Junior Management Science 7(2) (2022) 429-454

Junior Management Science

journal homepage: www.jums.academy

Carbon Risk in European Equity Returns

Fabian Alexander Meyer

Universität Wien

Abstract

Investors perceive climate change and the volatility of asset prices caused by the ongoing low carbon transition of the economy, so-called carbon risk, to have an impact on their portfolio performance. However, the extent of carbon risk's impact on asset prices is still largely unknown. This paper provides a comprehensive quantification of carbon risk in European equity prices and examines whether it constitutes a systematic risk factor. I construct a carbon risk factor to determine the unique share of return attributable to differences in carbon intensity. During the sample period less (more) carbon intensive firms offer higher (lower) returns, which leads to a significant positive return of the carbon risk factor. Moreover, the carbon factor is significantly related to the sample covariance matrix of returns and offers a carbon risk premium in the cross-section of returns. In combination with the enhanced explanatory power relative to standard asset pricing models, this indicates that carbon risk constitutes a systematic risk factor. Consequently, investors can estimate carbon risk exposures based on widely available stock returns and include stocks without explicit carbon emission information in their risk management and investment process.

Keywords: Carbon risk; carbon risk factor; factor model; asset pricing.

1. Introduction

Compared to pre-industrial times, human activities are assumed to have led to a 1°C increase in global temperature (IPCC, 2018). Scientists agree that maintaining this development has potentially devastating consequences for natural and human systems. Hence, experts call for collective efforts to limit the global temperature rise to 1.5°C. Especially the financial system and its participants, who provide funding for proven yet costly cleaner technologies, are identified as crucial stakeholders to achieve this goal (IPCC, 2018). In this context, the European Union (EU) increases its efforts regarding sustainable finance to promote "the transition to a low-carbon, more resource-efficient and sustainable economy" (European Commission, 2021). In 2019 the EU announced the *European Green Deal* as the spearhead of its measures with the main goal of reaching climate neutrality by 2050. This transition process will be financed with an investment plan amounting to at least one trillion Euros of public and private investments. These political efforts are accompanied by heightened climate awareness in the private sector, where carbon divestment and green investment movements are entering the mainstream. For example, the Institutional Investors Group on Climate Change (IIGCC) has 270 members from 16 European countries with more than 35 trillion Euros in assets under management who pledge to support the low carbon transition of the economy (IIGCC, 2021). The enormous magnitude of the outlined financial streams has the potential to affect asset prices. As climate change's financial impact is not restricted to certain industries or firms but affects the entire economic system, the underlying carbon risk might explain systematic broad stock market movements.

Although the question, how carbon risk affects stock returns, generates widespread interest in the nascent climate finance literature, current evidence is mixed and primarily focused on the United States of America (US) (e.g., Bolton & Kacperczyk, 2020; In, Park, & Monk, 2019; Pedersen, Fitzgibbons, & Pomorski, 2021). In contrast, academic articles on carbon risk's effect on asset prices focusing solely on Europe are scarce. Moreover, ambiguity exists regarding the appropriate measures of carbon risk. This motivates authors to use complex, composite measures of carbon risk (e.g., Görgen et al., 2020), which raises questions regarding the comparability of existing results.

Based on common carbon intensity measures of the STOXX Europe Total Market Index constituents in the pe-

DOI: https://doi.org/10.5282/jums/v7i2pp429-454

riod from December 2006 to June 2020, the main goal of this thesis is to examine the following research questions:

1. Is carbon risk a systematic risk factor in European equity returns?
2. What is the magnitude and direction of the relationship between carbon risk and asset prices?

A further aspect within the scope of this thesis is to highlight the carbon risk factor's usefulness for risk management purposes.

Ambiguous empirical results on the relationship between carbon emission and stock returns and contradicting theoretical arguments emphasize that carbon risk could affect stock returns in different ways.[1] First, considering the currently observed divestment of carbon intensive assets, investors in brown firms might hold stakes larger than optimal in them and require higher compensation for non-optimal risk-sharing (Heinkel, Kraus, & Zechner, 2001). However, even in the absence of discriminatory tastes, holding carbon intensive stocks exposes shareholders to policy and technological risks, as regulatory changes target especially those firms accelerating global warming. Consequently, rational investors should require compensation for bearing this additional risk. Recent research, furthermore, suggests that green assets exert hedging properties against climate risk (Choi, Gao, & Jiang, 2020; Engle, Giglio, Kelly, Lee, & Stroebel, 2020), which should lead to lower returns when no climate event materializes. Thus, based on discriminatory tastes and risk consideration, one could expect high emission firms to outperform cleaner ones. I refer to this as the *dirty alpha hypothesis* in the following.

Alternatively, stocks with low carbon intensity potentially perform better if this attribute relates positively to future profits and this is not correctly priced yet. This argument is based on the empirical observation that low carbon intensity positively predicts accounting and market performance (e.g., Busch & Lewandowski, 2018; Pedersen et al., 2021). Moreover, a positive outperformance of low emission stocks could also be caused by higher demands due to recent shifts in tastes (e.g., Pástor, Stambaugh, & Taylor, 2021). Such effects could be driven by the rising awareness of market participants for climate change. However, in this case one would not expect a long lasting effect once investors' preferences are correctly reflected in asset prices. In general, mispricing and unexpected changes in preferences might cause an outperformance of low emission stocks which is referred to as the *clean alpha hypothesis*.

To examine my research questions, the thesis proceeds as follows. Chapters 2 and 3 briefly review the theoretical basis for this thesis. First, the principles of carbon accounting are outlined before different specifications of factor models are discussed. Chapter 4 reviews the related literature and chapter 5 describes the data. The construction of the carbon factor and its properties are included in chapter 6. Chapter 7 contains a rigorous test whether the carbon risk factor is indeed a priced risk factor and chapter 8 proposes fields of application. The results of the various analyses are aggregated and discussed in chapter 9 before a short conclusion is drawn in chapter 10.

2. Carbon Accounting

2.1. Measures of CO_2 and Global Warming Potential

The greenhouse effect is a phenomenon caused by so-called greenhouse gases (GHG) which occur naturally in the Earth's atmosphere and maintain the habitability of the planet. However, human activities have led to a concentration of GHGs in the atmosphere and give rise to the anthropogenic greenhouse effect caused by positive *radiative forcing*, i.e., more radiation is received than emitted by the Earth's climatic system (e.g., Brohé, 2017). Global awareness for the issue arose when the United Nations (UN) launched their first international framework with the aim to "prevent dangerous anthropogenic interference with the climate system" (UNFCCC, 1992, p. 9). As a consequence, efforts were undertaken to make GHG emissions traceable and quantifiable through so-called GHG inventories. The process of measuring CO_2 equivalents at the entity level is commonly referred to as carbon or GHG accounting.[2]

Although carbon dioxide is frequently equated with greenhouse gas, carbon accounting also covers other GHGs. More precisely, the *Kyoto Protocol* identifies five additional relevant GHGs: methane (CH_4), nitrous oxide (N_2O), sulphur hexafluoride (SF_6), hydrofluorocarbons, and perfluorocarbons (UNFCCC, 2008). Each GHG has a different global warming potential (GWP) depending on its radiative efficiency and residence time in the atmosphere, which characterizes its radiative forcing effect over a specific time span. I.e., GWP measures how much energy is absorbed by one tonne of a GHG over a given time period relative to one tonne of carbon dioxide (United States Environmental Protection Agency, 2021). Consequently, carbon dioxide's GWP is normalized to one. Organizations or states then record the emission of each GHG individually before converting it into carbon accounting's main unit of account, metric tonnes of CO_2 equivalent (tCO_2e), based on GWP values. Due to the great complexity involved in measuring the impact of individual greenhouse gases on radiative forcing and changing atmospheric conditions, GWP values are regularly updated and therefore constitute a source of uncertainty for carbon accounting. A commonly proposed best practice is the usage of the most recent GWP values for 100 years as provided by UN's Intergovernmental Panel on Climate Change (e.g., WBCSD & WRI, 2013). In the following, I use these CO_2

[1] In the following hypotheses on the direction of carbon intensity's impact on stock prices, I do not strictly differentiate between risk-based explanations and characteristics relating to investors' biases. For a differentiation compare, for instance, Pukthuanthong, Roll, and Subrahmanyam (2019).

[2] For a comprehensive literature review of the different meanings of carbon accounting in different disciplines, see Stechemesser and Guenther (2012).

equivalents reported following the GHG protocol to determine the greenhouse gas emission of a company.[3]

2.2. Boundaries of Carbon Accounting

The GHG protocol (WBCSD & WRI, 2004), which sets out rules for the emission quantification and reporting, was first published in 2001 and updated in 2004 based on practitioners' feedback. Since then the GHG protocol has developed into a widely applied standard for the preparation of emission inventories at the corporate level and is endorsed by major data providers (Busch, Johnson, Pioch, & Kopp, 2018). Like traditional financial accounting, carbon accounting also considers varying legal and organizational structures of companies. Hence, reporting firms can define the organizational boundaries according to which the emission inventory is compiled. The two approaches proposed by the GHG protocol are the *equity share approach* and the *control approach*. With the equity share approach the company reports emissions according to its economic interest in the company, which typically coincides with the equity stake in an operation (WBCSD & WRI, 2004). As an alternative, the control approach only considers CO_2 emissions from controlled operations, rather than those the company has interest in. This approach can be further distinguished based on whether financial or operational control over the entities of interest exists.

After identifying the organizational boundaries of considered GHG emissions, a further differentiation between direct and indirect emissions is necessary. This is referred to as the operational boundaries or scopes of carbon accounting. The GHG protocol proposes three scopes (WBCSD & WRI, 2004): The direct CO_2 emissions of a company are subsumed in the scope 1 measure. Depending on the organizational boundaries, this encompasses all GHG emissions from sources controlled or owned by the firm. Among others, it includes the GHG emissions set free by the combustion of fossil fuels, transportation of materials or the processing of chemicals. Scope 2 and 3 cover indirect carbon emissions, i.e., those emissions that are caused by company activities "but occur at sources owned or controlled by another company" (WBCSD & WRI, 2004, p. 25). More precisely, scope 2 refers to emissions from the purchase of externally generated electricity. As it represents one of the largest sources of a firm's GHG emission, it is individually reported and identified as one of the foremost sources of GHG emission reductions (e.g., Brohé, 2017; Busch & Lewandowski, 2018). The optional category of scope 3 emission covers the remaining indirect CO_2 emissions along the firm's value chain. This includes up- and downstream emissions, for instance, emissions from purchased input goods or the usage of sold products. Due to its optional character and the inherent ambiguity of determining emissions for the entire value chain, scope 3 reporting typically only covers a small fraction of the real value (Busch, 2011).

[3]For simplicity I use the terms GHG emission, carbon emission, and CO_2 emission interchangeably in the following.

3. Factor Models

3.1. General Structure of Factor Models

Nowadays the capital asset pricing model's (CAPM) inability to individually explain the cross-section of expected returns is commonly accepted and part of finance textbooks (e.g., Bodie, Kane, & Marcus, 2018; Campbell, Lo, & MacKinlay, 1997). Rather, returns of individual assets or portfolios of multiple assets are likely to depend on more than one determinant at the same time and even to a varying degree over time. Since return time series are observed to follow similar patterns and the joint analysis of multiple asset returns quickly results in complex or even inefficient multivariate statistical analyses (e.g., Tsay, 2014), a large stream of the finance literature is concerned with the identification of broad underlying factors for dimension reduction. In general, three main types of factor models emerged in the finance literature (Zivot & Wang, 2003). The first two types can be classified as theoretical approaches which either use macroeconomic variables as factors or construct factor portfolios relating to firm characteristics. These types are therefore commonly referred to as *macroeconomic factor model* and *fundamental factor model*, respectively. Thirdly, *statistical factor models* exist which extract the non-observable factors from the realized stock returns (Zivot & Wang, 2003).

Generally, the structure introduced below underlies each of the three types of factor models, but for clarity the particularities of statistical factor models are highlighted in the next section. It is assumed that an asset's return generating process is a linear function of a limited number of common factors. For notational simplicity, only the cross-sectional regression formulation of the model is presented which has the form

$$\underset{(N\times 1)}{\boldsymbol{r_t}} = \underset{(N\times 1)}{\boldsymbol{\alpha}} + \underset{(N\times K)}{\boldsymbol{\beta}}\underset{(K\times 1)}{\boldsymbol{f_t}} + \underset{(N\times 1)}{\boldsymbol{\varepsilon_t}}, \forall t, \tag{1}$$

where $\boldsymbol{r_t} = \left(r_{1,t}, \ldots, r_{N,t}\right)'$ denotes the vector of either real or excess returns of assets i $(i = 1, \ldots, N)$ at time t $(t = 1, \ldots, T)$.[4] The vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)'$ represents the intercept, $\boldsymbol{\beta}$ is the matrix of factor loadings, and $\boldsymbol{f_t}$ refers to the common factor realizations at time t.[5] K denotes the number of common factors included in the model. The first two moments of $\boldsymbol{f_t}$ are described by $E[\boldsymbol{f_t}] = \overline{\boldsymbol{f}}$ and the covariance matrix $\boldsymbol{\Omega_f}$. Finally, $\boldsymbol{\varepsilon_t}$ denotes the vector of asset specific factors, or error terms. The covariance matrix of the error terms, $\boldsymbol{D}$, is assumed to be diagonal. Moreover, common factors and error terms are assumed to be uncorrelated among each other. In contrast, especially in the macroeconomic and fundamental factor model, common factors can

[4]If factors are observable, equation 1 can be interpreted as cross-sectional regression. This interpretation, however, is incorrect if factors are unobservable in statistical factor models (Tsay, 2010). For alternative representations of the general factor model, see Zivot and Wang (2003) or Tsay (2010).

[5]Pukthuanthong et al. (2019) further classify common factors into priced and unpriced common factors.

be correlated. The resulting covariance matrix of asset returns thus can be denoted by

$$\boldsymbol{\Omega_r} = \boldsymbol{\beta}\,\boldsymbol{\Omega_f}\,\boldsymbol{\beta}' + \boldsymbol{D}. \tag{2}$$

The model formulation presented in equation 1 is particularly common for macroeconomic and fundamental factor models. For instance, it is the basis for the well-known macroeconomic single factor model by Sharpe, the market model, and its notable later extensions to a fundamental factor model by Fama and French (1993), Carhart (1997), and Fama and French (2015). Especially the methodology by Fama and French (1993), where factor returns are derived from hedge portfolios, is mentioned as the archetype of a fundamental factor model (e.g., Tsay, 2010). Conceptually, the mentioned macroeconomic and fundamental factor models have two main differences compared to the statistical factor model. First, since returns and therefore factors are observable, factor loadings and idiosyncratic volatility can be estimated directly through time series regression (Zivot & Wang, 2003).[6] A second conceptual difference mentioned by Tsay (2010) lies in the indeterminacy of the number of factors in fundamental factor models because characteristics could be (dis-)aggregated to adjust the number of factors. Statistical models offer a remedy by dissociating the factors mathematically, but issues regarding the number of factors to be chosen and the lack of economic interpretability arise (Campbell et al., 1997). Linking fundamental and statistical factors through canonical correlation analysis in section 7.1 therefore closes the gap between economical intuition and computational rigor.

