UNIVERSITY OF GOTHENBURG

SCHOOL OF BUSINESS, ECONOMICS AND LAW

Bachelor's Thesis within the Bachelor's Programme in Financial Economics
15 credits

The Role of Environmental Factors in Explaining Return Variation in Equity Markets

Authors:
Alexander Nyman Sjöström
Erik Von der Lancken

Supervisor:
Henrik Petri
Spring 2025

Abstract

This study investigates whether environmental factors help explain return variation in equity markets, and if so, whether this relationship reflects a climate-related risk premium or investor preferences. We use the Refinitive ESG database (2025) to construct our own environmental score, which is then incorporated in a Green Minus Brown (GMB) factor. The time span for our study is 2015 to 2024. Our findings indicate that there is a varying time effect: prior to 2020, environmental performance had limited influence on returns, while after 2020, the GMB factor provided meaningful explanatory power for several portfolios. This suggests a growing market relevance for environmental characteristics in recent years. We also observed that green stocks tended to be most large-cap while brown firms were often small-cap. Our study contributes to the growing literature by showing that the explanatory power of environmental performance has grown over the years. However, our Fama–MacBeth regression results indicate that the GMB factor is not a priced risk factor in a cross-sectional sense. Further research is needed to assess whether these patterns hold across other markets and longer time periods.

Keywords: Asset pricing, investor preference, carbon risk premium, Green Minus Brown (GMB factor), Fama French model

Acknowledgement

We would like to thank our supervisor, Henrik Petri, for his insights and feedback throughout the thesis process. We would also express gratitude to our opponents, whose constructive comments helped us improve the quality and clarity of our work.

1. Introduction *4*
1.1Background 4
1.2 Problem and Aim 5
1.3 Contribution 6
1.4 Outline 7
2. Theoretical Framework *8*
2.1 Traditional Asset Pricing Models and Risk Factors 8
2.2 Environmental Factors: Risk Premium vs. Investor Preference 9
2.3 Time Variation in Environmental pricing 12
3. Literature Review *13*
4. Method *17*
4.1 Data Collection and Sample Preparation 17
4.2 Development of a Custom Environmental Scoring System 17
4.3 Justification for a Custom Environmental Score vs. Commercial Ratings 22
4.4 Assumption of Historical Environmental Profile Stability 23
4.5 Portfolio Construction: Green and Brown Portfolios 24
4.6 Regression Model, Additional Analysis and Robustness Checks 27
5. Empirical Results *30*
5.1 Base vs. extended model: Overall fit and gains 30
5.2 GMB Factor coefficients and significance across portfolios 33
5.3 Subperiod analysis: Pre-2020 vs. post-2020: 39
5.4 Robustness checks and additional results 41
5.5 Fama – MacBeth cross sectional regression (GMB Risk Premium) 42
5.6 Rolling 36 – month beta analysis (Time varying GMB exposure) 43
6. Discussion *47*
6.1 Results 47
6.2 Limitations 52
7. Conclusion *54*

1. Introduction

1.1Background

Sustainable investing has rapidly become one of the hottest topics in finance. It is projected that global assets managed with environmental, social and governance (ESG criteria) will surpass 50 trillion by 2025 (*The Bright Spots in a Complicated ESG Framework*, 2022), highlighting an increasing investor emphasis on climate change and social concerns. Importantly, many investors seek "green" investments that resonate with their ethical values without sacrificing financial returns. A substantial amount of research of over 2000 studies, shows that the incorporation of ESG factors typically does not harm performance; approximately 90% of these studies find a non-negative relationship between ESG and corporate financial results, where a significant number indicate positive impacts (Friede et al., 2015). This evidence suggests that, historically, firms with strong ESG profiles have performed at least as well as others, supporting the statement that ethical investing can be compatible with competitive returns. Yet, it remains unclear whether markets truly price environmental risk, or if observed returns come from investor sentiment and temporary trends.

At the same time, the rise of ESG has created a debate about whether environmental factors are priced in equity markets. Specifically, the question arises: do stocks' environmental characteristics (such as carbon emissions or sustainability scores) systematically influence their expected returns? Some analysts argue that firms with favourable environmental profiles, often referred to as "green" companies, tend to outperform due to effective risk management or increased investor interest. Over the past decade, there have been times where "green stocks" have surpassed the performance of their "brown" (high-carbon) counterparts. For example, starting around 2010, shares of low-carbon enterprises yielded greater returns compared to those of high emission companies (Görgen et al, 2020). This outperformance is often linked to a growing investor preference for sustainable investments. Such demand-driven price effects suggest that green stocks may become overvalued in relation to their underlying fundamentals, potentially leading to diminished future returns. On the other hand, the stocks of "brown" firms,

which some investors may avoid, could be undervalued and offer higher expected returns in the future, either as a risk premium for climate-related concerns or simply because they are less in favour. Despite the increasing importance of environmental factor and a surge in sustainable investments, researchers are still divided on whether environmental factors are priced due to risk or investor preferences.

1.2 Problem and Aim

Despite a growing body of research, existing evidence on whether environmental performance is priced in equity markets remains inconclusive. Some studies, such as Bolton and Kacperczyk (2020), suggest that firms earn higher returns because of a carbon-related risk premium. Conversely, other studies argue that green firms outperform due to shifts in investor preferences rather than exposure to fundamental risks (Pástor, Stambaugh, & Taylor, 2022)

These conflicting perspectives create uncertainty about whether green or brown stocks outperform, when adjusting for risk. Furthermore, it remains debated whether investor preference or a carbon risk premium is the dominant driving force behind this performance. Additional complexities, such as limited historical data, changing report standards, and evolving investor behavior, further complicate the relationship between environmental performance and expected returns. As awareness of climate-related risks grows, understanding whether these factors are priced becomes increasingly relevant for both asset pricing theory and portfolio management.

Building on the mixed empirical findings presented above, this thesis investigates whether firms' environmental performance is systematically reflected in their stock returns. To achieve the purpose of this study, we address the following research question:

- *Do differences in environmental performance explain variation in expected stock returns, and if so, are they associated by priced risk or investor preferences?*

We approach this question by constructing a Green Minus Brown (GMB) portfolio based on a composite environmental score that captures both positive and negative sustainability indicators.

We test the factor's explanatory power within a Fama-French-Carhart multifactor framework, using the time-series regressions on diversified U.S equity portfolios.

1.3 Contribution

This thesis contributes to the growing literature on climate finance and asset pricing in several ways. First, we extend the work of Görgen et al. (2020) by examining a more recent time period (2015-2024), allowing us to capture structural changes and environmental behaviour over today rather than a few years ago. Second, while prior studies often either rely on single indicators such as carbon emissions or ESG factors that rely heavily on Social & Governance, our study focuses only on the Environmental aspect. Our study develops a Green Minus Brown (GMB) portfolio based on a comprehensive set of 55 environmental indicators. This broader approach that focuses solely on environmental scores provides a more holistic measure of environmental performance, rather than ESG. Third we test whether this environmental factor is priced in equity markets by using the Fama-French-Carhart multifactor model. By doing so, we investigate whether environmental factors have the characteristics of a systematic risk factor, or whether observed return differences are better explained by investor preference.

Overall, the thesis provides new insights on how environmental factors relate to return generation and whether a distinct environmental risk factor can be justified in modern asset pricing. However, we are aware that our study has certain limitations which are important to interpret alongside our contribution. While we test whether environmental performance is priced as either risk or preference, this could exclude other perspectives. For instance, ESG characteristics may also reflect reputational signaling, data biases, or shifting regulatory environments rather than a systematic risk. Moreover, the construction of the environmental scores could involve subjective design choices. Also, by focusing solely on US equities our results may not generalise across geographics or sectors. Therefore our findings should be understood in light of the specific theories we apply and the data choices we have made.

1.4 Outline

The structure of the thesis will be as follows: Section 2 presents the theoretical framework, introducing traditional asset pricing models and risk factors and theories regarding risk premia and investor preference. Section 3 introducing existing literature on environmental factors and stock returns. Section 4 describes the method, such as data selection, scoring system and regression models. Section 5 presents the empirical results of the regressions and robustness tests. Section 6 discusses the interpretation of the results in relation to theory and literature. Finally, section 7 concludes the study, by summarizing key findings, limitations and suggestions for future research.

2. Theoretical Framework

2.1 Traditional Asset Pricing Models and Risk Factors

Classic finance theory assumes that investors are rewarded only for bearing systematic risks, those broad market factors that cannot be eliminated through diversification. The Capital Asset Pricing Model (CAPM) encapsulates this concept by linking a security's anticipated return to its sensitivity to the market portfolio, usually referred to as market beta. CAPM was developed in the 1960s by William Sharpe, Jack Treynor, John Lintner, and Jan Mossin. It is inspired by ideas from Harry Markovitz in the 1950s. According to the model the market factor is the only source of risk premium. This model predicts that assets having higher covariance with market returns will generate higher returns to compensate for the extra volatility that investors have to deal with. Empirical research has often demonstrated that market beta, only, does not explain the variations in stock returns among different securities well. As a result, multifactor models were created to account for this. Fama and French (1993) came up with a three-factor model that includes the market component together with two other sources of risk factors: firm size and value. They saw that there were both size and value premiums since small stocks and "value" companies (those with high book-to-market ratios) performed better than large-cap and growth stocks, respectively. Later, by including investment and profitability elements, Fama and French (2014) expanded the model into a five-factor framework. The introduction of these new variables significantly improved the explanatory capacity of asset pricing models and showed that returns are affected by multiple sources of risk, not only just one.

The Arbitrage Pricing Theory (APT) gives a more general theoretical basis for multifactor models. According to the model, returns on assets can be predicted by using the linear relationship between the expected return and several different variables that capture systematic risk. When a factor cannot be diversified away, the arbitrage theory argues that it should consist of a risk premium. Therefore any exposure to these risks that investors cannot eliminate through diversification, will lead to increases in expected returns. Unlike the single-factor Capital Asset Pricing Model (CAPM), APT accounts for multiple risk factors that can affect asset prices. The market also does not need to be efficient, only that there are no straightforward arbitrage

opportunities (Adam, 2024). If one of these underlying factors can be viewed as an "environmental risk factor", for example one that correlated with changes in climate policy or fluctuations in carbon prices, it would support the idea that these risks explain variations in returns. We can therefore hypothesize that a "low-carbon" factor could yield a positive premium (indicating higher returns for carbon-intensive firms) if the explanatory power of the model will increase by adding the environmental factor. We developed the following hypothesis:

H_1: Carbon Risk Premium Hypothesis

Firms with higher environmental risk exposure (brown stocks) earn higher average returns than firms with low environmental risk (green stocks), reflecting compensation for systematic clime-related risks.

2.2 Environmental Factors: Risk Premium vs. Investor Preference

An important question for the study will be whether the variations in returns associated with environmental factors is a result of true risk premia, or if they can be linked to investor preferences or mispricing. If companies with poor environmental practices ("Brown" firms) face greater future risk because of greater costs that could happen in the future, such as carbon taxes, environmental regulations or damages from climate change, then these firms carry greater fundamental risk. As a result, investors would require higher expected returns for investing in such stocks to compensate for the climate risks that they possess. From this point of view, environmental risk could be compared with other risk factors because exposure of carbon-intensive firms should result in a return premium. This is because firm's cash flows will be more exposed to disruptions from climate and policy changes. In other words, if carbon emission or environmental risks contribute with systematic risk that cannot be diversified, the risk will be reflected in the prices, as a return premium for brown assets. Therefore, one could anticipate that brown stocks on average would perform better than green stocks, adjusted for other risk factors.

In contrast, preferences (or mispricing) could be a sign that variations may come from investors' non-financial motivations or market inefficiencies rather than from compensatory risk. Some investors could find value in holding "green" assets, or might be under institutional pressure to

avoid “brown” assets, regardless of the financial returns. These investors are willing to sacrifice additional return in exchange for assets that align with their values. As a result, green stocks could be inflated in value (driven up in price), while brown stocks may be undervalued. This can lead to green assets exhibiting lower expected returns, not due to risk factors but because of increased demand. For example Pástor , Stambaugh, and Taylor (2022) equilibrium models show that green firms experience lower cost of capital, while brown assets face higher cost of capital. This difference arises because brown firms are mainly held by investors who prioritize returns. Moreover, the models suggest that shifts in investor preference can result in temporary variations in returns. For instance, if investors suddenly shift their demand to green assets, maybe because of increased awareness of climate issues, this will lead to a rise in green asset prices, providing one time gain for current investors. This can lead to periods where green assets outperform brown stocks even if in the long run brown stocks have higher long-run expected returns. Recent studies that have shown that green assets outperform brown ones, for example Görgen et al (2020) study, can be interpreted within this context. A sudden surge in environmental awareness among investors can drive up green stock prices, leading to short-term excess returns that do not reflect a long-term risk premium. This will lead to our second hypothesis:

H_2: Investor Preference Hypothesis

Green stocks earn lower expected returns than brown stocks due to investor’s non-pecuniary preferences, leading to higher valuations and a lower cost of capital for evironmentally friendly firms.

There are studies that have tried to explain these preference-based explanations, especially for those so-called “sins stocks”. Investors who prioritize social responsibility often refuse to invest in companies related to alcohol, tobacco, gambling, and other unethical practices. Hong and Kacperczyk (2009), concluded that these sin stocks on average yielded approximately 2.5% higher return than similar stocks, even after controlling for standard risk factors. The authors link this higher return as a result of “normal” investors that tend to avoid these stocks, which results in a lower price driven by investor aversion. A similar study, made by Heinkel, Kraus, and Zechner (2001) showed that if a significant portion of investors exclude companies from their portfolios, the stock prices will decrease until the expected returns are sufficiently attractive to

investors that don't have constraints. This suggests that as long as “brown firms”, in equilibrium, remain unpopular, they will consist of a risk premium to those investors that are willing to invest in them. This “exclusion effect” can offer a justification for a carbon premium that is a result of investor preferences and constraints rather than by fundamental risk. In this point of view, it can be interpreted that brown stocks may achieve higher returns specifically because certain investors choose not to invest in them, reflecting similar patterns with “sin stocks”.

