Does ESG Information Deliver Investment Value? A High-Dimensional Portfolio Perspective

May 27, 2026

Abstract

This paper assesses whether ESG information helps build more efficient portfolios. Instead of treating ESG as a monolith, we build on a high dimensional information set of more than 200 ESG characteristics and employ a host of robust portfolio construction methods designed to avoid overfitting in this setting. Our results show that ESG information is not redundant, increasing the portfolio Sharpe ratio by up to 25 percent in-sample, compared with using financial information alone. However, out-of-sample benefits are insignificant as estimation errors offset any information advantage. We also show that optimal use of ESG information does not necessarily imply positive sustainability, as portfolios contain both green tilts to factors like human rights and brown tilts to factors like sin stocks. Furthermore, using ESG metrics provides no measurable risk reduction compared with using financial characteristics alone. Our findings suggest that ESG integration fails to deliver financial benefits out of sample.

JEL classification: C52, C55, G1, G11, G12, G17

Keywords: : ESG, Sustainable Investing, Portfolio Construction

1 Introduction

Many institutional investors are integrating sustainability criteria – also known as Environmental, Social, and Governance or ESG criteria – in their investment decisions. ESG information is inherently high dimensional – there are hundreds of characteristics to capture different dimensions of environmental, social and governance features of companies. This paper assesses whether this information helps investors build more risk-return efficient portfolios. More specifically, we ask whether ESG information leads to incremental improvements when used in addition to standard financial information.

Our perspective corresponds to an *"ESG aware"* investor who does not gain non-pecuniary benefits from ESG characteristics but considers the information to gain financial benefits, i.e., improve his portfolio's Sharpe ratio (see Pedersen et al. (2021)). This has become a pressing issue in practice as the legitimacy of ESG investing is increasingly questioned. Consider the following quote from a US-based investment officer at a public pension fund: *"We are ESG Advocates. Our sole proxy for success is financial performance and risk mitigation. Having more information when making investment decisions can be informative and better than having less information."* Whether using additional information in the form of ESG characteristics adds value is precisely the question we address. Our analysis also aligns with recent proposals for *"rational sustainability"*, which *"involves treating ESG like any other factor"* (Edmans, 2023, 2024). This is exactly the perspective we adopt.

The finance literature has produced substantial evidence on the financial implications - or risk premia - of ESG tilts. However, the literature focusses on low dimensional ESG information and assesses its value in-sample. In contrast, we assess the value obtained out-of-sample when using high dimensional ESG information. While ESG information is notoriously high dimensional—with a large number of possible issues and different ways to measure them—empirical studies often treat ESG as a monolith and rely on aggregate ESG scores or on a handful of selected scores for particular issues. For example, Avramov et al. (2022) consider aggregate ratings, and Eskildsen et al. (2024) construct an aggregate measure of greenness by averaging across selected characteristics. Several papers narrow the focus to single issues such as carbon emissions (Bolton and Kacperczyk, 2023), employee satisfaction (Edmans et al., 2024; Edmans, 2011), or vice activities—also known as *sin stocks*—(Hong and Kacperczyk, 2009). Selecting specific ESG scores ignores the rich information set available to investors and introduces a risk of data-snooping, as it is unclear what criteria guided the selection of one or a few variables from the high-dimensional information set that is available.

Empirical tests typically remain in-sample. To evaluate whether a single ESG variable earns a premium, researchers mostly rely on mean-variance spanning tests – that is, they check whether the ESG factor delivers a significant alpha in a multifactor regression. The assessment is equivalent to asking whether – when looking back over the full history of returns – adding a tilt to the ESG variable along with standard financial variables – would have increased the Sharpe ratio of a portfolio that holds the ex-post optimal weights in each tilt, compared with a similar ex-post optimal portfolio that excludes the ESG tilt but includes relevant financial factors. Neither the portfolio including the ESG tilt nor the benchmark without the ESG tilt are achievable ex-ante. If the ESG information used in such tests is not redundant with respect to the financial factors but adds so much estimation risk that investors cannot construct better out-of-sample portfolios, such tests will still conclude that ESG information adds value, because they ignore the estimation risk.

To offer a new perspective on the value of ESG information, we assess the out-of-sample value of using high-dimensional ESG information. To account for the rich information set available to investors, we build a dataset of more than 200 ESG characteristics from different sources. Thereby, we account for the well-known fact that ESG information is notoriously high dimensional: for example, an influential review by the World Economic Forum[1] considered hundreds of ESG metrics.

Recent review papers emphasize the high dimensionality of ESG information as a key issue that has not been adequately accounted for in the literature. For example, Hoberg and Manela (2025) write that *"the way firms contribute to [...] ESG goals is likely higher dimensional than is currently represented in the literature"*. Similarly, Pastor et al. (2025) argue: *"ESG is not a monolith. In our theory, a firm's greenness is a single number, but that number implicitly embeds a multitude of sustainability-related characteristics."* We account for this multitude of ESG characteristics to reflect the rich information set available to investors.

To build our high-dimensional ESG dataset, we complement aggregate ESG metrics with granular, issue-level information (e.g., waste management). This captures how investors process ESG information in practice (Edmans et al., 2025)[2], further motivating our high-dimensional approach.

ESG information matters only if investors can use it *ex-ante* to build portfolios that earn higher Sharpe ratios out of sample. Yet, it is well known that portfolio optimisation introduces estimation error, often delivering disappointing out-of-sample performance (Kan and Zhou, 2007; DeMiguel et al., 2009b). The problem is magnified when the factor set is large, because the portfolio is more likely to load on factors whose performance over the estimation sample is driven by noise (Kan and Zhou, 1999; Harvey et al., 2016; Kan et al., 2024). Therefore, to effectively assess out-of-sample value, we rely on robust portfolio-construction techniques that impose diversification constraints to reduce the impact of noise in risk and return estimates on portfolio weights.

Our analysis thus relies on two main ingredients: a high dimensional dataset and robust portfolio methods.

To capture high dimensional ESG information, we build a dataset of 222 characteristics from various sources. We include both traditional ESG ratings based on analysts' opinions, and alternative ESG characteristics either coming from academic research (sin stocks, high employee satisfaction, climate exposure) or from commercial providers that rely on systematic textual analysis of news or firms' documents. As set of controls, we consider a standard 'factor-zoo' including 130 financial characteristics already used in the previous literature (Kozak et al., 2020).

[1] These include metrics on a variety of reporting frameworks, such as GRI, CDP, SASB, Natural Capital Protocol, and WBCSD, see (KPMG International, 2020)

[2] In a global survey of 509 equity portfolio managers, Edmans et al. (2025) find that investors assess sustainability in granular terms.

To build robust portfolios, we follow several methods that focus on imposing constraints to avoid picking up noise in our high dimensional dataset and avoiding overfitting of portfolio weights. Following Kozak et al. (2020), we first build a long/short portfolio for each characteristic weighting stocks in proportion to their rank. These single factor portfolios form the basic building blocks of our analysis. We then construct a multi-factor portfolio that combines these characteristics-based single factors in a way that maximises the Sharpe ratio. With lots of factors, estimating efficient weights of each factor will lead to overfitting noisy patterns. Robust methods impose a penalty for solutions that have a large sum of squared weights, i.e., that are heavily concentrated in few factors. This forces portfolios to remain diversified over multiple factors.

We employ different portfolio construction methods to allow for robustness of our conclusions. Our first approach is to estimate optimal weights under Bayesian priors for the highest achievable Sharpe ratio as in Kozak et al. (2020)[3], which we abbreviate as KNS. Our second approach, in the spirit of DeMiguel et al. (2009a), is to form norm-constrained mean variance portfolios and our third approach is to construct similarly constrained minimum variance portfolios. To allow for general conclusions, we do not settle on a single level of constraints in each of these three methods but instead consider three different levels of constraints for a total of nine robust portfolio strategies. Using such robust portfolio methods alleviates the concern that adding more information via a high dimensional set of ESG data mechanically increases estimation risk. Instead, the regularisation imposed in these robust portfolio methods allows us to efficiently exploit the high-dimensional data set. For comparison, we also consider a naïve optimisation where we maximise the Sharpe ratio without imposing constraints.

Our focus is on the incremental value of adding ESG information to standard financial characteristics. Therefore, our benchmarks are portfolios constructed from the 130 financial characteristics using the same portfolio approach. We then consider portfolios that add the 222 ESG characteristics to the 130 financial characteristics and assess the difference in Sharpe ratio with the benchmark. While our interest ultimately lies with the out-of-sample results, we report the incremental value both in-sample and out-of-sample. For the in-sample results, we use returns for the entire period to estimate the weights of the multifactor portfolios and compute their returns over this same period. For the out-of-sample results, we split our sample period into three subsamples and use the other two subsamples to estimate portfolio weights and then evaluate returns for the held-out sample.

Our main findings on the incremental value of ESG information are shown in Figure 1. The figure shows the difference in Sharpe ratio between the multifactor portfolio that combines ESG and Financial information and the corresponding portfolio that uses only Financials, across the ten different portfolio methods we employ. The figure shows that any incremental benefits that appear in-sample, disappear completely in the out-of-sample setting, even when using robust portfolio methods that avoid overfitting. This key result challenges the idea that integrating ESG information adds value.

[Figure 1 about here]

[3] See also MacKinlay (1995) and Barillas and Shanken (2018).

More broadly, our analysis delivers three key results:

First, we show that the apparent risk/return benefits of using ESG information disappear out-of-sample. Our in-sample results display substantial benefits of using ESG information, suggesting that it is not redundant with respect to a high dimensional set of financial information. However, these in-sample benefits decline substantially when imposing strong diversification constraints, implying that in-sample value added stems from a few concentrated bets. Out of sample benefits are insignificant across portfolio methods, suggesting that the additional estimation risk in the expanded information sets offsets any information advantage.

Second, we show that optimal use of ESG information involves both green and brown tilts. A data-driven ESG-aware investor ends up with substantial brown tilts as well as green tilts. For example, portfolios consistently tilt to sin stocks across different portfolio-construction methods. On the other hand, human-rights issues and water-resource use are examples of consistent green tilts. Purely considering ESG as an information advantage thus requires anti-ESG investing on some dimensions. This result shows that *"rational sustainability"* does not necessarily imply positive sustainability. Note that the Principles for Responsible Investment, a UN initiative, requires the use of ESG information but does not prescribe any particular direction for investment tilts. The commitment of PRI signatories states they "believe that environmental, social, and corporate governance (ESG) issues can affect the performance of investment portfolios," and the first principle they commit to simply states: *"We will incorporate ESG issues into investment analysis and decision-making processes."* In the sense of the PRI, ESG integration thus means considering environmental, social, and governance factors as part of investment analysis – not automatically tilting portfolios toward *"green"* investments or ESG leaders. Our tests are consistent with this understanding of responsible investing. We also find that alternative ESG metrics receive substantial weight relative to traditional analyst-opinion ratings.

Third, we find that a focus on risk mitigation does not yield stronger benefits of using ESG information. Motivated by investor beliefs that ESG information should be useful for risk mitigation (Larcker et al., 2025), we test strategies that focus solely on volatility reduction. Similar to the Sharpe ratio results, we find that adding ESG characteristics does not lead to significant volatility reduction out of sample.

Overall, our analysis suggests that integrating ESG information in efficient portfolio construction does not yield financial benefits and goes counter to industry practice that exclusively focuses on positive sustainability tilts[4].

The remainder of the paper proceeds as follows. Section 2 defines the data used to represent financial and ESG information in our analysis. Section 3 presents summary statistics that describe the return and diversification potential of ESG and financial factors. Section 4 details the methodology used to construct our multi-factor portfolios, while Section 5 reports the empirical results for those portfolios. Section 6 concludes.

[4]This is not a problem if investors have preferences for sustainability as in that case investors may enjoy non-pecuniary gains from integrating ESG in the construction of their portfolios. But it is important if investors use ESG only with the expectation to improve their financial performance.

2 Information Sets

We start by providing a description of the characteristics we use in our analysis, both for sustainability and for standard financial information.

2.1 ESG information

We have witnessed a proliferation of ESG metrics, as investors may care about a host of different ESG issues (e.g., toxic waste vs. employee safety) and different ways to measure them (e.g., litigation or news reports or company policies). The results is an entire 'zoo' of ESG metrics.

To represent this ESG-zoo we collected a large and diverse set of 222 characteristics. We draw on different sources to attain an exhaustive list of ESG topics and measurement methodologies. We consider both granular data on particular issues (say waste management) and aggregate ESG information, which constructs an overall score from combining the underlying issues – often weighting them by sector specific considerations about so-called materiality for company profitability or risk.

We require each characteristic to have at least 15 years of history to allow us to run our tests. In the world of ESG data, this is a demanding requirement. For comparison, Avramov et al. (2022, Table B1) report that even some major ESG ratings providers offer less than ten years of historical data with ratings data starting as late as 2016 or 2014. In contrast, our analysis requires data starting in 2008. The challenge of data availability for early years is exacerbated in our setting where we require granular data rather than only aggregate ratings. Despite this tough requirement for 15 years of historical data, we were able to identify a large variety of relevant sources[5].

To facilitate our analysis of green and brown tilts we sign each ESG characteristic in the direction of positive sustainability, so that all our ESG factors are long high sustainability stocks and short low sustainability stocks. We split our ESG characteristics in two categories – traditional opinion ratings and alternative data.

Our traditional opinion ratings include consensus ratings that give us information on the consensus opinion of major ESG ratings providers for 17 different ESG dimensions, made up of 1 overall rating, 4 category ratings (environment, employees, community and governance) and 12 underlying themes. We obtain these consensus ratings from CSRHub, which aggregates 973 different data sources including leading ESG rating providers (MSCI, LSEG, Sustainalytics), as well as open-source data (e.g., the CDP A list) and more specialized providers. To improve granularity, we add opinion ratings from a specific data provider, Moody's, that covers many subtopics within the usual ESG dimensions giving a total of 61 distinct ESG scores. Moody's obtains these scores with a discretionary assessment of the risks an ESG issue poses to a company's finances and the business's impact on the ESG issues, the so-called "double materiality".

[5]We invite ESG researchers who possess characteristics that match our data-quality criteria (sufficient granularity and at least 15 years of uninterrupted history starting from July 2008) to share their variables with us for inclusion in future revisions. Data can be sent to research@scientificbeta.com.

Amongst the alternative ESG data, we have 21 open-source metrics from academic sources. These include activity classifications into *sin stocks* and public rankings identifying leaders in *employee satisfaction*[6]. Including these characteristics is important as there is widely cited evidence for a link with investment performance. This category also includes 19 climate metrics from Sautner et al. (2023)[7] who use textual analysis of earnings call transcripts to capture firms' exposure, risk, or sentiment to different climate issues (opportunities, regulation, or physical impact). We consider an additional 110 metrics that combine textual data and analyst views. These metrics measure reputation and controversies extracted from media sources and company's self-reported performance on ESG issues, and cover topics and subtopics within the usual ESG dimensions. We obtain these metrics from Covalence, a Swiss provider of alternative ESG ratings. Finally, we complement our set of alternative ESG characteristics with 13 scores obtained from systematic textual analysis of company reports and web sources. These scores assess ESG performance, materiality and transparency and are published by GaiaLens a UK-based provider of "an AI-powered sustainability analytics platform for institutional investors".

To provide a sense of the ESG issues included in our dataset, we went through the relevant documentation to identify the distinct ESG issues covered by our dataset. We identify 49 distinct ESG issues that we list in table 1.

[Table 1 about here]

We also highlight that most of our ESG characteristics incorporate the so-called principle of "industry-specific materiality", the idea that the financial relevance of specific ESG issues varies across industries[8]. Our dataset accounts for such materiality considerations simply because most of the data providers assess materiality and construct their scores accordingly. However, we do not take a stance of what is material or not. Instead, we aim for a broad coverage of ESG issues and rely on providers' choices of what is material for their respective data.

Overall, our ESG dataset represents a comprehensive aggregation of issues and methodologies

[6]We follow Hong and Kacperczyk (2009) to create a dummy variable to capture stocks that operate in SIN industries (alcohol, tobacco, and gaming), and Edmans (2011) to capture firms that have high levels of employee's satisfaction.

[7]We download them directly from one of the authors' website.

[8]This principle recognizes that not all ESG characteristics matter equally for all companies—what is "material" for one industry might be immaterial for another. Research argues that integrating industry-materiality into ESG metrics is important to capture long-term investment value (Khan et al., 2016). As documented by Berg et al. (2024), major ESG raters typically account for industry-specific effects, albeit with different perspectives about what ESG issues are truly material. Therefore, aggregating scores from several providers, our consensus metrics from CSRHub allows us to capture the average opinion about the ESG issues that are deemed as material. Industry-materiality is also explicitly integrated in the Moody's ratings through the "double-materiality" approach, as well as in the methodologies of our two commercial providers of alternative ESG characteristics. GaiaLens relies on a proprietary approach to reflect industry-materiality in the weightings of its aggregate ESG scores. Covalence accounts for industry-materiality in its processes relying on the Sustainability Accounting Standards Board (SASB) map to flag the relevant news/data.

reflective of standard industry practices, providing a reflection of the extensive "ESG zoo" confronting investors interested in sustainability.

2.2 Financial information

As set of controls for the value of ESG information we construct factors based on purely financial information. To create our set of financial characteristics we follow Kozak et al. (2020), henceforth KNS, who compile a list of 130 financial characteristics in two distinct categories.

The first set, that we call Published Return Effects, includes 50 financial characteristics which are typically regarded by the academic literature as being useful to explain the cross-section of expected returns. KNS compile this list following standard definitions from Novy-Marx and Velikov (2016), McLean and Pontiff (2016), Kogan and Tian (2015), and Hou et al. (2015). We follow the same definitions provided in the appendix of KNS and generate our own set of characteristics using common financial datasets (CRSP and Compustat)[9]. To facilitate the description of the dataset, we categorize the characteristics in this category into seven themes based on standard factor definitions: Size, Value, Momentum, Profitability, Investment, Low risk, and Others. Each characteristic is assigned to one of these themes and given a sign (+1 or −1) such that a high (low) value implies a high (low) expected return, in line with the original evidence that identified the corresponding premium.

The second category, which we denote as General Firm Characteristics, includes 68 financial ratios obtained from the WRDS Financial Ratios Database, plus 12 additional return-based characteristics capturing past 12-month returns. While these characteristics are recognized in the academic literature, they are not generally linked to expected returns. For example, leverage is widely used in studies of capital structure, but there is no consistent evidence supporting its predictive power for returns. Nonetheless, KNS demonstrate that augmenting Published Return Effects with General Firm Characteristics improves mean-variance efficiency of portfolios out-of-sample. Hence, to assess whether ESG characteristics add value out-of-sample, we control for the full set of financial characteristics, including General Firm Characteristics. Additionally, the inclusion of General Firm Characteristics helps mitigate concerns about data mining. While published return effects are known to perform well ex-post[10], ex-ante investors may plausibly consider a broader set of financial signals. By expanding the set of financial characteristics beyond those with established links to returns, we reduce hindsight bias in the factor pool available for constructing multifactor portfolios.

3 Descriptive Statistics on Individual Factors

This section introduces the factors we construct from the information sets described in the previous section based on US stocks. First, we explain the methodology to construct the

[9] A full list of definitions is available in the appendix. Characteristics are made available on one of the authors website up until 2019. We replicate them in order to be able to extend our analysis to 2023.

[10] It is worth noting that McLean and Pontiff (2016) find that factors based on standard financial characteristics tend to have a lower premium after they are discovered (post-publication effect).

time-series of factor returns. We then describe how we adjust factor returns to improve comparability across factors. Finally, we provide descriptive statistics on factor returns. We focus on statistics that illustrate the potential of ESG factors to add value either by enhancing returns (alpha potential) or reducing risk (diversification potential). Since we are merely describing the factors' risk and return properties, we do not adopt an out-of-sample perspective in this initial analysis.

3.1 Factor construction and return adjustments

We compute factor returns for a period of 15 years, from August 2008 to June 2023. To build our factors we follow the methodology in KNS. We construct factors as monthly rebalanced long-short portfolios. Both legs of the factors are portfolios of US stocks in the CRSP universe. To make sure that our results are not driven by illiquid stocks, we exclude micro-cap stocks – stocks with market capitalization below 0.01% of the aggregate market capitalization at each point in time[11].

For all our characteristics, we then construct a factor which represents a portfolio that tilts to the relevant characteristic. First, we compute a normalised rank score for each stock. We then compute returns for a long-short portfolio that weights stocks in proportion to their rank score. Finally, we adjust factor returns to eliminate differences in market beta and volatility across factors.

We compute the weighting criterion as follows: First, for each characteristic 'i' of a stock 's' at month 't' denoted $c^i_{s,t}$, all stocks are sorted on the value of their characteristic. We only use information that is available when stocks are sorted, so characteristics are always lagged relative to returns. Each month, stocks are ranked by their characteristic in ascending order, from 1 to the number of all stocks n_t. Second, we obtain the rank-transformation of the characteristics as follows:

$$rc^i_{s,t} = \frac{rank(c^i_{s,t})}{n_t + 1} \tag{1}$$

Then, we normalize the rank-transformation by subtracting the cross-sectional average ($\overline{rc^i_t}$) and dividing by the sum across all stocks:

$$z^i_{s,t} = \frac{\left(rc^i_{s,t} - \overline{rc^i_t}\right)}{\sum_{s=1}^{n_t} \left|rc^i_{s,t} - \overline{rc^i_t}\right|}, \quad \text{with} \quad \overline{rc^i_t} = \frac{1}{n_t}\sum_{s=1}^{n_t} rc^i_{s,t} \tag{2}$$

We collect the relevant characteristic score for the entire cross section of stocks each month in the vector $\mathbf{Z}_{s,t}$.

To obtain factor returns, we multiply the characteristic score by the vector of stocks' monthly returns at month t ($\mathbf{R}_t$) to obtain the factor returns for characteristic s at month t ($F_{s,t}$).

[11] This gives us about 1,000 stocks each month to construct our factors.

We also obtain daily returns adjusting $\mathbf{Z}_{s,t}$ daily to make sure that each leg of the factors reflects a monthly rebalanced portfolio[12].

To facilitate the comparison across factors and focus on the ability of their ability to capture cross-sectional differences in returns (alphas), we make two adjustments to the factor returns $F_{s,t}$. First, we remove the impact on the returns that can be explained by the market and obtain the market-exposure adjusted returns $\dot{F}_{s,t}$

$$\dot{F}_{s,t} = F_{i,t} - \beta_s R^e_{M,t} \tag{3}$$

where $R^e_{M,t}$ is the market factor[13], and β_s is the exposure of the factor to the market factor estimated with OLS regressions of monthly factor returns on monthly observations of the market factor plus an intercept[14]. For our in-sample analysis we use the full sample to estimate market exposures, while for our out-of-sample results β_s is estimated only with the observations from the estimation period[15].

We then adjust for differences in volatility across factors taking the volatility of the market factor as reference

$$\ddot{F}_{s,t} = \dot{F}_{s,t} \frac{\sigma(R^e_M)}{\sigma(\dot{F}_{s,t})} \tag{4}$$

As for the market-exposure adjustment, the volatility adjustment is also done based on monthly observations[16] and using only information available over the estimation period[17]. Note that for the rest of the paper, when we discuss factor returns, we will refer to the adjusted factor returns $\ddot{F}_{s,t}$ unless otherwise specified.

These risk adjustments facilitate comparison across factors for different characteristics. When assessing potential for return enhancement, we immediately obtain alpha over the market at the same level of volatility for each factor. When considering potential for risk reduction, we only need to consider correlations across factors given similar volatility levels by construction.

[12] Every day of the month we adjust the weights in each leg of the factor to reflect differences in returns over the previous days in the month. This gives us a sequence of daily returns for each leg of the factor that reflects a monthly rebalanced portfolio.

[13] The return of the value-weighted market in excess of the risk-free rate obtained from Kenneth French's website.

[14] We also compute daily market adjusted returns for the factors but always using the same beta estimated with monthly returns to reduce the impact of differences in liquidity across factors.

[15] The estimation period for the OOS results is 10 years, we explain in details how we obtain it in section 4.

[16] For consistency, we adjust also daily factor returns with the ratio of volatilities estimated with monthly observations.

[17] Meaning that we use the full-sample for the in-sample results (including those reported in this section), and the shorter sub-samples for the out-of-sample results.

3.2 Descriptive statistics: Potential for risk and return improvements

For each factor we compute the in-sample alpha (α_s) as the annualized time-series average of $\ddot{F}_{s,t}$. Since the factors are constructed to have zero market beta, the average returns can be interpreted as abnormal returns above the compensation for market risk, or alpha.

