Impact of financed emissions constraints on the composition of an optimal investment portfolio

Timo Norrkniivilä

School of Science

Bachelor's thesis
Espoo 28.9.2023

Supervisor

Prof. Ahti Salo

Advisor

D.Sc. (Econ.) Matti Koivu

Aalto University
School of Science

Aalto University
School of Science

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of the bachelor's thesis

Author Timo Norrkniivilä

Title Impact of financed emissions constraints on the composition of an optimal investment portfolio

Degree programme Bachelor's Programme in Science and Technology

Major Mathematics and Systems Sciences **Code of major** SCI3029

Teacher in charge Prof. Ahti Salo

Advisor D.Sc. (Econ.) Matti Koivu

Date 28.9.2023 **Number of pages** 29+2 **Language** English

Abstract

Financial institutions contribute to significant indirect environmental effects through their outstanding loans and investments, which has led to a globally standardized methodology to calculate the financed emissions of investment portfolios. As global warming calls for societal decarbonization and companies around the world are required to measure and report an increasing amount of data of their environmental performance, financial institutions are faced with incentives to reduce their financed emissions.

We study how constraining the financed emissions of a stock portfolio affects its performance and industry composition. The stock returns and their dependencies are modelled with Student's t-distributions and a t-copula, which are fitted to realized market price data. Monte Carlo simulation is used to draw return scenarios that are used for calculating the expected portfolio return and risk (Conditional Value-at-Risk). The optimal portfolio is obtained by maximizing the expected return while constraining the maximum allowed risk and the financed emissions of the portfolio.

We find that the financed emissions constraints lead to a systematic decrease in portfolio return, but note that all optimized portfolios perform better than most benchmarks used.

Keywords Financed emissions, CSRD, Conditional Value-at-Risk, copula, Monte Carlo simulation

Aalto University
School of Science

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi

Tekniikan kandidaatintyön tiivistelmä

Tekijä Timo Norrkniivilä	
Työn nimi Päästörajoite-ehtojen vaikutus optimaalisen sijoitusportfolion rakenteeseen	
Koulutusohjelma Teknistieteellinen kandidaattiohjelma	
Pääaine Matematiikka ja systeemitieteet	**Pääaineen koodi** SCI3029
Vastuuopettaja Prof. Ahti Salo	
Työn ohjaaja KTT Matti Koivu	
Päivämäärä 28.9.2023 **Sivumäärä** 29+2	**Kieli** Englanti

Tiivistelmä

Rahoitusalalla toimivien instituutioiden lainat ja sijoitukset pitävät sisällään merkittäviä epäsuoria ympäristövaikutuksia, mikä on johtanut standardisoituun tapaan mitata sijoitusportfolioiden kautta rahoitettuja kasvihuonekaasupäästöjä (engl. financed emissions). Ilmaston lämpenemisen hillitsemiseksi asetettuja hiilineutraaliustavoitteita pyritään edistämään kansallisin ja kansainvälisin laein, jotka velvoittavat yrityksiä mittaamaan ja raportoimaan ympäristövaikutuksiaan entistä tarkemmin. Tämä luo rahoitusalan instituutioille kannustimia vähentää portfolioidensa päästöjä.

Tässä työssä tutkitaan, miten osakeportfolion tuotto ja toimialajakauma muuttuvat, kun portfolion päästöjä rajoitetaan. Osaketuottoja ja niiden välisiä riippuvuuksia mallinnetaan toteutuneisiin hintatietoihin sovitetuilla t-jakaumilla ja t-kopulalla. Sovitetuista jakaumista luodaan Monte Carlo -simuloinnilla tuottoskenaarioita, joiden perusteella lasketaan portfolion tuotto-odotus ja Conditional Value-at-Risk -riskitunnusluku. Optimaalinen portfolio muodostetaan maksimoimalla tuotto-odotus niin, että riskitunnusluku ja portfolion päästöt pidetään tiettyjen enimmäisarvojen alapuolella.

Eri rajoite-ehdoilla optimoitujen portfolioiden vertailu osoittaa, että päästöjen rajoittaminen vähentää systemaattisesti portfolion tuottoa, mutta lähes riippumatta käytetyistä rajoite-ehdoista optimoidut portfoliot suoriutuvat verrokkeja paremmin.

Avainsanat Päästörajoitteet, CSRD, Conditional Value-at-Risk, kopula, Monte Carlo -simulointi

Contents

1 Introduction

Excessive greenhouse gas (GHG) emissions are the root cause of global warming, which poses the threat of serious global socioeconomic impacts (Woetzel et al., 2020). Limiting the warming to 1.5 °C above preindustrial levels is considered crucial to prevent multiple climate tipping points (McKay et al., 2022), but requires accelerating decarbonization of the society (PCAF, 2022; Woetzel et al., 2020). Companies are among the largest GHG emitters globally, and thus play a vital role in mitigating the effects of climate change (Siddique et al., 2021), starting with the task of measuring and reporting carbon emissions. While ever more companies partake in public carbon disclosure voluntarily, in many countries, including the majority of EU countries, the UK and the USA, the reporting of GHG emissions is today mandatory for companies meeting certain criteria (European Parliament, 2014; Shanahan, 2020; EPA, 2009). To support and encourage environmental disclosure, the Carbon Disclosure Project (later renamed to CDP) was established in 2000 as a collective platform for reporting companies' GHG emissions, and has since broadened the scope to various other environmental disclosure measures. CDP may be considered a default carbon reporting medium for companies, as in 2022, companies that disclosed through CDP comprised over half of the global market capitalization (CDP, 2023).

GHG emissions are categorized by the GHG Protocol (WRI and WBCSD, 2004) into scope 1 (direct emissions), scope 2 (emissions from purchased or acquired electricity, steam, heat and cooling) and scope 3 (all other indirect emissions in the corporate value chain). One of the scope 3 categories includes financed emissions, the emissions that an institution finances through its loans and investments, and which typically comprise the largest portion of emissions in the financial sector (PCAF, 2022). While the disclosure of scope 3 emissions has previously not been covered by most carbon regulations, reporting them will be mandatory in the EU starting from 2024, following the introduction of the Corporate Sustainability Reporting Directive (CSRD) (European Commission, 2023). This has in recent years incentivized financial institutions to make efforts to reduce the financed emissions of their investment portfolios and to direct capital to support companies that work towards decarbonization. The question is whether one can reduce financed emissions without a deterioration in the financial performance of the portfolios.

In this thesis, to improve understanding of how the increased emission legislation affects the financial sector, we introduce financed emissions constraints to a stock portfolio optimization model and study their impact on the optimized portfolios. Risk management and stock portfolio optimization have seen major advances in the latest decades, but while relationships between companies' environmental and financial performance have been studied using various methods, very limited research combines environmental measures and modern portfolio optimization methods, creating a gap we aim to fill. Our analysis is restricted to the European and U.S. stock markets as the developed carbon reporting practices and legislation in these areas have been shown to improve carbon data quality (see for example Cai, 2022). These are also considered the most relevant areas from a Nordic investor's point of view.

The rest of the thesis is structured as follows: In Section 2, we discuss advances in portfolio optimization methodology, along with relevant research, and introduce results of previous studies addressing environmental and economic performance. Section 3 first discusses the choice of the companies used in the thesis, and then provides a detailed overview of the applied portfolio optimization model. In Section 4, we analyze the composition and performance of the optimized portfolios, while Section 5 provides a summary of the results along with possible caveats, and discusses future research opportunities.

2 Literature review

Central to the modern portfolio theory, initiated by Markowitz (1952), is finding a balance between the expected return and allowed risk of the portfolio of risk-bearing assets. According to McNeil et al. (2005), "most modern measures of the risk in a portfolio are statistical quantities describing the conditional or unconditional loss distribution of the portfolio over some predetermined horizon Δ". While Markowitz (1952) used variance as the risk measure, a multitude of other, more sophisticated risk measures based on loss distributions have been proposed since. Popular choices in modern finance research include Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR). With a fixed confidence level α, VaR is simply the α-quantile of the loss distribution, while CVaR is the conditional expectation of the losses exceeding the corresponding VaR.

VaR, even though a popular measure of risk, is known to have several undesirable characteristics. Lack of subadditivity, leading to VaR behaving poorly with respect to addition of risks, is perhaps the most notable of these (Artzner et al., 1999). Also, while VaR controls the probability of losses, it by definition fails to take into account the severity of them (Artzner et al., 1999). In addition, VaR is difficult to optimize when calculated from scenarios (Rockafellar and Uryasev, 2000). CVaR solves many of the problems associated with VaR and is thus often preferred over VaR in risk modelling. In this thesis, a linearization of the CVaR risk constraint, as introduced by Rockafellar and Uryasev (2000); Krokhmal et al. (2001), will be implemented.

The optimization problem may be defined in two ways: either maximize the portfolio return while constraining the acceptable level of risk, or minimize the risk while requiring a certain level of return. While we adopt the former approach, the choice between the two does not affect the optimization results; the same risk-return frontier is obtained both ways (Krokhmal et al., 2001).