In summary, there are two competing forces. If environmental factors represent a threat to their cash flows, asset pricing theory predicts a risk premium for brown stocks. On the other hand, if the preferences of investors for sustainability dominate, green assets could be considered as “overvalued”, while brown assets could be seen as “undervalued”, resulting in elevated valuations and diminishing returns for green assets and lower valuations and potentially higher future returns for brown stocks. These two forces are also not mutually exclusive; both risk and investor preferences can affect asset prices at the same time, and their significance can change over time. Our study will investigate which of these mechanisms that more accurately will show up in the data. If brown firms yield a better return, this would support the risk premium (or exclusion) hypothesis.
Conversely, if there are no return differences or a green premium is observed, it could mean that investor demand has fully adjusted prices in accordance with sustainability preferences.
Therefore, if there is a stronger systematic return linked to environmental performance, then an environmental risk factor could exist. If there is absence of such a factor it would imply that the observed return variations are likely temporary or influenced by other variables. Under the risk premium hypothesis, we expect brown stocks to yield higher average returns (a carbon premium) after controlling for other factors. Under the preference/mispricing hypothesis, any return differential could favor green stocks or be null, due to prices already reflecting sustainability preferences.

2.3 Time Variation in Environmental pricing

In recent years, the relevance of environmental performance in asset pricing has become increasingly time-dependent. Regulatory initiatives, growing public awareness of sustainability, and institutional shifts (such as ESG mandates) have all contributed to a dynamic pricing environment.

Research by Bolton and Kacperczyk (2021) show that there was no significant carbon premium during the 1990s, however they found a highly significant carbon premium during their sample period (2005- 2007). This study suggests that environmental factors may not have been priced consistently over time, but that their impacts intensified in later periods.

This temporal dynamic raises the possibility that the pricing of environmental risks and preferences is time-varying, and that recent years reflect a structural change in how markets value sustainability. This leads us to the following hypothesis:

H_3: Time-Variation Hypothesis

The pricing impact of environmental factors has increased over time. Specifically, the GMB return differential is significantly greater in the post-2020 period than in the pre-2020 period, reflecting changing investor behavior, increased climate awareness, and regulatory focus.

3. Literature Review

The research on the financial impacts of environmental factors has been growing rapidly in recent years. There is some early evidence on how non-financial considerations impact returns. One of these is Hong and Kacperczyk (2009) who studied so-called "sin stocks". They found that stocks that engage in unethical practices had a persistent return premium and outperformed otherwise similar stocks by approximately 2.5% per year. This supports the belief that investor exclusions can lead to mispricing that some investors take advantage of. It also leads to a discussion whether environmentally "brown" stocks (for example heavy carbon emitters) might get higher returns because some investors avoid them.

Heinkel et al. (2001) showed that since polluting firms are being held by fewer investors because green investors avoid them, it leads to lower stock prices and thus a higher cost of capital for these types of stocks. These studies laid the groundwork for the expectation of a "carbon premium", similar to the sin stock premium, that comes from exclusionary behaviour from investors. However, these models assume that both investors and firms are strictly categorized as either green or polluting, which oversimplifies the complexity of ESG investment landscape where preferences and sustainability levels vary.

Many empirical studies have lately examined whether carbon emissions and other environmental factors could be reflected in stock returns. One of these is a paper from Bolton and Kacperczyk (2020) that studied how carbon emissions could affect the cross-section of US stock returns. They concluded that companies with higher total carbon dioxide emissions, when controlling for size, book-to-market and other factors, earn higher returns. This reinforces the risk premium hypothesis, though the robustness of this effect during different market conditions remains debated.

Notably, they also find that investors exclude high carbon companies but unlike for "sin" stocks, limited risk sharing could not alone explain the return premiums for these carbon-stocks. Their result, however, was consistent, that investors require compensation for holding carbon-intensive assets. This supports the idea that excluding certain stocks can create a return premium for those

stocks, a similarity that could apply to “brown” stocks. A question that still remains unanswered is whether a carbon premium exists over time or diminishes as more investors incorporate climate risks. We aim to address this gap by testing whether such a premium exists using data from 2015-2024 and a broader environmental score.

Engle et al. (2019) examined how news about climate change affects stock returns through analysis of newspapers. They built a “climate news index” based on the frequency of climate related news and developed hedge portfolios to isolate exposure to this news factor. By going long or short based on the stocks performance during climate news, they showed that it is possible to both in sample and out of sample successfully hedge innovations in climate change news. This indicated that climate news is a priced source of risk, that news about climate move markets, and assets react differently to that risk. This result is important due to two reasons, first it proves that climate risk clearly affects returns, which in their study was an ESG-related factor that is similar to a risk factor. On the other hand, if climate risk can be hedged through diversification, it raises the question whether climate risk is a systematic risk and investors should demand a premium for it. If an investor can eliminate a climate risk from their portfolio at a low cost, then investors will not demand a large premium for it, since it can be avoided. The authors write that if climate risk can be hedged and many investors apply these investment strategies, then it will increase the price of those firms and expected return will be lower. While innovative, the reliance on news frequency may capture investor sentiment rather than fundamental risk factor.

Some studies have construct factors to test their explanatory power for environmental factors. One of these studies is made by Görgen et al. (2020), they develop a BMG factor to capture carbon risk in stock returns. They used a sample of 1600 firms between 2010 and 2017 and calculated each firm’s carbon beta, which is their sensitivity to the returns of a brown-minus-green factor portfolio. Görgen and the others find that the carbon factor significantly improves the explanatory power of the model. The carbon factor’s contribution to the explaining power was even comparable to other common risk factors, such as size, showing that carbon risk carries a substantial weight in explaining variation in returns. Moreover, the firms that improved their footprints and became greener tended to outperform those who went in the opposite direction and

became “browner”. Thus, reducing carbon risk appeared to be rewarded by the market. These findings show that investors pay attention to carbon exposures and can explain why some stocks earn higher or lower returns. However, one aspect of Görgen et al. (2020) study that is noteworthy is that during the sample period, the authors draw the conclusion that there was no clear carbon risk premium in that decade. Instead, they argue that it could be evidence that investors are unable to predict and quantify cabron risks and also because green firms are becoming green more faster than brown firms, making green firms outperform brown ones. The market seems to be pricing climate concerns rather than demanding higher returns for brown risk.

Pástor, Stambaugh, and Taylor (2022) offers an explanation for these events. They argue that green assets can have lower expected returns but still outperform brown assets when investors shift their preferences . They saw that green stocks outperformed brown stocks in the 2010s, with cumulative return difference of 175% over their 8.2-year sample period (2020-2020). Their model predicts that when investors care more about sustainability they will bid up green assets. As the preferences are priced, green stocks should in the long run yield lower returns than brown stocks. Pastor et al. claim that the strong performance of green stocks lately, are driven by large numbers of investors that shift their preferences toward green stocks, and are willing to pay more for it. This study is also evident to Bauer et al. (2023), where they also found that green firms have outperformed brown ones over the past decade in many countries. The absence of a premium does not necessarily refute an existence of a clime risk factor, but could indicate that investors preferences and future regulations are already influencing the prices and reducing the compensation for these in the market. Pastor et al (2022) also find that smaller stocks react more slowly to climate news compared to large stocks.

At last, Friede et al. (2015) has reviewed 2000 empirical studies which provide evidence of ESG and financial performance. They saw that 90% of all the studies find a non-negative linear relationship between ESG and financial performance. Instead, a majority of the findings presented a positive linear relationship. Friede et al. concludes that incorporation of an ESG factor historically has not penalized investors on average, in some cases it has even improved it. This summarises an important point in our study, there is little evidence that by acting

environmentally friendly won't necessarily lead to worse returns. The findings are mixed; for example the carbon premium of Bolton and Kacperczyk (2020) study indicate that environmental risk is rewarded with higher returns, whereas other studies have pointed out the importance of investor preferences and market adjustments to a lower carbon economy for stock returns. Our study will help us to navigate the conflicting findings in the literature, to determine whether environmental factors are a risk factor today or a reflection of investor tastes in the market.

4. Method

4.1 Data Collection and Sample Preparation

Sample Selection: This study focuses solely on U.S listed companies, aligning with the U.S. market data used for traditional asset pricing tests (Fama-French factors are derived from stocks listed in the U.S.). To avoid limiting the sample used in the study to one or a few indexes, this study includes all U.S. listed stocks covered in the LSEG (Refinitiv) environmental database. This approach yielded an initial universe of 2,943 stocks.

Environmental data and filtering: We collected 93 environmental variables per firm from LSEG's ESG dataset, covering three different types of data. (1) quantitative performance metrics (e.g., emissions, energy use), (2) binary indicators (e.g., environmental policies, coded true/false) and (3) controversy flags (e.g., involvement in environmental incidents). As disclosure varied a lot from firm to firm, a data completeness threshold was applied: firms missing more than 35% of variables were excluded. 65% data availability was required. This approach balanced data quality with inclusion and sample size, removing poorly covered firms. This yielded a sample of 2,445 companies. This filtering step reduced biases from firms lacking environmental transparency.

The sample of 2,445 U.S. stocks contains a robust set of environmental data which combines numerical metrics, binary indicators and controversy flags. This comprehensive dataset served as the basis for constructing the custom environmental score.

4.2 Development of a Custom Environmental Scoring System

To quantify overall environmental performance, a custom scoring methodology was constructed based on LSEG's 93 environmental variables. The rationale for developing a tailored score rather than relying on already existing ESG ratings is presented in the following subsection.

Variable Selection: From the 93 variables only 55 were ultimately selected based on material relevance, data availability and lack of redundancy. Variables with minimal cross-sectional variation or duplicative metrics were also excluded. The final set reflects a careful yet comprehensive representation of key environmental indicators.

Categorization of Metrics: The selected 55 variables were then grouped into nine different thematic categories, each capturing a unique dimension of environmental performance. This structure is suitable for easy interpretability and enables different weighting assigned to each category based on importance. The table down below details the categories and their assigned weight. For instance, the *GHG Emissions & Carbon Efficiency* category includes metrics on greenhouse gas output and intensity, while *Water & resource Efficiency* focuses on water consumption and sustainable resource use. Each metric is assigned to one category only once.

Table of Environmental Score Categories and Weights

Category	Weight	Description of Category Focus
GHG Emissions & Carbon Efficiency	20%	Greenhouse gas emissions levels and carbon intensity/efficiency
Environmental Governance & Disclosure	8%	Quality of environmental management, oversight, and transparency in reporting.
Emissions & Pollution Policy	10%	Existence and rigor of policies on emissions reduction and pollution control.
Biodiversity & Land Use	12%	Impacts on biodiversity and land use; conservation efforts.
Circular Economy & Product Innovation	10%	Initiatives in recycling, waste reduction, and eco-innovation in products or services.
Climate Risk Management	12%	Assessment and mitigation of climate-related risks (e.g., climate scenario analysis).
Water & Resource Efficiency	10%	Efficiency in water usage and sustainable resource utilization.
Supply Chain Environmental Integration	10%	Environmental standards imposed on supply chain and procurement processes.
Environmental Controversies	8%	Track record of environmental incidents and controversies.

Scoring of Individual Metrics: Each variable was scored according to its type: numerical, binary or controversy-related.

- **Numerical Metrics:** A percentile ranking approach was used to assign scores that ranged between 0 and 100, standardizing all metrics. The percentile rankings were formed in the direction of the action. For positively oriented metrics (e.g., renewable energy use), higher values resulted in higher percentiles and thus higher scores. Reversely, rankings for negative oriented metrics (e.g., emissions, waste) were reversed so lower values received higher scores. This method, with percentile rankings and standardization, ensures comparability, mitigates outlier influence and prevents any single metrics from disproportionately affecting the total score.

- **Binary Metrics:** True/false indicators (e.g., ISO 14001 certification) were scored using an inverse frequency weighting scheme. A score of 1 was assigned to the environmentally favorable response (typically "true", though "false" was used in cases where it reflected better environmental practice). This score was then scaled by the inverse frequency of the favorable response in the sample. For instance, if 85% of firms responded in the environmental-positive direction, the contribution of that metric for those firms was weighted as $1 - 0.85 = 0.15$, ensuring that common practices had lower impact on the total score. Instead, if only 5% of the firms responded in the environmental positive direction, the corresponding score would be $1 - 0.05 = 0.95$. This method increased the influence of rare but impactful actions at the same time as it reduced weight of, potentially less meaningful disclosures, or "minimum praxis". This rewarded firms for environmental leadership rather than routine compliance.

- **Controversy Metrics:** Different from the other metrics these acted as penalties. Firms began with full scores for the controversy category, with deductions applied for reported incidents based on presence, count, or severity. A small or single incident could reduce a score to 0.5, repeated or severe controversies could end in a score of zero. This ensured that negative events directly lowered the overall score, preventing high ratings for firms with serious environmental harm despite otherwise strong metrics.

Aggregation and Normalization: We next aggregated metric scores into category scores and then into the final composite score:

- **Category scores:** Metric scores were aggregated into nine thematic category scores per firm and for the first eight categories (controversies category excluded) metrics values were summed and normalized using min.max scaling to derive a numerical value within the 0-1 range. The lowest scoring firm in each category was assigned 0 and the highest scoring firm 1, with linear scaling in between. This standardization is corrected for differing metric counts or scales across categories. This ensured equal contribution potential in the combined score. The controversy category was already on a 0-1 scale and thus further normalization was unnecessary.

