ESG’S NULL EFFECT ON THE QUALITY PREMIUM

ANALYSING ESG INTEGRATION INTO THE QUALITY-FACTOR ON THE S&P500 FROM 2003 TO 2024

NIKOLA PAJKIC

LORINC RÁTFAI

Master Thesis

Stockholm School of Economics

2025

ESG's null effect on the Quality premium : Analysing ESG integration into the Quality-factor on the S&P500 from 2003 to 2024

Abstract:

On a full sample of ESG-rated S&P 500 stocks from 2003 to 2024, ESG overlays add no statistically significant alpha improvement to the Quality-minus-Junk strategy. Furthermore, ESG overlays also fail to improve returns within the Quality leg. Results are confirmed with factor-controls up to FF5+1, as well as several robustness tests. These findings challenge earlier research that suggest an "icing on the cake" effect for high-quality stocks.

Keywords:

ESG, QMJ, portfolio construction, factor investing

Authors:

Nikola Pajkic (25369)
Lorinc Rátfai (42922)

Tutor:

Magnus Dahlquist, Handelsbanken Professor of Finance

Examiner:

Gualtiero Azzalini, Assistant Professor, Department of Finance

Acknowledgments:

We would like to express our gratitude to Magnus for his valuable guidance and feedback.

Master Thesis
Master Programme in Finance
Stockholm School of Economics

Contents

1. Introduction

We investigate how integrating Environmental, Social and Governance (ESG) scores into the long and short leg of the Quality-minus-Junk (QMJ) factor (Asness, Frazzini and Pedersen, 2019) change alphas for ESG-rated S&P500 stocks between 2003-2024. Specifically, we construct a strategy, sorting firms based on both their ESG and Quality metrics, that goes long on "Green Quality" and short on "Brown Junk" stocks. Our results, in **Table 1**, indicate that a value-weighted long-short portfolio formed along these lines does not deliver statistically significant alpha improvements. We also fail to establish a statistically significant difference in alphas within high- (and low-) quality legs after ESG overlay, contrary to the "icing-on-the-cake" effect observed by Lu, Su and Wu (2024) in the Taiwanese markets.

Table 1 (snippet):

	GMB VW - QMJ
CAPM	0.17%
	(0.991)
FF5+1	0.28%
	(1.618)

Refer to Table 1 for full description.

Our research is relevant for asset-allocators and sustainability-driven investors, since it explores how ESG-considerations can drive higher returns while also advancing sustainability targets.

Through our methodology, we create a unique factor combination by sequentially sorting, first on quality- and then on ESG-scores for S&P500 stocks, creating four buckets of stocks (high/low-ESG, high/low-Quality), and then going long Green Quality and short Brown Junk. In line with most recent research in asset-pricing, we employ Fama-French 5+1 controls for assessing results. More concretely, we use Market, Size, Growth, Profitability, Investment, and Momentum factors as controls (refer to Appendix A1 for precise definitions). Our analysis focuses on the S&P 500 because of data availability and liquidity considerations as this provides us with a stock horizon less affected by trading costs.

1.1. Literature Review

Research at the intersection of quality investing and ESG is scarce. Lu et al. (2024) find an "icing-on-the-cake" effect in Taiwanese data, where returns are higher for high-ESG companies only among high-quality companies. Otero González, Durán Santomil, Vieito and Reboredo (2025) find, among US stocks during 2014-2020, that ESG generally destroys risk-adjusted return in Quality portfolios, which can be offset by investing in companies with strong economic moat (sustainable competitive advantage). Melas, Nagy and Kulkarni (2016) find that ESG correlates higher with quality than most other factors, so investors targeting a quality factor return can improve ESG exposure without substantial quality exposure reduction.

Research on quality investing is extensive. Specifically, Asness et al. (2019) show that although high-quality companies (as defined by profitability, growth and safety characteristics) command a price premium, their value is still not fully reflected. As a result, investors going long high- and short low-quality can earn alpha and generate higher risk-adjusted returns. Other papers, e.g. Otero González et al. (2025), employ different definitions of quality (i.e. ROIC, F-score, Gross Margin) and find the same effect. We adopt the definition of quality from Asness et al. (2019).

With regards to ESG research, Friede, Busch and Bassen (2015) find in their meta-analysis of over 2,000 empirical studies that roughly 90% of papers support a nonnegative relationship between ESG and company financial performance, the majority of which support a positive relationship. However, the evidence is less clear on the relationship between ESG performance and stock returns. Glossner (2018) finds that ignoring ESG risk has a negative impact on performance. La Torre, Mango, Cafaro and Leo (2020) find that investors pay a premium for high-ESG stocks in European markets. Meanwhile, Huang, Lu and Wee (2020) find no effect of ESG on firm value. We advance the research by investigating if the effect of ESG varies when conditioned on additional variables (i.e. on quality).

We build mainly on the first two papers: we differ from Lu et al. (2024) in that we emphasize risk adjustment and use US data, and from Otero González et al. (2025) by examining the influence of ESG on quality more granularly.

2. Hypotheses

H1	An ESG-conditioned QMJ portfolio, going long Green Quality stocks and short Brown Junk ones, yields higher alphas than standard QMJ on the S&P500 between 2003 and 2024
H2	For S&P500 stocks between 2003 and 2024, there exists alpha between high-ESG and low-ESG, exclusively for high quality stocks

For Hypothesis 1, we test the performance of a new portfolio formation strategy between 2003 and 2024 against standard QMJ as in Asness et al. (2019), each month going long green high-quality stocks and short brown low-quality ones. Since it has already been established that higher quality firms command alpha, we want to see whether this effect can further be exacerbated by introducing ESG screens.

Hypothesis 2 builds on Lu et al. (2024), who find a so-called "icing-on-the-cake" effect in incorporating ESG scoring into quality score-based investing in Taiwan. Namely, they observe that ESG only drives a raw return premium for high-Quality stocks. They suggest that the effect is driven by investor attention: for ESG performance to be properly priced, it needs to be noticed by investors, and since attention is a scarce resource for investors, they prefer to allocate it to Quality over Junk stocks. In our analyses, "icing-on-the-cake" refers to an alpha and not to a raw return premium.

To test our hypotheses, we first group stocks into four buckets based on their ESG and Quality scores and measure relevant performance metrics for each. We test the "icing-on-the-cake" effect. Next, we form a long-short portfolio based on two buckets (Green Quality minus Brown Junk) and compute performance metrics. Next, we compare these metrics to the relevant QMJ benchmark. We perform all analyses for both equal-weighted and value-weighted constructions, as well as size-sorted and non-size sorted quality classifications, to check for consistency and robustness.

3. Empirical Work

3.1. Hypothesis 1 – Green Quality minus Brown Junk portfolio

Table 1: Alpha differences between GMB and QMJ

	No size sort		Size sort	
	VW-QMJ	EW-QMJ	VW-QMJ	EW-QMJ
CAPM	0.17%	-0.13%	0.10%	-0.02%
	(0.991)	(-0.680)	(0.709)	(-0.090)
FF3	0.18%	-0.13%	0.12%	0.00%
	(1.068)	(-0.695)	(0.807)	(-0.017)
FF3+1	0.21%	-0.13%	0.15%	0.03%
	(1.210)	(-0.657)	(1.025)	(0.159)
FF5	0.26%	-0.05%	0.11%	-0.02%
	(1.487)	(-0.257)	(0.774)	(-0.139)
FF5+1	0.28%	-0.04%	0.15%	0.01%
	(1.618)	(-0.214)	(1.001)	(0.051)

This table shows the alpha of a strategy that goes long a Green-minus-Brown (GMB) portfolio and short a Quality-minus-Junk (QMJ) portfolio, with respect to several factor model specifications. GMB involves a monthly 3x2 sort, first on quality and then on ESG score, and going long the high-quality high-ESG stocks and short the low-quality low-ESG stocks. The factor model specifications are as such: CAPM has only the market (MKT) return, "FF3" and "FF5" refers to the Fama and French (1992, 2015) models with added size (SMB), value (HML), profitability (RMW) and investment (CMA) factors, and "+1" to the addition of momentum (UMD). Results are reported with and without first sorting on size (by splitting on median market equity), and under value-weights and equal-weights for the GMB portfolio. This analysis is done on S&P500 stocks from 2003 to 2024 that have ESG scores in LSEG Workspace. Alphas are presented in monthly percentage terms. T-statistics are shown in parentheses below coefficient estimates. 10% significance of a t-stat is denoted by *, 5% by ** and 1% by ***.

Hypothesis 1 is concerning the improvements in risk-adjusted return that could be achieved through integrating ESG into the QMJ factor by going long Green Quality stocks and short Brown Junk ones. (A strategy hereinafter referred to as Green-minus-Brown or GMB). Our results, in **Table 1**, show that such a strategy does not create incremental alpha compared to the standard QMJ framework.

Although, in **Table 2**, we find higher, statistically significant alphas for our new value-weighted GMB strategy compared to QMJ, we reject Hypothesis 1 since differences in alphas are not significant. **Table 1** presents the differences in alphas between the two strategies, as well as the statistical significance of the differences. Results are computed by taking the return difference between our new strategy (GMB) and the sample-replicated, non-size sorted QMJ, regressing returns on different multifactor models and obtaining alphas. Our results are robust to size-sorted QMJ construction as well.

Individual performance **(Table 2)** metrics for value-weighted (VW) GMB show 0.63 CAPM alpha, statistically significant at the 5% level, implying significant abnormal returns after accounting for market risk, which is comparable to QMJ's 0.46 CAPM alpha, also significant at 5%. Across most (FF3, FF3+1, FF5) multi-factor models VW GMB maintains positive and statistically significant alphas and significance at the 10% or 5% levels, though it shows no significance with FF5+1 factor controls. QMJ shows 5% significance only with FF3 controls, with significance disappearing after further factor controls. Refer to Appendix A1 for performance metric definitions and factor model specifications, and Appendix A2 for graphs of GMB and QMJ performance.

We find that the VW GMB strategy achieves a slightly positive monthly mean return (0.01) and Sharpe ratio (0.43), outperforming the equal-weighted (EW) version (mean 0, Sharpe 0.06) and QMJ (mean 0, Sharpe 0.21), suggesting modest risk-adjusted returns for ESG-focused portfolios. Maximum drawdown (MDD) is smaller for the value-weighted GMB (-0.29) compared to the equal-weighted one (-0.48) or QMJ (-0.33), indicating slightly lower downside risk. Sortino ratios also show VW GMB's downside-risk-adjusted returns stronger (0.64 vs 0.07 for EW and 0.28 for QMJ).

