}_{u,v}(\mathcal{Y}_{JPN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{US})$	0.050	0.000	0.050
$\text{NPV}_{u,v}(\mathcal{Y}_{US}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{EUR})$	0.360	0.360	0.335
$\text{NPV}_{u,v}(\mathcal{Y}_{EUR}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{US})$	0.195	0.110	0.185
$\text{NPV}_{u,v}(\mathcal{Y}_{US}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{CHN})$	1.000	1.000	1.000
$\text{NPV}_{u,v}(\mathcal{Y}_{CHN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{US})$	0.235	0.235	0.230
$\text{NPV}_{u,v}(\mathcal{Y}_{JPN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{EUR})$	0.085	0.085	0.085
$\text{NPV}_{u,v}(\mathcal{Y}_{EUR}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{JPN})$	1.000	1.000	1.000
$\text{NPV}_{u,v}(\mathcal{Y}_{JPN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{CHN})$	0.175	0.175	0.175
$\text{NPV}_{u,v}(\mathcal{Y}_{CHN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{JPN})$	0.425	0.415	0.375
$\text{NPV}_{u,v}(\mathcal{Y}_{EUR}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{CHN})$	0.355	0.340	0.310
$\text{NPV}_{u,v}(\mathcal{Y}_{CHN}) \succcurlyeq \text{NPV}_{u,v}(\mathcal{Y}_{EUR})$	0.315	0.315	0.29

Notes: This Table reports the *p*-values of the Lee et al. (2023) time stochastic dominance (TSD) test using the null hypothesis specified in Equation (10). The null hypothesis states that the expected discounted utility of the strategy in *i*, $\text{NPV}_{u,v}(\mathcal{Y}_i)$, $n = m = 2$, second-order time and second-order stochastic dominates that of *j*, $\text{NPV}_{u,v}(\mathcal{Y}_j)$. The reported *p*-values are those obtained from the LFC algorithm, Contact-set approach, and numerical delta method (NDM).

than in UK. Similar interpretations suffice for the other six instances. The conclusion is largely the same irrespective of the type of test-statistic employed.

Focusing on the contact-set approach, the test results in Table 8 show that the long-short portfolio strategy in US second-order stochastically dominates that in the UK for second-time order at 5% significance level. This means that a risk-averse investor having a monotone decreasing discounting function will assign higher NPV to long-short strategy in the US than long-short strategy in the UK. This finding remains the same at 1% significance level when the paired-test involves US and Japan. In addition, the long-short portfolio strategy in the UK second-order stochastically dominates that in Japan for second-time order at 10% significance level. Similarly at 10% level of significance, we find that the long-short strategy in Europe second-order stochastically dominates that in Japan for second-time order.

Based on the results in Tables 7 and 8, we document that the NPV of long-short strategy in US second-order stochastically dominates that in the UK and Japan for both first and second-time order at 5% and 1% significance levels, respectively. This can be interpreted as an evidence for a dynamic dollar realised gains associated with green-tilted investing for a risk-averse agent with a monotone decreasing discounting function. These insights might be challenging to observe from a reduced-form analysis that relies solely on mean values, as it hampers the ability to infer discounted utility or discounting factors when using the t-test or the conventional static first-order and second-order stochastic dominance test.

At first glance, it may seem surprising that in general the portfolio strategy for the US dominates all the other corresponding portfolio strategies based on high minus low in GRF selection. However, one possible reason for the outcome results may be the fact that the constituents of the Green Revenues Index for US (Russell 1000) are more dispersed in terms of GRF whereas the constituents for Europe and Japan may be closer together in terms of GRF.

5. Conclusion

To transition towards a fully green economy, it is essential for investors in financial markets to operate in an environment where stock prices accurately reflect the extent of green activities undertaken by the companies in which they invest. In this study, we employ a detailed database of green revenues to track the relationship between green activities and stock prices across global markets, specifically comparing the five largest economies.