The histograms in figure 2 display the full distribution of the alphas in the ESG and financial set[18]. Both distributions reveal significant potential to capture alpha. Indeed, both for financials and ESG factors, about half the factors have an absolute alpha above 5% annual. The alpha distribution for ESG factors has longer tails, indicating the existence of extremely high alphas within the set of ESG factors. Overall, the evidence in figure 2 suggests substantial potential to capture excess returns from ESG factors.

[Figure 2 about here]

We also analyse the potential of our ESG factors to diversify risk when combined into a portfolio. Figure 3 reports the average pairwise correlation for factors that belong to different groups, like traditional ESG ratings, alternative ESG metrics and different types of financial factors. The diagonal elements show the average correlation between the different factors in the same group and the off-diagonal elements show average correlation between factors belonging to two different groups.

ESG factors display low correlation between factors in the same group: the correlation is 0.38 on average across traditional ratings and 0.26 between factors for alternative ESG metrics. This suggests that ESG factors returns are not driven by a single common factor. Instead, they are exposed to different dimensions of risk. The average correlation between factors in the two ESG clusters is actually lower than the one observed for the factors in most of the Published Return Effects clusters. This large dispersion across ESG factors' returns is consistent with previous evidence from the ESG literature that different measures of sustainability tend to disagree (Berg et al., 2022)[19]. We also find that ESG factors tend to have low correlation with financial factors on average – the average pairwise correlation ranges from -0.18 (Alternative ESG factors with Size factors) to 0.31 (Alternative ESG factors with Low risk factors). This means that there is no single financial factor that can completely explain the returns the ESG factors[20]. The correlation structure across factors

[18] Appendix B reports summary descriptive statistics on the factors' CAPM alphas and also for the Fama-French 6 factors alphas. In both cases we find that in-sample analyses would suggests that our selection of ESG factors contains the potential to generate alpha.

[19] The fact that different ESG definitions lead to strong dispersion in returns is not limited to ESG factors, it is reflected also in 'real-world' implementation of ESG strategies. Indeed, Bruno et al. (2024) find that there is strong dispersion in returns of ETFs that integrate ESG information in the portfolio construction process.

[20] Additional analysis in the appendix shows that linear combinations of financial factors cannot fully explain the returns variation in the set of ESG factors. We analyse the minimal number of orthogonal dimensions, i.e., principal components (PCs) needed to explain a given proportion of the covariance matrix the financial factors only and of the Financial plus ESG factors. We find that adding ESG factors one need

shows that returns to ESG factors are not simply redundant with respect to financial factors, suggesting potential for diversification benefits when adding ESG factors to a set of financial factors.

[Figure 3 about here]

In summary, our evidence suggests that the ESG factor set has significant potential for enhancing mean-variance efficiency through either return contribution or risk diversification. Of course, this analysis is not sufficient to confirm the ability of ESG characteristics to improve portfolio efficiency relative to financial factors. For this, we need to build portfolios combining ESG and Financial factors jointly and testing results out-of-sample, as we do in the next section.

Instead of running these out-of-sample portfolio tests, we could simply show single factor or multi factor alphas from such an analysis. This is indeed the approach retained in many papers assessing the performance of selected ESG factors[21]. The results in this section show that testing factors separately, one could easily find some ESG factors that performed well in-sample. Indeed, the large number of ESG characteristics would offer a fertile ground for "*p*-hacking", selecting factors that happen to have significant alphas. To avoid this trap, our analysis in the next section turns to out-of-sample performance when using the entire available information set.

4 Methodology for Constructing Multifactor Portfolios

To test the value of accessing the entire information set including financial factors and ESG factors, we build multi factor portfolios and then evaluate their performance out-of-sample. This section describes the portfolio methods we draw on.

We first consider a case where estimation risk is ignored, before moving to robust portfolio methods that regularise portfolio weights to avoid loading on noise. Our naive base case is the standard mean-variance efficient (MVE) portfolio ($\boldsymbol{w}^*$) which is obtained by maximizing the Sharpe ratio

$$\max_{w} \frac{\boldsymbol{w}'\boldsymbol{\mu}}{\sqrt{\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}}} \tag{5}$$

more independent dimensions to explain a given percentage of variance. For instance, 36 PCs are needed to explain 94% of the variance generated by the financial factors alone, but we need 154 PCs to explain the same level of variance when we add ESG factors to the Financials. This is greater than the number of financial factors, so greater than the number of orthogonal dimensions that the financial factors can mathematically generate.

[21] Papers that test ESG investment strategies in-sample and controlling only for market exposure or for a sparse set of factors are for instance, Gompers et al. (2003), Henriksson et al. (2019), Giese et al. (2021), Faccini et al. (2023), or Coqueret et al. (2025).

Where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the vector of expectations and the covariance matrix of the market-adjusted returns. Using the sample estimates of for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ one obtains the naïve mean-variance multifactor portfolio.

$$\boldsymbol{w}^* = \widehat{\boldsymbol{\Sigma}}^{-1}\widehat{\boldsymbol{\mu}} \tag{6}$$

High dimensionality gives the advantage of a richer set of information, but it also increases the risk of overfitting. One way to reduce overfitting risks is to choose a restricted number of economically motivated factors (Ross, 1976; Fama and French, 2018). However, the literature has developed approaches that allow exploiting the richer information given by a high-dimensional set of factors while keeping estimation risk under control[22]. We draw on several well-documented approaches to construct portfolios that are robust to estimation risk in a high-dimensional setting. We rely on three different approaches to make sure that our results do not depend on a specific choice:

1. Bayesian (or shrinkage) approaches: shrink the portfolio weights by modelling restrictive priors.
2. Constrained mean-variance approaches: limit concentration by imposing constraints directly on the weights, but keeping the estimated expected returns as input.
3. Constrained minimum-volatility approaches: limit concentration by imposing constraints and removing the estimated expected returns from the inputs.

What all these approaches have in common is that they reduce the impact of estimation error on weights by avoiding extreme positions, i.e., imposing some regularisation on weights. Intuitively, these methods impose diversification across factors. We can think of the different methods as imposing different diversification constraints or equivalently, as different ways to impose scepticism about outsized gains that can be generated in-sample.

For each approach, we consider three versions that correspond to different degrees of diversification (weak, moderate, and strong). Again, this allows us to draw general conclusions that do not depend on a particular specification choice.

4.1 Bayesian multifactor portfolios

To mitigate overfitting, the typical Bayesian approach must set the subjective prior beliefs (or priors) that the investor (the economic agent that must solve the portfolio optimization problem) has on the estimated parameters. In other words, we need to model parameter uncertainty[23]. As highlighted in the literature, uncertainty about factor means is the main

[22]Notice that many of the studies in the literature on robust optimization focus on results that impose diversification at stock level. In our case, all our factors are well diversified portfolios, so our problem is not stock-level idiosyncratic risk, but factor's specific risk.

[23]There is a rich literature that studies the use of Bayesian approaches to construct optimal portfolios. Zellner and Chetty (1965), Klein and Bawa (1976), and Brown (1978) were among the first to advocate using subjective return distributions in portfolio choice problems. We will leverage on the methodology and the priors used in Kozak et al. (2020).

source of fragility in the estimation process (see Merton 1980). Therefore, we follow Kozak et al. (2020) (KNS), consider $\boldsymbol{\Sigma}$ as known, and focus on the following family of priors,

$$\boldsymbol{\mu} \sim \mathcal{N}\left(\mathbf{0}, \frac{\kappa^2}{\tau}\boldsymbol{\Sigma}^{\eta}\right) \tag{7}$$

Where τ is the trace of the covariance matrix of the market-adjusted returns ($\boldsymbol{\Sigma}$) and κ is a constant controlling the scale of $\boldsymbol{\mu}$[24]. In practice, this prior means that the in-sample estimates of the expected market-adjusted factor returns are shrunk towards the CAPM prediction ($\mathbf{0}$), and the confidence in the CAPM prediction is lower the greater the observed deviations from the CAPM predictions ($\boldsymbol{\Sigma}$). The intensity of the shrinkage is regulated by κ and η. Several papers have made different choices on η. For instance, Harvey et al. (2008) have used $\eta = 0$, while Pastor (2000) and Pastor et al. (2000) have used $\eta = 1$. As KNS show, these choices imply large expected Sharpe ratios for the low-eigenvalue principal components (PCs) of $\boldsymbol{\Sigma}$. This is economically implausible because, as shown in Kozak et al. (2018), large Sharpe ratios for low eigenvalue PCs translate in near-arbitrage opportunities (possibility of implausibly large Sharpe ratios). To guarantee the absence of near-arbitrage opportunities KNS shows that one must impose $\eta \geq 2$. Following KNS, we choose $\eta = 2$, because it gives us the least restrictive ('flattest') Bayesian prior that guarantees economic plausibility (absence of near-arbitrage opportunities). With $\eta = 2$ the optimal solution is

$$\widehat{\boldsymbol{w}} = (\boldsymbol{\Sigma} + \gamma \boldsymbol{I})^{-1}\, \overline{\boldsymbol{\mu}} \tag{8}$$

Where $\overline{\boldsymbol{\mu}}$ is the vector of estimated market-adjusted returns, and $\gamma = \frac{\tau}{\kappa^2 T}$ with T being the number of observations in the estimation sample. KNS also show that (8) maps into a penalized estimator that resembles the ridge estimator, with γ being the penalization parameter. Therefore, the penalization is inversely proportional to κ, which can be interpreted as the maximum expected root expected maximum squared Sharpe ratio[25] for $\eta = 2$,

$$E\left[\widehat{\boldsymbol{w}}'\boldsymbol{\Sigma}\widehat{\boldsymbol{w}}\right]^{1/2} = \kappa \tag{9}$$

This means that the choice of $\eta = 2$ also facilitates the economic interpretability of the parameter that regulates the shrinkage (κ). Intuitively, setting a low κ corresponds to investors who do not believe that large deviations of expected returns from the CAPM can be predicted *ex-ante*, which means that they will discount large in-sample Sharpe ratios as spurious and apply shrinkage. To set κ we follow three different approaches.

Weak (heuristic-based): Following the argument from Ross (1976) we impose the maximum Sharpe ratio to be equal to two times the market Sharpe ratio, as higher Sharpe ratios are considered unreasonable (see also MacKinlay 1995 and Cochrane and Saa-Requejo 2000).

[24] Under specific parameter conditions, KNS show that κ can be interpreted as the maximum achievable Sharpe ratio.

[25] Another way to interpret κ is the variance of the Stochastic Discount Factor (SDF), which has been shown in Hansen and Jagannathan (1991) to represent the upper bound on the achievable Sharpe ratio.

Moderate (data-driven): We use historical data preceding our testing sample to find the κ that maximizes out-of-sample predictive performance (approach used by Kozak et al. 2020). We find that the optimal κ (0.16) corresponds to about 36% of the market Sharpe ratio[26].

Strong (extreme): We also consider the extreme case in which the investor spreads the weights equally across all the factors included in the selection[27]. This can be considered an extreme form of shrinkage[28].

4.2 Constrained multifactor portfolios

A popular approach into control estimation risk is to impose constraints on the portfolio weights[29]. The objective of these constraints is similar to using priors in the Bayesian approach. The idea is that constraints should prevent extreme positions that can be driven by spurious in-sample realizations. The difference relative to the Bayesian approach is that instead of targeting directly a pre-determined maximum Sharpe ratio, we impose a minimum level of diversifications across the factors.

We consider two different objective functions. For the constrained mean-variance approach the objective is to maximize the Sharpe ratio of the multifactor portfolio, while for the minimum volatility approach the objective is to minimize volatility of the portfolio. In both cases, we optimize the objective using three different sets of constraints, which give us a total of six distinct constrained optimization approaches. First, we constraint the sign of the portfolios weights to be consistent with the sign of the average factor returns over the estimation period (directional constraint). Second, we impose commonly used constraints that the sum of all weights sums up to one (full investment constraint) and that weights are non-negative[30]. These constraints also ensure that all the constrained portfolios have the same level of leverage[31]. Finally, we impose a minimum level of diversification across factors with an inequality constraint on the effective number of factors (ENF):

$$\sum_{i}^{N_K} \frac{1}{w_i^2} \geq \theta \tag{10}$$

We obtain results for three levels of diversification constraint: weak ($\theta = 10$), moderate

[26]We report full details on the procedure in the appendix.

[27]The position in each factor is such that the investor goes long on factors that had a positive return in-sample and short on factors that had a negative return in-sample.

[28]Indeed, we show in the appendix that to replicate the in-sample Sharpe ratio of this portfolio using equation 9, we would need to set the max achievable Sharpe ratio (κ) very close to zero. MacKinlay and Pastor (2000, p.906) also argue that an Equally Weighted portfolio can be interpreted as a portfolio that imposes a mean-variance link under specific priors. For instance, shrinking the covariance matrix towards a diagonal matrix and the means to be proportional to diagonal elements of the matrix.

[29]See for instance Jagannathan and Ma (2003) and DeMiguel et al. (2009a).

[30]See for instance Jagannathan and Ma (2003).

[31]Different levels of leverage do not affect the comparisons of Sharpe ratios as the Sharpe ratio is homogenous of degree 0 with respect to leverage. However, leverage does affect the comparison of volatilities. Because in our tests on the effect of ESG on portfolio volatilities we only use constrained portfolios, these constraints ensure use that the effects that we find are not due to differences in leverage.

($\theta = 30$), and strong ($\theta = 50$). This constraint directly controls the level of concentration in the multifactor portfolio, making sure that the portfolio does not take extreme positions in one or few factors. Again, testing portfolios with three different minimum thresholds for the effective number of factors allows us to assess the value-added of ESG information across different portfolio specifications.

5 Empirical Analysis of Multifactor Portfolios

In this section we analyse the investment value of ESG information from a multifactor portfolio perspective in our stylized setting[32]. We set out to answer three questions. First, does ESG information provide robust improvements in mean-variance portfolio efficiency? More specifically, we test whether combining ESG factors to financial factors in an ex-ante optimal way improves the Sharpe ratio of multifactor portfolios out-of-sample. Second, we want to understand how ESG information influences the composition of the multifactor portfolios. Which ESG factors matter the most? Are there notable differences between green and brown tilts? How do positions in alternative ESG factors compare to traditional ESG factors? Finally, using portfolio approaches that focus on risk reduction, we test whether ESG information helps reduce risk.

5.1 Assessing the incremental value of ESG information

Our goal is to investigate whether ESG information has incremental value relative to Financial information from the perspective of portfolio efficiency. From the perspective of portfolio efficiency, integrating ESG factors in portfolio construction is valuable only if it improves the Sharpe ratio relative to using financial factors only.

For two generic sets of factors $\boldsymbol{F}_K$ and $\boldsymbol{F}_L$, Barillas and Shanken (2017) show that one can test whether $\boldsymbol{F}_K$ improves portfolio efficiency relative to $\boldsymbol{F}_L$ comparing the Sharpe ratios of the respective portfolios – the portfolio that maximizes the Sharpe ratio for that specific set of factors. In other words, we want to test

$$SR\left(P\left(\boldsymbol{F}_K\right)\right) > SR\left(P\left(\boldsymbol{F}_L\right)\right) \tag{11}$$

In our setting, consider that $\boldsymbol{F}_K$ includes financial and ESG information while $\boldsymbol{F}_L$ only includes financial information. When the set of factors in $\boldsymbol{F}_K$ achieves a higher Sharpe ratio than the set of factors in $\boldsymbol{F}_L$, it means that the factors in $\boldsymbol{F}_L$ do not span in factors in $\boldsymbol{F}_K$. In other words, $\boldsymbol{F}_K$ contains factors that are non-redundant relative to the factors contained in $\boldsymbol{F}_L$.

This setting is used widely when assessing the value of ESG information. What differentiates our work is that we consider a high-dimensional set of ESG factors instead of one ESG

[32]Our question is that of the informational value of ESG information in a stylised setting, not the development of realistic portfolio strategies. While we focus on out-of-sample results and consider a reasonably investable universe of stocks, we ignore many important implementation considerations. Most notably, we do not account for transaction costs and do not impose time consistent cross validation samples.

factor or few ESG factors when running this comparison. Perhaps even more importantly, comparisons of Sharpe Ratios are typically conducted in-sample. In contrast, we run out-of-sample tests[33]. The out-of-sample perspective is important since both the in-sample Sharpe ratio of the portfolio using only financials and that of the portfolio integrating financials and ESG factors contain perfect hindsight on risk and return parameters. A more relevant question is how Sharpe ratios compare in the presence of estimation risk.

We compare Sharpe ratios of multifactor portfolios constructed with different sets of factors. As a benchmark, we use only the financial factors to construct portfolios[34]. We then create portfolios that use both financial and ESG factors. Comparing the resulting Sharpe ratios allows concluding on the incremental value of ESG information. As an aside, we also provide results for portfolios that invest only in ESG factors.

To compute the p-values of the difference between Sharpe ratios of the multifactor portfolios with their benchmark, we use monthly time series of the volatility-scaled market adjusted returns to which we apply the Ledoit and Wolf (2008) approach. This approach relies on the stationary block bootstrap procedure proposed in Politis and Romano (1994)[35] which, unlike a classical bootstrap, allows us to account for possible autocorrelation in the time-series of our multifactor portfolios. We also compute p-values associated with the test whether Sharpe ratios are different from zero.

5.2 In and out-of-sample testing

We compute in-sample and out-of-sample performance statistics of the multifactor portfolios. To compute the in-sample results, the estimation sample for the inputs of our multifactor portfolios – the average factor returns $\widehat{\boldsymbol{\mu}}$ and the covariance matrix $\widehat{\boldsymbol{\Sigma}}$ – corresponds to the sample used for testing. In particular, we use the full sample of 15 years from August 2008 to June 2023. This means that we compute the weights only once and keep the weights of the multifactor portfolios constant across each month in the sample multiplying the weights by the matrix of factor returns over the full sample. This gives us 15 years of in-sample returns for each multifactor portfolio.

For the out-of-sample tests, the period used for estimation must be separate from the one used for performance evaluation. We use a 3-fold division, splitting the full sample in three sub-samples of five years. For each sub-sample we use the other two samples (representing a period of ten years) to estimate the inputs of the multifactor portfolio and compute the

[33]We do report in-sample performance for completeness, but we focus primarily on out-of-sample results.

[34]However, we also run our tests using a sparse set of factors, which contains the cross-sectional factors of the Fama and French (2018) model – SMB, HML, RMW, CMA, and Momentum. The results reported in the appendix, show that our conclusions on the incremental value of ESG information are robust to the use of a sparse set of factors as a benchmark.

[35]We run our bootstraps extracting 5,000 paths, and using an expected block length of 10 months, which is consistent with the length used in Kan et al. (2024). We estimate the optimal expected length of the block applying the Ledoit and Wolf (2008) procedure to the difference between the in-sample Financials + ESG multifactor naïve portfolio and the in-sample Financials multifactor naïve portfolio. Then to maintain consistency we use the same expected block length for all the other tests. However, we have checked that our results are also robust to using 1 month as expected block length, which corresponds to the IID bootstrap (results made available upon request).

portfolio weights. We keep the weights constant across each sub-sample multiplying them by the matrix of factor returns for that specific sub-sample. We repeat the operation for each sub-sample, and this gives us 15 years of out-of-sample returns for each our multifactor portfolios. Figure 4 illustrates this approach.

[Figure 4 about here]

5.3 Empirical results for multifactor portfolios

5.3.1 Increments in Sharpe ratios

Does including ESG factors improve the Sharpe ratio of portfolios? In table 2 we report in-sample and out-of-sample differences in Sharpe ratios between portfolios including ESG factors and their benchmark which only contains financial factors. Panel A shows in-sample results while Panel B shows out-of-sample results. In parentheses, we show the p-value associated with the Sharpe ratio difference, as obtained with the Ledoit and Wolf (2008) test. The first row of each panel reports results for portfolios constructed using only ESG factors, while the second row shows the results when using both financial and ESG factors.

We report results for portfolios across ten different portfolio construction approaches. These include one portfolio for naïve optimisation (1) and nine portfolios using robust portfolio construction: three portfolios representing the Bayesian shrinkage approach (2,3, and 4), three constrained mean-variance portfolios (5,6, and 7), and three constrained minimum-volatility portfolios (8,9 and 10). There are three different portfolios for each approach because we look at various levels of constraints.

The benchmark for each portfolio applies the same portfolio construction approach to the set of financial factors. Using the same portfolio construction approach ensures comparability between each portfolio and its benchmark: the only difference is the information set used. The last row in each panel reports the level of the Sharpe ratio of the benchmark as a reference point.

Focusing on the in-sample results, we find that using ESG factors alone is not able to generate significant improvements in Sharpe ratios relative the benchmark, no matter the portfolio construction used. However, the lack of large differences with the benchmark also indicates that multifactor portfolios using only ESG information perform about as well as portfolios that use only financial information.

Our main interest is in the portfolios that combine financial and ESG factors to assess the incremental improvements from using ESG information. Here we observe several instances of large and statistically significant improvements in in-sample Sharpe ratios.

By construction, the naïve approach achieves the largest possible Sharpe ratio in-sample. Since adding more factors increases the set of possible solutions, the combined portfolio must by construction have a Sharpe Ratio greater or equal than the portfolio using only financial factors. Adding ESG to the financial factors increases the annualised Sharpe ratio by a magnitude of 0.75 in-sample – clearly an economically important effect. This shows that the ESG factors are not spanned by the financial factors. When imposing a lower bound on the effective number of factors in mean variance and minimum variance portfolios, adding ESG information also leads to significant improvements of in-sample Sharpe ratios. Only in the case of Bayesian shrinkage, where diversification levels are fixed rather than set as a lower bound, do we fail to observe in-sample improvements. In the appendix we show that Bayesian shrinkage imposes the strongest level of diversification among the methods used[36]. These in-sample results are of limited relevance, however. Given that there is estimation risk, out-of-sample results are more relevant.

The out-of-sample changes in Sharpe ratio are shown in Panel B. The results show that out-of-sample, changes in Sharpe ratios are insignificant across the host of portfolio construction methods we employ. Adding ESG factors does not provide any improvements out-of-sample. We can conclude that the in-sample improvements are just an illusion. They disappear when estimation risk is present, even when using portfolio methods designed to tackle this risk optimally. The additional information inherent in ESG Factors does not deliver any benefits out-of-sample[37].

[Table 2 about here]

5.3.2 Out-of-sample performance decay

The stark difference between in-sample and out-of-sample results is evidence of the over-fitting risk that investors face. Of course, estimation error is always an issue for portfolio construction, as returns of financial assets are inherently noisy. Kan et al. (2024) show that the out-of-sample decay in risk-adjusted performance of portfolios is a relevant indicator for the level of noise.

Table 3 shows how Sharpe ratios change from the in-sample analysis to the out-of-sample analysis for the different portfolios we analyse. Panel A reports the in-sample annualised Sharpe ratios of the multifactor portfolios and Panel B reports out-of-sample Sharpe ratios. Panel C report the out-of-sample gap which is just the difference between in-sample and out-of-sample results.

The results in Panel C show substantial degradation of Sharpe ratio when going out of sample for the naïve portfolios. This is expected as these portfolios do not impose any diversification constraints and thus load maximally on estimation error. The robust portfolio methods lead to a substantial reduction in out-of-sample decay, with values corresponding to about 20 percent to 70 percent of performance decay observed for the naïve portfolios. This shows

[36] We use a measure of diversification that considers not just the distribution of the multifactor portfolio weights across factors but also how the factors correlate to each other. This measure tells us how the weights of the portfolio are distributed across different risk contributors. We find that the Shrinkage approaches generate the portfolios that are the most diversified across different sources of risk.

[37] In the appendix we also show that ESG factors cannot even provide statistically significant improvements relative to a sparse set of factors (SMB, HML, RMW, CMA, and Momentum).

that the portfolio methods successfully tame large amounts of estimation risk inherent in the high-dimensional datasets we analyse.

To assess the effectiveness of the robust portfolio methods, it is also worth considering the *level* of Sharpe ratio of the portfolios combining ESG and financial factors shown in Panel B. Portfolios using Bayesian shrinkage and constrained mean variance portfolios obtain Sharpe ratios that are significantly higher than zero, despite drawing on a set of 352 factors (130 financial factors and 222 ESG). This confirms that these methods perform well when faced with a high-dimensional dataset - even though it contains substantial amounts of noise as indicated by the large Sharpe ratio decay in naïve portfolios.

[Table 3 about here]

5.3.3 What are the ESG tilts in optimal portfolios?

Despite the lack of incremental value added by ESG information out-of-sample, it is interesting to look at the composition of the portfolios. Portfolio composition tells us what would have been the optimal direction of tilts for an investor.

