Empirical analysis of the trade-offs among risk, return, and climate risk in multi-criteria portfolio optimization

Sebastian Utz[1,2,3] · **Ralph E. Steuer**[4]

Abstract
This paper contains an empirical analysis that studies trade-offs among risk, return, and climate risk in asset management. Using a multi-criteria optimization approach to generate nondominated portfolios in a tri-criterion context, we document how it is possible in a portfolio to reduce climate risk substantially by allowing expected return to be reduced only slightly. The empirical tests conducted use the sample of stocks that were in the S&P 500 over the period 2001–2020. In demonstrating the versatility of our approach, six different linear measures of climate risk are employed.

Keywords Climate risk · Portfolio selection · Nondominated surface · Trade-offs

1 Introduction

Multi-criteria decision-making (MCDM) and its application to finance is a widely studied area of operations research (Steuer, 1986; Steuer & Na, 2003; Zopounidis & Doumpos, 2002). The famous problem in this area is the problem of portfolio selection as developed by Markowitz (1952, 1956, 1959). As formulated by Markowitz, portfolio selection is a bi-criterion problem with one objective being quadratic, to minimize risk, and the other being linear, to maximize return (i.e., expected return). The purpose of the formulation is to compute the set of all nondominated criterion vectors, which when graphed takes on the shape of a frontier which, in deference to finance, will be called the *efficient frontier*. Since an efficient frontier shows, and only shows, all criterion vector candidates for optimality, a decision maker finds his or her optimal portfolio by selecting his or her most preferred point on the efficient frontier. This is the well-studied case for bi-criterion risk–return, or

✉ Ralph E. Steuer
rsteuer@uga.edu

Sebastian Utz
sebastian.utz@wiwi.uni-augsburg.de

1 Department of Climate Finance, University of Augsburg, Universitaetsstr. 16, 86159 Augsburg, Germany

2 Centre for Climate Resilience, University of Augsburg, Universitaetsstr. 12, 86159 Augsburg, Germany

3 Sustainable Finance Research Platform, 10117 Berlin, Germany

4 Department of Finance, University of Georgia, Athens, GA 30602, USA

mean–variance (M–V), portfolio selection. However, a recent challenge is how to include a third criterion (i.e., an additional objective) thus causing the problem of portfolio selection to become a tri-criterion one. The additional objectives most often mentioned in this regard relate to sustainability as in the case in this paper with our concerns being about climate risk. Thus, with climate risk, our interest in portfolio selection is not just to minimize risk and to maximize return but to also minimize climate risk for which we use in our testing six different linear measures.

Hirschberger et al. (2013) present an MCDM algorithm that is able to compute the set of all nondominated criterion vectors of a tri-criterion portfolio selection problem with one quadratic and two linear criteria, which is what is needed in this paper. In this case, when graphed, the set of all nondominated criterion vectors forms what would be called in finance, and what we call in this paper, an *efficient surface*. Again, one's optimal portfolio is one's most preferred of the nondominated criterion vectors generated. But here, it is on the problem's efficient surface as opposed to being on an efficient frontier as it is in a bi-criterion case. While it is harder to compute an efficient surface than an efficient frontier, the big difference between tri-criterion and bi-criterion portfolio selection is that it is far more difficult without cognitive assistance to identify one's most preferred point on an efficient surface. A recent contribution that provides cognitive assistance that benefits the results of this paper is the *non-contour (NC)-efficient fronts* method described in Steuer and Utz (2023).

The importance in studying climate risk as a third objective in portfolio selection derives from at least four sources:

(1) The topic is important for a large proportion of investors since the sustainable investment market has grown rapidly over the last decade.[1]
(2) There continue to be differences of opinion in academia about the relationships between the financial and third criterion portfolio outcomes in the sustainable investment industry (Cornell, 2021; López Prol & Kim, 2022).
(3) Sustainable investing is considered as a tool for curbing negative externalities (Hong et al., 2023; Pástor et al., 2021; Riedl & Smeets, 2017).
(4) Climate change, with extreme weather events and average temperature increases, is now an important business and international agenda item.

From a theoretical point of view, sustainable investing contrasts with traditional investing in which investors care about only two characteristics (risk and return) as they make their investment decisions. Sustainable investors, while not ignoring risk and return, strive to put their money to work in ways that are consistent with their values if at all possible. Moreover, sustainable investors tend to remain committed to their sustainable holdings even when they perform poorly (Webley et al., 2001) indicating that they are willing to sacrifice proportions of their financial returns in order to support sustainability efforts (Hofmann et al., 2008; Dorfleitner & Utz, 2014; Mackenzie & Lewis, 1999; Pasewark & Riley, 2010). This is in accordance with the viewpoint that over and above the utility earned by the risk-return performance of a portfolio, sustainable investors gain further utility from the sustainability properties of a portfolio. In this way, as found by Levitt and List (2007), sustainable investors are usually willing to engage in trades between the utilities of wealth and morality.

In this paper, we conduct an empirical study that shows the trade-offs among risk, return, and climate risk for a real-world data set consisting of stocks that were in the S&P 500 over

[1] According to the Global Sustainable Investment Review 2022 technical report (www.gsi-alliance.org/members-resources/gsir2022), over $30 trillion out of the $115 trillion in global assets under management are presently overseen in a sustainable fashion.

the period 2001 to 2020. Our main strategy is to generate the efficient surfaces of different portfolio selection problems for the six different linear measures of climate risk employed.

On each efficient surface, we apply the NC-efficient fronts method of Steuer and Utz (2023) to study the trade-offs among the investment objectives of risk, return, and climate risk. The main idea of the approach is to determine by how much the climate risk of a portfolio can be reduced ceteris paribus if an investor relaxes portfolio expected return by a small amount, i.e., by a few basis points. We find, for instance, that a reduction in portfolio expected return of 5 basis points/month reduces climate risk by at least 17.8%. However, the rate of substitution between climate risk and expected return is not linear but decreases as reductions in expected return become larger. Thus, the first few basis points can be swapped for a comparably large reduction of climate risk. Therefore, our analysis on the climate risk dimension of portfolios shows that untapped room for improvement regarding the level of climate risk of portfolios in final investment decisions exists. In summary, the approach of this paper enables us to substantially reduce climate risk in portfolios while insufficiencies in financial performance can generally be avoided.

Because all speciality funds, Paris-aligned benchmark portfolios[2] included, rely on screens excluding certain investments, theory tells us that such funds cannot be expected to beat the market over the long haul as they are less diversified (Adler & Kritzman, 2008). Therefore, it is assumed that the exposure to risk for sustainable investment is higher than for unsustainable or traditional investment (Carswell, 2002). But is this true? While numerous studies have been conducted to test this supposition, the literature has been unable to reach a conclusion about the issue because of the diversity of results obtained from different studies. Several empirical studies (Humphrey & Tan, 2014; Kempf & Osthoff, 2007; Pizzutilo, 2017) document that additional sustainable screenings have no general negative effect on financial performance and risk. These studies consider investments in firms which to a certain extent are sustainable (e.g., they have high ESG ratings). ESG (environmental, social, and governance) investments only exclude the sustainably worst firms from a portfolio, and thus reduce the likelihood of being involved in an ethical or ecological scandal, and as shown in our tests, do not necessarily lead to lower financial returns.

2 Methodology and data

The empirical study of this paper shows the trade-offs that exist between the financial and non-financial sides of a tri-criterion portfolio selection problem when using six different climate risk measures from three data providers. In the following, we provide details about the methodological framework, the sample, and the dataset.

2.1 NC-efficient fronts approach

We study trade-offs in sustainable portfolio selection based upon the NC-efficient fronts approach of Steuer and Utz (2023). The idea of the approach is to make it easier to deal with the complexity of the three-dimensional trade-offs present in a climate risk problem by focusing on, after screening, risk vs. return first so as to obtain an M–V efficient frontier, and then after extracting the risk from the most preferred point on the M–V efficient frontier, looks at trade-offs between return and climate risk holding risk constant at the extracted level

[2] As described in “Understanding Paris-aligned indexes” published by State Street Global Advisors (https://www.ssga.com/library-content/pdfs/understanding-paris-aligned-indexes.pdf).

of risk. At the center of the approach are NC-efficient fronts. “NC-efficient” is in the name to avoid confusion with iso-quants or contours as NC-efficient fronts are lines like none other of which we are aware of in the literature. An NC-efficient front is a line on the efficient surface on which all portfolios have an expected return reduced by some fixed amount compared to portfolios on the M–V efficient frontier with the same variance.

Consider Fig. 1. The gray area is the projection of an efficient surface onto the M–V plane. The “northwest boundary” of the projection is the M–V efficient frontier of the problem, and the dashed line is a projection of an NC-efficient front onto the M–V plane. Since we measure the reduction in expected return in basis points (bp), we refer to an NC-efficient front, whose reduction in expected return is *bp*, as a *bp* NC-efficient front. As for the *bp* NC-efficient front of the dashed line, its $bp = \mu_{P_1^{ef}} - \mu_{P_1}$ as indicated on the graph, where P_1 is a portfolio on the *bp* NC-efficient front and P_1^{ef} is its corresponding portfolio (i.e., portfolio of identical variance) on the M–V efficient frontier. To be noted is that the M–V efficient frontier constitutes the far periphery of the efficient surface from the point of minimum climate risk minCR. Following from this, the point of minCR is the closest to us of all points on the efficient surface. In this way, the *bp* NC-efficient front of the dashed line bulges toward us and is closer to the viewer than the M–V efficient frontier.

