BIS Working Papers

No 1362

Embracing carbon uncertainty in portfolio construction

by Fan Dora Xia and Omar Zulaica

Monetary and Economic Department

June 2026

JEL classification: G11, G28, Q54, Q56

Keywords: carbon footprints, sovereign debt, portfolio optimization, risk parity

This publication is available on the BIS website (www.bis.org).

ISSN 1020-0959 (print)
ISSN 1682-7678 (online)

Embracing Carbon Uncertainty in Portfolio Construction

Fan Dora Xia[*] **Omar Zulaica**[*,1]

[*]Bank for International Settlements

Abstract

We propose a framework for constructing fixed-income portfolios of sovereign bonds that integrates financial and environmental considerations. Central to our approach is the introduction of *carbon returns*, a concept analogous to financial returns, modeled as random variables to capture the inherent uncertainty of future carbon emissions. Based on the financial and carbon return profiles of individual countries' sovereign bonds, we employ an algorithm inspired by Hierarchical Risk Parity (HRP) to construct portfolios that balance each country's contribution to the portfolio's tail risk, as measured by expected shortfall, of financial and carbon returns. Focusing on developed market sovereign bonds, our results demonstrate that it is possible to design portfolios that effectively align decarbonization objectives with financial performance, both in-sample and out-of-sample, while accommodating diverse investor preferences.

JEL Classification: G11, G28, Q54, Q56
Keywords: carbon footprints, sovereign debt, portfolio optimization, risk parity

[1] The authors (dora.xia@bis.org and omar.zulaica@bis.org) would like to thank Andrew Ang, Pierre Cardon, Ingo Fender, Luis Garcia-Feijoo, Philip Woolridge, and participants of the 47th EMEAP-BIS Forum on Financial Market, the Banco de España Meeting of Experts on International Reserve Management, the joint World Bank - Financial Analysts Journal Conference on Public Asset Management, and the Ninth Public Investors Conference for helpful comments, and Jimmy Shek for excellent research assistance. The views expressed in this publication are those of the authors and do not necessarily reflect the views of the BIS or its member central banks.

1 Introduction

Institutional investors, such as reserve managers, are increasingly recognizing the importance of integrating climate risk considerations into their portfolios. This reflects a broader move for institutional investors to embed sustainability—spanning environmental, social, and governance (ESG) goals—into portfolio construction. In 2020, (Fender et al. (2020)) showed that sustainability was a relevant objective for at least one third of central banks.[2] Fender et al. (2022), in turn, highlight ways in which such considerations can be put into practice in public investors' portfolios. Within this broader ESG agenda, climate-related risks have emerged as both the most measurable and policy salient dimension for fixed income investors, which motivates our focus in what follows.

Presumably, the integration of climate considerations into investment decision-making is driven by two key factors. The first are the risk implications of climate change, which necessitate action. These may arise from physical risks, such as the increasing frequency and severity of extreme weather events, or from transition risks linked to regulatory policies, technological advancements, and shifts in consumer preferences towards a low-carbon economy. The second factor is the potential for institutional investors to influence climate outcomes through their investment strategies, such as by raising the funding costs of environmentally damaging activities (Scatigna et al. (2021); Xia and Zulaica (2022)).[3] To illustrate this, Swinkels et al. (2025) argue that government bond investors can help close the financing gap for the sustainable development goals (SDGs) by directing capital toward governments with strong yet underfunded sustainability policies, and give practical examples of how investors can integrate SDG scores in portfolios of developed and emerging markets.

In practice, for investors with longer horizons, such objectives are often justified purely from a risk management perspective. For others, where mandates allow, actively influencing climate outcomes may also align with their broader goals (Carstens (2024)). Regardless, given the central role of government bonds in institutional investors' portfolios, reducing their carbon footprints has become a pressing practical concern. Two key questions arise when building decarbonized portfolios for sovereign securities: the first is how to measure the portfolio's carbon footprint; the second is how to blend the notion of a

[2] More recently, another survey of reserve managers indicates that more than a third of surveyed central banks have incorporated sustainability as a fourth reserve management objective, alongside the traditional goals of liquidity, safety, and return (HSBC (2025)).

[3] See de Bandt et al. (2025) for an in-depth discussion on climate-related risks and a survey of empirical analysis of their impact on the financial system, with a particular focus on banks.

carbon footprint with considerations of the portfolio's financial performance.

The existing literature frequently employs constrained optimization frameworks to decarbonize sovereign bond portfolios. These frameworks typically prioritize portfolio financial returns—such as minimizing tracking error relative to a benchmark or index-while incorporating carbon footprint considerations via carbon budgets. These budgets constrain the portfolio's carbon footprint, typically measured as the weighted sum of countries' carbon emissions, with the weights corresponding to each country's allocation within the portfolio. Countries' carbon emissions can be represented either in raw terms or scaled by factors such as GDP or population, and referred to as carbon intensities. Notable examples of such allocation methodologies include Cheng et al. (2022), Le Guenedal and Roncalli (2022), and Schwaiger et al. (2023).

A key limitation of this approach is the treatment of individual countries' carbon footprints as deterministic values, either based on historical data (ie, the latest observation of a carbon footprint, assumed to represent future behavior) or forward-looking values tied to a specific scenario (eg, NGFS scenarios which are also deterministic at the carbon footprint level). This perspective fails to account for the inherent uncertainty in any projection. It is akin to conducting an asset allocation exercise based on a single projected path of future returns for each asset (in this case, each sovereign bond), which could result in suboptimal outcomes unless actual conditions align perfectly with the projections.

In this paper, we propose a novel framework for incorporating carbon footprint considerations into sovereign bond allocations, treating the carbon footprint of individual countries as a random variable rather than a deterministic one. The key innovation of this framework is the introduction of the concept of *carbon returns* for individual countries' sovereign bonds.[4] Carbon returns are defined as the negative change in carbon emissions for a given country, paralleling the concept of financial returns, which are based on changes in bond prices. Like financial returns, carbon returns are random variables, and their distributions can be inferred from historical data or forecasting models. Using stochastic carbon returns instead of historical emissions levels or deterministic projections has two key advantages. It avoids simply favoring current low emitters and instead rewards actual reductions. It also enhances robustness, as return distributions are found to vary less across emissions measures than absolute levels.

To construct decarbonized sovereign portfolios, we adopt an approach inspired by

[4] While our analysis focuses on decarbonizing sovereign bond portfolios, the framework readily extends to other asset classes.

the Hierarchical Risk Parity (HRP) algorithm proposed by López de Prado (2016). The framework balance the contributions of financial and carbon risks from each country's sovereign bonds. Unlike traditional mean-variance optimization, the HRP algorithm offers a more robust and diversified approach to portfolio construction. It allocates risk effectively by leveraging the hierarchical structure of assets, thereby avoiding corner solutions. Additionally, HRP is resilient to estimation errors in covariance and correlation matrices. By utilizing clustering techniques and hierarchical weight assignment, it is less sensitive to noisy data compared to the covariance matrix inversion required in mean-variance optimization. Together, these features make HRP particularly well-suited to our application.

We apply this methodology to developed market sovereign bonds and demonstrate that it is possible to design portfolios that effectively align decarbonization objectives with financial performance, both in-sample and out-of-sample. Our focus on sovereign bonds from advanced economies reflects their significant share in institutional investors' sovereign bond portfolios (see for example, World Bank (2025)). Nonetheless, our framework can be equally applied to sovereign bonds from all countries.

The remainder of this paper is organized as follows. Section 2 introduces the concept of carbon returns and describes the data used to compute country-level carbon returns. Section 3 outlines how to decarbonize sovereign bond portfolios using carbon returns, with a primary focus on our framework inspired by the HRP algorithm. Section 4 presents the results of the HRP optimization, both in-sample and out-of-sample. Section 5 concludes.

2 Carbon returns

We introduce a novel concept termed *carbon return*, defined as the negative percentage change in a carbon footprint measure. The negative sign reflects the desirability of reduced carbon footprints, mirroring how investors seek positive financial returns. Given this framing, all standard metrics such as expected return, volatility, tail risk measures, and correlations can be calculated for carbon returns. This shifts the analysis from deterministic values to the distribution of outcomes, representing a significant departure in the way carbon footprints are currently incorporated into objective functions for portfolio construction.

We measure a country's environmental footprint for sovereign bonds using CO2 emissions. The approach is justified for two reasons. First, CO2 emissions generally account for approximately 80% of total greenhouse gas (GHG) emissions in advanced economies, which makes them a reliable proxy for overall carbon footprints in these countries. Indeed, CO2 provides a timely and accurate representation of a country's overall GHG footprint (Penninga and Zomerdijk (2024)). Second, data on other GHG are updated less frequently, often with significant publication delays, with some categories (eg, consumption-based measures) typically unavailable. Throughout, we use "CO2 emissions," "carbon emissions," and "carbon footprints" interchangeably.

2.1 Definition and motivation

For a given carbon footprint measure C_t, such as production-based carbon emissions, the carbon return R_t^c is defined as:

$$R_t^c = -\frac{C_t}{C_{t-1}}. \tag{1}$$

Analogous to financial returns, which capture the rate of appreciation or depreciation in a security's price, carbon returns represent the rate of reduction or increase in a country's carbon footprint. The negative sign ensures that a reduction in emissions is treated as a positive outcome, as an increase in emissions reflects an undesirable state. This metric can therefore be interpreted similarly to financial returns, where maximizing the expected return is a desirable objective.

This is a crucial contribution, as an often overlooked fact is the inherent uncertainty surrounding a country's future carbon footprint, which has led to the definition of multiple scenarios (eg, CO2 transition pathways from NGFS (2022)). In the existing literature, the problem of decarbonizing portfolios is frequently formulated as follows (see Le Guenedal and Roncalli (2022), for example):

$$\begin{aligned}
&\mathbf{w}^* = \arg\min \frac{1}{2}(\mathbf{w}-\mathbf{b})'\Sigma(\mathbf{w}-\mathbf{b}) - \gamma(\mathbf{w}-\mathbf{b})'\mu \\
&\text{s.t.} \quad \mathbf{w}'\mathbf{1} = 1; \quad \mathbf{c}'\mathbf{w} \leq c^+,
\end{aligned}$$

where $\mathbf{w}^* = (w_1^*, \ldots, w_n^*)$ is the target portfolio, $\mathbf{b} = (b_1, \ldots, b_n)$ is the benchmark portfolio, Σ is the covariance matrix of financial returns, $\mu = (\mu_1, \ldots, \mu_n)$ is the vector of expected financial returns, $\mathbf{c} = (c_1, \ldots c_n)$ collects carbon footprints of individual assets, and c^+ is the upper bound on the portfolio's carbon footprint. The objective is to construct a portfolio $\mathbf{w}^*$ that closely tracks the benchmark portfolio $\mathbf{b}$ while keeping its carbon footprint below the upper bound c^+.