To calculate the risk, three approaches to construct the loss distribution are usually considered: historical simulation, variance-covariance method and Monte Carlo simulation (Linsmeier and Pearson, 1996; McNeil et al., 2005). Historical simulation means simulating future asset returns by generating them from realized historical scenarios. It is straightforward to implement since one does not have to consider the nature of the underlying return distribution nor the dependency structure between risk factors: these will be assumed to exactly follow what has been observed in the past. However, the reliability of the method is compromised if the amount of historical data is limited or does not contain enough extreme events. The variance-covariance method offers a convenient analytical solution to risk measurement, with the downside of requiring two crude simplifications that are often unjustifiable in risk modelling: the assumption of normally distributed returns and a linear relationship between the loss distribution and the risk-factor changes (McNeil et al., 2005).

In this thesis, Monte Carlo simulation is used to calculate the risk and expected portfolio return. To this end, probability distributions are fitted to historical returns and a large number of random samples from them is generated. The success of Monte Carlo based methods depends largely on how well the chosen distributions describe the underlying phenomena; our model should be able to accurately capture the return distributions of the individual stocks, but also the dependency structure between them, since stock prices have a tendency to move together following prevailing market trends (Embrechts et al., 2003). For capturing the inter-asset dependency structure in the model, copulas have shown promising results as they help model the marginal return distributions separately from the dependency structure between assets, offering more freedom compared to isolated simulation methods (Shekhar and Trede, 2017; Kakouris and Rustem, 2014). A copula, in its simplest form, can be viewed as a multivariate distribution with uniformly distributed margins, making it a probability measure in the unit cube (Nelsen, 2005). The notion was first introduced by Sklar (1959) but only later gained traction in risk modelling applications.

Two important copula families, offering frequently used alternatives for risk modelling, are elliptical copulas and Archimedean copulas (Embrechts et al., 2003; McNeil et al., 2005). Archimedean copulas are important for their ability to model asymmetric tail dependence; as McNeil et al. (2005) point out, in finance it is often reasonable to assume that the dependence between large losses is greater than the dependence between large gains. Ang and Chen (2002) provide empirical support for this by showing correlation asymmetries in U.S. stocks and aggregate market. However, McNeil et al. (2005) show that extending Archimedean copulas to higher dimensions may be problematic and Kakouris and Rustem (2014) point out that simulation from them may be difficult.

Elliptical copulas, the copulas of elliptical distributions, have the benefit of being simple to implement and simulate from (Embrechts et al., 2003). Two commonly studied elliptical copulas include the Gaussian (normal) copula and t-copula, that can be seen as a generalized Gaussian copula, since t-copula approaches the Gaussian copula as its degrees of freedom parameter tends to infinity. While Gaussian copulas have been criticized for their insufficient capability to capture joint extreme moves of the risk factors (McNeil et al., 2005), t-copulas can be used to model stronger tail dependence, making them the preferred copula choice in many existing studies (Kole et al., 2007). For this reason, a t-copula will also be adopted in this thesis, leading to a modelling and simulation methodology loosely inspired by the implementations of Shekhar and Trede (2017); Vauhkonen (2022).

Anquetin et al. (2022) provide a prominent example of constructing carbon sensitive stock portfolios by studying the performance of portfolios that are penalized on all three emission scopes in a constrained mean-variance optimization framework. They are able to cut the emission intensities in half without a significant effect on the portfolio Sharpe ratio. In addition, they find that penalizing scope 3 emissions in addition to scope 1–2 leads to better performing portfolios than by only taking scopes 1–2 into account.

The mean-variance framework is also used by Acerbi (2022) who compares portfolios of S&P 500 stocks decarbonized with three different measures: carbon emissions, carbon intensity and carbon beta (a measure of climate transition risk exposure). The author finds that constraining the carbon intensity produces the best performing portfolio. Chakrabarty and Nag (2023) conduct a literature review of different mathematical models used to measure carbon risk. They note that several studies find higher stock returns in companies with higher carbon emissions. Both Tang and Luo (2014) and Siddique et al. (2021) suggest that high-quality carbon disclosure and the actual carbon performance of a company have a positive relationship, which may stem from impression management endeavors; good carbon performers are likely willing to disclose more information to show sustainability to stakeholders and the general public.

Outside the portfolio optimization framework, the relationship between environmental and financial performance has been analyzed by for example Siddique et al. (2021). The authors show a positive relationship between carbon disclosure and long-term financial performance, but a negative relationship in short-term. Delmas et al. (2015) find that reductions in GHG emissions lead to increased long-term market performance. Both findings are consistent with Horváthová (2010) who conducts a meta-analysis on numerous empirical studies addressing the relationship and concludes that "it takes time for environmental regulation to materialise in financial performance".

3 Methodology

3.1 Asset selection

Assets for the portfolio optimization problem are selected amongst European and U.S. companies that disclosed their GHG emissions through CDP in 2020. Historical market prices of the companies and market indexes (see below) from January 2005 through May 2023 are obtained from Yahoo Finance. Adjusted closing price is used instead of the raw closing price, as the former accounts for various corporate actions affecting the stock price and is thus considered a more accurate measure when comparing the performance of multiple assets (Ganti, 2020). Asset prices are taken from the last available dates of each month. The GHG emissions of the studied companies are provided by CDP (2020) and, in order to calculate an emission intensity for each company, financial statements at fiscal year-end 2019 are obtained from CreditEdge.

Only companies for which all mentioned data are available are considered. To equalize the country distribution and to allow feasible computing times, the list of U.S. companies is narrowed down to approximately a fifth of all possible companies via random sampling. It is also noted that a significant portion of German companies are eliminated due to gaps in market price data before 2010, leaving them underrepresented. Further, countries with less than 10 companies are left out of the analysis. Ultimately, 724 companies are selected to be studied (see Table 1 for their distribution by country). To assess the performance in contrast to the general market, major market indexes from the four most represented countries are used as a benchmark. These include S&P 500 (United States), CAC 40 (France), FTSE 100 (United Kingdom) and SMI (Switzerland).

Table 1: Details of the stock exchanges behind the market price data and the number of companies from each

Country	Stock exchange	Original currency	Companies
Austria	Wiener Börse	EUR	19
Belgium	Brussels Stock Exchange	EUR	28
Denmark	Nasdaq Copenhagen	DKK	22
Finland	Nasdaq Helsinki	EUR	25
France	Euronext Paris	EUR	109
Germany	Börse Berlin	EUR	12
Italy	Borsa Italiana	EUR	37
Netherlands	Euronext Amsterdam	EUR	28
Norway	Oslo Stock Exchange	NOK	20
Sweden	Nasdaq Stockholm	SEK	63
Switzerland	SIX Swiss Exchange	CHF	74
United Kingdom	London Stock Exchange	GBP	99
United States	New York Stock Exchange	USD	188

All data in a currency other than euro (see Table 1) are converted to euro, by dividing by the historical currency exchange rate at the given date. For each company, an emission intensity (in CO2 equivalent tonnes per a million euros) is calculated following the convention for listed enterprises outlined by PCAF (2022) as closely as possible: for each company, the sum of absolute scope 1 and scope 2 emissions is divided by the sum of total equity and debt. We note that several papers, including (Anquetin et al., 2022; Acerbi, 2022), use the company's revenue as the emission intensity denominator rather than the standard set by PCAF (2022). The standardized convention is a natural methodological choice as we calculate the financed emissions of portfolios, but may provide slightly deviating figures compared to the revenue-based approach.

The mean and median financed emission intensities are calculated separately for each industry (see Table 2). For comparison, the mean and median of the revenue-based emission intensity provided by CDP is also presented in Table 2. The results may be considered as expected, with the highest emissions observed in sectors such as fossil fuels and transportation, and the lower end being represented by services and biotech, health care & pharma. The infrastructure sector demonstrates a large difference in the median and mean emission intensity, which may indicate that the sector contains companies from a broad scale of activity groups.

Table 2: Mean and median financed emission intensity (FE) and the revenue-based emission intensity (RI) by industry, ordered by the median of the former. The total number of companies from each industry is in parenthesis, and all values are in tCO2e/MEUR.