Both VW GMB and QMJ strategies display consistently large and highly significant (1%) negative exposures to the value factor (HML). This reaffirms the strong "anti-value" tilt, typical of QMJ. Negative and significant SMB loadings indicate a tilt against small caps, although this is more pronounced for EW GMB and QMJ portfolios. All portfolios show strong and significant exposures to the momentum (UMD) and profitability (RMW) factors under more granular models, with high t-stats across FF5 and FF5+1 specifications. The robustness of these exposures for our GMB portfolio matches traditional QMJ, indicating that integrating ESG into quality selection does not erode the core return drivers of QMJ. The value-weighted GMB portfolio yields more consistent and statistically significant alphas than the equal-weighted one, suggesting that our strategy is more effective among larger stocks.

Table 2: Performance metrics of value- and equal-weighted Green-minus-Brown, and benchmark QMJ

	No size sort			With size sort		
	GMB VW L/S	GMB EW L/S	QMJ	GMB VW L/S	GMB EW L/S	QMJ
# observations	263	263	263	261	261	261
Mean	1%	0%	0%	0%	0%	0%
Median	1%	0%	1%	0%	0%	0%
Volatility	4%	4%	4%	4%	4%	3%
Min	-15%	-19%	-14%	-18%	-19%	-14%
Max	13%	20%	14%	19%	19%	11%
MDD	-29%	-48%	-41%	-36%	-43%	-26%
Sharpe ratio	0.43	0.06	0.27	0.12	-0.02	0.09
Sortino ratio	0.64	0.07	0.36	0.15	-0.03	0.13
CAPM alpha	0.63	0.33	0.48	0.36	0.24	0.31
	(2.54**)	(1.43)	(1.95*)	(1.7*)	(1.09)	(1.73*)
MKT	-0.16	-0.32	-0.19	-0.27	-0.31	-0.26
	(-2.87***)	(-6.09***)	(-3.48***)	(-5.84***)	(-6.29***)	(-6.61***)
FF3 alpha	0.49	0.17	0.32	0.24	0.12	0.19
	(2.33**)	(0.9)	(1.65*)	(1.35)	(0.63)	(1.3)
MKT	-0.04	-0.17	-0.06	-0.17	-0.2	-0.15
	(-0.83)	(-3.72***)	(-1.22)	(-3.87***)	(-4.43***)	(-4.2***)
SMB	-0.21	-0.35	-0.24	-0.22	-0.21	-0.29
	(-2.16**)	(-4.05***)	(-2.73***)	(-2.7***)	(-2.47**)	(-4.45***)
HML	-0.75	-0.71	-0.88	-0.61	-0.63	-0.53
	(-10.08***)	(-10.42***)	(-12.96***)	(-9.72***)	(-9.46***)	(-10.47***)
FF3+1 alpha	0.43	0.1	0.24	0.19	0.07	0.15
	(2.09**)	(0.53)	(1.3)	(1.11)	(0.4)	(1.07)
MKT	0	-0.12	0	-0.13	-0.17	-0.12
	(0.01)	(-2.48**)	(-0.01)	(-3.01***)	(-3.56***)	(-3.35***)
SMB	-0.15	-0.28	-0.16	-0.17	-0.16	-0.26
	(-1.53)	(-3.19***)	(-1.87*)	(-2.11**)	(-1.89*)	(-3.86***)
HML	-0.7	-0.64	-0.81	-0.57	-0.59	-0.5
	(-9.3***)	(-9.51***)	(-12.08***)	(-8.97***)	(-8.73***)	(-9.72***)
UMD	0.16	0.21	0.2	0.12	0.13	0.09
	(2.98***)	(4.41***)	(4.34***)	(2.76***)	(2.7***)	(2.65***)
FF5 alpha	0.36	0.05	0.14	0.11	-0.02	0.02
	(1.71*)	(0.27)	(0.73)	(0.64)	(-0.13)	(0.14)
MKT	-0.02	-0.15	-0.04	-0.13	-0.16	-0.12
	(-0.43)	(-3.12***)	(-0.78)	(-3.13***)	(-3.57***)	(-3.53***)
SMB	-0.08	-0.24	-0.04	-0.1	-0.09	-0.12
	(-0.72)	(-2.5**)	(-0.46)	(-1.18)	(-0.93)	(-1.68*)
HML	-0.83	-0.84	-0.94	-0.77	-0.84	-0.65
	(-8.25***)	(-9.1***)	(-10.55***)	(-9.14***)	(-9.43***)	(-9.84***)
RMW	0.34	0.28	0.52	0.27	0.29	0.45
	(2.82***)	(2.55**)	(4.92***)	(2.69***)	(2.72***)	(5.81***)
CMA	0.13	0.24	0.06	0.31	0.41	0.18
	(0.86)	(1.78*)	(0.43)	(2.52**)	(3.15***)	(1.84*)
FF5+1 alpha	0.31	-0.02	0.07	0.08	-0.06	-0.01
	(1.48)	(-0.09)	(0.36)	(0.43)	(-0.34)	(-0.1)
MKT	0.02	-0.1	0.02	-0.11	-0.13	-0.09
	(0.34)	(-2.02**)	(0.41)	(-2.41**)	(-2.87***)	(-2.75***)
SMB	-0.01	-0.16	0.04	-0.06	-0.04	-0.08
	(-0.14)	(-1.71*)	(0.44)	(-0.67)	(-0.44)	(-1.13)
HML	-0.75	-0.74	-0.84	-0.71	-0.78	-0.6
	(-7.37***)	(-8.04***)	(-9.47***)	(-8.32***)	(-8.61***)	(-8.97***)
RMW	0.35	0.29	0.53	0.28	0.29	0.46
	(2.95***)	(2.76***)	(5.25***)	(2.79***)	(2.82***)	(5.97***)
CMA	0.07	0.17	-0.02	0.27	0.37	0.14
	(0.47)	(1.26)	(-0.17)	(2.18**)	(2.83***)	(1.48)
UMD	0.16	0.2	0.21	0.11	0.11	0.1
	(3.05***)	(4.38***)	(4.76***)	(2.61***)	(2.48**)	(2.85***)

This table shows descriptive statistics, performance metrics and coefficients from regressions of excess return on various sets of factor portfolios, for a Green-minus-Brown portfolio and a baseline Quality-minus-Junk portfolio. GMB involves a monthly 3x2 sort, first on quality and then on ESG score, and going long the high-quality high-ESG stocks and short the low-quality low-ESG stocks. MDD refers to Maximum Drawdown, the biggest peak-to-trough decline in the cumulative return for the sample period. The factor model specifications are as such: CAPM has only the market (MKT) return, "FF3" and "FF5" refers to the Fama and French (1992, 2015) models with added size (SMB), value (HML), profitability (RMW) and investment (CMA) factors, and "+1" to the addition of momentum (UMD). Results are reported with and without first sorting on size (by splitting on median market equity), and under value-weights and equal-weights for the GMB portfolio. This analysis is done on S&P500 stocks from 2003 to 2024 that have ESG scores in LSEG Workspace. Mean, median, volatility, minimum, maximum returns and all alphas are reported in monthly percentage terms. Sharpe and Sortino ratios are annualized. T-stats are shown below coefficient estimates. 10% significance of a t-stat is denoted by *, 5% by ** and 1% by ***.

3.2. Hypothesis 2 – "Icing-on-the-cake"

For Hypothesis 2, we analyse, in **Table 3**, alpha differences between Green Quality (GQ) and Brown Quality (BQ), as well as Green Junk (GJ) and Brown Junk (BJ). The results indicate no statistically significant differences in alphas across various factor-control specifications, suggesting that an ESG overlay impacts neither the long nor short positions in QMJ. The absence of significant differences suggests a rejection of Hypothesis 2. Our results are robust to size-sorted QMJ construction.

Table 3: Alpha differences within QMJ legs ("icing on the cake")

	Without size sort				With size sort			
	Value-weighted		Equal-weighted		Value-weighted		Equal-weighted	
	GQ-BQ	GJ-BJ	GQ-BQ	GJ-BJ	GQ-BQ	GJ-BJ	GQ-BQ	GJ-BJ
CAPM	0.01%	0.14%	-0.07%	0.14%	0.12%	0.17%	0.01%	0.20%
	(0.051)	(0.633)	(-0.499)	(0.743)	(0.792)	(1.046)	(0.097)	(1.227)
FF3	-0.05%	0.20%	-0.09%	0.14%	0.11%	0.18%	0.02%	0.20%
	(-0.312)	(0.926)	(-0.733)	(0.746)	(0.767)	(1.129)	(0.171)	(1.200)
FF3+1	-0.05%	0.21%	-0.07%	0.15%	0.13%	0.21%	0.06%	0.21%
	(-0.309)	(0.993)	(-0.561)	(0.814)	(0.892)	(1.294)	(0.474)	(1.240)
FF5	-0.21%	0.30%	-0.14%	0.15%	0.00%	0.23%	-0.05%	0.22%
	(-1.233)	(1.368)	(-1.057)	(0.762)	(0.015)	(1.417)	(-0.352)	(1.314)
FF5+1	-0.21%	0.32%	-0.12%	0.16%	0.02%	0.26%	-0.01%	0.23%
	(-1.201)	(1.432)	(-0.886)	(0.829)	(0.152)	(1.603)	(-0.078)	(1.386)

This table shows the alpha of a strategy that goes long either a Green portfolio and short a Brown portfolio, either among Quality portfolios (GQ-BQ) or Junk portfolios (GJ-BJ), with respect to several factor model specifications. These portfolios arise from a monthly 3x2 sort, first on quality and then on ESG score. The factor model specifications are as such: CAPM has only the market (MKT) return, "FF3" and "FF5" refers to the Fama and French (1992, 2015) models with added size (SMB), value (HML), profitability (RMW) and investment (CMA) factors, and "+1" to the addition of momentum (UMD). Results are reported with and without first sorting on size (by splitting on median market equity), and under value-weights and equal-weights. This analysis is done on S&P500 stocks from 2003 to 2024 that have ESG scores in LSEG Workspace. Alphas are presented in monthly percentage terms. T-statistics are shown in parentheses below coefficient estimates. 10% significance of a t-stat is denoted by *, 5% by ** and 1% by ***.

In **Table 4** we report risk-adjusted return metrics and multi-factor alphas for each bucket across both value-weighted and equal-weighted scenarios to demonstrate the consistency and robustness of our results. Refer to Appendix A2 for graphs of bucket returns. Significant multi-factor alphas for GQ and BQ persist across CAPM, FF3, FF3+1, FF5 models (as well as FF5+1 for BQ) and both weighting schemes, with significant t-stats confirming that neither market, size, value, profitability, investment, or momentum exposures can fully explain the outperformance of these buckets. Results are robust to size-sorted QMJ construction (Appendix A3).

In every specification, Quality buckets (GQ, BQ) deliver outsized Sharpe, Treynor, and Sortino ratios compared to Junk buckets (GJ, BJ), highlighting superior risk-adjusted returns for high-quality groups regardless of ESG overlay. Moreover, high-quality buckets also exhibit higher multi-factor alphas, confirming the persistence of the Quality premium.