Going back to the idea of "rational sustainability" it could be the case that an investor considers ESG information purely to target a high Sharpe ratio and ends up with substantial green tilts in his portfolio. However, he could also end up with brown tilts if the data-driven portfolio method determines these to be optimal, in which case the portfolio would go against positive sustainability – at least on some dimensions. To provide some perspective on whether financially motivated investors would favour green tilts, we compare the green and the brown tilts in the portfolios we construct. In addition, we compare the weights in the traditional and in the alternative ESG factors to provide perspective on which type of data matters most when using ESG information to target a high Sharpe ratio.

To gain perspective on how the different information sets influence portfolio composition, we regroup factors into the following five categories:

- *Financial factors*: includes all the financial factors.
- *Traditional green factors*: the traditional ESG factors that enter in the multifactor portfolio with a positive sign
- *Traditional brown factors*: the traditional ESG factors that enter the multifactor portfolio with a negative sign
- *Alternative green factors*: the alternative ESG factors that enter in the multifactor portfolio with a positive sign
- *Alternative brown factors*: the alternative ESG factors that enter in the multifactor portfolio with a negative sign

For each category we report the sum of the absolute weights for the relevant factors. In our analysis above, we have three sets of weights of ex-ante portfolios that we construct for our out-of-sample analysis. To focus on a single set of weights, here we analyse the portfolios constructed with the full sample. These reflect the optimal tilts an investor would hold when building a portfolio at the end of the sample period that would be optimal ex-ante for the subsequent period. Results for the three sets of weights constructed for the different sub-samples are available in the appendix. To facilitate comparison of weights across portfolio methods, we normalise weights of the Bayesian shrinkage portfolios by leverage so that they sum to 100% like the other portfolios.

We report the portfolio compositions in figures 5a and 5b. Figure 5a focusses on the importance of green and brown ESG tilts compared to tilts to financial metrics. We find that all our portfolios allocate large weights to the ESG factors – between 30% and 65%. The portfolios incorporate both green and brown positions. Neither green nor brown positions dominate. This result shows that a purely pecuniary motivation to use ESG information does not lead to favouring green tilts. Instead, it leads to exploiting both green and brown tilts based on their empirical risk and return characteristics.

[Figure 5a about here]

Figure 5b focusses on comparing the weights given to alternative vs. traditional ESG metrics. We note that allocations to alternative ESG factors match and often exceeds those to traditional ESG factors. This suggests that the information captured by alternative ESG data is important relative to what is captured from traditional ESG information.

[Figure 5b about here]

Table 4 looks at factors with the largest absolute value in each of our multifactor portfolios.

Among the traditional ESG factors, some of the Moody's ratings tend to dominate the top positions with large positive tilts to the human rights rating and negative tilts to the rating for employee training programs. However, imposing strong shrinkage skews the portfolios toward certain factors based on consensus ratings (CSR Hub) for themes like human rights and product safety.

Among the alternative ESG factors, the most consistent result is a brown position on the sin stocks factor for all portfolios that use average returns as input[38]. We also find that the water score provided by GaiaLens tends to obtain a large green tilt across most portfolios. Commercial ratings that use a systematic NLP approach (Covalence) appear more prominently in the MinVol portfolios, suggesting that they may be more useful to diversify risk.

[Table 4 about here]

[38]This is consistent with the large alpha of the sin stock factor over our entire sample (reported in the appendix).

5.3.4 Does ESG information help to reduce risk?

In this last part of our analysis, we focus exclusively on the risk dimension. This is important as risk mitigation is a key motivation for integrating ESG factors portfolio decisions. In a recent survey Larcker et al. (2025) found that 61% of institutional investors believe that integrating ESG criteria in portfolio construction can help reduce volatility.

To investigate whether ESG factors do indeed help to reduce risk, we focus on multifactor portfolios that exclusively target risk reduction in the objective function. Therefore, we focus on our constrained minimum volatility portfolios. To test differences in volatility we use the Ledoit and Wolf (2011) approach, which is also based on the stationary block bootstrap of Politis and Romano (1994)[39].

We report the results in table 5. We find that adding ESG factors to the financial factors does not necessarily help reduce portfolio volatility - even in-sample. The only case where adding ESG factors reduces volatility significantly in-sample is when we impose strong diversification constraints (effective number of factors of at least 50). When imposing weaker diversification constraints, volatility reduction is not significant even in in-sample analysis. Out-of-sample, the risk reduction observed in-sample breaks down even for the most diversified portfolio, suggesting that correlation patterns that help reduce risk in-sample do not persist out of sample.

[Table 5 about here]

5.3.5 Robustness tests

In the appendix, we report a battery of robustness tests designed to verify that our main conclusions are not driven by particular factor selection or specification choices.

First, we assess robustness with respect to alternative subsets of ESG factors (see Appendix D.3). The goal of this analysis is to ensure that the lack of incremental out-of-sample value from ESG characteristics is not specific to the inclusion of certain factors. We consider four distinct ESG subsets: (i) Consensus ESG factors based on CSRHub scores,[40] (ii) ESG factors excluding aggregate pillar scores,[41] (iii) our selection of Traditional ESG factors, and finally (iv) our selection of Alternative ESG factors. In all these specifications, we obtain results that corroborate our main findings. Virtually all ESG subsets improve the in-sample Sharpe ratio, yet none of these improvements persist out of sample.

Second, we explore the robustness of our findings to alternative specifications of the financial benchmark (see Appendix D.4). In one specification, we restrict our benchmark to a sparse set of five standard cross-sectional factors from Fama and French (2018)—size (SMB), value (HML), profitability (RMW), investment (CMA), and momentum (MOM). In another specification, we adopt the high-dimensional set of financial factors proposed by Jensen et al.

[39] In this case, we impose an expected lag of the block of 10 months.

[40] CSRHub scores aggregate ratings across various ESG providers.

[41] We remove ESG scores that represent macro ESG pillars such as ESG, E, S, and G.

(2023), which has recently gained popularity in the literature.[42] Across both of these benchmarks, our main conclusion remains intact: ESG characteristics fail to produce statistically significant improvements in out-of-sample Sharpe ratios.

6 Conclusion

This study evaluates the incremental value of ESG information over financial metrics when constructing risk-return efficient portfolios. We depart from treating ESG as a monolith, instead allowing portfolios to draw flexibly on a high-dimensional set of ESG metrics. To avoid overfitting in this high-dimensional setting, we employ robust portfolio construction approaches, and we employ different portfolio methods to obtain general conclusions.

Our findings reveal a stark contrast between in-sample and out-of-sample performance. ESG factors demonstrate significant improvements in sample, except in the most constrained portfolios using Bayesian approaches. However, these improvements disappear out-of-sample, with no significant performance improvements across a variety of robust portfolio methods.

These findings highlight that – even though they provide additional information - ESG metrics may not improve out-of-sample performance. In the out-of-sample setting estimation risk cancels any improvements in risk-adjusted performance, even when using robust portfolio methods that avoid overfitting.

Despite failing to add value out-of-sample, ESG characteristics receive substantial portfolio allocations ex-ante. Optimal portfolios select both green and brown tilts of similar magnitude, contradicting the assumption that incorporating ESG information necessarily means adopting positions that are favourable to sustainability. In a nutshell, these results show that accounting for more information via ESG criteria does not lead to investing more sustainably. We also note that alternative ESG factors get weights of similar magnitude as traditional ratings, underlining the relevance of new approaches to ESG data generation.

Our results also provide a perspective on claims that ESG information is particularly suited for risk management. We show that ESG factors do not help reduce portfolio risk out-of-sample. This challenges a premise underlying many ESG investment strategies.

Overall, our findings show that – despite the massive amount of available information – ESG characteristics provide no informational advantage for constructing risk-return efficient portfolios beyond what financial information already delivers. Unlike previous research, this finding does not depend on considering a particular definition of ESG. Instead, it results from using a vast amount of ESG information from a highly granular dataset to build robust portfolios. Our results also illustrate that out-of-sample assessment leads to substantially different conclusions from traditional mean variance spanning tests.

[42] As emphasized by Bessembinder et al. (2025), commonly used high-dimensional factor sets often do not span each other.

References

Avramov, Doron, Si Cheng, Abraham Lioui, and Andrea Tarelli, 2022, Sustainable investing with esg rating uncertainty, *Journal of Financial Economics* 145, 642–664.

Barillas, Francisco, and Jay Shanken, 2017, Which alpha?, *Review of Financial Studies* 30, 1316–1338.

Barillas, Francisco, and Jay Shanken, 2018, Comparing asset pricing models, *Journal of Finance* 73, 715–754.

Berg, Florian, Julian F. Kölbel, and Roberto Rigobon, 2022, Aggregate confusion: The divergence of esg ratings, *Review of Finance* 26, 1315–1344.

Bessembinder, Hendrik, Aaron Burt, and Christopher Hrdlicka, 2025, Do popular high-dimensional factor sets span each other?, Working Paper 5150471, SSRN.

Bolton, Patrick, and Marcin Kacperczyk, 2023, Global pricing of carbon-transition risk, *Journal of Finance* 78, 3677–3754.

Brown, Stephen J., 1978, The portfolio choice problem: Comparison of certainty equivalence and optimal bayes portfolios, *Communications in Statistics—Simulation and Computation* 7, 321–334.

Bruno, Giovanni, Mikheil Esakia, and Felix Goltz, 2024, From esg confusion to return dispersion: Fund selection risk is a material issue for esg investors, Scientific Beta White Paper.

Cochrane, John H., and Jesús Saa-Requejo, 2000, Beyond arbitrage: Good-deal asset price bounds in incomplete markets, *Journal of Political Economy* 108, 79–119.

Coqueret, Guillaume, Thomas Giroux, and Olivier David Zerbib, 2025, The biodiversity premium, *Ecological Economics* 228, 108435.

DeMiguel, Victor, Lorenzo Garlappi, Francisco J. Nogales, and Raman Uppal, 2009a, A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms, *Management Science* 55, 798–812.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal, 2009b, Optimal versus naive diversification: How inefficient is the 1/{N} portfolio strategy?, *Review of Financial Studies* 22, 1915–1953.

Edmans, Alex, 2011, Does the stock market fully value intangibles? employee satisfaction and equity prices, *Journal of Financial Economics* 101, 621–640.

Edmans, Alex, 2023, The end of esg, *Financial Management* 52, 3–17.

Edmans, Alex, 2024, Rational sustainability, *Journal of Applied Corporate Finance* Forthcoming.

Edmans, Alex, Tom Gosling, and Dirk Jenter, 2025, Sustainable investing in practice: Objectives, beliefs, and limits to impact, FEB-RN Research Paper No. 18/2024; ECGI Finance Working Paper No. 1028/2024. Last revised: August 12, 2025.

Edmans, Alex, Darcy Pu, Chendi Zhang, and Lucius Li, 2024, Employee satisfaction, labor market flexibility, and stock returns around the world, *Management Science* 70, 4357–4380.

Eskildsen, Marc, Markus Ibert, Theis Ingerslev Jensen, and Lasse Heje Pedersen, 2024, In search of the true greenium, SSRN Working Paper.

Faccini, Renato, Rastin Matin, and George Skiadopoulos, 2023, Dissecting climate risks: Are they reflected in stock prices?, *Journal of Banking & Finance* 155, 106948.

Fama, Eugene F., and Kenneth R. French, 2018, Choosing factors, *Journal of Financial Economics* 128, 234–252.

Giese, Guido, Zoltán Nagy, and Linda-Eling Lee, 2021, Deconstructing esg ratings performance: Risk and return for e, s, and g by time horizon, sector, and weighting, *Journal of Portfolio Management* 47, 94–111.

Gompers, Paul, Joy Ishii, and Andrew Metrick, 2003, Corporate governance and equity prices, *Quarterly Journal of Economics* 118, 107–156.

Hansen, Lars Peter, and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225–262.

Harvey, Campbell, John Liechty, and Merrill Liechty, 2008, Bayes vs. resampling: A rematch, *Journal of Investment Management* 6, 29–45.

Harvey, Campbell R., Yan Liu, and Heqing Zhu, 2016, ... and the cross-section of expected returns, *Review of Financial Studies* 29, 5–68.

Henriksson, Roy, Joshua Livnat, Patrick Pfeifer, and Margaret Stumpp, 2019, Integrating esg in portfolio construction, *Journal of Portfolio Management* 45, 67–81.

Hoberg, Gerard, and Asaf Manela, 2025, The natural language of finance, *Foundations and Trends in Finance* 14, 244–365.

Hong, Harrison, and Marcin Kacperczyk, 2009, The price of sin: The effects of social norms on markets, *Journal of Financial Economics* 93, 15–36.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* 28, 650–705.

Jagannathan, Ravi, and Tongshu Ma, 2003, Risk reduction in large portfolios: Why imposing the wrong constraints helps, *Journal of Finance* 58, 1651–1683.

Jensen, Theis Ingerslev, Bryan Kelly, and Lasse Heje Pedersen, 2023, Is there a replication crisis in finance?, *Journal of Finance* 78, 2465–2518.

Kan, Raymond, Xiaolu Wang, and Xinghua Zheng, 2024, In-sample and out-of-sample sharpe ratios of multi-factor asset pricing models, *Journal of Financial Economics* 155, 103837.

Kan, Raymond, and Guofu Zhou, 1999, A critique of the stochastic discount factor methodology, *Journal of Finance* 54, 1221–1248.

Kan, Raymond, and Guofu Zhou, 2007, Optimal portfolio choice with parameter uncertainty, *Journal of Financial and Quantitative Analysis* 42, 621–656.

Khan, Mozaffar, George Serafeim, and Aaron Yoon, 2016, Corporate sustainability: First evidence on materiality, *Accounting Review* 91, 1697–1724.

Klein, Roger W., and Vijay S. Bawa, 1976, The effect of estimation risk on optimal portfolio choice, *Journal of Financial Economics* 3, 215–231.

Kogan, Leonid, and Mary H. Tian, 2015, Firm characteristics and empirical factor models: A model-mining experiment, International finance discussion paper, Federal Reserve Board.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2018, Interpreting factor models, *Journal of Finance* 73, 1183–1223.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2020, Shrinking the cross-section, *Journal of Financial Economics* 135, 271–292.

KPMG International, 2020, Moving towards common sustainability metrics, White Paper.

Larcker, David F., Laura-Elena Lee, Amit Seru, and Brian Tayan, 2025, 2024 institutional investor survey on sustainability.

Ledoit, Olivier, and Michael Wolf, 2008, Robust performance hypothesis testing with the sharpe ratio, *Journal of Empirical Finance* 15, 850–859.

Ledoit, Olivier, and Michael Wolf, 2011, Robust performance hypothesis testing with the variance, *Wilmott* 86–89.

MacKinlay, A. Craig, 1995, Multifactor models do not explain deviations from the capm, *Journal of Financial Economics* 38, 3–28.

MacKinlay, A. Craig, and Lubos Pastor, 2000, Asset pricing models: Implications for expected returns and portfolio selection, *Review of Financial Studies* 13, 883–916.

McLean, R. David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *Journal of Finance* 71, 5–32.

Merton, Robert C., 1980, On estimating the expected return on the market: An exploratory investigation, *Journal of Financial Economics* 8, 323–361.

Novy-Marx, Robert, and Mihail Velikov, 2016, A taxonomy of anomalies and their trading costs, *Review of Financial Studies* 29, 104–147.

Pastor, Lubos, 2000, Portfolio selection and asset pricing models, *Journal of Finance* 55, 179–223.

Pastor, Lubos, Robert F. Stambaugh, and Lucian A. Taylor, 2000, Comparing asset pricing models: An investment perspective, *Journal of Financial Economics* 56, 335–381.

Pastor, Lubos, Robert F. Stambaugh, and Lucian A. Taylor, 2025, Sustainable investing, SSRN Working Paper, Available at SSRN.

Pedersen, Lasse Heje, Shaun Fitzgibbons, and Lukasz Pomorski, 2021, Responsible investing: The esg-efficient frontier, *Journal of Financial Economics* 142, 572–597.

Politis, Dimitris N., and Joseph P. Romano, 1994, The stationary bootstrap, *Journal of the American Statistical Association* 89, 1303–1313.

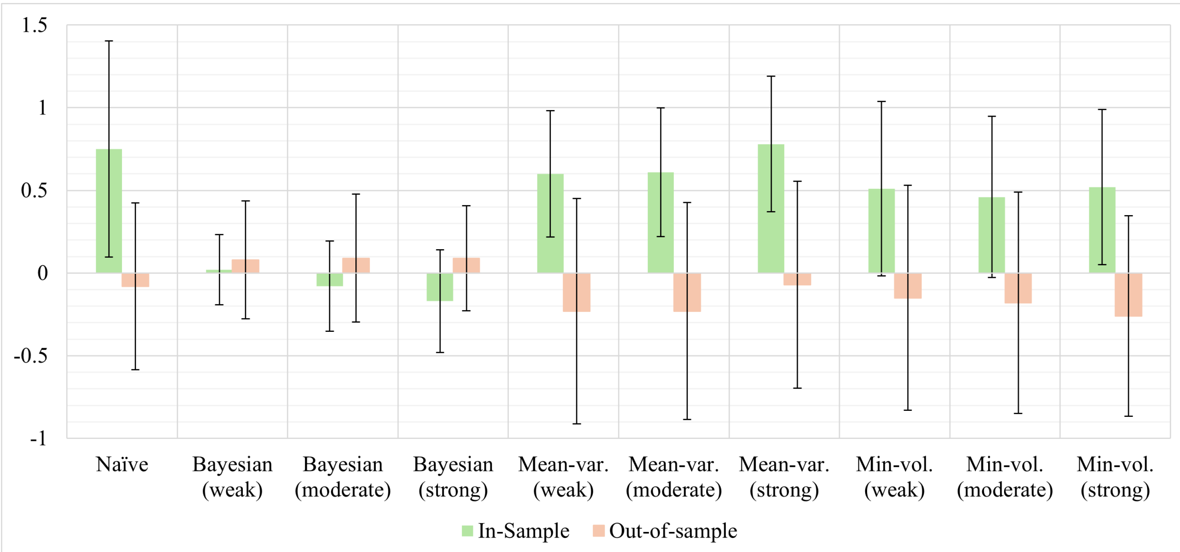
Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Sautner, Zacharias, Laurence Van Lent, Grigory Vilkov, and Ruishen Zhang, 2023, Firm-level climate change exposure, *Journal of Finance* 78, 1449–1498.

Zellner, Arnold, and V. Karuppan Chetty, 1965, Prediction and decision problems in regression models from the bayesian point of view, *Journal of the American Statistical Association* 60, 608–616.

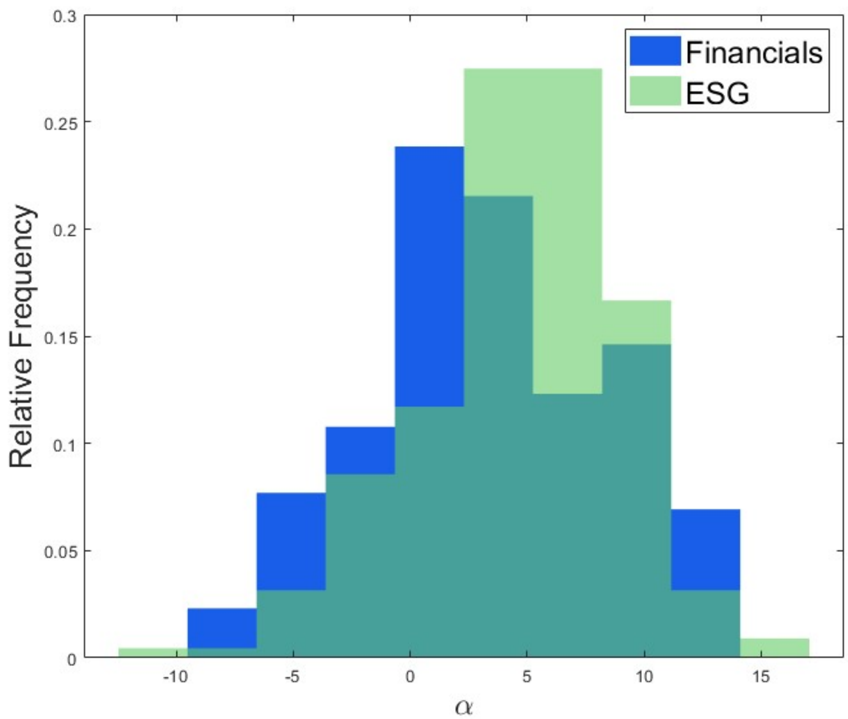
Figures and Tables

Figure 1: Incremental value (Sharpe ratio increase) of ESG information



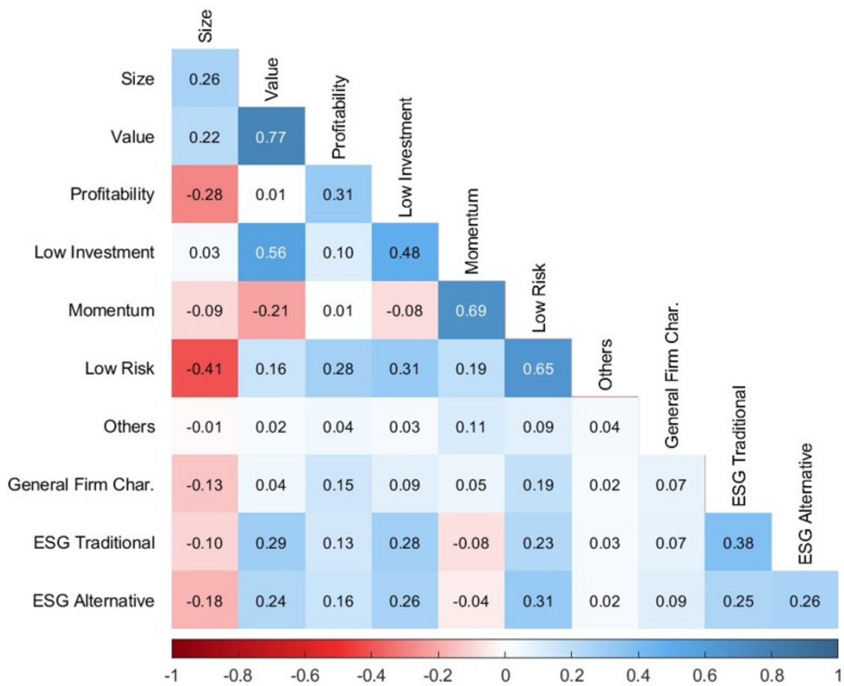
The chart reports differences in annualized Sharpe ratios (ΔSR) between various multifactor portfolios and their respective benchmarks, along with the bootstrapped 90%-confidence intervals of ΔSR. The multifactor portfolios are constructed using the set of factors that contains both ESG and Financial factors and 10 portfolio construction approaches - 1 Naive (no control for estimation error) and 9 robust approaches with three different methodologies (Bayesian, constrained Mean-variance, and constrained minimum-volatility) each with three levels of control for estimation error (weak, moderate, and strong). For each multifactor portfolio the benchmark is constructed using the same portfolio construction approach but using only the Financial factors. Green (orange) bars represent in-sample (out-of-sample) results. We compute 90%-confidence intervals via the stationary block bootstrap of Ledoit and Wolf (2008).

Figure 2: Distribution of alpha by factor category (annualized percent values)



This figure shows the relative frequency of the alphas of financial factors (blue area) and ESG factors (green area). Alphas are obtained as the average of the market adjusted factor returns ($F_{i,t} - \beta_i R^e_{M,t}$), scaled to have volatility equal to the volatility of the market returns in excess of the risk-free rate. We also annualise and multiply by 100 the alphas. We use monthly observations from August 2008 to June 2023.

Figure 3: Average pairwise correlation within and across groups of factors



This figure shows the average pairwise correlation across adjusted factors returns within and between 12 categories of factors. Adjusted factor returns are the factor returns minus the returns explained by their market exposure ($F_{i,t} - \beta_i R^e_{M,t}$), scaled to have volatility equal to that of the market returns in excess of the risk-free rate ($R^e_{M,t}$). We report the average Pearson correlation of the adjusted returns of each factor with each other factor within the same factor category (diagonal elements) or with factors in other factor categories (off-diagonal elements). Sample is from August 2008 to June 2023.