In connection with notation to be used shortly, for P^i a criterion vector on the efficient surface that is on the *bp* NC-efficient front, let its inverse image in $\mathbf{x}$-space be designated $\mathbf{x}^{bp,i}$, and for the corresponding point of P^i on the M–V efficient frontier with the same variance P_i^{ef}, let its inverse image be designated $\mathbf{x}^i$.

For portfolios whose criterion vectors are on an NC-efficient front, let us now calculate a sensitivity measure $\Delta\nu \in [0, 100]$. $\Delta\nu$ is designed to capture the percentage by which the *climate risk gap*, that is, the difference between the climate risk of the minimum climate risk portfolio and the climate risk of portfolios on the M–V efficient frontier, can be closed by a reduction of *bp* in expected return as a function of variance. For giving up *bp* of expected return of a portfolio $\mathbf{x}^i$ on the M–V efficient frontier, the sensitivity measure is defined as

$$\Delta\nu = \frac{\boldsymbol{\nu}^T\mathbf{x}^i - \boldsymbol{\nu}^T\mathbf{x}^{bp,i}}{\boldsymbol{\nu}^T\mathbf{x}^i - \nu_{\min}}(100) \tag{1}$$

where $\boldsymbol{\nu}$ is the objective function coefficient vector of the climate risk measure in question and $\nu_{\min}$ is the climate risk measure’s minimum value.

In (1), the numerator is the difference in climate risk between portfolios P_1^{ef} and P_1 both of which have the same variance in Fig. 1, and the denominator is the difference of the climate risks between the portfolios of minimum climate risk minCR and P_1^{ef}. The difference between the climate risks of the portfolios P_1^{ef} and P_1 represents the improvement in the climate risk objective if the investor is willing to give up $\mu_{P_1^{ef}} - \mu_{P_1}$ in expected return keeping variance fixed. Since this absolute number might be influenced by the units of different climate risk measures, we relate this difference to the maximal possible improvement in climate risk starting from portfolio P_1^{ef}, i.e., to the difference in the climate risks between the portfolios P_1^{ef} and minCR. The quotient $\Delta\nu$ tells the investor the proportion of the climate risk gap between the M–V efficient portfolio P_1^{ef} and the minimum climate risk portfolio minCR that can be closed by a choosing portfolio P_1.

Additionally, one can combine several lines of the measure $\Delta\nu$ in one figure to show the relative improvements in climate risk possible by varying variance for different levels of *bp* reductions in expected return (see Figure 3 in Steuer and Utz (2023)). Thus, such a graph illustrates the three-dimensional trade-offs in two dimensions with the horizontal axis

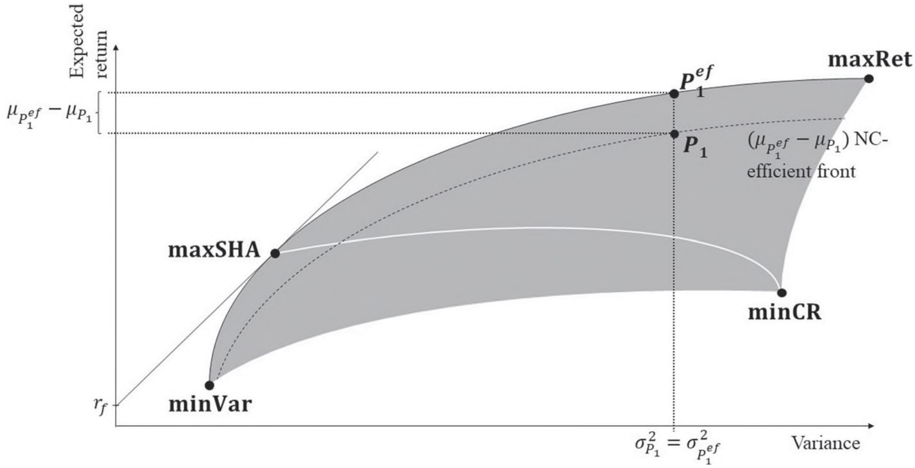

Fig. 1 In gray is the projection of an efficient surface onto the M–V plane. P_1^{ef} and P_1 are two portfolios on the efficient surface that have the same variance but differ in expected return by $\mu_{P_1^{ef}} - \mu_{P_1}$ as indicated. With $bp = \mu_{P_1^{ef}} - \mu_{P_1}$, the dashed line is the bp NC-efficient front. At the vertices of the projection are the minimum variance, maximum return, and minimum climate risk points of the efficient surface. Also shown is the maximum Sharpe ratio point with the white line being the line of maximum Sharpe ratio points over the efficient surface

varying variance, the vertical axis showing Δv, and the different lines representing different levels of bp reductions in expected return.

The empirical analysis in this paper is based on the calculation of a set of NC-efficient fronts. On each NC-efficient front, we consider different specific places called *portfolio points* at which the portfolios generated in this study are compared, where a *portfolio point* is a particular location on the efficient surface, or a portfolio possessing a particular composition.[3]

For the maxSHA and maxRet points on the M–V efficient frontier, they come from the following formulations

$$\max \left\{ \frac{\boldsymbol{\mu}^T \mathbf{x}}{\sqrt{\mathbf{x}^T \boldsymbol{\Sigma}\, \mathbf{x}}} \right\} \qquad\qquad \max \{\boldsymbol{\mu}^T \mathbf{x}\}$$
$$\text{s.t.} \quad \mathbf{x} \in S \qquad\qquad\qquad\qquad \text{s.t.} \quad \mathbf{x} \in S$$

where $\mathbf{x} \in \mathbb{R}^n$ is the vector of portfolio weights on the securities, $\boldsymbol{\mu} \in \mathbb{R}^n$ is the vector of expected returns, and $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$ is the covariance matrix. The portfolios maxSHA0 and maxRet0 as such lie on the portion of the periphery of the efficient surface furthest away from the minCR point. Concentrating on the maxSHA*bp* and maxRet*bp* points on the $bp = 5, 10, 20$ NC-efficient fronts, they come from the following formulations

[3] Other than for the equally-weighted and value-weighted portfolios, *portfolio points* are not portfolios. Portfolio points are just points on the efficient surface as it is their inverse images that are portfolios.

Table 1 This table provides a guide to the 19 portfolio points at which the problems of this paper are compared along with their abbreviations

Item	q-number	Portfolio point description	Abbreviation
1	1	Equally-weighted portfolio	1/n
2	2	Value-weighted portfolio	vw
3	3	minCR point	minCR
4	4, 8, 12, 16	minVar points on the four bp NC-efficient fronts	minVarbp
5	5, 9, 13, 17	maxSHA points on the four bp NC-efficient fronts	maxSHAbp
6	6, 10, 14, 18	minCR points on the four bp NC-efficient fronts	minCRbp
7	7, 11, 15, 19	maxRet points on the four bp NC-efficient fronts	maxRetbp

Of the 19, 17 are on the efficient surface. On each of the $bp = 0, 5, 10, 20$ NC-efficient fronts (where the 0 NC-efficient front is the M–V efficient frontier), there are 4 (the minVarbp, maxSHAbp, minCRbp, and maxRetbp) points. Of the 19 portfolio points, five are familiar with the other 14 new for addressing the tri-criterion concerns of this paper. The entries in the q-number column are designations and come into play starting at the end of this section

$$\max\left\{\frac{\boldsymbol{\mu}^T\mathbf{x}}{\sqrt{\mathbf{x}^T\boldsymbol{\Sigma}\,\mathbf{x}}}\right\} \quad \text{s.t.} \quad \mathbf{x}\in\mathcal{E}(bp)$$

$$\max\,\{\boldsymbol{\mu}^T\mathbf{x}\} \quad \text{s.t.} \quad \mathbf{x}\in\mathcal{E}(bp)$$

In these formulations, $\mathcal{E}(bp)$, not S, is the feasible region. Here $\mathcal{E}(bp)$ consists of only those $\mathbf{x}\in S$ that define the NC-efficient front for the bp-value in question, $bp = 5, 10, 20$. For example, solving $\max\{\boldsymbol{\mu}^T\mathbf{x}/\sqrt{\mathbf{x}^T\boldsymbol{\Sigma}\mathbf{x}}\}$ with $\mathcal{E}(bp = 10)$ generates the portfolio point whose abbreviation is maxSHA10. As another example, solving $\max\{\boldsymbol{\mu}^T\mathbf{x}\}$ with $\mathcal{E}(bp = 20)$ generates the portfolio point whose abbreviation is maxRet20.