Multi-factor alphas, as well as Sharpe, Treynor, and Sortino ratios are higher for value-weighted Quality portfolios, suggesting larger firms contribute more to risk-adjusted returns in these strategies. Alphas and t-stats remain robust across both weighting schemes, underscoring the persistence and reliability of the Quality premium.

Quality portfolios (GQ and BQ) exhibit negative HML loadings, indicating growth bias, while Junk portfolios (GJ and BJ) show positive HML, reflecting a value tilt. Junk portfolios also consistently display negative UMD loadings at the 1% significance level, which may be caused by persistent divestment from Junk companies by investors.

Market betas remain close to 1 for all buckets, while CAPM and multi-factor t-statistics for market exposures are universally highly significant, confirming that portfolio returns are primarily equity-risk driven.

Table 4: Performance metrics for value- and equal-weighted (non-size-sorted) quality buckets with ESG-overlay

	Value-weighted				Equal-weighted			
	Green Quality	Green Junk	Brown Quality	Brown Junk	Green Quality	Green Junk	Brown Quality	Brown Junk
Sharpe ratio	0.86	0.42	0.82	0.32	0.76	0.52	0.79	0.44
Treynor ratio	0.15	0.07	0.15	0.06	0.12	0.09	0.13	0.08
Sortino ratio	1.3	0.5	1.17	0.43	1.05	0.72	1.1	0.62
CAPM alpha	0.35	-0.27	0.34	-0.41	0.18	-0.14	0.25	-0.28
	(2.64***)	(-1.37)	(2.2**)	(-2.22**)	(1.91*)	(-0.73)	(2.12**)	(-1.39)
MKT	0.93	1.18	0.96	1.09	1.02	1.34	1.02	1.34
	(31.58***)	(26.87***)	(27.64***)	(26.6***)	(48.47***)	(30.74***)	(39.22***)	(29.46***)
FF3 alpha	0.25	-0.17	0.31	-0.36	0.16	-0.01	0.25	-0.15
	(2.28**)	(-1.01)	(2.08**)	(-2.13**)	(1.67*)	(-0.03)	(2.23**)	(-0.87)
MKT	1.03	1.1	0.97	1.08	1.04	1.22	1	1.22
	(38.32***)	(27.71***)	(27.12***)	(25.87***)	(45.98***)	(32.33***)	(36.55***)	(29.98***)
SMB	-0.32	0.08	0.06	-0.11	-0.08	0.27	0.17	0.28
	(-6.27***)	(1.04)	(0.91)	(-1.34)	(-1.92*)	(3.84***)	(3.39***)	(3.65***)
HML	-0.33	0.64	-0.31	0.42	-0.08	0.65	-0.17	0.63
	(-8.3***)	(10.95***)	(-5.84***)	(6.96***)	(-2.44**)	(11.66***)	(-4.16***)	(10.59***)
FF3+1 alpha	0.24	-0.11	0.29	-0.32	0.17	0.1	0.24	-0.06
	(2.14**)	(-0.67)	(1.97**)	(-1.9*)	(1.82*)	(0.69)	(2.15**)	(-0.37)
MKT	1.04	1.06	0.98	1.04	1.03	1.14	1	1.15
	(37.31***)	(26.26***)	(26.38***)	(24.41***)	(43.86***)	(32.43***)	(35.32***)	(29.11***)
SMB	-0.3	0.02	0.08	-0.15	-0.1	0.17	0.18	0.19
	(-5.85***)	(0.24)	(1.12)	(-1.88*)	(-2.22**)	(2.54**)	(3.48***)	(2.54**)
HML	-0.31	0.59	-0.29	0.39	-0.09	0.55	-0.16	0.55
	(-7.78***)	(10.08***)	(-5.44***)	(6.26***)	(-2.76***)	(10.87***)	(-3.87***)	(9.65***)
UMD	0.04	-0.16	0.04	-0.12	-0.04	-0.28	0.02	-0.25
	(1.47)	(-3.91***)	(1.11)	(-2.69***)	(-1.63)	(-8.03***)	(0.82)	(-6.22***)
FF5 alpha	0.19	-0.01	0.4	-0.31	0.1	0.06	0.24	-0.08
	(1.69*)	(-0.03)	(2.72***)	(-1.74*)	(1.07)	(0.41)	(2.08**)	(-0.49)
MKT	1.04	1.08	0.94	1.06	1.06	1.22	1	1.21
	(38.13***)	(27.27***)	(26.24***)	(24.9***)	(46.07***)	(31.46***)	(35.56***)	(29.03***)
SMB	-0.24	-0.08	-0.01	-0.16	-0.04	0.2	0.18	0.21
	(-4.41***)	(-1.02)	(-0.07)	(-1.84*)	(-0.77)	(2.51**)	(3.21***)	(2.52**)
HML	-0.33	0.76	-0.11	0.49	-0.17	0.67	-0.18	0.67
	(-6.25***)	(9.89***)	(-1.58)	(5.92***)	(-3.67***)	(8.9***)	(-3.26***)	(8.22***)
RMW	0.2	-0.41	-0.13	-0.13	0.1	-0.2	0.02	-0.17
	(3.21***)	(-4.45***)	(-1.55)	(-1.32)	(1.95*)	(-2.28**)	(0.34)	(-1.77*)
CMA	-0.02	-0.21	-0.41	-0.13	0.17	-0.02	0.02	-0.06
	(-0.26)	(-1.84*)	(-4.0***)	(-1.09)	(2.55**)	(-0.22)	(0.3)	(-0.48)
FF5+1 alpha	0.17	0.05	0.38	-0.27	0.12	0.16	0.23	0
	(1.55)	(0.3)	(2.59***)	(-1.54)	(1.24)	(1.15)	(2.01**)	(0.01)
MKT	1.05	1.04	0.95	1.03	1.05	1.14	1	1.14
	(37.47***)	(26.16***)	(25.89***)	(23.73***)	(44.35***)	(32.16***)	(34.6***)	(28.58***)
SMB	-0.23	-0.14	0.02	-0.2	-0.05	0.08	0.19	0.12
	(-4.02***)	(-1.8*)	(0.24)	(-2.32**)	(-1.13)	(1.18)	(3.3***)	(1.45)
HML	-0.31	0.69	-0.08	0.44	-0.19	0.53	-0.17	0.55
	(-5.66***)	(8.86***)	(-1.13)	(5.16***)	(-4.06***)	(7.64***)	(-2.95***)	(6.98***)
RMW	0.21	-0.42	-0.12	-0.14	0.1	-0.22	0.02	-0.19
	(3.27***)	(-4.69***)	(-1.5)	(-1.4)	(1.9*)	(-2.79***)	(0.36)	(-2.07**)
CMA	-0.04	-0.15	-0.43	-0.09	0.18	0.08	0.02	0.03
	(-0.47)	(-1.35)	(-4.19***)	(-0.75)	(2.79***)	(0.83)	(0.19)	(0.3)
UMD	0.05	-0.16	0.06	-0.11	-0.05	-0.29	0.02	-0.25
	(1.68*)	(-4.04***)	(1.64)	(-2.62***)	(-1.94*)	(-8.29***)	(0.8)	(-6.32***)

This table shows performance metrics and coefficients from regressions of excess return on various sets of factor portfolios, for the Green and Brown parts of Quality and Junk portfolios.These portfolios arise from a monthly 3x2 sort, first on quality and then on ESG score. The factor model specifications are as such: CAPM has only the market (MKT) return, "FF3" and "FF5" refers to the Fama and French (1992, 2015) models with added size (SMB), value (HML), profitability (RMW) and investment (CMA) factors, and "+1" to the addition of momentum (UMD). Treynor ratios are computed using CAPM betas. Results are reported under value-weights and equal-weights. This analysis is done on S&P500 stocks from 2003 to 2024 that have ESG scores in LSEG Workspace. Alphas are presented in monthly percentage terms. Sharpe, Treynor and Sortino ratios are annualized. T-statistics are shown in parentheses below coefficient estimates. 10% significance of a t-stat is denoted by *, 5% by ** and 1% by ***.

3.3. Construction methodology

In our analyses, each month we construct both value-weighted and equal-weighted portfolios. Stocks are sequentially sorted into four buckets: first into Quality and Junk (top and bottom 30 percentiles of QMJ scores, without size-sorting), and then each of those buckets is split into Green and Brown along their respective median ESG scores. This construction isolates the marginal effect of ESG on quality attributes. We sort first on quality and then on ESG scores in line with our initial aim of examining the incremental effect of ESG on existing Quality portfolios as opposed to devising new ways to form portfolios (which would argue for a simultaneous sort) or looking at the effect of quality on ESG portfolios (which would argue for sorting first on ESG).

Figure 1: Portfolio construction methodology

		Quality Score		
		>70%	70-30%	<30%
ESG Score	>50%	Green Quality		Green Junk
	<50%	Brown Quality		Brown Junk

This figure explains our portfolio construction methodology: first splitting along quality scores at the 30th and 70th percentile (to obtain Junk and Quality), and in turn each of those portfolios at the median into Green and Brown.

4. Setup

4.1. Data Sources

Our stock universe encompasses stocks 2003-2024 on the S&P500 that received an ESG score in LSEG Workspace, and is free of survivorship bias.

We work from five data sources: CRSP; the Global Stock Returns and Characteristics database from Jensen, Kelly and Pedersen (2023); LSEG Workspace; Quality Minus Junk: Factors, Monthly from the AQR Data Library (AQR, 2025); and Fama/French 5 Factors (2x3) from the Kenneth R. French data library (French, 2025).

We identify stocks (by their CRSP PERMNOs), current and "dead", that had been part of the S&P500 index between 2002 (the first year for which LSEG ESG scores are available) and December 2024 based on the CRSP S&P 500 List. Next, we obtain monthly quality scores (refer to Appendix A1 for definitions and Appendix A2 for implications), as well as relevant accounting data (i.e. (Dollar) Return per Share, Market Equity, Total Assets, Return on Equity, and Book Leverage) for our analysis from the dataset provided by Jensen et al. (2023), which shifts all relevant variables by 4 months to account for standard reporting delay.[1]

Our dataset includes the overall ESG score as well as the three pillar scores: Environmental, Social, and Governance. We obtain annual ESG scores from LSEG Workspace by first matching PERMNO identifiers for each stock with their corresponding CUSIP codes and linking historical ESG data specifically for our selected S&P 500 components. In line with the methodology of Jensen et al. (2023) for quality scores, we apply a four-month lag to all ESG scores to align timing with subsequent financial performance to ensure consistency in how we incorporate ESG factors into our analyses. Next, we create a panel dataset of stock specific characteristics (see Appendix A3), matching characteristics through time at the end of each month through interlinking

[1] There are 35 – predominantly REIT – stocks which were at some point part of the S&P500, however, the Jensen et al. (2023) database does not include.

PERMNOs and CUSIPs associated with stocks at different points in time, and forward filling (annual) ESG scores for 11 months.