Figure 4: Representation of the in and out of sample periods

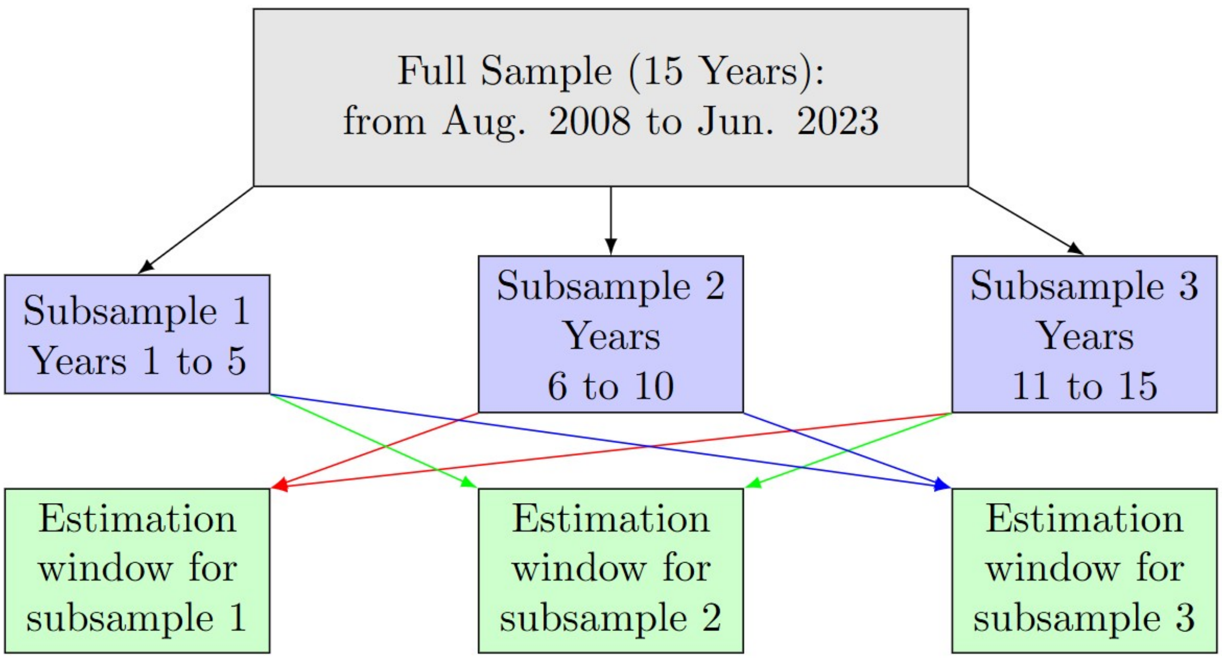
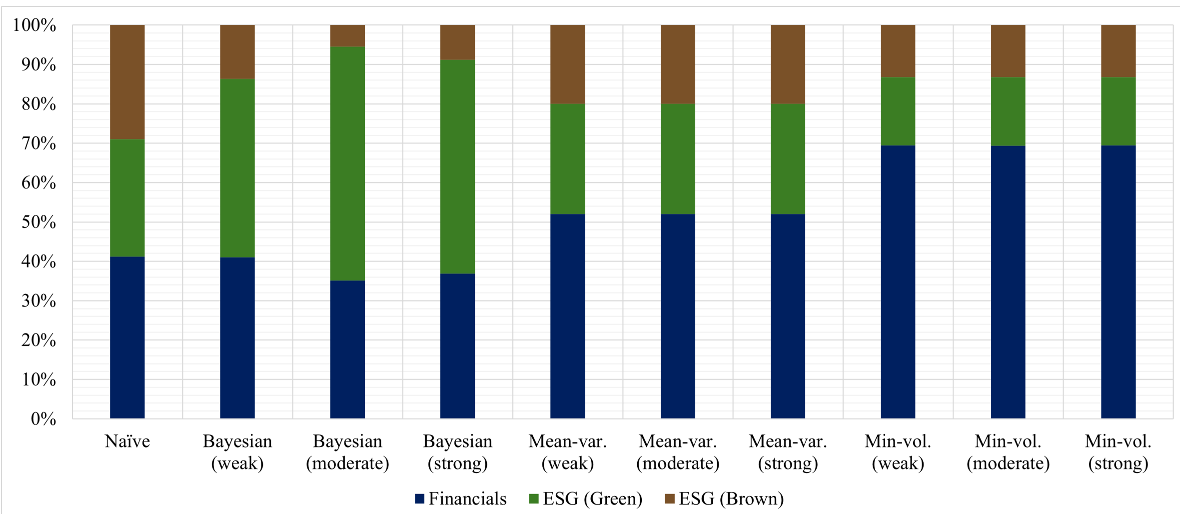
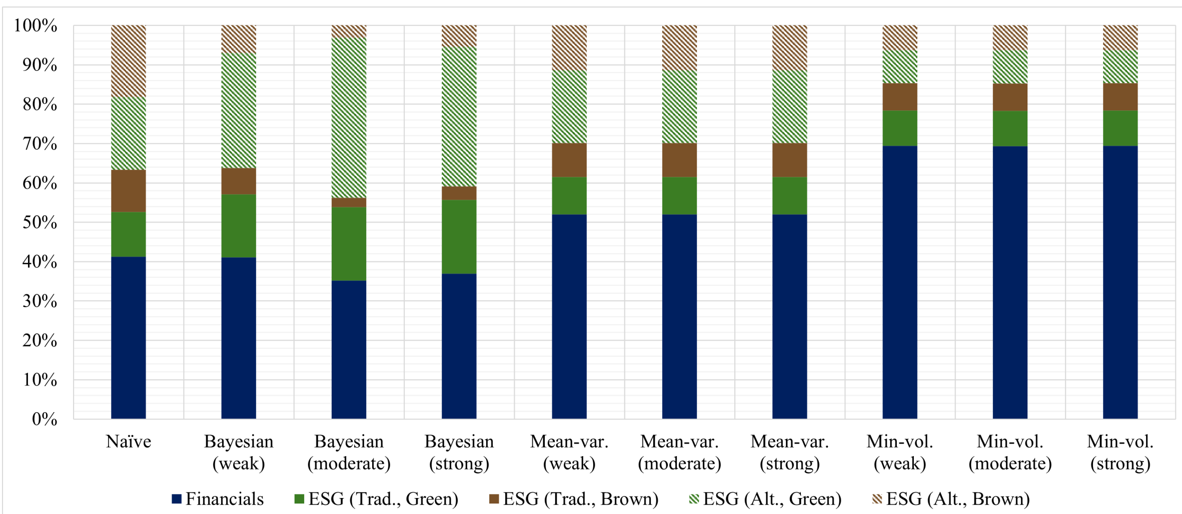


Figure 5a: Factor Bets in Portfolios that Optimally Combine ESG and Financial Factors: Green vs. Brown tilts (full sample results)



This figure displays the in-sample normalized factor weights across five factor categories for 10 multifactor portfolios. For each factor category, normalized weights are computed as the ratio of the sum of absolute weights within that category to the portfolio's total absolute factor weight (leverage). Factor categories comprise Financials (blue), ESG-Green (green) and ESG-Brown (brown). For ESG factors, green (brown) designation indicates that the factor weight in the multifactor portfolio corresponds to a long (short) position in the most sustainable leg paired with a short (long) position in the least sustainable leg of the factor. As for the multifactor portfolios' construction approaches, we report results for the Naïve approach, three Bayesian approaches with different levels of shrinkage (weak, moderate, and strong), three constrained mean-variance approaches with effective number of factors (ENF) constraint (10 for weak, 30 for moderate, and 50 for strong), and three minimum-volatility approaches with ENF constraint (10 for weak, 30 for moderate, and 50 for strong). Factor weights are estimated using the full sample covariance matrix and mean factor returns from August 2008 through June 2023.

Figure 5b: Bets in Portfolios that Optimally Combine ESG and Financial Factors: Alternative vs. traditional ESG metrics (full sample results)



This figure displays the in-sample normalized factor weights across five factor categories for 10 multifactor portfolios. For each factor category, normalized weights are computed as the ratio of the sum of absolute weights within that category to the portfolio's total absolute factor weight (leverage). Factor categories comprise Financials (blue), Traditional ESG-Green (solid green), Traditional ESG-Brown (solid brown), Alternative ESG-Green (green stripes), and Alternative ESG-Brown (brown stripes). For both Traditional and Alternative ESG factors, green (brown) designation indicates that the factor weight in the multifactor portfolio corresponds to a long (short) position in the most sustainable leg paired with a short (long) position in the least sustainable leg of the factor. As for the multifactor portfolios' construction approaches, we report results for the Naïve approach, three Bayesian approaches with different levels of shrinkage (weak, moderate, and strong), three constrained mean-variance approaches with effective number of factors (ENF) constraint (10 for weak, 30 for moderate, and 50 for strong), and three minimum-volatility approaches with ENF constraint (10 for weak, 30 for moderate, and 50 for strong). Factor weights are estimated using the full sample covariance matrix and mean factor returns from August 2008 through June 2023.

Table 1: Distinct ESG topics covered

E	**S**	**G**
Energy Efficiency and Climate Policies	Product Safety (Process and Use)	Governance and Board Independence
Emissions Reduction and GHG Management	Information to Customers	Controversy Identification and Resolution
Water Use and Pollution Control	Diversity, Equity, and Inclusion	Anti-Corruption Measures
Biodiversity Conservation and Ecosystem Management	Labor Relations and Fair Treatment	Cybersecurity and Data Privacy
Sustainable and Green Product Innovation	Training and Development for Employees	Transparency in Advocacy and Lobbying
Waste Management and Circular Economy	Health and Safety in Workplace	Equitable Tax Practices
Environmental Supply Chain Monitoring	Human Rights Compliance and Promotion	Corporate Political Contributions Transparency
Product Design and Lifecycle Impact	Supply Chain Social Compliance	Investor Rights and Engagement
Affordable and Clean Energy Development	Community Philanthropy and Engagement	Risk Management Frameworks
Physical Climate Risk Management	Employee Wellbeing and Benefits	Ethical Governance Practices
Ethical Treatment of Animals in Testing	Social Responsibility in Advertising	Transparency in Supply Chain Labor Practices
Sustainable Urban Development	Fair Trade and Supplier Relations	Legal and Regulatory Compliance
Recyclable and Eco-Friendly Packaging	Modern Slavery Prevention	
Conflict Mineral Sourcing	Access to Essential Products and Services	
Sustainable Agriculture Practices	Protection of Cultural Heritage	
Renewable Energy Investments	Gender Equality Initiatives	
	Child and Forced Labor Abolition	
	Inclusive Product and Service Design	
	Public Health Impact of Products	
	Community Health and Wellness Programs	
	Equity in Access to Digital Resources	

Table 2: Differences in Sharpe ratios (annualized)

Panel A: In-sample tests

Set of base factors	Multifactor portfolios									
	Naïve	Bayesian			Mean-variance constrained			Min-volatility constrained		
		weak	moderate	strong	weak	moderate	strong	weak	moderate	strong
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ESG only	0.07	-0.05	-0.19	-0.30	-0.04	-0.03	0.04	-0.16	-0.21	-0.09
	(0.85)	(0.79)	(0.39)	(0.22)	(0.89)	(0.91)	(0.89)	(0.63)	(0.53)	(0.78)
Financials + ESG	**0.75**	0.02	-0.08	-0.17	**0.60**	**0.61**	**0.78**	0.51	0.46	**0.52**
	(0.06)	(0.84)	(0.62)	(0.34)	(0.02)	(0.02)	(<.01)	(0.10)	(0.11)	(0.07)
Benchmark (Financials only) SR:	3.08	1.38	0.99	0.96	2.34	2.33	2.16	1.71	1.76	1.70

Panel B: Out-of-sample tests

Set of base factors	Multifactor portfolios									
	Naïve	Bayesian			Mean-variance constrained			Min-volatility constrained		
		weak	moderate	strong	weak	moderate	strong	weak	moderate	strong
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ESG only	-0.49	0.01	0.01	0.04	-0.35	-0.35	-0.18	-0.31	-0.34	-0.42
	(0.35)	(0.97)	(0.98)	(0.87)	(0.45)	(0.44)	(0.68)	(0.49)	(0.43)	(0.31)
Financials + ESG	-0.08	0.08	0.09	0.09	-0.23	-0.23	-0.07	-0.15	-0.18	-0.26
	(0.76)	(0.65)	(0.66)	(0.6)	(0.55)	(0.57)	(0.84)	(0.69)	(0.63)	(0.45)
Benchmark (Financials only) SR:	0.70	0.66	0.61	0.64	0.99	0.99	0.83	0.64	0.67	0.75

This table reports annualized differences (ΔSR) in annualized Sharpe ratios (in ***bold*** *if p-value $\leq$ 0.1) between various multifactor portfolios and a benchmark, along with two-tailed p-values (in brackets) for testing ΔSR=0. The last row of each panel reports the benchmark's SR. The multifactor portfolios are constructed using two sets of factors and 10 construction approaches. The set of factors are ESG and Financials plus ESG. As for the multifactor portfolios' construction approaches, we report results for the Naïve approach, three Bayesian approaches with different levels of shrinkage (weak, moderate, and strong), three constrained mean-variance approaches with effective number of factors (ENF) constraint (10 for weak, 30 for moderate, and 50 for strong), and three minimum-volatility approaches with ENF constraint (10 for weak, 30 for moderate, and 50 for strong). The benchmark portfolios only use our 130 financial factors and the indicated construction approach. In Panel A (in-sample results), portfolio weights are estimated from the full-sample covariance matrix and average returns (August 2008 to June 2023). In Panel B (out-of-sample results), the full sample is divided into three non-overlapping five-year periods, and weights for each five-year sub-sample are estimated from the other two sub-samples. We compute p-values via the stationary block bootstrap of Ledoit and Wolf (2008) with a block length of 10 months.*

Table 3: In-sample VS out-of-sample Sharpe ratios

Set of base factors	Multifactor portfolios									
	Naïve	Bayesian			Mean-variance constrained			Min-volatility constrained		
		weak	moderate	strong	weak	moderate	strong	weak	moderate	strong
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: In-sample Sharpe ratio										
Financials	**3.08**	**1.38**	**0.99**	**0.96**	**2.34**	**2.33**	**2.16**	**1.71**	**1.76**	**1.70**
	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)
ESG	**3.15**	**1.33**	**0.80**	**0.66**	**2.29**	**2.29**	**2.20**	**1.55**	**1.55**	**1.61**
	(<.01)	(<.01)	(0.01)	(0.03)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)
ESG + Financials	**3.83**	**1.40**	**0.91**	**0.79**	**2.94**	**2.94**	**2.94**	**2.22**	**2.22**	**2.22**
	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)
Panel B: Out-of-sample Sharpe ratio										
Financials	**0.70**	**0.66**	**0.61**	**0.64**	**0.99**	**0.99**	**0.83**	**0.64**	**0.67**	**0.75**
	(0.08)	(0.06)	(0.06)	(0.05)	(<.01)	(<.01)	(<.01)	(0.04)	(0.03)	(0.02)
ESG	0.21	**0.67**	**0.62**	**0.68**	0.64	0.64	0.66	0.33	0.33	0.33
	(0.54)	(0.05)	(0.04)	(0.02)	(0.12)	(0.12)	(0.11)	(0.32)	(0.32)	(0.33)
ESG + Financials	**0.62**	**0.74**	**0.70**	**0.73**	**0.76**	**0.76**	**0.76**	0.49	0.49	0.49
	(0.09)	(0.02)	(0.02)	(0.02)	(0.05)	(0.04)	(0.04)	(0.12)	(0.12)	(0.12)
Panel C: In-sample minus out-of-sample Sharpe ratio (decay)										
Financials	**-2.39**	**-0.72**	**-0.38**	**-0.32**	**-1.34**	**-1.34**	**-1.33**	**-1.08**	**-1.09**	**-0.96**
	(<.01)	(<.01)	(0.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)
ESG	**-2.94**	**-0.66**	-0.18	0.02	**-1.65**	**-1.65**	**-1.54**	**-1.23**	**-1.23**	**-1.29**
	(<.01)	(<.01)	(0.14)	(0.83)	(<.01)	(<.01)	(<.01)	(0.01)	(0.01)	(0.01)
ESG + Financials	**-3.21**	**-0.66**	**-0.21**	-0.06	**-2.18**	**-2.18**	**-2.18**	**-1.73**	**-1.73**	**-1.73**
	(<.01)	(<.01)	(0.05)	(0.48)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)

This table reports annualized Sharpe ratios (Panel A and B) and OOS gap (Panel C) of various multifactor portfolios (in bold if p-value $\leq$ 0.1), along with two-tailed p-values (in brackets). The multifactor portfolios are constructed using three sets of factors and 10 construction approaches. The set of factors are Financials, ESG, and Financials plus ESG. As for the multifactor portfolios' construction approaches, we report results for the Naïve approach, three Bayesian approaches with different levels of shrinkage (weak, moderate, and strong), three constrained mean-variance and three min-volatility approaches with effective number of factors (ENF) constraint (10 for weak, 30 for moderate, and 50 for strong). In Panel A, portfolio weights are estimated from the full-sample covariance matrix and average returns (August 2008 to June 2023), and the p-values test SR$\neq$0. In Panel B, the full sample is divided into three non-overlapping five-year periods, and weights for each five-year sub-sample are estimated from the other two sub-samples, and the p-values test SR$\neq$0. Panel C reports the difference between the IS and the OOS SR (OOS gap), and the p-value test that the OOS SR is different from the IS SR. We compute p-values via the stationary block bootstrap of Ledoit and Wolf (2008) with a block length of 10 months.

Table 4: Top three factors (by absolute weights) and normalized weights

Multifactor portfolio	Factor set					
	Financials		ESG Traditional		ESG Alternative	
Naïve	Short Interest	1.3%	MOODYS Community implementation	1.0%	GAIA transparency score	1.2%
	R&D/Sales	1.1%	MOODYS information to customer	-0.9%	GAIA water score	1.0%
	Invest. Growth	1.1%	MOODYS Community leadership	-0.8%	Sin Stocks	-1.0%
Bayesian (weak)	Invest. Growth	1.2%	CSR Human Rights	1.1%	SAUT. Climate Change reg sent.	1.5%
	Comp. Issuance	0.9%	MOODYS Governance leadership	-0.7%	Sin Stocks	-1.1%
	Sales/Equity	0.9%	CSR Product	0.7%	SAUT. Climate Change reg pos. sent.	1.0%
Bayesian (moderate)	Comp. Issuance	0.7%	CSR Human Rights	0.8%	SAUT. Climate Change reg sent.	1.0%
	Invest. Growth	0.7%	CSR Community	0.7%	Sin Stocks	-0.8%
	Sales/Inv.Cap	0.7%	CSR Product	0.6%	SAUT. Climate Change reg pos. sent.	0.7%
Bayesian (strong)	NA	NA	NA	NA	NA	NA
	NA	NA	NA	NA	NA	NA
	NA	NA	NA	NA	NA	NA
Mean-variance (weak)	Accrual Ratio	-4.4%	MOODYS train programs	-2.5%	GAIA water score	3.0%
	Ind.Rev+Leverage	3.2%	MOODYS Human Rights results	1.6%	COV Labor controversy.	2.8%
	Price/Sales	-2.9%	MOODYS employee community inolvmnt.	-1.4%	Sin Stocks	-2.4%
Mean-variance (moderate)	Accrual Ratio	-4.4%	MOODYS train programs	-2.5%	GAIA water score	3.0%
	Ind.Rev+Leverage	3.2%	MOODYS Human Rights results	1.6%	COV Labor controversy.	2.8%
	Price/Sales	-2.9%	MOODYS employee community inolvmnt.	-1.4%	Sin Stocks	-2.4%
Mean-variance (strong)	Accrual Ratio	-4.4%	MOODYS train programs	-2.5%	GAIA water score	3.0%
	Ind.Rev+Leverage	3.2%	MOODYS Human Rights results	1.6%	COV Labor controversy.	2.8%
	Price/Sales	-2.9%	MOODYS employee community inolvmnt.	-1.4%	Sin Stocks	-2.4%
Min-vol. (weak)	Equity/Inv.Cap	4.3%	MOODYS train programs	-1.8%	COV Labor controversy.	1.7%
	Op.Marg.PreDep	-3.6%	MOODYS Governance overall.	-1.8%	COV Mgmt controversy.	-1.3%
	Cash Flow/Price	-3.4%	MOODYS Human Rights results	1.2%	SAUT. Climate Change expos. sent.	1.2%
Min-vol. (moderate)	Equity/Inv.Cap	4.3%	MOODYS train programs	-1.9%	COV Labor controversy.	1.7%
	Op.Marg.PreDep	-3.6%	MOODYS Governance overall.	-1.8%	COV Mgmt controversy.	-1.3%
	Cash Flow/Price	-3.3%	MOODYS Human Rights results	1.2%	SAUT. Climate Change expos. sent.	1.2%
Min-vol. (strong)	Equity/Inv.Cap	4.3%	MOODYS train programs	-1.8%	COV Labor controversy.	1.7%
	Op.Marg.PreDep	-3.6%	MOODYS Governance overall.	-1.8%	COV Mgmt controversy.	-1.3%
	Cash Flow/Price	-3.3%	MOODYS Human Rights results	1.2%	SAUT. Climate Change expos. sent.	1.2%

The table above reports the top three factors by absolute weight and the respective absolute weight in the factor normalized by leverage for our 30 in-sample multifactor portfolios (3 sets of factors by 10 portfolio construction approaches).

Table 5: Differences in volatility of Minimum-volatility portfolios relative to the benchmark

Set of base factors	Panel A: In-sample			Panel B: Out-of-sample		
	Min-volatility constrained			Min-volatility constrained		
	weak	moderate	strong	weak	moderate	strong
	(1)	(2)	(3)	(4)	(5)	(6)
ESG	**0.80**	**0.67**	**0.49**	0.91	0.90	0.75
	(<.01)	(<.01)	(<.01)	(0.14)	(0.14)	(0.26)
Financials + ESG	0.00	-0.13	**-0.51**	0.27	0.26	-0.01
	(0.94)	(0.11)	(<.01)	(0.34)	(0.37)	(0.96)
Financials only (benchmark) Volatility	0.93	1.06	1.44	1.12	1.13	1.40

This table reports annualized differences in standard deviation of the returns (volatility) of various multifactor portfolios relative to the volatility of the benchmark portfolio (last row), along with two-tailed p-values (in brackets) for testing whether the volatility of the multifactor portfolio is different from the volatility of the benchmark. The multifactor portfolios are constructed using two sets of base factors, ESG only, and Financials plus ESG. The set of base factors of the benchmark is Financials only. For each set of base factors, we adopt 3 portfolio's construction approaches targeting minimization of portfolio's volatility, with three different levels of effective number of factors constraint (10 for weak, 30 for moderate, and 50 for strong). In Panel A (in-sample results), portfolio weights are estimated from the full-sample covariance matrix and average returns (August 2008 to June 2023). In Panel B (out-of-sample results), the full sample is divided into three non-overlapping five-year periods, and weights for each five-year sub-sample are estimated from the other two sub-samples. We compute p-values via the stationary block bootstrap of Ledoit and Wolf (2011) with a block length of 10 months. We report the volatility in bold if p-value $\leq$ 0.1.

Appendix

A Description of characteristics

A.1 Published return effects

We report here a brief description of all the characteristics used to construct the factors in our Published Returns Effects category, divided by factor-theme. A more detailed description can be found in the internet appendix of Kozak et al. (2020) (KNS), which we follow to construct our characteristics. All the characteristics are constructed using the CRSP/Compustat database. Within the square brackets we report 1 (-1) if the factor is long (short) the stocks with high value of the characteristics and short (long) the stocks with low value of the characteristics. Finally, we report the main reference.

Size: characteristics related to the size of the company or the stock price.

1. Size (size) [-1]: Equity market capitalization of the company updated annually (Fama and French, 1993).
2. Price (price) [-1]: Stock price (Blume and Husic, 1973).

Value: characteristics related to stock valuation, typically comparing book value or other valuation measures to market capitalization.

3. Value-annual (value) [-1]: Book-to-market ratio updated annually (Fama and French, 1993).
4. Dividend Yield (divp) [-1]: Dividend to stock price ratio (Naranjo et al., 1998).
5. Earnings/Price (ep) [1]: Earnings-to-price ratio (Basu, 1977).
6. Sales-to-Price (sp) [1]: Revenues-to-market equity ratio (Barbee Jr et al., 1996).
7. Value-monthly (valuem) [1]: book-to-market ratio updated monthly (Asness and Frazzini, 2013).

Profitability: return predictors related to company profitability or operational efficiency.

8. Gross Profitability (prof) [1]: Ratio of gross profits to total assets (Novy-Marx, 2013).
9. Asset Turnover (aturnover) [1]: Ratio of sales to total assets (Soliman, 2008).
10. Gross Margins (gmargins) [1]: Ratio of gross profits to sales (Novy-Marx, 2013).
11. Cash Flow / Market Value of Equity (cfp) [1]: Cash flow relative to the market value of equity (Lakonishok et al., 1994).
12. Return on Assets (annual) (roaa) [1]: Net income relative to total assets, updated annually (Chen et al., 2011).

13. Return on Equity (annual) (roea) [1]: Net income relative to book value of equity, updated annually (Haugen and Baker (1996)).
14. Return on Book Equity (roe) [1]: Net income relative to book value of equity, updated monthly (Chen et al., 2011).
15. Return on Market Equity (rome) [1]: Net income relative to market value of equity, updated (Chen et al., 2011).
16. Return on Assets (roa) [1]: Net income relative to total assets, updated quarterly (Chen et al., 2011).