	FE		RI	
Industry	Mean	Median	Mean	Median
Transportation services (19)	835.28	724.33	731.96	604.71
Power generation (12)	835.14	513.31	2524.4	1088.9
Fossil Fuels (15)	1246.1	367.58	1056.5	478.83
Apparel (2)	328.34	328.34	187.91	187.91
Materials (57)	686.74	323.97	682.58	275.51
Food, beverage & agriculture (37)	105.25	63.345	91.948	53.250
Hospitality (19)	136.80	44.797	84.846	29.077
Retail (47)	74.260	44.628	36.862	15.355
Manufacturing (137)	82.608	41.691	50.031	24.131
Biotech, health care & pharma (61)	31.319	21.701	34.017	34.333
Infrastructure (55)	194.99	20.139	324.25	36.175
Services (263)	41.046	7.0823	76.559	9.7339
All data (724)	179.21	31.473	237.42	27.345

3.2 Modelling returns

Log-returns will be used to measure changes in asset prices. This naturally follows from using the logarithmic asset prices as risk factors, a standard practice in financial risk management (McNeil et al., 2005; Manzanares and Schwartzlose, 2009). Let $K_{i,t}$ be the price of asset i at time t. Then, the (1-period) log-return of asset i at t is given by

$$r_{i,t} = \ln \left(\frac{K_{i,t}}{K_{i,t-1}} \right) = \ln K_{i,t} - \ln K_{i,t-1} \tag{1}$$

In this thesis, one-month intervals are used to calculate the returns. This has two clear advantages over daily returns, arguably the most common convention in stock return modelling. As McNeil et al. (2005) point out, volatility clustering (i.e. the tendency of extreme price changes appearing successively) of the returns is less apparent in longer intervals. This makes it more reasonable to assume the independence and homoscedasticity of monthly returns compared to daily returns so that one does not have to account for conditional variance with methods such as GARCH filtering. Furthermore, a longer interval makes missing data a lesser problem. We construct an international portfolio, using price data from multiple stock exchanges that are closed on different days in concordance to different national holidays; while it is relatively straightforward to use the last available price value of each month, trying to fill the gaps in daily data would be a much more tedious task and possibly affect the reliability of the results. The downside to using a longer interval is that there are fewer data points, which can make it difficult to obtain a reliable fit, unless a much longer training period is used. The number of data points can be preserved by using overlapping periods, but as this leads to additional serial dependencies that have to be factored in in order to stationarize the return series, it is generally not recommended (McNeil et al., 2005).

The choice of the training period (i.e. the data used to fit the distributions) can significantly affect the return distribution and thus has to be considered carefully for more reliable optimization results. Naturally, more recent market behaviour can be assumed to be more indicative of future performance of an asset, so in choosing length of the training period, the challenge lies in obtaining enough data, but simultaneously avoiding data that can be considered outdated and thus misleading. One also has to consider the frequency of extreme market events that have occurred during the training period and keep in mind that using heavy-tailed distributions, especially those with infinite variance, tend to exaggerate the frequency of extreme values.

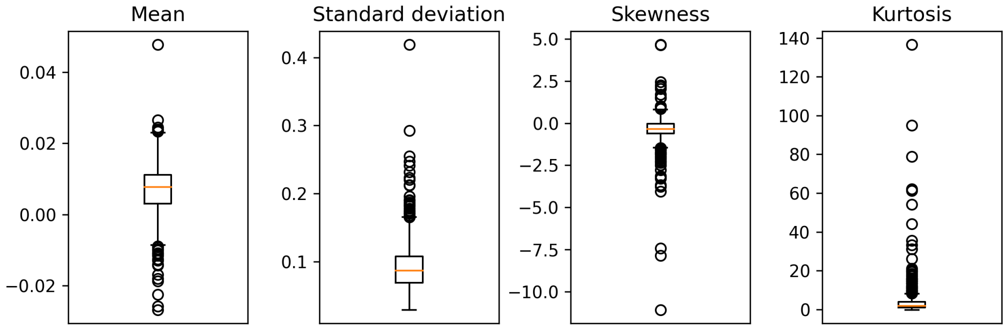
We use monthly market prices from January 2005 until December 2019 for fitting the distributions (referred to as the in-sample or the training period), and prices from December 2019 until May 2023 for out-of-sample performance analysis of the optimized portfolios. The length (in months) of the in-sample period is denoted by $t_0 = 179$ (note that $t = 1$ corresponds to February 2005, the first possible instance to calculate the return). While the 2007-2008 financial crisis is included in the training period, the effects of the COVID-19 pandemic starting early 2020, and the Russian invasion of Ukraine in 2022, are deliberately excluded and instead used for testing. As the latter two events have contributed to noticeable shifts in market trends, the out-of-sample analysis is expected to provide an overview of the robustness of the portfolios to crises.

3.3 Univariate distributions of the return series

We use Student's t-distribution for modelling the log-returns of individual assets, as it is widely used in existing research to capture the heavy tails that are often observed in financial risk management settings (Officer, 1972; McNeil et al., 2005). Note that in this thesis, t-distribution refers to the location-scale t-distribution $lst(\mu, \tau^2, \nu)$, where μ is the location (expected value) of the distribution, τ is a scaling parameter and ν is the degrees of freedom parameter. The link to the standard t-distribution $t(\nu)$ with ν degrees of freedom is

$$Y = \mu + \tau X, \text{ where } Y \sim lst(\mu, \tau^2, \nu), \ X \sim t(\nu), \tag{2}$$

which shows the convention of defining μ and τ as properties of the distribution itself. The t-distribution approaches the normal distribution as the degrees of freedom parameter tends to infinity, and can thus be seen as a generalization of the normal distribution, providing more freedom in fitting.



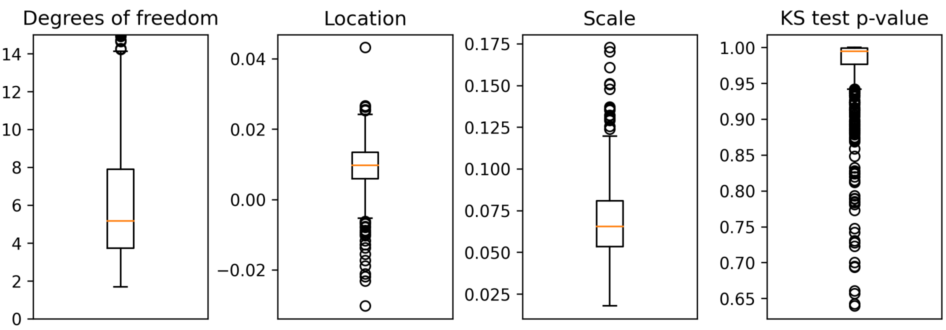
	Mean	Standard deviation	Skewness	Kurtosis
100 %	0.047727	0.41897	4.6612	136.67
75 %	0.011059	0.10777	−0.041060	3.8510
50 %	0.007705	0.086820	−0.33752	1.8987
25 %	0.003064	0.069161	−0.61810	0.90534
0 %	−0.026869	0.029033	−11.083	−0.34813
mean	0.006936	0.092856	−0.40647	3.9023

Figure 1: Distribution of the mean, standard deviation, skewness and kurtosis of the monthly log-return series

By visual inspection, the return series $\boldsymbol{r_i} \in \mathbb{R}^{t_0}$ of each individual asset are fairly symmetrically distributed and follow a bell-like shape matching that of a t-distribution (examples in Figure 3). To further assess their characteristics, the first four moments are calculated for each series (see Figure 1, with standard deviation instead of variance describing the second moment). The percentage values in the table refer to the corresponding quantiles such that the 0 %-quantile is the minimum value and the 100 %-quantile the maximum. The kurtosis values describe excess kurtosis, where the normal distribution has 0 kurtosis. It is apparent that most of the log-return series are leptokurtic, showing signs of heavier tails than a normal distribution would be capable of modelling. Most series have a skewness close to zero, but outliers exhibiting both positive and negative skewness are observed. Over 75 % of the assets exhibit a negative skewness, but as most of the values land very close to 0, a symmetric t-distribution is assumed to provide a sufficient fit. The distributional hypothesis can be expressed as

$$r_{i,t} \sim lst(\mu_i, \tau_i^2, \nu_i). \tag{3}$$

A t-distribution is fitted to each return series $\boldsymbol{r_i}$ separately, using the `t.fit()` method in the `scipy.stats` Python package, obtaining maximum likelihood estimators for the distribution parameters. We denote by n the number of assets in the data; $\boldsymbol{\mu} \in \mathbb{R}^n$ are the location parameters, $\boldsymbol{\tau} \in \mathbb{R}^n_+$ the scale parameters and $\boldsymbol{\nu} \in \mathbb{R}^n_+$ are the degrees of freedom of the fitted t-distributions. The Kolmogorov-Smirnov (KS) test is performed for each of the fitted distributions to assess goodness of fit. The distributions of the fitted parameter estimators and the p-values of the KS tests are presented in Figure 2. Note that the boxplot of the degrees of freedom parameters has been rescaled to allow viewing the majority of data in more detail. Notably, the minimum p-value of the KS test among the assets is 0.64, which is well above any commonly used significance levels and thus does not justify rejecting the distributional hypothesis for any of the assets.



100 %	$8.9854 \cdot 10^6$	0.043198	0.17278	1.0000
75 %	7.8821	0.013370	0.080734	0.99921
50 %	5.1524	0.009707	0.065391	0.99446
25 %	3.7121	0.005870	0.53395	0.97625
0 %	1.6818	-0.030191	0.017979	0.63983
mean	56172	0.009217	0.068712	0.97321

Figure 2: Distribution of the parameter estimators in fitted t-distributions, and the p-values of the Kolmogorov-Smirnov tests

As an example, Figure 3 provides visualizations of the return distributions of 4 companies in our data, along with the fitted t-distribution. Empirical CDF of the returns and the CDF of the t-distribution are also shown; the KS test measures goodness of fit based on the largest vertical distance between the two.