- **Category Weighting and Composite Score:** Each firm's final environmental score was calculated as a weighted average of its nine normalized category score. The weighting which was outlined in the table of environmental score categories and weights reflects both sustainability materiality and relevance in current environmental data and ESG discourse. GHG Emissions & Carbon Efficiency received the highest weight of 20% due to its central and direct impacting role in environmental impact assessments. Categories such as Climate Risk Management and Biodiversity & Land Use were weighted at 12% each, highlighting their growing importance and future impacts. Pollution Policy, Circular Economy, Water & resource Efficiency and Supply Chain Environmental Integration were mid-weighted at 10%. The categories Environmental Governance & Disclosure and Environmental Controversies received 8% each, acknowledging their supportive or corrective roles rather than direct impact on environmental outcomes. The final environmental score for firm *i* was derived as:

$$Environmental\ Score_i = \sum_{c=1}^{9} w_c * Category\ Score_{i,c}$$

where w_c is the weight for category c, and $Category\ Score_{i,c}$ is firm *i*'s normalized score in that category. Lastly, scores were rescaled to a 0-100 range for interpretability.

Initial Implementation: The data gathering and scoring system was initially developed in Microsoft Excel, which enabled transparent manual inspection of raw data, scoring formulas and further calculations on intermediate level. Excel facilitated the application of percentile rankings, diverse conditional logic and binary weight scaling through their comprehensive formula approach. This environment supported validation and control of scoring logic and ensured methodological accuracy. Once final environmental scores were derived for each firm, they were used to construct portfolios. Further analyses were then conducted using Python to open up for more advanced statistical modeling.

4.3 Justification for a Custom Environmental Score vs. Commercial Ratings

A custom environmental performance score was developed to highlight key limitations of commercial already existing ESG scores or ratings. Research has shown that ESG scores provided from different institutions are frequently opaque, inconsistent and divergent in methodology. Unlike credit ratings, which correlate closely across providers (e.g., 0.99 between S&P and Moody's), ESG scores and ratings for the same firm can often differ significantly. An MIT Sloan Study reported an average correlation of just 0.61 across six major ESG rating providers. (Financier Worldwide. 2022, October). This misalignment, named "aggregate confusion" by Berg et al. (2022), arises from varying indicator sets, weightings, evaluation criteria and view of penalties from non-disclosure companies.
Dimson et al. demonstrate that a firm which is top rated by one agency may receive a significantly lower score from another, depending on the emphasis placed on differing ESG dimensions. (Dimson, E., Marsh, P., & Staunton, M. 2020).

Similarly, Christensen et al. (2022) found that more ESG disclosure can lead to increase in rating divergence since agencies interpret the same information in different ways. Another argument for the justification of developing a custom environmental score is due to the lack of transparency and isolation on environmental performance specifically. Proprietary methodologies often use unknown metrics and weights, making replication or adjustments practical impossible for users. (Financier Worldwide. 2022, October).

Moreover, the isolation of environmental performance is very rare since most ratings aggregate social and governance factors along with environmental factors into a composite score which complicates the process to evaluate the environmental factors even more.

To address these issues, the custom score was specifically designed to challenge four issues: (i) transparency, with clear defined metrics and weights which is easy to use and interpret; (ii) comparability, with a standardised scoring framework applied uniformly across firms; (iii) thematic focus, by isolating metricses on environmental performance from broader ESG factors; and (iv) analytical relevance, by aligning category weights with priorities after importance. This is a central action to this study.

This approach responds to academic challenges regarding more reliable and transparent ESG metrics (Berg et al., 2022; Dimson et al., 2020). It provides a consistent measure of environmental performance driven by research that avoids the subjectivity and opacity of third-party ratings.

4.4 Assumption of Historical Environmental Profile Stability

Since the environmental metrics reflect the firm's most recent fiscal year (2024), an assumption was required in order to classify stocks as “green” or “brown” over the full sampling period between 2015 and 2024. It was assumed that environmental performance is persistent which means that firms ranking in the top or bottom decile today were likely to be among the greenest or brownest historically.

This assumption is well supported by the gradual nature of environmental changes. Pioneers in sustainability often maintain strong practices due to their long-term investments and organisational cultures, while laggards often face structural barriers. Therefore, environmental ranking can be seen as relatively stable.

A study found that 68.5% of firms in the ESG top decile remained there after revisions of the rating which suggested stability at the extremes. Berg, F., Fabisik, K., & Sautner, Z. (2020).

Effects from the industry also reinforce this pattern since utilities and extractives consistently exhibit higher emissions than tech or finance firms. As such, without major transformations, large shifts in rank are unlikely. Firms with early ESG adoptions often sustain those actions and practices and therefore will outperform over time. (Eccles, R. G., Ioannou, I., & Serafeim, G. 2014). This assumption is acknowledge as a limitation as it introduces potential forward-looking bias. To diminish this, only firms in the top and bottom decile (10%) were used. The middle 80% of the sample was excluded where ranking is more fluid. This action increases the likelihood that classified firms maintained fairly consistent environmental performance during the sample period.

4.5 Portfolio Construction: Green and Brown Portfolios

Taking advantage of the costum scoring system, two portfolios were constructed to reflect the extremes in environmental performance. Firms in the top 10% of scores (245 firms) formed the green portfolio and those in the bottom 10% (247 firms) formed the brown portfolio. Sorting was based on environmental scores as of end 2024 and a slight asymmetry arose from firms falling exactly on the decile boundary.

Return Data Collection and Filtering: Monthly adjusted return data was collected for all firms included in the green and brown portfolio from January 2015 to December 2024 which consist of 120 months. Factor data came from the Kenneth French library and individual stock returns were sourced from LSEG's Refinitiv database. A 30% missing data threshold was applied to ensure reliability. Firms missing over 36 of the 120 months were excluded from the study. This removed firms with limited listings, newly IPO firms or early delistings. After this threshold was applied the green portfolio remained at 245 firms while the brown portfolios was reduced to 113 due to higher attrition, often from small-cap or delisted firms. No replacements were made for excluded firms to preserve the integrity of the environmental score consistency despite resulting portfolio size imbalance.

Addressing Portfolio Size Imbalance: Although the green portfolio had more firms included (245 vs. 113), equal-weighted returns for the portfolios were used to ensure comparability. Each stock contributed 1/N of its respective portfolio return each month. While the brown portfolio may experience slightly higher idiosyncratic volatility due to its smaller size, both portfolios remained sufficiently large and diverse for further robust analysis. This approach ensures an unbiased, balanced comparison between the greenest and brownest firms.

Portfolio Return Calculation and GMB Factor Construction: For each month, returns were averaged across firms with available data. Firms with missing returns that month were excluded, not counted as zero returns. Portfolio membership varied in a dynamic way with IPOs and delistings. The monthly return for the two portfolios was defined as:

$$R_{G,t} = \frac{1}{N_{G,t}} \sum_{i \in G} R_{i,t}, \qquad R_{B,t} = \frac{1}{N_{B,t}} \sum_{i \in B} R_{i,t},$$

The green minus brown (GMB) factor was calculated as:

$$GMB_t = R_{G,t} - R_{B,t}$$

This equation represents the spread in monthly return between green and brown firms. This results in a positive component when green firms outperform brown firms and vice versa. To align with well established asset pricing models, the GMB series was converted to excess returns by subtracting the monthly U.S. T-bill rate (as a proxy for risk-free rate) which was sourced from the Kenneth French data library. The resulting GMB-RF serves as a custom factor representing an environmental return premium. In order to conduct asset pricing analysis, monthly return data for traditional risk factors and benchmark test assets were collected. Monthly returns for the Fama-French three factors together with the momentum factor (MOM) were sourced from Kenneth French's data library.

- Market Risk premium (MKT-RF): Value-weighted market return minus the one-month T-bill rate.

- Size Factor (SMB): Return from a small-cap mimus large-cap stocks portfolio, capturing the size premium.
- Value Factor (HML): Return of high book-to-market (value) minus low book-to-market (growth) stocks.
- Momentum Factor (MOM): Return of prior winners minus losers which uses the 12-1 month specification (also called UMD).

All factors were used as monthly excess returns with the risk-free rate included separately and the custom green minus brown (GMB) factors was added to form a five-factor model which included following factors: MKT-RF, SMB, HML, MOM, GMB. To ensure consistent interpretation of intercepts (alphas) and comparability with standard asset pricing frameworks, GMB was expressed in excess form. Factor data, like the firm's monthly returns, spans from beginning of 2015 to end 2024.

Benchmark Portfolios: As test assets, the 25 Fama-French portfolios based on size and book-to-market ratio (B/M ratio 5x5 matrix) were used. This data was gathered from the same source as the factor returns. These portfolios are constructed on an annual basis by sorting NYSE stocks into five size quintiles and five B/M quintiles, yielding 25 combinations in total (e.g., Small-Value or Large-Growth). Each of the 25 portfolios is equally-weighted and updated monthly. This combination of 25 portfolios ensure exposure to diverse risk profiles, from small-growth to large-value firms. This is standard in empirical asset pricing. Since these portfolios are not constructed using environmental data at all, they provide an out-of-sample test of the explanatory power of the GMB factor.

For clarity, portfolios are named 1 to 25. Four "corner" portfolios highlight the naming logic:

- Portfolio 1 = Small size, Low B/M ratio (small growth)
- Portfolio 5 = Small size, High B/M ratio (small-value)
- Portfolio 21 = Large size, Low B/M ratio (large-growth)
- Portfolio 25 = Large size, High B/M ratio (large-value)

All 25 portfolio returns cover January 2015 to December 2024, consistent with the factor window.

Data Processing: All series were merged into one time series with 120 monthly observations and final data processing and cleaning were executed in Python. This prepared the data for upcoming regression analysis.

4.6 Regression Model, Additional Analysis and Robustness Checks

Regression Model and estimation Procedure

In order to evaluate the explanatory power of the GMB factor, a time-series regression framework aligned with Fama and French (1993) was applied to the 25 benchmark portfolios. Including the GMB factor allows us to test H_1 (Carbon Risk Premium Hypothesis), where a negative GMB beta indicates that the portfolio performs better when brown stocks outperform green stocks, consistent with brown stocks earning a risk premium for higher environmental risk exposure. For each portfolio, the following model was estimated:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{MKT,i}(MKT - RF)_t + \beta_{SMB,i}SMB_t + \beta_{HML,i}HML_t + \beta_{MOM,i}MOM_t + \beta_{GMB,i}GMB_t$$

Here $R_{i,t} - R_{f,t}$ is the portfolio's excess return and the right hand side of the equation includes five excess return factors. Each regression was run individually for $p = 1, \ldots, 25$ using ordinary least squares method (OLS) on the 120 monthly observations.

Estimation Procedure: To estimate the regressions and extract coefficients, t-statistics, intercepts and adjusted R^2 Python's statistical models library was used.

Interpretation of Results: Factor loadings (β) indicate exposure to its respective risk factor. Specifically, a significant $\beta_{GMB,i}$ suggest exposure, or sensitivity, to the GMB factor. In other words, a portfolio with positive GMB beta performs better in green-dominated return months.

A positive and significant GMB beta supports H_2 (Investor Preference Hypothesis), suggesting that portfolios earn higher returns when green stocks outperform, potentially due to investor preferences for sustainable firms driving up valuations and lowering their cost of capital. The intercept, alpha, captures unexplained average return and alphas statistically significant close to zero indicate a well-specified model.

Multicollinearity Diagnostics: A correlation matrix was computed in order to assess whether the GMB factor overlaps with existing factors. A correlation of greater than 0.8 is a clear indication of multicollinearity. Variance Inflation Factors (VIFs) were also computed to further investigate potential multicollinearity which uses following equation:

$$VIF_j = \frac{1}{1 - R_j^2}$$

Where R_j^2 is the R^2 from regressing factor j on all others. The empirical accepted threshold for non-present multicollinearity is a VIF less than 5. Since the intercept is constant, it was excluded from the VIF analysis.

Implementation: All regression and further analysis, tables and figures were conducted in Python throughout the full sample and outputs were aggregated and organized clearly to ensure consistency and enable further analysis.

Additional Analyses and Robustness Checks: Three complementary analyses were conducted to evaluate the dynamics and robustness of the GMB factor.

1. **Sub-Period Regressions:**
   The full sample was split into two sub-periods. Pre-2020 (January 2015-December 2019) and post-2020 (January 2020-December 2024), with 2020 chosen for two main reasons. It resulted in equal sub-periods and the onset of the COVID-19 pandemic and a surge in ESG investment awareness. The model was estimated once again separately for each 25 Fama-French portfolios over both sub-periods using the five-factors. This tested whether the explanatory power of the GMB factor varied over time and its dynamics,

this analysis serves as a robustness check for time stability of the factor. These sub-period address H_3 (Time-Variation Hypothesis), by examining whether the influence of the GMB factor, reflecting the return spread between green and brown firms, intensifies in the post-2020 period, consistent with growing climate awareness and shifts in investor behavior.

2. **Fama-MacBeth Cross-Sectional Regression:**
   To evaluate whether the GMB factor is priced in the cross-sectional dimension a two-stage Fama-MacBeth (1973) regression was applied. At first stage in the regression, betas from time-series regressions were used as exposure and in the following stage the 25 portfolio's excess returns were regressed on their respective factor betas for each month. The slope on the GMB beta represents the price of GMB risk (λ). Taking the average of the monthly lambda-values and testing for significance provides a test of whether GMB commands a return premium. This analysis complements the time-series regressions by highlighting if GMB returns are compensated risk.

3. **Rolling GMB Betas:**
   Rolling 36-month regressions were derived for the four "corner" portfolios presented earlier (1,5,21 and 25) to examine the time-variation in GMB exposure for small-growth, small-value, large-growth and large-value portfolios. Starting from January 2015 to December 2017, the window rolled monthly until the end of the sample in December 2024, estimating GMB betas for each month over the full sample. This method captures evolving market dynamics and investor preferences toward environmental characteristics over time.

All robustness tests were conducted in Python which enabled flexible and consistent estimation across sub-samples and other modified specifications. All together, these analyses test the sensitivity, stability and structural variation along with the impact of the GMB factor, enhancing confidence in its empirical relevance.