3.2. Statistical Factor Model and Asymptotic Principal Component Analysis

In statistical factor models both factor realizations and factor loadings of assets are latent and must be estimated based on the return vector of N assets, $\boldsymbol{r_t}$. In comparison to the fundamental factor model, statistical factor models therefore have the advantage of only requiring return data and the absence of multicollinearity (Alexander, 2001). The traditional statistical factor model is based on an orthogonal factor structure and assumes that $\boldsymbol{r_t}$ is generated by a stationary process with mean $\overline{\boldsymbol{r}}$ and a covariance matrix of $\boldsymbol{\Omega_r}$ (Tsay, 2010). As above, it is assumed that few linear combinations, the unobserved K common factors $\boldsymbol{f_t} = \left(f_{1,t}, \ldots, f_{K,t}\right)$, can be used to explain a large fraction of $\boldsymbol{\Omega_r}$. The remaining unexplained share of variance is linearly explained by the vector of error terms $\boldsymbol{\varepsilon_t}$, which is assumed to be independent of $\boldsymbol{f_t}$. However, for identifiability it is now assumed that $E[\boldsymbol{f_t}] = 0$ and the factor covariance matrix equals the $K \times K$ identity matrix $\boldsymbol{I_K}$. Since the common factors explain all cross-covariances of asset returns, the return covariance matrix simplifies to

$$\boldsymbol{\Omega_r} = \boldsymbol{\beta}\,\boldsymbol{I_K}\,\boldsymbol{\beta}' + \boldsymbol{D} = \boldsymbol{\beta}\boldsymbol{\beta}' + \boldsymbol{D}, \tag{4}$$

compared to equation 2. Moreover, now the matrix of factor loadings, $\boldsymbol{\beta}$, must be full column rank, otherwise one or more factors are redundant (Tsay, 2014). As shown in equation 5 below, the only formula-related change in the model relative to the general factor model in equation 1 occurs in the constant; i.e., $\boldsymbol{\alpha}$ is replaced by $\overline{\boldsymbol{r}}$. This is commonly done without loss of generality since the factors are computed from the covariance matrix (Campbell et al., 1997). Equation 5, however, uniquely identifies $\boldsymbol{f_t}$ and $\boldsymbol{\beta}$ only up to an orthogonal transformation which can hamper the interpretation of the factors (Zivot & Wang, 2003).

$$\underset{(N\times 1)}{\boldsymbol{r_t}} - \underset{(N\times 1)}{\overline{\boldsymbol{r}}} = \underset{(N\times K)}{\boldsymbol{\beta}}\underset{(K\times 1)}{\boldsymbol{f_t}} + \underset{(N\times 1)}{\boldsymbol{\varepsilon_t}}, \forall t \tag{5}$$

If the number of time periods exceeds the number of assets (i.e., $T > N$), a common method for factor estimation is principal component analysis (PCA). However, in many financial applications N exceeds T. In this case, the return covariance matrix of the sample cannot be inverted anymore, i.e., it becomes singular and restricts the usage of PCA (Zivot & Wang, 2003). Asymptotic principal component analysis (APCA), a method proposed by Connor and Korajczyk (1986, 1988), provides a remedy and is shown to be asymptotically ($N \rightarrow \infty$) equivalent to the traditional factor analysis under certain assumptions.[7] It applies traditional PCA to the sample cross-sectional covariance matrix of demeaned stock returns $\boldsymbol{R}$,

$$\underset{(T\times T)}{\boldsymbol{\Omega_R}} = \frac{1}{N}\underset{(T\times N)}{\boldsymbol{R}'}\underset{(N\times T)}{\boldsymbol{R}}. \tag{6}$$

The eigenvectors of the K largest eigenvalues of $\boldsymbol{\Omega_R}$ are then used to consistently estimate the common factors (Tsay, 2014).

4. Related Literature

A broad stream of literature studies the impact of environmental, social, and governance (ESG) ratings in financial markets. Due to the crucial importance of global warming, particularly the environmental perspective of ESG with focus on carbon emission has generated widespread interest. My thesis is primarily linked to this dynamically growing literature on carbon risk's effect on asset prices. Although the literature in the field is primarily empirical, also notable theoretical contributions have been made.

[6] This feature is extensively used in the following chapters to determine the relevance of the carbon factor. Refining equation 1 for individual assets, i, and explicitly incorporating the period's risk-free rate, $r_{f,t}$, the resulting model can be represented by

$$r_{i,t} - r_{f,t} = \alpha_i + \boldsymbol{\beta_i'}\boldsymbol{f_t} + \varepsilon_{i,t}. \tag{3}$$

[7] See Tsay (2014) for a discussion of the assumptions of APCA and the presented formula.

An early theoretical contribution regarding the effects of environmentally friendly investing on asset prices in equilibrium is made by Heinkel et al. (2001). The authors argue that if the market is segmented into green and non-green investors, the latter require compensation for limited risk-sharing when holding polluting firms shunned by green investors. Hence, Heinkel et al.'s model predicts higher expected returns for non-green firms. An alternative equilibrium model incorporating ESG considerations is proposed by Pástor et al. (2021). According to their model, green stocks command negative CAPM alphas, whereas brown stocks have positive abnormal returns. In the model's base setting, the underperformance of green stocks is explained by the investors' ESG preferences. The investors' equilibrium asset choice is described by three-fund separation where each portfolio consists of the risk-free asset, the market portfolio, and an ESG portfolio. As a first extension of their model, Pástor et al. introduce a priced ESG factor which explains the above mentioned CAPM alphas. If ESG concerns increase unexpectedly (i.e., positive realization of the ESG factor), green firms perform better due to higher demand for green products and greater utility derived by investors holding them. In this case, green firms outperform brown ones in spite of lower expected returns. In a second extension of their model for climate risk, Pástor et al. introduce an additional risk-based rationale for brown stocks' higher expected returns. Green stocks can be seen as hedges against climate shocks, which investors dislike, and therefore should offer lower returns than their browner counterparts.[8] This leads to the conclusion that in the presence of climate risk sensitivities investors should additionally hold a climate-hedging portfolio (four-fund separation). Finally, in a further recent theoretical contribution, Pedersen et al. (2021) extend the traditional CAPM for ESG preferences and present the ESG-Sharpe ratio frontier which conceptualizes the trade-off between risk, return, and ESG preferences. Similar to Pástor et al. (2021), they find a four-fund separation solution to be optimal. However, funds are allocated between the risk-free asset, the market portfolio, the minimum variance portfolio, and the optimal ESG portfolio in their model. Moreover, the authors show that higher ESG assets can either generate lower or higher expected returns relative to their conventional counterparts. The latter is the case if ESG characteristics predict higher future profits and a large fraction of investors is unaware of this relationship. That means prices of these profitable assets are not bid up yet. In an empirical examination of their predictions based on the S&P 500 index, Pedersen et al. measure the environmental dimension as the scope 1 and 2 carbon intensity provided by Trucost for the period from January 2009 to March 2019. They examine whether ESG measures help to generate abnormal returns by testing their relation to future fundamentals and investors' demands. Lower carbon intensity is shown to positively predict accounting returns (measured by return on net operating assets) and increase institutional ownership, which results in higher valuations for greener assets. Moreover, carbon intensity-sorted quintile hedge portfolios generate weakly significant excess returns and positive alphas relative to common factor models.

Besides the mentioned theoretical contributions, a broad empirical literature has emerged recently. Using a similar approach as Pedersen et al. (2021), In et al. (2019) study the performance of carbon intensity hedge portfolios. Their sample ranges from 2005 to 2015 and covers 736 US stocks. As main measure of carbon risk they use scope 1 to 3 carbon intensity provided by Trucost. Overall, they find economically and statistically significant returns when going long in low carbon intensity stocks and shorting high intensity stocks. Bolton and Kacperczyk (2020) examine whether carbon emissions have an impact on the cross-section of US stock returns. Their sample spans the period from 2005 to 2017 and includes more than 3,400 firms for which Trucost provides scope 1 to 3 CO_2 emission data. Controlling for known firm-specific return predictors (e.g., market value, book-to-market ratio, and profitability), Bolton and Kacperczyk do not find a significant cross-sectional relationship between emission intensity and stock return. However, both absolute amounts of emission and emission growth rates are significantly related to returns for all three scopes. This points to the existence of a carbon return premium by which investors are compensated for bearing carbon risk associated with high emission firms. Moreover, this premium cannot be explained by known risk factors such as the Fama-French factors. Testing whether the carbon premium is driven by institutional investors' withdrawal of funds from high emission firms and resulting limited risk-sharing (e.g., as in Heinkel et al. (2001)), Bolton and Kacperczyk only find weak evidence for systematic effects. While investors divest from companies with high scope 1 carbon intensity to a statistically significant degree, this effect is not observable for scope 2 and 3 carbon intensity. Consequently, they do not seem to avoid high emission companies per se.

Another stream of empirical literature does not only ask whether but also when carbon risks materialize. Recent research suggests that carbon risk affects asset prices when the impact of climate change becomes tangible reality. For a sample of 74 stock exchanges worldwide, Choi et al. (2020) find that in times of abnormally warm local temperatures the attention to global warming rises and low carbon stocks outperform. Moreover, they show that especially retail investors, who are prone to limited attention problems, react to climatic events by replacing stock ownership in high emission firms with low emission firms. Instead of looking at climate disasters directly, Engle et al. (2020) propose a dynamic equity hedging approach whose returns hedge against climate news innovations. Using environmental scores of US firms from MSCI and Sustainalytics in the time from 2009 to 2016, they find significant covariations between the hedge portfolios and the climate change news indices. In a related article, Ilhan, Sautner, and Vilkov (2020) add further evidence for the existence of priced carbon risk. Studying the impact of climate policy uncertainty in option prices of S&P 500 firms

[8] For empirical evidence regarding the hedging properties of green stocks, see Engle et al. (2020) and Choi et al. (2020).

in the time from 2009 to 2016, they show that protection against left-tail risk is more costly for carbon intensive firms. This effect tends to be even more systematic on the sector-level, when firm-specific risk is diversified away. Moreover, using proxies of climate change attention as in Engle et al. (2020) and Choi et al. (2020), Ilhan et al. also find a stronger positive effect of carbon intensities on the implied volatility slope and hence on the option price when climate change awareness is higher.

Most closely related to my thesis is the paper by Görgen et al. (2020). They study carbon risk in asset prices based on a global sample of 1,657 firms in the period from 2010 to 2017. Joining environmental data from four databases, they construct a carbon risk factor using the Fama and French (1993) methodology. Besides carbon intensity, which they classify as the *value chain* dimension of carbon risk, Görgen et al. also consider additional *public perception* and *adaptability* measures as firm characteristics in the calculation of their factor. Among others, the latter two dimensions include aggregate environmental scores and related subscores. Employing a panel regression approach, they show that brown firms provide higher returns on average. Although their carbon risk factor explains variation in stock returns, they do not find evidence of an associated risk premium. Lastly, Hübel and Scholz (2020) use the same factor construction methodology and create an environmental, social, and governance factor, respectively. Based on Datastream's aggregate ESG scores of the annual STOXX Europe Total Market Index constituents from 2003 to 2016, they observe an outperformance of firms with low environmental rating. Furthermore, they show how firms without ESG ratings can be incorporated in investors' risk management by computing exposures of individual stocks to their three factors.

5. Data

The thesis focuses on European equities which are proxied by the STOXX Europe Total Market Index in the time from December 2006 to June 2020. I download annual constituents as per year-end from Thomson Reuters Datastream to avoid potential survivorship bias over the sample period. The index covers at least 95% of the free-float market capitalization in each of the 17 European countries included in the index (Deutsche Börse Group, 2021).[9] Hence, this sample reflects the investable universe of most investors and ensures the practical relevance of my results. All financial data retrieved from Datastream are US dollar (USD) denominated to accord with other data sources. As the direct usage of financial data from Thomson Reuters can lead to incorrect inference compared to other data providers,[10] I follow the guidelines on static and dynamic screens by Ince and Porter (2006) and Schmidt, von Arx, Schrimpf, Wagner, and Ziegler (2019) for data cleaning. Static screens refer to time-invariable information, whereas dynamic screens address information which changes over time. Applied static screens comprise, for example, the exclusion of non-equity assets and foreign or non-primary listed stocks. The used dynamic screens aim at resolving potential liquidity biases arising from incorrectly padded zero-returns and Datastream's decimalization, which leads to distortions for pennystock returns. For a comprehensive overview of the applied data screens see table 19 in appendix A. Similar to related literature (e.g., Görgen et al., 2020; In et al., 2019), I exclude firms labeled as *financials* according to the Industry Classification Benchmark (ICB). Financials tend to have low levels and a small dispersion of carbon intensities, despite potentially high carbon risk exposures due to their business model. For instance, reinsurance companies are among the firms most vulnerable to climate change in spite of low carbon intensities. Furthermore, firms – especially financials – have considerable freedom regarding reported carbon emissions depending on their defined organizational boundaries. Restricting the sample to non-financial firms results in 1,710 unique stocks from ten industries. The financial data set is complemented by USD denominated European monthly factor return data provided by French (2020). As risk-free rate the one month US treasury bill rate included in French's data set is used.

In the following, I use carbon intensity as measure of carbon risk. Carbon intensity is defined as metric tonnes of scope 1 and 2 CO_2 emission equivalents per one million USD of net sales:

$$\text{Carbon Intensity} = \frac{\text{Scope 1 \& Scope 2 Emission}}{\text{Net Sales}} \left[\frac{tCO_2e}{\text{million USD}}\right]. \tag{7}$$

Hence, it reflects how much tonnes of CO_2 the company requires to generate one million USD of net sales. Due to its relative nature, carbon intensity is more robust towards firm size or process adjustments, and accounts for changes in emission from variations in production over the business cycle (Busch & Lewandowski, 2018). The used emission figures from Datastream comprise company reported and estimated carbon emissions. A study by Busch et al. (2018) finds that not only company reported emissions but also scope 1 and 2 estimates are mostly homogeneous across major data providers. Since scope 3 emissions are less widely available and inconsistencies in emission estimates exist (Busch et al., 2018), I do not include this emission scope in my analyses.

Table 1 shows the industry composition of my sample including stocks with and without emission information and the time series average summary statistics of carbon intensity. The largest industries in my sample are industrials (25.26%) and consumer discretionary (20.47%), whereas none of the remaining eight industries has a share above 10%. The mean of carbon intensity exceeds the median in each of the ten industries and the third quartile in seven industries. This indicates the presence of highly carbon intensive firms in all

[9] Countries included are: Austria, Belgium, Poland, Denmark, Finland, France, Germany, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the UK.

[10] For instance, Ince and Porter (2006) do not find the widely observed momentum effect in US returns using raw data from Datastream.

Table 1: Summary statistics of carbon intensity and industry composition.