The NC-efficient fronts of a problem are roughly parallel to the M–V efficient frontier in a curved sense on the efficient surface. But of course they are not contours because the third objective varies over them. The greater the bp-value, the further an NC-efficient front is from the M–V efficient frontier and the closer it is to the minCR point. Now, for the minCRbp points on the M-V efficient frontier and other NC-efficient fronts, the formulation is

$$\min\,\{\boldsymbol{\nu}^T\mathbf{x}\} \quad \text{s.t.} \quad \mathbf{x}\in\mathcal{E}(bp) \tag{2}$$

where $\boldsymbol{\nu}$ is the coefficient vector of the climate risk measure. This is not only for $bp = 5, 10, 20$, but for $bp = 0$ as well, because S does not work in this case. This is because if S were used in place of $\mathcal{E}(bp)$, (2) would generate the minCR point.

For items 1 and 2 in the Table 1 we have the equally-weighted (1/n) and value-weighted (vw) portfolios as portfolio points. The 1/n portfolio is included because of the strong showing of that portfolio point in the study by DeMiguel et al. (2009b), as none of the other portfolios in that paper were able to exhibit better out-of-sample risk-adjusted returns than the 1/n portfolio. However, with both the 1/n and vw portfolios being inefficient, there is no meaningful way to represent these portfolios on the display of Fig. 1 and hence dots for them are not shown.

Whereas the path of maxRetbp points as we cross the NC-efficient fronts is the portion of the periphery of the efficient surface directly connecting the maxRet0 and minCR points, a typical path of maxSHAbp portfolio points as we cross the NC-efficient fronts is the white line in Fig. 1. It is not possible to indicate a typical path of minCR*bp* portfolio points as one crosses the NC-efficient fronts even though the paths of maxSHA*bp* and minCR*bp* points wind up at the same minCR point. The reason is this: While one knows about where the maxSHA0 point is on an M–V efficient frontier by just looking at it (tangency point on the M–V efficient frontier), the location of the minCR0 point on an M–V efficient frontier could be anywhere. Thus, one would never know where the path of minCR*bp* points starts on the M–V efficient frontier to be able to draw it to the minCR point. In each problem the path could have a completely different originating point. This is why no attempt is made to indicate a path of minCR*bp* portfolio points in Fig. 1.

2.2 Sample and dataset

We now discuss the sample and dataset used in our study. The main sample of our study consists of all firms that have been included in the S&P500 for at least one month during the period from 2001 and 2020. In total, these are 1,049 firms. We optimize portfolios on a monthly basis using only firms that were S&P500 constituents in the respective months. For each month and all sample firms, we estimate the expected returns and the covariance matrices based on the previous 120 monthly returns which we derived from price data from Bloomberg. Moreover, we also downloaded market capitalizations from Bloomberg. In addition to financial data, we use different measures for climate risk following the climate finance literature (Bolton & Kacperczyk, 2021; Sautner et al., 2023; Görgen et al., 2020).

In detail, we use the following six measures: From Trucost, we obtained (1) the total environmental impact ratio (Total Impact Ratio) measured as the total direct and indirect external environmental cost divided by revenue[4] and (2) total emission intensity (Total Intensity) measured as the greenhouse gases emitted by the direct operations of and suppliers to a firm divided by revenue. Both measures are to be minimized to reduce climate risk. From our second data source, i.e., the measures described in Sautner et al. (2023), we use (3) the Climate Change Exposure (CC Expo) measure, (4) the Climate Change Risk (CC Risk) measure, and (5) the Climate Change Sentiment (CC Senti) measure. The first two of these measures are to be minimized to reduce climate risk, the latter is to be maximized. Finally, we use (6) Carbon β from Görgen et al. (2020) that is to be minimized to reduce climate risk. The measures of climate risk from Trucost are available from 2005 to the end of 2019, the climate change measures of Sautner et al. (2023) are available from 2001 to the end of 2020, and Carbon β of Görgen et al. (2020) is available from the end of 2012 to the end of 2019.

Table 2 shows the distribution of the climate risk measures of the firm-level time-series average quantities. The climate change measures of Sautner et al. (2023) show concentration of the numbers close to zero. Moreover, the distributions of CC Expo and CC Risk are skewed to the right which indicates a relatively high number of low values, and only a few very high values. This observation also holds for the total impact ratio and the total intensity measure. Carbon β ranges from -3.035 to 4.964 with a mean of -0.168 and is less skewed and less leptokurtic than the other measures. We also provide the correlation matrix of how the different climate risk measures relate to one another in Table 3. The correlation of

[4] The Trucost definition of this variable reads as follows: “Direct external environmental impacts are those impacts that a company has on the environment through its own activities; indirect environmental impacts of a company result from the goods and services that they purchase. Trucost applies a monetary value to environmental impact quantities, which represents the global average damage of each environmental impact. All values employed are secondary - the synthesis of existing published and unpublished literature. The sum of all the direct and indirect external environmental costs of the company is expressed as a percentage of revenue."

Table 2 This table shows the mean, standard deviation (sd), minimum (min), 5%-quantile (q0.05), median, 95%-quantile, maximum (max), skewness (skew), and kurtosis (kurt) of the firm-level time-series average values of the six considered climate risk measures where CC stands for climate change

	mean	sd	min	q0.05	median	q0.95	max	skew	kurt
Tot Imp Rat	4.027	8.599	0.237	0.288	1.424	17.964	75.724	4.734	26.797
Tot Int	37.912	112.934	0.372	0.853	7.457	153.288	1235.398	5.859	40.294
CC Expo	0.091	0.203	0.000	0.011	0.029	0.462	1.659	4.707	25.055
CC Risk	0.003	0.007	0.000	0.000	0.001	0.016	0.075	4.503	24.311
CC Senti	−0.000	0.000	−0.005	−0.000	0.000	0.000	0.003	−3.252	60.445
Carbon β	−0.168	0.779	−3.035	−1.247	−0.237	1.002	4.964	1.108	5.294

Total Impact Ratio (Tot Imp Rat) and Total Intensity (Tot Int, in 1000) from Trucost, Climate Change Exposure (CC Expo), Risk (CC Risk), and Sentiment (CC Senti) (see Sautner et al., 2023, all values are multiplied by 100), and Carbon β (see Görgen et al., 2020)

Table 3 This table shows the correlation matrix of the firm-level time-series average values of the six considered climate risk measures

	Tot Imp Rat	Tot Int	CC Expo	CC Risk	CC Senti	Carbon β
Tot Imp Rat	1.000	0.859	0.092	0.041	−0.112	−0.025
Tot Int	0.859	1.000	0.060	0.030	−0.113	−0.032
CC Expo	0.092	0.060	1.000	0.842	−0.235	0.034
CC Risk	0.041	0.030	0.842	1.000	−0.236	0.043
CC Senti	−0.112	−0.113	−0.235	−0.236	1.000	0.044
Carbon β	−0.025	−0.032	0.034	0.043	0.044	1.000

measures within provider are seen to deviate substantially from zero whereas the correlation of measures between providers is seen to be close to zero. The Climate Change Sentiment measure is negatively correlated with all measures except Carbon β. In general, the negative correlation is reasonable since this number has to be maximized to reduce climate risk. Since the correlations between the measures of different providers are rather low, it makes sense to consider all of them in our test since they might consider different aspects of climate risk. In using all of them, we are able to study the trade-offs of financial performance and climate risk for different aspects of climate risk.

2.3 Empirical strategy

In the following, we describe the experiments that we have conducted to demonstrate the thesis of this paper. In using a rolling-sample approach to carry out our experiments, we consider subsamples using 121 consecutive monthly return observations. This enables $L = 178$ subsamples for the Total Intensity and Total Impact Ratio measures, $L = 215$ subsamples for the CC Expo, CC Risk, and CC Senti measures, and $L = 84$ subsamples for Carbon β.

Since the subsamples are repetitive, we just focus on how one works on one of the datasets to illustrate. Each subsample uses 121 consecutive monthly return observations with the last at time t. From the first 120 months we form expected return vector $\boldsymbol{\mu}_*$ and covariance matrix $\boldsymbol{\Sigma}_*$ from which we compute the M–V efficient frontier of the dataset at time $(t-1)$. After this is done, next is needed $\boldsymbol{\nu}_*$, the climate risk vector at time $(t-1)$ for the stocks in the dataset. The vector does not come from historical data. It comes from the latest available numbers of the respective variable of climate risk.