5. Empirical Results

This section aims to present the empirical findings from the portfolio regressions and additional analyses. The base asset pricing model employs the four standard factors – market excess return (MKT-RF), the size factor "small minus big" (SMB), the value factor "high minus low" which refers to book-to-market ratio (HML) and the momentum factor (MOM). Each of these factors is expected to load in each respective conventional direction: MKT-RF should have a positive coefficient for all portfolios (higher market returns would imply higher expected return for all portfolios), SMB should load positively for portfolios of smaller firm (and negatively for portfolios of larger firms), HML should load positively for portfolios that are value-oriented, i.e. high book-to-market ratio, and negatively for growth oriented portfolios (low B/M ratio). MOM is expected to be positive for portfolios with exposure to momentum, meaning portfolios that continue to perform well following past performance trends.

The extended model introduces a fifth factor, the "green minus brown factor" (GMB) which is defined as the return of a "green" stock portfolio minus the return of a "brown" stock portfolio. The GMB factor is included to assess whether positive environmental performance has incremental explanatory power for stock returns. A GMB factor that is positive indicates that the portfolio tends to perform better when green stocks outperform brown stocks whereas a negative factor would mean that the portfolio does better when brown stocks outperform green stocks. The estimation is conducted using time series regression on monthly return data.

5.1 Base vs. extended model: Overall fit and gains

Figure 1 presents one important component of the regression results, the adjusted R^2 gain when introducing the GMB factor, for each of the 25 portfolios for the whole sampling period. The adjusted R^2 measures the proportion of return variance explained by the model while accounting for model complexity. Across the full sample period (2015-2024), the GMB factor led to a modest improvement for most portfolios. The average increase in adjusted R^2 is approximately 0.0058 (in absolute terms). That means on average the extended model explains approximately 0.58 percentage points more of the return variation than without the GMB factor. The

improvement is consistently positive or close to zero for most of the portfolios. In a few cases the adjusted R^2 is essentially zero or even slightly negative which indicates that the additional factor did not help explain returns after accounting for the degrees of freedom. At the same time, several portfolios did see the GMB factor contribute significantly in a positive way.

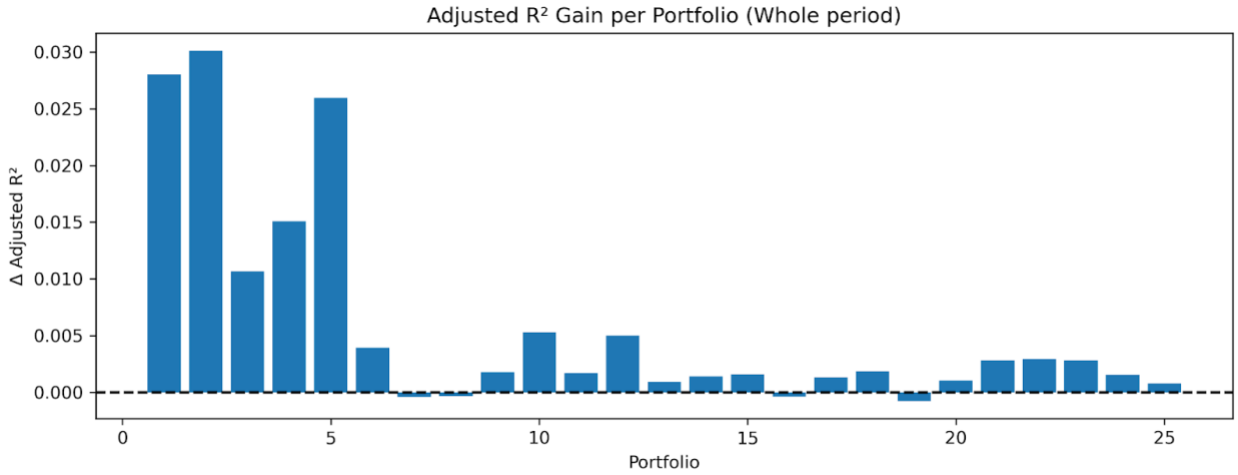

Figure 1: Adjusted R2 gained over full sampling period for each portfolio when adding the GMB factor

As shown in figure 1 some portfolios see an increase of the explanatory power of 1.5-3 percentage points (0.015-0.03) over the whole sample. Only 4 out of 25 portfolios saw negative development in adjusted. The portfolios with the largest gains in explanatory power suggest that the GMB factor is capturing return patterns not accounted for by the standard factors in those cases. Importantly, the improvement in model fit from the environmental factor is not uniform over time. The adjusted R^2 gains are substantially stronger in the post-2020 period than in the pre-2020 period. Figure 2 and 3 presents the adjusted R^2 gains for each 25 portfolios for both subperiods (pre- and post 2020). In the earlier subperiod between 2015 and 2020 the average adjusted R^2 gain is only about 0.0032 (Figure 2) whereas in the later subperiod the explanatory power increased with 0.0078 (figure 3). In other words, in the more recent five-year period the GMB factor added about 0.78 percentage points of the explanatory power on average compared with only about 0.32 percentage points prior to 2020.

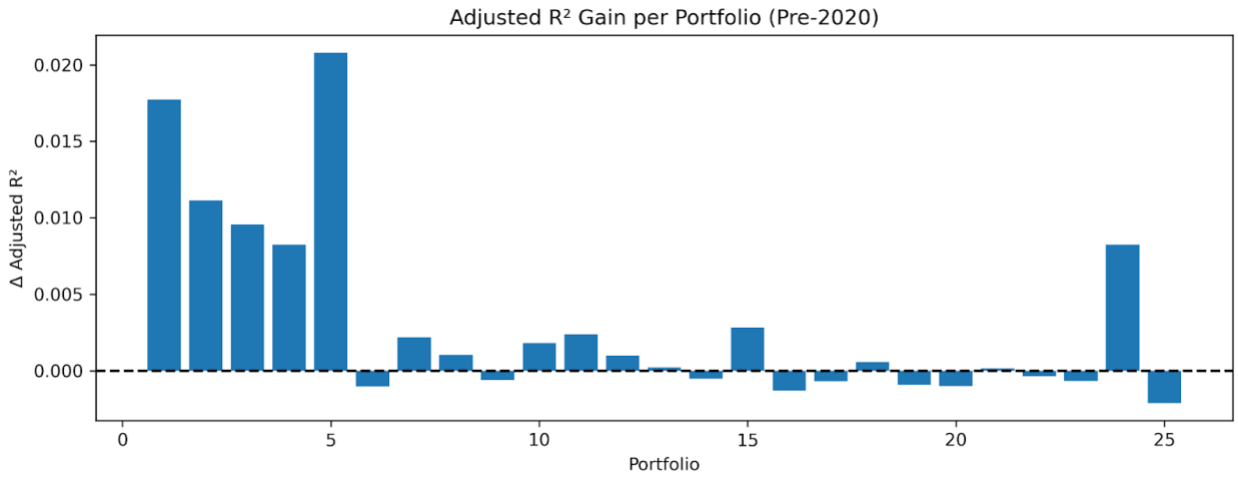

Figure 2: Adjusted R2 gained Pre-2020 for each portfolio when adding the GMB factor

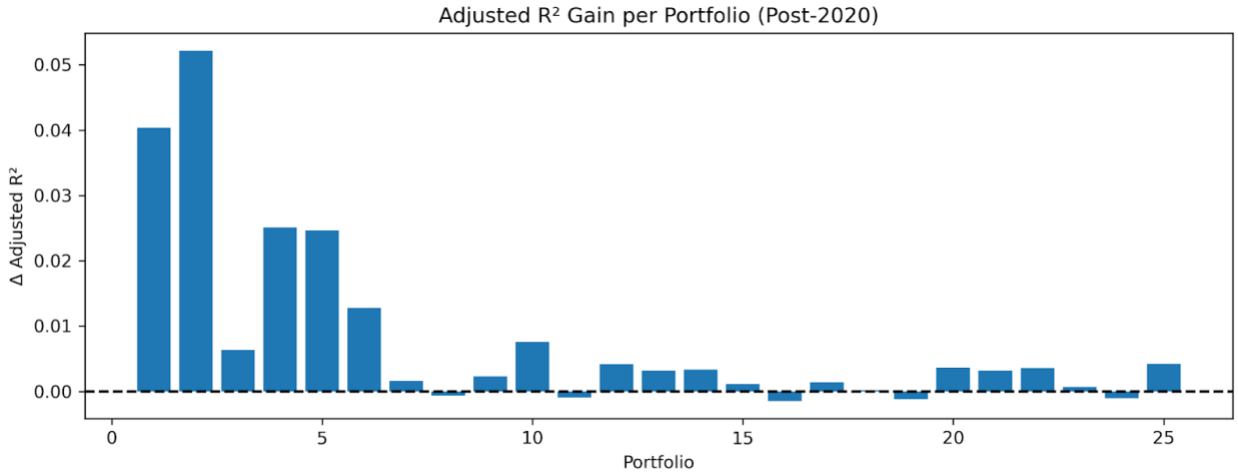

Figure 3: Adjusted R2 gained Post-2020 for each portfolio when adding the GMB factor

This pattern is evident in the distribution of adjusted R^2 improvements by portfolio. When figure 2 and 3 is compared it clearly presents itself that post-2020 only 5 out of 25 portfolios experienced negative gain in adjusted R^2 while in the pre-2020 was 9 out of 25. Also, the highest gain noted in a portfolio for adjusted R^2 pre 2020 was noted just above 2 percentage points (portfolio 5) compared to above 5 percentage points for portfolio 2 post 2020. Thus, the inclusion of the environmental factor appears to have become more relevant in explaining returns in the later years of the sample. It can be observed that the portfolios with highest gain in adjusted R^2 tend to be those in the smaller size categories (size 1 firms).

5.2 GMB Factor coefficients and significance across portfolios

The regression results for the GMB factor coefficients (betas), indicating a mix of positive and negative exposures across the 25 portfolios. Table 1 reports the estimated GMB beta and table 2 respective t-statistics for each factor and portfolio.

Portfolio	MKT-RF_beta	SMB_beta	HML_beta	MOM_beta	GMB_excess_return_beta
P1	0,965364307	1,155382357	-0,352674424	-0,174727316	-0,467177518
P2	1,014234584	1,041623968	-0,266842652	-0,191426384	-0,468740399
P3	0,992605709	1,084741407	0,008768212	-0,111680072	-0,270553898
P4	0,899815318	0,860820482	0,248490731	-0,195593559	-0,276437618
P5	0,885996254	0,866713737	0,372630671	-0,127407732	-0,391474132
P6	1,164435791	1,179310208	-0,304171452	0,011179672	-0,170322406
P7	1,02861705	1,056567855	0,018234536	-0,107012701	0,003185641
P8	1,055616939	0,957806002	0,308766716	-0,026472461	-0,013031556
P9	0,938885991	0,996268045	0,627764611	-0,077646938	0,09956334
P10	1,24438724	0,792654544	0,8611159	-0,043701584	-0,221323284
P11	1,097995389	1,006848612	-0,242989591	-0,025625495	0,105622259
P12	0,996301291	0,744449655	0,108809369	-0,10669852	0,141929214
P13	0,994982329	0,622542609	0,367817702	-0,082856515	0,071444941
P14	1,072656567	0,677827525	0,627874077	-0,097489943	0,0908557
P15	1,217213727	0,646904786	0,905802747	-0,114744738	-0,125978935
P16	1,054572716	0,514234573	-0,24021178	0,004519943	0,036261519
P17	1,047843523	0,387657195	0,189882559	-0,038523991	0,077702061
P18	1,035277708	0,315067715	0,4393823	-0,046531389	0,089251483
P19	1,09969183	0,366489989	0,515302397	-0,154965619	-0,021296851
P20	1,124599171	0,375479347	0,888594637	-0,09770618	-0,100569688
P21	1,025697061	0,104627201	-0,195409084	-0,004271537	0,090917355
P22	0,925928682	0,126041592	0,174028983	-0,035450922	0,091830861
P23	0,909371787	0,063395171	0,355233805	-0,06097218	0,092618905
P24	0,932112063	0,076488587	0,64750039	-0,042397096	0,087228816
P25	1,104184717	0,175073639	0,793335902	-0,034526692	-0,088971509

Table 1: Coefficient results from the regressions, including the GMB factor beta

Portfolio	MKT-RF_tstat	SMB_tstat	HML_tstat	MOM_tstat	GMB_excess_return_tstat
P1	11,20045655	6,873891203	-3,554987255	-1,67414011	-3,975100399
P2	13,16617014	6,933690417	-3,009510695	-2,052151176	-4,462467976
P3	14,33744649	8,034404528	0,11003369	-1,332162597	-2,865962584
P4	19,82329467	9,724504486	4,756113863	-3,558471655	-4,466231974
P5	13,01270979	6,527456303	4,754819677	-1,545319029	-4,216582554
P6	23,77289141	12,34601935	-5,395162202	0,188487271	-2,550115922
P7	32,05150289	16,88202377	0,493638454	-2,753699298	0,072796953
P8	35,32814425	16,43707222	8,97768882	-0,731636404	-0,319839751
P9	32,37414653	17,61546471	18,80623649	-2,211038543	2,517715209
P10	19,50563263	6,371194737	11,72696676	-0,565702607	-2,544211876
P11	27,75450699	13,05057394	-5,336300348	-0,534923619	1,957988472
P12	33,84251863	12,96700717	3,211129863	-2,993070084	3,535619011
P13	34,09892572	10,9402399	10,95159943	-2,344976513	1,795636771
P14	38,18937257	12,37466533	19,42110672	-2,86634295	2,372223956
P15	24,82407655	6,765175368	16,04942539	-1,932526893	-1,884195104
P16	28,45201822	7,114268532	-5,630545116	0,100706155	0,717470772
P17	35,51221684	6,736923781	5,590955324	-1,078199524	1,931236985
P18	34,85246522	5,438923841	12,85105529	-1,293627283	2,203503583
P19	23,85906875	4,077341057	9,713242856	-2,776542684	-0,338859417
P20	23,56902054	4,035173978	16,17957309	-1,691033919	-1,545726165
P21	48,05003358	2,513342215	-7,953145519	-0,165251462	3,123511117
P22	31,73947326	2,215483363	5,182786931	-1,003543537	2,308514014
P23	30,40039255	1,086741697	10,31743158	-1,683278388	2,270696219
P24	24,5167083	1,031627953	14,79631811	-0,920908855	1,682579865
P25	23,02672669	1,872163042	14,37365371	-0,59460965	-1,360702265

Table 2: T-statistics results from the regressions, including the GMB factor t-statistics

As shown in the tables some portfolios have positive GMB loadings while others have negative loadings, reflecting heterogeneous exposure to the "green" factor. For instance, eight out of the ten portfolios with smallest size show negative beta where seven of the betas are sizable and significant meaning these portfolios returns tend to increase when green stocks outperform brown stocks. In contrast, out of the 15 portfolios with the biggest size (size 3,4 and 5) only four portfolios show a negative beta.