	Count		Carbon Intensity			
Industry	Absolute	%	1st Q	Median	Mean	3rd Q
Basic Materials	128	7.49	165.50	415.56	648.91	798.16
Consumer Discretionary	350	20.47	8.17	19.08	106.22	47.22
Consumer Staples	127	7.43	28.87	48.85	73.94	80.41
Energy	101	5.91	63.43	224.01	475.04	427.66
Health Care	144	8.42	13.70	24.04	44.92	40.86
Industrials	432	25.26	17.51	32.86	304.00	82.70
Real Estate	133	7.78	21.36	44.04	66.43	69.33
Technology	144	8.42	7.47	12.43	28.44	22.51
Telecommunication	75	4.39	10.77	23.29	53.07	34.35
Utilities	76	4.44	331.96	866.59	1,753.90	1,509.55
Total	1,710	100				

industries. In particular, the striking difference between the median and the mean for industrials is primarily driven by the inclusion of large cement producers in this industry sector, e.g., LafargeHolcim. Among the ten industries, utilities, basic materials, and energy can be identified as especially carbon intensive industries. In contrast, less carbon intensive industries are consumer discretionary, telecommunication, and technology.

Table 2 shows the resulting annual sample used in the subsequent sections. As the emission data are compiled from company reports, I rebalance my sample at the end of June of each year to account for potential look-ahead bias arising from the lagged availability of company filings. The row *Rated* refers to stocks with emission information available at the end of the previous year, $t-1$, and price information in July of year t. In contrast, the *Unrated* row includes stocks without emission information. The main focus of my analyses is the last decade from July 2010 to June 2020, whose start is indicated by the vertical line in the table. The preceding period from July 2007 to June 2010 is used for auxiliary purposes in rolling regressions to generate estimates in later tests of the factors. While the sample's total number of stocks is relatively stable and ranges between 1,212 and 1,323, the number of rated stocks grows gradually from 364 in 2007 to 1,020 in 2019. Although the selected time span is relatively short compared to common asset pricing studies, the restriction to this time period appears meaningful for the following reasons. First, carbon accounting tends to be a relatively new phenomenon as discussed in section 2. While Thomson Reuters offers the longest time series of emission data among major data providers going back to 2002 (Busch et al., 2018), previous to the fiscal year 2006 the data availability drops sharply. Moreover, most other major data providers start their coverage during or only shortly before my sample period (Busch et al., 2018). Thus, including the earlier, additional years could introduce systematic differences and lead to false inference. Using the years after 2009 as main sample period is also consistent with other studies and ensures the comparability of my results (e.g., Engle et al., 2020; Görgen et al., 2020; Pedersen et al., 2021). Secondly, research by Engle et al. (2020) suggests that public awareness for climate change has gradually increased since the change of the millennium and intensified during the sample period. Key events mentioned in their study include the UN Climate Change Conferences in Copenhagen (12/2009), Doha (12/2012), and Paris (12/2015). The latter culminated in the notorious *Paris Agreement*. Especially in Europe, recently also grassroots movements such as *Fridays for Future* manifest this heightened awareness.

Prior to the subsequent main analyses I check for a selection bias caused by the exclusion of financials. For this purpose, I test the screened sample relative to the Fama and French (1993) 3-factor model extended for the momentum factor proposed by Carhart (1997).[11] Similar to the factor model outlined in section 3.1, the regression equation for an individual portfolio i is denoted by

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i^{MKT} f_t^{MKT} + \beta_i^{HML} f_t^{HML} + \beta_i^{SMB} f_t^{SMB} + \beta_i^{MOM} f_t^{MOM} + \varepsilon_{i,t}. \tag{8}$$

In the equation, the superscript MKT, HML, SMB, and MOM refer to the widely applied market, value, size, and momentum factor, respectively.[12] Table 3 shows the fit of my sample relative to this 4-factor model. Both the full sample and the sample of stocks with emission information available (*Rated*) are well explained by the 4-factor model with an adjusted R^2 close to one. Despite the exclusion of financials, the market betas are close to one and the alphas are close to zero and statistically not significant. I conclude that neither the exclusion of financials nor the restriction to firms with emission information introduce a systematic bias.

[11] In the literature this model constellation is frequently referred to as Carhart 4-factor model. For brevity, I frequently refer to this model as the 4-factor model in the following.

[12] For a comprehensive description of the factors and their construction see, e.g., French (2020).

Table 2: Number of stocks in the sample in July of year t.

t	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Rated	364	649	637	664	692	687	704	715	732	749	743	832	1,020
Unrated	936	632	575	569	561	547	535	571	583	567	580	489	264
Total	1,300	1,281	1,212	1,233	1,253	1,234	1,239	1,286	1,315	1,316	1,323	1,321	1,284

Note. The vertical line represents the start of the main sample period. The period from July 2007 to June 2010 is used for auxiliary purposes.

Table 3: 4-factor model for the value-weighted portfolio of all stocks and rated stocks from 07/2010 to 06/2020.

	MKT	HML	SMB	MOM	Alpha	Adj. R^2
All	1.006***	-0.192***	-0.094***	0.033	-0.011	98.88
	(93.02)	(-9.23)	(-3.55)	(1.55)	(-0.23)	
Rated	1.005***	-0.204***	-0.138***	0.025	-0.022	98.82
	(90.11)	(-8.78)	(-4.85)	(1.17)	(-0.45)	

Note. *p<0.1, **p<0.05, ***p<0.01; t-statistics in parentheses; heteroskedasticity robust standard errors are used; alpha is reported as monthly percentage.

6. Understanding Carbon Risk in Asset Prices

6.1. Return of Carbon Intensity-Sorted Portfolios

In order to establish an understanding of the relation between carbon intensity and returns, I construct carbon intensity-sorted quintile portfolios. The value-weighted portfolios are annually updated at the end of June based on the reported carbon intensity of the previous year end. The first quintile (*Clean*) contains the firms with the lowest carbon intensities, whereas the fifth quintile contains the firms with the highest emissions intensities (*Dirty*). Figure 1 plots the cumulative returns of the quintile portfolios adjusted for the value-weighted return of the market portfolio in the time from July 2007 to June 2020. The vertical red dotted line represents the start of the main sample period in mid 2010. While more carbon intensive portfolios performed better in the earlier years, the low emission portfolios outperformed clearly in the more recent years. Notable is also the inverse development of the top and bottom quintile starting in mid 2008. The cleanest 20% of stocks show a relatively linear positive development, whereas the opposite is true for the dirtiest quintile.

Over the sample period a monotonous relationship between excess returns and carbon-intensity quintiles emerges. Table 4 reports the risk and return profile of the quintiles. Panel A shows the results from the main sample period from 07/2010 to 06/2020 and panel B additionally includes the preceding three years for comparison. The latter sample includes the great financial crisis and is characterized by lower average excess returns and greater volatility. In both panels the extreme portfolios have a higher standard deviation. While in panel A this does not affect the monotonous relationship in Sharpe ratios, in panel B the second quintile offers the most favorable risk-return trade-off. The calculated Sharpe measures are also consistent with MacKinlay (1995) who finds that over long periods the annualized Sharpe ratios of diversified portfolios lie well below one. Overall, only the monthly excess return of the first quintile in panel A is statistically different from zero at the 5% significance level. At least weak significance is found for the two adjacent portfolios.

6.2. Construction of the SCOPE12 Factor

After observing a monotonous negative relationship between returns and carbon intensities, I proceed by constructing a carbon risk factor. Since the carbon intensity is calculated based on the scope 1 and 2 emissions, I refer to the constructed carbon factor as *SCOPE12*. I apply the Fama and French (1993) methodology for factor construction, where a characteristic-based return predictor is converted into a factor by calculating the return differential of a zero-investment portfolio. Comparable to Fama and French's HML factor, SCOPE12 controls for size and goes long (short) in stocks with low (high) carbon intensity. I annually double sort stocks into six portfolios based on their market capitalization in June of year t and their carbon intensity at the end of the previous year $t-1$. I apply the median size in my sample and the terciles of carbon intensity as breakpoints. The annual update at the end of June aims at reducing potential look-ahead bias compared to related studies (e.g., Görgen et al., 2020; Hübel & Scholz, 2020). I then calculate the value-weighted monthly simple returns, r_t, for the four relevant portfolios for factor construction: small size and high carbon intensity (SH), small size and low carbon intensity (SL), big size and high carbon intensity (BH), and big and low carbon intensity (BL). The SCOPE12 factor goes long in the equally-weighted portfolio of firms with the lowest carbon intensity and shorts the equally-weighted portfolio of firms with the highest carbon intensity. Hence, the return on the carbon risk factor in period t, $f_t^{SCOPE12}$, is defined as the return of the following portfolio,

$$f_t^{SCOPE12} = \frac{1}{2}(r_t^{SL} + r_t^{BL}) - \frac{1}{2}(r_t^{SH} + r_t^{BH}). \quad (9)$$

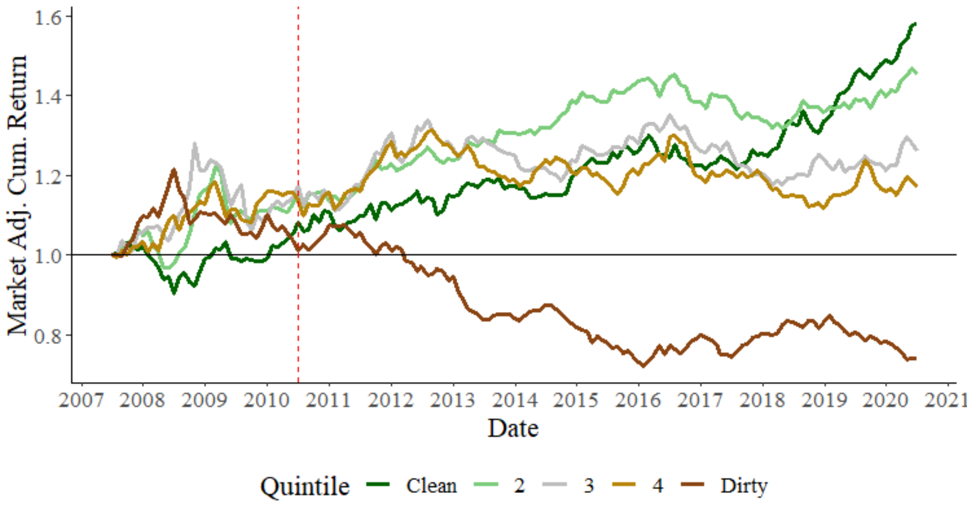

Figure 1: Market adjusted cumulative returns of the carbon intensity-sorted, value-weighted quintile portfolios in the period from 07/2007 to 06/2020.

Table 4: Risk and return of carbon intensity-sorted quintile portfolios.

Panel A. Period from 07/2010 to 06/2020.					
Quintile	Clean	2	3	4	Dirty
Excess return	0.94**	0.81*	0.69*	0.65	0.37
Standard deviation	4.98	4.62	4.27	4.81	5.50
t-statistic	2.07	1.91	1.76	1.48	0.73
Sharpe ratio	0.65	0.61	0.56	0.47	0.23

Panel B. Period from 07/2007 to 06/2020.					
Quintile	Clean	2	3	4	Dirty
Excess return	0.52	0.47	0.38	0.33	0.04
Standard deviation	5.84	5.06	4.80	5.55	6.31
t-statistic	1.12	1.15	1.00	0.74	0.07
Sharpe ratio	0.31	0.32	0.28	0.21	0.02

Note. *p<0.1, **p<0.05, ***p<0.01; excess return is reported as monthly percentage; the Sharpe ratio is annualized.

To ensure the robustness of my results, I also consider alternative specifications of the carbon factor. First, I address potential objections that the above factor simply reflects the differential between well-performing, low emission technology companies and low-performing, high emission industries. For this purpose, I replicate the SCOPE12 factor with industry adjusted carbon intensity terciles. That is, I keep the size sort unchanged and repeat the second sort into carbon intensity terciles for each industry individually. Then I allocate the stocks to the four relevant portfolios as above. The resulting factor, *IND_SCOPE12*, is significantly correlated ($\rho = 0.63$) with the SCOPE12 factor. In addition to this, I also consider the factor construction without an additional size control. The clean minus dirty (CMD) factor goes long in the value-weighted portfolio of the least carbon intensive tercile and short in the most carbon intensive tercile. The correlation between the CMD and SCOPE12 factor is 0.92, suggesting that SCOPE12's additional size sort has only limited impact on the performance of the factor. For comparison, figure 2 below depicts the cumulative returns of the three factors in the period from July 2007 to June 2020. The vertical red dotted line again represents the start of the main sample. While the SCOPE12 factor is more volatile and experienced a pronounced drawdown during the great financial crisis, it performed well since mid 2008 generating average monthly returns of 0.323% ($t = 2.137$, $\sigma = 1.885\%$). IND_SCOPE12, in contrast, generated a mean return of 0.201% ($t = 2.205$) with a lower standard deviation of 1.140% per month. The CMD factor, which neither adjusts for size nor for industry, offered an average monthly return of 0.395% ($t = 2.218$, $\sigma = 2.224\%$). As indicated by the t-statistics in parentheses, all three average monthly factor returns are statistically

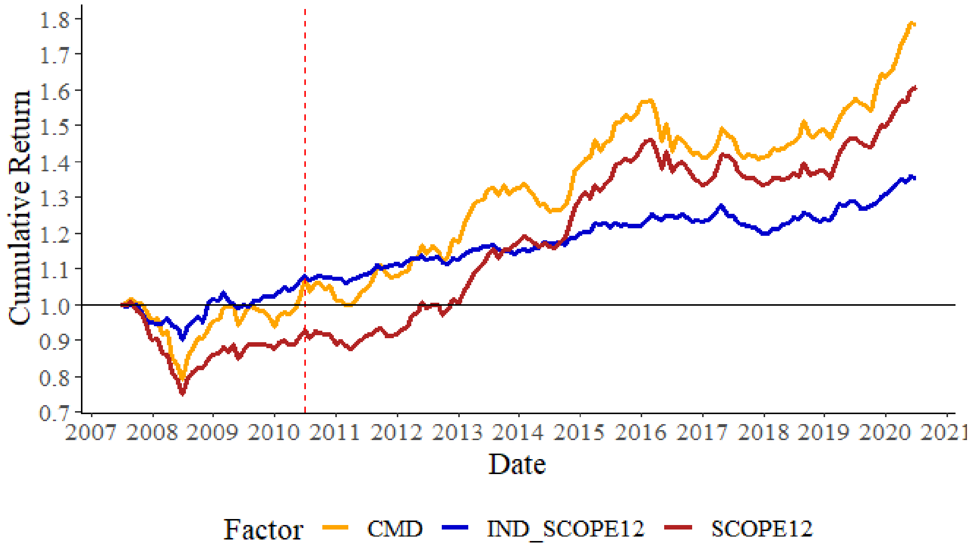

Figure 2: Cumulative returns of the CMD, SCOPE12, and IND_SCOPE12 factor in the period from 07/2007 to 06/2020.

significant at the 5% level.

Despite the close resemblance of SCOPE12 and the two alternative factors, I only consider the SCOPE12 factor in my subsequent analyses for the following reasons. First, carbon risks are more likely to materialize based on absolute carbon intensities rather than emission intensity differences within industries. Hence, the economic intuition appears weaker for the industry-sorted factor. Secondly, related research finds that ESG ratings positively depend on firm size (Kaiser, 2020) because larger firms might be able to invest in greener technologies. As a consequence, double sorting rules out potential size effects present in my data and the CMD factor.