Now, with $\boldsymbol{\mu}_*$, $\boldsymbol{\Sigma}_*$, $\boldsymbol{\nu}_*$ and knowledge of all (σ_i^2, μ_i) combinations along the M–V efficient frontier, we solve the following optimization model

$$\begin{aligned} &\min\{\boldsymbol{\nu}_*^T \mathbf{x}\} \\ \text{s.t. } &\mathbf{x}^T \boldsymbol{\Sigma}_* \mathbf{x} = \sigma_i^2 \\ &\boldsymbol{\mu}_*^T \mathbf{x} \geq \mu_i - bp \\ &\mathbf{x} \in S \end{aligned}$$

where "min" is replaced by "max" in the case of the CC Senti climate risk measure, and all securities have an upper bound of 5% in S, for enough (σ_i^2, μ_i) combinations to construct the $bp = 5, 10, 20$ NC-efficient fronts, as the M–V efficient frontier (that is, the 0 NC-efficient front) has already been computed. Then from the four frontiers, we obtain the portfolios that are the inverse images of the 16 portfolio points on them. Combining them with the inverse

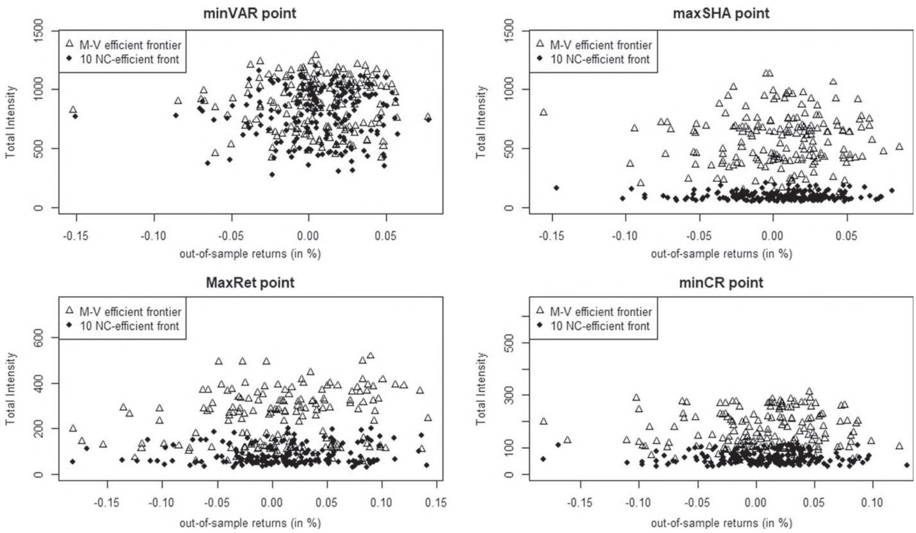

Fig. 2 For the $L = 178$ optimizations of Total Intensity, we have the above four plots for the four portfolio points indicated. Using triangles, we have the Total Intensity, out-of-sample return combinations at the four portfolio points along the M–V efficient frontier (i.e., along the 0 NC-efficient front). Using dots, we have the Total Intensity, out-of-sample return combinations at the four portfolio points along the 10 NC-efficient front. Note the amount by which the cloud of dots is below the cloud of triangles as a result of the 10bp/mth reduction in expected return

image of the minCR point, the 1/n equally-weighted portfolio, and the vw value-weighted portfolio, we have the $\mathbf{x}_{t-1,q}$ portfolios of the $q = 1, \ldots, 19$ portfolio points of the dataset. With the subsample's 121-st monthly return observation giving us the subsample's out-of-sample return vector R_t, we are able to compute the subsample's 19 out-of-sample realized portfolio returns $R_t^T \mathbf{x}_{t-1,q}$ for the t-th time period. Furthermore, for $q = 1, \ldots, 19$, the $\boldsymbol{\nu}_*^T \mathbf{x}_{t-1,q}$ give us the climate risks of these portfolios. Doing this for all subsamples and all six datasets provides us with the inputs needed for the statistical tests that follow.

3 Results

In the following section, we show that the amounts of climate risk that M–V portfolio selection is unable to bring to the attention of an investor, but can be identified by application of the Hirschberger et al. (2013) and Steuer and Utz (2023) approaches. In particular, the results show that the reduction of climate risk does not come with a substantial reduction in financial out-of-sample performance. Figure 2 illustrates a summary of the empirical results for four different portfolio points, i.e., the minimum variance portfolio, the maximum Sharpe ratio portfolio, the maximum return portfolio, and the minimum climate risk portfolios on the M–V efficient frontier and on the 10 NC-efficient front.

Figure 2 displays climate risk on the vertical axis (measured by Total Intensity from Trucost) and the out-of-sample returns, i.e., the realized portfolio return in the month after the portfolio construction, on the horizontal axis. The "triangles" represent the climate risk-return combinations of the M–V efficient frontier portfolio points, while the "dots" represent the ones of the 10 NC-efficient front portfolio points. It is easy to observe that the 10 NC-efficient front portfolio points have substantially lower climate risk than the M–V efficient frontier portfolios, since the cloud of dots is located lower than the one of the triangles. If the financial performance of the 10 NC-efficient front portfolios would be substantially lower than that of the M–V efficient frontier portfolios, we would expect to see a shift of the cloud of the 10 NC-efficient front portfolios to the left. However, we do not observe such a shift. Therefore, the figure provides anecdotal evidence that the 10 NC-efficient front portfolios perform apparently better in terms of climate risk and not worse in the financial dimension. In the following, we show further statistics that support the conclusion from the illustrating example.

3.1 Climate risk performance and portfolio characteristics at the portfolio points

In this investigation, we provide a summary of portfolio characteristics {i.e., volatility, return, climate risk, and the sensitivity measure $\Delta\nu$ defined in Eq. (1)} of the M–V efficient frontier and the 5, 10, 20 NC-efficient front portfolios. Each front is characterized by up to 100 portfolios, and the basis of the reported statistics are all portfolios on all fronts. Table 4 shows the mean, minimum, and maximum values of these portfolios. Each panel in the table displays the results for another measure of climate risk. The volatility of the portfolio return (displayed in the first row of each panel) appears to increase on average from the M–V efficient frontier to the NC-efficient fronts for all climate risk measures except of Panel 6 (Carbon β). The average expected returns of the portfolios appear rather constant across the different fronts. However, the climate risk objective reduces clearly on the NC-efficient fronts compared to the M–V efficient frontier. The reduction in climate risk is also documented by increasing $\Delta\nu$ measures. Since $\Delta\nu$ specifies how much less climate risk can be attained by heading into the efficient surface from a given point on the M–V efficient frontier for a given relaxation in expected return, variance held constant, we observe that in Panel 1, the reduction of the expected return by 10bp generates portfolios that on average have a $\Delta\nu = 48.85\%$. This means, that the reduction of the expected return by 10 bp helps to close the climate risk gap from M–V efficient frontier portfolios to the climate risk of the minCR portfolio point by almost 50%. While a 20bp reduction in expected returns yields a $\Delta\nu$ measure of about 60–68% for the Trucost climate risk measures and about 77–92% for the Sautner et al. (2023) climate risk measures, the $\Delta\nu$ for the Carbon β in Panel 6 is on average 32%.

In Table 4 the second, third, and fourth sets of columns show the reduction of climate risk that the procedure by Hirschberger et al. (2013) is able to identify as a function of different levels of expected return relaxation. Going into the table in more detail, the 5 NC-efficient fronts experience an average reduction in climate risk of about 37.8%, 47.3%, 53.5%, 89.7%, 53.5%, and 17.8%. In other words, the 5bp/mth relaxation enables us to eliminate on average 17.8–89.7% of the climate risk that would otherwise be included in the portfolios by not using the tri-objective model to generate portfolios.

Continuing with the 10bp reduction, we observe the $\Delta\nu$ values to increase. However, although these portfolios are twice the distance away from the M–V efficient frontier in terms of expected return as the 5 NC-efficient front portfolios, the $\Delta\nu$ do not double. For instance, an expected-return relaxation of 10bp/mth, or 1.2%/year, leads to a 1.624 reduction in total impact ratio on average $(4.952 - 3.328)$ over no relaxation and a $\Delta\nu$ of 48.85%. While relaxations up to 1.2%/yr would probably be within the leeway given to fund managers by investors caring for climate risk, relaxations beyond this might soon become too much. However, for sake of discussion, should we double the relaxation to 20bp/mth, we would only experience a further reduction of .66 in the Total Impact Ratio measure. This is less

Table 4 Summary statistics for volatility (Vola), return (Ret), level of climate risk (CR), and the sensitivity measure $\Delta\nu$ on the M–V efficient frontier and 5, 10, 20 NC-efficient fronts for the six different measures of climate risk of this paper