Firms seeking to position themselves within the portfolios of investors worldwide, regardless of their alignment with green initiatives, can gain valuable insights from our analysis. For those firms aiming to be perceived as viable green investment opportunities without compromising their competitive edge, relocating operations to regions such as the UK, Europe, or China may prove advantageous. These regions are more conducive to green investment strategies, offering an environment that supports the transition towards sustainability. Conversely, firms that wish to attract investors while simultaneously continuing or expanding their involvement in brown economic activities might find greater success operating in the US. For investors exposed to such firms, it is crucial to recognize that in economies where the majority of firms are transitioning towards a green economy, achieving green investment objectives becomes more feasible and can potentially be leveraged as a marketing tool for financial products. On the other hand, in countries where there is a large discrepancy between green and brown firms, investors may consider strategic trading approaches, such as going long on green firms and short on brown firms, in anticipation of upcoming regulatory shifts. Alternatively, investors might opt to take a position that goes long on brown firms and short on green firms, should external events—such as presidential elections—signal a potential retreat from green finance.

Firms in the UK and Europe demonstrate a higher level of environmental sustainability compared to firms in the US and Japan. As a result, investors subject to legal constraints, such as pension funds in France, or those influenced by stakeholder-driven objectives focused on fostering a more sustainable society, as seen in Scandinavian countries, may increasingly include firms from the UK and Europe in their portfolios to meet these social and regulatory requirements. Economies that are more supportive of green activities, such as the UK and Europe, exhibited greater resilience in their climate change initiatives during the COVID-19 pandemic.

We comparatively analyse the dynamics of the GRF over the entire period of our study. We conclude overall that companies in the UK, Europe and China have more green exposure than the companies in the US and Japan. By monitoring dollar realised returns portfolio strategies, we observed that a high minus low portfolio based on the GRF measure leads to US being the dominating force in terms of dollar generation. Thus, investors can use the GRF database to design strategies that still produce significant returns while taking into consideration the green revenues levels of the companies representing the major economies.

Portfolios including US and/or UK on one side and China and/or Japan on the other will be more diversified given the lower multiple correlation coefficient for those groupings. Our findings indicate some polarization for firms in the US and UK on one side and for firms in the China and Japan on the other.

Using the latest advances in stochastic order dominance for portfolio strategies, we compare the performances of long-short green portfolios that can be organised for the major five economies. Given the larger wedge in terms of green revenues for firms in the US relative to firms in the other economies, it is perhaps not surprising that in dollar terms the US portfolio is the most dominant whilst the Japan portfolio is the least dominant. As green activities cover more firms from more countries worldwide further research could be carried out along similar lines exposed in our paper to construct green portfolios with equity of firms from those other countries.

Policymakers may find the results of our study valuable and start thinking about measures to support activities related to helping the climate change agenda. In this context, governments could consider implementing taxes on the profits generated by firms engaged in environmentally harmful (brown) activities, and transfer those taxes to alleviate the financial losses induced to firms following green policies. Such measures could be introduced over a defined period, for example, five years, during which firms that change their operations to comply with the climate change agenda can offset their losses made by comparing to operating as before, in a standard non-green way. Over time, this approach may encourage more firms to shift to operating in a greener way, as they would not face immediate financial detriment while transitioning.

Furthermore, policy makers could look at the results obtained by investors into foreign firms. To further advance the climate change agenda, regulators could also implement taxes on profits obtained from investing in non-green firms in other countries. This may stop domestic investors following dollar profits with disregard to climate change agenda. Such a policy may contribute to a segregation of economic activities that would encourage the "green capital" to return to the green economies, while the "brown capital" will remain in the brown economies.

Notes

1. For a given company, any activity generating green revenues is mapped to one or more micro sectors and then an aggregate green revenue measure is computed. The GRDM relies on a three-tiers green system by dividing the activities of a company, over the 133 micro-sectors as following: Tier 1 includes activities with significant and clear environmental benefits(such as Solar); Tier 2 covers activities with limited but net positive environmental benefits (such as Water Utilities); Tier 3 includes activities with net neutral or negative environmental benefits (such as Nuclear).
2. The stocks in the quintile portfolios are weighted by market capitalisations.

Acknowledgements

We are grateful to many colleagues, Lazar Simeonidis, Di Wu, as well as participants at Essex Finance Centre (EFiC) 2024 Conference in Banking and Corporate Finance in Colchester and the European Financial Management Association 2024 conference in Lisbon, for their comments that helped us to improve the quality of the paper. We are also very grateful to FTSE Russell (part of the LSEG now) for their support with the green revenues data and clarifications about their methodology.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

No funding was received.

References

Alexander, G., & Baptista, A. (2004). A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model. *Management Science*, *50*(9), 1261–1273. https://doi.org/10.1287/mnsc.1040.0201

Andersen, T. G., Bollerslev, T., Christoffersen, P. F., & Diebold, F. X. (2006). Volatility and correlation forecasting. In *Handbook of Economic Forecasting*, vol. 1, 777–878.