Table 5 reports the descriptive statistics and Pearson correlation coefficients of the SCOPE12 factor with the Fama-French 3-factor and 5-factor model extended for momentum. The analysis is based on monthly returns in the period from 07/2007 to 06/2020. Panel A reports the results for the 4-factor model underlying my further analyses. Additionally, panel B shows the relation of the SCOPE12 factor to the Fama and French (2015) 5-factor model augmented with momentum.[13] Although I do not further consider the Fama-French 5-factor model in later analyses, I include it here to ensure the robustness of the carbon factor and rule out potentially strong interdependencies between the factors. The restriction to the 4-factor model as benchmark appears reasonable as CMA and RMW themselves are relatively new factors which still raise concerns regarding their robustness and economic rationale (e.g., Blitz, Hanauer, Vidojevic, & van Vliet, 2018). Moreover, inevitable additional dependencies between the factors would further complicate the interpretation of the results. Particularly, the highly negative correlation of -0.78 between HML and RMW raises concerns regarding potential redundancies.

Among the downloaded factors only the returns on HML, MOM, and RMW are at least weakly statistically significant. While common factors do not need to be uncorrelated in fundamental factor models (Tsay, 2010), problems of multicollinearity might arise if the SCOPE12 factor covaries strongly with one or more of the standard factors used in the asset pricing literature. Overall, the SCOPE12 factor seems to be only weakly correlated ($|\rho| < 0.3$) with the factors of either model. Of course, looking at the correlation matrix does not detect constellations in which collinearity exists between three or more variables and pairwise correlations are low (Kennedy, 2008). Therefore, I additionally orthogonalize the factors in section 6.4 and eliminate all linear relations between the factors. Nevertheless, the observed low correlations in both panels provide a first indication that the SCOPE12 factor is not merely a replication of known factors but might extend the explanatory power of standard factor models.

6.3. Explanatory Power of the SCOPE12 Factor

6.3.1. Carbon Intensity-Sorted Quintile Portfolios

A frequent criterion for a factor to be deemed relevant is that it enhances the explanatory power in a time series regression relative to previously proposed factors. I therefore test whether the carbon factor can improve the adjusted R^2 compared to the Carhart 4-factor model. The underlying model is the model presented in equation 8 extended with the SCOPE12 factor as additional explanatory variable. The resulting model used in the following subsections can be writ-

[13] While SMB refers to the original size factor constructed based on the book-market ratio sorted portfolios, I denote the Fama and French (2015) size factor as $SMB^{\dagger}$. $SMB^{\dagger}$ is constructed as the average of the three size factors based on the book-market, profitability, and investment sort. RMW and CMA refer to the Fama and French (2015) profitability and investment factor, respectively.

Table 5: Descriptive statistics and correlations of the factors from 07/2007 to 06/2020.

Panel A. Carhart 4-factor model and SCOPE12.

				Correlations			
	Mean	t-stat	σ	MKT	SMB	HML	MOM
MKT	0.215	0.47	5.677				
SMB	0.049	0.320	1.917	0.014			
HML	-0.340*	-1.689	2.512	0.495	-0.049		
MOM	0.815***	2.631	3.870	-0.477	-0.034	-0.532	
SCOPE12	0.323**	2.137	1.885	-0.230	0.072	-0.274	0.053

Panel B. Fama-French 5-factor + momentum model and SCOPE12.

				Correlations					
	Mean	t-stat	σ	MKT	SMB†	HML	MOM	RMW	CMA
MKT	0.215	0.47	5.677						
SMB†	0.072	0.48	1.876	0.028					
HML	-0.340*	-1.689	2.512	0.495	0.013				
MOM	0.815***	2.631	3.870	-0.477	-0.063	-0.532			
CMA	-0.047	-0.403	1.463	-0.213	-0.209	0.406	0.120		
RMW	0.431***	3.425	1.573	-0.362	-0.051	-0.784	0.401	-0.348	
SCOPE12	0.323**	2.137	1.885	-0.230	0.066	-0.274	0.053	-0.083	0.141

Note. *p<0.1, **p<0.05, ***p<0.01; mean is reported as monthly percentage.

ten as

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i^{MKT} f_t^{MKT} + \beta_i^{HML} f_t^{HML} + \beta_i^{SMB} f_t^{SMB} + \beta_i^{MOM} f_t^{MOM} + \beta_i^{SCOPE12} f_t^{SCOPE12} + \varepsilon_{i,t}. \tag{10}$$

Since it is lively debated in the finance literature whether asset pricing tests should be conducted with portfolios or single assets (e.g., Pukthuanthong et al., 2019), I first test the SCOPE12 factor with portfolio returns as left-hand-side variable. That is, I use the carbon intensity-sorted and value-weighted quintile portfolio returns from above. As before, the first quintile contains the clean companies, i.e., the 20% of stocks with the lowest carbon intensity. In contrast, the fifth quintile contains the most carbon intensive or dirty stocks. Table 6 reports the median carbon intensity of the five portfolios during the main sample period. As the carbon intensities tend to decrease in all quintiles during the sample period, I report the average of the annual median figures. While the median carbon intensities of the first three portfolios are relatively close to each other in absolute terms, the gap widens for the two upper quintiles. The median in the dirtiest quintile is even more than six times as high as in the fourth quintile and nearly 95 times larger than in the least carbon intensive portfolio.

Table 7 reports the results of the quintile portfolio regressions on the 4-factor model (panel A) and the extended 5-factor model (panel B). In all regressions the market betas are close to one and statistically significant indicating a broad exposure to the market. With respect to the HML factor there is a contrarian exposure between the fifth quintile and the other four quintiles in both panels. While all quintile portfolios load negatively on the SMB factor, only the betas of the three intermediate carbon intensity portfolios are statistically significant. Lastly, the momentum factor reliably explains returns only in the third portfolio and all regression intercepts are insignificant in both panels. Although the traditional 4-factor model cannot be rejected based on this observation, the alphas' absolute distance from zero decreases in all quintiles through the addition of SCOPE12. Panel B shows the corresponding SCOPE12 factor loadings. The second quintile has a marginally higher carbon factor beta than the first quintile. For the higher carbon intensity portfolios the factor loading decreases monotonically and becomes the largest in magnitude for the fifth quintile with −0.729. The third portfolio's SCOPE12 beta of close to zero indicates that no significant exposure to carbon risk exists in this quintile.

Looking at the change in adjusted R^2, ΔR^2, reported in the last column of panel B, the SCOPE12 factor adds explanatory power in each of the quintiles except for the third one. To check whether this difference is statistically significant, the last column in panel B also reports the result of the heteroskedasticity robust F-statistic for nested models (e.g., Wooldridge, 2016). Except for the third quintile, the F-test is highly significant and confirms the SCOPE12 factor's ability to enhance explanatory power. Especially for the high carbon intensity portfolios, the additional adjusted R^2 turns out larger with 1.82% in the fourth quintile and 3.80% in the fifth quintile.

To ensure the robustness of my results, the portfolio regression analysis is repeated for the IND_SCOPE12 and the

Table 6: Median carbon intensities in the carbon intensity- sorted quintile portfolios from 07/2010 to 06/2020.

Quintile	Clean	2	3	4	Dirty
Median	6.50	17.95	36.24	91.97	616.73

Table 7: 4-factor model and 5-factor model with SCOPE12 for carbon intensity-sorted quintile portfolios from 07/2010 to 06/2020.

Panel A. 4-factor model.

Quintile	MKT	HML	SMB	MOM	Alpha	Adj. R^2
Clean	1.072***	-0.338***	-0.018	0.066	0.094	93.68
	(31.34)	(-6.75)	(-0.25)	(1.34)	(0.72)	
2	0.985***	-0.293***	-0.096*	-0.032	0.138	95.89
	(40.87)	(-6.12)	(-1.81)	(-0.83)	(1.45)	
3	0.929***	-0.337***	-0.242***	0.076*	-0.058	94.13
	(33.10)	(-5.81)	(-3.79)	(1.70)	(-0.51)	
4	1.022***	-0.345***	-0.204***	0.003	-0.087	92.96
	(39.74)	(-5.17)	(-2.68)	(0.06)	(-0.69)	
Dirty	1.036***	0.229***	-0.091	0.014	-0.188	93.10
	(37.92)	(3.17)	(-1.23)	(0.25)	(-1.43)	

Panel B. 5-factor model.

Quintile	MKT	HML	SMB	MOM	SCOPE12	Alpha	Adj. R^2	ΔR^2
Clean	1.079***	-0.272***	-0.034	0.061	0.276***	-0.009	94.30	0.62***
	(31.78)	(-5.36)	(-0.54)	(1.31)	(4.15)	(-0.07)		
2	0.992***	-0.225***	-0.112**	-0.037	0.282***	0.032	96.67	0.78***
	(46.29)	(-5.41)	(-2.43)	(-0.94)	(4.42)	(0.36)		
3	0.929***	-0.345***	-0.240***	0.077*	-0.033	-0.046	94.09	-0.04
	(33.28)	(-5.66)	(-3.78)	(1.73)	(-0.60)	(-0.41)		
4	1.010***	-0.451***	-0.178***	0.012	-0.445***	0.079	94.78	1.82***
	(40.41)	(-6.25)	(-2.70)	(0.29)	(-5.46)	(0.66)		
Dirty	1.017***	0.055	-0.048	0.028	-0.729***	0.085	96.90	3.80***
	(48.81)	(1.00)	(-0.86)	(0.71)	(-11.82)	(0.89)		

Note. *p<0.1, **p<0.05, ***p<0.01; t-statistics in parentheses; heteroskedasticity robust standard errors used. The last column reports the significance of the heteroskedasticity consistent F-test for nested models.

CMD factor, respectively. Table 20 in appendix B contains the regression table for both factors as additional explanatory variables. In contrast to SCOPE12, the industry-sorted carbon factor adds more explanatory power for the least carbon intensive portfolio. The CMD factor performs similar to SCOPE12 for carbon intensive portfolios and explains more variance in the less carbon intensive portfolios. Overall, the results from above remain largely unchanged with significant carbon betas and F-tests in the same portfolios as before.

6.3.2. Impact of SCOPE12 on Single Stock Returns

The previous section shows that the SCOPE12 factor performs well when applied to carbon intensity-sorted portfolios. To be included in these portfolios, however, carbon emission information must be available for the respective stock. As shown in table 2, such information is not available for a large fraction of firms. Furthermore, a frequently mentioned concern with portfolio formation is that it renders stock-specific effects unobservable. Therefore, I repeat the time series regression on the 5-factor model (see equation 10) with individual stocks' excess returns as dependent variable. In comparison to the previous portfolio analysis, I break down the ten year period into two subperiods of five years from 07/2010 to 06/2015 with 868 observations and from 07/2015 to 06/2020 with 897 available complete time series. This procedure allows for approximately 30% larger subsamples in each of the periods because a complete ten year return time series is only available for 667 firms.

Table 8 reports the absolute and relative frequency of significant factor loadings for each of the conventional significance levels. Results for the period from 07/2010 to 06/2015 are shown in panel A and the subsequent period in panel B. The proportion of significant SCOPE12 regression coefficients is relatively similar in both panels, albeit slightly higher in the first five year period. In the first five year subperiod more stocks have a significant exposure to SCOPE12 than to the momentum factor. Furthermore, SCOPE12 performs only slightly worse than the HML factor during this time span. In panel B, the carbon factor is only marginally less significant than the momentum factor. In sum, SCOPE12 helps to explain the variation of individual stock returns.

The above analysis jointly examines the full sample of emission rated and unrated stocks. However, concern exists that the carbon factor's significance is driven by the emission rated stocks, which are also included in the factor construction. Therefore, a further differentiation into *rated* and *unrated* subsamples appears reasonable. Since some assets' carbon emissions are not continuously available and the universe of stocks with CO_2 emission information is gradually extended by data providers, I classify stocks with less than three years (out of ten years) of emission information as unrated. Tables 21 and 22 in appendix C report the subperiod performances of the SCOPE12 factor for rated and unrated stocks, respectively. Looking at the significant factor betas, it can be seen that the SMB factor is of higher relevance in the unrated sample. This is consistent with the frequent observation that firms without ESG-related ratings tend to be smaller on average (e.g., Hübel & Scholz, 2020). Although the SCOPE12 factor is a bit less significant in the unrated sample, the patterns remain largely unchanged. It is to be noted, however, that the other factor candidates are also less significant in the unrated subsample, except for SMB and MOM in the first subperiod. Looking at panel B in table 22, the carbon factor's significance levels closely resembles those of MOM and HML. These solid results in the unrated sample suggest that the carbon factor can be generalized beyond stocks with available emission information, which creates a sound foundation for the more rigorous single stock analyses in chapter 7.

6.4. Orthogonal Factor Model

6.4.1. Orthogonalization of the Common Factors

Despite only low or moderate correlations between the factors, as shown in table 5, equation 2 indicates that the variance of asset returns not only depends on factor loadings but also on the covariances of factor returns. I therefore orthogonalize the common factors and reassess the sensitivities of asset returns to the unique variation of the underlying factor. This helps to determine how much of the systematic variance is indeed attributable to the respective factor. In order to disentangle the linear relationships in the 5-factor model, I use the democratic orthogonalization procedure suggested by Klein and Chow (2013). Their method builds on Löwdin's symmetric orthogonalization and assigns equal shares of common variance to the respective orthogonalized factors. This results in the identical R^2 in a multiple regression as if the original factors were used. Moreover, using symmetric orthogonalization ensures maximal similarity with the original factors and a unique result in comparison to sequential procedures.[14]

I apply the symmetric orthogonalization to the downloaded Carhart 4-factors and the self-calculated SCOPE12 factor.[15] Table 9 reports the distributional properties of the orthogonalized factors and their correlations with the original factors. The similarity of the orthogonal factors with the original factors is very high with most correlation coefficients close to unity. That is, correlations are 0.970, 0.998, 0.896, 0.941, and 0.976 for the MKT, SMB, HML, MOM, and SCOPE12 factor, respectively. Only HML, which is the factor with the highest cross-correlations in table 5, is slightly lower correlated with its orthogonal counterpart. With regard to the orthogonal factors' mean returns, the general direction and magnitude remains similar to the original factors. The largest change can be observed for the mean return of the market factor, which changes from 0.215% (MKT) to 0.621% ($\text{MKT}^{\perp}$) per month. Analogously to the original factors, only momentum and the carbon factor have mean returns different from zero at the 5% level. By construction, the standard deviations are unchanged compared to the original factors and the correlations among the orthogonally transformed factors themselves are zero.

I use the orthogonally transformed factor returns as explanatory variables for the value-weighted returns of the carbon intensity-sorted quintile portfolios in table 10. Since the variation of the explanatory variable is now unique to it, ordinary least squares (OLS) regression estimates tend to be more precise given a correct allocation of the jointly explained variance (Kennedy, 2008). In comparison with the results of the original factors in panel B of table 7, I find changes in the signs and magnitudes of the regression coefficients. The regression alphas and (adjusted) R^2, however, remain unchanged to the previous analysis by construction. First, market betas tend to be lower for the uncorrelated factors but are still close to one, except for the third quintile. The greatest changes can be observed with regard to $\text{HML}^{\perp}$ and $\text{MOM}^{\perp}$. In contrast to the original value factor, the loading on $\text{HML}^{\perp}$ is statistically and economically highly significant only for the most carbon intensive portfolio. Furthermore, the orthogonal momentum beta is negative and highly significant for most portfolios. The observed changes suggest that the original factors tend to overestimate the portfolios' exposure to systematic risk arising from the market and value factor, while underestimating the systematic risk from momentum. In contrast, factor loadings on $\text{SMB}^{\perp}$ and $\text{SCOPE12}^{\perp}$ closely resemble the original factors. The $\text{SCOPE12}^{\perp}$ betas are all highly significant and monotonically decrease from the lowest to the highest carbon intensity portfolio. Especially the dirty quintile loads even stronger on the orthogonal carbon factor than the original factor.