	Efficient frontier			5 NC-efficient front			10 NC-efficient front			20 NC-efficient front		
	mean	min	max	mean	min	max	mean	min	max	mean	min	max
Panel 1: Trucost Total Impact Ratio												
Vola	0.033	0.018	0.080	0.042	0.021	0.113	0.042	0.021	0.112	0.042	0.021	0.103
Ret	0.025	0.011	0.039	0.030	0.014	0.045	0.029	0.014	0.045	0.028	0.014	0.044
CR	4.952	2.109	8.191	4.032	1.281	11.221	3.328	1.154	10.926	2.668	0.953	11.087
$\Delta\nu$				37.798	2.066	69.157	48.850	4.190	76.330	60.821	2.061	83.969
Panel 2: Trucost Total Intensity												
Vola	0.033	0.018	0.080	0.042	0.021	0.113	0.042	0.021	0.113	0.042	0.021	0.107
Ret	0.025	0.011	0.039	0.030	0.014	0.045	0.029	0.014	0.045	0.028	0.014	0.044
CR	470.549	161.802	892.772	347.844	74.228	1246.501	249.535	65.762	1195.404	189.55	51.271	1224.518
$\Delta\nu$				47.275	3.318	88.610	59.357	7.229	90.391	68.398	4.671	92.449
Panel 3: Sautner et al. (2023) *Climate Change Exposure*												
Vola	0.035	0.019	0.089	0.050	0.021	0.148	0.050	0.021	0.143	0.050	0.021	0.142
Ret	0.025	0.011	0.040	0.028	0.010	0.047	0.028	0.009	0.047	0.027	0.009	0.046
CR	0.001	0.000	0.002	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.001
$\Delta\nu$				53.506	8.528	77.904	65.895	18.148	84.719	77.763	5.255	92.927
Panel 4: Sautner et al. (2023) *Climate Change Risk*												
Vola	0.035	0.019	0.089	0.051	0.021	0.143	0.051	0.021	0.142	0.050	0.021	0.137
Ret	0.025	0.011	0.040	0.028	0.010	0.047	0.027	0.009	0.047	0.026	0.009	0.046
CR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$\Delta\nu$				89.677	7.992	100	90.779	15.08	99.999	92.112	9.572	100

Table 4 continued

	Efficient frontier			5 NC-efficient front			10 NC-efficient front			20 NC-efficient front		
	mean	min	max	mean	min	max	mean	min	max	mean	min	max
Panel 5: Sautner et al. (2023) *Climate Change Sentiment*												
Vola	0.035	0.019	0.089	0.050	0.021	0.148	0.050	0.021	0.143	0.050	0.021	0.142
Ret	0.025	0.011	0.040	0.028	0.010	0.047	0.028	0.009	0.047	0.027	0.009	0.046
CR	0.001	0.000	0.002	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.000	0.001
$\Delta\nu$				53.506	8.528	77.904	65.895	18.148	84.719	77.763	5.255	92.927
Panel 6: Görgen et al. (2020) *Carbon* β												
Vola	0.033	0.018	0.085	0.032	0.018	0.074	0.032	0.018	0.074	0.032	0.018	0.071
Ret	0.023	0.009	0.037	0.023	0.010	0.037	0.023	0.009	0.036	0.022	0.009	0.035
CR	−0.130	−0.438	0.242	−0.491	−0.678	0.099	−0.588	−0.795	−0.135	−0.705	−0.934	−0.172
$\Delta\nu$				17.805	0.393	25.886	24.350	0.738	35.178	32.222	1.212	46.883

The values reported in this table are based on 178 (215 and 84) monthly generated frontiers with 100 portfolios on each front for the six different climate risk measures. The respective statistics (mean, min, and max) are based on all the generated numbers on the particular frontiers than that gained in the first basis point of relaxation. What we are witnessing is the trade-off situation changing from a large amount of climate risk for a small amount of relaxation, to a small amount of climate risk for a large amount of relaxation. The explanation is curvature of the efficient surface. This suggests that a sweet spot might exist between a 5 and 10bp relaxation.

3.2 Financial performance at the portfolio points

With the empirical force of the gains in climate risk of the previous subsection clear, investors might ask how financial portfolio characteristics are affected. Therefore, we consider diversification and risk-adjusted returns of NC-efficient front portfolios in the following.

First, our measure for diversification is the industry weight, i.e., the sum of all weights of the stocks in the industry, in a portfolio on a particular NC-efficient front. For this exercise, we follow the ICB industry classification with the 11 industries Basic Materials, Consumer Discretion, Consumer Staples, Energy, Financials, Health Care, Industrials, Real Estate, Technology, Telecommunications, and Utilities. Table 5 shows the average industry weights for the M–V efficient frontier and the 5, 10, 20 NC-efficient fronts. The six panels show the distribution of the average weights for the six measures of climate risk.

The table shows that the weights of the industries Energy and Utilities decreased with a higher reduction of the expected return (and more focus on climate risk). However, with the exception of the Energy industry in Panels 3, 4, and 5, even the portfolios that focus more on a reduction of climate risk (i.e., the portfolios on the NC-efficient fronts) invest a reasonable proportion of their wealth in each industry. The Health Care industry is the only industry with increasing weights for all measures of climate risk. Besides, the weights of the Financials industry also increase for the Trucost climate risk measures and Carbon β, and remain almost unchanged for the experiments using the Sautner et al. (2023) measures. Overall, the diversification of portfolios with respect to the industry distribution appears to be not substantially affected by the application of the tri-criterion portfolio model introduced by Hirschberger et al. (2013).

The next quantity we analyze is risk-adjusted portfolio return. We do this by employing the five-factor model of Fama and French (2015) and noting the α's generated and their significance levels. The news from our tests is that the α's generated show (in the majority of the cases) no evidence of risk-adjusted returns of any materiality having to be given up. That is, what we show is that the significant gains in climate risk of the previous subsection can be achieved without diminutions in monthly risk-adjusted returns of any significance.

In carrying out our Fama and French tests, we compare the out-of-sample returns at each of the 19 portfolio points of each dataset with the returns of well-documented risk factors on financial markets. In doing so, we can determine whether the portfolios that result from our optimizations generate market-rate compensations for the risks taken. The well-documented risk factors taken into account are those of the five-factor Fama-French model, i.e., excess returns on the market portfolio $\mathbf{R}_{Mkt} - \mathbf{R}_f$ (with $\mathbf{R}_f$ the vector of risk-free rates), the return on the small minus big portfolio $\mathbf{R}_{SMB}$, the return on the high minus low portfolio $\mathbf{R}_{HML}$, the return on the robust minus weak portfolio $\mathbf{R}_{RMW}$, and the return on the conservative minus aggressive portfolio $\mathbf{R}_{CMA}$. In this way, we estimate the average abnormal return of portfolio point q over the period with out-of-sample returns by α_q using the following model

$$\begin{aligned}\mathbf{R}_q = \alpha_q \mathbf{1} + \beta_{Mkt}(\mathbf{R}_{Mkt} - \mathbf{R}_f) + \beta_{SMB}\mathbf{R}_{SMB} + \\ + \beta_{HML}\mathbf{R}_{HML} + \beta_{RMW}\mathbf{R}_{RMW} + \beta_{CMA}\mathbf{R}_{CMA} + \boldsymbol{\epsilon}_q.\end{aligned} \quad (3)$$

Table 5 This table reports average sector portfolio weights on the M–V efficient frontier and on the 5, 10, 20 NC-efficient fronts

Mean weight	Eff. frontier	5 NC-efficient front	10 NC-efficient front	20 NC-efficient front
Panel 1: Trucost Total Impact Ratio				
Basic Materials	0.029	0.021	0.026	0.026
Consumer Discretion	0.187	0.140	0.136	0.121
Consumer Staples	0.118	0.109	0.102	0.087
Energy	0.081	0.066	0.061	0.047
Financials	0.057	0.105	0.112	0.139
Health Care	0.204	0.313	0.314	0.314
Industrials	0.074	0.051	0.051	0.052
Real Estate	0.019	0.029	0.035	0.044
Technology	0.111	0.109	0.112	0.12
Telecommunications	0.028	0.019	0.019	0.019
Utilities	0.091	0.039	0.034	0.032
Panel 2: Trucost Total Intensity				
Basic Materials	0.029	0.022	0.025	0.024
Consumer Discretion	0.187	0.154	0.137	0.120
Consumer Staples	0.118	0.126	0.113	0.092
Energy	0.081	0.047	0.044	0.034
Financials	0.057	0.088	0.115	0.158
Health Care	0.204	0.329	0.335	0.339
Industrials	0.074	0.044	0.047	0.044
Real Estate	0.019	0.026	0.032	0.034
Technology	0.111	0.109	0.110	0.12
Telecommunications	0.028	0.018	0.015	0.012
Utilities	0.091	0.036	0.027	0.022
Panel 3: Sautner et al. (2023) *Climate Change Exposure*				
Basic Materials	0.026	0.040	0.046	0.050
Consumer Discretion	0.173	0.124	0.125	0.131
Consumer Staples	0.122	0.120	0.108	0.102
Energy	0.068	0.001	0.001	0.001
Financials	0.043	0.028	0.030	0.041
Health Care	0.225	0.350	0.359	0.354
Industrials	0.080	0.084	0.086	0.089
Real Estate	0.022	0.038	0.042	0.047
Technology	0.121	0.167	0.160	0.142
Telecommunications	0.029	0.026	0.026	0.029
Utilities	0.090	0.022	0.017	0.015