Investment: return predictors related to investments or asset growth.

17. Share Issuance (annual) (nissa) [-1]: Issuance of new shares updated annually (Pontiff and Woodgate, 2008).
18. Asset Growth (growth) [-1]: Growth in assets (Cooper et al., 2008).
19. Investment (inv) [-1]: Changes in investments in production assets and inventories (Chen et al., 2011).
20. Investment-to-Capital (invcap) [-1]: Ratio of capital expenditures to fixed assets (Xing, 2008).
21. Investment Growth (igrowth) [-1]: Growth in capital expenditures (Xing, 2008).
22. Sales Growth (sgrowth) [-1]: annual growth in sales (Lakonishok et al., 1994).
23. Share Issuance (monthly) (nissm) [-1]: Issuance of shares updated on a monthly basis (Pontiff and Woodgate, 2008).
24. Composite Issuance (ciss) [-1]: Combined issuance of shares (Daniel and Titman, 2006).
25. Share Repurchases (repurch) [1]: Repurchase of shares (Ikenberry et al., 1995).
26. Debt Issuance (debtiss) [-1]: Issuance of long-term debt (Spiess and Affleck-Graves, 1999).
27. Growth in LTNOA (gltnoa) [-1]: Growth in long-term net operating assets minus adjustments (Fairfield et al., 2003).

Momentum: return predictors related to past stock returns.

28. Momentum-6m (mom) [1]: Stock return over the past six months (Jegadeesh and Titman, 1993).
29. Industry Momentum (indmom) [1]: Stock return of the industry (Moskowitz and Grinblatt, 1999).
30. Momentum-1 year (mom12) [1]: Stock return over the past year (Jegadeesh and Titman, 1993).

Low Risk: return predictors related to measures of risk.

31. Idiosyncratic Volatility (ivol) [-1]: Company-specific volatility (Ang et al., 2006).
32. Beta Arbitrage (beta) [-1]: Beta with respect to the market index (Fama and MacBeth, 1973).

Others: return predictors that do not fit clearly into one of the previous themes.

33. Industry Relative Reversals (Low Volatility) (indrrevlv) (-1): Trend reversals with low volatility (Da et al., 2014).
34. Value profitability (valprof) [1]: Combination of value and profitability (Novy-Marx, 2013).
35. Piotroski's F-score (F-score) [1]: A score of financial health of the firm (Piotroski, 2000).
36. Accruals (accruals) [-1]: Changes in assets and liabilities (Sloan, 1996).
37. Net Operating Assets (noa) [-1]: Net operating assets (Hirshleifer et al., 2004).
38. Value-Momentum (valmom) [1]: Combination of value and past performance (Novy-Marx, 2013).
39. Value-Momentum-Profitability (valmomprof) [1]: Combination of value, performance, and profitability (Novy-Marx, 2013).
40. Leverage (lev) [-1]: Ratio of total assets to the market value of equity (Bhandari, 1988).
41. Seasonality (season) [1]: Seasonal effects on returns; average monthly return in the same calendar month over the last 5 years. (Heston and Sadka, 2008).
42. Share Volume (shvol) [-1]: Reflects trading volume (Datar et al., 1998).
43. Short Interest (shortint) [-1]: Short interest (Asquith et al., 2005).
44. PEAD-SUE (sue) [1]: Standardized unexpected earnings (Foster et al. (1984)).
45. Firm Age (age) [-1]: Age of the company (Barry and Brown (1984)).
46. Momentum-Reversal (momrev) [-1]: Combination of past performance and trend reversals (Jegadeesh and Titman, 1993).
47. Long-term Reversals (lrrev) [-1]: Long-term trend reversals (De Bondt and Thaler, 1985).
48. Industry Relative Reversals (indrrev) [-1]: Trend reversals relative to the industry (Da et al., 2014).
49. Industry Momentum-Reversal (indmomrev) [1]: Combination of past performance and trend reversals in the industry (Moskowitz and Grinblatt, 1999).
50. Short-term Reversal (strev) [-1]: Short-term reversals (Jegadeesh, 1990).

A.2 General firm characteristics

We report here a brief description of all the General Firm Characteristics (with id in parentheses). The general firm characteristics from 1 to 68 are from the WRDS Financial Ratio dataset, from 69 to 80 are constructed from CRSP. All the general firm characteristics factors are long the stocks with high value of the characteristics and short the stocks with low value of the characteristics.

1. P/E-Diluted, Excl. EI (pe_exi) – Valuation. Price-to-Earnings, excl. Extraordinary Items (diluted).
2. P/E-Diluted, Incl. EI (pe_inc) – Valuation. Price-to-Earnings, incl. Extraordinary Items (diluted).
3. Price/Sales (ps) – Valuation. Multiple of Market Value of Equity to Sales.
4. Price/Cash flow (pcf) – Valuation. Multiple of Market Value of Equity to Net Cash Flow from Operating Activities.
5. Enterprise Value Multiple (evm) – Valuation. Multiple of Enterprise Value to EBITDA.
6. Book/Market (bm) – Valuation. Book Value of Equity as a fraction of Market Value of Equity.
7. Shiller's Cyclically Adjusted P/E Ratio (capei) – Valuation. Multiple of Market Value of Equity to 5-year moving average of Net Income.
8. Dividend Payout Ratio (dpr) – Valuation. Dividends as a fraction of Income Before Extra. Items.
9. Net Profit Margin (npm) – Profitability. Net Income as a fraction of Sales.
10. Operating Profit Margin Before Depreciation (opmbd) – Profitability. Operating Income Before Depreciation as a fraction of Sales.
11. Operating Profit Margin After Depreciation (opmad) – Profitability. Operating Income after Depreciation as a fraction of Sales.
12. Gross Profit Margin (gpm) – Profitability. Gross Profit as a fraction of Sales.
13. Pre-tax Profit Margin (ptpm) – Profitability. Pretax Income as a fraction of Sales.
14. Cash Flow Margin (cfm) – Financial Soundness. Income before Extraordinary Items and Depreciation as a fraction of Sales.
15. Return on Assets (roa_fr) – Profitability. Operating Income Before Depreciation as a fraction of average Total Assets based on most recent two periods.
16. Return on Equity (roe_fr) – Profitability. Net Income as a fraction of average Book Equity based on most recent two periods, where Book Equity is defined as the sum of Total Parent Stockholders' Equity and Deferred Taxes and Investment Tax Credit.

17. Return on Capital Employed (roce) – Profitability. Earnings Before Interest and Taxes as a fraction of average Capital Employed based on most recent two periods, where Capital Employed is the sum of Debt in Long-term and Current Liabilities and Common/Ordinary Equity.

18. After-tax Return on Average Common Equity (aftret_eq) – Profitability. Net Income as a fraction of average of Common Equity based on most recent two periods.

19. After-tax Return on Invested Capital (aftret_invcapx) – Profitability. Net Income plus Interest Expenses as a fraction of Invested Capital.

20. After-tax Return on Total Stockholders' Equity (aftret_equity) – Profitability. Net income as a fraction of average of Total Shareholders' Equity based on most recent two periods.

21. Pre-tax return on Net Operating Assets (pretret_noa) – Profitability. Operating Income After Depreciation as a fraction of average Net Operating Assets (NOA) based on most recent two periods, where NOA is defined as the sum of Property Plant and Equipment and Current Assets minus Current Liabilities.

22. Pre-tax Return on Total Earning Assets (pretret_earnat) – Profitability. Operating Income After Depreciation as a fraction of average Total Earnings Assets (TEA) based on most recent two periods, where TEA is defined as the sum of Property Plant and Equipment and Current Assets.

23. Common Equity/Invested Capital (equity_invcap) – Capitalization. Common Equity as a fraction of Invested Capital.

24. Long-term Debt/Invested Capital (debt_invcap) – Capitalization. Long-term Debt as a fraction of Invested Capital.

25. Total Debt/Invested Capital (totdebt_invcap) – Capitalization. Total Debt (Long-term and Current) as a fraction of Invested Capital.

26. Interest/Average Long-term Debt (int_debt) – Financial Soundness. Interest as a fraction of average Long-term debt based on most recent two periods.

27. Interest/Average Total Debt (int_totdebt) – Financial Soundness. Interest as a fraction of average Total Debt based on most recent two periods.

28. Cash Balance/Total Liabilities (cash_lt) – Financial Soundness. Cash Balance as a fraction of Total Liabilities.

29. Inventory/Current Assets (invt_act) – Financial Soundness. Inventories as a fraction of Current Assets.

30. Receivables/Current Assets (rect_act) – Financial Soundness. Accounts Receivables as a fraction of Current Assets.

31. Total Debt/Total Assets (debt_at) – Solvency. Total Liabilities as a fraction of Total Assets.

32. Short-Term Debt/Total Debt (short_debt) – Financial Soundness. Short-term Debt as a fraction of Total Debt.

33. Current Liabilities/Total Liabilities (curr_debt) – Financial Soundness. Current Liabilities as a fraction of Total Liabilities.

34. Long-term Debt/Total Liabilities (lt_debt) – Financial Soundness. Long-term Debt as a fraction of Total Liabilities.

35. Free Cash Flow/Operating Cash Flow (fcf_ocf) – Financial Soundness. Free Cash Flow as a fraction of Operating Cash Flow, where Free Cash Flow is defined as the difference between Operating Cash Flow and Capital Expenditures.

36. Avertising Expenses/Sales (adv_sale) – Other. Advertising Expenses as a fraction of Sales.

37. Profit Before Depreciation/Current Liabilities (profit_lct) – Financial Soundness. Operating Income before D&A as a fraction of Current Liabilities.

38. Total Debt/EBITDA (debt_ebitda) – Financial Soundness. Gross Debt as a fraction of EBITDA.

39. Operating CF/Current Liabilities (ocf_lct) – Financial Soundness. Operating Cash Flow as a fraction of Current Liabilities.

40. Total Liabilities/Total Tangible Assets (lt_ppent) – Financial Soundness. Total Liabilities to Total Tangible Assets.

41. Long-term Debt/Book Equity (dltt_be) – Financial Soundness. Long-term Debt to Book Equity.

42. Total Debt/Total Assets (debt_assets) – Solvency. Total Debt as a fraction of Total Assets.

43. Total Debt/Capital (debt_capital) – Solvency. Total Debt as a fraction of Total Capital, where Total Debt is defined as the sum of Accounts Payable and Total Debt in Current and Long-term Liabilities, and Total Capital is defined as the sum of Total Debt and Total Equity (common and preferred).

44. Total Debt/Equity (de_ratio) – Solvency. Total Liabilities to Shareholders' Equity (common and preferred).

45. After-tax Interest Coverage (intcov) – Solvency. Multiple of After-tax Income to Interest and Related Expenses.

46. Cash Ratio (cash_ratio) – Liquidity. Cash and Short-term Investments as a fraction of Current Liabilities.

47. Quick Ratio (Acid Test) (quick_ratio) – Liquidity. Quick Ratio: Current Assets net of Inventories as a fraction of Current Liabilities.

48. Current Ratio (curr_ratio) – Liquidity. Current Assets as a fraction of Current Liabilities.

49. Capitalization Ratio (capital_ratio) – Capitalization. Total Long-term Debt as a fraction of the sum of Total Long-term Debt, Common/Ordinary Equity and Preferred Stock.

50. Cash Flow/Total Debt (cash_debt) – Financial Soundness. Operating Cash Flow as a fraction of Total Debt.

51. Inventory Turnover (inv_turn) – Efficiency. COGS as a fraction of the average Inventoriesbased on the most recent two periods.

52. Asset Turnover (at_turn) – Efficiency. Sales as a fraction of the average Total Assets basedon the most recent two periods.

53. Receivables Turnover (rect_turn) – Efficiency. Sales as a fraction of the average of Accounts Receivables based on the most recent two periods.

54. Payables Turnover (pay_turn) – Efficiency. COGS and change in Inventories as a fraction of the average of Accounts Payable based on the most recent two periods.

55. Sales/Invested Capital (sale_invcap) – Efficiency. Sales per dollar of Invested Capital.

56. Sales/Stockholders Equity (sale_equity) – Efficiency. Sales per dollar of total Stockholders' Equity.

57. Sales/Working Capital (sale_nwc) – Efficiency. Sales per dollar of Working Capital, defined as difference between Current Assets and Current Liabilities.

58. Research and Development/Sales (RD_SALE) – Other. R&D expenses as a fraction of Sales.

59. Accruals/Average Assets (Accrual) – Other. Accruals as a fraction of average Total Assets based on most recent two periods.

60. Gross Profit/Total Assets (GProf) – Profitability. Gross Profitability as a fraction of Total Assets.

61. Book Equity (be) – Other. Firm size as measured by total book equity.

62. Cash Conversion Cycle (Days) (cash_conversion) – Liquidity. Inventories per daily COGS plus Account Receivables per daily Sales minus Account Payables per daily COGS.

63. Effective Tax Rate (efftax) – Profitability. Income Tax as a fraction of Pretax Income.

64. Interest Coverage Ratio (intcov_ratio) – Solvency. Multiple of Earnings Before Interest and Taxes to Interest and Related Expenses.

65. Labor Expenses/Sales (staff_sale) – Other. Labor Expenses as a fraction of Sales.

66. Dividend Yield (divyield) – Valuation. Indicated Dividend Rate as a fraction of Price.

67. Price/Book (ptb) – Valuation. Multiple of Market Value of Equity to Book Value of Equity.

68. Trailing P/E to Growth (PEG) ratio (PEG_trailing) – Valuation. Price-to-Earnings, excl. Extraordinary Items (diluted) to 3-Year past EPS Growth.
69. Return in Month $t-1$ (ret_lag1) – Other. Past one-month returns in month $t-1$.
70. Return in Month $t-2$ (ret_lag2) – Other. Past one-month returns in month $t-2$.
71. Return in Month $t-3$ (ret_lag3) – Other. Past one-month returns in month $t-3$.
72. Return in Month $t-4$ (ret_lag4) – Other. Past one-month returns in month $t-4$.
73. Return in Month $t-5$ (ret_lag5) – Other. Past one-month returns in month $t-5$.
74. Return in Month $t-6$ (ret_lag6) – Other. Past one-month returns in month $t-6$.
75. Return in Month $t-7$ (ret_lag7) – Other. Past one-month returns in month $t-7$.
76. Return in Month $t-8$ (ret_lag8) – Other. Past one-month returns in month $t-8$.
77. Return in Month $t-9$ (ret_lag9) – Other. Past one-month returns in month $t-9$.
78. Return in Month $t-10$ (ret_lag10) – Other. Past one-month returns in month $t-10$.
79. Return in Month $t-11$ (ret_lag11) – Other. Past one-month returns in month $t-11$.
80. Return in Month $t-12$ (ret_lag12) – Other. Past one-month returns in month $t-12$.

A.3 ESG Characteristics

We report a short description of all the ESG characteristics that we employ for our analyses. We also report the source, and in brackets the sign of the relation between the value of the characteristic and the level of sustainability of the company; 1 (-1) indicates that high value of the characteristic is associated with high (low) levels of sustainability. We use this sign also to determine the long and the short leg of the factors constructed with ESG information.

1. **SINHK_SIN** (-1): A dummy variable for companies with significant revenues from sin industries (Tobacco, alcohol, and gaming). Source - Sin Stocks Hong and Kacperczyk (2009).
2. **BC_BC** (1): A dummy variable for the 'Best companies for employees to work for'. Source - Best Companies Edmans (2011).
3. **COV_CONS** (1): Overall controversy score reflecting current ESG controversies intensity, calculated using negative news volume with 20% monthly erosion factor. Source - Covalence.
4. **COV_DISC** (1): Overall disclosure score combining company-reported ESG indicators and corporate communications, normalized to 0-100 scale. Source - Covalence.
5. **COV_E** (1): Environmental category score combining disclosure and reputation analysis across environmental dimensions. Source - Covalence.
6. **COV_ESG** (1): Overall ESG score averaging Environmental, Social and Governance category scores. Source - Covalence.

7. **COV_ECON** (1): Economic and environmental impacts score based on news analysis and company disclosures. Source - Covalence.
8. **COV_EDISC** (1): Environmental disclosure score based on reported environmental indicators and communications. Source - Covalence.
9. **COV_EREP** (1): Environmental reputation score based on stakeholder perceptions and news analysis. Source - Covalence.
10. **COV_G** (1): Governance category score combining disclosure and reputation analysis across governance dimensions. Source - Covalence.
11. **COV_GCON** (1): Governance controversy score based on negative governance-related news. Source - Covalence.
12. **COV_GDISC** (1): Governance disclosure score based on reported governance indicators and communications. Source - Covalence.
13. **COV_GREP** (1): Governance reputation score based on stakeholder perceptions and news analysis. Source - Covalence.
14. **COV_REPS** (1): Overall reputation score based on positive and negative news analysis with 2% monthly erosion factor. Source - Covalence.
15. **COV_S** (1): Social category score combining disclosure and reputation analysis across social dimensions. Source - Covalence.
16. **COV_SCON** (1): Social controversy score based on negative social-related news. Source - Covalence.
17. **COV_SDISC** (1): Social disclosure score based on reported social indicators and communications. Source - Covalence.
18. **COV_SREP** (1): Social reputation score based on stakeholder perceptions and news analysis. Source - Covalence.
19. **COV_ESG_ENV1** (1): Environmental impact of products score combining disclosure and reputation analysis. Source - Covalence.
20. **COV_ESG_ENV2** (1): Resources management score combining disclosure and reputation analysis. Source - Covalence.
21. **COV_ESG_ENV3** (1): Emissions, effluents and waste score combining disclosure and reputation analysis. Source - Covalence.
22. **COV_ESG_GOV10** (1): Sustainability strategy score combining disclosure and reputation analysis. Source - Covalence.
23. **COV_ESG_GOV11** (1): Remuneration score combining disclosure and reputation analysis. Source - Covalence.
24. **COV_ESG_GOV8** (1): Management score combining disclosure and reputation analysis. Source - Covalence.

25. **COV_ESG_GOV9** (1): Shareholders score combining disclosure and reputation analysis. Source - Covalence.
26. **COV_ESG_SOC4** (1): Human rights score combining disclosure and reputation analysis. Source - Covalence.
27. **COV_ESG_SOC5** (1): Social impact of products score combining disclosure and reputation analysis. Source - Covalence.
28. **COV_ESG_SOC6** (1): Labor practices score combining disclosure and reputation analysis. Source - Covalence.
29. **COV_ESG_SOC7** (1): Community impact score combining disclosure and reputation analysis. Source - Covalence.
30. **COV_ESGHLD_ENV1** (1): ESG controversy score for environmental impact of products. Source - Covalence.
31. **COV_ESGHLD_ENV2** (1): ESG controversy score for resources management. Source - Covalence.
32. **COV_ESGHLD_ENV3** (1): ESG controversy score for emissions, effluents and waste. Source - Covalence.
33. **COV_ESGHLD_GOV10** (1): ESG controversy score for sustainability strategy. Source - Covalence.
34. **COV_ESGHLD_GOV8** (1): ESG controversy score for management. Source - Covalence.
35. **COV_ESGHLD_GOV9** (1): ESG controversy score for shareholders. Source - Covalence.
36. **COV_ESGHLD_SOC4** (1): ESG controversy score for human rights. Source - Covalence.
37. **COV_ESGHLD_SOC5** (1): ESG controversy score for social impact of products. Source - Covalence.
38. **COV_ESGHLD_SOC6** (1): ESG controversy score for labor practices. Source - Covalence.
39. **COV_ESGHLD_SOC7** (1): ESG controversy score for community impact. Source - Covalence.
40. **COV_ESGREP_ENV1** (1): ESG reputation score for environmental impact of products. Source - Covalence.
41. **COV_ESGREP_ENV2** (1): ESG reputation score for resources management. Source - Covalence.
42. **COV_ESGREP_ENV3** (1): ESG reputation score for emissions, effluents and waste. Source - Covalence.

43. **COV_ESGREP_GOV10** (1): ESG reputation score for sustainability strategy. Source - Covalence.
44. **COV_ESGREP_GOV11** (1): ESG reputation score for remuneration. Source - Covalence.
45. **COV_ESGREP_GOV8** (1): ESG reputation score for management. Source - Covalence.
46. **COV_ESGREP_GOV9** (1): ESG reputation score for shareholders. Source - Covalence.
47. **COV_ESGREP_SOC4** (1): ESG reputation score for human rights. Source - Covalence.
48. **COV_ESGREP_SOC5** (1): ESG reputation score for social impact of products. Source - Covalence.
49. **COV_ESGREP_SOC6** (1): ESG reputation score for labor practices. Source - Covalence.
50. **COV_ESGREP_SOC7** (1): ESG reputation score for community impact. Source - Covalence.
51. **COV_DISCLO_ENV1** (1): Disclosure score for environmental impact of products based on reported indicators. Source - Covalence.
52. **COV_DISCLO_ENV2** (1): Disclosure score for resources management based on reported indicators. Source - Covalence.
53. **COV_DISCLO_ENV3** (1): Disclosure score for emissions, effluents and waste based on reported indicators. Source - Covalence.
54. **COV_DISCLO_GOV10** (1): Disclosure score for sustainability strategy based on reported indicators. Source - Covalence.
55. **COV_DISCLO_GOV11** (1): Disclosure score for remuneration based on reported indicators. Source - Covalence.
56. **COV_DISCLO_GOV8** (1): Disclosure score for management based on reported indicators. Source - Covalence.
57. **COV_DISCLO_GOV9** (1): Disclosure score for shareholders based on reported indicators. Source - Covalence.
58. **COV_DISCLO_SOC4** (1): Disclosure score for human rights based on reported indicators. Source - Covalence.
59. **COV_DISCLO_SOC5** (1): Disclosure score for social impact of products based on reported indicators. Source - Covalence.
60. **COV_DISCLO_SOC6** (1): Disclosure score for labor practices based on reported indicators. Source - Covalence.

61. **COV_DISCLO_SOC7** (1): Disclosure score for community impact based on reported indicators. Source - Covalence.
62. **COV_SDG_1** (1): SDG mapping score for No Poverty (Goal 1) combining disclosure and reputation analysis. Source - Covalence.
63. **COV_SDG_10** (1): SDG mapping score for Reduced Inequalities (Goal 10) combining disclosure and reputation analysis. Source - Covalence.
64. **COV_SDG_11** (1): SDG mapping score for Sustainable Cities and Communities (Goal 11) combining disclosure and reputation analysis. Source - Covalence.
65. **COV_SDG_12** (1): SDG mapping score for Responsible Consumption and Production (Goal 12) combining disclosure and reputation analysis. Source - Covalence.
66. **COV_SDG_13** (1): SDG mapping score for Climate Action (Goal 13) combining disclosure and reputation analysis. Source - Covalence.
67. **COV_SDG_14** (1): SDG mapping score for Life Below Water (Goal 14) combining disclosure and reputation analysis. Source - Covalence.
68. **COV_SDG_15** (1): SDG mapping score for Life on Land (Goal 15) combining disclosure and reputation analysis. Source - Covalence.
69. **COV_SDG_16** (1): SDG mapping score for Peace, Justice and Strong Institutions (Goal 16) combining disclosure and reputation analysis. Source - Covalence.
70. **COV_SDG_17** (1): SDG mapping score for Partnerships for the Goals (Goal 17) combining disclosure and reputation analysis. Source - Covalence.
71. **COV_SDG_2** (1): SDG mapping score for Zero Hunger (Goal 2) combining disclosure and reputation analysis. Source - Covalence.
72. **COV_SDG_3** (1): SDG mapping score for Good Health and Well-being (Goal 3) combining disclosure and reputation analysis. Source - Covalence.
73. **COV_SDG_4** (1): SDG mapping score for Quality Education (Goal 4) combining disclosure and reputation analysis. Source - Covalence.
74. **COV_SDG_5** (1): SDG mapping score for Gender Equality (Goal 5) combining disclosure and reputation analysis. Source - Covalence.
75. **COV_SDG_6** (1): SDG mapping score for Clean Water and Sanitation (Goal 6) combining disclosure and reputation analysis. Source - Covalence.
76. **COV_SDG_7** (1): SDG mapping score for Affordable and Clean Energy (Goal 7) combining disclosure and reputation analysis. Source - Covalence.
77. **COV_SDG_8** (1): SDG mapping score for Decent Work and Economic Growth (Goal 8) combining disclosure and reputation analysis. Source - Covalence.
78. **COV_SDG_9** (1): SDG mapping score for Industry, Innovation and Infrastructure (Goal 9) combining disclosure and reputation analysis. Source - Covalence.