The carbon footprint metric $\mathbf{c}$ is typically based on historical data, making it a backward-looking approach. To address this limitation, authors such as Schwaiger et al. (2023) have introduced forward-looking versions of these variables. However, these approaches often rely on a single, deterministic scenario for each path, effectively ignoring the uncertainty inherent in carbon footprint forecasts. This is analogous to constructing an investment portfolio ignoring the variance of asset returns and their correlations, which can result in suboptimal outcomes.

We model carbon returns as random variables drawn from a distribution when constructing portfolios, analogous to financial returns. This explicitly accounts for uncertainty in carbon footprints.

Shifting focus to stochastic carbon returns from historical emission levels or deterministic emission projections offers two key advantages. First, approaches based on carbon footprint levels often result in portfolio shifts toward current low emitters, rather than countries with more aggressive, forward-looking transition plans. As argued by Angelini (2024), reallocating capital based solely on emission levels is unlikely to drive meaningful change. Low emitters already have low emissions, while high emitters may face higher financing costs, potentially hindering their adoption of sustainable technologies. In contrast, capital allocation based on emission reductions incentivizes decarbonization. Second, as we will show later, while the absolute level of carbon emissions depends on the type of emissions, the historical distribution of carbon returns tends to exhibit greater consistency. This consistency ensures that portfolios constructed using carbon returns are more robust to variations in the measurement of carbon emissions.

How can carbon returns be used to construct decarbonized portfolios? By design, carbon returns are defined in a manner analogous to financial returns. As a result, key metrics commonly used for financial returns—such as expected return, volatility, expected shortfall, and correlation—are equally applicable to carbon returns. Furthermore, frameworks designed to optimize within the financial return space can be directly applied to the carbon return space.

2.2 Data on carbon emissions

To apply the concept of carbon returns, it is essential to select an appropriate metric for measuring country-level carbon footprints. In the literature, a common approach is to use carbon emissions either in their absolute form or normalized by factors that reflect the size of the economy, such as GDP or population. The normalized measure, often referred to as a carbon intensity, accounts for economic size and enables easier cross-country comparisons.

As the absolute *level* of carbon emissions is more directly relevant for assessing climate risk, we focus on raw emissions data in our analysis. This is because, when footprints are taken percentage changes, the size effect of raw carbon emissions is largely accounted for. Nonetheless, our carbon return concept can also be applied to carbon intensities. In such cases, the resulting carbon return would reflect the rate of decarbonization as well as the growth rate of GDP and population.

When analyzing a country's carbon emissions, it is also important to determine the type of emissions to consider. We focus on production-based CO2 emissions because they involve fewer imputations and have longer, more consistent histories than consumption-based estimates or variants with land-use change. Nonetheless, while absolute levels of emissions vary depending on the accounting method, the distribution of historical carbon returns remains relatively consistent, which reveals a collateral benefit of using our framework (see Table 5). See Appendix A for a detailed description of the different types of carbon emissions available.

To source the above information, we rely on the Global Carbon Project (GCP). The GCP tracks trends in carbon emissions and removals globally and is widely recognized as the most comprehensive database of its kind. It covers over 190 countries and provides historical data on production-based emissions starting from the 1800s, as well as consumption-based emissions from 1990. [5]We procure the data in bulk via Our World in Data (OWID) (link). See Figure 11 in Appendix B for the time series of production-based carbon emissions for the countries included in our analysis. Our analysis focuses on a time series spanning 1982 to 2021, providing 40 years of carbon emissions data for 23 advanced economies. These economies include Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Singapore, South Korea, Spain, Sweden, Switzerland, the United Kingdom, and the United States.[6]

[5] Other widely used sources of carbon-emissions data include the United Nations Greenhouse Gas Inventory, the OECD Air Emissions database, and the Trucost Environmental Dataset available on the S&P Global Marketplace. We use the GCP as our primary dataset because it offers broad coverage and is freely and easily accessible. The distribution of historical carbon returns is consistent across these sources (see Table 6).

[6] Note that the span of the carbon return data differs from that of the financial return data, primarily due to differences in availability. However, as long as both historical data samples used are deemed informative about the future distribution of the respective returns, the mismatch in data spans does not pose a significant issue. This is particularly relevant in the context of portfolio optimization, where the investment manager much choose the distribution which is most representative of the future.

2.3 Empirical analysis of carbon returns

Similar to financial returns, portfolio construction exercises involving carbon returns should be inherently forward looking and based on estimates of the future distributions of these returns. For our analysis, we use historical data as an input to estimate the distributions of carbon returns, a common practice in asset allocations for financial returns. Nonetheless, our framework can readily accommodate results from forward looking estimates that do not rely on historical data–this discussion is nonetheless out of the scope of this paper. We analyze historical carbon return data next.

The left-hand panel of Figure 1 presents the historical carbon emissions of the United States across different accounting methods. The corresponding distributions of carbon returns are shown in the right-hand panel. Note that carbon emissions data are available at an annual frequency, and carbon returns are therefore annual (”year on year”) changes in carbon emissions.

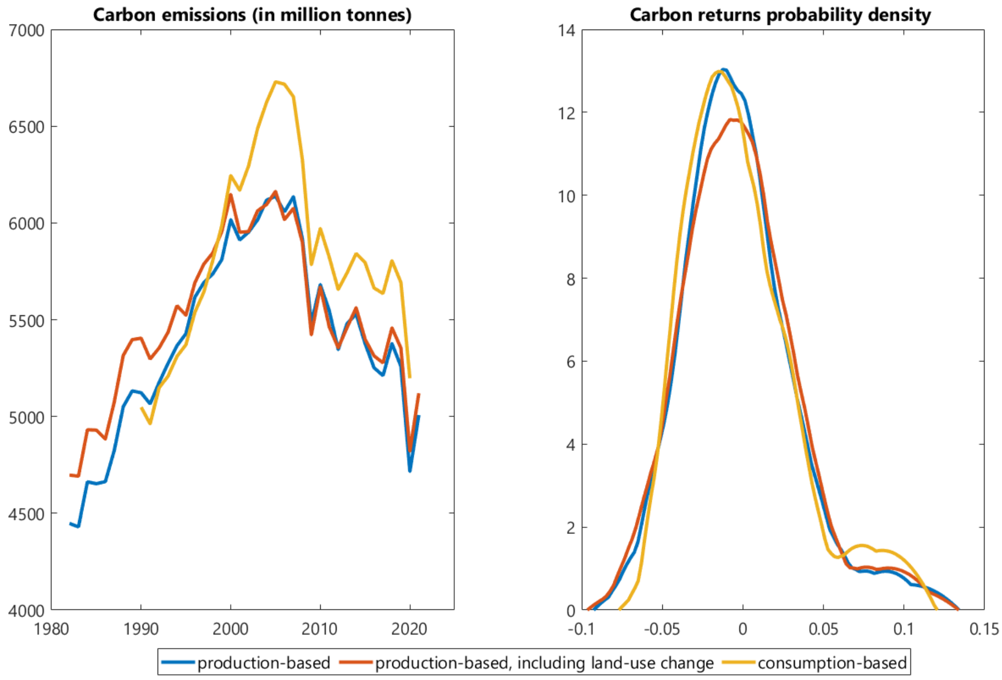
While the absolute levels of carbon emissions vary significantly depending on the accounting method—raising questions about which elements to include—the distributions of carbon returns demonstrate remarkable similarity (Figure 1, right-hand panel). This observation suggests that the choice of carbon accounting method may have a limited impact on the distribution of carbon returns.

We evaluate this hypothesis by formally testing whether carbon returns calculated based on different types of carbon emissions are statistically identical in Table **??** of Appendix B. The results indicate that the null hypothesis—that carbon returns based on production-based emissions and consumption-based emissions are drawn from identical distributions—cannot be rejected at conventional confidence levels for all countries in our sample. Similarly, for production-based emissions that consider land-use change, the null hypothesis of identical distributions with production-based or consumption-based emissions is not rejected at 10% significance level, except in the cases of Australia and New Zealand.

For our portfolio construction exercise, we select production-based carbon emissions from GCP as our baseline measure. Similar to our choice between CO2 and GHG data, production-based emissions are preferred vis-à-vis alternatives due to their greater precision and longer time series availability.

We estimate carbon returns for each of the 23 countries in our sample. Table 1 reports their mean return, volatility, value-at-risk and expected shortfall (both at the 97.5% confidence level). Among the countries analyzed, European economies such as Germany,

Figure 1: Carbon emissions and carbon returns for the United States



Note: The left-hand panel displays carbon emissions (in million tonnes) across various emission types in the United States, while the right-hand panel depicts the probability density of the corresponding carbon returns. The blue, red, and yellow lines represent production-based carbon emissions, production-based emissions including land-use change, and consumption-based emissions, respectively.

France, Italy, and the United Kingdom tend to exhibit higher carbon returns, on average (second column), and less extreme values in the tails (fourth and fifth columns). These findings reflect the outcome of relatively favorable decarbonization trends in Europe, likely driven by effective domestic policies and structural changes towards a low-carbon economy. This is consistent with the observation in Crippa et al. (2023) that the European Union saw the most significant decrease in greenhouse-gas emissions among major emitters. Between 1990 and 2024, the EU's GHG emissions were nearly 35% lower in 2024 than in 1990, illustrating a decoupling of territorial emissions from economic growth. Over the same period, the United States' emissions fell by only about 5%.