Focusing on the carbon factor, the similar results when using $\text{SCOPE12}^{\perp}$ instead of SCOPE12 as explanatory variable provide further evidence that the original carbon factor

[14] The uniqueness prove goes back to Aiken, Erdos, and Goldstein (1980).

[15] In the following the orthogonal factors are identified by the superscript perpendicular symbol, $\perp$.

Table 8: Significance levels of the single stock factor betas.

Panel A. Single stock regressions for the period 07/2010 to 06/2015.

Significance	10%		5%		1%	
$N = 868$	Abs.	%	Abs.	%	Abs.	%
MKT	847	97.58	842	97.00	806	92.86
HML	251	28.92	179	20.62	82	9.45
SMB	317	36.52	230	26.50	115	13.25
MOM	110	12.67	61	7.03	15	1.73
SCOPE12	201	23.16	137	15.78	65	7.49

Panel B. Single stock regressions for the period 07/2015 to 06/2020.

Significance	10%		5%		1%	
$N = 897$	Abs.	%	Abs.	%	Abs.	%
MKT	863	96.21	846	94.31	795	88.63
HML	220	24.53	156	17.39	71	7.92
SMB	373	41.58	278	30.99	154	17.17
MOM	198	22.07	128	14.27	47	5.24
SCOPE12	182	20.29	117	13.04	47	5.24

Note. Heteroskedasticity robust standard errors used.

Table 9: Descriptive statistics of orthogonal factors and correlation with original factors from 07/2007 to 06/2020.

				Correlations with original factors				
	Mean	t-statistic	σ	MKT	SMB	HML	MOM	SCOPE12
$MKT^{\perp}$	0.621	1.366	5.677	0.970				
$SMB^{\perp}$	0.041	0.264	1.917		0.998			
$HML^{\perp}$	-0.225	-1.120	2.512			0.896		
$MOM^{\perp}$	0.928***	2.995	3.870				0.941	
$SCOPE12^{\perp}$	0.346**	2.294	1.885					0.976

Note. *p<0.1, **p<0.05, ***p<0.01; mean is reported as monthly percentage.

Table 10: Orthogonalized 5-factor model for carbon intensity-sorted quintile portfolios from 07/2010 to 06/2020.

Quintile	$MKT^{\perp}$	$HML^{\perp}$	$SMB^{\perp}$	$MOM^{\perp}$	$SCOPE12^{\perp}$	Alpha	Adj. R^2
Clean	0.991***	0.014	-0.005	-0.117***	0.151**	-0.009	94.30
	(34.55)	(0.37)	(-0.08)	(-3.10)	(2.49)	(-0.07)	
2	0.926***	0.056*	-0.086*	-0.199***	0.147**	0.032	96.67
	(48.21)	(1.73)	(-1.85)	(-6.10)	(2.57)	(0.36)	
3	0.847***	-0.055	-0.223***	-0.071*	-0.107**	-0.046	94.09
	(33.80)	(-1.16)	(-3.51)	(-1.93)	(-2.17)	(-0.41)	
4	0.942***	-0.050	-0.173**	-0.154***	-0.495***	0.079	94.78
	(37.71)	(-0.84)	(-2.59)	(-4.48)	(-6.75)	(0.66)	
Dirty	1.038***	0.448***	-0.056	-0.272***	-0.874***	0.085	96.90
	(60.77)	(11.33)	(-1.01)	(-8.54)	(-16.63)	(0.89)	

Note. *p<0.1, **p<0.05, ***p<0.01; t-statistics in parentheses; heteroskedasticity robust standard errors used.

only has weak linear relationships with the other four factors and provides unique explanatory power. Since all factors are mutually uncorrelated, the presented analysis also rules out more complex linear relationships among three or more factors in contrast to the correlation analysis in table 5.

6.4.2. Decomposition of the Coefficient of Determination

While the previous section provides an overview of the portfolio returns' sensitivities towards the underlying factors, the risk can be closer examined decomposing the coefficient of determination. Due to the absence of covariance between the explanatory variables, fractions of R^2 can be attributed to the individual factors. Asset i's total coefficient of determination, R_i^2, is therefore the sum of the contributions of the K individual factors, $R_{i,k}^2$ (Klein & Chow, 2013). The resulting formula is

$$R_i^2 = \sum_{k=1}^{K} R_{i,k}^2 \,, \quad \text{where } R_{i,k}^2 = \left(\frac{\hat{\beta}_{k,i}^{\perp} \hat{\sigma}_k^{\perp}}{\hat{\sigma}_i} \right)^2 . \tag{11}$$

The sample estimate of asset i's factor loading on the kth orthogonalized factor is denoted by $\hat{\beta}_{k,i}^{\perp}$. The variables $\hat{\sigma}_k^{\perp}$ and $\hat{\sigma}_i$ denote the sample estimate of the standard deviation of orthogonal factor k's and asset i's returns, respectively.

I apply the systematic risk decomposition to the carbon intensity-sorted quintile portfolios and report the five factors' provision of additional explanatory power in table 11. As mentioned above, the cumulative portion of explained variance in portfolio returns is equivalent to the original factors. As expected for diversified portfolios, MKT$^{\perp}$ contributes most of the explanatory power for each of the quintile portfolios with about 90% of the explained variance. In contrast, the R^2 contributions of HML$^{\perp}$, SMB$^{\perp}$, and MOM$^{\perp}$ lie below 1% in most regressions. Only in the most carbon intensive portfolio, the quintile with the highest R^2, HML$^{\perp}$ and MOM$^{\perp}$ have higher explanatory power with 4.04% and 1.99%, respectively. Of even higher importance in this portfolio, however, is the SCOPE12$^{\perp}$ factor which explains 6.53% of the return's variance. It can be seen that SCOPE12$^{\perp}$ is particularly important in more carbon intensive portfolios, whereas in quintiles 1 to 3 its additional explanatory power is lower. The carbon factor therefore seems to primarily explain return variation of equities especially exposed to the low carbon transition of the European economy.

6.5. Maximum Squared Sharpe Ratios

An alternative to running time series regressions on the factor candidates is the right-hand-side approach in form of spanning regressions proposed by Fama and French (2018). In a spanning regression each factor is individually regressed on the model's remaining factors. In case the regression constant is significantly different from zero, the factor used as dependent variable enhances the right-hand-side model's explanation of average returns. Hence, this approach allows for an additional perspective on whether the carbon factor should be added to the Carhart 4-factor model.

Based on the original results by Gibbons, Ross, and Shanken (1989), Fama and French (2018) show that

$$\hat{\boldsymbol{\alpha}}' \boldsymbol{\Omega}_{\boldsymbol{\varepsilon}}^{-1} \hat{\boldsymbol{\alpha}} = SR^2(\boldsymbol{\Pi}, \boldsymbol{f}) - SR^2(\boldsymbol{f}) = SR^2(\boldsymbol{\alpha}), \tag{12}$$

holds. Where $\boldsymbol{\Pi}$ is the combination of assets' excess returns and alternative factors' returns, which are currently not included in the model. $\boldsymbol{f}$ denotes the factors of the model under consideration. On the left-hand-side of the equation, $\hat{\boldsymbol{\alpha}}$ is a $N \times 1$ vector of estimated regression intercepts from the time series regression of $\boldsymbol{\Pi}$ on $\boldsymbol{f}$. The resulting estimate of the $N \times N$ residual covariance matrix is $\boldsymbol{\Omega}_{\varepsilon}$. Hence, the left-hand-side expression in equation 12 is solely determined by the choice of the factors and their resulting maximum squared Sharpe ratio (SR). Since the squared Sharpe ratio of the intercept, $SR^2(\boldsymbol{\alpha})$, has to be minimized, the factor model with the highest squared Sharpe ratio, $SR^2(\boldsymbol{f})$, is chosen. $SR^2(\boldsymbol{\Pi}, \boldsymbol{f})$ denotes the maximum squared Sharpe ratio which can be achieved by the combination of $\boldsymbol{\Pi}$ and $\boldsymbol{f}$.

The formulation in equation 12 also provides the basis for the estimation of the marginal contribution an additional factor k has to the model's maximum squared Sharpe ratio (Fama & French, 2018):

$$\frac{\hat{\alpha}_k^2}{\hat{\sigma}_{\varepsilon,k}^2} = SR^2(\boldsymbol{f}, k) - SR^2(\boldsymbol{f}) \tag{13}$$

In this equation, $\hat{\alpha}_k$ denotes the constant and $\hat{\sigma}_{\varepsilon,k}$ the residual standard deviation resulting from the time series regression of factor k on the remaining factors of the model. Their squared ratio equates the difference between the maximum squared Sharpe ratio of the model of $\boldsymbol{f}$ and k combined and $\boldsymbol{f}$ individually. The left side of equation 13 indicates that the contribution of factor k to the model's squared Sharpe ratio depends on two factors. First, if $\hat{\alpha}_k$ is close to zero, the returns of factor k are well explained by the other factors and its contribution is small. Secondly, the contribution to the squared Sharpe ratio is small if the other factors only explain little variation in factor k's return ($\hat{\sigma}_{\varepsilon,k}$ is large). As mentioned above, the statistical significance of a factor's marginal contribution is measured by the size of the t-statistic for the intercept in a spanning regression.

Table 12 reports the results from the spanning regressions of the Carhart 4-factor model (panel A) and the extended 5-factor model (panel B) in the time from July 2010 to June 2020. The last column reports the marginal contribution of each factor to the maximum squared Sharpe ratio when added to the model's other factors. The marginal contributions and regression coefficients of the traditional four factors are similar across both models, indicating that the SCOPE12 factor adds a unique portion to the 4-factor maximum squared Sharpe ratio. Among the traditional four factors, only MKT and MOM have significant intercepts and therefore reliably contribute to the Sharpe measure in both models. Although the regression constant is larger for MKT than MOM in both panels, MKT's larger residual variance leads to a smaller Sharpe ratio contribution in comparison

Table 11: R^2 decomposition of carbon intensity-sorted quintile portfolios from 07/2010 to 06/2020.

	Decomposed R^2						
Quintile	$MKT^{\perp}$	$HML^{\perp}$	$SMB^{\perp}$	$MOM^{\perp}$	$SCOPE12^{\perp}$	R^2	$1-R^2$
Clean	93.85	0.00	0.00	0.45	0.24	94.54	5.46
2	94.85	0.09	0.10	1.51	0.26	96.81	3.19
3	93.08	0.10	0.77	0.23	0.16	94.34	5.66
4	90.99	0.07	0.36	0.84	2.75	95.00	5.00
Dirty	84.44	4.04	0.03	1.99	6.53	97.03	2.97

Note. (Decomposed) R^2 values are given as percentages.

Table 12: Spanning regressions and $SR^2(\boldsymbol{f})$ decomposition for the 4-factor and 5-factor model from 07/2010 to 06/2020.

Panel A. 4-factor model.

	α	MKT	HML	SMB	MOM	R^2	σ_ε	$\alpha^2/\sigma^2_\varepsilon$
MKT	1.223***		0.704***	0.071	-0.356**	24.90	4.217	0.084
	(2.949)		(3.766)	(0.303)	(-2.202)			
HML	-0.112	0.155***		-0.019	-0.350***	35.60	1.977	0.003
	(-0.556)	(3.766)		(-0.175)	(-4.996)			
SMB	0.090	0.011	-0.014		0.009	0.10	1.675	0.003
	(0.530)	(0.303)	(-0.175)		(0.138)			
MOM	0.883***	-0.113**	-0.506***	0.018		30.63	2.376	0.138
	(3.875)	(-2.202)	(-4.996)	(0.138)				

Panel B. 5-factor model.

	α	MKT	HML	SMB	MOM	SCOPE12	R^2	σ_ε	$\alpha^2/\sigma^2_\varepsilon$
MKT	1.299***		0.646***	0.084	-0.349**	-0.224	25.34	4.205	0.095
	(3.056)		(3.236)	(0.357)	(-2.157)	(-0.831)			
HML	0.049	0.129***		0.006	-0.309***	-0.402***	41.80	1.879	0.001
	(0.248)	(3.236)		(0.060)	(-4.547)	(-3.497)			
SMB	0.061	0.013	0.005		0.008	0.078	0.56	1.671	0.001
	(0.346)	(0.357)	(0.060)		(0.114)	(0.729)			
MOM	0.863***	-0.111**	-0.493***	0.015		0.050	30.70	2.374	0.132
	(3.653)	(-2.157)	(-4.547)	(0.114)		(0.329)			
SCOPE12	0.373**	-0.027	-0.239***	0.059	0.019		18.83	1.448	0.066
	(2.518)	(-0.831)	(-3.497)	(0.729)	(0.329)				

Note. *p<0.1, **p<0.05, ***p<0.01; α and R^2 are reported as a percentage.

to MOM. In both models, the HML and SMB factor's intercept is statistically non-distinguishable from zero. That is, HML's returns can be explained through highly significant loadings on the MKT, MOM, and SCOPE12 factor. Although SMB cannot be explained by the other factors, SMB's average returns are close to zero during the sample period and therefore cannot be reliably distinguished from zero either. Focusing on the SCOPE12 factor's spanning regression in the last row of panel B, a statistically significant intercept of 0.373% per month can be found. Among the explanatory variables only HML reliably explains the carbon factor's returns. Besides, SCOPE12's spanning regression has the lowest residual standard deviation. This results in the carbon factor having the third largest marginal contribution to the 5-factor model's squared Sharpe ratio with $\frac{\hat{\alpha}^2}{\hat{\sigma}^2_\varepsilon} = 0.066$ behind MOM and MKT.

Overall, the spanning regression analysis suggests that the SCOPE12 factor should be included in the 5-factor model as it provides a statistically significant increase in the model's $SR^2(\boldsymbol{f})$. Although SMB and HML appear redundant in the above analysis, the results of spanning regressions tend to be sensitive to the chosen time period (Fama & French, 2018). Hence, I provide a more rigorous examination of which factors should be considered in a factor model in section 7.1.

7. Factor Identification

7.1. Covariance Matrix of Returns and Factor Returns

The analyses presented so far provide a first indication for the relevance of a carbon factor in order to explain the cross-section of European stock returns. Being aware of potential data-snooping objections against additionally introduced factors (e.g., Campbell et al., 1997), I conduct additional rigorous tests of my factor. I follow the protocol for the identification of genuine risk factors proposed by Pukthuanthong et al. (2019). According to the authors, a genuine risk factor must fulfill the following requirements: a) it must relate to the covariance matrix of stock returns, b) it must be priced in the cross-section of stock-returns, and c) it should offer a risk-reward ratio consistent with risk pricing.