79. **COV_DISCLO_SDG1** (1): Disclosure score for SDG 1 (No Poverty) based on reported indicators. Source - Covalence.
80. **COV_DISCLO_SDG10** (1): Disclosure score for SDG 10 (Reduced Inequalities) based on reported indicators. Source - Covalence.
81. **COV_DISCLO_SDG11** (1): Disclosure score for SDG 11 (Sustainable Cities and Communities) based on reported indicators. Source - Covalence.
82. **COV_DISCLO_SDG12** (1): Disclosure score for SDG 12 (Responsible Consumption and Production) based on reported indicators. Source - Covalence.
83. **COV_DISCLO_SDG13** (1): Disclosure score for SDG 13 (Climate Action) based on reported indicators. Source - Covalence.
84. **COV_DISCLO_SDG14** (1): Disclosure score for SDG 14 (Life Below Water) based on reported indicators. Source - Covalence.
85. **COV_DISCLO_SDG15** (1): Disclosure score for SDG 15 (Life on Land) based on reported indicators. Source - Covalence.
86. **COV_DISCLO_SDG16** (1): Disclosure score for SDG 16 (Peace, Justice and Strong Institutions) based on reported indicators. Source - Covalence.
87. **COV_DISCLO_SDG17** (1): Disclosure score for SDG 17 (Partnerships for the Goals) based on reported indicators. Source - Covalence.
88. **COV_DISCLO_SDG2** (1): Disclosure score for SDG 2 (Zero Hunger) based on reported indicators. Source - Covalence.
89. **COV_DISCLO_SDG3** (1): Disclosure score for SDG 3 (Good Health and Well-being) based on reported indicators. Source - Covalence.
90. **COV_DISCLO_SDG4** (1): Disclosure score for SDG 4 (Quality Education) based on reported indicators. Source - Covalence.
91. **COV_DISCLO_SDG5** (1): Disclosure score for SDG 5 (Gender Equality) based on reported indicators. Source - Covalence.
92. **COV_DISCLO_SDG6** (1): Disclosure score for SDG 6 (Clean Water and Sanitation) based on reported indicators. Source - Covalence.
93. **COV_DISCLO_SDG7** (1): Disclosure score for SDG 7 (Affordable and Clean Energy) based on reported indicators. Source - Covalence.
94. **COV_DISCLO_SDG8** (1): Disclosure score for SDG 8 (Decent Work and Economic Growth) based on reported indicators. Source - Covalence.
95. **COV_DISCLO_SDG9** (1): Disclosure score for SDG 9 (Industry, Innovation and Infrastructure) based on reported indicators. Source - Covalence.
96. **COV_REP_SDG1** (1): Reputation score for SDG 1 (No Poverty) based on stakeholder perceptions and news analysis. Source - Covalence.

97. **COV_REP_SDG10** (1): Reputation score for SDG 10 (Reduced Inequalities) based on stakeholder perceptions and news analysis. Source - Covalence.
98. **COV_REP_SDG11** (1): Reputation score for SDG 11 (Sustainable Cities and Communities) based on stakeholder perceptions and news analysis. Source - Covalence.
99. **COV_REP_SDG12** (1): Reputation score for SDG 12 (Responsible Consumption and Production) based on stakeholder perceptions and news analysis. Source - Covalence.
100. **COV_REP_SDG13** (1): Reputation score for SDG 13 (Climate Action) based on stakeholder perceptions and news analysis. Source - Covalence.
101. **COV_REP_SDG14** (1): Reputation score for SDG 14 (Life Below Water) based on stakeholder perceptions and news analysis. Source - Covalence.
102. **COV_REP_SDG15** (1): Reputation score for SDG 15 (Life on Land) based on stakeholder perceptions and news analysis. Source - Covalence.
103. **COV_REP_SDG16** (1): Reputation score for SDG 16 (Peace, Justice and Strong Institutions) based on stakeholder perceptions and news analysis. Source - Covalence.
104. **COV_REP_SDG17** (1): Reputation score for SDG 17 (Partnerships for the Goals) based on stakeholder perceptions and news analysis. Source - Covalence.
105. **COV_REP_SDG2** (1): Reputation score for SDG 2 (Zero Hunger) based on stakeholder perceptions and news analysis. Source - Covalence.
106. **COV_REP_SDG3** (1): Reputation score for SDG 3 (Good Health and Well-being) based on stakeholder perceptions and news analysis. Source - Covalence.
107. **COV_REP_SDG4** (1): Reputation score for SDG 4 (Quality Education) based on stakeholder perceptions and news analysis. Source - Covalence.
108. **COV_REP_SDG5** (1): Reputation score for SDG 5 (Gender Equality) based on stakeholder perceptions and news analysis. Source - Covalence.
109. **COV_REP_SDG6** (1): Reputation score for SDG 6 (Clean Water and Sanitation) based on stakeholder perceptions and news analysis. Source - Covalence.
110. **COV_REP_SDG7** (1): Reputation score for SDG 7 (Affordable and Clean Energy) based on stakeholder perceptions and news analysis. Source - Covalence.
111. **COV_REP_SDG8** (1): Reputation score for SDG 8 (Decent Work and Economic Growth) based on stakeholder perceptions and news analysis. Source - Covalence.
112. **COV_REP_SDG9** (1): Reputation score for SDG 9 (Industry, Innovation and Infrastructure) based on stakeholder perceptions and news analysis. Source - Covalence.
113. **CSR_OVERALL** (1): Overall CSR Hub rating on a 0-100 scale that measures a company's overall ESG performance by aggregating ratings across all categories. Source - CSR Hub.

114. **CSR_COMM** (1): Community category rating that evaluates a company's commitment and effectiveness within local, national and global communities, including citizenship, charitable giving, and volunteerism. Source - CSR Hub.

115. **CSR_EMP** (1): Employees category rating that assesses policies and performance in diversity, labor relations, labor rights, compensation, benefits, and employee training, health and safety. Source - CSR Hub.

116. **CSR_ENV** (1): Environment category rating that measures a company's interactions with the environment, including resource use, emissions, and environmental policies. Source - CSR Hub.

117. **CSR_GOV** (1): Governance category rating that evaluates disclosure of policies, board independence, executive compensation, and ethical leadership practices. Source - CSR Hub.

118. **CSR_COMM_DEV** (1): Community Development & Philanthropy rating measuring community citizenship through charitable giving and community impact management. Source - CSR Hub.

119. **CSR_PROD** (1): Product rating evaluating product responsibility, development, design, and impacts on customers and society. Source - CSR Hub.

120. **CSR_RIGHTS** (1): Human Rights & Supply Chain rating assessing commitment to human rights conventions and supply chain management. Source - CSR Hub.

121. **CSR_BEN** (1): Compensation & Benefits rating measuring workforce loyalty through fair compensation and benefits. Source - CSR Hub.

122. **CSR_LABOR** (1): Diversity & Labor Rights rating evaluating workplace policies for fair treatment and diversity. Source - CSR Hub.

123. **CSR_H_C** (1): Training, Health & Safety rating assessing workplace safety and employee development programs. Source - CSR Hub.

124. **CSR_CLIMATE** (1): Energy & Climate Change rating measuring effectiveness in addressing climate change and energy efficiency. Source - CSR Hub.

125. **CSR_ENV_POL** (1): Environment Policy & Reporting rating evaluating environmental policies and reporting practices. Source - CSR Hub.

126. **CSR_RESOURCE** (1): Resource Management rating assessing efficiency in resource use and supply chain management. Source - CSR Hub.

127. **CSR_BOARD** (1): Board rating evaluating board membership, independence, and effectiveness. Source - CSR Hub.

128. **CSR_ETHICS** (1): Leadership Ethics rating measuring stakeholder relationship management and ethical decision-making culture. Source - CSR Hub.

129. **CSR_TRANS** (1): Transparency & Reporting rating assessing corporate transparency and alignment with sustainability goals. Source - CSR Hub.

130. **MOODYS_HR** (1): Overall Human Resources score evaluating company practices in labor relations, working conditions, and human rights in the workplace. Source - Moody's ESG Equitics.
131. **MOODYS_HRL** (1): Leadership score in Human Resources, assessing management commitment and effectiveness in implementing HR policies and systems. Source - Moody's ESG Equitics.
132. **MOODYS_HRI** (1): Implementation score in Human Resources, evaluating the concrete measures and processes put in place. Source - Moody's ESG Equitics.
133. **MOODYS_HRR** (1): Results score in Human Resources, measuring the effectiveness and outcomes of HR policies and practices. Source - Moody's ESG Equitics.
134. **MOODYS_HR1** (1): Promotion of labor relations score, focusing on social dialogue and employee participation. Source - Moody's ESG Equitics.
135. **MOODYS_HR1_1** (1): Labor relations sub-score specifically focused on the respect of trade unions rights and collective bargaining. Source - Moody's ESG Equitics.
136. **MOODYS_HR2** (1): Encouraging employee participation score, evaluating employee involvement in company decisions and information sharing. Source - Moody's ESG Equitics.
137. **MOODYS_HR2_3** (1): Career management sub-score evaluating career development opportunities and training. Source - Moody's ESG Equitics.
138. **MOODYS_HR2_4** (1): Training programs sub-score assessing the quality and accessibility of professional development. Source - Moody's ESG Equitics.
139. **MOODYS_HR3** (1): Responsible management of restructuring score, evaluating how company handles workforce changes. Source - Moody's ESG Equitics.
140. **MOODYS_HR3_2** (1): Sub-score on measures to limit and mitigate impacts of restructuring on employees. Source - Moody's ESG Equitics.
141. **MOODYS_ENV** (1): Overall Environmental score evaluating company's environmental strategy and management system. Source - Moody's ESG Equitics.
142. **MOODYS_ENVL** (1): Leadership score in Environmental management, assessing commitment and strategy. Source - Moody's ESG Equitics.
143. **MOODYS_ENVI** (1): Implementation score in Environmental practices, evaluating concrete measures and processes. Source - Moody's ESG Equitics.
144. **MOODYS_ENVR** (1): Results score in Environmental performance, measuring outcomes and effectiveness. Source - Moody's ESG Equitics.
145. **MOODYS_ENV1** (1): Environmental strategy and eco-design score. Source - Moody's ESG Equitics.
146. **MOODYS_ENV1_1** (1): Sub-score on incorporation of environmental considerations in product/service design. Source - Moody's ESG Equitics.

147. **MOODYS_ENV2** (1): Pollution prevention and control score. Source - Moody's ESG Equitics.

148. **MOODYS_ENV2_2** (1): Sub-score on management of water resources and water pollution. Source - Moody's ESG Equitics.

149. **MOODYS_ENV2_4** (1): Sub-score on management of atmospheric emissions. Source - Moody's ESG Equitics.

150. **MOODYS_ENV2_7** (1): Sub-score on waste management practices. Source - Moody's ESG Equitics.

151. **MOODYS_CS** (1): Overall Community Involvement score evaluating company's impact on local communities. Source - Moody's ESG Equitics.

152. **MOODYS_CSL** (1): Leadership score in Community Involvement, assessing strategy and commitment. Source - Moody's ESG Equitics.

153. **MOODYS_CSI** (1): Implementation score in Community Involvement initiatives. Source - Moody's ESG Equitics.

154. **MOODYS_CSR** (1): Results score in Community Involvement, measuring effectiveness of programs. Source - Moody's ESG Equitics.

155. **MOODYS_CS1** (1): Promotion of social and economic development score. Source - Moody's ESG Equitics.

156. **MOODYS_CS1_1** (1): Sub-score on local social and economic development initiatives. Source - Moody's ESG Equitics.

157. **MOODYS_CS1_2** (1): Sub-score on management of impacts on local communities. Source - Moody's ESG Equitics.

158. **MOODYS_CS3** (1): Contribution to general interest causes score. Source - Moody's ESG Equitics.

159. **MOODYS_CS3_1** (1): Sub-score on charitable giving and philanthropy. Source - Moody's ESG Equitics.

160. **MOODYS_CS3_2** (1): Sub-score on employee involvement in community projects. Source - Moody's ESG Equitics.

161. **MOODYS_CG** (1): Overall Corporate Governance score. Source - Moody's ESG Equitics.

162. **MOODYS_CGL** (1): Leadership score in Corporate Governance, assessing board effectiveness. Source - Moody's ESG Equitics.

163. **MOODYS_CGI** (1): Implementation score in Corporate Governance practices. Source - Moody's ESG Equitics.

164. **MOODYS_CGR** (1): Results score in Corporate Governance performance. Source - Moody's ESG Equitics.

165. **MOODYS_CG1** (1): Board of Directors score. Source - Moody's ESG Equitics.

166. **MOODYS_CG1_1** (1): Sub-score on board independence and effectiveness. Source - Moody's ESG Equitics.

167. **MOODYS_CG2** (1): Audit and internal controls score. Source - Moody's ESG Equitics.

168. **MOODYS_CG2_1** (1): Sub-score on internal control systems and risk management. Source - Moody's ESG Equitics.

169. **MOODYS_CG3** (1): Shareholders rights score. Source - Moody's ESG Equitics.

170. **MOODYS_CG3_1** (1): Sub-score on protection of shareholder rights. Source - Moody's ESG Equitics.

171. **MOODYS_CG4** (1): Executive remuneration score. Source - Moody's ESG Equitics.

172. **MOODYS_CG4_1** (1): Sub-score on alignment of executive compensation with stakeholder interests. Source - Moody's ESG Equitics.

173. **MOODYS_CIN** (1): Overall Corporate Business Behavior score. Source - Moody's ESG Equitics.

174. **MOODYS_CINL** (1): Leadership score in Business Behavior, assessing ethics and compliance. Source - Moody's ESG Equitics.

175. **MOODYS_CINI** (1): Implementation score in Business Behavior practices. Source - Moody's ESG Equitics.

176. **MOODYS_CINR** (1): Results score in Business Behavior performance. Source - Moody's ESG Equitics.

177. **MOODYS_CIN1** (1): Product safety and customer relations score. Source - Moody's ESG Equitics.

178. **MOODYS_CIN1_1** (1): Sub-score on product safety management. Source - Moody's ESG Equitics.

179. **MOODYS_CIN2** (1): Information to customers score. Source - Moody's ESG Equitics.

180. **MOODYS_CIN2_1** (1): Sub-score on responsible marketing and customer information. Source - Moody's ESG Equitics.

181. **MOODYS_HRTS** (1): Overall Human Rights score. Source - Moody's ESG Equitics.

182. **MOODYS_HRTSL** (1): Leadership score in Human Rights protection. Source - Moody's ESG Equitics.

183. **MOODYS_HRTSI** (1): Implementation score in Human Rights practices. Source - Moody's ESG Equitics.

184. **MOODYS_HRTSR** (1): Results score in Human Rights performance. Source - Moody's ESG Equitics.
185. **MOODYS_HRTS1** (1): Fundamental human rights protection score. Source - Moody's ESG Equitics.
186. **MOODYS_HRTS1_1** (1): Sub-score on prevention of human rights violations. Source - Moody's ESG Equitics.
187. **MOODYS_HRTS2** (1): Civil and political rights score. Source - Moody's ESG Equitics.
188. **MOODYS_HRTS2_1** (1): Sub-score on protection of civil liberties. Source - Moody's ESG Equitics.
189. **MOODYS_HRTS2_4** (1): Sub-score on non-discrimination practices. Source - Moody's ESG Equitics.
190. **MOODYS_OVERALL** (1): Overall ESG score combining all dimensions of environmental, social and governance performance. Source - Moody's ESG Equitics.
191. **GAIA_OVERALL** (1): Overall ESG score that combines Environmental, Social and Governance pillars using dynamic materiality weightings. Scores range from 0-100. Source - GaiaLens.
192. **GAIA_TRANSP** (1): Overall transparency score reflecting how well populated a company's ESG data is compared to peers. Used as a proxy for confidence in the overall score. Source - GaiaLens.
193. **GAIA_E** (1): Environmental pillar score based on themes including energy production & consumption, environmental impacts, GHG emissions, waste management and water management. Source - GaiaLens.
194. **GAIA_S** (1): Social pillar score incorporating assessment of workforce practices, community impact and customer relationships. Source - GaiaLens.
195. **GAIA_G** (1): Governance pillar score evaluating business ethics, model & innovation, risk & crisis management and leadership. Source - GaiaLens.
196. **GAIA_ETHICS** (1): Theme score focused on business ethics, measuring company practices and policies related to ethical business conduct. Source - GaiaLens.
197. **GAIA_CUST** (1): Theme score evaluating customer-related practices, including customer welfare, privacy and product safety. Source - GaiaLens.
198. **GAIA_ENVIRO** (1): Theme score assessing environmental impacts across company operations and supply chain. Source - GaiaLens.
199. **GAIA_GHG** (1): Theme score specifically focused on greenhouse gas emissions and climate-related practices. Source - GaiaLens.
200. **GAIA_LEADER** (1): Theme score evaluating leadership and management practices including board composition and executive conduct. Source - GaiaLens.

201. **GAIA_WASTE** (1): Theme score assessing waste management practices and circular economy initiatives. Source - GaiaLens.

202. **GAIA_WATER** (1): Theme score focused on water management practices including water usage, recycling and impact on water resources. Source - GaiaLens.

203. **GAIA_WORK** (1): Theme score evaluating workforce-related practices including diversity, health & safety, and labor relations. Source - GaiaLens.

204. **SAUTNER_CC_EXP** (1): Overall climate change exposure score capturing how frequently climate change-related topics are discussed in earnings calls, measured using a machine learning keyword discovery algorithm that identifies relevant bigrams. Source - Climate Exposure Sautner et al. (2023).

205. **SAUTNER_CC_RISK** (1): Climate change risk exposure capturing discussions specifically related to risks and uncertainties of climate change in earnings calls. Source - Climate Exposure Sautner et al. (2023).

206. **SAUTNER_CC_POS** (1): Positive climate change sentiment score measuring discussions of climate change topics in conjunction with positive tone words in earnings calls. Source - Climate Exposure Sautner et al. (2023).

207. **SAUTNER_CC_NEG** (1): Negative climate change sentiment score measuring discussions of climate change topics in conjunction with negative tone words in earnings calls. Source - Climate Exposure Sautner et al. (2023).

208. **SAUTNER_CC_SENT** (1): Overall climate change sentiment score combining positive and negative tone measures around climate discussions in earnings calls. Source - Climate Exposure Sautner et al. (2023).

209. **SAUTNER_OP_EXP** (1): Climate change opportunity exposure capturing discussions related to business opportunities from climate change, including renewable energy and green technologies. Source - Climate Exposure Sautner et al. (2023).

210. **SAUTNER_OP_RISK** (1): Risk exposure specifically related to climate change opportunities, focusing on uncertainties around green business opportunities. Source - Climate Exposure Sautner et al. (2023).

211. **SAUTNER_OP_POS** (1): Positive sentiment score specifically for climate change opportunity discussions. Source - Climate Exposure Sautner et al. (2023).

212. **SAUTNER_OP_NEG** (1): Negative sentiment score specifically for climate change opportunity discussions. Source - Climate Exposure Sautner et al. (2023).

213. **SAUTNER_OP_SENT** (1): Overall sentiment score specifically for climate change opportunity discussions. Source - Climate Exposure Sautner et al. (2023).

214. **SAUTNER_RG_EXP** (1): Regulatory climate change exposure capturing discussions related to climate regulations and policies. Source - Climate Exposure Sautner et al. (2023).

215. **SAUTNER_RG_RISK** (1): Risk exposure specifically related to climate change regulations and policies. Source - Climate Exposure Sautner et al. (2023).

216. **SAUTNER_RG_POS** (1): Positive sentiment score for discussions around climate change regulations. Source - Climate Exposure Sautner et al. (2023).

217. **SAUTNER_RG_NEG** (1): Negative sentiment score for discussions around climate change regulations. Source - Climate Exposure Sautner et al. (2023).

218. **SAUTNER_RG_SENT** (1): Overall sentiment score for discussions around climate change regulations. Source - Climate Exposure Sautner et al. (2023).

219. **SAUTNER_PH_EXP** (1): Physical climate change exposure capturing discussions related to physical climate impacts like extreme weather events. Source - Climate Exposure Sautner et al. (2023).

220. **SAUTNER_PH_POS** (1): Positive sentiment score for discussions around physical climate change impacts. Source - Climate Exposure Sautner et al. (2023).

221. **SAUTNER_PH_NEG** (1): Negative sentiment score for discussions around physical climate change impacts. Source - Climate Exposure Sautner et al. (2023).

222. **SAUTNER_PH_SENT** (1): Overall sentiment score for discussions around physical climate change impacts. Source - Climate Exposure Sautner et al. (2023).

B Additional descriptive statistics

B.1 Additional descriptive statistics of alpha-potential

For each factor we compute the in-sample alpha (α_s) as the annualized time-series average of the adjusted factor returns $\ddot{F}_{s,t}$ (see description in the main document). To characterize the distribution of a set of factors containing K factors, in table B.1 we report the average alpha of the factor set ($\overline{\alpha} = \frac{1}{K}\sum_{s=1}^{K}\alpha_s$), the average absolute alpha ($\overline{|\alpha|} = \frac{1}{K}\sum_{s=1}^{K}|\alpha_s|$), the average alpha standard error ($\overline{s.e.(\alpha)} = \frac{1}{K}\sum_{s=1}^{K}s.e.(\alpha_s)$), the cross-sectional standard deviation of the alphas in the factor set ($cs.std(\alpha) = \left(\frac{1}{K}\sum_{s=1}^{K}(\alpha_s - \overline{\alpha})^2\right)^{1/2}$), and the percentage of statistically significant alphas above (% St.Sign. ($\alpha > 0$)) and below 0 (% St.Sign.($\alpha < 0$)). These statistics help us to understand how the ESG factors compare to the Financial factors in terms of their ability to capture market-adjusted returns in-sample.

We report the results in table B.1. over the sample period, we see that the average ESG factor generated a positive annualized alpha of 4.7%, which means that taking a positive exposure to sustainability would have generated a positive market-adjusted return. However, we are more interested in the average absolute alpha (column 2), because our reference investor is an ESG-aware investor, who is willing to use ESG information without restriction on the direction of the exposure to sustainability. We find that the average absolute alpha for the ESG factors is 5.5% annually, which is close to the 5.8% observed for the factors in the published return effect set and higher than the 4.4% observed for the general firm characteristics set. In addition, the average standard error in the ESG set (4.8) is almost the same as the one observed for the Published return effects and General firm characteristics set. We also note that the percent of ESG factors with statistically significant alphas (positive and negative) is substantially lower than the one observed for the Published return effects factors but larger than the one of the General firm characteristics factors.

[Table B.1 about here]

We report in Table B.2. the top 5 factors by absolute value of CAPM alpha over the full sample of our analysis (August 2008 to June 2023) for three clusters of factors – financials, ESG traditional, and ESG alternative.

[Table B.2 about here]

We also report the summary of the Fama-French 6 factors alphas[43] in table B.3. Even correcting for exposure these standard 6 factors, one could still find many ESG factors that generate a significant alpha in-sample. Table B.4 reports the largest FF6 alphas by factor cluster.

[Table B.3 about here]

[Table B.4 about here]

B.2 Additional summary descriptive statistics of diversification-potential

In figure B.1 we report the minimal number of orthogonal dimensions – principal components (PCs) – needed to explain a given proportion of the covariance matrix the Financial factors only and of the Financial plus ESG factors. We can see that adding ESG to the Financial factors alters the structure of the in-sample covariance matrix[44]. Adding ESG factors one need more independent dimensions to explain a given percentage of variance. For instance, 36 PCs are needed to explain 94% of the variance generated by the financial factors alone, but we need 154 PCs to explain the same level of variance when we add ESG factors to the

[43] These are the factors included in Fama and French (2018) – SMB, HML, RMW, CMA, and Momentum. We downloaded the time series of the monthly returns of these factors from K. French website: `https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

[44] Because of the large number of factors in our sets, to guarantee invertibility of the covariance matrix, we use the Ledoit and Wolf (2004) covariance matrix shrunk with a Wishart prior, meaning that the elements of the covariance matrix are shrunk towards those of the identity matrix. In addition, we use daily market-adjusted returns of the factors, which keeps to a minimum the impact of shrinkage.

Financials. This is greater than the number of financial factors, so greater than the number of orthogonal dimensions that the financial factors can mathematically generate.

[Figure B.1 about here]

In other words, the in-sample results suggest that ESG factors add truly independent dimensions to the set of financial factors, which provides the impression that ESG factors have the potential to improve risk diversification.