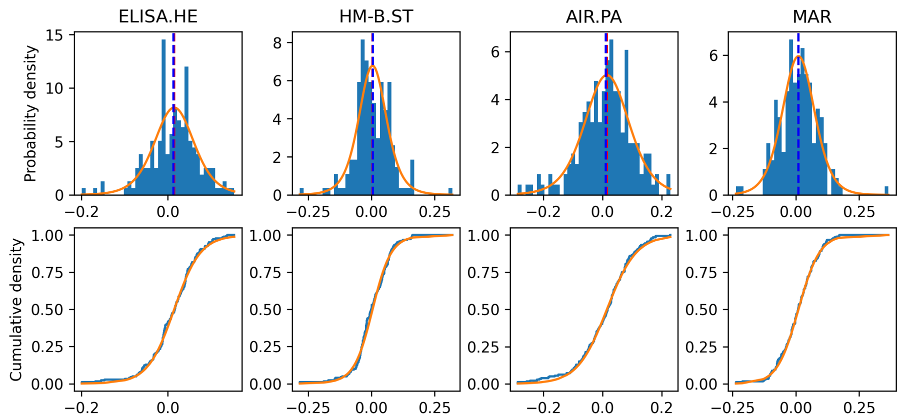


Figure 3: [Top] Distribution of log-returns of Elisa Oyj, Hennes & Mauritz AB, Airbus SE and Marriott International, Inc. from Feb 2005 to Dec 2019, the empirical mean (blue dashed), the fitted t-distribution (orange) and its location (red dashed). [Bottom] ECDF of the returns (blue) and CDF of the fitted t-distributions (orange).

3.4 Modelling dependency with a t-copula

The dependencies between the return series are modelled with a t-copula which is here defined as in (Embrechts et al., 2003). Let $t^n_{\nu,R}$ be the CDF of the standard multivariate t-distribution in n dimensions with ν degrees of freedom and the linear correlation matrix $R \in \mathbb{R}^{n\times n}$. The multivariate t-distribution itself is denoted by $t(\nu, R)$. Also, let t_ν^{-1} be the inverse CDF of the standard univariate t-distribution with ν degrees of freedom. The standard n-dimensional t-copula couples together a multivariate t-distribution and n standard t-distributions with equal degrees of freedom. It is defined as

$$C^t_{\nu,R} : [0,1]^n \to [0,1], \quad C^t_{\nu,R}(\boldsymbol{u}) = t^n_{\nu,R}(t_\nu^{-1}(u_1), ..., t_\nu^{-1}(u_n)), \tag{4}$$

where $\boldsymbol{u}$ follows the standard uniform distribution. Now denote the cumulative distribution function (CDF) of a certain return distribution $lst(\mu_i, \tau_i^2, \nu_i)$ by

$$F_i : \mathbb{R} \to [0,1]. \tag{5}$$

By the probability integral transform (PIT) theorem, $F_i(r_{i,t})$ follows approximately the standard uniform distribution. Thus, by performing a PIT to each of the return series $\boldsymbol{r_i}$ by applying F_i, we obtain a collection of n vectors $\boldsymbol{u_i} \in \mathbb{R}^{t_0}$ that approximately follow the standard uniform distribution and can be used for fitting the copula. This collection is denoted by $u \in [0,1]^{n\times t_0}$.

Maximum likelihood estimation (MLE) is used to find the optimal degrees of freedom $\hat{\nu}$ and the corresponding correlation matrix $R(\hat{\nu}) = \hat{R}$ of the t-copula $C^t_{\hat{\nu},\hat{R}}(\boldsymbol{u})$. The likelihood function, subject to maximization, is (McNeil et al., 2005)

$$\mathcal{L}(\nu|u) = \sum_{t=1}^{t_0} \ln f_{\nu,R}(t_\nu^{-1}(u_{i,t}), ..., t_\nu^{-1}(u_{n,t})) - \sum_{t=1}^{t_0}\sum_{i=1}^{n} \ln f_\nu(t_\nu^{-1}(u_{i,t})), \tag{6}$$

The maximum likelihood estimator for the degrees of freedom is

$$\hat{\nu} = \underset{\nu}{\operatorname{argmax}}\, \mathcal{L}(\nu|u). \tag{7}$$

The high dimensionality in our setting makes (6) relatively expensive computationally; note that also the correlation matrix has to be separately computed for each value of ν. However, we assume that integer degrees of freedom provide sufficient accuracy, and maximize $\mathcal{L}(\nu|u)$ via numerical testing over integer solutions ν only. As Figure 4 shows, the likelihood function attains its maximum at $\nu = 11$ and as such, $\hat{\nu} = 11$ is selected as the degrees of freedom of the t-copula.

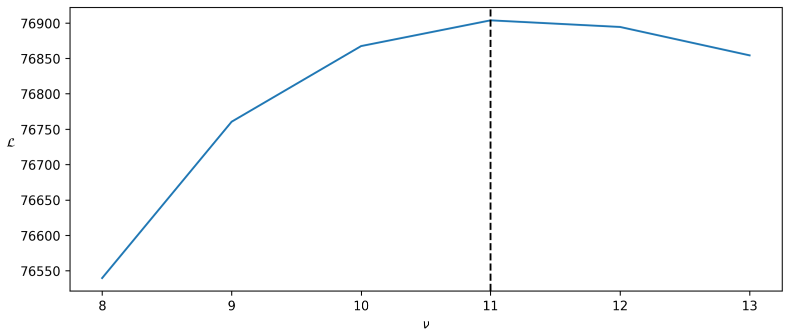


Figure 4: Value of the likelihood function for $\nu \in [8, 9, ..., 13]$.

3.5 Scenario creation

Future returns are simulated by generating random samples from the fitted copula, equivalent to generating samples from the multivariate t-distribution $t(\hat{\nu}, \hat{R})$ and performing a PIT to these samples (see Equation 4). Implementing the latter approach, we generate N random samples from $t(\hat{\nu}, \hat{R})$ and transform each element $\tilde{q}_{j,i}$ in this collection with $t_{\hat{\nu}}(\tilde{q}_{j,i})$. This gives a collection $\tilde{u} \in \mathbb{R}^{N \times n}$ of uniformly distributed samples, which are further transformed into log-return scenarios using the inverse CDF of the corresponding univariate marginal distributions:

$$\tilde{r}_{j,i} = F_i^{-1}(\tilde{u}_{j,i}). \tag{8}$$

Each row $\tilde{\boldsymbol{r}}_j \in \mathbb{R}^n$ represents an individual scenario of monthly log-returns of each asset. At this point we note that due to randomness of the data, and especially the infinite variance property of the marginal distributions with $\nu_i \in (1, 2]$, the simulated data may contain extreme outliers that are not truly representative of the underlying market behaviour. As an example, note that since $e^4 \approx 55$, a value $\tilde{r}_{j,i} = 4$ would correspond to a 55-fold increase in the asset price during one month, which is deemed nearly impossible under normal circumstances. We opt for a simple approach to mitigate the undesired effects of the outliers in optimization; any simulated scenarios containing values above 4 or below -4 are removed.

For the optimization problem, the portfolio risk and expected return are inferred for a 1-year period. From (1) we get

$$\ln K_{i,t} = \ln K_{i,t-1} + r_{i,t}, \tag{9}$$

and then the 12-month cumulative return of asset i is given by

$$\ln K_{i,12} - \ln K_{i,0} = \sum_{t=1}^{12} r_{i,t}, \tag{10}$$

where $\ln K_{i,0}$ is the initial value of the asset. The individual returns of asset i are assumed to be independent and identically distributed as characterized by the marginal distributions $lst(\mu_i, \tau_i^2, \nu_i)$. Furthermore, the scenarios $\tilde{\boldsymbol{r}}_{\boldsymbol{j}}$ are assumed to be independent, which allows us to create 12-month return scenarios by simply adding together 12 monthly scenarios at a time; the resulting collection of 12-month log-return scenarios will be denoted by $\tilde{s}$. The number of scenarios $\tilde{\boldsymbol{s}}_{\boldsymbol{j}}$ is $J \leq \frac{N}{12}$, where the inequality is caused by the removal of outliers. We generate $N = 120000$ monthly scenarios, and after the aggregation, 58 outlying scenarios are removed to obtain $J = 9942$ yearly scenarios. Figure 5 shows the distribution of the simulated scenarios of 4 companies, demonstrating that the scenario count is sufficient to provide fair convergence of the distribution of yearly scenarios.

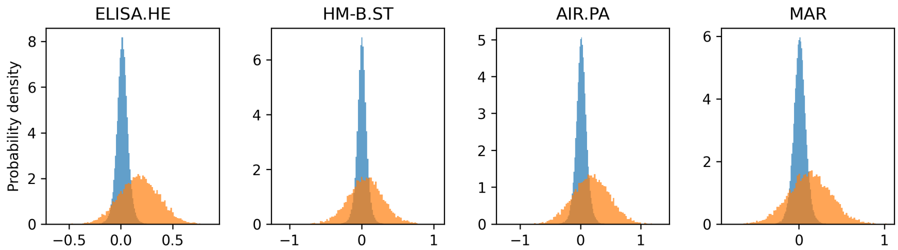


Figure 5: Distribution of the the simulated 1-month log-return scenarios (blue) and the 12-month log-return scenarios aggregated from the 1-month scenarios (orange) of Elisa Oyj, Hennes & Mauritz AB, Airbus SE and Marriott International, Inc.