This section constitutes the first stage in Pukthuanthong et al.'s protocol and aims to determine factors that systematically move asset prices. This is achieved by extracting statistical factors from the sample covariance matrix and relating them to the fundamental factors[16] via canonical correlation analysis. Facing the problem of a large number of observations N and a smaller number of time periods T, I extract principal components (PC) from the transformed $T \times T$ sample covariance matrix of European stock returns using the asymptotic principal component analysis as described in section 3.2. For the full ten year period from July 2010 to June 2020 ($T = 120$), I have complete return time series for $N = 667$ stocks. I construct the $\boldsymbol{\Omega_R}$ matrix (compare equation 6) based on this data and retain the first ten PCs ($L = 10$), which account for approximately 90% of the volatility in the covariance matrix. The threshold is in line with Pukthuanthong et al. (2019) and ensures that most of the covariances are explained by the calculated statistical factors. Consequently, equation 4 holds approximately.

The next step is to link those ten PCs with the five factor candidates via canonical correlations. The rationale for the usage of canonical correlations is the rotational indeterminacy of statistical factor models mentioned in section 3.2, which restricts their direct interpretation. Hence, checking for correlations between linear combinations of the extracted PCs and linear combinations of the factor candidates provides a reasonable proxy whether a factor is linked to the sample covariance matrix (Pukthuanthong et al., 2019). If a factor candidate is not significantly canonically correlated with the eigenvectors maximizing the explained variance, it is rejected as a genuine risk factor and does not enter the cross-sectional regression in the second stage of the protocol. The canonical correlation analysis between the ten eigenvectors extracted from the $\boldsymbol{\Omega_R}$ matrix and the five factor candidates ($K = 5$) results in $\min(K, L)$ five canonical variate pairs and hence five canonical correlations. For each canonical variate pair, the weights of the linear combinations of both the eigenvectors and the factor candidates are chosen such that their resulting canonical correlation is maximized. Moreover, additional constraints ensure the uncorrelatedness of the canonical variate pairs. Panel A of table 13 reports the canonical correlations and their respective F-statistics obtained from Wilks' lambda test. The test's null hypothesis is that the correlations in the current and all following rows are equal to zero. It can be rejected for the first four canonical correlations. Especially the first canonical correlation is large and relatively close to one. Hence, the hypothesis of independence of the two multivariate sets of variables can be rejected overall. Panel B builds on this result and reports the significance level for the factor candidates. The reported t-statistics are obtained from the regression of each canonical variate for the set of principal components on all five factor candidates over the whole sample period. Since there are five canonical variate pairs, I run five regressions in total and calculate the arithmetic average of the absolute t-statistics for each factor. The last row in panel B repeats the analysis and calculates only the mean absolute t-statistics of canonical correlations which are statistically significant at the 5% level. Therefore, it only considers the t-statistics of the first four regressions. Since an insignificant canonical correlation indicates that no relationship between the two sets of variables exists, only significant correlations are considered in the following. Moreover, as absolute t-statistics are reported, one-tailed cut-off levels are used in the following discussion. Pukthuanthong et al. (2019) argue that a genuine risk factor should have an average absolute t-statistic of statistically significant canonical correlations in excess of the one-tailed 2.5% cut-off level ($t \approx 1.96$). Focusing on the average t-statistics from significant correlations, only the MOM factor is below the approximate critical rejection level of 1.96. However, its t-statistic ($t = 1.705$) exceeds the threshold for the one-tailed 5% significance level ($t \approx 1.65$). Besides the market risk factor ($t = 5.669$), the SCOPE12 factor has the second highest average absolute t-statistic ($t = 3.537$), which indicates that the carbon factor should be considered in the cross-sectional regressions.

Pukthuanthong et al. (2019) mention the potential presence of nonstationarity as a relevant restriction to the factor extraction from the sample covariance matrix. In order to avoid excluding a relevant risk factor and to account for fluctuations in the relative importance of the five factor candidates over time, I also conduct APCA and calculate canonical correlations for the two five year subperiods from 07/2010 to 06/2015 and 07/2015 to 06/2020, respectively. The results from the additional analyses are reported in tables 23 and 24 in appendix D. Considering the shorter time periods offers larger samples of complete return time series, which again allows to check for robustness of the asymptotic PCs from the full sample period. In general, the canonical correlations and significance levels of the subperiods are close to the full ten year period. However, as expected and in line with Pukthuanthong et al. (2019), I find that the factor candidates vary in relative importance between the subsamples. Particularly, the momentum factor is highly significant with an average absolute t-statistic of 2.225 in the more recent subperiod. Since my time series is rather short and due to the weak significance of MOM for the full time period, I do not want to incorrectly exclude the momentum factor. Based on my results, I conclude that each of the five factor candidates is sufficiently related to the sample covariance matrix. Hence, all five factors are incorporated in the second step

[16] Strictly speaking, MKT is a macroeconomic factor. However, for simplicity I refer to all five factor candidates (MKT, HML, SMB, MOM, and SCOPE12) as fundamental factors in the following.

Table 13: Canonical correlations with asymptotic principal components and significance of factor candidates for the period from 07/2010 to 06/2020 ($N = 667$).

Panel A. Canonical correlations.

Canonical variate	Canonical correlation	F-stat	df1	df2	p-value
1	0.847***	7.423	50	482.2	0
2	0.715***	4.757	36	399	0
3	0.557***	3.095	24	310.9	0
4	0.440**	2.060	14	216	0.02
5	0.186	0.653	6	109	0.69

Panel B. Significance of factor candidates.

	MKT	SMB	HML	MOM	SCOPE12
Mean absolute t-stat	4.659	1.936	2.569	1.529	2.849
Mean absolute t-stat (significant corr.)	5.669	2.344	2.843	1.705	3.537

Note. *p<0.1; **p<0.05; ***p<0.01; t-statistics in panel B are calculated with heteroskedasticity robust standard errors.

cross-sectional regressions.

7.2. Cross-Sectional Regressions and the Carbon Risk Premium

After ensuring that the factor candidates are related to the covariance matrix of stock returns, I conduct the second stage of Pukthuanthong et al.'s protocol for factor identification and check whether the factor candidates command risk premiums by running a variant of Fama and MacBeth (1973) cross-sectional regressions. Fama and MacBeth's methodology offers a widely used way to generate standard errors adjusted for cross-sectional correlations (Cochrane, 2005). Although there is an ongoing debate on whether the cross-sectional regressions should use single stock or portfolio returns as dependent variable, I use individual returns for the following analyses.[17] Among the foremost arguments against the usage of portfolio returns "is that diversification into portfolios can mask cross-sectional phenomena in individual stocks that are unrelated to the portfolio grouping procedure" (Jegadeesh, Noh, Pukthuanthong, Roll, & Wang, 2019, p. 274).

In the first step of the Fama-MacBeth procedure, the first-pass regression, I estimate the betas by running the time series regression shown in equation 10 for each stock i individually. For the time series regression either a rolling window regression approach or a complete time series regression with full sample betas can be used (Cochrane, 2005). While Pukthuanthong et al. compute betas based on all available return observations, I follow the rolling regression approach (e.g., Fama & MacBeth, 1973; Görgen et al., 2020). As the available historical return time series varies across the sample assets and the cross-section is smaller compared to other asset pricing papers, I use 36-month rolling window regressions starting in July 2007 to estimate the five beta coefficients. This approach ensures the comparability of betas across assets at each point in time and retains a larger cross-section than longer regression windows.

In the second step, I test whether individual stocks' risky excess returns linearly depend on their covariances with the five factor candidates (Jegadeesh et al., 2019). For this analysis, the beta estimates are treated as explanatory variables and the following second-pass or cross-sectional regression is performed for each month in the time from July 2010 to June 2020 ($t = 1, \ldots, 120$):

$$\begin{aligned} r_{i,t} - r_{f,t} = \alpha_t + \gamma_t^{MKT}\hat{\beta}_{i,t}^{MKT} + \gamma_t^{HML}\hat{\beta}_{i,t}^{HML} + \gamma_t^{SMB}\hat{\beta}_{i,t}^{SMB} \\ + \gamma_t^{MOM}\hat{\beta}_{i,t}^{MOM} + \gamma_t^{SCOPE12}\hat{\beta}_{i,t}^{SCOPE12} + \varepsilon_{i,t}, \\ \text{where } i = 1, \ldots, N. \end{aligned} \tag{14}$$

The hat on the betas denotes that the estimates from the first-pass regression are used and the subscript t indicates the time-variability of the beta coefficients. The regression coefficients, γ_t, are interpreted as the risk premiums per month for each of the five factor candidates. Due to the rolling regression approach, the size of the cross-section varies over time. On average, I have beta coefficients of 1,146 stocks in the cross-sectional regressions. The risk premium for each factor is then calculated as the mean over all cross-sectional regression estimates as

$$\overline{\gamma} = \frac{1}{T}\sum_{t=1}^{T}\hat{\gamma}_t. \tag{15}$$

The Fama-MacBeth procedure then determines the variance

[17] See also Pukthuanthong et al. (2019) for a rationale for the usage of individual assets as left-hand-side variables.

of this estimate as

$$\sigma^2_{\overline{\gamma}} = \frac{1}{T}\sum_{t=1}^{T}\frac{(\hat{\gamma}_t - \overline{\gamma})^2}{T-1}. \tag{16}$$

On the basis of the estimates from equations 15 and 16, I compute the t-statistics according to the usual formula to determine the risk premiums' statistical significance. However, as the betas from the first-pass regression are estimates rather than exact values, the Fama-MacBeth approach is prone to an errors-in-variables (EIV) problem. In line with Pukthuanthong et al. (2019), I therefore repeat the cross-sectional regression with double-sorted portfolio betas as explanatory variables to alleviate the EIV problem. Annually at the end of June, I sort stocks into decile portfolios based on market value. Then, within each of the deciles, I conduct a second sort based on beta percentiles. Pukthuanthong et al. (2019) suggest the usage of deciles in the second step but I additionally report the result for betas sorted into quintiles to account for the smaller size of my sample. Note that this second sort is repeated for each factor individually. Within each of the 10×10 or 10×5 portfolios, I assign the arithmetic average beta to the stocks in the respective portfolio. The average betas are updated each month based on my rolling 36-month estimates to account for retired stocks.

The results from the computations outlined above are reported in table 14. The risk premiums are relatively stable in sign and magnitude across all three regressions. However, the explanatory power is reduced through the EIV correction as the adjusted R^2 drops from 19.14% in the first column to about 4% in the regressions with the double-sorted betas. Across all three models only the intercept is highly significant. In contrast, the other factors lack significant risk premiums, except for SCOPE12 in the EIV corrected models. SCOPE12's monthly risk premium of 0.155% in the 10×5 sorted model is weakly significant, whereas the risk premium of 0.195% in the 10×10 sorted model is highly significant. Overall, the results from the EIV corrected regressions point to the existence of a positive risk premium for holding low carbon intensity firms in my sample. Nevertheless, the results should be interpreted cautiously as the inevitable EIV problem constitutes a serious violation of the OLS assumptions and alternative approaches to treat this problem are discussed in the literature (e.g., Jegadeesh et al., 2019).

7.3. Hedge Portfolio Returns

The previous two sections of this chapter examined one necessary and one sufficient condition of factor testing. However, Pukthuanthong et al. (2019) suggest an additional test to determine whether the risk-reward trade-off is within reasonable limits for the respective risk factor candidate. This test rules out the case that the SCOPE12 factor reflects systematic empirical regularities which are in fact driven by market inefficiency or behavioral biases (Campbell et al., 1997). If deviations from the CAPM or alternative multifactor models are driven by a genuine risk factor, they should be bounded by the factor's relation to variance (MacKinlay, 1995). For non-risk explanations no such relation exists and the Sharpe ratio can be unbounded in theory. Based on theoretical and empirical arguments, MacKinlay (1995) proposes an annual Sharpe ratio of 0.6 as upper bound for a risk factor. Although this threshold is motivated by the historical excess return and standard deviation of the US stock market, I use it as a reasonable proxy for my European sample. If a risk factor delivers a Sharpe ratio significantly exceeding MacKinlay's threshold, this would provide an indication that the underlying factor might not be consistent with a risk-based pricing explanation (Pukthuanthong et al., 2019).

In contrast to the previous section, this robustness check is performed based on hedge portfolios. I construct an equally-weighted hedge portfolio for each of my five factor candidates by going long (short) in the stock quintile with the highest (lowest) factor beta. The betas used for monthly sorting are estimated as in the first-pass time series regression in section 7.2. Panel A in table 15 reports the mean return of the five quintile hedge portfolios. With 0.557% per month the SCOPE12 hedge portfolio offers the largest return, and – except for HML – all hedge portfolios generate positive returns. For none of the portfolios, however, returns are statistically different from zero.

The threshold proposed by MacKinlay (1995) refers to a long-only portfolio. Therefore, the zero-investment hedge portfolios must be combined with a representative long-only portfolio to test whether they exceed the 0.6 bound (Pukthuanthong et al., 2019). For this purpose, the market excess return is added to the hedge portfolio returns, which is referred to as *combined returns* in the following. The resulting mean and volatility of the combined returns are reported in panel B of table 15. Combining the momentum and SCOPE12 hedge portfolio returns with the market excess return offers significant returns at the 5% and 10% level, respectively. The last column reports the statistics of the market excess return for comparison. Based on the combined returns the annual Sharpe ratio is calculated for each factor, which ranges from 0.21 (HML) to 0.68 (MOM). The t-statistic in the last row of panel B indicates whether the Sharpe ratio is statistically different from 0.6. Sharpe ratios for MKT, HML, and SMB are significantly smaller than the threshold, whereas MOM's and SCOPE12's are not statistically different from it. This indicates that the traded versions of the factor candidates yield Sharpe ratios consistent with risk-based explanations. I therefore proceed by retaining all five factor candidates.

8. Carbon Risk Management with Exposure Portfolios

8.1. Construction of Industry Adjusted Exposure Portfolios

Based on the finding that the SCOPE12 factor can be considered a genuine equity risk factor, I address two key issues socially responsible investors face: First, investors are aware of potential negative financial implications carbon risks pose to their portfolio and actively want to manage those instead of withdrawing from problematic stocks (Krueger, Sautner,

Table 14: Estimated risk-premiums for factor candidates from 07/2010 to 06/2020.

	No EIV Correction	EIV Correction	
		10 × 5	10 × 10
MKT	-0.070	-0.286	-0.262
	(-0.147)	(-0.913)	(-0.819)
HML	-0.248	-0.174	-0.195
	(-1.050)	(-1.279)	(-1.358)
SMB	0.066	0.112	0.119
	(0.408)	(1.387)	(1.433)
MOM	0.384	0.243	0.251
	(1.551)	(1.344)	(1.430)
SCOPE12	0.222	0.155*	0.195**
	(1.516)	(1.814)	(2.240)
Intercept	0.910***	1.093***	1.070***
	(5.341)	(3.798)	(3.830)
Adj. R^2	19.14	3.52	4.15

Note. *p<0.1, **p<0.05, ***p<0.01; t-statistics in parentheses; heteroskedasticity robust standard errors are used; risk premiums are expressed as percentages per month.

Table 15: Risk and return of beta-sorted hedge portfolios from 07/2010 to 06/2020.

Panel A. Quintile hedge portfolios.

	MKT	HML	SMB	MOM	SCOPE12
Mean	0.134	-0.130	0.115	0.493	0.557
σ	5.413	5.098	4.663	4.260	4.574
t (mean)	0.27	-0.28	0.27	1.27	1.33

Panel B. Quintile hedge portfolios combined with market excess return.