C Details on Multifactor Portfolio Methods

C.1 Bayesian methods – details on the derivation of the formula

The first approach that we use to deal with the high-dimensionality challenge is the Bayesian approach. To mitigate overfitting, we incorporate economically motivated priors in the optimal allocation problem. We follow KNS to define the structure of the prior, and focus on shrinking expected returns as they have the greater impact on weights. We specify priors on market-adjusted factor returns ($\boldsymbol{\mu}$) as follows

$$\boldsymbol{\mu} \sim \mathcal{N}\left(\mathbf{0}, \frac{\kappa^2}{\tau}\boldsymbol{\Sigma}^{\eta}\right) \tag{C.1}$$

Where τ is the trace of the covariance matrix of the market-adjusted returns ($\boldsymbol{\Sigma}$) and κ is a constant controlling the scale of $\boldsymbol{\mu}$. We will show that κ can be interpreted as the maximum achievable Sharpe ratio. This means that the in-sample estimates of the expected returns are shrunk towards the CAPM predictions and the confidence in the CAPM predictions is lower the greater the observed deviation from the CAPM predictions.

To find the optimal weights one first computes the posterior expected factor returns ($\widehat{\boldsymbol{\mu}}$) and then substitutes them in the optimal mean-variance weight ($\boldsymbol{w}^* = \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$). Assuming a normal distribution for both the prior and the likelihood, the posterior expected return is

$$\widehat{\boldsymbol{\mu}} = \left(\boldsymbol{\Sigma} + \gamma\boldsymbol{\Sigma}^{2-\eta}\right)^{-1}\boldsymbol{\Sigma}\overline{\boldsymbol{\mu}} \tag{C.2}$$

Where $\overline{\boldsymbol{\mu}}$ is the vector of the factors' mean returns and $\gamma = \frac{\tau}{\kappa^2 T}$.

To find the optimal weights ($\widehat{\boldsymbol{w}}$), KNS shows that imposing $\eta = 2$ provides a solution that is a ridge-like estimator

$$\widehat{\boldsymbol{w}} = \left(\boldsymbol{\Sigma} + \gamma\boldsymbol{I}\right)^{-1}\overline{\boldsymbol{\mu}} \tag{C.3}$$

Where we can interpret γ as the ridge-like penalization that regularize/shrinks the weights.

C.2 Bayesian methods – details on hyperparameter tuning in the Data-driven approach

For the data-driven shrinkage approach we need to tune the hyperparameter (κ) regulating the level of shrinkage.

To guarantee that our results are fully out-of-sample, for the tuning we use only data that is outside the testing sample. The sample used for tuning is from August 1975 to July 2008 (tuning sample). We use only portfolios whose returns are available over the entire tuning-sample, meaning only our 130 Financial factors. Separating the tuning from the testing sample ensures that our out-of-sample (OOS) tests' results are truly OOS.

As in KNS, we use volatility-scaled market-beta-adjusted daily factor returns observations ($\ddot{F}_{i,t}$)[45] and we use a 3-fold cross-validation (CV). We split the tuning sample in 3 equal subsamples. Then, for each of the three subsamples we compute $\widehat{\boldsymbol{w}}$ using equation (C.3) using the average return and the covariance matrix of volatility-scaled market-adjusted returns estimated over the other 2 subsamples, and using a set of 1000 equally-spaced values for the parameter κ that goes from 0.01 to 10. For each $\widehat{\boldsymbol{w}}$ we evaluate the OOS fit of the resulting model of the expected returns computing the OOS cross-sectional R-squared as follows

$$R^2_{OOS} = 1 - \frac{(\boldsymbol{\mu}_2 - \boldsymbol{\Sigma}_2\widehat{\boldsymbol{w}})'\,(\boldsymbol{\mu}_2 - \boldsymbol{\Sigma}_2\widehat{\boldsymbol{w}})}{\boldsymbol{\mu}_2'\boldsymbol{\mu}_2} \tag{C.4}$$

Where the subscript 2 indicates the sample moments of the withheld sample. For each level of κ we take the average R^2_{OOS} across the three subsamples and find the optimal κ^* for which the average R^2_{OOS} is maximized.

We report the graphical representation of the relation between the hyperparameter κ and the average cross-sectional R-squared, both in and out-of-sample (Figure C.1), and the effect of shrinkage on the mean-variance-efficient (MVE) portfolio weights $\widehat{\boldsymbol{w}}$ (Figure C.2).

[Figure C.1 about here]

[Figure C.2 about here]

C.3 Constrained methods – details on mean-variance constrained approaches

For each of the mean-variance constrained approaches, we find the optimal set of weights $\boldsymbol{w}$ that solves the following problem using the MATLAB 'fmincon' function

[45] See main document for details on how we compute $\ddot{F}_{i,t}$.

$$
\max_{\boldsymbol{w}} \quad \frac{\boldsymbol{w}'\boldsymbol{\mu}^*}{\sqrt{\boldsymbol{w}'\boldsymbol{\Sigma}^*\boldsymbol{w}}}
$$

$$
\text{s.t.} \quad \begin{cases} \boldsymbol{w}'\boldsymbol{e} = 1 \\ \boldsymbol{w} \geq \boldsymbol{0} \\ \sum_{i=1}^{N} \frac{1}{w_i^2} \geq \theta \end{cases} \tag{C.5}
$$

Where $\boldsymbol{\mu}^*$ and $\boldsymbol{\Sigma}^*$ are respectively the vector of the sample averages and the covariance matrix of the volatility-scaled market adjusted returns $\ddot{F}_{i,t}$ multiplied by 1 (-1) if the sample average $\boldsymbol{\mu}$ is positive (negative). Both $\boldsymbol{\mu}^*$ and $\boldsymbol{\Sigma}^*$ are estimated over the estimation window and the sign-adjustment of $\ddot{F}_{i,t}$ also concerns only the estimation window. The vector $\boldsymbol{e}$ is a vector of ones, $\boldsymbol{I}$ is an $N \times N$ identity matrix (N is the number of factors), and θ is the minimum allowed effective number of factors, which can be equal to 10, 30, or 50. This approach imposes to invest only in factors that had a positive adjusted return over the estimation window while limiting the concentration.

C.4 Constrained methods – details on minimum-volatility constrained approaches

For each of the minimum-volatility constrained approaches we solve the following problem using the MATLAB *CVX toolbox* from Grant and Boyd (2008, 2014)

$$
\min_{\boldsymbol{w}} \quad \sqrt{\boldsymbol{w}'\boldsymbol{\Sigma}^*\boldsymbol{w}}
$$

$$
\text{s.t.} \quad \begin{cases} \boldsymbol{w}'\boldsymbol{e} = 1 \\ \boldsymbol{w} \geq \boldsymbol{0} \\ \sum_{i=1}^{N} \frac{1}{w_i^2} \geq \theta \end{cases} \tag{C.6}
$$

Where $\boldsymbol{\Sigma}^*$ is the covariance matrix of the volatility-scaled market adjusted returns $\ddot{F}_{i,t}$ multiplied by 1 (-1) if the sample average $\boldsymbol{\mu}$ is positive (negative). To estimate $\boldsymbol{\Sigma}^*$ we use only the estimation window and the sign-adjustment of $\ddot{F}_{i,t}$ also concerns only the estimation window. The vector $\boldsymbol{e}$ is a vector of ones, $\boldsymbol{I}$ is an $N \times N$ identity matrix (N is the number of factors), and θ is the minimum allowed effective number of factors, which can be equal to 10, 30, or 50. This approach imposes to invest only in factors that had a positive adjusted return over the estimation window while limiting the concentration.

D Additional Empirical Analyses

D.1 Effective number of risk contributors

Here we report an analysis that allows us to compare the level of diversification across factors achieved by our multifactor portfolios, which allows us to handle the fact that we deal with portfolios that have both long and short positions in the factors.

We consider a measure of portfolio diversification (or de-concentration) that takes into account not only the weights but also how much each factor contributes to the total risk of the portfolio through its volatility and correlation with the other factors[46].

We follow Maillard et al. (2010) (MRT) and compute the marginal risk contribution (marginal contribution to the total risk of the portfolio) of each factor for a portfolio $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)$ as

$$\partial_{w_i}\sigma(\boldsymbol{w}) = \frac{\partial\sigma(\boldsymbol{w})}{\partial w_i} = \frac{w_i\sigma_i^2 + \sum_{j\neq i} w_j\sigma_{ij}}{\sigma(\boldsymbol{w})} \tag{D.1}$$

Where $\sigma(\boldsymbol{w})$ is the volatility of the of the portfolio, σ_i^2 the variance of $\ddot{F}_i$ and σ_{ij} the covariance between $\ddot{F}_i$ and $\ddot{F}_j$. MRT also show that the total risk contribution of $\ddot{F}_i$ is $RC_i = w_i\partial_{w_i}\sigma(\boldsymbol{w})$ and that the sum of the risk contributions of all the factors included in the portfolio must equal the total volatility of the portfolio. It follows ϕ_i, the sum of the ratio of RC_i over the total portfolio risk is equal to 1

$$\sum_i \frac{RC_i}{\sigma(\boldsymbol{w})} = \sum_i \phi_i = 1 \tag{D.2}$$

Because the sum of ϕ_i is always equal to 1, we can compare different portfolios even if the sums of their weights differ. As a measure of diversification, we compute the Effective Number of Risk Contributors (ENRC)

$$ENRC = \frac{1}{\sum_i {\phi_i}^2} \tag{D.3}$$

This follows the idea of risk parity, if all the factors equally contribute to the overall risk of the portfolio the risk of the portfolio will be perfectly distributed across all the risk contributors, hence the ENRC will equal the number of factors in the portfolio, otherwise it will be lower.

We report the ENRC of our in-sample multifactor portfolios in figure D.1. We note that multifactor portfolios based on Bayesian/shrinkage approaches improve diversification across factors (ENRC) much more than the other approaches.

[46]However, because we normalize volatilities of all the factors, the only important aspect of the risk contribution will be correlation.

[Figure D.1 about here]

D.2 Equivalent max expected Sharpe ratio

We report an additional analysis that allows us to evaluate the levels of diversification across factors in the different multifactor portfolios.

Because the in-sample Sharpe ratio is a monotonically increasing function of the max expected Sharpe ratio, we can find the equivalent max Sharpe ratio of each of our multifactor portfolios as the parameter $\widehat{\kappa}$ that equalizes the ex-post in-sample Sharpe ratio of the portfolio obtained with the formula $\widehat{\boldsymbol{w}} = (\boldsymbol{\Sigma} + \gamma \boldsymbol{I})^{-1} \overline{\boldsymbol{\mu}}$ to the in-sample Sharpe ratio of our multifactor portfolios. This gives us an alternative way (relative to the ENRC) to compare the level of shrinkage of the different portfolios.

We report the equivalent max Sharpe ratios in table D.1. For the Naïve portfolios the penalization must be zero, this means that we are looking at the portfolio of an investor that allows his expectation on the maximum achievable Sharpe ratio to be arbitrarily large. For the Constrained Mean-variance portfolios $\widehat{\kappa}$ is between 5 and 8 times the Sharpe ratio of the market portfolio (0.46), which is much larger than the limit allowed for the Shrinkage portfolios.

[Table D.1 about here]

D.3 Robustness of results on Sharpe ratio improvements to different ESG sub-selections

We show in table D.2 that our finding that adding ESG factors does not improve Sharpe ratios of a Financials only multifactor portfolios OOS is robust to using different ESG sub-selections. We look at four sub-selections of the ESG factors; the first sub-selection includes only the Consensus ESG factors[47], the second excludes ESG scores that aggregate multiple sub-categories[48], the third selection focuses on the Traditional ESG factors, and finally one that includes only the Alternative ESG sub-selections. In table D.2 we report only the annualized Sharpe ratios improvements of the combined multifactor portfolio (Financials plus ESG sub-selection) relative to the financial multifactor benchmark.

We report both in-sample (panel A) and out-of-sample (panel B) improvements in Sharpe ratios. While we find some improvements in-sample, none of the results are significant out-of-sample.

[Table D.2 about here]

[47]Based on CSRHub scores.

[48]We remove ESG scores that represent macro ESG pillars such as ESG, E, S, and G.

D.4 Robustness of results on Sharpe ratio improvements to different benchmarks

We show that our finding that adding ESG factors does not improve Sharpe ratios of a financial factors only multifactor portfolios OOS is robust to using different benchmarks.

In table D.3 we use a sparse benchmark that includes only the 5 cross-sectional factors from Fama and French (2018) – Size (SMB), Value (HML), Profitability (RMW), Investment (CMA), and Momentum (MOM)[49]. As sparsity is already a form of regularization, we form the sparse benchmark using just the naïve portfolio construction approach and we use this portfolio as benchmark for all the multifactor portfolios based on the ESG and Financials plus ESG sets of factors. Note that the Financials set of factors used for the Financials plus ESG multifactor portfolios uses our high-dimensional set of 130 financial factors, thus in this case it is only the benchmark that changes relative to our base case results reported in the main document.

[Table D.3 about here]

We also assess the robustness of our results to changes of the high-dimensional set of financial factors[50]. We use the high-dimensional set of financial factors from Jensen et al. (2023) – we call this set of financial factors JKP factors[51]. We use the JKP factors both to construct the benchmark and to construct the combined Financials plus ESG multifactor portfolios. We report the results in table D.4.

Results in both tables confirm our main finding – ESG factors do not add value OOS.

[Table D.4 about here]

Appendix References

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *Journal of Finance* 61, 259–299.

Asness, Clifford, and Andrea Frazzini, 2013, The devil in hml's details, *Journal of Portfolio Management* 39, 49–68.

[49] We download the time series of Fama-French factors' returns from Kenneth French database: `https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

[50] This is important to check as Bessembinder, Burt, and Hrdlicka (2025) find that popular high-dimensional sets of financial factors do not span each other. This implies that the result that our ESG set of factors do not help to improve mean-variance efficiency OOS could change when using a different set of financial factors as benchmark.

[51] The JKP factors are available at `https://jkpfactors.com/`. Note that this set of factors is constructed with a different approach relative to the one that we use; they are also long-short portfolios, but they are based on the top and bottom terciles of the distribution based on the characteristic and each leg is a value-weighted portfolio with the market equity windsorized at the NYSE 80$^{\text{th}}$ percentile.

Asquith, Paul, Parag A. Pathak, and Jay R. Ritter, 2005, Short interest, institutional ownership, and stock returns, *Journal of Financial Economics* 78, 243–276.

Barbee Jr, William C., Sandip Mukherji, and Gary A. Raines, 1996, Do sales–price and debt–equity explain stock returns better than book–market and firm size?, *Financial Analysts Journal* 52, 56–60.

Basu, Sanjoy, 1977, Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis, *Journal of Finance* 32, 663–682.

Bhandari, Laxmi Chand, 1988, Debt/equity ratio and expected common stock returns: Empirical evidence, *Journal of Finance* 43, 507–528.

Blume, Marshall E., and Frank Husic, 1973, Price, beta, and exchange listing, *Journal of Finance* 28, 283–299.

Chen, Long, Robert Novy-Marx, and Lu Zhang, 2011, An alternative three-factor model, Working Paper 1418117, SSRN.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609–1651.

Da, Zhi, Qianqiu Liu, and Ernst Schaumburg, 2014, A closer look at the short-term return reversal, *Management Science* 60, 658–674.

Daniel, Kent, and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *Journal of Finance* 61, 1605–1643.

Datar, Vinay T., Narayan Y. Naik, and Robert Radcliffe, 1998, Liquidity and stock returns: An alternative test, *Journal of Financial Markets* 1, 203–219.

De Bondt, Werner F.M., and Richard Thaler, 1985, Does the stock market overreact?, *Journal of Finance* 40, 793–805.

Edmans, Alex, 2011, Does the stock market fully value intangibles? employee satisfaction and equity prices, *Journal of Financial Economics* 101, 621–640.

Fairfield, Patricia M., J. Scott Whisenant, and Teri Lombardi Yohn, 2003, Accrued earnings and growth: Implications for future profitability and market mispricing, *Accounting Review* 78, 353–371.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2018, Choosing factors, *Journal of Financial Economics* 128, 234–252.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Grant, Michael, and Stephen Boyd, 2014, *CVX: Matlab Software for Disciplined Convex Programming, version 2.1*.

Grant, Michael C., and Stephen P. Boyd, 2008, Graph implementations for nonsmooth convex programs, in *Recent Advances in Learning and Control*, 95–110 (Springer London).

Heston, Steven L., and Ronnie Sadka, 2008, Seasonality in the cross-section of stock returns, *Journal of Financial Economics* 87, 418–445.

Hirshleifer, David, Kewei Hou, Siew Hong Teoh, and Yinglei Zhang, 2004, Do investors overvalue firms with bloated balance sheets?, *Journal of Accounting and Economics* 38, 297–331.

Hong, Harrison, and Marcin Kacperczyk, 2009, The price of sin: The effects of social norms on markets, *Journal of Financial Economics* 93, 15–36.

Ikenberry, David, Josef Lakonishok, and Theo Vermaelen, 1995, Market underreaction to open market share repurchases, *Journal of Financial Economics* 39, 181–208.

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881–898.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Jensen, Theis Ingerslev, Bryan Kelly, and Lasse Heje Pedersen, 2023, Is there a replication crisis in finance?, *Journal of Finance* 78, 2465–2518.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2020, Shrinking the cross-section, *Journal of Financial Economics* 135, 271–292.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541–1578.

Ledoit, Olivier, and Michael Wolf, 2004, A well-conditioned estimator for large-dimensional covariance matrices, *Journal of Multivariate Analysis* 88, 365–411.

Ledoit, Olivier, and Michael Wolf, 2008, Robust performance hypothesis testing with the sharpe ratio, *Journal of Empirical Finance* 15, 850–859.

Maillard, Sébastien, Thierry Roncalli, and Jérôme Teïletche, 2010, The properties of equally weighted risk contribution portfolios, *Journal of Portfolio Management* 36, 60.

Moskowitz, Tobias J., and Mark Grinblatt, 1999, Do industries explain momentum?, *Journal of Finance* 54, 1249–1290.

Naranjo, Andy, M. Nimalendran, and Mike Ryngaert, 1998, Stock returns, dividend yields, and taxes, *Journal of Finance* 53, 2029–2057.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Piotroski, Joseph D., 2000, Value investing: The use of historical financial statement information to separate winners from losers, *Journal of Accounting Research* 1–41.

Pontiff, Jeffrey, and Artemiza Woodgate, 2008, Share issuance and cross-sectional returns, *Journal of Finance* 63, 921–945.

Sautner, Zacharias, Laurence Van Lent, Grigory Vilkov, and Ruishen Zhang, 2023, Firm-level climate change exposure, *Journal of Finance* 78, 1449–1498.

Sloan, Richard G., 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings?, *Accounting Review* 289–315.

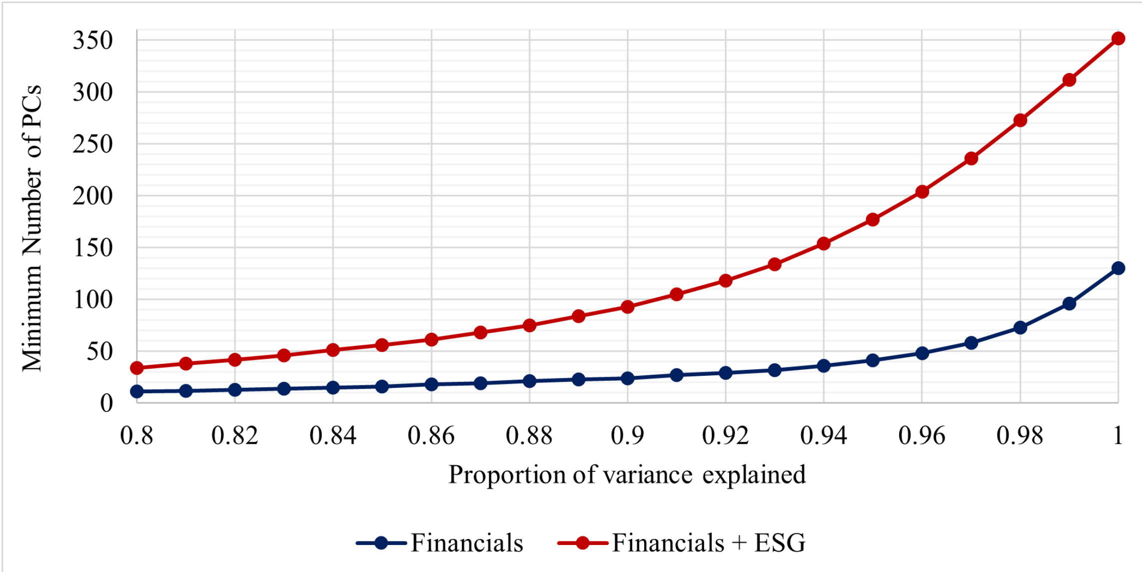
Soliman, Mark T., 2008, The use of dupont analysis by market participants, *Accounting Review* 83, 823–853.

Spiess, D. Katherine, and John Affleck-Graves, 1999, The long-run performance of stock returns following debt offerings, *Journal of Financial Economics* 54, 45–73.

Xing, Yuhang, 2008, Interpreting the value effect through the q-theory: An empirical investigation, *Review of Financial Studies* 21, 1767–1795.

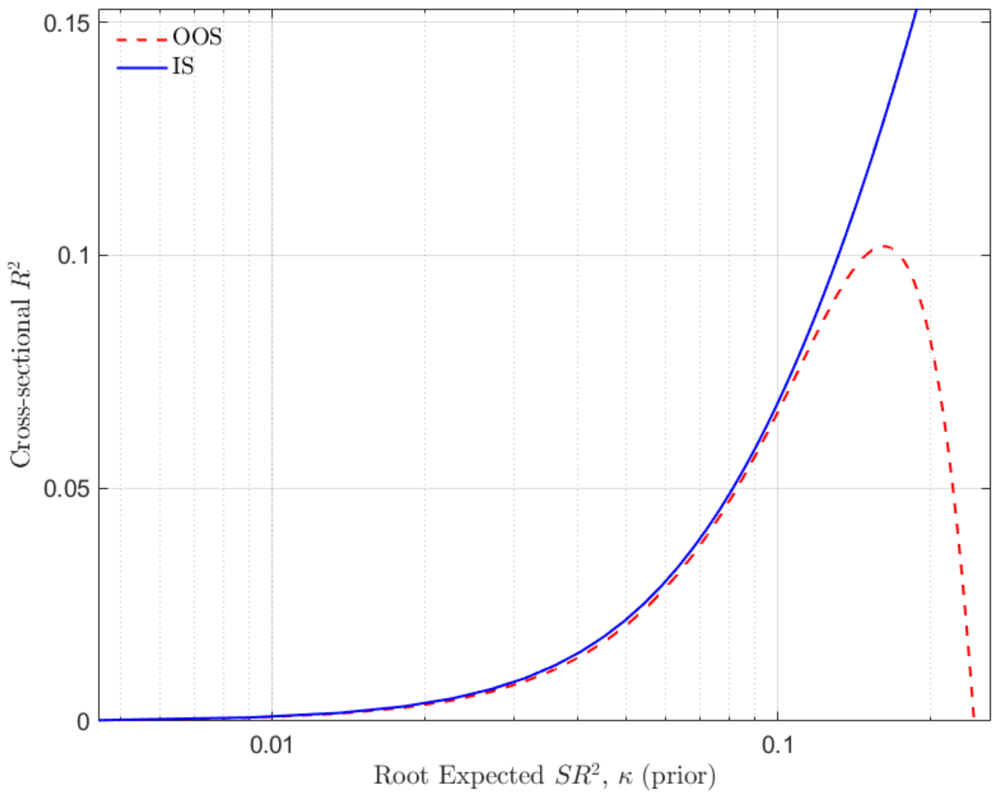
Appendix: Figures and Tables

Figure B.1: Number of PCs needed to explain a given proportion of variance



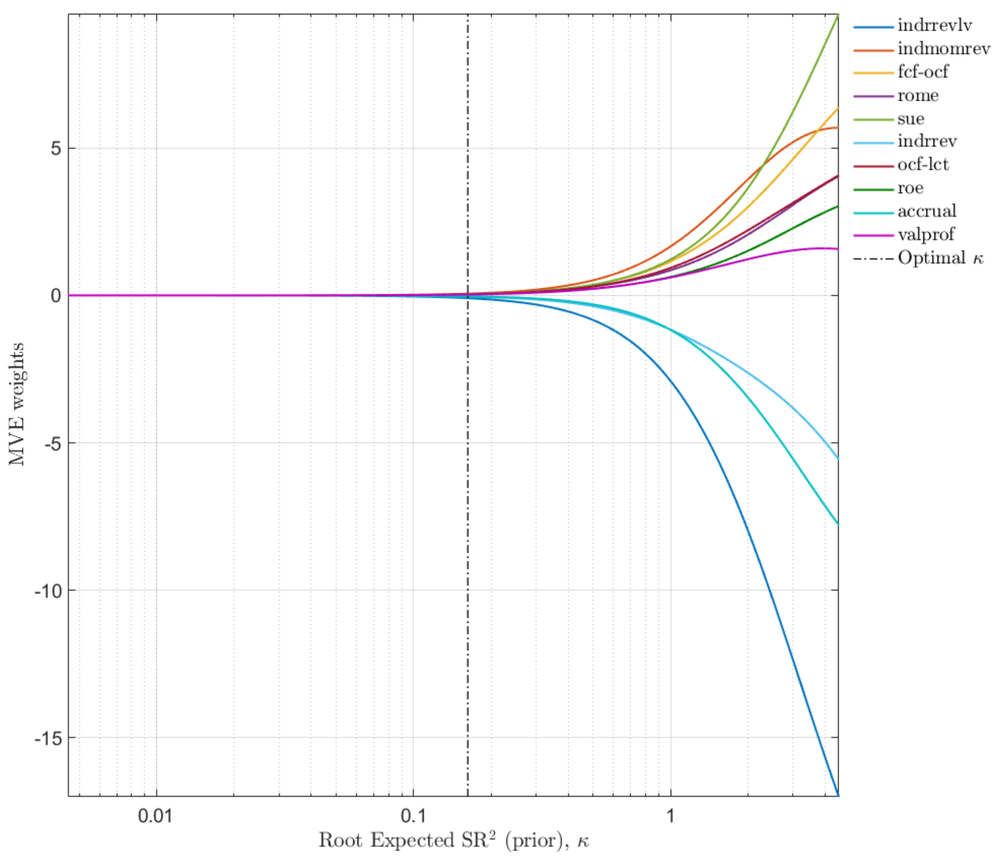
This figure shows on the x-axis the proportion of variance explained and on the y-axis the minimum of principal components (PC) needed to explain a given proportion of variance of a set of factors. We consider four sets of factors: Financials, Financials + ESG Traditional, Financials + ESG Alternative, and Financials + ESG. To compute the covariance matrix, we use the formula of the shrunk covariance matrix from the (Ledoit and Wolf, 2004). We use daily observations of market-adjusted factor returns. To adjust the daily factor returns for market exposure we estimate the market beta using monthly observations. The period is from August 2008 to June 2023.