Table 5 continued

Mean weight	Eff. frontier	5 NC-efficient front	10 NC-efficient front	20 NC-efficient front
Panel 4: Sautner et al. (2023) *Climate Change Risk*				
Basic Materials	0.026	0.040	0.040	0.041
Consumer Discretion	0.173	0.102	0.102	0.104
Consumer Staples	0.122	0.119	0.117	0.114
Energy	0.068	0.002	0.004	0.007
Financials	0.043	0.025	0.03	0.037
Health Care	0.225	0.325	0.317	0.301
Industrials	0.080	0.088	0.089	0.092
Real Estate	0.022	0.034	0.034	0.035
Technology	0.121	0.174	0.174	0.171
Telecommunications	0.029	0.027	0.028	0.031
Utilities	0.090	0.062	0.065	0.067
Panel 5: Sautner et al. (2023) *Climate Change Sentiment*				
Basic Materials	0.026	0.040	0.046	0.050
Consumer Discretion	0.173	0.124	0.125	0.131
Consumer Staples	0.122	0.120	0.108	0.102
Energy	0.068	0.001	0.001	0.001
Financials	0.043	0.028	0.030	0.041
Health Care	0.225	0.350	0.359	0.354
Industrials	0.080	0.084	0.086	0.089
Real Estate	0.022	0.038	0.042	0.047
Technology	0.121	0.167	0.160	0.142
Telecommunications	0.029	0.026	0.026	0.029
Utilities	0.090	0.022	0.017	0.015
Panel 6: Görgen et al. (2020) *Carbonβ*				
Basic Materials	0.024	0.010	0.006	0.006
Consumer Discretion	0.250	0.230	0.232	0.222
Consumer Staples	0.110	0.156	0.164	0.176
Energy	0.056	0.078	0.070	0.056
Financials	0.046	0.058	0.069	0.078
Health Care	0.161	0.224	0.227	0.240
Industrials	0.058	0.030	0.028	0.025
Real Estate	0.017	0.001	0.000	0.000
Technology	0.126	0.138	0.139	0.138
Telecommunications	0.034	0.001	0.001	0.001
Utilities	0.116	0.073	0.063	0.056

For the model, we download from Kenneth French's website[5] the monthly returns of the risk factors from the North-American category for the S&P sample. We ran 114 regression models (19 portfolio points times 6 measures for climate risk) to obtain from Equation (3) the risk-adjusted abnormal returns α_q and coefficients of the risk factors (β_{Mkt}, β_{SMB}, β_{HML}, β_{RMW}, and β_{CMA}) for the portfolios of all portfolio points.

In Table 6 we show the 114 monthly α-results (i.e., abnormal returns) of the regressions. The portfolio points 1/n, vw, minVar0, maxSHA0, and maxRet0 are available for investors who ignore climate risk as an objective, i.e., who only base their decisions on the financial characteristics. That is, they are only for the portfolio points on the M–V efficient frontier. While slightly negative, they are not statistically different from zero (except the maxSHA0 portfolio point).

The minCR portfolio is essentially never located on the M–V efficient frontier as it is the portfolio that ignores expected return and variance and only concentrates on minimizing climate risk. For this portfolio, the results of the experiments show that this portfolio generates risk-adjusted abnormal returns close to zero. Another portfolio that is new to the decision-making process in the M–V framework is minCR0 portfolio. It is the portfolio of minimum climate risk of all portfolios on the M–V efficient frontier. While the α's of this portfolios are slightly negative, they are again not significantly less than zero.

The other rows in the table are for the portfolio points on the $bp = 5$, 10, 20 NC-efficient fronts and give us a sense of what else is on the efficient surface as we shift away from the M–V efficient frontier to regions of lower climate risk. The 14 other portfolio points could be generated when applying Hirschberger et al. (2013) to obtain the efficient surface. Related to these 14 portfolio points are 84 abnormal return estimates of which only three (two, five) portfolio points generate an out-of-sample risk-adjusted portfolio return that is significantly smaller than zero at the 1%-level (5%-, 10%-level). These portfolios have a significantly lower return than an investment alternative generating market return at the same profile. These results reveal that the majority of the NC-efficient front portfolio points (about 88%) generate a return that is not significantly different from the market-rate return. Therefore, we conclude that reducing climate risk does not to a large extent harm financial return.

Even if we were to take the most negative α for a minCR-portfolio point on an NC-efficient front portfolio, the one of -0.463 (minCR05, Carbon β), that would only amount to -5.56 bp/yr, well within any leeway of any climate risk committed investor would certainly grant to a fund manager for pursuing low climate risk. Thus, the meaning of the columns is that we find no evidence that there is any material loss in risk-adjusted returns when using tri-criterion portfolio selection, rather than M–V optimization. In summary, Table 6 shows us that the portfolios of almost all portfolio points on the efficient surface generate performances that compensate investors with market-rate returns.

4 Robustness checks

In this section, we conduct robustness checks on the portfolio results of the previous section. We do this to see if anything unusual crops up that might cast doubt on the interpretation of this paper, which is that current methods for computing optimal climate risk portfolios fail to notice considerable amounts by which climate risk can be reduced without affecting risk-adjusted returns.

[5] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Table 6 This table reports the abnormal returns (α) at the portfolio points in our experiments

Fama and French (2015) α	Tot Imp Rat	Tot Int	CC Expo	CC Risk	CC Senti	Carbon β
1/n	−0.095	−0.095	−0.072	−0.072	−0.072	−0.143**
vw	−0.171***	−0.171***	−0.171***	−0.171***	−0.171***	−0.140***
minCR	0.103	−0.100	−0.008	0.019	0.072	−0.040
minVar0	−0.184	−0.184	−0.017	−0.017	−0.017	−0.114
maxSHA0	−0.309***	−0.309***	−0.170	−0.17	−0.170	−0.131
minCR0	−0.177	−0.214	−0.075	−0.044	−0.063	−0.352
maxRet0	−0.104	−0.104	0.011	0.011	0.011	−0.568*
minVar05	−0.186	−0.184	−0.015	−0.011	−0.015	−0.106
maxSHA05	−0.322***	−0.347***	−0.133	−0.158	−0.144	−0.086
minCR05	−0.189	−0.105	0.039	0.003	0.003	−0.463*
maxRet05	0.015	−0.064	0.060	−0.042	−0.043	−0.554*
minVar10	−0.189	−0.187	−0.016	−0.005	−0.019	−0.112
maxSHA10	−0.317***	−0.307**	−0.116	−0.153	−0.121	−0.038
minCR10	−0.138	−0.128	−0.001	−0.068	−0.017	−0.453*
maxRet10	0.032	0.078	0.027	−0.044	−0.020	−0.541*
minVar20	−0.196	−0.190	−0.025	−0.008	−0.031	−0.126
maxSHA20	−0.246**	−0.247*	−0.096	−0.142	−0.090	−0.032
minCR20	−0.143	−0.011	0.034	0.046	0.018	−0.310
maxRet20	0.088	0.152	0.063	−0.043	0.017	−0.416

From the five-factor Fama-French model (Eq. (3)) with Newey and West (1987) standard errors, in this table we report the average one-month alpha (in basis points) at each portfolio point of each dataset along with accompanying t-statistics with * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ indicating significance levels

In these robustness checks, other performance measures (following DeMiguel, Garlappi, Nogales, & Uppal, 2009a; DeMiguel, Garlappi, & Uppal, 2009b) are applied to compare the portfolios generated at NC-efficient front portfolio points against the 1/n portfolio under the assumption that the 1/n portfolio can be viewed as a satisfactory benchmark for normalcy that performs as well as any other portfolio.

The performance measures used for comparing the portfolios of the experiments against the 1/n portfolio are out-of-sample variance, out-of-sample Sharpe ratio, out-of-sample certainty-equivalent (CEQ) return, and portfolio turnover. In summary, the results of all below robustness tests provide supporting evidence for the conclusions drawn from the main analyses.

(1) Out-of-sample variance. Our first additional performance measure is the out-of-sample variance of the portfolio returns. Therefore, we test the hypothesis that the out-of-sample variance of a particular portfolio point is equal to the out-of-sample variance of the 1/n portfolio point, i.e., $H_0 : \ \Delta\sigma^2 := \sigma_q^2 - \sigma_{1/n}^2 = 0,\ q \in \{2, \ldots, 19\}$. The reason for this notation is that the 1/n portfolio is portfolio $q = 1$ in Table 1. The out-of-sample variance is given by

$$\hat{\sigma}_q^2 = \frac{1}{L-1}\sum_{t=1}^{L}(R_{t,q}^T \mathbf{x}_{t-1,q} - \hat{\mu}_q)^2, \tag{4}$$

where L is the number of out-of-sample periods. Statistical inference is derived by applying the bootstrap approach suggested by Ledoit and Wolf (2011).

Results of our out-of-sample variance tests are given in the first set of columns of Table 7 for the Trucost Total Impact Ratio. Results for the other measures of climate risk are not shown as they are in effect qualitatively the same. The results of the tests show that the out-of-sample variance of the minVar*bp* and the maxSHA*bp* portfolio points for all levels of *bp* are significantly lower than the out-of-sample variance of the 1/n portfolio. Additionally, the out-of-sample variance of the minCR*bp* portfolio points is not significantly different from the out-of-sample variance of the 1/n portfolio. Moreover, the absolute levels of the out-of-sample variance appear also to be rather invariant to changes in the *bp*-relaxation.