What's notable is that three portfolios with negative beta are the "corner" portfolios, portfolio 10, 15 and 25. That's the portfolio with the highest B/M ratio for each level of size (B/M level 5 for size level 3,4 and 5). However, many portfolios have GMB betas that are close to zero, indicating little to no actual differential performance connected to the GMB spread. In terms of statistical significance, the majority of GMB coefficients are significantly different from zero at the conventional 5% level for the full sample, although it's not a vast majority. Specifically in the full period regression, 15 out of 25 portfolios have GMB coefficient t-statistics that are equal to or greater than 1.96 (indicating significance at the 5% level or better). Among these significant

coefficients, there are instances of both signs for the direction of the betas. The negative loaded betas mostly belong to the smaller market cap portfolios and the positive betas to the larger market cap portfolios and the presence of both positive and negative significant betas highlights that the GMB factor is capturing an effect that benefits some portfolios and detracts from others. Figure 4 and 5 provides a heatmap of the GMB t-statistics and coefficients across the 5x5 portfolio grid which clearly highlights the pattern mentioned earlier. As seen in the figures below negative significant betas appear only in the first two rows in the smallest size portfolios. The values not highlighted in the heatmaps indicating non statistical significance.

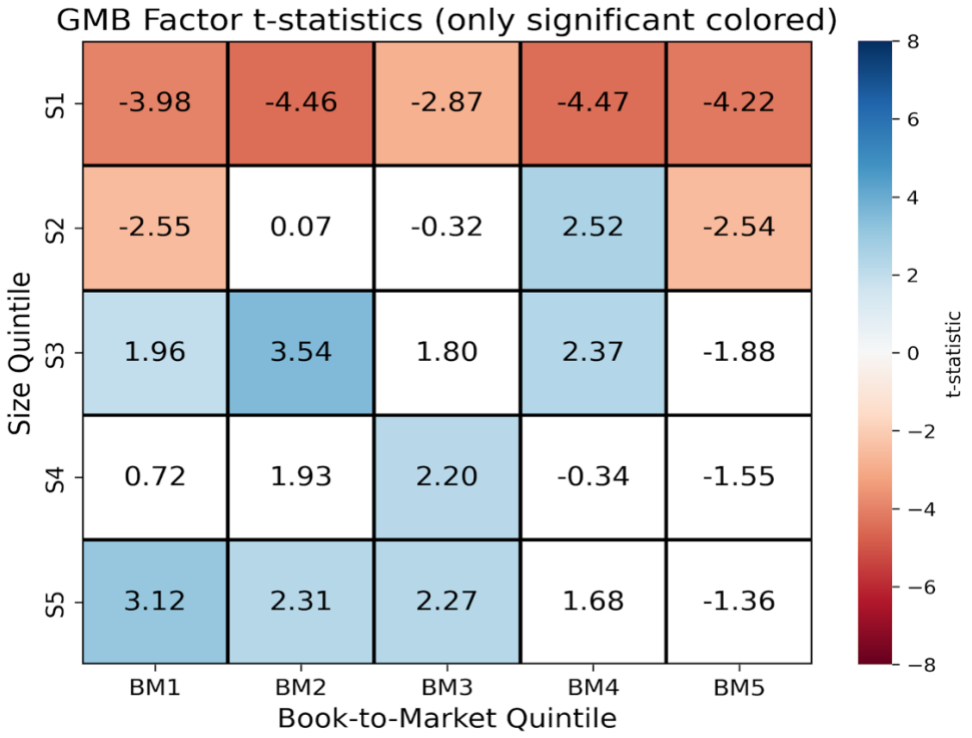

Figure 4: GMB t-statistics for FF 25 portfolios plotted through a heatmap

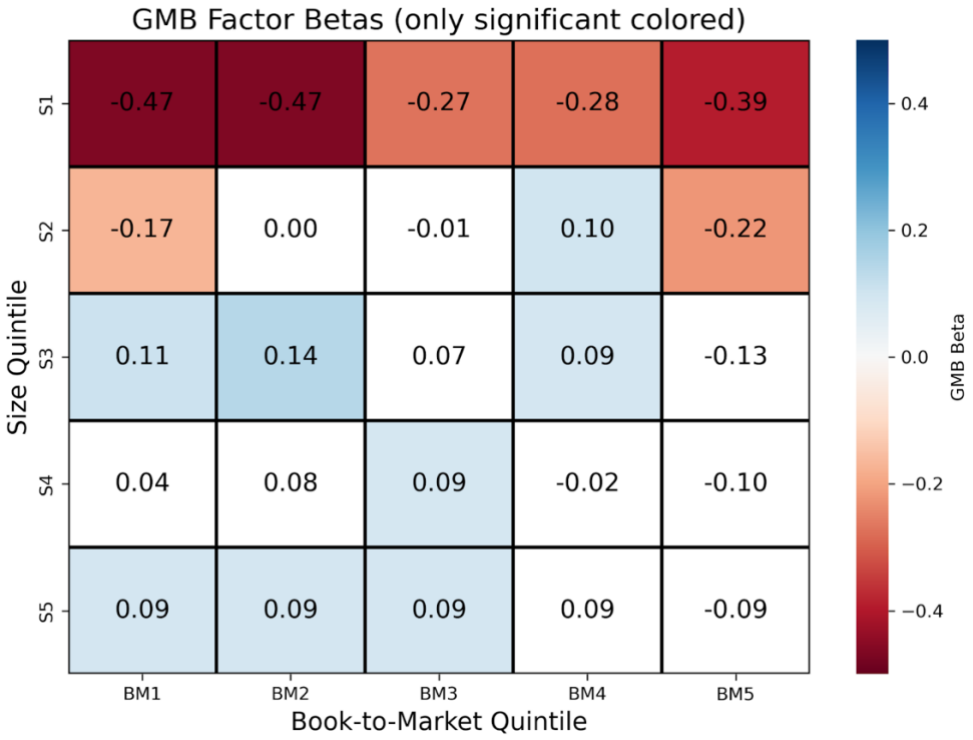

Figure 5: GMB betas for FF 25 portfolios plotted through a heatmap

The magnitude of most of the GMB coefficients in economic terms is almost zero even though it is significant in statistical terms. The statistically significant betas are generally on the order of around one tenth (in absolute value) which indicates less economic significance. For instance, a portfolio with a positive GMB beta of 0.25 would imply that a 1% higher return of the GMB factor in a given month is associated with an additional 0.25% return for that portfolio (all else equal). At the same time a negative beta of -0.25 would mean that the portfolio underperforms by 0.25% in months when green outperforms brown by 1%. As mentioned, most of the portfolios have GMB betas closer to zero, indicating smaller economic exposure to the GMB factor.
To show cross-sectional patterns figure 6 and 7 highlights the average GMB t-statistic by size and value quintile.

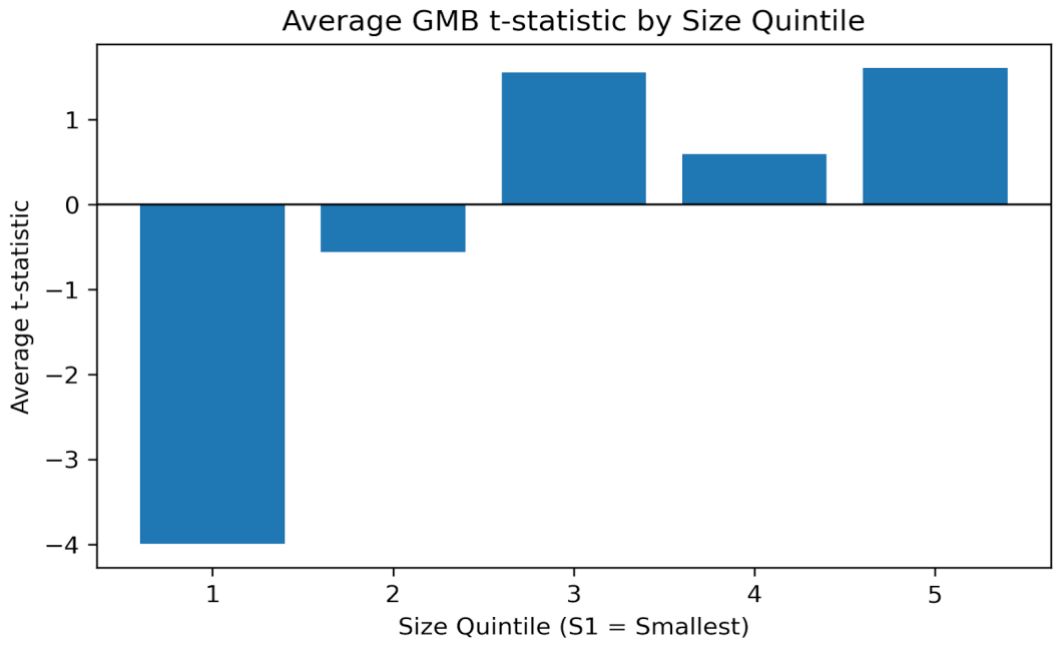

Figure 6: Average GMB t-statistics for FF 25 portfolios sorted on size

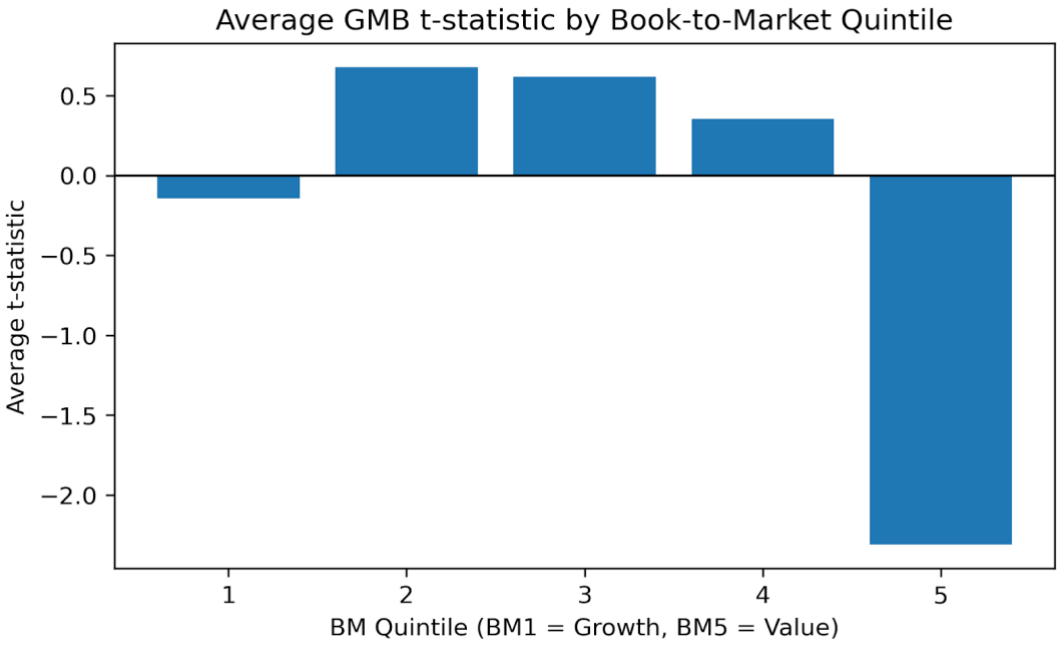

Figure 7: Average GMB t-statistics for FF 25 portfolios sorted on B/M ratio

Figure 8 and 9 shows the corresponding GMB factor betas by size and quintile.