For our factor controls, we use U.S. market, SMB, HML, and UMD returns from the AQR database, and RMW and CMA returns from the Kenneth French library. For more details on the construction of these factors, refer to Appendix A1.

4.2. ESG discussion

Our data source from LSEG Workspace employs a data-driven approach to generate ESG scores, using materiality matrices and weighting schemes. The data goes back to 2002, however, our analysis begins in 2003 since only a few companies were rated before that (Appendix A2). We opt for LSEG data due to its availability and comprehensive coverage. Instead of ESG scores with controversy ratings, we use unadjusted ESG scores to reduce subjectivity.

ESG measures are used to evaluate companies' standing with respect to their environmental and social impact, as well as governance structure based on reported data. The measure incorporates many different metrics such as emissions, resource use, workforce diversity, and governance structure.

Using ESG measures for investments may help investors with finding sustainable and resilient businesses, and it may also align their portfolios with mandates and/or moral values. Atz, Whelan, Holt, Casey Clark, Salazar, Liu, and Bruno (2021) observe that studies conducted between 2015-2020 find a positive relationship between ESG and financial performance in 58% of cases. For the relationship between ESG and risk-adjusted returns, they find this number to be 59%.

Moreover, firms with higher ESG ratings can access cheaper financing through their lower cost of capital and more favourable risk profile (Feldhütter and Pedersen 2024), although, this conclusion remains contested, some claiming that it only affects firms with certain (i.e. high growth and/or high WACC) characteristics (Arouri, Mrabet, Temimi and Smaoui, 2025). Furthermore, investors may demand better reporting related to ESG as a result of regulatory pressure and stakeholder expectations (Bernow, Godsall, Klempner and Merten, 2019).

While incorporating ESG measures into investment strategies may have many benefits, there are certain limitations that need to be considered. One of the most important limitations of ESG measures is the divergence in ratings. As documented by Berg, Kölbel and Rigobon (2022), correlation between ESG ratings across different agencies is between 0.38 and 0.71. Secondly, ESG ratings may be limited due to incomplete data as many firms have differences in what they report, making comparability difficult. Furthermore, the weightings employed by each ESG data provider may differ considerably resulting in varying scores.

These elements introduce a degree of subjectivity, which LSEG seeks to mitigate through regular updates to its methodology. The scores measure relative performance by benchmarking companies within their industry groups for environmental and social factors (and by country of incorporation for governance). To account for the varying significance of ESG factors across industries, the scoring system incorporates materiality adjustments via a proprietary magnitude matrix. This matrix dynamically modifies weights based on disclosure levels and the industry impact of each ESG theme.

In practice, LSEG's ESG scoring involves several steps. Initially, it converts raw data points into standardized percentile ranks. Next, it calculates scores for ten thematic areas, which are grouped into three main categories: environmental, social, and governance. These category scores are weighted according to industry-specific assessments derived from materiality analyses.

4.3. ESG Persistence

Our analysis starts with establishing the persistence of ESG and ESG Pillar Scores since predictability is a prerequisite for a forward-looking, rational market to be able to price ESG in stocks. We confirm that both the overall ESG Score, and its individual pillars demonstrate persistence over time, meaning a significant difference remains between the highest and lowest ESG-rated stocks.

Table 5 presents the persistence of ESG scores. We observe that while the difference in ESG (Overall and Pillar) scores shrinks between the highest and lowest deciles, they stay far enough apart to reject the null hypothesis of no difference 10 years into the future using Newey-West t-statistics, which accounts for heterogeneity.

A look deeper shows that stocks corresponding to the lowest deciles tend to improve their scoring the most, average overall scores more than doubling (from ~20 to ~50), suggesting a positive correlation between scoring and future improvements. This, however, does not result in the disappearance of differences between the highest and lowest deciles: the t-stat for the difference in means remains highly significant beyond 1%, its N-W t-stat only dropping to 15.51 in 10 years from 24.84. This observation holds for all ESG Pillar scores as well as the overall measure.

For our methodology, we divide stocks from our dataset (S&P 500 2003-2024) into ten fixed deciles based on their annual ESG (Pillar) Scores at the given month, ensuring a statistically significant difference exists between the top and bottom deciles. Next, we track how these ratings evolve from one year up to ten years. Since Asness et al. (2019) already established the persistence of Quality Scores, we limit our analysis purely to the ESG dimension.

Table 5: Persistence of ESG measures

		P1 (Low)	P2	P3	P4	P5	P6	P7	P8	P9	P10 (High)	H-L	*H-L t-stat*
ESG	t	20.02	30.10	36.14	41.10	45.74	50.47	55.40	60.87	67.40	77.91	57.88	*24.84*
	t + 12 M	24.20	33.32	38.72	43.01	46.95	51.75	55.89	61.01	67.30	76.26	52.02	*19.60*
	t + 36 M	29.66	38.08	42.80	46.57	50.18	53.58	58.00	62.03	67.52	76.15	46.17	*17.87*
	t + 60 M	35.72	43.28	46.78	50.43	52.46	56.00	60.21	64.40	67.87	76.07	39.66	*15.60*
	t + 120 M	47.55	50.70	53.20	53.26	58.68	60.54	62.70	66.47	68.87	75.58	27.84	*15.51*
Env.	t + 120 M	42.76	41.79	41.29	44.68	49.34	55.84	60.41	62.10	67.80	78.03	35.27	*13.59*
Soc.	t + 120 M	48.45	52.25	51.91	55.61	59.84	63.64	65.28	66.40	71.56	80.34	31.88	*13.37*
Gov.	t + 120 M	48.72	53.89	60.52	61.60	60.85	63.04	62.80	63.78	67.99	71.74	22.48	*11.07*

The graph shows the persistence of the ESG Scores and its three pillars. Each year, stocks are sorted into 10 deciles based on their ESG (Pillar) Scores and then the average of each decile is calculated. Next, the average Scores for each decile's components going forward by 12, 36, 60, and 120 months is tracked. The difference between the Highest and Lowest deciles is reported on the right of each row alongside the t-statistic which tests the hypothesis that the difference is 0. Source: Refinitiv Workspace

4.4. Establishing QMJ in stock universe

We construct two Quality portfolios to test the robustness of the QMJ factor in our sample. The first portfolio does not apply a size control, while the second follows Asness et al. (2019) and controls for firm size, by first sorting into small and big companies based on median market cap breakpoint. This design allows us to assess robustness and whether size adjustment is necessary within the S&P 500 universe, which already consists of large-cap stocks.

As shown in **Table 6**, both methods result in similar factor loadings, as well as risk-adjusted return metrics, although the non-size-sorted QMJ shows minor, 10%, significance for CAPM and 3F alphas. To verify these findings, in **Table 7**, we examine the monthly differences in alphas between the two portfolios and find non-size-sorted QMJ outperforming under CAPM with 10% significance, however, this result disappears under further factor-controls. This suggests that size effects in a sample containing large stocks do not offset the exclusion of high- or low-quality stocks from either portfolio leg. Additionally, the overall performance of the QMJ strategies appear similar to that of the Profitability pillar, indicating the factor's high reliance on it.

Table 6: Alpha differences between non-size-sorted and size-sorted QMJ

	QMJ - QMJ size sorted
CAPM	0.22%
	(1.676*)
FF3	0.20%
	(1.483)
FF3+1	0.20%
	(1.486)
FF5	0.14%
	(1.059)
FF5+1	0.14%
	(1.044)

This table shows the alpha of a strategy that goes long a non-size-sorted Quality-minus-Junk portfolio and shorts a size-sorted Quality-minus-Junk portfolio, with respect to several factor model specifications. These portfolios arise from a monthly sort, either immediately into Quality, neutral and Junk, or first into Small and Big (based on median market equity breakpoints), and going long Quality and short Junk. The factor model specifications are as such: CAPM has only the market (MKT) return, "FF3" and "FF5" refers to the Fama and French (1992, 2015) models with added size (SMB), value (HML), profitability (RMW) and investment (CMA) factors, and "+1" to the addition of momentum (UMD). This analysis is done on S&P500 stocks from 2003 to 2024 that have ESG scores in LSEG Workspace. Alphas are presented in monthly percentage terms. T-statistics are shown in parentheses below coefficient estimates. 10% significance of a t-stat is denoted by *, 5% by ** and 1% by ***.

In our analysis, each month we assign the top and bottom 30 percentiles of stocks to High- and Low-Quality buckets. Next, with monthly rebalancing, we calculated performance metrics using model specifications up to Fama-French's 3-factor model plus UMD. We employ a conventional 30/40/30 portfolio split to avoid stocks that fluctuate around the median, while ensuring that the long/short portfolios' legs contain enough stocks for reliable statistical tests.

Table 7: Performance metrics of non-size-sorted and size-sorted QMJ

	Panel A: without size sort				Panel B: with size sort			
	QMJ	Prof	Growth	Safety	QMJ	Prof	Growth	Safety
Excess return	0.32	0.36	0.06	-0.04	0.06	0.22	0.14	-0.15
	(1.29)	(1.47)	(0.28)	(-0.21)	(0.29)	(1.07)	(0.77)	(-0.77)
CAPM alpha	0.48	0.46	0.1	0.25	0.26	0.3	0.23	0.17
	(1.95*)	(1.85*)	(0.49)	(1.26)	(1.28)	(1.48)	(1.29)	(1.02)
FF3 alpha	0.32	0.29	0.01	0.14	0.12	0.18	0.15	0.07
	(1.65*)	(1.62)	(0.05)	(0.77)	(0.78)	(1.23)	(0.91)	(0.5)
FF3+1 alpha	0.24	0.24	-0.04	0.05	0.04	0.14	0.09	0
	(1.3)	(1.35)	(-0.24)	(0.27)	(0.31)	(0.97)	(0.55)	(-0.03)
MKT	0	0.05	0.07	-0.17	-0.06	0.01	0.01	-0.23
	(-0.01)	(1.03)	(1.41)	(-4.09***)	(-1.71*)	(0.17)	(0.12)	(-6.76***)
SMB	-0.16	-0.09	-0.1	-0.18	-0.12	0.04	-0.08	-0.12
	(-1.87*)	(-1.07)	(-1.14)	(-2.33**)	(-1.75*)	(0.58)	(-1.05)	(-1.92*)
HML	-0.81	-0.91	-0.4	-0.41	-0.67	-0.78	-0.34	-0.39
	(-12.08***)	(-13.93***)	(-5.89***)	(-6.72***)	(-12.63***)	(-14.6***)	(-5.75***)	(-7.83***)
UMD	0.2	0.15	0.15	0.25	0.21	0.11	0.17	0.21
	(4.34***)	(3.27***)	(3.23***)	(5.97***)	(5.68***)	(2.91***)	(4.19***)	(6.2***)
Sharpe	0.27	0.31	0.06	-0.04	0.06	0.23	0.16	-0.17
Info ratio	0.29	0.3	-0.05	0.06	0.07	0.22	0.12	-0.01
Adjusted R2	0.47	0.49	0.19	0.44	0.54	0.5	0.23	0.53

This table shows performance metrics and coefficients from regressions of excess return on various sets of factor portfolios, for Quality-minus-Junk portfolios. These portfolios arise from either a monthly sort into Quality, neutral and Junk (panel A) or a monthly 2x3 sort, into Small and Big (based on median market equity) and then into Quality and Junk. The factor model specifications are as such: CAPM has only the market (MKT) return, "FF3" refers to the Fama and French (1992) model with added size (SMB) and value (HML) factors, and "+1" to the addition of momentum (UMD). Information ratios are computed using FF3+1 alphas. Treynor ratios are calculated using CAPM betas. Panel A reports results for quality portfolios formed simply on a monthly sort into Quality, neutral and Junk (and going long Quality, short Junk), while panel B first sorts into Small and Big (based on median market equity) and then into quality the same way as panel A.This analysis is done on S&P500 stocks from 2003 to 2024 that have ESG scores in LSEG Workspace. Returns are in excess of the US Treasury bill rate. Alphas and excess returns are presented in monthly percentage terms. Sharpe, Treynor and Sortino ratios are annualized. T-statistics are shown in parentheses below coefficient estimates. 10% significance of a t-stat is denoted by *, 5% by ** and 1% by ***.