Table 1: Summary statistics of carbon returns by country

Country	Mean Return	Volatility	Value-at-Risk (97.5%)	Expected Shortfall (97.5%)
Australia	-1.36%	2.26%	-6.55%	6.73%
Austria	-0.49%	5.06%	-11.69%	15.73%
Belgium	0.52%	4.90%	-9.85%	12.64%
Canada	-0.64%	3.04%	-6.11%	6.49%
Denmark	1.01%	8.95%	-20.66%	21.51%
Finland	0.39%	8.98%	-18.71%	19.85%
France	0.94%	3.78%	-7.80%	9.26%
Germany	1.05%	3.15%	-5.47%	5.53%
Greece	-0.45%	5.66%	-11.58%	13.94%
Ireland	-1.01%	4.31%	-8.19%	9.59%
Italy	0.24%	4.23%	-8.67%	8.74%
Japan	-0.43%	3.93%	-11.27%	13.36%
Netherlands	0.26%	4.90%	-9.05%	11.03%
New Zealand	-1.89%	4.44%	-11.68%	13.04%
Norway	-0.69%	3.56%	-7.86%	7.96%
Portugal	-1.30%	7.48%	-19.13%	25.34%
Singapore	-2.51%	19.11%	-41.27%	54.87%
South Korea	-3.92%	5.30%	-12.94%	13.81%
Spain	-0.49%	5.95%	-11.71%	13.89%
Sweden	1.53%	4.63%	-10.60%	12.37%
Switzerland	0.17%	4.41%	-10.71%	12.03%
United Kingdom	1.12%	3.82%	-6.03%	6.29%
United States	-0.22%	3.29%	-5.72%	6.18%

Note: The table summarizes the statistics of carbon returns calculated using production-based carbon emissions data from 1982 to 2021. The top 10 values for mean return and value-at-risk, as well as the bottom 10 values for volatility and expected shortfall, are highlighted.

3 Portfolio construction with carbon returns

This section examines how conventional portfolio construction methods can be applied to incorporate sovereign-level carbon return data. First, we demonstrate the use of the classical mean-variance approach, introduced by Markowitz (1952), which is arguably the most widely used portfolio construction method, and we highlight its limitations. Second, we introduce a machine learning based method known as Hierarchical Risk Parity (HRP), proposed by López de Prado (2016). We then adapt HRP to be able to construct portfolios that integrate *both* carbon returns and financial returns. Details on financial return data can be found in Appendix C.

3.1 Mean-variance optimization

In a standard mean-variance setting, an investor seeks to minimize total portfolio volatility for a given level of expected (financial) returns. The investor's objective is to construct an optimal portfolio, $\mathbf{w}^* = (w_1^*, \ldots, w_n^*)$, where w_i^* represents the weight allocated to country i. When applied to carbon returns, the goal of this vector of weights $\mathbf{w}^*$ is to minimize the variance of carbon returns while achieving a target level of expected decarbonization.

The mathematical formulation of this optimization problem is as follows:

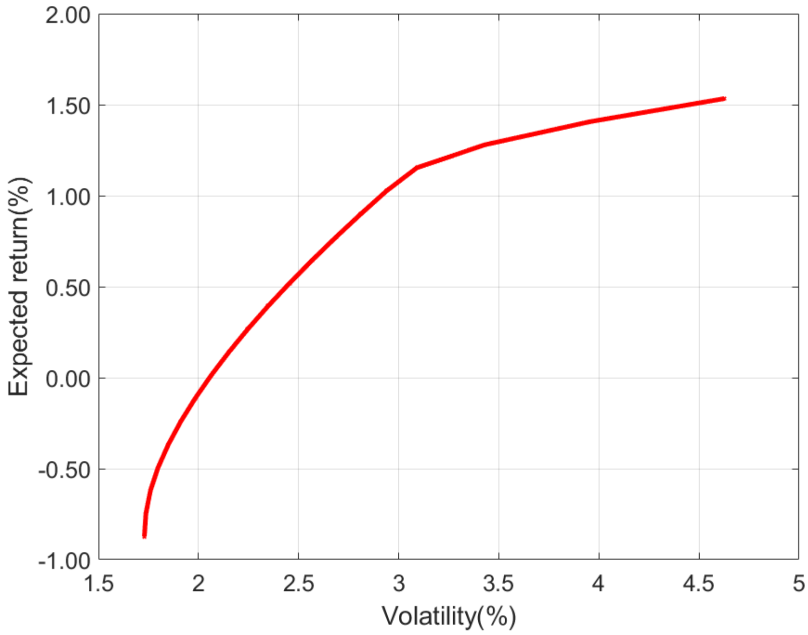
$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \mathrm{Var}[R_{p,t}^c] \quad \text{s.t.} \quad \mathbf{w'1} = 1, \quad \mathbf{w} \geq 0, \quad \mathrm{E}[R_p^c] = \bar{r}^c,$$

where $\mu_t^c = (\mu_{1,t}^c, \ldots, \mu_{n,t}^c)$ is a vector collecting the expected carbon returns for individual countries, Σ is the cross-country covariance matrix of carbon returns, the portfolio's expected decarbonization is given by $\mathrm{E}[R_p^c] = \mathbf{w}'\mu_t^c$, and the portfolio's decarbonization variance is given by $\mathrm{Var}[R_{p,t}^c] = \mathbf{w}'\Sigma\mathbf{w}$. Here, the constraint $\mathbf{w'1} = 1$ ensures that the portfolio weights add up to 100%, while permitting long positions only ($\mathbf{w} \geq 0$). The third constraint specifies a target level of expected portfolio decarbonization, $\bar{r}^c$, which corresponds to global emissions reduction goals.

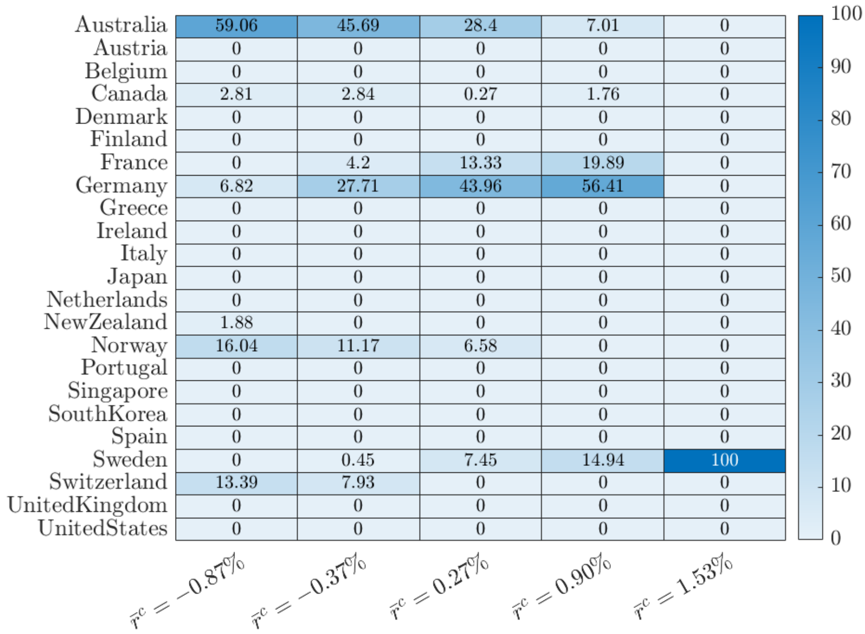
The empirical results (Figure 2) demonstrate that historical emission patterns yield an expected decarbonization rate ranging from -1% to 1.5% (top panel, y-axis). At the leftmost edge of the efficient frontier, portfolios exhibit negative expected decarbonization rates, indicating that these allocations, on average, lead to an increase in the carbon footprint. While such portfolios minimize emission-change volatility, they are environmentally undesirable as they imply a worsening of emissions over time, on average. As the right side of the efficient frontier, the portfolios can achieve higher levels of expected decarbonization but at a cost of greater volatility in carbon returns.

Figure 2: Results from mean-variance optimization

(a) Efficient frontier



(b) Weights for selected portfolios



Note: The top panel illustrates the efficient frontier derived from mean-variance optimization. The bottom panel displays the portfolio weights allocated to individual countries for selected portfolios along the efficient frontier, corresponding to varying levels of the expected decarbonization rate.

An examination of the portfolio weights, shown in the bottom panel of Figure 2 for five selected portfolios along the frontier, provides insights into the composition of these allocations. Portfolios with low volatility tend to overweight countries such as Australia, whose emissions profile is relatively stable (”low risk”) but has often trended upward in absolute terms. As we move toward the right side of the efficient frontier, portfolios increasingly allocate to countries such as Sweden, Germany, and France. These countries exhibit stronger emissions reductions but are characterized by greater year-to-year variability, which is not necessarily desirable.

Finally, as clearly observed on the table, portfolios on the frontier include only six countries at most. Corner solutions—where a small number of assets (in this case, countries) dominate the portfolio—are indeed a well-documented outcome of mean-variance optimization. However, such allocations may be less appealing (or unrealistic) to practitioners, who prioritize diversification alongside environmental objectives. To address this, the next section introduces a method allowing investors to achieve greater diversification while considering both carbon *and* financial returns.

Nonetheless, for those less concerned about portfolio concentration, the mean-variance algorithm can still be adapted to incorporate both financial and carbon returns. For example, investors can analyze the mean-variance frontier for carbon returns while setting a predetermined level (or ”threshold”) of expected financial returns. Alternatively, they can examine the mean-variance frontier for financial returns while ensuring a specific level of decarbonization. A third option is to follow a similar framework to that proposed in Pedersen et al. (2021), where we would see ESG scores replaced by a moment of the carbon return distribution.

3.2 HRP optimization

As discussed above, mean-variance optimization often results in portfolios that are heavily concentrated in a few assets, which may not align with the preferences of some institutional investors seeking diversification. Moreover, the framework’s reliance on accurate estimates of expected returns and return covariances poses a significant challenge, particularly when dealing with carbon returns.

An alternative is the HRP portfolio optimization framework introduced by López de Prado (2016). Unlike mean-variance optimization, HRP focuses on risk diversification through hierarchical clustering and risk allocation. It avoids common issues such as concentration in a small number of assets and instability caused by noisy or ill-conditioned covariance matrices. HRP achieves this by combining principles of machine learning (hierarchical clustering) with risk-based asslocation.

3.2.1 Standard HRP

The standard HRP algorithm consists of three key blocks. First, it calculates a distance matrix based on return correlations, transforming the correlation structure into a hierarchical relationship among assets. Using this distance matrix, hierarchical clustering is performed to group assets into clusters, which are then arranged into a tree-like structure (dendrogram). Second, the covariance matrix is reordered into a quasi-diagonal form, ensuring that assets with stronger relationships are grouped together. Finally, risk is allocated hierarchically: risk is distributed within clusters and across clusters, using inverse-variance weighting. This hierarchical approach ensures that risk is evenly distributed across the portfolio, resulting in a diversified allocation that does not rely on precise return or covariance estimates. For a detailed discussion on the implementation of HRP, see López de Prado (2018).