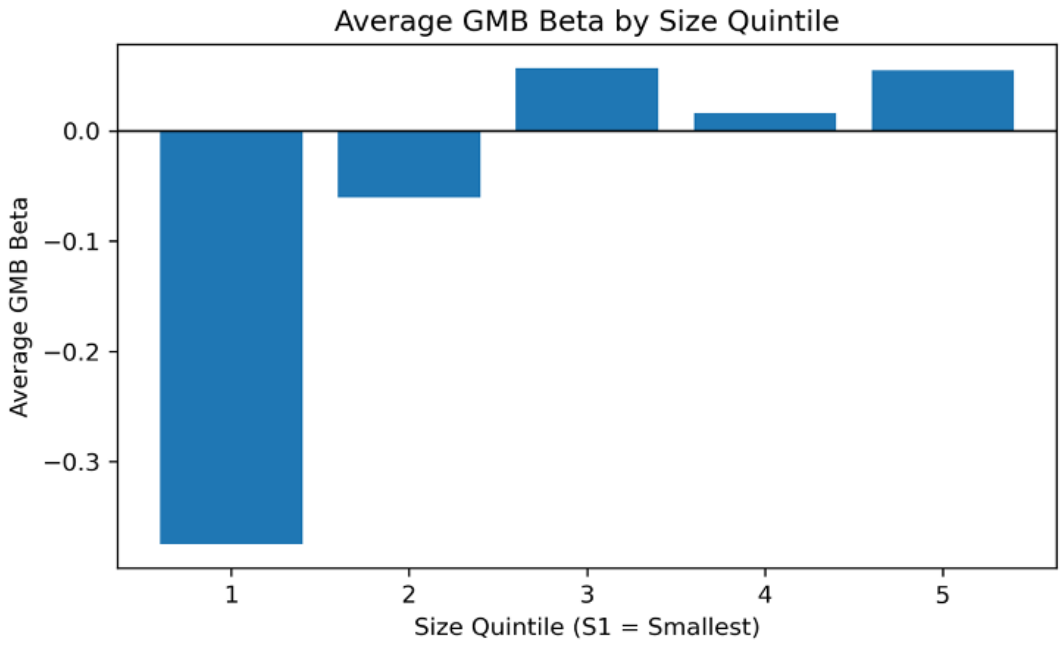

Figure 8: Average GMB beta for FF 25 portfolios sorted on size

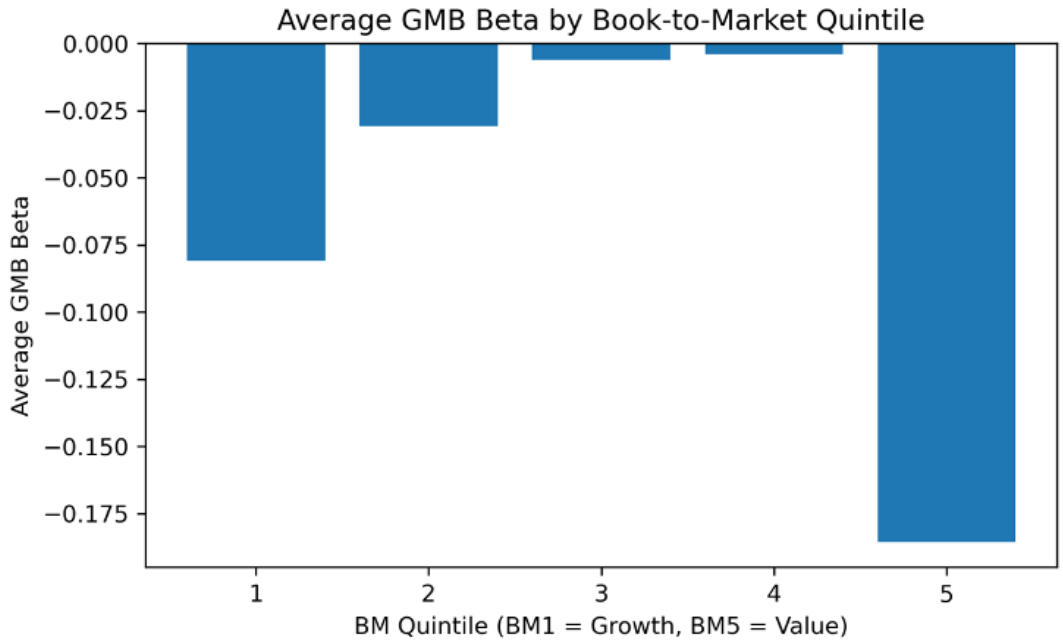

Figure 9: Average GMB beta for FF 25 portfolios sorted on B/M ratio

Portfolios including stocks with the smallest market capitalization (size 1), on average, have convincing negative beta with an average of -0.375 and an average t-statistic of -3.998. This result is, on average, statistically significant at a 0.1% level and all of the five small-cap portfolios have significant betas at the 0.5% confidence level. The average beta among these portfolios of -0.375 is also economic significant and implies that when green stocks outperform brown stocks by 1%, these portfolios tend on average to underperform by 0.375%, indicating a inverse sensitivty to environmental performance.

At the same time, the largest size quintile has almost zero exposure with an average beta of 0.0547 and an average t-statistic of 1.605, which means no economic or statistical significance. Across the book-to-market ratio dimension, growth-oriented portfolios with low B/M ratio (B/M 1 quintile) have neither any statistical or economic significance on average with a beta of -0.081 and t-statistic of -0.145. Value-oriented stocks with high B/M ratios (B/M 5 quintile) have an average GMB factor beta of -0.186 and an average t-statistic of -2.310 which indicates both statistically and economic significance.

These are average tendencies by portfolios with certain characteristics and only a subset shows statistically strong factor effects. This confirms that while GMB factor has significant impact in most of the cases, it is not universally significant across all tested portfolios.

5.3 Subperiod analysis: Pre-2020 vs. post-2020:

The GMB factor effect is examined in two subperiods, January 2015 to December 2019 (Pre-2020) and January 2020 to December 2024 (Post-2020), to analyse changes over time. To further illustrate the regression outcome and the difference in both statistically and economic significance between the two subperiods figure 10 and 11 plots a heatmap for all 25 portfolios betas for both the subperiods.

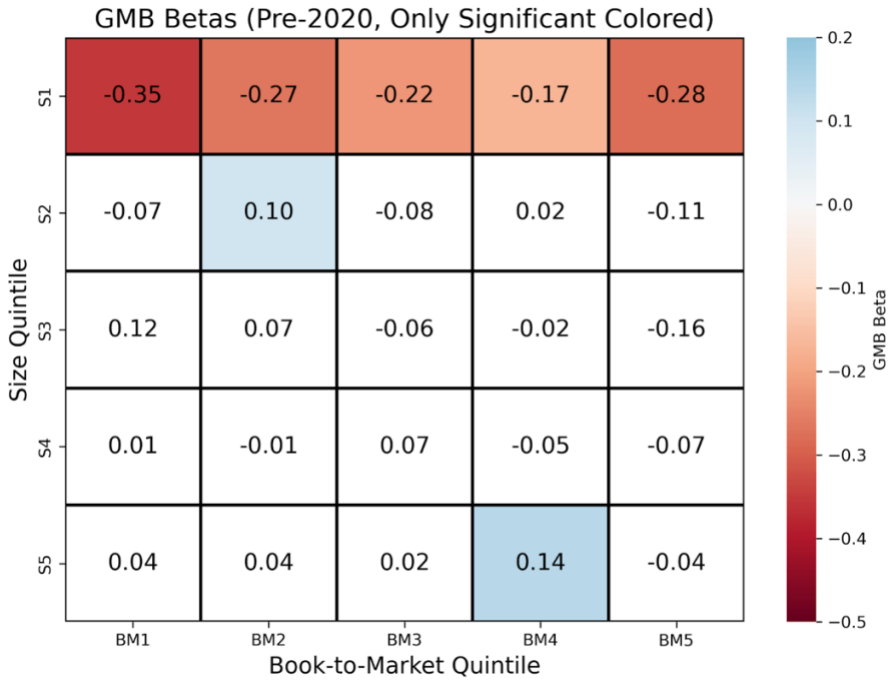

Figure 10: GMB betas for subperiod Pre-2020 plotted through a heatmap

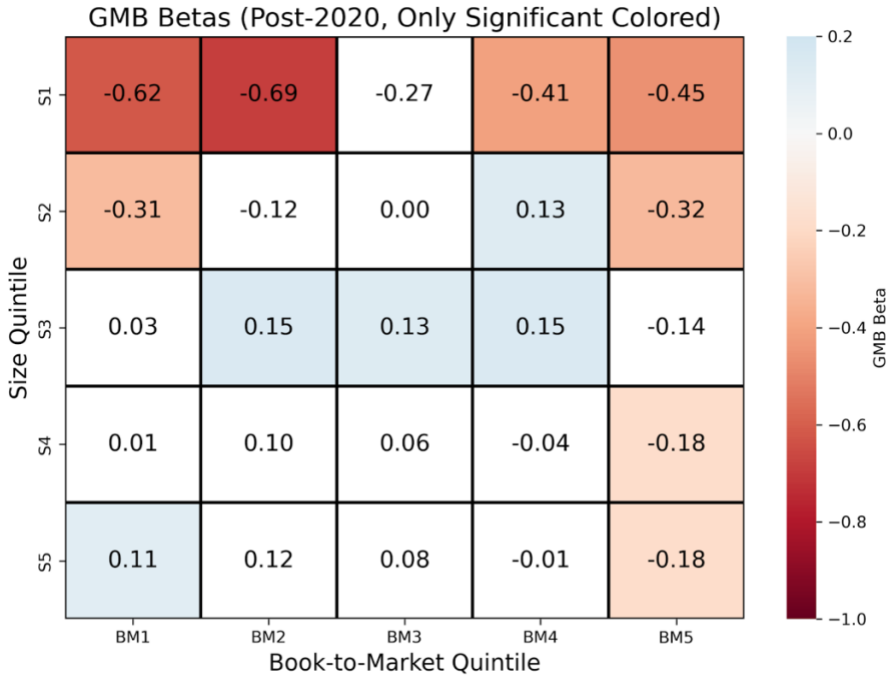

Figure 11: GMB betas for subperiod Post-2020 plotted through a heatmap

The regression outcomes reveal a clear shift in the influence of the GMB factor between the two subperiods. In the pre-2020 subsample the GMB coefficients are generally small and only a few cases of statistical significance. Only seven out of 25 portfolios show a significant GMB factor beta where five out of those seven belong to the size 1 quintile. The average impact on adjusted R^2 when introducing the new factor is only 0.0032 on average for that earlier subperiod and many portfolios had essentially no contribution from GMB in that period. In contrast, during the post-2020 period a greater number of portfolios display meaningful GMB betas. Specifically, 13 out of 25 portfolios show statistically significant betas where the economic significance seems stronger than for the whole period. The mean of all 25 portfolios betas (absolute value) changed by 97,2% comparing the pre-2020 mean beta of -0.0545 to post-2020 mean beta of -0.1075. The average adjusted R^2 improvement when introducing the GMB factor increased with 0.0078 post-2020 reflects these gains in explanatory power. The difference in adjusted between the two subperiods are illustrated in figure 2 and 3 earlier in this section. The empirical finding is that the GMB factor's contribution to the explanatory power to return variation is materially higher in the recent subperiod and this time variation is an important result, indicating that any pricing effect related to "green" vs "brown" firms has strengthened in the latter half of the decade.

5.4 Robustness checks and additional results

Several checks to ensure the reliability of the results were conducted. Table 3 indicates the correlation between all five explanatory variables used in the model: MKT-RF, SMB, HML, MOM and GMB_excess_return.

	MKT-RF	SMB	HML	MOM	GMB_excess_return
MKT-RF	**1,0000**	0,3009	0,0320	-0,3996	-0,0548
SMB	0,3009	**1,0000**	0,1139	-0,3521	-0,5840
HML	0,0320	0,1139	**1,0000**	-0,3042	-0,1742
MOM	-0,3996	-0,3521	-0,3042	**1,0000**	0,1601
GMB_excess_return	-0,0548	-0,5840	-0,1742	0,1601	**1,0000**

Table 3: Correlation between all five factors in the model plotted through a matrix

The correlation matrix reveals that all pairwise correlations are below 0.80, which is the level of concern for multicollinearity. The greatest absolute correlation is observed between SMB and

GMB_excess_return at -0.5840. Other notable correlations can be seen between MKT-RF and MOM at -0.3996 and MOM and SMB at -0.3521.

To further assess potential multicollinearity among the explanatory variables a VIF (Variance Inflation Factor) analysis was conducted. The VIF analysis was computed for each of the five explanatory factors, along with the intercept. The results are as follows: the VIF for the intercept came out at 1.16, MKT-RF is 1.27, SMB is 1.80, HML is 1.14 and GMB_excess_return is 1.69. All values fall well below the conventional threshold of 5.0 and are close to 1. This indicates that the explanatory variables used in the regression exhibit low correlation with each other which ensures stability in the estimation of their respective coefficients. The fifth introduced factor GMB has the second highest VIF value after SMB indicating that the latest introduced factor GMB has less correlation with the other factors than SMB.

5.5 Fama – MacBeth cross sectional regression (GMB Risk Premium)

The Fama – MacBeth two pass regression was conducted using our 25 portfolios sorted on size and value quintiles. This was conducted to estimate the cross-sectional prices of risk associated with the five-factor model. The average risk premium for the market excess return factor, MKT-RF, was 0.000017, with a standard deviation of 0.0813 which results in a t-statistic of 0.0023. The SMB and HML factor also came out with average risk premium of essentially zero, both with very low t-statistics. The momentum factor (MOM) showed a higher average risk premium of 0.0179 and a t-statistic of 2.2391 making it the only factor passing the 1.96 level for statistical significance. The GMB factor yielded an average risk premium of 0.003949 with a standard deviation of 0.0660 producing a t-statistic of 0.6555. These results indicate that among the factors in the model only momentum exhibited a statistically significant cross sectional risk premium at the 5% conventional level. This means that the GMB factor does not exhibit a statistically significant cross-sectional risk premium in the Fama – MacBeth framework, as its average risk premium estimate (0.003949) is associated with a t-statistic of only 0.6555, which is well below the 1.96 threshold. Therefore, the GMB factor cannot be interpreted as a traditional priced risk factor based on these results.

5.6 Rolling 36 – month beta analysis (Time varying GMB exposure)

To examine the time varying exposure to the GMB factor a 36 – month rolling beta was computed for selected portfolios. Figure 12 through 15 illustrate the rolling GMB betas for four representative "corner" portfolios: Portfolio 1 (Smallest size, lowest B/M. Size 1, B/M 1), portfolio 5 (Size 1, B/M 5), portfolio 21 (Size 5, B/M 1) and portfolio 25 (size 5, B/M 5).