Andriosopoulos, D., Doumpos, M., Pardalos, P. M., & Zopounidis, C. (2019). Computational approaches and data analytics in financial services: A literature review. *Journal of the Operational Research Society*, *70*(10), 1581–1599. https://doi.org/10.1080/01605682.2019.1595193

Ardia, D., & Boudt, K. (2015). Testing equality of modified Sharpe ratios. *Finance Research Letters*, *13*, 97–104. https://doi.org/10.1016/j.frl.2015.02.008

Auzepy, A., Bannier, C. E., & Martin, F. (2023). Are sustainability-linked loans designed to effectively incentivize corporate sustainability? A framework for review. *Financial Management*, *52*(4), 643–675. https://doi.org/10.1111/fima.12437

Ballestero, E., Bravo, M., Pérez-Gladish, B., Arenas-Parra, M., & Plà-Santamaria, D. (2012). Socially responsible investment: A multicriteria approach to portfolio selection combining ethical and financial objectives. *European Journal of Operational Research*, *216*(2), 487–494. https://doi.org/10.1016/j.ejor.2011.07.011

Ban, G., El Karoui, N., & Lim, A. (2018). Machine learning and portfolio optimization. *Management Science*, *64*(3), 1136–1154. https://doi.org/10.1287/mnsc.2016.2644

Bao, Y. (2009). Estimation risk-adjusted Sharpe ratio and fund performance ranking under a general return distribution. *Journal of Financial Econometrics*, *7*(2), 152–173. https://doi.org/10.1093/jjfinec/nbn022

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

Bolton, P., & Kacperczyk, M. (2022). Global pricing of carbon-transition risk. *Journal of Finance*, *78*(6), 3677–3754.

Brodie, J., Daubechies, I., De Mol, C., Giannone, D., & Loris, I. (2009). Sparse and stable Markowitz portfolios. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(30), 12267–12272. https://doi.org/10.1073/pnas.0904287106

Brooks, C., & Schopohl, L. (2021). Green accounting and finance: Advancing research on environmental disclosure, value impacts and management control systems. *The British Accounting Review*, *53*(1), 100973. https://doi.org/10.1016/j.bar.2020.100973

Capponi, A., & Rubtsov, A. (2022). Systemic risk-driven portfolio selection. *Operations Research*, *70*(3), 1598–1612. https://doi.org/10.1287/opre.2021.2234

Chapple, L., Clarkson, P. M., & Gold, D. L. (2013). The cost of carbon: Capital market effects of the proposed emission trading scheme (ETS). *Abacus*, *49*(1), 1–33. https://doi.org/10.1111/abac.12006

Chatterji, A., Durand, R., Levine, D., & Touboul, S. (2016). Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strategic Management Journal*, *37*(8), 1597–1614. https://doi.org/10.1002/smj.2407

Choi, B., & Luo, L. (2021). Does the market value greenhouse gas emissions? Evidence from multi-country firm data. *British Accounting Review*, *53*(1), 100909. https://doi.org/10.1016/j.bar.2020.100909

Clarkson, P. M., Li, Y., Pinnuck, M., & Richardson, G. D. (2015). The valuation relevance of greenhouse gas emissions under the European Union carbon emissions trading scheme. *European Accounting Review*, *24*(3), 551–580. https://doi.org/10.1080/09638180.2014.927782

Cornell, B. (2021). ESG preferences, risk and return. *European Financial Management*, *27*(1), 12–19. https://doi.org/10.1111/eufm.12295

Cornuéjols, G., Peña, J., & Tütüncü, R. (2018). *Optimization methods in Finance* (2nd Ed.). Cambridge University Press.