	Combined portfolio returns					
	MKT	HML	SMB	MOM	SCOPE12	$r_{MKT} - r_f$
Mean	0.747	0.483	0.728	1.106**	1.170*	0.613
σ	9.825	7.902	7.688	5.671	6.723	4.866
t (mean)	0.83	0.67	1.04	2.14	1.91	1.38
SR	0.26***	0.21***	0.33***	0.68	0.60	0.44*
t (SR)	-3.69	-4.25	-2.98	0.83	0.03	-1.79

Note. *p<0.1, **p<0.05, ***p<0.01; mean is reported as monthly percentage; the Sharpe ratio (SR) is annualized; the last row in panel B reports the t-statistic of the Sharpe ratio against the MacKinlay (1995) threshold of 0.6.

& Starks, 2020). Hence, a main goal for risk management must be to help investors identify stocks exposed to carbon risk while maintaining a broad market exposure. Secondly, critics correctly point out that carbon emission information is mainly available for large firms because emission disclosure is still limited. This restricts investors who are aware of carbon risks in their investable universe. To mitigate this problem and to allow investors to identify carbon risks, widely available stock returns can be used. Based on the estimated exposure to the SCOPE12 factor, which equals the regression beta, both firms with and without CO_2 emission information could be considered in the management of carbon risk. For comparability with my previous analyses, I consider quintile exposure portfolios in the following.[18]

For the estimation of return-based exposures I run single stock, 36-month rolling window regressions on the 5-factor model (equation 10) starting in July 2007. I proceed by sorting stocks into industry adjusted and value-weighted quintile portfolios based on beta. For instance, the clean portfolio

[18] Note that by construction a positive (negative) SCOPE12 beta corresponds to less (more) carbon intensive stocks. To avoid confusion and ensure consistency with previous analyses, I therefore classify stocks in the highest (lowest) exposure quintile as clean (dirty).

Table 16: Average industry composition of exposure portfolios.

	Quintile						
Industry	Clean	2	3	4	Dirty	Total	%
Basic Materials	19	19	19	19	18	93	8.1
Consumer Discretionary	45	45	45	45	44	223	19.5
Consumer Staples	18	18	18	18	17	88	7.7
Energy	13	13	13	13	13	64	5.6
Health Care	20	19	19	19	19	96	8.4
Industrials	63	63	63	63	62	313	27.3
Real Estate	17	17	17	17	16	84	7.3
Technology	18	17	17	17	17	85	7.4
Telecommunication	10	10	10	9	9	47	4.1
Utilities	11	11	11	11	10	53	4.6
Total	233	229	229	229	226	1,146	100

Note. Values are rounded.

contains the 20% of stocks with the largest SCOPE12 beta in each industry. As in previous analyses, the portfolios are rebalanced annually at the end of June. The chosen approach offers diversification across industries whilst maintaining an equal industry composition of the quintile portfolios. It can therefore be interpreted as a variant of the best-in-class approach widely employed by ESG funds. In contrast, sorting portfolios merely based on exposure to the SCOPE12 factor would result in an unequal weighting of industries and contradict the investors' wish for a broad market exposure. Analogous to the first-pass regression in section 7.2, I have on average beta estimates of 1,146 stocks in my sample. 414 of these 1,146 companies lack explicit carbon intensity information provided by Datastream for the respective year, which equals more than one-third of the overall sample size. Hence, the investable universe is considerably extended through the consideration of SCOPE12 exposures in comparison to relying on reported carbon emissions alone. Table 16 shows the average industry compositions of the quintile portfolios. The percentage distribution across industries reported in the last column is similar to the complete sample presented in table 1. Consequently, the quintile portfolios offer a similar industry exposure as holding the market. As above, consumer discretionary and industrials constitute the largest industries and jointly account for nearly half of the sample.

To check for the resemblance of the exposure-sorted portfolios and the previously considered carbon intensity-sorted portfolios, I compare the inverse percentile ranks of the estimated exposures[19] with the actual carbon intensities. I find a significant positive Spearman rank correlation of approximately 0.3. This is comparable to the results of Hübel and Scholz (2020) who follow a similar approach for their environmental, social, and governance factor. Differences between the carbon intensities and exposures might arise from the additional industry sort, the change of a firm's carbon risk during the three year estimation period, and carbon risks potentially captured by beta but not by the reported emissions. For instance, sensitivities to SCOPE12 might reflect that certain firms got bad press for environmental reasons and therefore behave like high carbon stocks, which is not captured by the reported carbon intensities. This aspect also suggests that the usage of exposures could be more robust to greenwashing and strategic reporting of environmental risks (Hübel & Scholz, 2020).

8.2. Performance of Exposure Portfolios

Financial and risk considerations rank highly among investors' core motives to incorporate carbon risks in their investment process (Krueger et al., 2020). The risk-return profiles of the industry adjusted exposure portfolios for all stocks are reported in panel A of table 17. Panel B shows the exposure-sorted portfolios consisting of the above mentioned stocks without explicit emission information on Datastream. However, due to the smaller size of the unrated subsample these portfolios are *not* industry adjusted.

Both panels show a similar risk-return profile as the portfolios based on emission intensities in table 4. In both panels the excess return decreases monotonically from the cleanest to the dirtiest quintile. Additionally, I find that the extreme portfolios have the highest standard deviations. As investors are particularly concerned to reduce both portfolio and tail risks associated with carbon risk (Krueger et al., 2020), looking at exposures can identify particularly volatile and hence undesirable stocks. In both panels, the higher volatility in the clean portfolio results in a marginally smaller Sharpe ratio compared to the second portfolio. Although unrated stocks in panel B offer higher average excess returns in all quintiles, they are also more volatile. This results in Sharpe ratios of similar magnitude in both panels. To illustrate the performance of the portfolios over time, appendix E contains a plot of the cumulative returns of the portfolios from panel A and B in figure 3 and 4, respectively.

[19] As noted, high carbon stocks have low SCOPE12 betas.

Table 17: Risk and return of industry- and SCOPE12 exposure-sorted quintile portfolios from 07/2010 to 06/2020.

Panel A. Performance of quintile portfolios of all stocks.

	Clean	2	3	4	Dirty
Excess Return	0.85*	0.80*	0.67*	0.60	0.49
Standard deviation	5.17	4.53	4.44	4.72	5.52
t-statistic	1.81	1.92	1.65	1.39	0.98
Sharpe ratio	0.57	0.61	0.52	0.44	0.31

Panel B. Performance of quintile portfolios of unrated stocks.

	Clean	2	3	4	Dirty
Excess Return	1.01**	0.97**	0.84*	0.83*	0.55
Standard deviation	5.31	4.92	4.68	5.08	5.15
t-statistic	2.08	2.15	1.96	1.78	1.18
Sharpe ratio	0.66	0.68	0.62	0.56	0.37

Note. *p<0.1; **p<0.05; ***p<0.01; excess return is reported as monthly percentage; the Sharpe ratio is annualized.

Table 18: 4-factor and 5-factor model for industry- and SCOPE12 exposure-sorted quintile portfolios.

Panel A. 4-factor model.

Quintile	MKT	HML	SMB	MOM	SCOPE12	Alpha	Adj. R^2
Clean	1.093***	-0.217***	0.028	0.038		0.061	95.64
	(39.82)	(-4.27)	(0.55)	(0.92)		(0.58)	
2	0.970***	-0.242***	-0.233***	0.027		0.111	96.48
	(53.72)	(-6.08)	(-4.65)	(0.87)		(1.37)	
3	0.937***	-0.175***	-0.167**	0.048		0.003	94.80
	(31.39)	(-3.23)	(-2.60)	(0.95)		(0.03)	
4	0.996***	-0.195***	-0.129**	0.017		-0.086	95.95
	(51.39)	(-4.45)	(-2.34)	(0.42)		(-1.02)	
Dirty	1.103***	-0.047	0.188**	0.029		-0.252*	91.90
	(31.41)	(-0.72)	(2.28)	(0.44)		(-1.74)	

Panel B. 5-factor model.

Quintile	MKT	HML	SMB	MOM	SCOPE12	Alpha	Adj. R^2	ΔR^2
Clean	1.097***	-0.180***	0.019	0.035	0.152**	0.004	95.79	0.15**
	(40.86)	(-3.41)	(0.39)	(0.88)	(2.45)	(0.04)		
2	0.970***	-0.241***	-0.233***	0.027	0.005	0.109	96.45	-0.03
	(53.97)	(-5.66)	(-4.67)	(0.87)	(0.11)	(1.36)		
3	0.933***	-0.208***	-0.159**	0.051	-0.139**	0.055	94.96	0.16**
	(31.39)	(-3.58)	(-2.50)	(1.04)	(-2.55)	(0.50)		
4	0.989***	-0.263***	-0.112**	0.022	-0.283***	0.020	96.70	0.75***
	(53.33)	(-5.62)	(-2.54)	(0.68)	(-4.04)	(0.25)		
Dirty	1.086***	-0.198***	0.225***	0.041	-0.634***	-0.015	94.72	2.82***
	(38.32)	(-3.13)	(3.32)	(0.73)	(-9.00)	-0.12		

Note. *p<0.1, **p<0.05, ***p<0.01; t-statistics in parentheses; heteroskedasticity robust standard errors used. The last column reports the significance of heteroskedasticity consistent F-test for nested models.

Finally, I check whether the exposure-sorted portfolios generate abnormal returns and whether the SCOPE12 factor adds explanatory power. For this purpose, I regress the value-weighted and industry adjusted portfolio returns from panel A in table 17 on the 4-factor model and 5-factor model, respectively. Again, I report the results of the heteroskedasticity consistent F-test for nested models in the last column. The magnitude and the significance of the regression coefficients in table 18 is similar to the carbon intensity-sorted quintile portfolios in table 7. The results of the 4-factor regression in panel A indicate that all portfolios have market betas close to one and a tilt towards growth stocks. The latter, however, is not significant for the dirty quintile. While exposure to the size factor is mixed across the quintiles, none of the portfolios loads significantly on the MOM factor. With regard to abnormal returns, only for the dirty portfolio a weakly significant monthly alpha of -0.25% is found. This suggests that the portfolio of high carbon stocks underperforms after adjusting for common risk factors. One explanation for this finding could be that stocks in this quintile are overvalued relative to low carbon stocks and therefore offer lower returns. For example, investors perceive carbon intensive sectors such as the oil, the automotive, and the electric utilities industry to be overvalued considering climate change (Krueger et al., 2020). Although the presented quintiles have the same industry composition, the observed negative alpha in the dirty quintile might reflect the investors' perception of incorrect carbon risk pricing. Panel B shows that the SCOPE12 factor captures this abnormal return and leads to smaller alphas in most portfolios. In comparison to panel A, the other four regression coefficients remain largely unchanged. As expected, the additional SCOPE12 factor loadings decrease monotonically from the clean to the dirty portfolio. Comparable to table 7, the additional explanatory power is higher and more statistically significant for the more dirty portfolios. The change in adjusted R^2 is only non-significant for the second quintile.

Overall, the results indicate that the SCOPE12 factor adds explanatory power after adjusting the portfolios for a broad industry exposure. My findings also suggest that the SCOPE12 factor indeed captures the fraction of systematic risk attributable to priced carbon risk and therefore should be considered by investors to avoid negative financial outcomes. Lastly, using exposures rather than actual emission information seems to yield comparable risk-return results while extending the investment universe. These aspects emphasize the potential usefulness of the SCOPE12 factor for investors' carbon risk management.

9. Discussion

9.1. SCOPE12 as Genuine Risk Factor

The guiding research question of this thesis is to determine whether carbon risk is a systematic equity risk factor in European stocks. Previous research primarily uses time series regressions and the additional explanatory power of newly introduced factors as criterion for their relevance (e.g., Hübel & Scholz, 2020). However, publishing factors which merely add R^2 or reduce alpha relative to a benchmark model has proven to be too little of a hurdle in the factor model literature and resulted in "a zoo of new factors" (Cochrane, 2011, p. 1047). These supposed improvements can have several causes (e.g., data-snooping) and do not necessarily reflect progress in the identification of superior factor models (Campbell et al., 1997). Hence, recently new methodologies were proposed to identify systematic risk factors which should be considered in factor models (e.g., Fama & French, 2018; Klein & Chow, 2013; Pukthuanthong et al., 2019). Based on the combination of various methodologies for testing the carbon factor, SCOPE12 can be identified as a systematic risk factor.

First, the carbon factor is shown to be only weakly related to other accepted factors and explains returns when tested with both portfolios and individual stocks. The magnitude of the regression coefficients and additional R^2 is consistent with related research by Hübel and Scholz (2020) and Görgen et al. (2020). Similar to Görgen et al. (2020), I find that especially the variance in more carbon intensive portfolios – irrespective of whether actual ratings or carbon factor exposures are considered – is well captured by adding the SCOPE12 factor. This observation becomes particularly striking when factors are democratically orthogonalized. The unique R^2 from adding SCOPE12 is about 25 times higher for the most carbon intensive quintile than the clean one. Overall, this points to the existence of a large common variation of highly carbon intensive stocks due to the low carbon transition of the European economy.

The necessary condition for SCOPE12 to be considered as a risk factor is shown through the significance of the relation between SCOPE12's returns and the principal components of the sample covariance matrix (Pukthuanthong et al., 2019). This result is reliable because the ten extracted PCs in my analysis explain approximately 90% of the variance in stock returns and therefore provide a good proxy. Then I test for factor risk premiums using the methodology of Fama and MacBeth (1973) with single stock returns as dependent variable. There is evidence for the existence of a positive risk premium related to the carbon factor in the EIV corrected cross-sectional regressions. That means, investors are compensated for holding cleaner stocks. Although the chosen methodology is consistent with related literature, the result that only SCOPE12 offers a risk premium should be treated with certain caution. The treatment of EIV issues is complex and the t-statistics are shown to be sensitive to the beta sort chosen. Furthermore, alternative EIV correction approaches for cross-sectional regressions based on individual stocks and portfolios as dependent variable exist (e.g., Fama & MacBeth, 1973; Jegadeesh et al., 2019). Those methods, however, come with their own restrictions and tend to have data requirements exceeding the natural limitations of the presented sample. Apart from that, the observed weak significance of other regression coefficients than the intercepts is consistent with related papers (e.g., Görgen et al., 2020; Pukthuanthong et al., 2019). In contrast to my results, however, Görgen et al. (2020) do not find statistically significant evidence that carbon risk is priced in the cross-section of returns. A potential cause for this difference might be their consideration of additional return predictors as control variables. Lastly, SCOPE12 offers a reward-to-risk ratio that appears to be consistent with risk pricing limits suggested by the finance literature (e.g., MacKinlay, 1995).

As a genuine risk factor, SCOPE12 also becomes interesting for investors who seek to manage the carbon risk of their holdings even in the absence of explicit emission information. The factor construction through double sorting based on size and carbon intensities allows to extract the unique portion of return caused by emission differences irrespective of potential size effects. Considering that stocks without ESG information tend to be significantly smaller on average (e.g., Hübel & Scholz, 2020), this feature appears particularly desirable for beta estimation. Measuring stocks' exposure to the SCOPE12 factor through time series regression therefore helps socially responsible investors in two ways. First, it allows them to considerably extend their universe of investable equities (i.e., small caps), and secondly, they can identify assets to be included in carbon risk management. With regard to risk, I find that the most extreme quintiles also have the highest standard deviations. This makes these portfolios also interesting from a risk management perspective as investors are particularly concerned about both portfolio and tail risks related to climate change (Krueger et al., 2020). Regarding the dirty exposure portfolio, I find weakly significant evidence that the Carhart 4-factor model cannot sufficiently explain its returns. The identified negative 4-factor alpha, however, can be attributed to the SCOPE12 factor and emphasizes the carbon factors relevance. Overall, my analyses suggest that investors can achieve similar results by using a return-based exposure measure instead of directly reported carbon intensities.