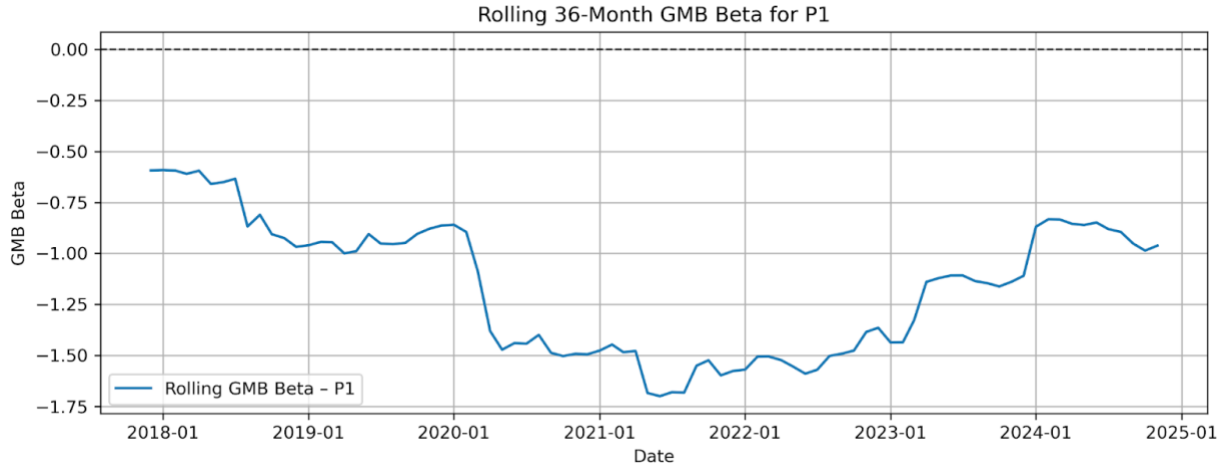

Figure 12: Rolling 36-month GMB betas for portfolio 1 (Small-Growth)

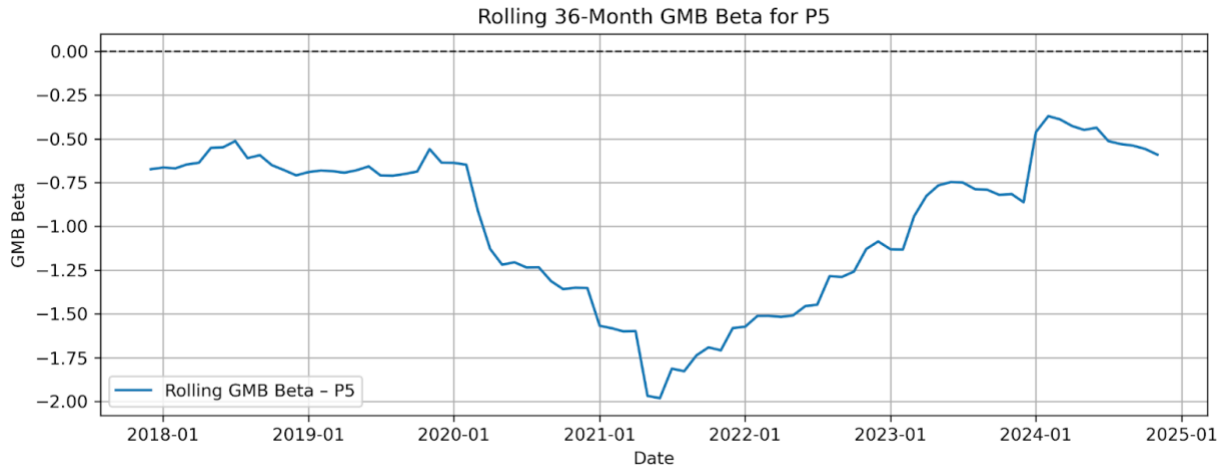

Figure 13: Rolling 36-month GMB betas for portfolio 5 (Small-Value)

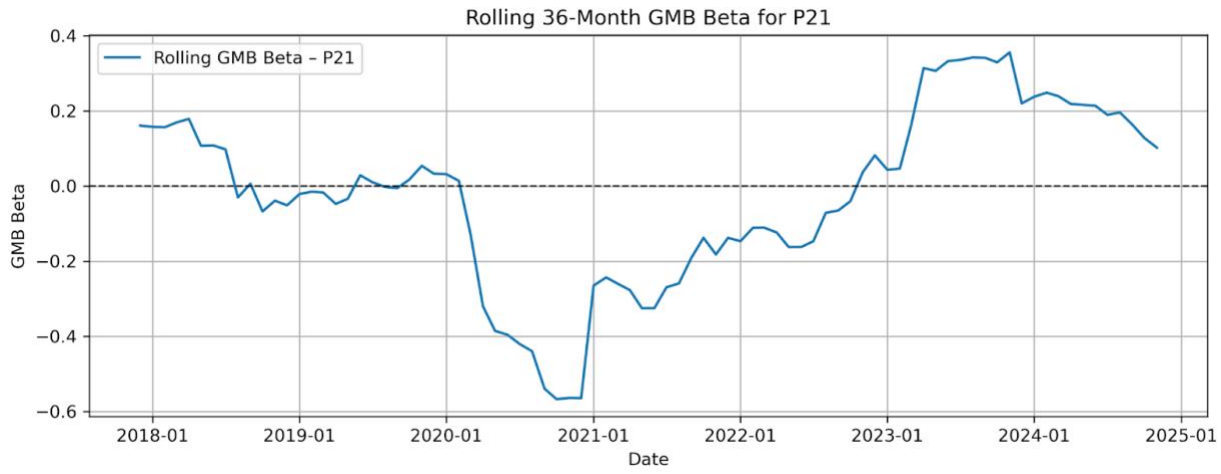

Figure 14: Rolling 36-month GMB betas for portfolio 21 (Large-Growth)

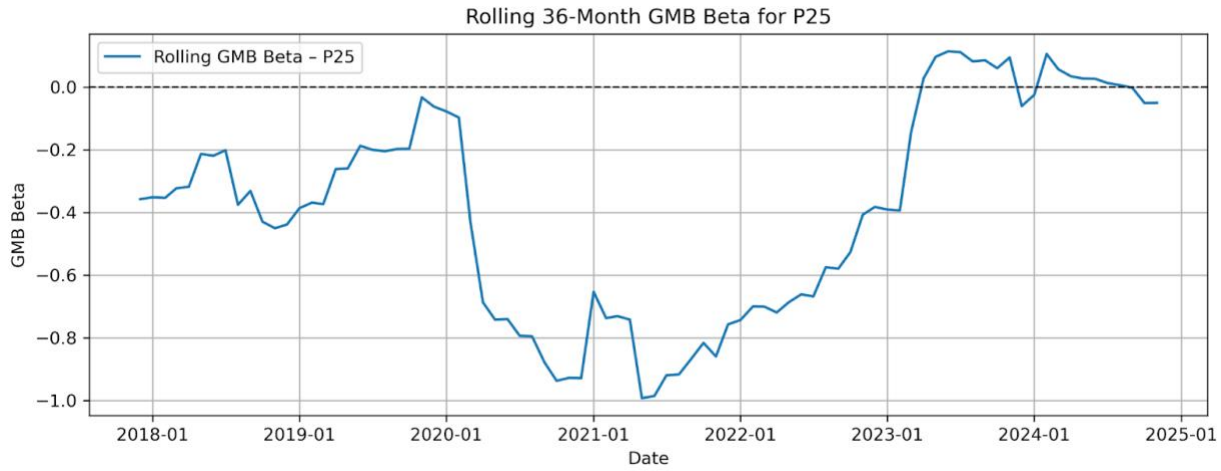

Figure 15: Rolling 36-month GMB betas for portfolio 25 (Large-Value)

These figures display how each portfolio's GMB factor direction evolves over time from 2018 through 2024 since it's a three-year rolling beta. From these plots it's clearly shown that all GMB betas vary substantially over the sample period for all portfolios with a general pattern of more negative GMB exposure during 2020 and 2021 followed by a partial or full reversal in the following years. However, clear differences can be seen across the size and B/M dimensions. Portfolio 1 and 5, which is the portfolios including stocks of the smallest sizes exhibit consistently negative GMB betas throughout the period. For portfolio 1 the rolling beta ranges approximately between -0.5 and -1.7 where the negative peak came around 2020.

Thereafter the GMB beta gradually moves upward or becomes less negative ending around -1.0 in December 2024. Portfolio 5 (Smallest size, highest B/M ratio) behaves similarly with a negative GMB beta under the whole period but becomes somewhat more extreme in magnitude. The rolling beta for portfolio 5 varies from around -0.3 to -2.0 with its highest peak in the beginning of 2024 and its lowest peak in the middle of 2021. Despite this rebound, both small cap portfolios always remained a GMB beta below 0 in the sample, indicating a persistent tilt toward "brown" stock. Between the two small cap portfolios the portfolio with highest B/M ratio (portfolio 5) generally shows a more extreme negative GMB beta than the Small sized growth portfolio (portfolio 1), especially during the 2020-2021 interval.

In contrast, the portfolios including larger cap stocks, portfolios 21 and 25 show less negative and more variable GMB betas. Portfolio 21 which is the large cap, low B/M ratio portfolio begins the sample with a slightly positive GMB beta of just under 0.20 indicating a mild exposure to the "green" factor. This beta then clearly declines into the negative area under and around zero before really making its way down in the negative territory under the year 2020. In late 2020 it reaches its lowest point of around -0.58 and after 2021, portfolio 21's GMB exposure steadily increases. The rolling beta crosses into positive territory by the end of 2022 and continues to rise until peaking at almost 0.4 in the end of 2023 before declining marginally.
On the other hand, portfolio 25 with large cap and high B/M ratio stocks starts with a slightly negative beta in 2018 with around -0.40%. This GMB exposure becomes more negative during 2020 and hits its lowest beta in mid-2021 of -1.0. From thereon portfolio 25's beta trends steadily upwards and by 2023 it had risen to approximately -0.4, to keep on going upwards and even touching positive territory. From early 2023 until the end of the sample the portfolio 25's GMB beta is around zero, indicating a neutral GMB exposure in 2024.

Comparing across these corner portfolios indicating a visible pattern related to size and book-to-market characteristics. The small-size portfolios 1 and 5 consistently throughout the sample have negative GMB loadings with portfolio 5 being the most negative. At the same time, the large-size portfolios 21 and 25 show higher GMB betas on average, with the large-size-growth portfolio 21 achieving substantially positive GMB exposure in 2023 and 2024.

Within each size category, the value-oriented portfolios with high B/M ratios (portfolio 5 and 25) tend to exhibit more negative GMB betas than their growth-counterparts (portfolio 1 and 21). What all four portfolios have in common is that they experienced their most negative beta around 2021 which was then followed by a partial rebound. By the end of the sample in 2024, only the large-size growth portfolio (21) has a clearly positive GMB beta. Portfolio 25 (large-size value portfolio) was approximately zero by the end of the sample and both the small-sized portfolios 1 and 5 remained substantially negative. These results show that GMB factor exposures are dynamic over time and experience systematic variation across the size and B/M ratio spectrum, even though none of the four corner portfolios shows a persistently positive GMB loading throughout the whole sample.

6. Discussion

6.1 Results

Our empirical findings reveal a nuanced answer to the question of whether environmental factors are priced in equity markets. Overall, the addition of the Green Minus Brown (GMB) factor provided a modest improvement in explaining portfolio returns over 2015-2024, but this effect was not consistent over time. In the first half of our sample (2015-2019), environmental performance showed little influence on returns, GMB factor loadings were small and mostly insignificant, indicating that the return difference between green and brown stocks were highly limited during this period (see Figure 10 & 11). It suggests that initially, markets did not systematically incorporate environmental factors into return sensitivities. Yet, in the other half subperiod (2020-2024) we observe a clear shift: more portfolios showed meaningful exposure, both statistically and economical, to the GMB factor, where more than 50% of the portfolios having statistically significant GMB betas. That is, the impact of companies' environmental performance on stocks grew significantly stronger in the recent years of our sample. This time-pattern is an important finding, implying that any effect regarding green vs. brown companies have strengthened in the past half of the decade. This supports our third hypothesis, the Time-Variation Hypothesis, which proposes that the relevance impact of environmental factors has increased over time. Our data confirms this, as the GMB factor became more statistically and economically significant for our 25 test portfolios after 2020.

This development can be linked to major shifts in investor behaviour and market conditions around 2020. We explicitly split the sample at year-end 2019 to identify potential impacts of rising environmental behaviour awareness and the COVID-19 pandemic on market dynamics. The evidence does confirm that the period after 2020 coincided with significantly heightened relevance of environmental factors. A likely explanation for this is that investor preferences for sustainable investments increased around this time, leading to large capital inflows into "green" assets. Such demand-driven dynamics can bid up the prices of environmentally friendly firms and/or depress the prices of environmentally harmful firms, thereby affecting realized returns.

This interpretation aligns with the investor preference (mispricing) hypothesis advanced in recent literature. For example, Pástor, Stambaugh, and Taylor (2022) argue that green firms can outperform not due to compensation for higher fundamental risk, but due to investors' increased demand for sustainability, which drives their valuations up. When investors reallocate their portfolios toward more sustainable assets, their realized returns may temporarily rise if the shift was unexpected, but the future expected return may be driven down as a result.

Our findings in the post-2020 subsample, where several portfolios exhibited positive and significant GMB betas, suggest that "green" stocks played a larger role in return variation across portfolios. This "green premium" in recent years (i.e. green stocks outperforming brown) could imply that prices have adjusted toward sustainability preferences, as one would expect under preference-driven equilibrium where investors accept slightly lower returns on green assets in exchange for non-financial utility. This is in line with our second hypothesis, the Investor Preference Hypothesis, which suggests that green stocks earn lower expected returns not because of heightened fundamental risks, but because investors are willing to pay a premium for sustainable and environmental friendly assets.

On the other hand, our results from the earlier period show more evidence consisten with a carbon risk premium narrative. In the 2015-2019 window, there were few signs that green stocks outperformed, if anything the data hints that "brown" firms may have earned slightly higher returns, as reflected by the fact that several portfolios displayed negative GMB betas during this period (though the betas were not always significant). This pattern could be interpreted as consistent with the hypothesis of a "carbon premium", which says that investors demand higher returns as a compensation for the exposure of climate risks. As discussed earlier, several studies support this view. For example, Bolton and Kacperczyk (2020) find that U.S companies with higher carbon emissions earned higher subsequent returns, even after controlling for size, value, and other factors. Such a return can be interpreted either as a risk premium for carbon exposure or as a consequence of investor exclusion (where "brown" firms trade at a lower price due to investor exclusion and thus offering higher yield).