4.5. "Icing on the cake" for individual stocks

Lu et al. (2024) verify the existence of "icing on the cake" by regressing individual returns on ESG scores, a dummy for Quality, and the interaction of the two. They show that ESG scores have no significant influence on returns *per se*, rather, solely through the Quality interaction term (hence, "icing-on-the-cake"). We redo their analysis with some modifications but fail to find (**Table 8**) a positively significant coefficient on the interaction term between ESG score and Quality (in other words, we fail to find "icing-on-the-cake"). Instead, we find some weak evidence of an analogous effect among Junk stocks, suggesting that there may nonetheless be value in a strategy that goes long Green Quality and short Brown Junk.

Neither the coefficient on ESG score nor the interaction between ESG score and a Quality dummy are significant in any model specification, which goes against the "icing-on-the-cake" result found by Lu et al. (2024). Notably, the positive interaction between ESG and Junk is robust up to Model 3, which suggests that Green Junk may do better than Brown Junk.

Repeating the main analysis using the individual quality and ESG pillars, we again find that ESG alone and the ESGxQuality interaction are never significant, adding to the robustness of our finding that there is no "icing-on-the-cake". The ESGxJunk interaction itself is positively significant across all models for the Social pillar and only significant (10%) in Model 1 for Growth and Environmental pillars, which speaks against the robustness of the “anti-icing-on-the-cake" effect identified in the main regression.

The Quality effect itself materialises in most of our results: the Junk dummy is almost never significant, and the Quality dummy is positively significant in Models 2 and 3 for most regressions. Strangely, the opposite effect appears in Model 4 for Profitability and Safety: a significantly negative Quality dummy coefficient and a significantly positive Junk dummy coefficient seem to speak in favour of Junk-minus-Quality instead. This, however, arises due to our use of firm fixed effects which are estimated over the whole sample and render the results redundant from an investing standpoint.

Table 8: "Icing on the cake" regressions on individual stock returns

	Main regression			
	(1)	(2)	(3)	(4)
Intercept	0.00899	0.00008	0.0006	0.01559
	(16.51083***)	(0.15215)	(1.1884)	(5.41301***)
ESG	-0.00051	0.00002	0.00002	0.00003
	(-0.88255)	(0.03363)	(0.03605)	(0.03692)
Q	0.00121	0.00187	0.00187	-0.00095
	(1.35504)	(2.17078**)	(2.16386**)	(-0.73633)
ESG*Q	0.00039	0.00016	0.00016	-0.00044
	(0.44026)	(0.20617)	(0.20188)	(-0.52433)
J	0.00026	-0.00071	-0.00071	0.00225
	(0.24746)	(-0.74212)	(-0.75048)	(1.45396)
ESG*J	0.00271	0.00188	0.00187	0.00137
	(2.19305**)	(1.79045*)	(1.80083*)	(1.39252)
MKT		1.09505	1.03979	
		(38.89771***)	(49.92945***)	
Log(assets)		-0.0009	-0.0009	-0.01215
		(-1.921*)	(-1.92916*)	(-4.34804***)
Lev		0.00144	0.00144	0.00255
		(2.85796***)	(2.92323***)	(2.49577**)
ROE		-0.00138	-0.00138	-0.00255
		(-2.27029**)	(-2.31154**)	(-2.45781**)
SMB			0.08382	
			(2.45241**)	
HML			0.13221	
			(3.04063***)	
RMW			0.04594	
			(1.28396)	
CMA			0.04481	
			(1.14788)	
UMD			-0.07918	
			(-4.46077***)	
# obs	47497	47497	47497	47497
Adj. R2	0.00009	0.27464	0.27943	0.28533

This table shows the results of a pooled OLS regression of individual firm excess returns on ESG score, a Quality dummy variable (equal to 1 if Quality score is above 70th percentile), the interaction between the two, and equivalently for a Junk dummy (1 if below 30th percentile) and its interaction with ESG score, under various model specifications. Model (1) is the regression as just described, Model (2) controls for market excess return, log assets, leverage, and ROE, Model (3) adds the size, value, profitability, investment and momentum factor portfolios, and Model (4) is Model (2) with firm and time fixed effects. The regressions are repeated substituting overall Quality score for its pillars (Profitability, Growth, Safety), and similarly for ESG score (Environmental, Social, Governance). All non-market variables are lagged by one month, winsorised at the 1st and 99th percentile, and standardised to have a mean of zero and a standard deviation of one. Standard errors are clustered on firm and month. T-statistics are shown in parentheses below coefficient estimates. 10% significance of a t-stat is denoted by *, 5% by ** and 1% by ***.

Table 8 (continued)

	Profitability				Growth				Safety			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Intercept	0.00811	-0.00087	-0.00035	0.01594	0.00911	0.00013	0.00066	0.01526	0.00923	0.00041	0.00093	0.01491
	(14.95909***)	(-1.50014)	(-0.62876)	(4.82541***)	(17.22663***)	(0.25038)	(1.3015)	(5.33554***)	(16.37961***)	(0.73782)	(1.74329*)	(5.40177***)
ESG	0.00021	0.0005	0.00057	0.00015	-0.00044	0.00007	0.00006	-0.0001	0.00035	0.00024	0.00029	0.00032
	(0.33804)	(0.87334)	(1.00562)	(0.1862)	(-0.77997)	(0.13983)	(0.11258)	(-0.12316)	(0.62399)	(0.47023)	(0.56641)	(0.41211)
Q	0.00238	0.00305	0.00305	-0.00125	0.00112	0.00141	0.00142	-0.00067	-0.0002	-0.00007	-0.00006	-0.00341
	(2.78704***)	(3.23737***)	(3.22502***)	(-0.70922)	(1.17446)	(1.59204)	(1.60092)	(-0.60081)	(-0.24838)	(-0.08958)	(-0.08484)	(-2.80573***)
ESG*Q	0.00007	0.00015	0.00006	0.00033	0.0008	0.00047	0.00054	0	-0.0004	0.00036	0.00024	0.00001
	(0.07754)	(0.18873)	(0.07566)	(0.31995)	(0.82823)	(0.58419)	(0.67316)	(0.00315)	(-0.51088)	(0.51441)	(0.33109)	(0.00936)
J	0.00192	0.00118	0.00118	0.0042	-0.00008	-0.00047	-0.00047	-0.00031	0.00085	0.00019	0.0002	0.00349
	(2.04195**)	(1.18682)	(1.19175)	(2.07236**)	(-0.07856)	(-0.54618)	(-0.55278)	(-0.22979)	(0.8639)	(0.22038)	(0.23193)	(2.09228**)
ESG*J	0.0005	0.00012	-0.00002	0.00036	0.00191	0.0015	0.00148	0.00128	0.00057	0.00099	0.00095	-0.00001
	(0.42784)	(0.12035)	(-0.02267)	(0.29577)	(1.7983*)	(1.60679)	(1.59702)	(1.41741)	(0.56548)	(1.09183)	(1.05287)	(-0.01065)
MKT		1.09514	1.03984			1.0951	1.03983			1.09518	1.03986	
		(38.94334***)	(50.00647***)			(38.91253***)	(49.96373***)			(38.93276***)	(49.99593***)	
Log(assets)		-0.00093	-0.00093	-0.01219		-0.00106	-0.00106	-0.01176		-0.00106	-0.00105	-0.0119
		(-1.93165*)	(-1.92996*)	(-4.43709***)		(-2.30827**)	(-2.32485**)	(-4.31829***)		(-2.22649**)	(-2.2363**)	(-4.46386***)
Lev		0.00137	0.00137	0.00255		0.00127	0.00127	0.00282		0.0012	0.0012	0.00209
		(2.63226***)	(2.694***)	(2.50381**)		(2.61443***)	(2.66327***)	(2.78661***)		(2.42712**)	(2.4629**)	(2.07691**)
ROE		-0.0013	-0.0013	-0.00246		-0.00118	-0.00118	-0.00301		-0.00091	-0.00091	-0.0028
		(-2.09669**)	(-2.13448**)	(-2.37875**)		(-2.00599**)	(-2.03473**)	(-2.91292***)		(-1.64901*)	(-1.67173*)	(-2.76042***)
SMB			0.0838				0.08385				0.08374	
			(2.45524**)				(2.45406**)				(2.44996**)	
HML			0.13206				0.13236				0.13215	
			(3.03447***)				(3.04484***)				(3.03551***)	
RMW			0.04595				0.04597				0.04573	
			(1.28671)				(1.28542)				(1.2798)	
CMA			0.04479				0.04456				0.04468	
			(1.14561)				(1.14063)				(1.14216)	
UMD			-0.07932				-0.07911				-0.0793	
			(-4.47037***)				(-4.45343***)				(-4.46941***)	
# obs	47497	47497	47497	47497	47497	47497	47497	47497	47497	47497	47497	47497
Adj. R2	0.00004	0.27463	0.27942	0.28536	0.00001	0.27457	0.27936	0.28523	-0.00006	0.27449	0.27927	0.28548

Table 8 (continued)