Figure C.1: Shrinkage parameter selection κ



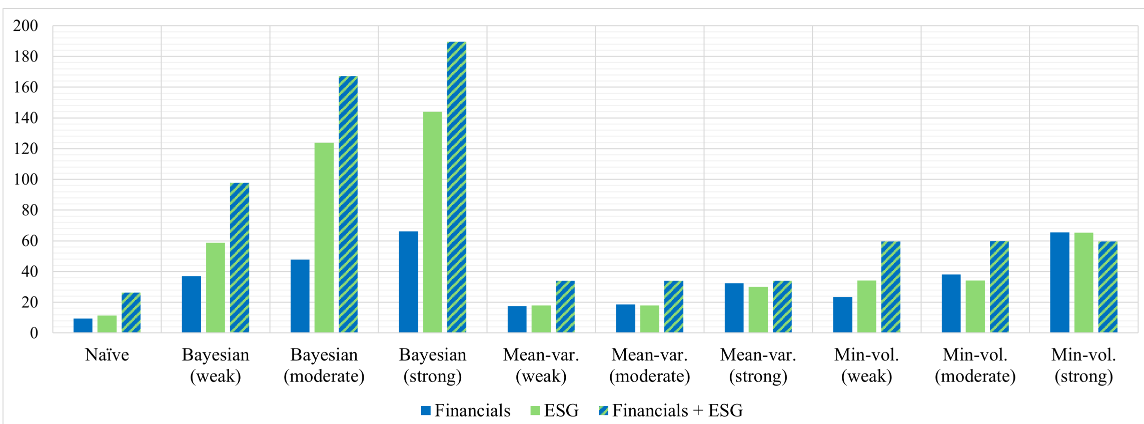
The figure represents the relation between the average cross-sectional R-squared across the three subsamples of the 3-fold cross-validation and the shrinkage parameter κ (red dashed line) and the relation between the in-sample cross-sectional R-squared and κ. We use the volatility-scaled market-adjusted factor returns from August 1975 to July 2008.

Figure C.2: Effect of shrinkage on mean-variance efficient (MVE) weights



The figure represents the relation between the shrinkage parameter κ (x-axis) and the ten largest weights (in absolute value) of the in-sample MVE portfolio (y-axis). The x-axis is on a logarithmic scale. We use the volatility-scaled market-adjusted factor returns from August 1975 to July 2008. Full description of the characteristics can be found in Appendix A

Figure D.1: Effective number of risk contributors (in-sample portfolios)



This figure displays the effective number of risk contributors (ENRC) of multifactor portfolios for three set of base factors – Financials, ESG, and Financials plus ESG. Multifactor portfolios are constructed with 10 portfolio construction approaches - 1 Naive (no control for estimation error) and 9 robust approaches with three different methodologies (Bayesian, constrained Mean-variance, and constrained minimum-volatility) each with three levels of control for estimation error (weak, moderate, and strong). Weights of all the multifactor portfolios are estimated in-sample (from August 2008 to June 2023).

Table B.1: Summary statistics on CAPM alphas (annualized) per factor category

Factor category	$\overline{\alpha}$	$\overline{\lvert\alpha\rvert}$	$\overline{s.e.(\alpha)}$	$cs.std(\alpha)$	%St.Sign. ($\alpha > 0$)	%St.Sign. ($\alpha < 0$)
	(1)	(2)	(3)	(4)	(5)	(6)
Published return effect						
All	4.0	5.8	4.9	5.8	34%	2%
Size	-4.8	4.8	4.7	0.6	0%	0%
Value	-0.4	3.0	5.8	3.7	0%	0%
Profitability	7.1	7.2	4.8	4.7	67%	0%
Low Investment	8.0	8.4	5.3	5.1	64%	0%
Momentum	2.5	2.5	4.1	2.0	0%	0%
Low risk	9.9	9.9	4.6	2.9	100%	0%
Others	1.9	4.5	4.6	5.3	11%	6%
General firm char.						
All	3.0	4.4	4.8	4.5	19%	0%
ESG						
All	4.7	5.5	4.8	4.4	26%	0%
ESG Traditional	4.5	5.6	5.0	4.9	28%	0%
ESG Alternative	4.7	5.5	4.7	4.1	24%	1%

The CAPM alpha of each factor is estimated as the time-series average of the market adjusted factor returns ($F_{i,t} - \beta_i R^e_{M,t}$) scaled to have standard deviations equal to the standard deviation of the excess market returns. The market beta is estimated with an OLS regression using monthly observations. We then annualize and multiply by 100 each alpha. We obtain standard errors (s.e.) of the alphas from the same OLS regression, and adjust for heteroskedasticity and serial correlation. We scale the s.e. multiplying them by the same coefficient used for the alphas, we annualize and multiply them by 100. Statistical significance is obtained from the same OLS regression looking at the p-value. All the statistics that we report are cross-sectional. We report the average alpha (column 1), the average absolute alpha (column 2), the average standard errors (s.e.) of the alphas (column 3), the standard deviation of the alphas (column 4), and the percentage of positive (negative) alphas that are statistically significant (p-value below 0.1) in column 5 (6). The sample period is from August 2008 to June 2023.

Table B.2: Top five factors by absolute value of CAPM Alpha in each factor-cluster

Cluster	ID	Description	CAPM Alpha
Financials	ciss	Composite issuance	13.90
	sale_invcap	Sales/Invested Capital	13.34
	sale_equity	Sales/Stockholders Equity	13.11
	igrowth	Investment growth	12.78
	ivol	Idiosyncratic volatility	11.97
ESG Traditional	CSR_RIGHTS	Human Rights & Supply Chain rating assessing commitment to human rights conventions and supply chain management.	14.47
	CSR_PROD	Product rating evaluating product responsibility, development, design, and impacts on customers and society.	12.06
	CSR_COMM	Community category rating that evaluates a company's commitment and effectiveness within local, national and global communities, including citizenship, charitable giving, and volunteerism.	11.89
	CSR_OVERALL	Overall CSR Hub rating on a 0-100 scale that measures a company's overall ESG performance by aggregating ratings across all categories.	11.75
	CSR_CLIMATE	Energy & Climate Change rating measuring effectiveness in addressing climate change and energy efficiency.	11.14
ESG Alternative	SAUTNER_RG_SENT	Overall sentiment score for discussions around climate change regulations.	17.07
	SINHK_SIN	Companies with significant revenues from sin industries (Tobacco, alcohol, and gaming)	-12.47
	GAIA_TRANSP	Overall transparency score reflecting how well populated a company's ESG data is compared to peers. Used as a proxy for confidence in the overall score.	12.32
	COV_REP_SDG9	Reputation score for SDG 9 (Industry, Innovation and Infrastructure) based on stakeholder perceptions and news analysis	12.25
	SAUTNER_RG_POS	Positive sentiment score for discussions around climate change regulations.	11.83

We report the cluster, the ID, a short description, and the annualized CAPM alpha of each of the top 5 factors by absolute value of CAPM alpha within the cluster. For each factor the CAPM Alpha is computed as the intercept of the linear regression of the monthly factor returns on the returns of the market in excess of the risk-free rate - $F_{i,t} = \alpha_i + \beta_i \left(r_{mkt,t} - r_{rf,t}\right) + \varepsilon_{i,t}$*. For comparability, we scale the alphas multiplying them by the ratio of the volatility of the excess market returns to the volatility of the residuals* $\sigma(r_{mkt,t} - r_{rf,t})/\sigma(\varepsilon_{i,t})$*. We then annualize the alpha and multiply them by 100. The period is from August 2008 to June 2023.*

Table B.3: Summary statistics on FF6 alphas (annualized) per factor category

Factor category	$\overline{\alpha}$	$\overline{\lvert\alpha\rvert}$	$\overline{s.e.(\alpha)}$	$cs.std(\alpha)$	% St.Sign. ($\alpha > 0$)	% St.Sign. ($\alpha < 0$)
	(1)	(2)	(3)	(4)	(5)	(6)
Published return effect						
All	3.4	5.0	4.4	5.8	22%	2%
Size	-0.9	2.7	4.9	3.9	0%	0%
Value	-0.1	1.6	4.5	2.6	0%	0%
Profitability	1.8	3.1	4.4	3.2	0%	0%
Low Investment	8.8	10.1	4.6	6.5	73%	0%
Momentum	2.4	2.4	4.3	2.1	0%	0%
Low risk	6.6	6.6	4.1	5.0	50%	0%
Others	2.1	4.2	4.2	5.7	11%	6%
General firm char.						
All	1.1	3.7	4.3	4.4	9%	4%
ESG						
All	1.9	3.9	4.5	4.6	10%	4%
ESG Traditional	1.5	4.5	4.6	5.2	12%	5%
ESG Alternative	2.1	3.6	4.5	4.2	9%	3%

The Fama-French 6 factors (FF6) alpha of each factor is estimated as the time-series average of the FF6 adjusted factor returns scaled to have standard deviations equal to the standard deviation of the excess market returns. The FF6 factors betas are estimated with an OLS regression using monthly observations. We then annualize and multiply by 100 each alpha. We obtain standard errors (s.e.) of the alphas from the same OLS regression, and adjust for heteroskedasticity and serial correlation. We scale the s.e. multiply them by the same coefficient used for the alphas, we annualize and multiply them by 100. Statistical significance is obtained from the same OLS regression looking at the p-value. All the statistics that we report are cross-sectional. We report the average alpha (column 1), the average absolute alpha (column 2), the average standard errors (s.e.) of the alphas (column 3), the standard deviation of the alphas (column 4), and the percentage of positive (negative) alphas that are statistically significant (p-value below 0.1) in column 5 (6). The sample period is from August 2008 to June 2023.

Table B.4: Top five factors by absolute value of FF6 Alpha in each factor-cluster

Cluster	ID	Description	FF6 Alpha
Financials	igrowth	Investment growth	15.57
	noa	Net operating assets	14.97
	repurch	Stock repurchases	14.44
	nissa	Net stock issuance	13.32
	ciss	Composite issuance	12.47
ESG Traditional	CSR_RIGHTS	Human Rights & Supply Chain rating assessing commitment to human rights conventions and supply chain management.	11.39
	CSR_BEN	Compensation & Benefits rating measuring workforce loyalty through fair compensation and benefits.	10.95
	MOODYS_CG2	Audit and internal controls score	-10.67
	MOODYS_CG2_1	Sub-score on internal control systems and risk management	-10.67
	MOODYS_CGL	Leadership score in Corporate Governance, assessing board effectiveness	-10.45
ESG Alternative	SAUTNER_RG_SENT	Overall sentiment score for discussions around climate change regulations.	16.94
	BC_BC	Best companies for employees to work for	12.06
	COV_DISCLO_GOV9	Disclosure score for shareholders based on reported indicators	11.22
	GAIA_TRANSP	Overall transparency score reflecting how well populated a company's ESG data is compared to peers. Used as a proxy for confidence in the overall score.	10.86
	SAUTNER_RG_POS	Positive sentiment score for discussions around climate change regulations.	9.90

We report the cluster, the ID, a short description, and the annualized FF6 alpha of each of the top 5 factors by absolute value of FF6 alpha within the cluster. For each factor the FF6 Alpha is computed as the intercept of the linear regression of the monthly factor returns on the returns of the 6 factors used in Fama and French (2018) – excess market returns, SMB, HML, RMW, CMA, and Momentum. For comparability, we scale the alphas multiplying them by the ratio of the volatility of the excess market returns to the volatility of the residuals $\sigma(r_{mkt,t} - r_{rf,t})/\sigma(\varepsilon_{i,t})$. We then annualize the alpha and multiply them by 100. The period is from August 2008 to June 2023.

Table D.1: **Equivalent Max Expected Sharpe ratios ($\hat{\kappa}$).**

	Multifactor portfolios		
	Financials only	**ESG only**	**Financials + ESG**
Naïve	∞	∞	∞
Bayesian			
Weak	0.91	0.91	0.91
Moderate	0.16	0.16	0.16
Strong	0.00	0.00	0.00
Mean-variance constrained			
Weak	3.88	2.58	3.58
Moderate	3.82	2.58	3.58
Strong	2.98	2.34	3.58
Minimum-volatility constrained			
Weak	1.56	1.21	1.98
Moderate	1.67	1.21	1.98
Strong	1.54	1.29	1.98

This table reports the equivalent max expected Sharpe ratio of our multifactor portfolios ($\hat{\kappa}$), which we define as the $\hat{\kappa}$ that equalizes the Sharpe ratio of the portfolio obtained with the equation $\hat{w} = (\Sigma + \gamma I)^{-1}\,\mu$, with $\gamma = \kappa$, to the in-sample Sharpe ratio of our portfolios.

Table D.2: Annualized differences in Sharpe ratios relative to the Financials only benchmark (ESG sub-selections)

Set of base factors	Number of factors in the set	Multifactor portfolios									
		Naïve	Bayesian			Mean-variance constrained			Min-volatility constrained		
			weak	moderate	strong	weak	moderate	strong	weak	moderate	strong
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: In-sample differences in Sharpe ratios											
Financials + CSRHub	17	0.09	0.00	0.00	0.00	0.01	0.01	0.02	0.01	0.00	0.00
		(0.41)	(0.94)	(0.98)	(0.98)	(0.91)	(0.90)	(0.71)	(0.26)	(0.79)	(0.43)
Financials + non-aggregate ESG	201	**0.71**	0.04	-0.06	-0.16	**0.6**	**0.61**	**0.77**	0.5	0.46	**0.51**
		(0.08)	(0.73)	(0.67)	(0.37)	(0.02)	(0.01)	(<.01)	(0.11)	(0.11)	(0.08)
Financials + Traditional ESG	78	**0.37**	0.07	0.04	0.01	**0.31**	**0.32**	**0.44**	0.21	0.16	0.21
		(0.04)	(0.43)	(0.74)	(0.96)	(0.03)	(0.02)	(<.01)	(0.23)	(0.31)	(0.22)
Financials + Alternative ESG	144	**0.74**	-0.02	-0.1	-0.19	**0.45**	**0.46**	**0.60**	0.31	0.26	0.30
		(0.01)	(0.85)	(0.51)	(0.28)	(0.05)	(0.04)	(0.01)	(0.25)	(0.35)	(0.27)
Panel B: Out-of-sample differences in Sharpe ratios											
Financials + CSRHub	17	-0.15	0.02	0.06	0.08	0.01	0.01	0.02	0.01	0.00	0.01
		(0.29)	(0.79)	(0.42)	(0.40)	(0.87)	(0.86)	(0.77)	(0.24)	(0.93)	(0.27)
Financials + non-aggregate ESG	201	-0.13	0.08	0.09	0.10	-0.27	-0.26	-0.11	-0.17	-0.20	-0.28
		(0.64)	(0.68)	(0.66)	(0.57)	(0.5)	(0.5)	(0.77)	(0.65)	(0.59)	(0.41)
Financials + Traditional ESG	78	-0.26	0.09	0.13	0.17	-0.30	-0.29	-0.15	-0.09	-0.12	-0.20
		(0.45)	(0.31)	(0.26)	(0.22)	(0.40)	(0.38)	(0.64)	(0.76)	(0.68)	(0.46)
Financials + Alternative ESG	144	0.02	0.05	0.04	0.03	-0.05	-0.04	0.11	-0.09	-0.12	-0.20
		(0.91)	(0.75)	(0.85)	(0.87)	(0.82)	(0.83)	(0.58)	(0.72)	(0.64)	(0.43)

*This table reports differences (ΔSR) in annualized Sharpe ratios (in **bold** if p-value $\leq$ 0.1) between various multifactor portfolios and a benchmark, along with two-tailed p-values (in brackets) for testing ΔSR=0. The multifactor portfolios are constructed using three sets of factors and 10 construction approaches. The set of factors are Financials plus Consensus factors (from CSRHub scores), Financials plus Traditional ESG factors, and Financials plus Alternative ESG factors. As for the multifactor portfolios' construction approaches, we report results for the Naïve approach, three Bayesian approaches with different levels of shrinkage (weak, moderate, and strong), three constrained mean-variance approaches with effective number of factors (ENF) constraint (10 for weak, 30 for moderate, and 50 for strong), and three minimum-volatility approaches with ENF constraint (10 for weak, 30 for moderate, and 50 for strong). The benchmark portfolios use the Financials only set of factors. In Panel A (in-sample results), portfolio weights are estimated from the full-sample covariance matrix and average returns (August 2008 to June 2023). In Panel B (out-of-sample results), the full sample is divided into three non-overlapping five-year periods, and weights for each five-year sub-sample are estimated from the other two sub-samples. We compute p-values via the stationary block bootstrap of Ledoit and Wolf (2008) with a block length of 10 months.*

Table D.3: Annualized differences in Sharpe ratios – sparse Fama-French (2018) benchmark

Panel A: In-sample tests

Set of base factors	Multifactor portfolios									
	Naïve	Bayesian			Mean-variance constrained			Min-volatility constrained		
		weak	moderate	strong	weak	moderate	strong	weak	moderate	strong
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ESG only	**2.25**	0.42	-0.10	-0.24	**1.39**	**1.39**	**1.30**	**0.65**	**0.65**	**0.71**
	(<.01)	(0.13)	(0.71)	(0.38)	(<.01)	(<.01)	(<.01)	(0.09)	(0.08)	(0.06)
Financials + ESG	**2.93**	**0.50**	0.01	-0.11	**2.04**	**2.04**	**2.04**	**1.32**	**1.32**	**1.32**
	(<.01)	(0.05)	(0.97)	(0.66)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)	(<.01)
Benchmark SR:	0.9									

Panel B: Out-of-sample tests

Set of base factors	Multifactor portfolios									
	Naïve	Bayesian			Mean-variance constrained			Min-volatility constrained		
		weak	moderate	strong	weak	moderate	strong	weak	moderate	strong
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ESG only	-0.41	0.05	0.00	0.06	0.02	0.02	0.04	-0.30	-0.29	-0.29
	(0.18)	(0.88)	(0.99)	(0.85)	(0.97)	(0.97)	(0.95)	(0.52)	(0.53)	(0.55)
Financials + ESG	0.00	0.12	0.08	0.11	0.14	0.14	0.14	-0.13	-0.13	-0.13
	(0.99)	(0.61)	(0.76)	(0.65)	(0.75)	(0.74)	(0.75)	(0.76)	(0.76)	(0.76)
Benchmark SR:	0.62									

This table reports differences (ΔSR) in annualized Sharpe ratios (in ***bold*** *if p-value $\leq$ 0.1) between various multifactor portfolios and a benchmark, along with two-tailed p-values (in brackets) for testing ΔSR=0. The last row of each panel reports the benchmark's SR. The multifactor portfolios are constructed using two sets of factors and 10 construction approaches. The set of factors are ESG and Financials plus ESG. As for the multifactor portfolios' construction approaches, we report results for the Naïve approach, three Bayesian approaches with different levels of shrinkage (weak, moderate, and strong), three constrained mean-variance approaches with effective number of factors (ENF) constraint (10 for weak, 30 for moderate, and 50 for strong), and three minimum-volatility approaches with ENF constraint (10 for weak, 30 for moderate, and 50 for strong). The benchmark portfolios use only the 5 cross-sectional factors from Fama and French (2018) (SMB, HML, RMW, CMA, and MOM). In Panel A (in-sample results), portfolio weights are estimated from the full-sample covariance matrix and average returns (August 2008 to June 2023). In Panel B (out-of-sample results), the full sample is divided into three non-overlapping five-year periods, and weights for each five-year sub-sample are estimated from the other two sub-samples. We compute p-values via the stationary block bootstrap of Ledoit and Wolf (2008) with a block length of 10 months.*

Table D.4: Annualized differences in Sharpe ratios – Jensen, Kelly, and Pedersen (2023) factors as set of financials

Panel A: In-sample tests

Set of base factors	Multifactor portfolios									
	Naïve	Bayesian			Mean-variance constrained			Min-volatility constrained		
		weak	moderate	strong	weak	moderate	strong	weak	moderate	strong
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ESG only	0.14	**-0.50**	**-0.54**	**-0.63**	-0.23	-0.22	-0.18	-0.33	-0.35	-0.38
	(0.71)	(0.03)	(0.02)	(0.02)	(0.52)	(0.51)	(0.56)	(0.35)	(0.33)	(0.29)
Financials + ESG	**0.81**	-0.19	**-0.30**	**-0.41**	**0.66**	**0.66**	**0.79**	**0.52**	**0.51**	**0.41**
	(0.04)	(0.14)	(0.05)	(0.06)	(<.01)	(0.01)	(<.01)	(0.01)	(0.02)	(0.07)
Benchmark (Financials only) SR:	3.01	1.83	1.34	1.29	2.52	2.52	2.38	1.88	1.90	1.99

Panel B: Out-of-sample tests

Set of base factors	Multifactor portfolios									
	Naïve	Bayesian			Mean-variance constrained			Min-volatility constrained		
		weak	moderate	strong	weak	moderate	strong	weak	moderate	strong
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ESG only	-0.73	-0.41	-0.29	-0.26	-0.52	-0.52	-0.50	-0.75	**-0.74**	-0.65
	(0.20)	(0.23)	(0.30)	(0.30)	(0.29)	(0.29)	(0.30)	(0.10)	(0.09)	(0.15)
Financials + ESG	-0.20	-0.14	-0.09	-0.10	-0.20	-0.20	-0.20	-0.40	-0.39	-0.31
	(0.46)	(0.40)	(0.58)	(0.54)	(0.53)	(0.51)	(0.53)	(0.27)	(0.27)	(0.37)
Benchmark (Financials only) SR:	0.94	1.08	0.91	0.94	1.16	1.16	1.16	1.07	1.06	0.98

*This table reports annualized differences (ΔSR) in annualized Sharpe ratios (in **bold** if p-value $\leq$ 0.1) between various multifactor portfolios and a benchmark, along with two-tailed p-values (in brackets) for testing ΔSR=0. The last row of each panel reports the benchmark's SR. The multifactor portfolios are constructed using two sets of factors and 10 construction approaches. The set of factors are ESG and JKP-factors plus ESG. The JKP-factors is a set of 130 factors used in Jensen et al. (2023) downloaded from `https://jkpfactors.com/`. As for the multifactor portfolios' construction approaches, we report results for the Naïve approach, three Bayesian approaches with different levels of shrinkage (weak, moderate, and strong), three constrained mean-variance approaches with effective number of factors (ENF) constraint (10 for weak, 30 for moderate, and 50 for strong), and three minimum-volatility approaches with ENF constraint (10 for weak, 30 for moderate, and 50 for strong). The benchmark portfolios use the JKP-factors and the indicated construction approach. In Panel A (in-sample results), portfolio weights are estimated from the full-sample covariance matrix and average returns (August 2008 to June 2023). In Panel B (out-of-sample results), the full sample is divided into three non-overlapping five-year periods, and weights for each five-year sub-sample are estimated from the other two sub-samples. We compute p-values via the stationary block bootstrap of Ledoit and Wolf (2008) with a block length of 10 months.*