Our findings in the pre-2020 subperiod can be linked to this framework. While our Fama-MacBeth regression did not show a statistically signifiant carbon risk premium in a cross-sectional sense, the strong negative significant GMB betas observed in several portfolios suggest that brown stocks contributed more to return variation during this time. It suggests that when sustainability concerns were less in focus, environmentally "brown" stocks may have indeed delivered relatively higher returns. This could be linked to the "sin-effect" documented by Hong and Kacperczyk (2009), in which companies that engaged in controversial practices earned on average a 2.5% annual premium due to exclusion of norm-constrained investors. Similarly studies like Heinkel et al. (2001) showed that if a subset of investors avoid polluters those stocks will be held by fewer investors and require a higher expected return (higher cost of capital) in the market to attract demand. Our results from the earlier period are in line with this framework: limited investors' demand for brown firms likely led to higher required returns for those firms.

In summary, our two periods show different results and can be linked to different hypotheses. The first few years of our study align with our first hypothesis about the carbon risk premium or the "exclusion effect", which states that firm with higher environmental risk exposure, referred to as "brown stocks", may earn higher returns as compensation for the systematic environmentally realted risks or due to investor exclusion effects. Whereas the later period aligns with our second hypothesis, an investor preference-driven scenario, highlighting an evolution in how environmental characteristics influence return dynamics.

Importantly, we find that this transition was statistically evident in our model. The inclusion of the GMB factor led to a small improvement in model fit before 2020 (on average adding 0.32 percentage points to adjusted R^2) but delivered a relatively larger improvement after 2020 (0.78 percentage points on average. Several portfolios that previously showed no relationship with environmental performance began to exhibit significant GMB loadings in the later period. This indicates that environmental performance has become a more significant factor in explaining return variation across portfolios in recent years. However, our Fama – MacBeth cross-sectional regression results do not support the view that GMB is a priced risk factor, as the estimated risk premium was small and statistically insignificant.

This suggests that while the GMB factor helps explain return variation across portfolios, it does not command a consistent premium in expected returns.

Furthermore, we observed that the impact of the GMB factor varied across different types of portfolios. In our 25 diversified test portfolios (sorted by size and book-to-market characteristics) the exposures to the "green" factor was different, some portfolios had positive GMB betas while others had negative. This finding suggests that certain categories of stocks are effectively "long green, short brown", while others are the opposite. What is interesting is that smaller cap portfolios tended to have negative GMB betas (in about 8 of the 10 smallest portfolios), meaning those small-cap portfolios performed better when brown stocks outperformed green. Conversely, most large-cap portfolios showed either positive or only slightly negative GMB betas, indicating large firms generally did better when green stocks led the market. This could be linked to Pástor et al. (2022) findings that small firms react more slowly to climate news. It also implies that smaller companies are often brown firms with poorer environmental scores, whereas larger firms tend to have better sustainability profiles. Large-cap stocks may be more rewarded by the market by becoming more green than small cap stocks which could be because they are often more monitored by the media and investors, aligning our findings with Pástor et al. (2020). We find that many of the strongest negative GMB exposures were concentrated in small, high book-to market (value) stocks, for example the "corner" portfolios represent the smallest size and highest B/M firms had significantly negative GMB loadings. These are likely companies in traditionally "brown" industries (e.g. heavy industry or energy) which are both small and value-oriented. Conversely, portfolios with positive GMB betas may correspond to large growth companies with strong environmental ratings (for instance, big tech or firms in cleaner sectors).

Our results also give us additional insight into how the GMB factor overlaps with traditional factors. We observed a negative correlation between GMB and the SMB (size) factor in our model (-0.58). This suggests that part of what the environmental factor captures is related to firm size, smaller firms (which often had lower environmental scores) behaved somewhat like "brown" firms, while larger firms tend to be closer to green firms. There were also signs of some correlation with the HML (value) factor, evidenced by the negative GMB betas for high B/M portfolios.

These relationships show that environmental performance is not entirely independent of other factors. It correlates with firm characteristics such as size and value, which are established risk factors.

This raises an interesting point if the explanatory power of the GMB factor is driven by established risk premia, like small cap-or value stocks, then one could argue that environmental performance is capturing existing risk exposures rather than introducing an entirely distinct risk factor in the model. However, our multivariate regression is controlled for size and value, so the GMB effects we detect are beyond what could be explained by SMB and HML. The correlations between the GMB factor and traditional factors were not that strong so they could not be seen as multicollinearity. This is further supported by our VIF (Variance Inflation Factor) analysis, where the GMB factor had a VIF of 1.69 which is well below the standard threshold of 5, indicating low multicollinearity. GMB had the second-highest VIF after SMB (1.80), confirming that it introduces relatievly unique information beyond traditional factors.

Therefore, while GMB overlaps to some extent with known factors, the environmental factor also captures an additional dimension of return variation, especially in recent years. This shows that an environmental factor can enhance the explanatory power of asset pricing models, whether due to a new type of risk or because investor preferences change.

To summarize, we see evidence supporting both perspectives of investor preference and a risk premium. On one hand, the risk-based view would argue that if environmental risk is systematic, then “brown” stocks should carry higher expected returns as compensation. We did find a (weak) brown premium in earlier years and significant GMB betas for many value-and small-cap portfolios, which supports the view that environmental risk might have existed but was only partially priced and creating excess returns for brown stocks.

On the other hand, the investor preference view implies that as more investors care about sustainability, prices begin to reflect personal values and ethical goals, which are not financial risk and return. The result of this can be that green assets being “overvalued”, yielding lower future returns and brown stocks “undervalued” and yielding higher subsequent returns. Such

preferences are likely to have created advantages during this shift post 2020, where investors bid up green stocks as a potential effect of an investor preference shift. This may have temporarily raised green returns and depressed brown prices and potentially enable for a carbon premium going forward if brown stocks remain relatively undervalued.

6.2 Limitations

In our study there are several limitations of our study that have to be acknowledged. Our sample period is between 2015 and 2024 which is a relatively short period of time. It also captures a decade of rapid ESG adoption and external shocks to the market such as the COVID-19 pandemic. These effects may not reflect long term market conditions, and the observed results post 2020 could be partly driven by temporary shifts in the market. Furthermore, climate risks and risk factors often play out over decades, so a ten year window could be insufficient to capture a persistent carbon risk premium. Longer term data and pre 2010 data could give us different results.

Our environmental scores are also exposed to measurement concerns. While the Refinitive ESG score includes several indicators, by constructing our own environmental score and by picking 55 environmental indicators from the Refinitive ESG, it will reflect assumptions about what good environmental performance is. These assumptions, or potential inconsistencies across data providers, may bias the classification of green and brown firms. Additionally, ESG disclosure quality has improved in recent years which may influence the reliability between early versus late scores. Although we use relative rankings in GMB constructions to minimize this effect, inconsistencies in data quality will still remain a limitation. Also, our focus on US equities will give us a geographical bias. Our results may not be the same in other different parts of the world, with different environmental behaviour norms.

Another limitation of our study is the unequal number of green firms and brown firms in their respective portfolio, 245 vs. 113. We used an equal-weighted approach to ensure each firm contributes equally within its respective portfolio, the difference in portfolio sizes still may introduce structural asymmetries. Specifically, the green portfolio will benefit from greater

diversification, which could lead to a potential lower idiosyncratic risk compared to the brown portfolio. This imbalance may influence return volatility and potentially affect the interpretation of the GMB factor's performance.

7. Conclusion

This thesis aims to examine whether environmental factors help explain stock return variation, and whether this pricing reflects a climate risk premium or investor preferences. Based on our analysis of U.S stocks from 2015 to 2024, we see that environmental performance has become increasingly significant in explaining return variation, especially after 2020. While the earlier years of the study showed little significant return difference between green firms and brown firms, the GMB factor improved fit in later years, indicating that the role of environmental performance in stock variation has grown over time. We also saw that larger companies tended to be more green and small-cap stocks tended to be more brown.

Overall, our findings provide empirical support for all three hypotheses. The early part of the sample (2015-2019) aligns with our first hypothesis, consistent with a carbon risk premium or exclusion effect, while the later years (2020-2024) support the second hypothesis, indicating increasing influence of investor preferences. This shift in explanatory power and significant GMB betas between these periods confirms our third hypothesis, suggesting that environmental factors have become more relevant in explaining stock return variation over time. However, our cross-sectional Fama–MacBeth regression results do not provide evidence that the GMB factor is consistently priced as a traditional risk premium.

This study will contribute to the literature on sustainable finance by studying how structural changes and environmental factors affect stock returns in more recent years. Instead of relying on a single indicator, such as carbon emissions or broad ESG scores, that rely heavily on Social & Governance factors, our study will focus solely on the environmental effect. A broader focus of 55 indicators will provide a better measurement of the environmental performance. Furthermore, by incorporating the Fama French Carhart multifactor model we saw that environmental factors have had a growing impact on stock returns in the later period.

For investors, this suggests that sustainability considerations are becoming more important which investors will have to take into account in portfolio construction and investment decisions.

Our study has some limitations. Our sample is restricted to only one market and a relatively short time frame, which is marked by shocking events such as the COVID-19 pandemic. We also rely on our own environmental score that may contain subjective assumptions about what good environmental behaviour is.

Research in the future could extend our analysis by analyzing the international market instead of focusing solely on U.S equities, to see if this behaviour is consistent in all markets. It would also be valuable to analyze which environmental indicators affects stock returns the most. For example, does carbon exposure affect stock returns the most, or are other factors more important in explaining returns?

In conclusion our study shows that environmental factors can help explain return variation in equity markets, though not necessarily as a priced risk factor in the classical sense. Instead they appear to be more of a growing combination of financial performance and sustainability investor preferences.

References

Adam, H. (2024). *Arbitrage Pricing Theory (APT) Formula and How It's Used*. Investopedia. Retrieved April 18, 2025, from https://www.investopedia.com/terms/a/apt.asp

Bauer, M. D., Huber, D., Rudebusch, G. D., & Wilms, O. (2023, November 21). *Where is the carbon premium? Global performance of green and brown stocks*. Brookings Institution. https://www.brookings.edu/articles/where-is-the-carbon-premium-global-performance-of-green-and-brown-stocks/

Berg, F., Fabisik, K., & Sautner, Z. (2020). *Rewriting the past: The role of corporate sustainability ratings (*ECGI Finance Working Paper No. 708/2020). European Corporate Governance Institute. https://www.ecgi.global/sites/default/files/working_papers/documents/bergfabisiksautnerfinal.pdf

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

Bolton, P., & Kacperczyk, M. (2020). *Do investors care about carbon risk?* ECGI Finance Working Paper No. 711/2020. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3398441

Carhart, M. M. (1997). *On persistence in mutual fund performance*. Journal of Finance, 52(1), 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Christensen, D., Serafeim, G., & Sikochi, A. (2020). *Why is corporate virtue in the eye of the beholder? The case of ESG ratings* (Working Paper No. 20-084). Harvard Business School. https://www.hbs.edu/ris/Publication%20Files/20-084_6c5b0248-d117-4049-baad-c0e1877eb537.pdf

CRSP. (2025). *Center for Research in Security Prices (CRSP) monthly stock file* [Data set]. University of Chicago Booth School of Business. https://www.crsp.org

Dimson, E., Marsh, P., & Staunton, M. (2020). Divergent ESG ratings. https://doi.org/10.17863/CAM.55949

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). *The impact of corporate sustainability on organizational processes and performance*. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1964011

Engle, R., Giglio, S., Lee, H., Kelly, B., & Stroebel, J. (2019). *Hedging climate change news* (NBER Working Paper No. 25707). National Bureau of Economic Research. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3317570

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3–56. https://www.bauer.uh.edu/rsusmel/phd/Fama-French_JFE93.pdf

Fama, E. F., & French, K. R. (2014). *A five-factor asset pricing model*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2287202

Financier Worldwide. (2022, October). ESG ratings: Key considerations for stakeholders. Financier Worldwide Magazine. https://www.financierworldwide.com/esg-ratings-key-considerations-for-stakeholders?utm_source

French, K. R. (2025). *Fama/French & momentum factor data* [Data set]. Kenneth R. French - Data Library. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment, 5*(4), 210–233. https://www.tandfonline.com/doi/full/10.1080/20430795.2015.1118917#d1e141

Görgen, M., Jacob, A., Nerlinger, M., Riordan, R., Rohleder, M., & Wilkens, M. (2020). Carbon risk. *Working Paper.* (Available via SSRN and Federal Reserve Bank of San Francisco). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2930897

Heinkel, R., Kraus, A., & Zechner, J. (2001)**.** The effect of green investment on corporate behavior. *Journal of Financial and Quantitative Analysis, 36*(4), 431–449. https://www.researchgate.net/publication/226649051_The_Effect_of_Green_Investment_on_Corporate_Behavior

Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics, 93*(1), 15–36. https://www.sciencedirect.com/science/article/abs/pii/S0304405X09000634

Microsoft. (2025). *STOCKHISTORY function* [Function]. Microsoft Excel. https://support.microsoft.com/excel

Pástor, L., Stambaugh, R. F., & Taylor, L. A. (2022)**.** Dissecting green returns. *Journal of Financial Economics, 146*(2), 403–424. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3864502

Refinitiv. (2025). *Refinitiv ESG database* [Data set]. London Stock Exchange Group.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*(3), 425–442. https://doi.org/10.2307/2977928

The bright spots in a complicated ESG framework. (2022, July 25). The World Economic Forum. Retrieved April 16, 2025, from https://www.weforum.org/stories/2022/07/still-reason-for-optimism-about-esg-investing/