3.6 Constructing the optimization problem

3.6.1 Loss function

The relative change in asset value – rate of return – forms the basis for the portfolio loss function. The expected rate of return of asset i during the optimization period is approximated from the simulated scenarios as the average of exponentiated log-returns

$$\tilde{y}_i = \frac{\sum_{j=1}^{J} e^{\tilde{s}_{j,i}}}{J} \approx \mathbb{E}\left[\frac{K_{i,t}}{K_{i,t-12}}\right], \tag{11}$$

which gives a vector $\tilde{\boldsymbol{y}} \in \mathbb{R}^n$ of expected 12-month return rates. Let $\boldsymbol{x} \in \mathbb{R}^n$ be the decision variable in the optimization model, describing the amount invested in each asset. The total initial value of the portfolio is $V = \sum_{i=1}^{n} x_i$. As it is assumed that a continuous amount can be invested to each asset, and transaction costs are not accounted for in the model, the choice of V does not affect the optimal portfolio composition. We will use the value $V = 1$ which allows interpreting the losses and investments proportionally. The loss function of the portfolio, describing the negative change of the portfolio value during the optimization period, will now take the form

$$L(\boldsymbol{x}|\tilde{\boldsymbol{y}}) = -\boldsymbol{x}^T\tilde{\boldsymbol{y}} + V. \tag{12}$$

Since V defined as a constant, it may be omitted from the objective function and the term to be minimized will be $-\boldsymbol{x}^T\tilde{\boldsymbol{y}}$, the negative expected portfolio value at the end of the investment horizon.

3.6.2 Conditional Value-at-Risk constraint

In our scenario-based approach, the CVaR risk constraint is easily linearized along the lines proposed by Rockafellar and Uryasev (2000) and Krokhmal et al. (2001). We first consider the function

$$F_\alpha(\boldsymbol{x}, \zeta) = \zeta + \frac{1}{1-\alpha} \int_{\boldsymbol{y} \in \mathbb{R}^n} [L(\boldsymbol{x}|\boldsymbol{y}) - \zeta]^+ p(\boldsymbol{y}) d\boldsymbol{y}, \tag{13}$$

where ζ can be seen as a threshold parameter for the loss, and $[t]^+ = \max\{t, 0\}$. Random sampling of return scenarios makes it possible to approximate $F_\alpha(\boldsymbol{x}, \zeta)$ via the discretization

$$\tilde{F}_\alpha(\boldsymbol{x}, \zeta) = \zeta + \frac{1}{J(1-\alpha)} \sum_{j=1}^{J} [L(\boldsymbol{x}|\tilde{\boldsymbol{s}}_{\boldsymbol{j}}) - \zeta]^+, \tag{14}$$

where, as a result of the applied sampling method, it is assumed that each scenario $\tilde{\boldsymbol{s}}_{\boldsymbol{j}}$ has equal probability. Note that the interpretation for $\frac{1}{J}\sum_{j=1}^{J}[L(\boldsymbol{x}|\tilde{\boldsymbol{s}}_{\boldsymbol{j}}) - \zeta]^+$ is the empirical mean of the losses exceeding ζ. For a fixed $\boldsymbol{x}$, the empirical CVaR $\hat{\phi}_\alpha$ corresponding to the threshold α is then found by minimizing $\tilde{F}_\alpha$

$$\hat{\phi}_\alpha(\boldsymbol{x}) = \min_{\zeta \in \mathbb{R}} \tilde{F}_\alpha(\boldsymbol{x}, \zeta). \tag{15}$$

Frequently used values of the confidence level α include 0.90, 0.95 and 0.99 (Rockafellar and Uryasev, 2000) and with a higher confidence level, only more extreme losses are taken into account. We use a confidence level $\alpha = 0.99$ to obtain a sufficient safety margin. As an additional remark, the corresponding Value-at-Risk (VaR) is defined to be the left endpoint of the set

$$Z_\alpha = \operatorname*{argmin}_{\zeta \in \mathbb{R}} \tilde{F}_\alpha(\boldsymbol{x}, \zeta). \tag{16}$$

It is noted that Z_α may reduce to a single point, but in general is a nonempty, closed and bounded interval (Rockafellar and Uryasev, 2000). Now, let ω describe the maximum tolerated CVaR of the loss distribution so that

$$\hat{\phi}_\alpha(\boldsymbol{x}) \leq \omega. \tag{17}$$

We refer to (15) and note that by definition $\hat{\phi}_\alpha(\boldsymbol{x}) \leq \tilde{F}_\alpha(\boldsymbol{x}, \zeta)$ for all values of ζ. Thus, by introducing ζ as a variable in the optimization problem, the CVaR constraint can be expressed as

$$\tilde{F}_\alpha(\boldsymbol{x}, \zeta) \leq \omega \tag{18}$$

and linearized by introducing dummy variables z_j as the summands:

$$\zeta + \frac{1}{J(1-\alpha)} \sum_{j=1}^{J} z_j \leq \omega \tag{19}$$

$$z_j \geq L(\boldsymbol{x}|\tilde{\boldsymbol{s}}_{\boldsymbol{j}}) - \zeta, \quad z_j \geq 0, \quad j = 1, ..., J, \quad \zeta \in \mathbb{R}. \tag{20}$$

3.6.3 Financed emissions and diversification constraints

Let $\boldsymbol{E} \in \mathbb{R}^n$ be the vector of calculated emission intensities (in tCO2e/MEUR) of each individual company (see Section 3.1). We can express the financed emissions of the portfolio as the simple product $\boldsymbol{E}^T \boldsymbol{x}$. Let the maximum allowed financed emissions be γ, giving the financed emissions constraint

$$\boldsymbol{E}^T \boldsymbol{x} \leq \gamma \tag{21}$$

In addition, to even out the composition of the portfolio, we implement the following diversification constraints:

1. A single asset may constitute at most 2.5 % of the portfolio value ($\beta = 0.025$)
2. A single industry may constitute at most 20 % of the portfolio value ($\beta_A = 0.20$)
3. A single country may constitute at most 20 % of the portfolio value ($\beta_B = 0.20$)

Let I be the identity matrix in n dimensions, $A \in \mathbb{R}^{U \times n}$ be the industry indicator matrix, and $B \in \mathbb{R}^{W \times n}$ be the country indicator matrix, where $U = 12$ is the number of different industries and $W = 13$ is the number of different countries in the data. The industry indicator matrix is defined by

$$A_{u,i} = \begin{cases} 1, & \text{if company } i \text{ belongs to industry } u \\ 0, & \text{otherwise} \end{cases}, \; u = 1, ..., U, \; i = 1, ..., n,$$

and the country indicator matrix follows a similar definition. This leads to the diversification constraints

$$\begin{aligned} I\boldsymbol{x} &\leq \beta \\ A\boldsymbol{x} &\leq \beta_A \\ B\boldsymbol{x} &\leq \beta_B, \end{aligned} \tag{22}$$

where the right sides correspond to constant vectors in n dimensions.

3.6.4 Solving the problem

In order to make the optimization problem feasible for all $\omega \geq 0$, $\gamma \geq 0$, we incorporate into the model a constant asset, which describes the cash amount not exposed to risk. The corresponding return scenarios are described with a vector consisting of 1s only, which is added as a column in $\tilde{s}_j$. The amount invested in the constant asset is not constrained; the corresponding values in I, A B and $\boldsymbol{E}$ are set to 0. The optimization problem takes the following form:

$$\begin{aligned}
&\min_{\boldsymbol{x},\zeta,\boldsymbol{z}} \ -\boldsymbol{x}^T\tilde{\boldsymbol{y}} \\
\text{s. t.}\quad &\zeta + \frac{1}{J(1-\alpha)}\sum_{j=1}^{J} z_j \leq \omega \\
&z_j \geq -\boldsymbol{x}^T\tilde{\boldsymbol{s}}_{\boldsymbol{j}} + V - \zeta \\
&\sum_{i=1}^{n} x_i = V \\
&\boldsymbol{E}^T\boldsymbol{x} \leq \gamma \\
&I\boldsymbol{x} \leq \beta \\
&A\boldsymbol{x} \leq \beta_A \\
&B\boldsymbol{x} \leq \beta_B \\
&x_i \geq 0, \quad i = 1, ..., n, \quad z_j \geq 0, \quad j = 1, ..., J, \quad \zeta \in \mathbb{R},
\end{aligned}$$

where ζ and $\boldsymbol{z}$ are obtained as byproducts of the optimization. To implement and solve this optimization problem, the `modeling` module from the CVXOPT Python package is used for its capability of interpreting individually defined constraints that are not necessarily in standard form. Seven risk levels and three emission constraint levels are used for optimization:

$$\omega \in \{0.01, 0.03, 0.05, 0.10, 0.15, 0.20, 0.30\}$$
$$\gamma \in \{\infty, 30, 15\}$$

The emission constraint level 30 tCO2e/MEUR is selected as it is roughly equal to the median emission intensity (see Table 2), and the level 15 tCO2e/MEUR is chosen to halve the threshold. Note that when V and $\boldsymbol{x}$ are defined to describe proportional (unitless) values, γ should be interpreted as an emission intensity (tCO2e/MEUR), whereas by associating a unit of 1 million euros to the values of V and $\boldsymbol{x}$, γ can be interpreted as the absolute financed emissions (tCO2e) of the portfolio. The latter interpretation will be adopted in the next section when applicable.