(2) Out-of-sample Sharpe ratio. In comparing portfolios using this metric, we test the hypothesis that there is no difference between the out-of-sample Sharpe ratio of portfolio q and the out-of-sample Sharpe ratio of the 1/n portfolio, i.e., H_0: $\frac{\hat{\mu}_q}{\hat{\sigma}_q} - \frac{\hat{\mu}_{1/n}}{\hat{\sigma}_{1/n}} = 0,\ q \in \{2, \ldots, 19\}$. In H_0, $\hat{\mu}_q$ is the sample mean of out-of-sample returns given by

$$\hat{\mu}_q = \frac{1}{L}\sum_{t=1}^{L} R_{t,q}^T \mathbf{x}_{t-1,q} \tag{5}$$

where L is the number of out-of-sample periods and $\hat{\sigma}_q$ is out-of-sample standard deviation given by Equation (4).

To test the hypothesis, we compute two-sided p values using the studentized circular block bootstrap process of Ledoit and Wolf (2008) with $B = 1000$ bootstrap resamples and block size $b = 5$. The performance of bootstrap methods has been shown to be more efficient than the standard test suggested by Jobson and Korkie (1981) after correcting as in Memmel (2003). Results of our out-of-sample Sharpe ratio tests are given in the second set of columns of Table 7 for the Trucost Total Impact Ratio.

The Sharpe ratio tests reveal that no NC-efficient front portfolio underperforms the 1/n portfolio point at the 10% significance level on this measure. The column labeled $\Delta\widehat{\mathrm{SHA}}$

Table 7 This table reports on the differences in out-of-sample (ofs) variances, Sharpe ratios, certainty-equivalent returns (CEQs), and turnover rates between our portfolio points and the 1/n portfolio for the climate risk measure Total Impact Ratio

	ofs variance		ofs Sharpe		ofs CEQ		
	$\Delta\widehat{\text{var}}$	p value	$\Delta\widehat{\text{SHA}}$	p value	$\Delta\widehat{\text{CEQ}}$	p value	turnover
1/n	0.002		0.000		0.007		0.065
vw	− 0.309	(0.000)	− 0.003	(0.878)	− 0.001	(0.852)	0.016
minCR	0.756	(0.000)	− 0.065	(0.190)	− 0.002	(0.753)	0.186
minVar0	− 0.804	(0.000)	− 0.006	(0.886)	− 0.002	(0.858)	0.288
maxSHA0	− 0.467	(0.000)	− 0.040	(0.305)	− 0.003	(0.935)	0.337
minCR0	0.170	(0.261)	− 0.027	(0.540)	− 0.001	(0.660)	0.364
maxRet0	0.525	(0.001)	− 0.020	(0.663)	0.000	(0.422)	0.224
minVar05	− 0.798	(0.000)	− 0.007	(0.879)	− 0.002	(0.858)	0.288
maxSHA05	− 0.436	(0.001)	− 0.043	(0.251)	− 0.003	(0.948)	0.337
minCR05	0.090	(0.592)	− 0.031	(0.461)	− 0.001	(0.728)	0.362
maxRet05	0.468	(0.002)	− 0.005	(0.916)	0.001	(0.303)	0.248
minVar10	− 0.788	(0.000)	− 0.008	(0.849)	− 0.002	(0.863)	0.289
maxSHA10	− 0.411	(0.001)	− 0.041	(0.277)	− 0.003	(0.946)	0.331
minCR10	0.048	(0.745)	− 0.026	(0.508)	− 0.001	(0.714)	0.377
maxRet10	0.423	(0.005)	− 0.002	(0.957)	0.001	(0.301)	0.257
minVar20	− 0.771	(0.000)	− 0.011	(0.801)	− 0.002	(0.870)	0.296
maxSHA20	− 0.369	(0.002)	− 0.027	(0.442)	− 0.002	(0.888)	0.334
minCR20	0.065	(0.671)	− 0.029	(0.477)	− 0.001	(0.733)	0.371
maxRet20	0.410	(0.005)	0.005	(0.907)	0.002	(0.236)	0.259

The test of the out-of-sample variance considers the difference $\Delta\widehat{\text{var}} = \widehat{\sigma}_q^2 - \widehat{\sigma}_{1/n}^2$ and the bootstrapped p values (in parentheses) of the differences are computed in accordance with Ledoit and Wolf (2011). The test of the out-of-sample Sharpe ratio considers the difference $\Delta\widehat{\text{SHA}} = \widehat{\text{SHA}}_q - \widehat{\text{SHA}}_{1/n}$ and the bootstrapped p values (in parentheses) of the differences are also computed in accordance with Ledoit and Wolf (2008). The test of the out-of-sample CEQ considers the difference $\Delta\widehat{\text{CEQ}} = \widehat{\text{CEQ}}_q - \widehat{\text{CEQ}}_{1/n}$ and the p values (in parentheses) of the differences as computed in accordance with Greene (2002). Finally, turnover is reported in the last column of the table

reports the bootstrapped differences in Sharpe ratios between the different portfolios and the 1/n portfolio, with the numbers in parentheses being bootstrapped p values. A positive $\Delta\widehat{\text{SHA}}$ value represents a greater level of compensation (in terms of financial return) per unit of risk for the portfolio undergoing comparison than for the 1/n portfolio, and thus indicates better risk-adjusted performance. The statistical inference (reported as p values) adds to our findings by showing that the risk-adjusted performance of NC-efficient front portfolios is insignificantly different from that of the 1/n portfolio.

(3) Out-of-sample certainty-equivalent (CEQ) return. Here we compare the portfolios generated at the 18 portfolio points on the NC-efficient fronts against the 1/n portfolio using certainty-equivalent returns, where the CEQ return of an uncertain payoff is the risk-free rate at which an investor is indifferent between the two. In this way, the higher the CEQ return of a portfolio, the better its performance all other things equal. The CEQ return of a portfolio at portfolio point q is given by

$$\widehat{\mathrm{CEQ}}_q = \lambda_\mu \hat{\mu}_q - \frac{1}{2}\hat{\sigma}_q^2 \tag{6}$$

where $\hat{\mu}_q$ and $\hat{\sigma}_q^2$ are as in (5) and (4), and λ_μ is a risk tolerance parameter. Following DeMiguel et al. (2009b), we report results for the $\lambda_\mu = 1$ case. For determining whether the out-of-sample CEQ returns of two strategies q and 1/n are statistically distinguishable[6], we apply one-sided tests with null hypothesis H_0: $\widehat{\mathrm{CEQ}}_q - \widehat{\mathrm{CEQ}}_{1/n} = 0$ and alternative hypothesis $\mathrm{H}_a : \widehat{\mathrm{CEQ}}_q - \widehat{\mathrm{CEQ}}_{1/n} < 0$.

Results of our CEQ tests are given in Table 7 and they support the conclusion that NC-efficient front portfolios do not suffer financially, i.e., their risk-adjusted performances are on a par with those of the 1/n portfolio. To illustrate the table, the $\Delta\widehat{\mathrm{CEQ}} = -0.003$ entry for the maxSHA0 point asserts that the certainty equivalent of this point is only 0.3bp lower (statistically insignificant) than the riskless equivalent payoff of the 1/n portfolio. A high certainty-equivalent return of a risky asset indicates that the asset generates high returns, and thus, an investor would only substitute the gamble for a riskless alternative if the riskless alternative payoff were higher. Thus, a positive or insignificantly negative $\Delta\widehat{\mathrm{CEQ}}$ value occurs if the performance of the portfolio is not worse than that of the 1/n portfolio. We observe this result (positive or insignificantly negative $\Delta\widehat{\mathrm{CEQ}}$ values) at all portfolio points for all datasets. Tests for other values of λ_μ were conducted but are not reported because they show nothing new.

(4) Portfolio turnover. With this metric, we assess the amount of trading associated with the different portfolios. We calculate the monthly turnover of a portfolio as the percentage of the total wealth of a portfolio that is traded on average between two reporting dates. Formally, for generic portfolio q, portfolio turnover is given by

$$\mathrm{Turnover}_q = \frac{1}{L-1}\sum_{t=1}^{L-1}\sum_{i=1}^{n}\left|x^i_{q,t+1} - x^i_{q,t^+}\right| \tag{8}$$

where n is the number of securities, x^i_{q,t^+} is the weight on asset i at time $t+1$ before rebalancing, and $x^i_{q,t+1}$ is the weight on asset i at time $t+1$ after rebalancing. We report turnover because of concerns about this issue in actively managed portfolios, particularly when there is rebalancing each month. Results of our turnover testing are given in the last column of Table 7.