HRP offers several advantages over traditional optimization techniques, making it particularly suitable for our application. By avoiding the inversion of the covariance matrix, HRP is robust to estimation errors and numerical instability, which is especially important when dealing with noisy data, as is the case for carbon returns due to their limited historical data. Moreover, HRP naturally promotes diversification by hierarchically allocating risk across clusters, reducing the risk of over-concentration in a few assets. This feature is particularly valuable for institutional investors managing large and complex portfolios. Additionally, HRP does not rely on explicit return forecasts, making it a practical and reliable tool in scenarios where expected returns are difficult to estimate—an important consideration given the challenges of forecasting carbon returns.

3.2.2 Modified HRP algorithm

Inspired by the HRP algorithm, we propose a similar approach that integrates clustering techniques with risk-based asslocation methods. A modification we make to the standard HRP algorithm is the use of expected shortfall instead of volatility as the measure of risk. Given that carbon risk is characterized by extreme but infrequent events with significant impact, we believe that focusing on tail risk management is crucial to climate risk management. Expected shortfall (ES), a tail-risk measure, is therefore more appropriate in capturing these extreme risks.

The risk contribution in the case of ES can be calculated as follows: for a portfolio with weights $\mathbf{w}$, the risk contribution of asset i is given by:

$$\text{Risk Contribution}_i^{\text{ES}_\alpha}(\mathbf{w}) = -w_i \mathbb{E}[r_i \mid \mathbf{w}'\mathbf{r} \leq \text{VaR}_\alpha(\mathbf{w})],$$

where $\text{VaR}_\alpha(\mathbf{w})$ is the value-at-risk at the confidence level $(1-\alpha)$ for the portfolio $\mathbf{w}$. The above equation can be approximated as:

$$\text{Risk Contribution}_i^{\text{ES}_\alpha}(\mathbf{w}) \approx -w_i \frac{1}{[\alpha T]} \sum_{k=1}^{[\alpha T]} r_{ki}^{\text{sort}}$$

where r_{ki} represents the returns of asset i in the k^{th} portfolio, with portfolio returns sorted in ascending order over time. The calculation can be performed using either historical data (e.g., for financial returns) or simulated data (e.g., for carbon returns).

Populating return distributions

Before optimizing using our modified HRP algorithm, we require robust return distributions for carbon returns. However the historical time series of carbon returns are relatively short, often limited to 30 to 40 annual observations for most countries, as shown in Section 2. Copula methods (see for instance, Embrechts et al. (2001) and McNeil et al. (2015)) allow the marginal distribution of each country's carbon returns to be modeled separately from their dependence structure, providing a flexible way to capture non-linear relationships and potential tail co-movements across countries.

The approach is well established in financial risk management, particularly in settings where available data are limited but understanding joint downside risks remains important (Cherubini et al. (2004) and Patton (2012) are two examples). In this context, the t-copula is especially suitable as it accommodates stronger joint tail dependence than the Gaussian copula, a relevant feature when modeling transition risks that may materialize simultaneously across jurisdictions. Simulating from the fitted copula therefore provides a richer distribution of plausible carbon return outcomes while preserving empirically observed cross-country dependence patterns (ie, preserving the original, empirical marginals). Importantly, the objective is not to generate new information but to obtain more stable estimates of the distribution implied by the available data, particularly in the tails where observations are sparse.

Our process therefore begins with estimating the parameters of the joint cumulative distribution function (CDF) of carbon returns. A t-copula is chosen due to its ability to model tail dependencies, which are essential when modeling extreme events. Via maximum likelihood, we estimate the rank correlation matrix (Spearman's rho) and the degrees of freedom for the t-copula. Once calibrated, we use these to simulate 10,000 scenarios of joint carbon returns.

To preserve the empirical characteristics of each country's individual carbon return distribution, we use non-parametric kernel density estimation (KDE) to construct the marginal distributions using the original samples. Then, for each country, the simulated joint distribution from the copula is transformed into the marginal distribution by applying the inverse CDF of the country's KDE-based distribution. This ensures that the simulated data retains the properties of the original marginal distributions, while providing a much larger sample size. The outcome is greater visibility into the tails of each country's carbon return distribution, which is allows us to better estimate expected shortfall (or "footprint shortfall", in our case).

Modified HRP optimization

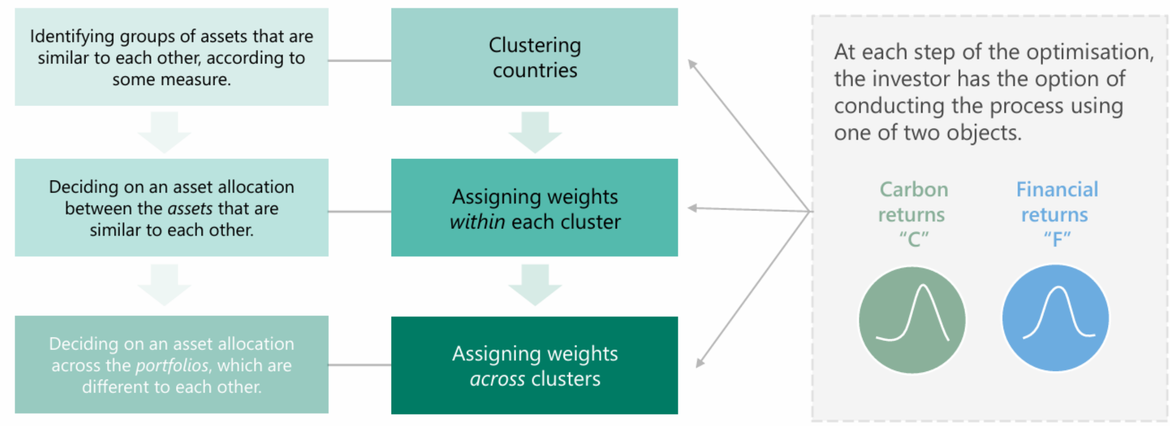
With the carbon return distributions established, the modified HRP algorithm can be applied to construct portfolios using various configurations. The optimization process involves three steps.

1. **Clustering countries:** Countries are first clustered based on their return correlations. Instead of the commonly used Pearson correlation matrix, we use rank-based correlation measures, as they are more robust in capturing tail dependence and are consistent with our above estimation of return distributions.

2. **Intra-cluster risk allocation:** Second, within each cluster, we assign weights to ensure that assets contribute equally to the cluster's overall ES.

3. **Inter-cluster risk allocation:** Weights are allocated across clusters by equalizing their ES contributions to the overall portfolio.

Figure 3: Optimization workflow



Note: The chart illustrates the workflow of our proposed optimization approach, inspired by the HRP algorithm, which integrates clustering techniques with risk-based allocation methods.

In each of the three steps, summarized in Figure 3, we can utilize either financial returns (labeled "F") or carbon returns ("C"). This results in eight distinct implementations, as listed in Table 2, each reflecting varying emphases on carbon and financial data. The configuration (F, F, F) represents the standard HRP algorithm, focusing exclusively on financial metrics without accounting for climate risk considerations. Conversely, (C, C, C) fully prioritizes the carbon return distribution. The remaining configurations represent combinations that balance both financial and carbon returns. The results of these eight optimization approaches are discussed in detail in the next section.

4 Optimization results

4.1 In-sample results

We first apply our optimization approach to the full set of historical data, in order to obtain all HRP-optimal portfolios. We proceed as follows:

1. Using all available historical returns, we cluster countries into J groups.

2. For each group $j \in \{1, \dots, J\}$, we construct an ES-parity portfolio with weight vector $\mathbf{w}_j$, where each element represents the weight assigned to a country within group j.

Table 2: Possible optimization implementations

	Clustering	**Intra-cluster weight**	**Inter-cluster weight**
1	F	F	F
2	F	F	C
3	F	C	F
4	C	F	F
5	F	C	C
6	C	F	C
7	C	C	F
8	C	C	C

Note: This table lists the eight different optimization approaches that can be implemented within our framework. "F" = financial returns, "C" = carbon returns. The second to fourth columns correspond to the three steps in the optimization workflow (see Figure 3).

3. We compute the historical return series for each group and solve a second ES-parity optimization to obtain the optimal weight assigned to group j in the final portfolio, denoted by $w_j^{(p)}$.

4. The final ex-ante allocation to the countries within group j is given by $w_j^{(p)}\,\mathbf{w}_j$. These weights are then used to compute financial and carbon returns, which serve as the basis for evaluating in-sample performance.

We examine the results from the two extreme cases of this in-sample optimization: (F,F,F) and (C,C,C), which correspond to the top and bottom rows of Table 2. For (F,F,F), countries are clustered, and weights are assigned both within and across clusters based solely on financial returns. Similarly, for (C,C,C), clustering and weight assignments are all based on carbon returns.

The clustering of countries by financial or carbon returns reveals some interesting contrasts, as illustrated in Figure 4. When clustered by financial returns, some expected patterns emerge. European countries generally group together; however, non-euro countries such as Norway, Sweden, and Switzerland tend to cluster separately (albeit not far) from Euro Area countries. Small, open economies—including Australia, New Zealand, Canada, Singapore, South Korea, and the United Kingdom—are grouped into the same cluster, some of which are commodity currencies. Additionally, there is a distinct cluster comprising so-called safe-haven governments, which includes Japan and the United States. In contrast, clustering based on carbon returns results in very different groupings, which are harder to rationalize without background on each country's environmental footprint. For instance, Denmark, Finland, and Sweden are grouped together, likely reflecting their significant and early reductions in carbon emissions. These countries have experienced a declining trend in carbon emissions since the 2000s, much earlier than many other nations (see Figure 11).