4 Results

4.1 Optimization results

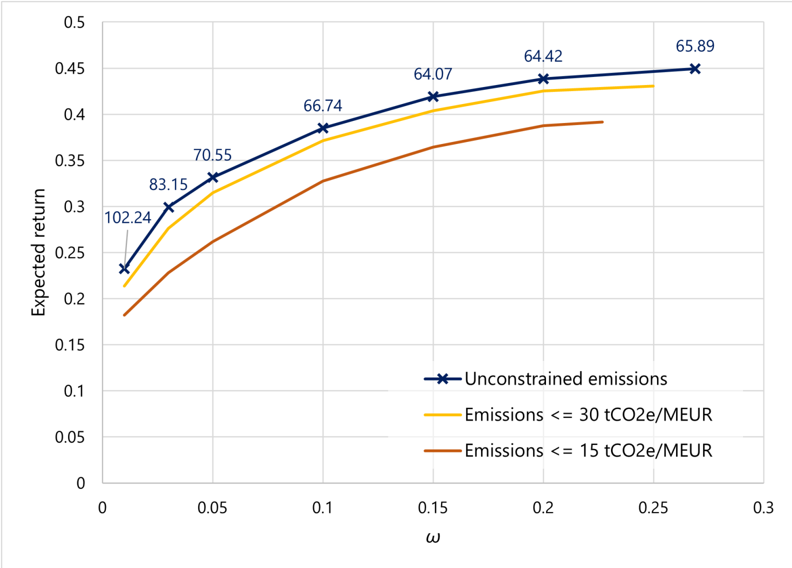


Figure 6: Expected risk-return frontiers of the optimized portfolios with three different maximum emission intensity levels. Above the portfolios with unconstrained emissions is the calculated emission intensity of the portfolio.

Solving the optimization problem with these 3 emission constraints and 7 risk constraints gives the efficient risk-return frontiers presented in Figure 6. The horizontal axis corresponds to the calculated portfolio risk (0.99-CVaR), obtained as the mean of the losses in the worst 1 % scenarios. It is noted that the highest risk tolerance constraint $\omega = 0.3$ is not active for any value of γ. With the lowest values of ω, all of the portfolio share is not exposed to risk, meaning that in the optimal portfolio composition, some of value is designated to the constant asset. The emission contribution of the constant asset is also assumed to be zero. This should be taken into account when comparing the performance of different portfolios. All parameter combinations resulting in constant asset investments are in Table 3.

Table 3: Parameter combinations resulting in some of the portfolio value not being exposed to risk, and the amount not exposed to risk in these portfolios

γ	$\omega = 0.01$	$\omega = 0.03$	$\omega = 0.05$
∞	0.144	0	0
30	0.245	0.079	0
15	0.361	0.205	0.135

Unconstrained, the financed emissions of most portfolios settle to around 65 tCO2e (see Figure 6). Thus, with most risk thresholds, the constraints restrict financed emissions to slightly less than a half (with $\gamma = 30$) and a quarter (with $\gamma = 15$), compared to the portfolio with unconstrained emissions. The portfolios with lower risk thresholds have higher financed emissions, which may happen by chance, or indicate that on average, companies with lower emissions exhibit a higher volatility. As an example, large fossil fuel companies often have very robust and well defined business models with a fairly predictable operating income.

With unconstrained emissions and CVaR, the expected 12-month return is around 45 %. Constraining the financed emissions to 30 tCO2e reduces the expected return by 3–5 % (in relative terms) with $\omega \geq 0.05$, and by around 8 % with $\omega \in \{0.01, 0.03\}$. Constraining the emissions to 15 tCO2e reduces the expected return by 12–15 % with $\omega \geq 0.1$ and by 21–24 % with $\omega \leq 0.05$. Taking into account the constant asset investments (Table 3), the reduction in expected performance is fairly constant across different values of ω.

We find that the industry composition of the portfolios changes relatively little from restricting the financed emissions, which may be partly caused by the industry diversification constraints already in place. With lower risk thresholds, the industry composition is observed to be more diverse, while more diversification constraints are active with higher allowed risk. For the detailed industry composition of 12 optimized portfolios, see Appendix B.

4.2 Out-of-sample portfolio performance

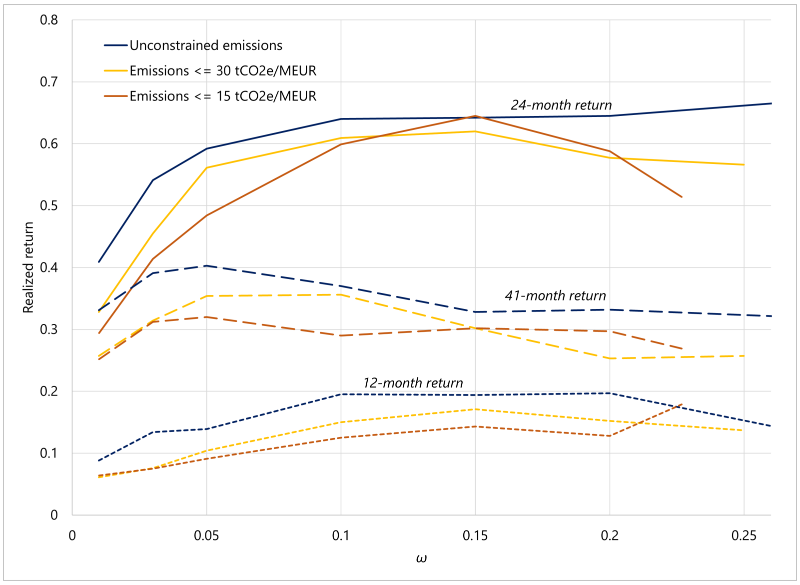


Figure 7: Realized risk-return frontiers of the optimized portfolios in 3 separate time horizons: 12 months (densely dashed), 24 months (solid line) and 41 months (sparsely dashed)

All optimized portfolios are exposed to the realized market prices of the out-of-sample period. Their performance is compared in 3 time horizons: 12 months, 24 months and 41 months (the whole out-of-sample period). This results in the realized risk-return frontiers depicted in Figure 7.

The 12-month and 24-month return frontiers exhibit a positive risk-return relationship with risk constraints up to $\omega = 0.15$, with the 24-month return showing a visibly stronger correlation. Increasing the CVaR threshold above 0.15 causes a decline in the 12-month and 24-month return of most portfolios. In the 41-month period, portfolios with $\omega = 0.05$ show on average the highest returns, with mostly a negative risk-return relationship on risk thresholds higher than that. We may conclude that in the 2020–2023 sample, long-term performance favors lower risk thresholds than the performance in a shorter period.

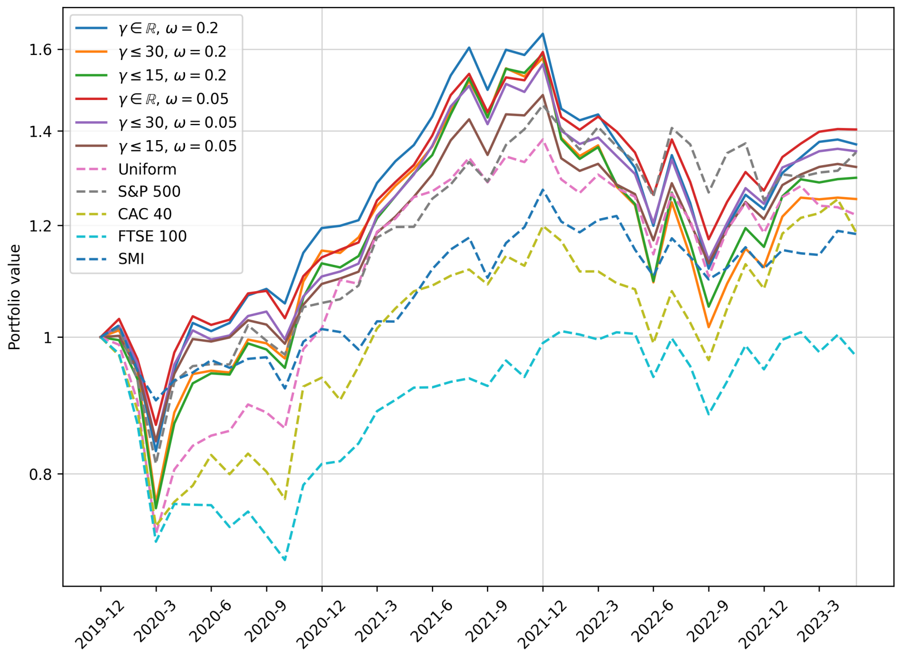


Figure 8: Out-of-sample performance of 6 optimized portfolios in comparison to an equally weighted portfolio and 4 market indexes

Figure 8 shows the change in the monthly value of the optimized portfolios with $\omega \in \{0.05, 0.2\}$, a portfolio with equal weights, and the market indexes. By visual inspection, the portfolios with lower risk constraints seem more robust to large market shifts, generally exhibiting smaller losses during market downswings and smaller returns during upswings. The observations indicate moderate predictability of portfolio volatility. All optimized portfolios in Figure 8 outperform the equally weighted portfolio and three out of the four market indexes during the out-of-sample period. In addition, three of the six portfolios outperform S&P 500, the best performing market index. The best performing portfolio during the out-of-sample period (unconstrained emissions, $\omega = 0.05$) exhibits a return of 40.3 %, while the S&P 500 return is 35.2 % and the equally weighted portfolio return is 22.1 %.