The average monthly turnover of the 1/n portfolio is observed to be 6.5%, and is thus similar to the ones reported in DeMiguel et al. (2009b). The value-weighted portfolio has a turnover smaller than the 1/n portfolio. All other turnover rates at the NC-efficient front portfolios exceed the turnover rate of the 1/n portfolio. Since in active portfolio management there always seem to be shares bought and sold between reporting dates, this leads to an

[6] Let $\bar{\mu}_q, \bar{\mu}_r, \sigma_q, \sigma_r$ and $\sigma_{q,r}$ denote the calculated means, standard deviations, and covariances of the out-of-sample rewards of portfolios q and r over a sample of size L. We evaluate the p values of the difference using the results of the asymptotic properties of the test statistic $f(\eta) = (\lambda_\mu \bar{\mu}_q - \sigma_q^2) - (\lambda_\mu \bar{\mu}_r - \sigma_r^2)$ for two different portfolio strategies r and q and the estimators for means and variances pooled in $\eta = (\lambda_\mu \bar{\mu}_q, \lambda_\mu \bar{\mu}_r, \sigma_q^2, \sigma_r^2)$ following Greene (2002), who shows

$$\sqrt{M}\left(f(\hat{\eta}) - f(\eta)\right) \sim \mathrm{N}\left(0, \frac{\partial f'}{\partial \eta}\Theta\frac{\partial f}{\partial \eta}\right) \quad \text{with} \quad \Theta = \begin{pmatrix} \sigma_q^2 & \sigma_{q,r} & 0 & 0 \\ \sigma_{q,r} & \sigma_r^2 & 0 & 0 \\ 0 & 0 & 2\sigma_q^4 & 2\sigma_{q,r}^2 \\ 0 & 0 & 2\sigma_{q,r}^2 & 2\sigma_r^4 \end{pmatrix}. \tag{7}$$

increase in turnover compared to passive strategies, in which only small adjustments to existing weights are normally required. Nevertheless, with regard to the turnover rates for the active portfolio management approaches described in DeMiguel et al. (2009b), our turnover rates are modest and do not jeopardized this research.

5 Conclusion

Whereas most research on sustainability and climate risk in investing has focused on how sustainability funds perform on returns, this paper is different in that we also focus on whether sustainability funds do all they can to offer investment solutions with the least climate risk content for their investors. This is important to serious, or "motivated" to use the term from the ESG paper by Pedersen et al. (2021), climate risk investors. Our findings are that significant amounts of climate risk are overlooked, but can be reduced in portfolios without significant losses in risk-adjusted returns provided appropriate methods are used. The situation is not obvious. Fund managements are unaware of what they are not taking care of, and clients are unaware of what reduction in climate risk they are missing. But here we spell it out as we can't see why both sides wouldn't be interested in the perspective.

Reduction potential is overlooked because it is totally not seen by the standard procedures used for constructing optimal portfolios (portfolio approaches only considering financial objectives). Nobody picks up a $100 bill if nobody sees it. But to reliably uncover amounts of climate risk that are very hard to know exist, one needs to treat a serious investor's interest in climate risk as, beyond risk and return, a third objective. With a third objective, the efficient frontier becomes an efficient surface. This is important because on the efficient surface are all points of Pareto optimal risk, return, climate risk trade-off. Since one's point of optimal three-way trade-off could be anywhere on the efficient surface, NC-efficient fronts as applied in this paper are used for exploring the efficient surface and this is how portfolios with significantly less climate risk without significantly diminished risk-adjusted returns can be found. The difficulty of the approach is that efficient surfaces and NC-efficient fronts require techniques of parametric quadratic programming to construct as shown in Hirschberger et al. (2013) and Steuer and Utz (2023), techniques known in operations research but not yet well known in finance.

Acknowledgements The authors are grateful to Markus Hirschberger, Yue Qi, Maximilian Wimmer, and to the seminar participants at the University of Zurich; the WU Vienna; the KLU Hamburg; the Green Summit, Vaduz; the Computational Management Science Conference, Trondheim; the Mini-Conference on Large-Scale Portfolio Investment (Beyond Markowitz Portfolio Theory) organized by Ignacy Kaliszewski at the Polish Academy of Sciences, Warsaw; and GOR 2023, Hamburg for their helpful comments.

Declarations

Conflict of interest Both authors declare that they have no Conflict of interest.

References

Adler, T., & Kritzman, M. (2008). The cost of socially responsible investing. *Journal of Portfolio Management, 35*(1), 52–56. https://doi.org/10.3905/JPM.2008.35.1.52

Bolton, P., & Kacperczyk, M. (2021). Do investors care about carbon risk? *Journal of Financial Economics, 142*(2), 517–549. https://doi.org/10.1016/j.jfineco.2021.05.008

Carswell, A. (2002). Crisis of conscience. *Australian CPA, 72*(3), 26–27.

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

DeMiguel, V., Garlappi, L., Nogales, F. J., & Uppal, R. (2009a). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *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? *The Review of Financial Studies, 22*(5), 1915–1953. https://doi.org/10.1093/rfs/hhm075

Dorfleitner, G., & Utz, S. (2014). Profiling German-speaking socially responsible investors. *Qualitative Research in Financial Markets, 6*(2), 72–82. https://doi.org/10.1108/QRFM-07-2012-0024

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

Görgen, M., Jacob, A., Nerlinger, M., Riordan, R., Rohleder, M., & Wilkens, M. (2020). Carbon risk. (SSRN Working Paper 2930897).

Greene, W. H. (2002). *Econometric analysis*. New York: Prentice-Hall.

Hirschberger, M., Steuer, R. E., Utz, S., Wimmer, M., & Qi, Y. (2013). Computing the nondominated surface in tri-criterion portfolio selection. *Operations Research, 61*(1), 169–183. https://doi.org/10.1287/opre.1120.1140

Hofmann, E., Hoelzl, E., & Kirchler, E. (2008). A comparison of models describing the impact of moral decision making on investment decisions. *Journal of Business Ethics, 82*(1), 171–187. https://doi.org/10.1007/s10551-007-9570-6

Hong, H., Wang, N., & Yang, J. (2023). Welfare consequences of sustainable finance. *The Review of Financial Studies, 36*(12), 4864–4918. https://doi.org/10.1093/rfs/hhad048

Humphrey, J. E., & Tan, D. T. (2014). Does it really hurt to be responsible? *Journal of Business Ethics, 122*(3), 375–386. https://doi.org/10.1007/s10551-013-1741-z

Jobson, J. D., & Korkie, B. M. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *The Journal of Finance, 36*, 889–908.

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

Ledoit, O., & Wolf, M. (2011). Robust performance hypothesis testing with the variance. *Wilmott magazine*, pp 86-89.

Ledoit, O., & Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance, 15*, 850–8859.

Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives, 21*(2), 153–174. https://doi.org/10.1093/acprof:oso/9780195328325.003.0015

López Prol, J., & Kim, K. (2022). Risk-return performance of optimized ESG equity portfolios in the NYSE. *Finance Research Letters, 50*, 103312. https://doi.org/10.1016/j.frl.2022.103312

Mackenzie, C., & Lewis, A. (1999). Morals and markets: The case of ethical investing. *Business Ethics Quarterly, 9*(3), 439–452. https://doi.org/10.2307/3857511

Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance, 7*(1), 77–91. https://doi.org/10.2307/2975974

Markowitz, H. M. (1956). The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics Quarterly, 3*(1–2), 111–133.

Markowitz, H. M. (1959). *Portfolio selection: Efficient diversification of investments*. New York: Wiley.

Memmel, C. (2003). Performance hypothesis testing with the Sharpe ratio. *Finance Letters, 1*, 21–23.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*(3), 703–708. https://doi.org/10.2307/1913610

Pasewark, W. R., & Riley, M. E. (2010). It's a matter of principle: The role of personal values in investment decisions. *Journal of Business Ethics, 93*(2), 237–253. https://doi.org/10.1007/s10551-009-0218-6

Pástor, L., 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

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

Pizzutilo, F. (2017). Measuring the under-diversification of socially responsible investments. *Applied Economics Letters, 24*(14), 1005–1018. https://doi.org/10.1080/13504851.2016.1248279

Riedl, A., & Smeets, P. (2017). Why do investors hold socially responsible mutual funds? *The Journal of Finance, 72*(6), 2505–2550. https://doi.org/10.1111/jofi.12547

Sautner, Z., Van Lent, L., Vilkov, G., & Zhang, R. (2023). Firm-level climate change exposure. *The Journal of Finance, 78*(3), 1449–1498. https://doi.org/10.1111/jofi.13219

Steuer, R. E. (1986). *Multiple criteria optimization: Theory, computation and application*. New York: Wiley.
Steuer, R. E., & Na, P. (2003). Multiple criteria decision making combined with finance: A categorized bibliographic study. *European Journal of Operational Research, 150*(3), 496–515. https://doi.org/10.1016/S0377-2217(02)00774-9
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
Webley, P., Lewis, A., & Mackenzie, C. (2001). Commitment among ethical investors: An experimental approach. *Journal of Economic Psychology, 22*(1), 27–42. https://doi.org/10.1016/S0167-4870(00)00035-0
Zopounidis, C., & Doumpos, M. (2002). Multi-criteria decision aid in financial decision making: Methodologies and literature review. *Journal of Multi-Criteria Decision Analysis, 11*(4–5), 167–186. https://doi.org/10.1002/mcda.333