	Environmental				Social				Governance			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Intercept	0.00901	0.00008	0.0006	0.01596	0.00899	0.00007	0.0006	0.01577	0.00896	0.00008	0.0006	0.01694
	(16.48402***)	(0.16343)	(1.19171)	(5.14708***)	(16.51754***)	(0.1435)	(1.18325)	(5.45756***)	(16.51976***)	(0.14874)	(1.18773)	(5.60856***)
ESG	-0.0007	-0.00024	-0.00019	-0.00097	-0.00073	-0.0003	-0.00031	-0.00081	-0.00001	0.00024	0.00024	0.00098
	(-1.35353)	(-0.53013)	(-0.43606)	(-1.38153)	(-1.35496)	(-0.61351)	(-0.63128)	(-1.20781)	(-0.01322)	(0.47077)	(0.46947)	(1.42146)
Q	0.00116	0.00185	0.00185	-0.00095	0.00122	0.0019	0.00189	-0.0009	0.00132	0.00201	0.00201	-0.00092
	(1.304)	(2.16307**)	(2.15984**)	(-0.73752)	(1.37205)	(2.20522**)	(2.19812**)	(-0.69314)	(1.48127)	(2.34325**)	(2.33549**)	(-0.71318)
ESG*Q	0.0004	0.00003	-0.00004	-0.0006	-0.00008	-0.00038	-0.00037	-0.00061	0.00102	0.00117	0.00119	0.00061
	(0.45809)	(0.03621)	(-0.04891)	(-0.71366)	(-0.09095)	(-0.47687)	(-0.45909)	(-0.75173)	(1.14679)	(1.52431)	(1.52611)	(0.71898)
J	0.00017	-0.00077	-0.00077	0.0022	0.00031	-0.00066	-0.00066	0.00232	0.00012	-0.00088	-0.00088	0.00222
	(0.1664)	(-0.7986)	(-0.80421)	(1.42028)	(0.29888)	(-0.6968)	(-0.70394)	(1.50102)	(0.10829)	(-0.89062)	(-0.89726)	(1.43179)
ESG*J	0.00205	0.00141	0.00133	0.00073	0.0026	0.00187	0.0019	0.0015	0.00141	0.00083	0.00081	0.00059
	(1.76076*)	(1.47606)	(1.41068)	(0.76736)	(2.30758**)	(1.92972*)	(1.97919**)	(1.64554*)	(1.18137)	(0.8002)	(0.7897)	(0.67037)
MKT		1.09511	1.03982			1.09508	1.0398			1.09513	1.03982	
		(38.90405***)	(49.93796***)			(38.90159***)	(49.9293***)			(38.90016***)	(49.94657***)	
Log(assets)		-0.00079	-0.00079	-0.01184		-0.00074	-0.00074	-0.01202		-0.00089	-0.00089	-0.01234
		(-1.67131*)	(-1.67894*)	(-4.26073***)		(-1.59516)	(-1.60039)	(-4.29745***)		(-1.98264**)	(-1.99214**)	(-4.4042***)
Lev		0.00141	0.00141	0.00261		0.00138	0.00138	0.00256		0.00152	0.00152	0.00259
		(2.78534***)	(2.85365***)	(2.57911***)		(2.72473***)	(2.78527***)	(2.52145**)		(3.0403***)	(3.11587***)	(2.54173**)
ROE		-0.00134	-0.00134	-0.00258		-0.00133	-0.00133	-0.00255		-0.0014	-0.0014	-0.00261
		(-2.20471**)	(-2.24477**)	(-2.49307**)		(-2.18336**)	(-2.22309**)	(-2.45678**)		(-2.32369**)	(-2.36871**)	(-2.52966**)
SMB			0.08391				0.08387				0.084	
			(2.45443**)				(2.45414**)				(2.45739**)	
HML			0.13206				0.13216				0.13179	
			(3.03759***)				(3.04002***)				(3.03003***)	
RMW			0.04608				0.04592				0.04606	
			(1.28695)				(1.28383)				(1.28815)	
CMA			0.04469				0.04494				0.04517	
			(1.14428)				(1.15047)				(1.15795)	
UMD			-0.07925				-0.0792				-0.07935	
			(-4.46235***)				(-4.46302***)				(-4.47223***)	
# obs	47497	47497	47497	47497	47497	47497	47497	47497	47497	47497	47497	47497
Adj. R2	0.00001	0.27457	0.27936	0.28533	0.00009	0.27463	0.27942	0.28535	0.00003	0.27463	0.27942	0.28539

We regress monthly excess returns of individual companies in our sample on ESG scores, dummies for Quality and Junk, and their interactions as explanatory variables, under four different model specifications for robustness: no controls (Model 1), Lu et al. (2024) style controls (Model 2), Model 2 with added FF5+1 controls (Model 3), and Model 2 with firm and time fixed effects (Model 4). Next, we repeat the analysis for the different ESG and Quality pillars to get a deeper understanding on what drives the relation between ESG, Quality, and excess returns. All regressions use two-way clustered standard errors to account for correlation in residuals across firms and months.

$$R_{i,t} - r_f = a + b_1 \times ESG_{i,t-1} + b_2 \times Q_{i,t-1} + b_3 \times ESG_{i,t-1} \times Q_{i,t-1}$$
$$+ b_4 \times J_{i,t-1} + b_5 \times ESG_{i,t-1} \times J_{i,t-1} + controls + \varepsilon_{i,t}$$

$Q_{i,t}$ corresponds to a dummy for quality, being equal to 1 if the firm is above the 70th percentile of quality scores and 0 otherwise, and $J_{i,t}$ is a dummy for junk, equal to 1 if the firm is below 30th percentile quality and 0 otherwise. We differ from Lu et al. (2024) by including a dummy variable for Junk as well as its interaction with the ESG score, because the counterfactual to Quality is not only Junk but also the middle 40 percentiles: a Junk dummy allows us to also examine the interaction of ESG and junk.

In line with Lu et al. (2024), Models 2 and 3 control for Market Return, Log Assets, Return on Equity (ROE), and Leverage, with the latter three sourced from Jensen et al. (2023), and Market Return from AQR's dataset (Appendix A1). ROE is calculated as Net Income divided by Book Equity. Leverage is measured as Book Assets divided by Book Equity. Model 3 additionally controls for size (SMB), value (HML), profitability (RMW), investment (CMA) and momentum (UMD), refer to Appendix A1.

Our justifications for these controls are as follows: Market Return controls for market movements and mitigates the risk of attributing general market effects to ESG or quality; firms with higher ROE typically demonstrate stronger financial health and may allocate more resources to ESG activity and have higher Profitability pillar scores; lower leverage firms generally have a healthier capital structure which enables them greater flexibility for ESG investments and have higher Safety pillar scores; Log Assets, reflecting firm size, are included since larger firms tend to have more resources for ESG improvements and achieve greater transparency. These controls enhance the robustness of our results and ensure that our analysis captures the distinct contributions of ESG and Quality scores.

Non-market controls are winsorised at the 99^{th} and 1^{st} percentiles to limit the influence of outliers in later standardisation. Next, all non-market variables are lagged by one month, and standardised (except quality dummies) to have mean 0 and standard deviation 1. This makes interpretation of coefficients uniform and more coherent.

5. Conclusion

In summary, we find that applying an ESG-overlay to the baseline QMJ factor, going long Green Quality and short Brown Junk, does not deliver statistically significant alpha improvements for S&P 500 stocks between 2003-2024. We also find no evidence of alpha improvements through ESG-overlay, “icing-on-the-cake”, within the Quality leg of QMJ. We confirm our results with factor-controls up to FF5+1, as well as several robustness tests. Our findings challenge earlier research by Lu et al. (2024) suggesting an “icing-on-the-cake” effect for high-quality stocks.

5.1. Limitations

There are certain limitations to our thesis. Most importantly, our portfolios are constrained to companies which are ESG rated which introduces data limitations for two reasons. Firstly, ESG scores started being reported around 2000 which limits the analysable horizon. Secondly, many companies are not ESG-rated, potentially making our conclusions biased towards rated firms. Notably, rated firms have higher returns than non-rated ones in our sample (Appendix A2). Another, obvious, limitation is that our analysis is constrained to S&P500 data, thus rendering the universality of our findings uncertain.

5.2. Further Research

There remains considerable scope for further inquiry into this topic. Notably, the positive interaction we observe between ESG scores and Junk in determining returns warrants further study. Furthermore, given the limited size and scope of our data and research, future research could examine whether results hold under alternative quality definitions (e.g. ROIC or Piotroski’s F-score as in Otero González et al. (2025)), as well as ESG scores from different providers. Future studies could also extend the analysis to other markets.

6. References

AQR Capital Management. Quality Minus Junk: Factors, Monthly. (2025). Aqr.com. https://www.aqr.com/Insights/Datasets/Quality-Minus-Junk-Factors-Monthly

Arouri, H., Mrabet, Z., Temimi, A. & Smaoui, H. (2025). The Differential Effects of ESG Performance on Cost of Capital: Growth Opportunities and Non-Linear Impacts. Preprint. http://dx.doi.org/10.2139/ssrn.5141021

Asness, C. S., Frazzini, A., & Pedersen, L. H. (2019). Quality minus junk. Review of Accounting Studies, 24(1), 34–112. https://doi.org/10.1007/s11142-018-9470-2

Atz, U., Whelan, T., Holt, T.V., Casey Clark, C., Salazar, P., Liu, Z., & Bruno, C. (2021). ESG AND FINANCIAL PERFORMANCE: Uncovering the Relationship by Aggregating Evidence from 1,000 Plus Studies Published between 2015 – 2020.

Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate Confusion: The Divergence of ESG Ratings. European Finance Review, 26(6), 1315–1344. https://doi.org/10.1093/rof/rfac033

Bernow, S., Godsall, J., Klempner, B., & Merten, C. (2019, August 7). More than values: The value-based sustainability reporting that investors want. Www.mckinsey.com. https://www.mckinsey.com/capabilities/sustainability/our-insights/more-than-values-the-value-based-sustainability-reporting-that-investors-want

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. The Journal of Finance (New York), 47(2), 427. https://doi.org/10.2307/2329112

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

Feldhütter, Peter, and Lasse Heje Pedersen. "Is Capital Structure Irrelevant with ESG Investors?" The Review of financial studies 38.8 (2025): 2362–2385. Web.

French, K. R. (2025). Kenneth R. French - Data Library. Dartmouth.edu. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. Journal of Sustainable Finance & Investment, 5(4), 210–233. https://doi.org/10.1080/20430795.2015.1118917

Glossner, S. (2018). The Price of Ignoring ESG Risks.

Huang, P., Lu, Y., & Wee, M. (2020). Corporate governance analysts and firm value: Australian evidence. Pacific-Basin Finance Journal, 63, Article 101430. https://doi.org/10.1016/j.pacfin.2020.101430

JENSEN, T. I., KELLY, B., & PEDERSEN, L. H. (2023). Is There a Replication Crisis in Finance? The Journal of Finance (New York), 78(5), 2465–2518. https://doi.org/10.1111/jofi.13249

La Torre, M., Mango, F., Cafaro, A., & Leo, S. (2020). Does the ESG Index Affect Stock Return? Evidence from the Eurostoxx50. Sustainability, 12(16), 6387. https://doi.org/10.3390/su12166387

Lu, C.-W., Wu, H.-L., & Su, Y.-H. (2024). The icing on the cake: ESG effect on the quality factor portfolios. Finance Research Letters, 70, Article 106304. https://doi.org/10.1016/j.frl.2024.106304

Melas, D., Nagy, Z. & Kulkarni, P. (2016). Factor Investing and ESG Integration. MSCI Research Insights, November 2016. https://www.msci.com/research-and-insights/paper/factor-investing-and-esg-integration

OpenAI. (2025). ChatGPT. Retrieved from ChatGPT website: https://chatgpt.com/Perplexity. (2025).