With lower risk thresholds, the relative order of different emission constraints is observed to largely match the expectation, so that portfolios with $\gamma = 30$ outperform those with $\gamma = 15$, but the opposite is observed with some of the higher risk thresholds. The portfolios with unconstrained emissions perform best in all considered time horizons, except for two occasions where a portfolio with constrained emissions exhibits the largest return. Any discrepancies only become apparent in longer periods and with higher allowed risk, which conforms to expectations and indicates they are random by nature.

During the 41-month period and the risk level $\omega = 0.05$, constraining the financed emissions to 30 tCO2e leads to 12.2 % reduction, and the constraint 15 tCO2e in a 20.6 % reduction in portfolio return. With the risk level $\omega = 0.2$, the corresponding reductions are 23.8 % and 10.5 %. While the exact performance reduction varies, its magnitude is similar across the risk levels. With most risk levels, the relative performance reduction is slightly larger than the corresponding reduction in expected return (see Figure 6).

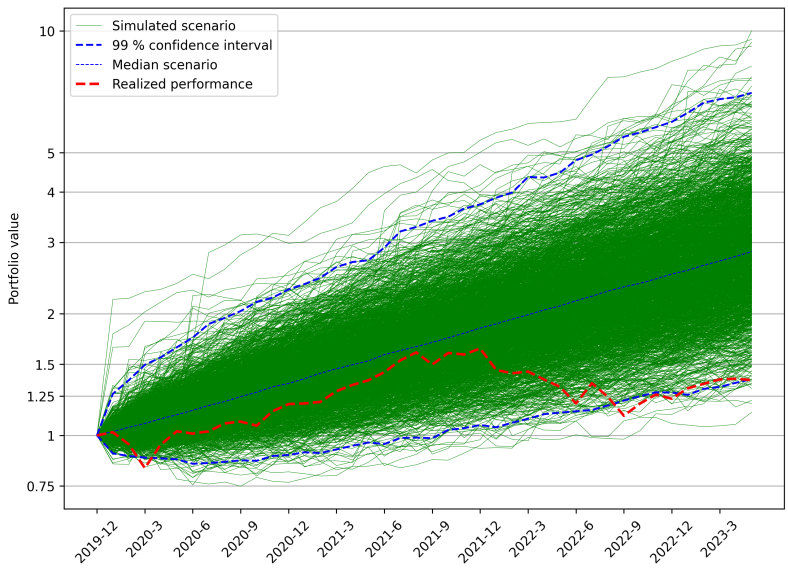


Figure 9: 2000 41-month scenarios of the optimal portfolio with $\gamma = \infty$ and $\omega = 0.1$, the 99 % confidence intervals and the median calculated from the scenarios, and the realized performance of the portfolio in out-of-sample testing

To compare the expected and realized portfolio performance in more detail, Figure 9 shows 2000 individual 41-month performance scenarios of the optimal portfolio with unconstrained emissions and risk constraint $\omega = 0.1$. This portfolio performed consistently well in all previously compared time horizons (see Figure 7). Note that the performance of individual scenarios depends on how the monthly scenarios are aggregated, and the scenarios presented here do not correspond to the 12-month scenarios used for optimization, but should provide a fairly accurate display of the distribution of scenarios and the expected return over a longer period. Figure 9 also shows the 99 % confidence interval and median calculated from the scenarios, along with the realized performance of the portfolio during the out-of-sample period.

In all simulated scenarios, a positive 41-month return is observed, with the maximum corresponding to a ten-fold increase in portfolio value. It is apparent that the realized performance is significantly weaker than the expected performance, which is largely a result of the unusual market events caused by the COVID-19 pandemic and the Russian invasion of Ukraine during the period. This acts as a reminder that the past performance of financial instruments should not be seen as a reliable indicator of future performance.

5 Summary

In this thesis, we studied how constraining the financed emissions of a stock portfolio affects the performance and industry composition of the portfolio. Unlike in earlier research on carbon sensitive mean-variance portfolios, we used a CVaR risk constraint and a Monte Carlo simulation framework where Student's t-distributions were used for modelling stock returns and a t-copula for modelling their dependencies. The optimization proved effective as the optimized portfolios clearly demonstrated an improved out-of-sample performance compared to an equally weighted portfolio of the same companies and the general market, represented by major market indexes.

Financed emissions of the portfolios with different risk levels were constrained to approximately match the median and the lower quartile emission intensity of the studied companies. This systematically reduced the portfolio return, which is in contrast with the results of Anquetin et al. (2022). Besides a different optimization and emission calculation methodology, Anquetin et al. (2022) consider a significantly larger amount of companies, which may be a significant factor allowing one to find better performing companies. Our results could also be seen as a loose indicator of a correlation between high emissions and high returns, which, as Chakrabarty and Nag (2023) point out, may be caused by various types of indirect causality.

Possible self-selection bias caused by studying only companies that reported to CDP is addressed by for example Siddique et al. (2021); Luo (2019). They find no significant self-selection bias in the CDP data, which means that the results are likely to be fairly representative of the broader market. Furthermore, the reliability of the GHG figures can be considered sufficient, as scope 1 and 2 emissions are subject to large-scale regulations, and because CDP extensively reviews the quality of reported data. CDP also provides estimates to replace missing or fallacious emission figures, which led to us using the estimated information whenever possible.

The data for studying the out-of-sample performance of the portfolios was impacted by the COVID-19 pandemic. This creates an additional model variate as one has to consider the differences in how well companies and industries coped with the pandemic. Consequently, the results presented do not necessarily translate to regular market performance, and a different time period could be considered for comparison. Additionally, out-of-sample performance was studied over a fairly short 41-month period, which does not suffice for drawing conclusions about long-term asset performance. Thus, the results do not directly contradict the hypothesis that environmental regulation has positive financial impact in long-term, as induced by earlier research.

We notice that our model tends to predict very high long-term returns of the optimal portfolios. This is likely the result of outlying scenarios drawn from return distributions with heavy right tails. While we expect scenarios drawn from symmetric distributions to contain equally many outliers in the upper and lower tails, it is possible to obtain series that contain only very few immensely large outliers in the right tail, which increases the expected return with no effect on the expected risk, as CVaR only takes into account the heaviness of the lower tail. We note that using hundreds of assets makes it difficult to assess the accuracy of every fitted distribution, but using left-skewed marginal distributions could provide an improvement to the model by mitigating the effects of outliers.

Our results can be used as an indicator of the impact of extending mandatory GHG reporting, but caution needs to be applied if drawing conclusions related to environmental performance, as GHG emissions alone should not be seen as a comprehensive indicator of a company's climate risk. However, along with the increasing legislation of environmental disclosure in the EU and worldwide, one may expect a growing amount of standardized environmental data to be available in the near future, which will make it possible to consider more accurate environmental measures, creating several new research opportunities on the relationship between financial and environmental performance.

References

Pietro Acerbi. Portfolio optimization and climate risk: decarbonization of the S&P 500 index. Master's thesis, University of Padova, 2022.

Andrew Ang and Joseph Chen. Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, 63(3):443–494, 2002.

Théophile Anquetin, Guillaume Coqueret, Bertrand Tavin, and Lou Welgryn. Scopes of carbon emissions and their impact on green portfolios. *Economic Modelling*, 115:105951, 2022. doi: https://doi.org/10.1016/j.econmod.2022.105951.

Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

Yuxin Cai. Study of carbon disclosure and its differences between different countries based on the case of energy company Shell. *BCP Business Management*, 29: 472–485, 2022. doi: https://doi.org/10.54691/bcpbm.v29i.2313.

CDP. Full GHG Emissions Dataset (Version 3.0), 2020.

CDP. Companies, 2023. Accessed: 2023-08-08. Available: https://www.cdp.net/en/companies.

Siddhartha P. Chakrabarty and Suryadeepto Nag. Risk measures and portfolio analysis in the paradigm of climate finance: a review. *SN Business & Economics*, 3(3):69, 2023.

Magali Delmas, Nicholas Nairn-Birch, and Jinghui Lim. Dynamics of environmental and financial performance: The case of greenhouse gas emissions. *Organization Environment*, 28(4):374–393, 2015.

Paul Embrechts, Filip Lindskog, and Alexander McNeil. Modelling dependence with copulas and applications to risk management. In Svetlozar T. Rachev, editor, *Handbook of Heavy Tailed Distributions in Finance*, pages 329–384. North Holland, 2003.

United States Environmental Protection Agency EPA. Mandatory Reporting of Greenhouse Gases (40 CFR part 98), 2009. Accessed: 2023-08-08. Available: https://www.ecfr.gov/current/title-40/chapter-I/subchapter-C/part-98.

European Commission. Corporate sustainability reporting, 2023. Accessed: 2023-08-08. Available: https://finance.ec.europa.eu/capital-markets-union-and-financial-markets/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en.