9.2. Outperformance of Low Carbon Stocks

The second research question addresses the direction and magnitude of the price effect of carbon risk and relates to a strand of literature finding mixed results. In the European sample less carbon intensive portfolios provide higher returns than more carbon intensive portfolios. This finding supports the clean alpha hypothesis mentioned in the introduction. Sorting stocks into both carbon intensity and SCOPE12 exposure portfolios, I find a monotonous decline of excess returns from the clean to the dirty portfolio in all my analyses. Especially the cleaner portfolios generate excess returns significantly different from zero in the last decade, whereas the carbon intensive portfolios do not. The observed relationship between excess returns and carbon risk on the portfolio level is also consistent with SCOPE12's positive risk premium found in the cross-sectional regressions.

Moreover, I generally do not find significant abnormal returns for the mentioned long-only quintile portfolios based on the Carhart 4-factor model. As stated above, only the dirty exposure portfolio has a weakly significant negative alpha relative to this model and provides some indication for a risk adjusted underperformance of high carbon stocks. In contrast, long-short strategies seem to generate returns which cannot be explained by the common risk factors. In the time from July 2007 to June 2020, the outperformance of cleaner assets results in a highly significant monthly average return of 0.32% on the SCOPE12 factor, which is constructed as the return on a size adjusted hedge portfolio going long (short) in less (more) carbon intensive firms. As shown in the spanning regressions, this return cannot be explained by the 4-factor model and generates a significant monthly alpha of 0.37% during the main sample period. This result is consistent with Pedersen et al. (2021) and In et al. (2019) who both find significant average excess returns on carbon intensity-sorted hedge portfolios for their US samples. Moreover, both studies report (weakly) significant positive alphas when regressing the hedge portfolio returns on standard factor models.

While the average returns are higher on less carbon intensive stocks during my sample period, the outperformance is not constant over time. Rather, the plot of the portfolios' market adjusted cumulative returns indicates that cleaner stocks started to outperform their higher emission peers in mid 2008. The performance differential between the clean and the dirty portfolio gradually expanded further during the main sample period. Interestingly, the onset of the outperformance of clean stocks corresponds closely to the period in which most of the spikes in the climate change news index by Engle et al. (2020) occur. If the observed outperformance is driven by unexpected changes in tastes for greener assets due to higher climate risk awareness, the observed pattern could be consistent with the ESG factor proposed by Pástor et al. (2021). That is, Pástor et al. (2021, p. 8) state that "if one computes average returns over a sample period when ESG concerns strengthen more than investors expected,(. . .) then green stocks outperform brown stocks, contrary to what is expected". Choi et al. (2020) find global evidence for such effects. They show that greener assets perform better when awareness for climate change is high during times of abnormally hot weather. Similarly, Pedersen et al. (2021) find higher institutional holdings and valuations for less carbon intensive assets based on their US sample. Lastly, institutional investors perceive certain carbon intensive sectors to be overvalued, as prices do not correctly reflect climate change risks yet. Vice versa, they see potentially cleaner sectors (e.g., battery or renewable energy producers) as undervalued (Krueger et al., 2020). A promising avenue for future research would therefore be to test whether the observed pattern during my sample period is due to a learning period in which the prices of stocks with less (more) carbon intensity are adjusted upwards (downwards). If the sample indeed reflects an adjustment period, one would expect a reversal in the performance pattern in a longer time series. Such effects would then be consistent with the dirty alpha hypothesis.

9.3. Limitations

Considering the complexity of capturing the extent of GHG emissions, current climate finance research is limited by the availability of accurate proxies. Although scope 1 and 2 emissions tend to be the most reliable and consistent measures of carbon emission across data providers, they only provide an incomplete picture of actual carbon emissions. Relevant additional information such as a firm's other indirect emissions along the value chain, scope 3 emissions, "are rarely reported by companies and are at best noisily estimated and inconsistent across different data providers" (Pedersen et al., 2021, pp. 12-13). Especially the option to choose between different organizational boundaries for carbon accounting bears the potential that the true emissions caused by a company are misrepresented. Moreover, the current accounting standards do not properly reflect the business model of financial firms (Görgen et al., 2020). The decision to exclude financials based on this observation, nevertheless, poses a main limitation of my research as it restricts the generalizability of the presented results.

In addition to this, Campbell et al. (1997) point out the fundamental factor literature's proneness to data-snooping issues. Fundamental factors are usually constructed based on characteristics which are empirically found to be relevant in explaining the cross-section of returns. Hence, the relevance of a particular measure might be overstated and not indicative for future periods. Carbon intensity measures – as the one used in this thesis – are widely used in the climate finance literature and are frequently found to positively relate to accounting and market performance (e.g., Pedersen et al., 2021). Moreover, relative emission measures tend to be more likely to yield statistically significant results (Busch & Lewandowski, 2018). Although the presented factor has a sound theoretical motivation and therefore the likelihood of data dredging is reduced (Campbell et al., 1997), such objections usually can only be overcome with very long time series.

With the ongoing development of the carbon accounting standards and reporting best practices, repeating the presented analyses with more comprehensive measures of carbon emission and a longer time series appears promising. Doing so would equally address both key limitations of my thesis.

10. Conclusion

Large-scale reforms of the European financial system initiated by the EU and additional commitments by institutional investors shift the capital market's focus onto climate change considerations. Especially the aspect of climate change mitigation through carbon emission reduction has entered mainstream financial decision making. This heightened awareness in the past decade bears the potential to systematically influence the cross-section of asset prices as global warming affects the entire economy, not just individual industries. In this context, the question of carbon risk's impact on asset prices generated some interest in the nascent climate finance literature. However, the currently available mixed evidence is primarily restricted to US or global samples. This thesis contributes to the literature by quantifying the carbon risk in European equity returns, determining whether it constitutes a systematic risk factor, and highlighting the practical application of the carbon factor for risk management.

Using carbon intensity as a firm characteristic, I calculate the SCOPE12 carbon factor. The factor is constructed as a double-sorted hedge portfolio which is long in cleaner stocks and short in dirty stocks. It provides a significantly positive average return during the sample period and the onset of its positive performance corresponds closely with heightened awareness for climate change in the media. In combination with the observed monotonous negative relationship between return and both carbon intensity and SCOPE12 exposure across quintile portfolios, this provides evidence for the clean alpha hypothesis.

I also find that SCOPE12 can be considered as a genuine risk factor because it is related to the sample covariance matrix of returns, commands a (weakly) significant positive risk premium, and offers a risk-reward trade-off within reasonable limits. This finding is corroborated by complementary analyses which confirm that the carbon factor cannot be explained by alternative factor candidates. SCOPE12 therefore explains a unique portion of systematic variation in European stock returns.

Lastly, current research emphasizes that investors want to manage carbon risks in their portfolio but lack appropriate tools and best practices. Especially the insufficient disclosure of GHG emissions poses a serious restriction for the identification of carbon risks. Using return-based SCOPE12 factor exposures instead addresses this issue and considerably expands the universe of stocks which can be considered for investment and risk management. With the gradually growing importance of climate change considerations in financial markets and institutional investors' awareness for its financial implications, strategically managing carbon risk with the SCOPE12 factor could therefore bear positive effects.

References

Aiken, J. G., Erdos, J. A., & Goldstein, J. A. (1980). On Löwdin Orthogonalization. *International Journal of Quantum Chemistry*, *18*, 1101–1108.

Alexander, C. (2001). *Market Models: A Guide To Financial Data Analysis*. Hoboken, New Jersey: Wiley.

Blitz, D., Hanauer, M. X., Vidojevic, M., & van Vliet, P. (2018). Five Concerns with the Five-Factor Model. *Journal of Portfolio Management*, *44*(4), 71–78.

Bodie, Z., Kane, A., & Marcus, A. J. (2018). *Investments* (11th ed.). New York: McGraw-Hill Education.

Bolton, P., & Kacperczyk, M. (2020). Do Investors Care about Carbon Risk? *Journal of Financial Economics (forthcoming)*.

Brohé, A. (2017). *The Handbook of Carbon Accounting*. New York: Routledge.

Busch, T. (2011). Which emissions do we need to account for in corporate carbon performance? Response to Murray and colleagues. *Journal of Industrial Ecology*, *15*(1), 160–163.

Busch, T., Johnson, M., Pioch, T., & Kopp, M. (2018). Consistency of Corporate Carbon Emission Data. *University of Hamburg / WWF Deutschland, Hamburg*.

Busch, T., & Lewandowski, S. (2018). Corporate Carbon and Financial Performance: A Meta-analysis. *Journal of Industrial Ecology*, *22*(4), 745–759.

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton, New Jersey: Princeton University Press.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, *52*(1), 57–82.

Choi, D., Gao, Z., & Jiang, W. (2020). Attention to Global Warming. *The Review of Financial Studies*, *33*(3), 1112–1144.

Cochrane, J. H. (2005). *Asset Pricing* (2nd ed.). Princeton, New Jersey: Princeton University Press.

Cochrane, J. H. (2011). Presidential Address: Discount Rates. *The Journal of Finance*, *66*(4), 1047-1108.

Connor, G., & Korajczyk, R. A. (1986). Performance Measurement with the Arbitrage Pricing Theory: A New Framework For Analysis. *Journal of Financial Economics*, *15*, 373–394.

Connor, G., & Korajczyk, R. A. (1988). Risk and Return in an equilibrium APT: Application of a New Test Methodology. *Journal of Financial Economics*, *21*, 255–289.

Deutsche Börse Group. (2021). *STOXX® Index Methodology Guide (Portfolio Based Indices)*. Retrieved from https://www.stoxx.com/document/Indices/Common/Indexguide/stoxx_index_guide.pdf

Engle, R. F., Giglio, S., Kelly, B., Lee, H., & Stroebel, J. (2020). Hedging Climate Change News. *The Review of Financial Studies*, *33*(3), 1184–1216.

European Commission. (2021). *Overview of sustainable finance.* Retrieved from https://ec.europa.eu/info/business-economy-euro/banking-and-finance/sustainable-finance/overview-sustainable-finance_en

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*, 3–56.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*, 1–22.

Fama, E. F., & French, K. R. (2018). Choosing factors. *Journal of Financial Economics*, *128*, 234–252.

Fama, E. F., & MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, *81*(3), 607–636.

French, K. R. (2020). *Current Research Returns.* Retrieved from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#International

Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A Test of the Efficiency of a Given Portfolio. *Econometrica*, *57*(5), 1121–1152.

Görgen, M., Jacob, A., Nerlinger, M., Riordan, R., Rohleder, M., & Wilkens, M. (2020). Carbon Risk. *Working Paper Augsburg University*.

Hübel, B., & Scholz, H. (2020). Integrating sustainability risks in asset management: the role of ESG exposures and ESG ratings. *Journal of Asset Management*, *21*, 52–69.

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

IIGCC. (2021). *About us.* Retrieved from https://www.iigcc.org/about-us/

Ilhan, E., Sautner, Z., & Vilkov, G. (2020). Carbon Tail Risk. *The Review of Financial Studies (forthcoming)*.

In, S. Y., Park, K. Y., & Monk, A. (2019). Is 'Being Green' Rewarded in the Market?: An Empirical Investigation of Decarbonization and Stock Returns. *Stanford Global Project Center Working Paper*.

Ince, O. S., & Porter, R. B. (2006). Individual Equity Return Data from Thomson Datastream: Handle With Care! *The Journal of Financial Research*, *29*(4), 463–479.

IPCC. (2018). *Intergovernmental Panel on Climate Change. Special report: Global warming of 1.5°C - Summary for Policymakers.* Retrieved from https://report.ipcc.ch/sr15/pdf/sr15_spm_final.pdf

Jegadeesh, N., Noh, J., Pukthuanthong, K., Roll, R., & Wang, J. (2019). Empirical tests of asset pricing models with individual assets: Resolving the errors-in-variables bias in risk premium estimation. *Journal of Financial Economics*, *133*(2), 273 – 298.

Kaiser, L. (2020). ESG Integration: value, growth and momentum. *Journal of Asset Management*, *21*, 32–51.

Kennedy, P. (2008). *A Guide to Econometrics* (6th ed.). Malden, MA: Wiley-Blackwell.

Klein, R. F., & Chow, V. K. (2013). Orthogonalized factors and systematic risk decomposition. *The Quarterly Review of Economics and Finance*, *53*, 175–187.

Krueger, P., Sautner, Z., & Starks, L. T. (2020). The Importance of Climate Risks for Institutional Investors. *The Review of Financial Studies*, *33*(3), 1067–1111.

MacKinlay, A. C. (1995). Multifactor models do not explain deviations from the CAPM. *Journal of Financial Economics*, *38*(1), 3 – 28.

Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics (forthcoming)*.

Pástor, L., Stambaugh, R. F., & Taylor, L. A. (2021). Sustainable Investing in Equilibrium. *Journal of Financial Economics (forthcoming)*.

Pukthuanthong, K., Roll, R., & Subrahmanyam, A. (2019). A Protocol for Factor Identification. *The Review of Financial Studies*, *32*(4), 1573–1607.

Schmidt, P. S., von Arx, U., Schrimpf, A., Wagner, A. F., & Ziegler, A. (2019). Common risk factors in international stock markets. *Financial Markets and Portfolio Management*, *33*(3), 213–241.

Stechemesser, K., & Guenther, E. (2012). Carbon accounting: a systematic literature review. *Journal of Cleaner Production*, *36*, 17–36.

Tsay, R. S. (2010). *Analysis of Financial Time Series* (3rd ed.). Hoboken, New Jersey: Wiley.

Tsay, R. S. (2014). *Multivariate Time Series Analysis - With R and Financial Applications*. Hoboken, New Jersey: Wiley.

UNFCCC. (1992). United Nations Framework Convention on Climate Change. New York.

UNFCCC. (2008). *Kyoto Protocol Reference Manual - On Accounting of Emissions and Assigned Amount.* Retrieved from https://unfccc.int/resource/docs/publications/08_unfccc_kp_ref_manual.pdf

United States Environmental Protection Agency. (2021). *Understanding Global Warming Potentials.* Retrieved from https://www.epa.gov/ghgemissions/understanding-global-warming-potentials

WBCSD, & WRI. (2004). *The Greenhouse Gas Protocol - A Corporate Accounting and Reporting Standard (Revised Edition).* Retrieved from https://ghgprotocol.org/sites/default/files/standards/ghg-protocol-revised.pdf

WBCSD, & WRI. (2013). *Required Greenhouse Gases in Inventories - Accounting and Reporting Standard Amendment.* Retrieved from https://ghgprotocol.org/sites/default/files/standards_supporting/RequiredgasesandGWPvalues_1.pdf

Wooldridge, J. M. (2016). *Introductory Econometrics - A Modern Approach* (6th ed.). Boston: Cengage Learning.

Zivot, E., & Wang, J. (2003). *Modeling Financial Time Series with S-PLUS*. New York: Springer.