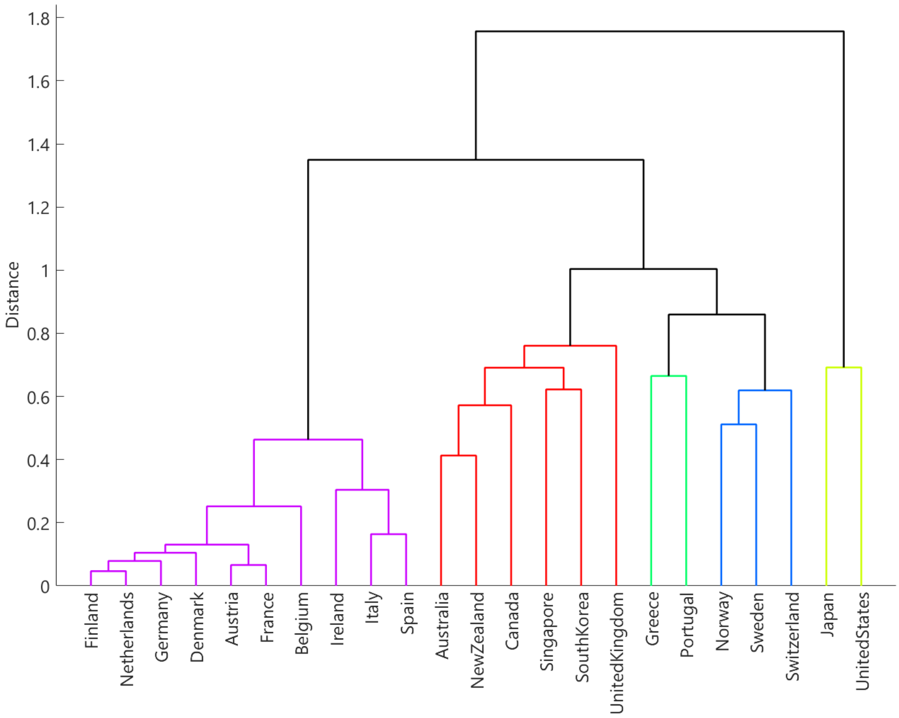
Consistently, portfolio weights for individual countries vary significantly when optimization is based solely on financial returns compared to carbon returns (Figure 5). For (F,F,F), it is unsurprising that the United States, Japan, and Switzerland hold the top three positions in portfolio share, partly reflecting their contributions to a lower expected shortfall. Together, these three countries account for nearly 50% of the total portfolio. In contrast, when optimization is based on carbon returns, the weights in (C,C,C) portfolio are more evenly distributed. The highest allocations are assigned to Norway, Australia, Sweden, and the United States, collectively accounting for less than 25% of the portfolio.

To evaluate whether our algorithm achieves its objective, we compare the financial return and carbon return profiles of portfolios (F,F,F) and (C,C,C). Figure 6 depicts the distributions of financial returns (left-hand panel) and carbon returns (right-hand panel) for the two portfolios. As desired, the (F,F,F) portfolio (shades of yellow) exhibits a less pronounced left tail for financial returns but a more pronounced left tail for carbon returns. In contrast, the (C,C,C) portfolio (shades of green) shows the opposite pattern.

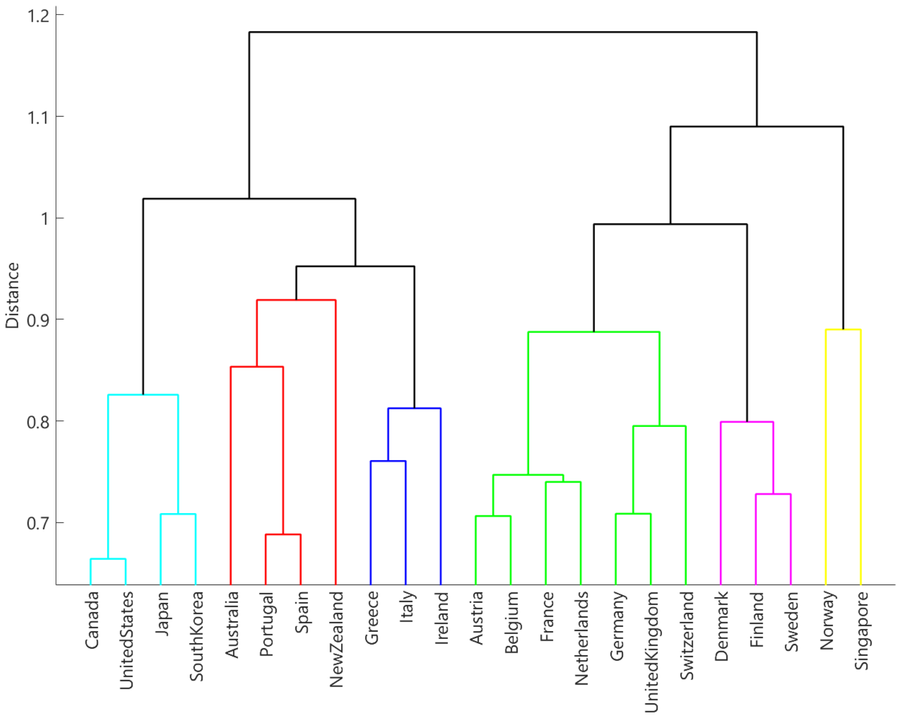
We compute additional summary statistics for the two portfolios' returns, and present them in Table 3. While the two portfolios are similar in terms of average financial and carbon returns, the (F,F,F) portfolio demonstrates a lower volatility of financial returns, whereas the (C,C,C) portfolio exhibits lower volatility in carbon returns (be reminded that variance is *not* part of our objective function). Furthermore, when focusing on downside risk—measured by the 5th, 2.5th, and 1st percentiles, as well as the probability of negative returns—the (F,F,F) portfolio shows lower downside risk for financial returns, while the (C,C,C) portfolio shows lower downside risk for carbon returns. These findings confirm that the portfolios prioritizing financial (or carbon) returns are more effective at mitigating downside risk in their respective domains, consistent with the intended design of our algorithm.

Figure 4: Clustering countries by returns' correlation

(a) By financial returns



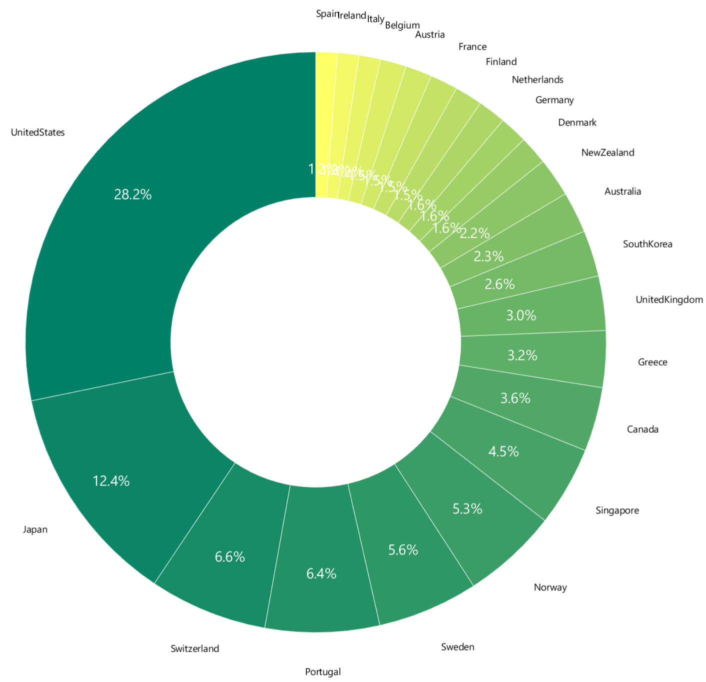
(b) By carbon returns



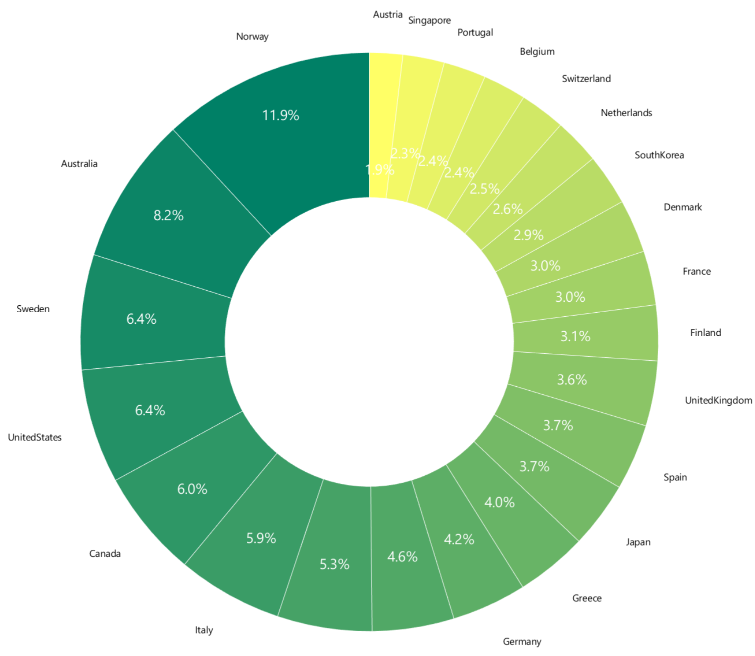
Note: The top panel shows the dendrogram when countries are clustered based on financial returns, while the bottom panel shows the dendrogram for clustering based on carbon returns.

Figure 5: Weights of individual countries for in-sample optimized portfolios

(a) (F,F,F) optimization



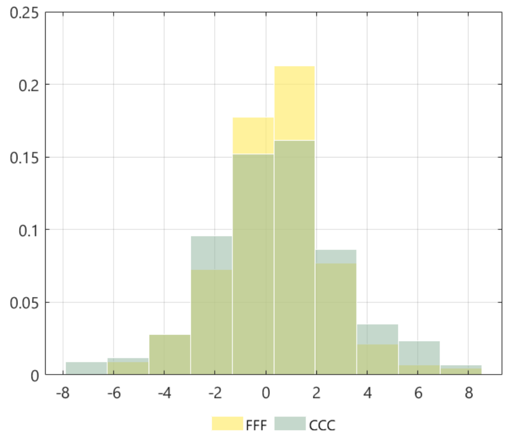
(b) (C,C,C) optimization



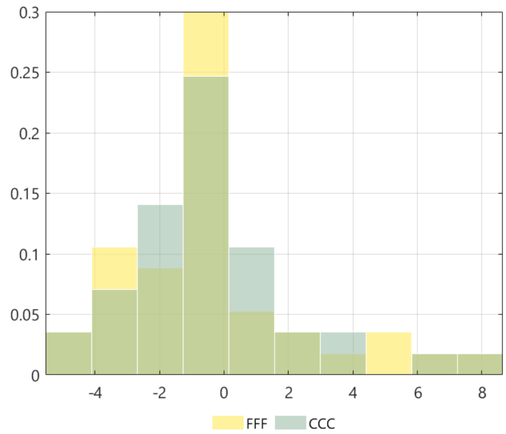
Note: The top pie chart shows the weight of individual countries in the (F,F,F) optimized portfolio, while bottom pie chart shows the weight of individual countries in the (C,C,C) optimized portfolio.