Otero-González, L., Durán Santomil, P., Vieito, J. P., & Reboredo, J. C. (2025). How to improve quality investing. Business Research Quarterly, 28(2), 453–473. https://doi.org/10.1177/23409444231202810

Perplexity AI. Retrieved from www.perplexity.ai website: https://www.perplexity.ai/

7. AI Declaration

ChatGPT and PerplexityAI have been used to assist in coding. This has improved the thesis by cutting down time spent on technicalities like for example reshaping panels, merging datasets with different datetime formats, or formatting outputs. A risk with doing so is that we would not understand the way the code works later, or that the code generated provides an error-free but otherwise incorrect result. We mitigate this by trying to pose the problem in a more abstract fashion and then manually adapt the result to work with our use case, and always double check the results before moving on. As part of this process, we have gotten better at coding ourselves and our reliance on ChatGPT and PerplexityAI for such help has gone down over time.

8. APPENDIX

Appendix A1. Definitions

Definitions of performance metrics

For a given stock's monthly return r_i, the annualized performance metrics are given by:

$$Sharpe_i = \frac{\overline{r_i - r_f}}{\sigma_i} \times \sqrt{12}$$

$$Treynor_i = \frac{\overline{r_i - r_f}}{\beta_i} \times 12$$

$$Sortino_i = \frac{\overline{r_i - r_f}}{\sigma_{downside,i}} \times \sqrt{12}$$

$$Information_i = \frac{\propto_i}{\sigma_i} \times \sqrt{12}$$

$\sigma_{downside,i}$ is the downside standard deviation, that is, the standard deviation only of negative monthly returns. Wherever a metric calls for the use of alpha or beta, we use the one estimated from a CAPM regression of monthly excess return on only the market excess return. The risk-free rate used is the monthly Treasury bill rate.

Factor model specifications

CAPM: $r_{i,t} - r_{f,t} = \propto_{CAPM} + \beta_{CAPM} \times MKT_t$

FF3: $r_{i,t} - r_{f,t} = \propto_{FF3} + \beta_{FF3,MKT} \times MKT_t + \beta_{FF3,SMB} \times SMB_t + \beta_{FF3,HML} \times HML_t$

FF3+1: $r_{i,t} - r_{f,t} = \propto_{FF3+1} + \beta_{FF3+1,MKT} \times MKT_t + \beta_{FF3+1,SMB} \times SMB_t + \beta_{FF3+1,HML} \times HML_t + \beta_{FF3+1,UMD} \times UMD_t$

FF5: $r_{i,t} - r_{f,t} = \propto_{FF5} + \beta_{FF5,MKT} \times MKT_t + \beta_{FF5,SMB} \times SMB_t + \beta_{FF5,HML} \times HML_t + \beta_{FF5,RMW} \times RMW_t + \beta_{FF\ ,CMA} \times CMA_t$

FF5+1: $r_{i,t} - r_{f,t} = \propto_{FF5+1} + \beta_{FF5+1,MKT} \times MKT_t + \beta_{FF5+1,SMB} \times SMB_t + \beta_{FF5+1,HML} \times HML_t + \beta_{FF5+1,RMW} \times RMW_t + \beta_{FF5+1,CMA} \times CMA_t + \beta_{FF5+1,UMD} \times UMD_t$

Factor portfolios

We obtain factor portfolio returns for the market (MKT), small-minus-big (SMB), high-minus-low (HML), and momentum (UMD, up-minus-down) from the Quality Minus Junk: Factors, Monthly database from the AQR Data Library (AQR, 2025).

MKT is the monthly value-weighted market return in excess of US Treasury bills, where the market is all available US common stocks in the merged CRSP/XpressFeed database.

SMB is constructed using a 2x3 conditional sort on size (total market equity, with a NYSE median breakpoint) and value (book equity / market equity), and going long the three small portfolios and short the three big portfolios:

$$SMB = \left(\frac{1}{3}\right) \times (Small\ Value + Small\ Neutral + Small\ Growth) \\ - \left(\frac{1}{3}\right) \times (Big\ Value + Big\ Neutral + Big\ Growth)$$

HML is constructed using the same 2x3 conditional sort as for SMB however the portfolio is long the two value portfolios and short the two growth portfolios:

$$HML = \left(\frac{1}{2}\right) \times (Small\ Value + Big\ Value) - \left(\frac{1}{2}\right) \times \ (Small\ Growth + Big\ Growth)$$

UMD is constructed using a 2x3 (is it?) sort on size and the past 12 months (excluding the most recent month) of returns, long the high-return portfolios and short the low-return portfolios:

$$UMD = \left(\frac{1}{2}\right) \times (Small\ High + Big\ High) - \left(\frac{1}{2}\right) \times (Small\ Low + Big\ Low)$$

Size is defined as total market equity, and the breakpoint is the median NYSE market equity. Value is defined as book equity divided by market equity. Size and value breakpoints are refreshed in June of each calendar year.

The remaining two factor portfolio returns, robust-minus-weak (RMW) and conservative-minus-aggressive (CMA) are from "Fama/French 5 Factors (2x3)" in the Kenneth R. French data library (French, 2025).

RMW is constructed using a 2x3 simultaneous sort on size and operating profitability, going long the two robust portfolios and short the two weak portfolios:

$$RMW = \left(\frac{1}{2}\right) \times (Small\ Robust + Big\ Robust) - \left(\frac{1}{2}\right) \times (Small\ Weak + Big\ Weak)$$

Where OP is defined as:

$$OP = \frac{Revenue - COGS - SG\&A - Int.Exp.}{Book\ equity}$$

CMA is constructed using a 2x3 simultaneous sort on size and investment (% change in total assets from year t-2 to year t-1), going long the two conservative portfolios and short the two aggressive portfolios:

$$CMA = \left(\frac{1}{2}\right) \times (Small\ Conservative + Big\ Conservative)$$
$$+ \left(\frac{1}{2}\right) \times (Small\ Aggressive + Big\ Aggressive)$$

Quality scores

Jensen et al. (2023) calculates quality-scores following Asness et al. (2019). A firm's Quality score is constructed using:

1. Profitability score
    a. Gross Profit over Assets
    b. Return on Equity
    c. Return on Assets
    d. Cash-flow over Assets
    e. Gross Margin
    f. Low Accruals
2. Growth score
    a. Five-year growth in Gross Profit over Assets
    b. Five-year growth in Return on Equity
    c. Five-year growth in Return on Assets
    d. Five-year growth in Cash-flow over Assets
    e. Five-year growth in Gross Margin
3. Safety score
    a. Negative Market Beta
    b. Low Leverage
    c. Low Bankruptcy Risk:
        i. Ohlson's O-score
        ii. Altman's Z-score
    d. Low Earnings Volatility

Individual variables are ranked cross-sectionally in ascending order, then z-scored/standardized (by subtracting the cross-sectional mean and dividing by cross-sectional standard deviation) to have a mean of zero and standard deviation of one in the cross-section. To aggregate to the pillar level, add up the variable z-scores and z-score the sum. To aggregate to an overall quality score, do the same with the pillars.

Appendix A2. Graphs

Level of quality on the S&P 500

In both the original Asness et al. (2019) formulation, and in Jensen, Kelly and Pedersen's (2023) dataset, quality scores are standardized to have a mean of 0 and standard deviation of 1 in the cross-section. One interesting byproduct of our using scores computed by Jensen et al. (2023), but not their entire universe of companies, is that this is not guaranteed to apply. In the below figures we report these statistics for each month, for overall, rated and unrated S&P.

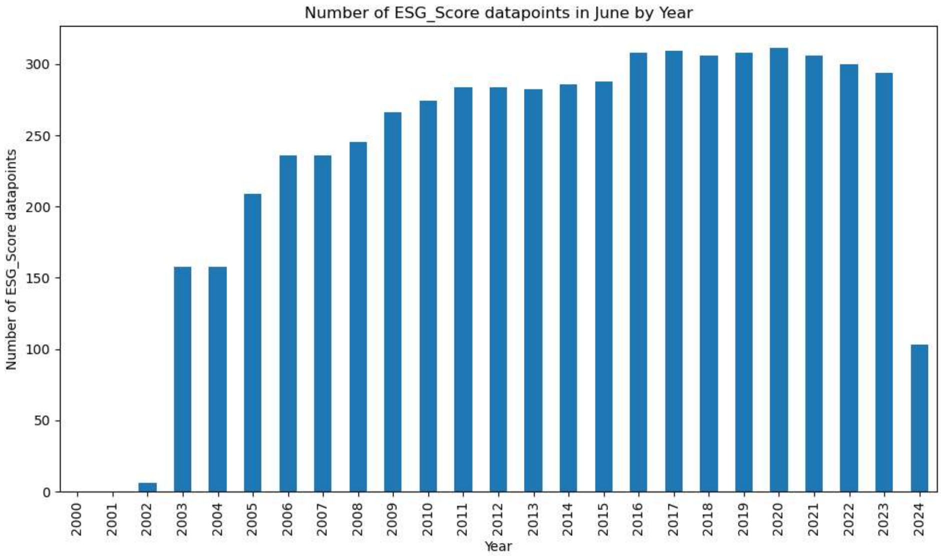
Figure 2: Cross-sectional means and standard deviations of QMJ scores

The graphs show the cross-sectional means and standard deviations of quality scores in our sample, for all S&P500 stocks, and for those that have and do not have ESG Scores in LSEG Workspace. These all differ from zero and one respectively because we obtain scores from the Global Stock Returns and Characteristics database from Jensen et al. (2023), which computes them on a bigger universe.

We nonetheless form our portfolios by comparing quality scores to 30th and 70th percentile cutoffs within our sample, meaning that we still have a relative quality effect. It is the same relative quality effect as if the scores had been initially computed on the correct universe. But we gain the advantage of being able to observe the quality of our sample relative to the Jensen et al. (2023) broader universe. Namely, we can state that the stocks on the S&P 500 seem to be of relatively higher quality, about the same growth, and far higher profitability.