European Parliament. Directive 2014/95/EU of the European Parliament and of the Council, 2014. Accessed: 2023-08-08. Available: http://data.europa.eu/eli/dir/2014/95/oj.

Akhilesh Ganti. Adjusted Closing Price. *Investopedia*, 2020. Accessed: 2023-06-02. Available: https://www.investopedia.com/terms/a/adjusted_closing_price.asp.

Eva Horváthová. Does environmental performance affect financial performance? A meta-analysis. *Ecological Economics*, 70(1):52–59, 2010.

Iakovos Kakouris and Berç Rustem. Robust portfolio optimization with copulas. *European Journal of Operational Research*, 235(1):28–37, 2014.

Erik Kole, Kees Koedijk, and Marno Verbeek. Selecting copulas for risk management. *Journal of Banking Finance*, 31(8):2405–2423, 2007.

Pavlo A. Krokhmal, Stanislav Uryasev, and Jonas Palmquist. Portfolio optimization with conditional value-at-risk objective and constraints. *Journal of Risk*, 4(2): 43–68, 2001.

Thomas Linsmeier and Neil Pearson. Risk measurement: An introduction to value at risk. *EconWPA, Finance*, 1996.

Le Luo. The influence of institutional contexts on the relationship between voluntary carbon disclosure and carbon emission performance. *Accounting & Finance*, 59(2): 1235–1264, 2019.

Andrés Manzanares and Henrik Schwartzlose. Risk control, compliance monitoring and reporting. In Ulrich Bindseil, Fernando González, and Evangelos Tabakis, editors, *Risk Management for Central Banks and Other Public Investors*, pages 157–206. Cambridge University Press, 2009.

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

David I. Armstrong McKay, Arie Staal, Jesse F. Abrams, Ricarda Winkelmann, Boris Sakschewski, Sina Loriani, Ingo Fetzer, Sarah E. Cornell, Johan Rockström, and Timothy M. Lenton. Exceeding 1.5°C global warming could trigger multiple climate tipping points. *Science*, 377(6611), 2022. doi: https://doi.org/10.1126/science.abn7950.

Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative Risk Management*. Princeton University Press, 2005.

Roger B. Nelsen. Copulas and quasi-copulas: An introduction to their properties and applications. In Erich Peter Klement and Radko Mesiar, editors, *Logical, Algebraic, Analytic and Probabilistic Aspects of Triangular Norms*, pages 391–413. Elsevier Science B.V., Amsterdam, 2005.

R. R. Officer. The distribution of stock returns. *Journal of the American Statistical Association*, 67(340):807–812, 1972.

PCAF. *The Global GHG Accounting and Reporting Standard Part A: Financed Emissions. Second Edition.* 2022.

R. Tyrrell Rockafellar and Stanislav Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2(3):21–41, April 2000.

Gary Shanahan. Managing government regulation for carbon reporting and action. In Aarti Krishnan and Simon Maxwell, editors, *Counting carbon in global trade: Why imported emissions challenge the climate regime and what might be done about it*, pages 63–66. ODI, London, 2020.

Chirag Shekhar and Mark Trede. Portfolio optimization using multivariate t-copulas with conditionally skewed margins. *Review of Economics & Finance*, 9:29–41, 2017.

Md Abubakar Siddique, Md Akhtaruzzaman, Afzalur Rashid, and Helmi Hammami. Carbon disclosure, carbon performance and financial performance: International evidence. *International Review of Financial Analysis*, 75:101734, 2021. doi: https://doi.org/10.1016/j.irfa.2021.101734.

Abe Sklar. Functions de repartition an dimension set leursmarges. *Publications de L'In-stitut de Statistique de L'Universite de Paris*, 8:229–231, 1959.

Qingliang Tang and Le Luo. Does voluntary carbon disclosure reflect underlying carbon performance? *Journal of Contemporary Accounting and Economics*, 10(3): 191–205, 2014.

Ilmari Vauhkonen. Value-at-risk analysis of EU emission allowance options. Master's thesis, Aalto University. School of Science, Espoo, 2022.

Jonathan Woetzel, Dickon Pinner, Hamid Samandari, Hauke Engel, Mekala Krishnan, Brodie Boland, and Carter Powis. *Climate Risk and Response: Physical Hazards and Socioeconomic Impacts.* McKinsey Global Institute, 2020.

World Resources Institute (WRI) and World Business Council for Sustainable Development (WBCSD). *The Greenhouse Gas Protocol: A Corporate Accounting and Reporting Standard.* 2004.

A Correlation coefficients by industry

For each industry pair, the mean of the correlation coefficients in the copula correlation matrix $\hat{R}$ are presented in Table A1. The coefficients are calculated between the PIT-transformed log-return series (see Section 3.4). The industries are Apparel (App.), Biotech, health care & pharma (Bio.), Food, beverage & agriculture (Foo.), Fossil Fuels (Fos.), Hospitality (Hos.), Infrastructure (Inf.), Manufacturing (Man.), Materials (Mat.), Power generation (Pow.), Retail (Ret.), Services (Ser.) and Transportation services (Tra.).

Table A1: Mean of correlation coefficients between industry pairs in the copula correlation matrix.

	App.	Bio.	Foo.	Fos.	Hos.	Inf.	Man.	Mat.	Pow.	Ret.	Ser.	Tra.
App.	0.681	0.190	0.199	0.220	0.244	0.249	0.297	0.305	0.170	0.216	0.261	0.219
Bio.	0.190	0.187	0.163	0.145	0.184	0.187	0.210	0.203	0.149	0.169	0.187	0.168
Foo.	0.199	0.163	0.212	0.159	0.184	0.190	0.207	0.204	0.173	0.179	0.187	0.160
Fos.	0.220	0.145	0.159	0.360	0.161	0.183	0.226	0.231	0.171	0.168	0.194	0.190
Hos.	0.244	0.184	0.184	0.161	0.281	0.219	0.259	0.253	0.141	0.221	0.228	0.207
Inf.	0.249	0.187	0.190	0.183	0.219	0.268	0.256	0.252	0.188	0.212	0.235	0.206
Man.	0.297	0.210	0.207	0.226	0.259	0.256	0.328	0.312	0.168	0.243	0.270	0.257
Mat.	0.305	0.203	0.204	0.231	0.253	0.252	0.312	0.323	0.171	0.234	0.265	0.250
Pow.	0.170	0.149	0.173	0.171	0.141	0.188	0.168	0.171	0.302	0.158	0.171	0.143
Ret.	0.216	0.169	0.179	0.168	0.221	0.212	0.243	0.234	0.158	0.239	0.217	0.201
Ser.	0.261	0.187	0.187	0.194	0.228	0.235	0.270	0.265	0.171	0.217	0.251	0.222
Tra.	0.219	0.168	0.160	0.190	0.207	0.206	0.257	0.250	0.143	0.201	0.222	0.269

B Industry distribution

The portion of the portfolio value constituted by each industry in the optimized portfolios with different emission constraints γ is presented in Table B1 for four risk constraints ω. Note that with $\omega = 0.05,\ \gamma = 15$, the industries do not comprise all of the portfolio value (see Table 3).

Table B1: Contribution of each industry (ordered by their median emission intensity in descending order) to the portfolio value with different constraint parameters (total number of companies from each industry in parenthesis).

ω		0.3			0.2	
Industry γ	∞	30	15	∞	30	15
Transportation services (19)	0	0	0	0	0	0
Power generation (12)	0	0	0	0	0	0
Fossil Fuels (15)	0	0	0	0.022	0	0
Apparel (2)	0	0	0	0	0	0
Materials (57)	0.025	0	0	0.073	0.017	0
Food, beverage & agriculture (37)	0.075	0.075	0.075	0.075	0.075	0.092
Hospitality (19)	0	0	0.025	0	0	0.011
Retail (47)	0.175	0.125	0.1	0.152	0.108	0.097
Manufacturing (137)	0.2	0.2	0.2	0.2	0.2	0.2
Biotech, health care & pharma (61)	0.2	0.2	0.2	0.2	0.2	0.2
Infrastructure (55)	0.125	0.2	0.2	0.1	0.2	0.2
Services (263)	0.2	0.2	0.2	0.2	0.2	0.2
ω		0.1			0.05	
Industry γ	∞	30	15	∞	30	15
Transportation services (19)	0	0	0	0	0	0
Power generation (12)	0	0	0	0	0	0
Fossil Fuels (15)	0.004	0	0	0.022	0	0
Apparel (2)	0	0	0	0	0	0
Materials (57)	0.05	0.018	0.033	0.05	0.035	0.030
Food, beverage & agriculture (37)	0.125	0.099	0.075	0.127	0.075	0.046
Hospitality (19)	0.023	0.027	0.016	0.075	0.072	0.025
Retail (47)	0.1	0.117	0.076	0.112	0.092	0.062
Manufacturing (137)	0.179	0.2	0.2	0.138	0.2	0.194
Biotech, health care & pharma (61)	0.2	0.183	0.2	0.186	0.2	0.2
Infrastructure (55)	0.119	0.155	0.2	0.090	0.127	0.108
Services (263)	0.2	0.2	0.2	0.2	0.2	0.2