Figure 6: Histogram of return distribution for in-sample optimized portfolios

(a) Financial return



(b) Carbon return



Note: The left-hand panel displays the histogram of financial returns for the (F,F,F) and (C,C,C) optimized portfolios, while the right-hand panel presents the histogram of carbon returns for the same portfolios. The yellow bars represent the (F,F,F) optimized portfolio, and the green bars represent the (C,C,C) optimized portfolio.

Table 3: Return summary statistics for in-sample optimized portfolios

	Financial returns		Carbon returns	
	(F,F,F)	(C,C,C)	(F,F,F)	(C,C,C)
Mean return	0.4%	0.4%	-0.4%	-0.4%
Volatility	2.0%	2.7%	2.9%	2.7%
95% Value-at-Risk	-3.2%	-3.9%	-4.2%	-4.2%
97.5% Value-at-Risk	-4.1%	-5.8%	-5.0%	-4.8%
99% Value-at-Risk	-4.7%	-6.4%	-5.5%	-5.1%
Prob. of negative returns	41.4%	44.1%	70.0%	67.5%

Note: The table presents summary statistics for financial and carbon returns of the (F,F,F) and (C,C,C) optimized portfolios, based on in-sample optimization.

4.2 Out-of-sample results

We now move on to evaluate the out-of-sample performance of our algorithm to assess its robustness and predictive value. We begin the exercise in January 2008, using financial return data through December 2007, and carbon emissions data through 2006 to account for the 1-year reporting lag. We then conduct the following steps iteratively:

1. Using all historical returns available up to month t, we cluster countries into J groups. In the case of carbon returns, a t-copula is used to generate the more populated version of the multivariate return distribution (see section 3.2.2) before usage.

2. For each group $j \in \{1, \ldots, J\}$, we construct an ex-ante ES-parity portfolio with weight vector $\mathbf{w}_j$, where each element represents the weight assigned to a country within group j.

3. We compute the historical return series for each group and solve a second ES-parity optimization to obtain the ex-ante portfolio weight for group j, denoted $w_j^{(p)}$.

4. The final ex-ante allocation to countries within group j is $w_j^{(p)}\mathbf{w}_j$. These weights are used to compute one-month-ahead returns, which form the basis for evaluating out-of-sample performance.

5. We roll the window forward by one month and repeat steps 1–4. This procedure adds one financial return observation each month and one carbon-return observation at each year-end.

We begin by comparing the (F,F,F)–optimized portfolio and the (C,C,C)–optimized portfolio. Rather than analyzing the detailed portfolio weights, which evolve over time, we focus on the return performance of these two portfolios (in both F and C spaces). The financial returns of both portfolios are of similar magnitude, with the (C,C,C) portfolio exhibiting slightly higher volatility but also higher returns overall (Figure 7, left-hand panel). Regarding financial tail risk, the results align with the in-sample findings: the return distribution histogram (Figure 8, left-hand panel) indicates slightly fatter tails for the (C,C,C) portfolio. This observation is further supported by the cumulative distribution (Figure 9, left-hand panel), where the (C,C,C) portfolio consistently has higher cumulative probability–a fatter tail–than the (F,F,F) portfolio for negative financial returns.

Turning to carbon returns, the (C,C,C)-optimized portfolio demonstrates superior performance. It is less likely to experience significant return drops (country-level rises in carbon emissions) and more likely to achieve substantial gains (i.e., decarbonization), resulting in much higher cumulative carbon returns (Figure 7, right-hand panel). With regard to tail risks, both the carbon return distribution histogram and cumulative distribution function indicate that the (C,C,C) portfolio has lower left-tail risk (Figure 8 and Figure 9, right-hand panels)–a desirable property when prioritizing carbon emissions in portfolio optimization.

A summary comparison of return statistics for the two portfolios are presented in Table 4 and reinforce our findings. The (F,F,F) portfolio exhibits lower volatility and reduced downside risk in terms of ex-post financial returns. Conversely, when focusing on carbon returns, the (C,C,C) portfolio outperforms (F,F,F) along these dimensions.

We also compare the performance of the remaining portfolios. For brevity, We report report only the 97.5% Value-at-Risk (Figure 10). Broadly, portfolios that place greater weight on financial returns (i.e., with more "F"s) exhibit lower financial tail risk, whereas portfolios that emphasize carbon returns (i.e., with more "C"s) display lower carbon tail risk.

Taken together, these results indicate that our algorithm performs well both in-sample and out-of-sample. Portfolios designed to mitigate the downside risk of financial (or carbon) returns achieve this objective. Investors can therefore choose among the eight possible implementations based on more subjective criteria, such as those best aligning with their policy objectives and/or investment preferences.For example, investors primarily concerned with financial returns may prefer portfolios with more "F"s (e.g., F, F, F), whereas those focused on carbon returns may prefer portfolios with more "C"s (e.g., C, C, C).

Figure 7: Cumulative return for out-of-sample optimized portfolios

(a) Financial return

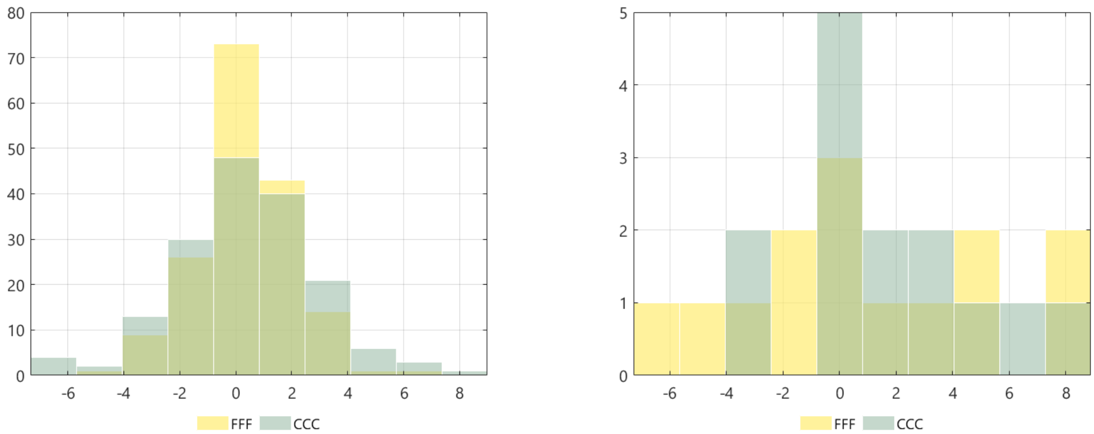
(b) Carbon return

Note: The left-hand panel displays the out-of-sample cumulative financial returns for the (F,F,F) and (C,C,C) optimized portfolios, while the right-hand panel presents the out-of-sample cumulative carbon returns for the same portfolios. The yellow lines represent the (F,F,F) optimized portfolio, and the green lines represent the (C,C,C) optimized portfolio.

Figure 8: Histogram of return distribution for out-of-sample optimized portfolios

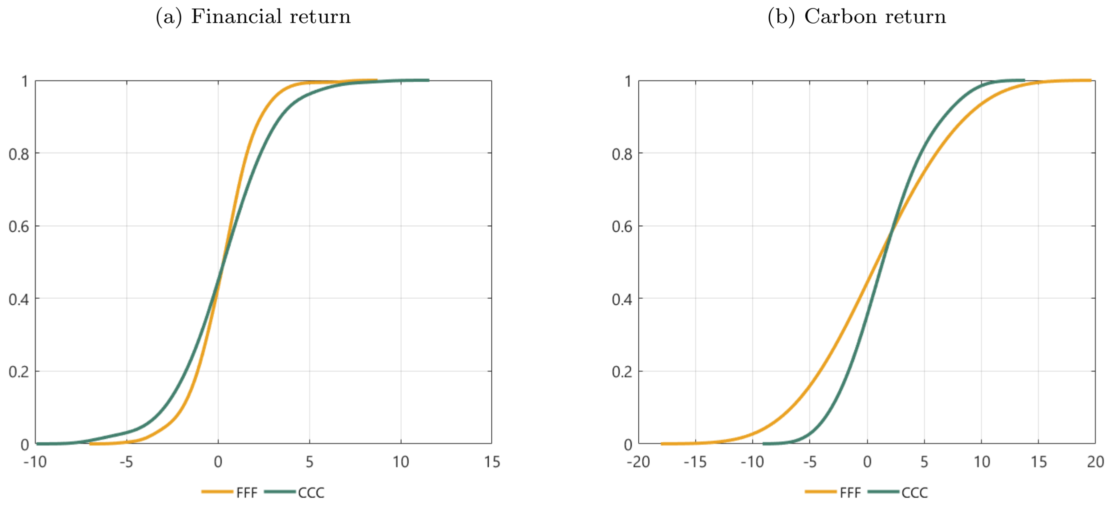
(a) Financial return

(b) Carbon return



Note: The left-hand panel displays the out-of-sample histogram of financial returns for the (F,F,F) and (C,C,C) optimized portfolios, while the right-hand panel presents the out-of-sample histogram of carbon returns for the same portfolios. The yellow bars represent the (F,F,F) optimized portfolio, and the green bars represent the (C,C,C) optimized portfolio.

Figure 9: Cumulative return distribution function for out-of-sample optimized portfolios



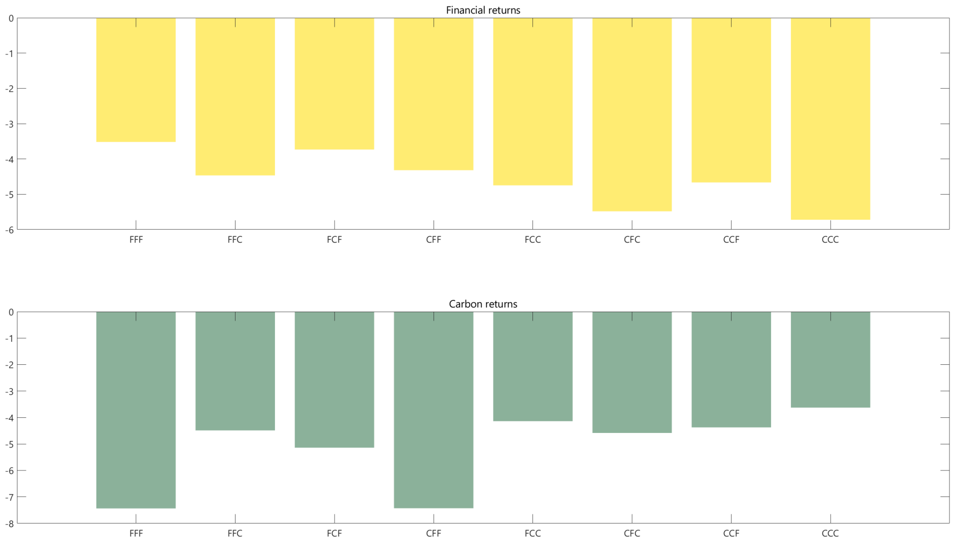
Note: The left-hand panel displays the out-of-sample cumulative financial return distribution for the (F,F,F) and (C,C,C) optimized portfolios, while the right-hand panel presents the out-of-sample cumulative carbon return distribution for the same portfolios. The yellow lines represent the (F,F,F) optimized portfolio, and the green lines represent the (C,C,C) optimized portfolio.