Crisci, A. (2023). A note on regression diagnostics for generalized estimating equations: Empirical study on environmental disclosure determinants. *Journal of the Operational Research Society*, *74*(4), 1042–1048. https://doi.org/10.1080/01605682.2022.2053310

DeMiguel, V., Garlappi, L., Nogales, F., & Uppal, R. (2009a). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norm. *Management Science*, *55*(5), 798–812. https://doi.org/10.1287/mnsc.1080.0986

DeMiguel, V., Garlappi, L., & Uppal, R. (2009b). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial Studies*, *22*(5), 1915–1953. https://doi.org/10.1093/rfs/hhm075

DeMiguel, V., & Nogales, F. J. (2009). Portfolio selection with robust estimation. *Operations Research*, *57*(3), 560–577. https://doi.org/10.1287/opre.1080.0566

DeMiguel, V., Plyakha, Y., Uppal, R., & Vilkov, G. (2013). Improving portfolio selection using option-implied volatility and skewness. *Journal of Financial and Quantitative Analysis*, *48*(6), 1813–1845. https://doi.org/10.1017/S0022109013000616

Dietz, S., & Matei, N. A. (2016). Spaces for agreement: A theory of time-stochastic dominance and an application to climate change. *Journal of the Association of Environmental and Resource Economists*, *3*(1), 85–130. https://doi.org/10.1086/683684

Dimson, E., Marsh, P., & Staunton, M. (2020). Divergent ESG ratings. *Journal of Portfolio Management*, *47*(1), 75–87. https://doi.org/10.3905/jpm.2020.1.175

Dümbgen, L. (1993). On nondifferentiable functions and the bootstrap. *Probability Theory and Related Fields*, *95*(1), 125–140. https://doi.org/10.1007/BF01197342

Edmans, A. (2023). The end of ESG. *Financial Management*, *52*(1), 3–17. https://doi.org/10.1111/fima.12413

Fang, Z., & Santos, A. (2019). Inference on directionally differentiable functions. *Review of Economic Studies*, *86*(1), 377–412.

Gasser, S. M., Rammerstorfer, M., & Weinmayer, K. (2017). Markowitz revisited: Social portfolio engineering. *European Journal of Operational Research*, *258*(3), 1181–1190. https://doi.org/10.1016/j.ejor.2016.10.043

Goldfarb, D., & Iyengar, G. (2003). Robust portfolio selection problems. *Mathematics of Operations Research*, *28*(1), 1–38. https://doi.org/10.1287/moor.28.1.1.14260

Griffin, P. A., Lont, D. H., & Pomare, C. (2021). The curious case of Canadian corporate emissions valuation. *British Accounting Review*, *53*(1), 100922. https://doi.org/10.1016/j.bar.2020.100922

Hallerbach, W., Spronk, J., & Soppe, A. (2004). A framework for managing a portfolio of socially responsible investments. *European Journal of Operational Research*, *153*(2), 517–529. https://doi.org/10.1016/S0377-2217(03)00172-3

Hamilton, S., Jo, H., & Statman, M. (1993). Doing well while doing good? The investment performance of socially responsible mutual funds. *Financial Analysts Journal*, *49*(6), 62–66. https://doi.org/10.2469/faj.v49.n6.62

Heinkel, R., Kraus, A., & Zechner, J. (2001). The effect of green investment on corporate behavior. *Journal of Financial and Quantitative Analysis*, *36*(4), 431–449. https://doi.org/10.2307/2676219

Hong, H., & Li, J. (2018). The numerical delta method. *Journal of Econometrics*, *206*(2), 379–394. https://doi.org/10.1016/j.jeconom.2018.06.007

Johnston, D. M., Sefcik, S. E., & Soderstrom, N. S. (2008). The value relevance of greenhouse gas emissions allowances: An exploratory study in the related United States SO2 market. *European Accounting Review*, *17*(4), 747–764. https://doi.org/10.1080/09638180802481615

Kempf, A., & Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. *European Financial Management*, *13*(5), 908–922. https://doi.org/10.1111/j.1468-036X.2007.00402.x

Klassen, R. D., & McLaughlin, C. P. (1996). The impact of environmental management on firm performance. *Management Science*, *42*(8), 1199–1214. https://doi.org/10.1287/mnsc.42.8.1199

Kolm, P. N., Tütüncü, R., & Fabozzi, F. J. (2014). 60 years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, *234*(2), 356–371. https://doi.org/10.1016/j.ejor.2013.10.060

Konar, S., & Cohen, M. A. (2001). Does the market value environmental performance? *Review of Economics and Statistics*, *83*(2), 281–289. https://doi.org/10.1162/003465301151143815

Larcker, D. F., Tayan, B., & Watts, E. M. (2022). Seven myths of ESG. *European Financial Management*, *28*(4), 869–882. https://doi.org/10.1111/eufm.12378

Lassance, N., DeMiguel, V., & Vrins, F. (2022). Optimal portfolio diversification via independent component analysis. *Operations Research*, *70*(1), 55–72. https://doi.org/10.1287/opre.2021.2140