Table 4: Return summary statistics for out-of-sample optimized portfolios

	Financial returns		Carbon returns	
	(F,F,F)	(C,C,C)	(F,F,F)	(C,C,C)
Mean return	0.2%	0.3%	0.9%	1.6%
Volatility	1.7%	2.5%	4.6%	3.2%
95% Value-at-Risk	-2.6%	-3.8%	-7.2%	-3.6%
97.5% Value-at-Risk	-3.5%	-5.7%	-7.4%	-3.6%
99% Value-at-Risk	-3.9%	-6.9%	-7.4%	-3.6%
Prob. of negative returns	42.9%	44.6%	50.0%	35.7%

Note: The table presents summary statistics for financial and carbon returns of the (F,F,F) and (C,C,C) optimized portfolios, based on in-sample optimization.

Figure 10: 97.5% Value-at-Risk (%) for different out-of-sample optimized portfolios



Note: Note: The figure reports the 97.5% Value-at-Risk the out-of-sample optimized portfolios. The optimization methods are defined in Table 2. The top panel presents financial returns Value-at-Risk, and the bottom panel presents carbon returns Value-at-Risk.

5 Concluding remarks

This paper proposes a practical framework for decarbonizing sovereign bond portfolios that integrates financial and environmental objectives through the concept of *carbon returns*. By defining carbon returns as changes in a country's carbon footprint, we move beyond carbon budget approaches that treat sovereign footprints as deterministic inputs (whether based on the latest historical observation or on a single scenario path). Instead, we treat transition outcomes as *random variables*, consistent with the notion that transition risk is uncertain, state-dependent (hence the talk of pathways), and potentially characterized by tail events.

Methodologically, we adapt the Hierarchical Risk Parity approach to sovereign portfolios and emphasize tail-risk management by using expected shortfall as the risk metric. The framework combines clustering and risk-based allocation in three steps and accommodates either financial or carbon return inputs at each stage, generating eight implementable portfolio variants. To address limited historical information on carbon returns, we enrich the return space using copula-based simulations and rely on rank-based dependence measures in clustering to better capture non-linear dependence structures. This is particularly relevant for transition risk, where co-movements may strengthen in adverse states of the world.

Empirically, using developed market sovereign bonds, we show that the framework delivers the intended risk reductions both in-sample and out-of-sample. Portfolios that place greater emphasis on financial returns deliver smaller downside risk in the financial space, while portfolios that place greater emphasis on carbon returns deliver smaller downside risk in carbon space. More broadly, portfolios with more “F” inputs tend to exhibit lower financial tail risk, whereas portfolios with more “C” inputs tend to exhibit lower carbon tail risk. These results confirm that the framework can be used to design sovereign portfolios that are robust, diversified, and explicitly targeted to the investor’s chosen notion of downside risk—financial, environmental, or a combination.

For central bank reserve managers and other investors with strong constraints on safety, liquidity, and return, the framework offers a tractable way to embed climate objectives without facing mean-variance corner solutions or complex benchmark engineering (i.e., using tracking-error based optimization methods). There are three benefits to these type of investor: (i) the HRP structure supports diversified allocations and reduces sensitivity to noisy estimates; (ii) the expected-shortfall objective focuses attention on highly adverse transition outcomes; and (iii) the eight configurations provide a menu of choices to map different mandates into portfolios.

Our findings also point to two other practical priorities. First, climate-aligned portfolio methodologies and benchmark frameworks could be strengthened by explicitly distinguishing between (a) deterministic alignment constraints and (b) a distributional, risk-based treatment of transition uncertainty. Second, the carbon return framework offers a bridge between imperfect measurement and actionable portfolio construction. Debates about carbon footprints often hinge on *levels*—production vs. consumption accounting, land-use adjustments, or the choice of denominator (GDP, population). Those choices can materially change country rankings in levels and make portfolio outcomes sensitive to methodological conventions. By shifting attention from levels to *changes* (i.e., decarbonization rates and their distributions), carbon returns reduce the extent to which results are driven by a single accounting convention and instead focus portfolio design on an economically intuitive object: whether and how reliably countries are reducing (or have been reducing) emissions over time.

Several avenues for future work remain. First, extending the analysis to a broader set of countries, maturities, currencies, and asset classes would help generalize the results;

after all, even though our framework is applied to advanced economy sovereign bonds, it readily extends to other asset classes (eg, corporates) and jurisdictions (eg, emerging markets). Second, incorporating forward-looking information more systematically—for both carbon and financial returns— for instance, by using scenario narratives coupled with return distributions, could enhance the utility of each portfolio alternative. Third, integrating portfolio dimensions such as liquidity and transaction costs would enhance the practical applicability of these methods. Finally, studying the interaction between portfolio construction and policy dynamics would shed light on how investor preferences and allocation choices can support credible decarbonization pathways.

Overall, the evidence suggests that investors can construct diversified sovereign portfolios that materially reduce downside risk in either financial or carbon dimensions while maintaining robust performance. By embracing uncertainty in emissions outcomes and treating transition risk as a distributional, tail-sensitive object, our framework offers a tractable, transparent, and adaptable path toward decarbonizing sovereign bond investments.

Appendix

A Different types of carbon emissions

Measuring a country's carbon emissions can be done through various methods, each with its own approach and focus. NGFS (2024) classifies emissions to three types based on emission allocations.

The most commonly applied allocation is *production-based carbon emissions*. It measures carbon emissions from all goods and services produced within a country territory. The goods and services can be consumed domestically or overseas via export. While this metric corresponds to the UN Framework Convention for Climate Change, and is recommended by the Kyoto Protocol, it favours countries, mostly advanced economies, that have relocated their carbon-heavy industries.

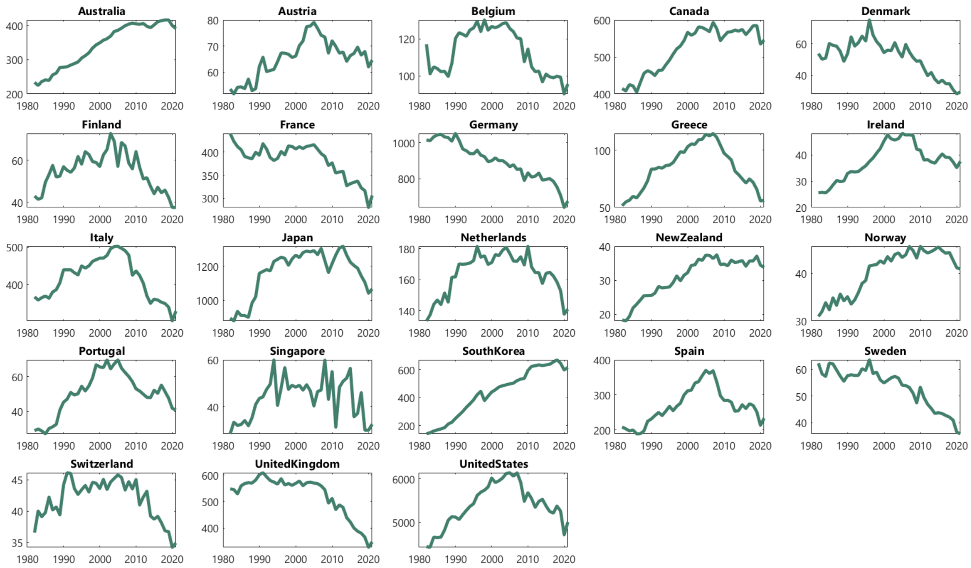
Production emissions are reported on two kinds: including and excluding emissions through land-use change. *Production-based emissions including land-use change* is an adjusted version of production-based emissions. The adjustment reflects the impact on carbon emissions from activities that alter the way land is used, including deforestation, reforestation, afforestation, and changes in agricultural practices. These activities can significantly impact the amount of carbon dioxide and other greenhouse gases released or absorbed by the land. Accounting for these land-use-related emissions and removals helps provide a comprehensive view of a country's total emissions from activities within its borders. While these emissions often do not play a significant role in a country's emissions, it can be non-negligible for certain countries. For example, countries like Canada and Finland have significant forest cover, which can act as carbon sinks. However, changes in land-use can reverse this. That said, according to the Partnership for Carbon Accounting Financials (PCFA) standard, data providers and climate experts disagree on the methods collecting emissions related to land use.

The other allocation type is *consumption-based carbon emissions*. It corresponds to carbon emissions from all goods and services consumed within a country's territory. It includes emissions from goods and services produced domestically and imports. The data is calculated by adjusting production-based emissions adjusted for trade: consumption-based emissions equal production-based emissions, minus emissions embedded in exports, plus emissions embedded in imports. The consumption-based emissions data tend to have narrower coverage than production-based emissions data. It is documented by few countries and is often estimated by data providers. A challenge is lack of high-quality trade data.

A third type of allocation is *government emissions.* Those are emissions from government activities, and therefore under direct control of a national government. This type of emissions is not included in our analysis as it is not a one of the metrics recommended by PCFA to measure issuer-level carbon emissions.

B Additional figures and tables

Figure 11: Production-based carbon emissions (in million tonnes)



Note: The chart shows production-based carbon emissions (in million tonnes) across countries, spanning the period from 1982 to 2021.

Table 5: Comparison of carbon returns based on different types of carbon emissions

Country	Prod vs Prod, land	Prod vs Cons	Prod, land vs Cons
Australia	0.06	0.80	0.06
Austria	0.90	0.70	0.62
Belgium	0.97	0.17	0.19
Canada	0.30	0.27	0.72
Denmark	0.80	0.24	0.37
Finland	0.61	0.78	0.40
France	0.87	0.98	0.87
Germany	0.84	0.90	0.80
Greece	0.41	0.65	0.66
Ireland	0.32	0.70	0.85
Italy	0.63	0.80	0.91
Japan	0.96	0.78	0.50
Netherlands	0.85	0.85	0.85
New Zealand	0.02	0.53	0.04
Norway	0.60	0.60	0.84
Portugal	0.45	0.40	0.39
Singapore	0.97	0.94	0.94
Korea	1.00	0.80	0.78
Spain	0.61	0.79	0.91
Sweden	0.54	0.61	1.00
Switzerland	0.80	0.88	0.85
United Kingdom	0.97	0.58	0.60
United States	0.89	0.62	0.49

Note: The table reports p-values from the Mann-Whitney U test (Mann and Whitney (1947)) comparing carbon returns calculated based on different types of carbon emissions. P-values below 0.05 are highlighted in red. Prod: production-based emissions; Prod, land: production-based emissions including land-use change; Cons: consumption-based emissions.

Table 6: Comparison of production-based carbon returns from different sources of data

Country	GCP vs UN	GCP vs OECD	GCP vs S&P
Australia	0.77	0.88	0.09
Austria	0.98	0.96	0.86
Belgium	0.99	0.99	0.23
Canada	0.85	1.00	0.09
Denmark	0.95	0.98	0.77
Finland	0.99	0.99	0.76
France	0.85	0.90	0.62
Germany	0.90	0.94	0.76
Greece	0.83	0.93	0.33
Ireland	0.79	0.81	0.13
Italy	0.76	0.94	0.31
Japan	0.93	0.94	0.42
Netherland	0.96	0.95	0.45
New Zealand	0.81	0.98	0.20
Norway	0.89	0.99	0.25
Portugal	0.97	0.98	0.86
Singapore			
Korea		0.96	0.16
Spain	0.70	0.90	0.16
Sweden	0.86	0.86	0.46
Switzerland	0.87	0.77	0.54
United Kingdom	1.00	0.84	0.29
United States	0.88	1.00	0.14

Note: The table reports p-values from the Mann-Whitney U test (Mann and Whitney (1947)) comparing production-based carbon returns calculated based on different data sources of carbon emissions. P-values below 0.05 are highlighted in red. Empty cells indicate insufficient numbers of observation.

C Data on financial returns

For financial returns data, we begin by collecting monthly sovereign yield data from Bloomberg. We start the price data from 2001 and includes 23 advanced economies in our sample, ensuring a balanced panel. We then apply a standard bootstrapping procedure to construct zero-coupon yield curves from the observed data. This method eliminates coupon effects from the securities, simplifying return calculations. The procedure is systematically applied across all countries in the sample.

To keep the analysis tractable, each country is represented by a single security on the curve. We select the 5-year maturity, as it is widely regarded as a liquid benchmark in many sovereign yield curves. To compute 5-year constant-maturity monthly returns, we price 5-year and 4-year 11-month (approximately 4.92-year) zero-coupon bonds for each country at monthly intervals. The return of the 5-year bond for country i at month t is calculated as: $R_{t,i}^{5\text{yr}} = \left(\frac{P_{t,i}^{4.92\text{yr}}}{P_{t,i}^{5\text{yr}}}\right) - 1$, where $P_{t,i}^{5(4.92)\text{ yr}}$ represents the price of the 5 (4.92)-year zero-coupon bond prevailing at month t. This approach enables the creation of a monthly return time series that captures changes in the underlying yield curve for each country.

Since the analysis is conducted from the perspective of a global investor, all local currency returns are converted into a single numeraire. We adopt the USD as the common currency, consistent with standard practices in global sovereign bond indices. To achieve this, we collect monthly exchange rate data from Bloomberg for each currency in the sample. The exchange rate adjustment is then applied to the local currency bond returns, ensuring that currency gains or losses are accounted for. This adjustment provides a more accurate representation of the potential gains or losses experienced by a globally diversified (and unhedged) investor.

Table 7 summarizes the return and risk characteristics of monthly returns in USD on zero-coupon, 5-year, constant-maturity government bonds for our 23 countries over the period from September 2001 to May 2023.

Table 7: Summary statistics of monthly financial returns by country

Country	Mean Return	Volatility	Value-at-Risk (97.5%)	Expected Shortfall (97.5%)
Australia	0.51%	3.40%	-7.60%	-8.91%
Austria	0.35%	2.86%	-6.00%	-7.12%
Belgium	0.37%	2.91%	-5.91%	-7.38%
Canada	0.37%	2.54%	-5.13%	-6.39%
Denmark	0.34%	2.88%	-5.51%	-7.19%
Finland	0.35%	2.81%	-5.51%	-6.89%
France	0.35%	2.84%	-5.84%	-6.93%
Germany	0.33%	2.81%	-5.60%	-6.78%
Greece	1.29%	14.93%	-20.97%	-27.49%
Ireland	0.46%	3.76%	-6.94%	-11.03%
Italy	0.46%	3.45%	-8.90%	-9.56%
Japan	0.06%	2.79%	-5.85%	-6.90%
Netherlands	0.34%	2.82%	-5.76%	-6.92%
New Zealand	0.60%	3.67%	-8.23%	-10.03%
Norway	0.29%	3.38%	-6.30%	-7.50%
Portugal	0.56%	4.55%	-9.03%	-13.04%
Singapore	0.35%	2.01%	-4.24%	-5.19%
South Korea	0.42%	3.65%	-8.07%	-9.50%
Spain	0.43%	3.37%	-6.79%	-9.60%
Sweden	0.29%	3.21%	-6.00%	-6.93%
Switzerland	0.43%	3.02%	-5.29%	-6.70%
United Kingdom	0.24%	2.67%	-5.85%	-7.87%
United States	0.30%	1.34%	-2.74%	-3.23%

Note: The table summarizes the statistics of financial returns calculated using zero-coupon yield curve data from September 2001 to May 2023.

References

Angelini, Paolo, "Portfolio decarbonisation strategies: questions and suggestions," *Bank of Italy Occasional Paper*, 2024, (840).

Carstens, Agustín, "Balancing reserve management objectives in the current economic and financial environment," Speech at the Regional High-level Reserve Management Conference, Riyadh, Saudi Arabia April 2024.

Cheng, Gong, Eric Jondeau, and Benoît Mojon, "Building portfolios of sovereign securities with decreasing carbon footprints," *BIS Working Papers No 1038*, 2022.

Cherubini, Umberto, Elisa Luciano, and Walter Vecchiato, *Copula Methods in Finance*, Wiley, 2004.

Crippa, Monica, Diego Guizzardi, Federico Pagani, Manjola Banja, Marilena Muntean, Edwin Schaaf, Wolfgang Becker, Fabio Monforti-Ferrario, Roberta Quadrelli, A Risquez Martin et al., "GHG emissions of all world countries," *Publications Office of the European Union, Luxembourg*, 2023, *10*, 953322.

de Bandt, Olivier, Laura-Chloé Kuntz, Nora Pankratz, Fulvio Pegoraro, Haakon Solheim, Gregory Sutton, Azusa Takeyama, and Fan Dora Xia, "The effects of climate change-related risks on banks: A literature review," *Journal of Economic Surveys*, 2025, *39* (4), 1553–1594.

Embrechts, Paul, Alexander J. McNeil, and Daniel Straumann, "Modelling Dependence with Copulas and Applications to Risk Management," in Svetlozar T. Rachev, ed., *Handbook of Heavy Tailed Distributions in Finance*, Elsevier, 2001, pp. 329–384.

Fender, Ingo, Mike McMorrow, and Omar Zulaica, "Sustainable Management of Central Banks' Foreign Exchange (FX) Reserves," *The INSPIRE Sustainable Central Banking Toolbox, Policy Briefing Paper 06, Grantham Research Institute on Climate Change and the Environment*, 2022.

_ , _ , Vahe Sahakyan, and Omar Zulaica, "Reserve management and sustainability: the case for green bonds," *BIS Working Paper No. 849*, 2020.

Guenedal, Théo Le and Thierry Roncalli, "Portfolio construction with climate risk measures," *Available at SSRN 3999971*, 2022.

HSBC, "HSBC Reserve Management Trends: Global instability tests Central Banks' resources," *HSBC Market & Regulatory Insights*, 2025.

López de Prado, Marcos, "Building diversified portfolios that outperform out-of-sample," *Journal of Portfolio Management*, 2016.

_ , *Advances in financial machine learning*, John Wiley & Sons, 2018.

Mann, Henry B and Donald R Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *The annals of mathematical statistics*, 1947, pp. 50–60.

Markowitz, Harry, "Portfolio Selection," *The Journal of Finance*, 1952, *7* (1), 77–91.

McNeil, Alexander J., Rüdiger Frey, and Paul Embrechts, *Quantitative Risk Management: Concepts, Techniques and Tools*, revised edition ed., Princeton University Press, 2015.

NGFS, "NGFS Climate Scenarios for central banks and supervisors," *Available in the NGFS website*, 2022.

_ , "Climate-related risks and impact in sovereign investments: data, metrics and implementation issues for central banks," *Network for Greening the Financial System Technical Document*, 2024.

Patton, Andrew J., "A Review of Copula Models for Economic Time Series," *Journal of Multivariate Analysis*, 2012, *110*, 4–18.

Pedersen, Lasse Heje, Shaun Fitzgibbons, and Lukasz Pomorski, "Responsible investing: The ESG-efficient frontier," *Journal of financial economics*, 2021, *142* (2), 572–597.

Penninga, Olaf and Casper Zomerdijk, "Mind the risk – smart decarbonization in government bond portfolios," *Rebeco Quant Chart*, 2024.

Scatigna, Michela, Fan Dora Xia, Anna Zabai, and Omar Zulaica, "Achievements and challenges in ESG markets," *BIS Quarterly Review, December*, 2021.

Schwaiger, Katharina, Riyadh Ali, and Andrew Ang, "Net-Zero Sovereign Bond Portfolios," *Evolving Practices in Public Investment Management*, 2023, pp. 167–184.

Swinkels, Laurens, Jan Anton van Zanten, Bruno Rein, and Rikkert Scholten, "Financing the sustainable development goals: exploring the role of government bond investors," *Financial Analysts Journal*, 2025, pp. 1–15.

World Bank, "Reserve Management Survey Report 2025: insights into public asset management - The Fifth Edition," *World Bank Group*, 2025.

Xia, Dora and Omar Zulaica, "The term structure of carbon premia," *Proceedings of the Eighth Public Investors Conference*, 2022.