Ledoit, O., & Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, *15*(5), 850–859. https://doi.org/10.1016/j.jempfin.2008.03.002

Ledoit, O., & Wolf, M. (2017). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *Review of Financial Studies*, *30*(12), 4349–4388. https://doi.org/10.1093/rfs/hhx052

Lee, K., Linton, O., & Whang, Y.-J. (2023). Testing for time stochastic dominance. *Journal of Econometrics*, *235*(2), 352–371. https://doi.org/10.1016/j.jeconom.2022.03.012

Lioui, A. (2022). Understanding the carbon price(s) of risk. *SSRN Electronic Journal*, Technical report, SSRN. https://doi.org/10.2139/ssrn.4174385

Liu, Y., Rekkas, M., & Wong, A. (2012). Inference for the Sharpe ratio using a likelihood-based approach. *Journal of Probability and Statistics*, *2012*, 1–24. 2012:Article ID 878561. https://doi.org/10.1155/2012/878561

Lo, A. W. (2002). Statistics of Sharpe ratios. *Financial Analysts Journal*, *58*(4), 36–52. https://doi.org/10.2469/faj.v58.n4.2453

Marohn, M., & Auer, B. R. (2024). A note on Steuer and Utz's (2023) multi-objective optimization approach for generating sustainability-efficient fronts. *European Journal of Operational Research*, *316*(2), 792–797. https://doi.org/10.1016/j.ejor.2024.03.023

Matos, P. (2020). *ESG and responsible institutional investing around the world: A critical review*. CFA Institute Research Foundations.

Matsumura, E. M., Prakash, R., & Vera-Muñoz, S. C. (2014). Firm-value effects of carbon emissions and carbon disclosures. *Accounting Review*, *89*(2), 695–724. https://doi.org/10.2308/accr-50629

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

Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2022). Dissecting green returns. *Journal of Financial Economics*, *146*(2), 403–424. https://doi.org/10.1016/j.jfineco.2022.07.007

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

Popescu, I. (2007). Robust mean-covariance solutions for stochastic optimization. *Operations Research*, *55*(1), 98–112. https://doi.org/10.1287/opre.1060.0353

Qi, J., Rekkas, M., Wong, A., et al. (2018). Highly accurate inference on the Sharpe ratio for autocorrelated return data. *Journal of Statistical and Econometric Methods*, *7*(1), 21–50.

Quaye, E., Tunaru, D., & Tunaru, R. (2024). Green-adjusted share prices: A comparison between standard investors and investors with green preferences. *Journal of Financial Stability*, *74*, 101314. forthcoming. https://doi.org/10.1016/j.jfs.2024.101314

Ramelli, S., Ossola, E., & Rancan, M. (2021). Stock price effects of climate activism: Evidence from the first global climate strike. *Journal of Corporate Finance*, *69*, 102018. https://doi.org/10.1016/j.jcorpfin.2021.102018

Sahamkhadam, M., & Stephan, A. (2024). Socially responsible multiobjective optimal portfolios. *Journal of the Operational Research Society*, *75*(10), 2065–2076. https://doi.org/10.1080/01605682.2024.2303075

Steuer, R. E., & Utz, S. (2023). Non-contour efficient fronts for identifying most preferred portfolios in sustainability investing. *European Journal of Operational Research*, *306*(2), 742–753. https://doi.org/10.1016/j.ejor.2022.08.007

Temple-West, P. (2023). US retreats from ESG investing as green bonds grow elsewhere. *Financial Times*. https://www.ft.com/patrick-temple-west?

Wang, J., & Zheng, N. (2020). Measures of correlation for multiple variables. arXiv preprint arXiv:1401.4827.

Yenipazarli, A., & Vakharia, A. (2015). Pricing, market coverage and capacity: Can green and brown products co-exist? *European Journal of Operational Research*, *242*(1), 304–315. https://doi.org/10.1016/j.ejor.2014.09.039

Zerbib, O. D. (2019). The effect of pro-environmental preferences on bond prices: Evidence from green bonds. *Journal of Banking & Finance*, *98*, 39–60. https://doi.org/10.1016/j.jbankfin.2018.10.012

Zerbib, O. D. (2022). A sustainable capital asset pricing model (S-CAPM): Evidence from environmental integration and sin stock exclusion. *Review of Finance*, *26*(6), 1345–1388. https://doi.org/10.1093/rof/rfac045