Availability of ESG scores

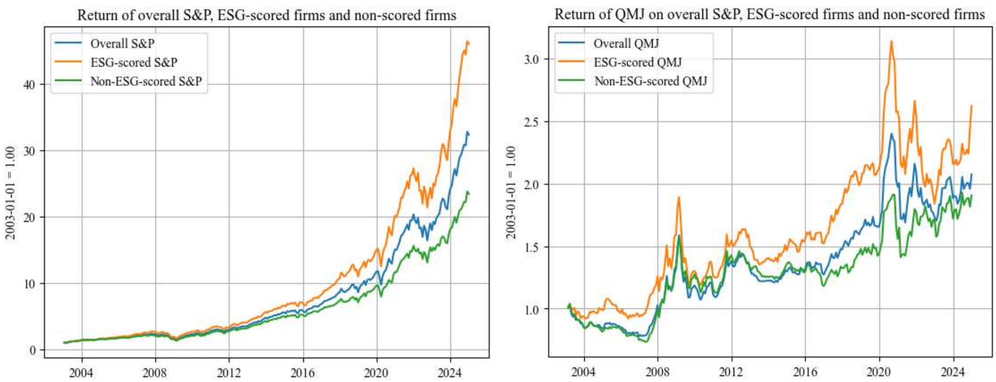
Figure 3: ESG scores per year



This figure shows the total number of LSEG Workspace ESG-scored S&P500 companies we have in our dataset for each year.

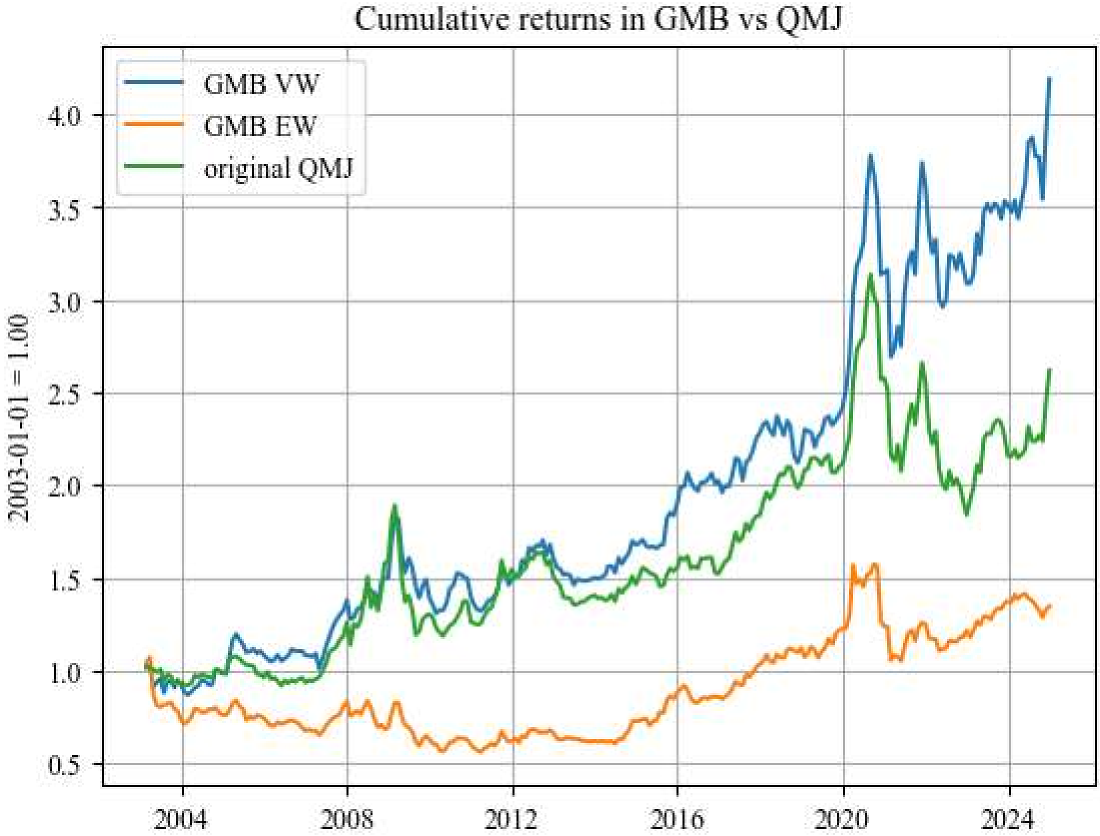
Cumulative returns

Figure 4: Cumulative return of S&P and QMJ in overall S&P, ESG-scored and non-ESG-scored firms



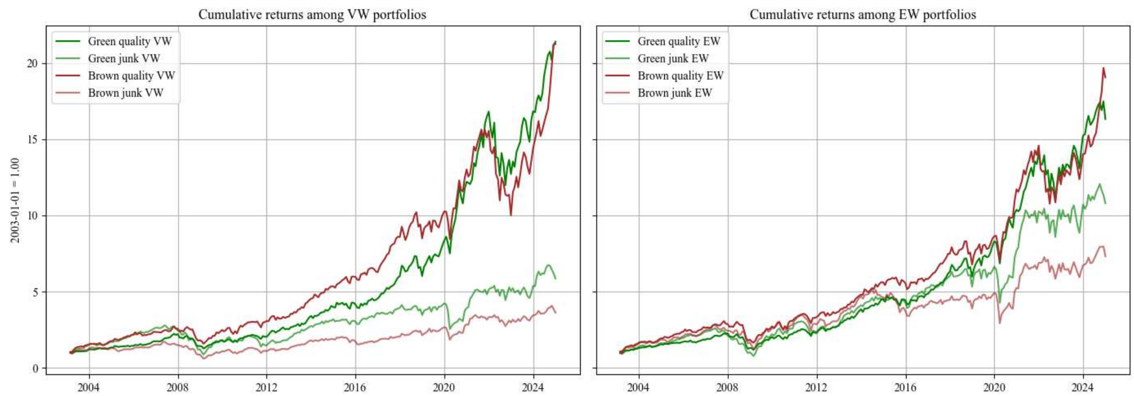
The graphs show the cumulative return of the overall S&P500 and the non-size-sorted Quality-minus-Junk benchmark strategy, both overall in our sample and for stocks that do and do not have ESG scores in LSEG Workspace.

Figure 5: Cumulative return of Green-minus-Brown versus Quality-minus-Junk



This graph shows the cumulative return of a non-size-sorted Quality-minus-Junk benchmark in comparison to our new strategy which goes long Green Quality and short Brown Junk stocks, in both a value- and equal-weighted case.

Figure 6: Cumulative returns in Green-minus-Brown legs



These graphs show the cumulative returns of the portfolios that we obtain from a sort, first on quality scores, then on ESG scores, both for the value- and equal-weighted case.

Appendix A3. Tables

Table 4 (b): Performance metrics for value- and equal-weighted (size-sorted) quality buckets with ESG-overlay

	Value-weighted				Equal-weighted			
	Green Quality	Green Junk	Brown Quality	Brown Junk	Green Quality	Green Junk	Brown Quality	Brown Junk
Sharpe ratio	0.76	0.51	0.61	0.39	0.71	0.57	0.66	0.44
Treynor ratio	0.12	0.08	0.11	0.07	0.11	0.09	0.12	0.08
Sortino ratio	1.02	0.7	0.84	0.55	0.97	0.8	0.92	0.63
CAPM alpha	0.13	-0.19	0.01	-0.36	0.08	-0.08	0.07	-0.29
	(2.04**)	(-1.2)	(0.1)	(-1.91*)	(1.03)	(-0.5)	(0.55)	(-1.36)
MKT	0.96	1.27	1.01	1.23	1.05	1.3	1.04	1.35
	(68.18***)	(36.9***)	(32.97***)	(30.02***)	(58.25***)	(35.93***)	(36.39***)	(29.06***)
FF3 alpha	0.12	-0.07	0.01	-0.25	0.1	0.05	0.08	-0.14
	(1.83*)	(-0.59)	(0.04)	(-1.61)	(1.29)	(0.47)	(0.64)	(-0.86)
MKT	0.98	1.16	1	1.14	1.02	1.16	1.01	1.22
	(63.92***)	(41.26***)	(30.97***)	(30.21***)	(52.55***)	(41.24***)	(33.01***)	(30.02***)
SMB	-0.04	0.29	0.13	0.18	0.1	0.38	0.16	0.32
	(-1.4)	(5.53***)	(2.07**)	(2.58***)	(2.7***)	(7.13***)	(2.78***)	(4.13***)
HML	-0.06	0.53	-0.21	0.55	0.03	0.56	-0.11	0.66
	(-2.6***)	(12.9***)	(-4.43***)	(9.96***)	(1.2)	(13.49***)	(-2.41**)	(11.13***)
FF3+1 alpha	0.13	0.01	0	-0.2	0.14	0.14	0.07	-0.07
	(2.05**)	(0.15)	(0)	(-1.28)	(1.74*)	(1.37)	(0.58)	(-0.42)
MKT	0.97	1.09	1.01	1.1	1	1.1	1.02	1.16
	(61.28***)	(43.09***)	(29.83***)	(28.77***)	(51.13***)	(43.2***)	(31.86***)	(28.88***)
SMB	-0.05	0.2	0.13	0.12	0.07	0.29	0.17	0.23
	(-1.86*)	(4.33***)	(2.11**)	(1.75*)	(1.8*)	(6.14***)	(2.85***)	(3.15***)
HML	-0.07	0.45	-0.21	0.5	0.01	0.48	-0.1	0.59
	(-3.06***)	(12.41***)	(-4.21***)	(9.08***)	(0.19)	(13.09***)	(-2.2**)	(10.21***)
UMD	-0.04	-0.23	0.02	-0.16	-0.09	-0.24	0.02	-0.22
	(-2.29**)	(-9.2***)	(0.44)	(-4.14***)	(-4.55***)	(-9.34***)	(0.65)	(-5.35***)
FF5 alpha	0.03	0.02	0.03	-0.21	0.01	0.12	0.06	-0.1
	(0.54)	(0.16)	(0.22)	(-1.33)	(0.12)	(1.03)	(0.43)	(-0.57)
MKT	0.99	1.15	0.99	1.13	1.04	1.16	1.02	1.2
	(68.33***)	(41.07***)	(29.84***)	(29.08***)	(55.29***)	(40.86***)	(32.4***)	(28.9***)
SMB	0.04	0.19	0.11	0.16	0.19	0.3	0.18	0.29
	(1.49)	(3.36***)	(1.63)	(1.98**)	(4.98***)	(5.14***)	(2.78***)	(3.37***)
HML	-0.13	0.55	-0.15	0.63	-0.06	0.55	-0.16	0.77
	(-4.6***)	(10.06***)	(-2.31**)	(8.34***)	(-1.66*)	(9.93***)	(-2.59***)	(9.44***)
RMW	0.21	-0.27	-0.03	-0.05	0.23	-0.23	0.04	-0.06
	(6.28***)	(-4.18***)	(-0.34)	(-0.58)	(5.18***)	(-3.43***)	(0.49)	(-0.57)
CMA	0.13	0	-0.13	-0.17	0.17	0.05	0.11	-0.22
	(3.03***)	(-0.01)	(-1.34)	(-1.52)	(3.21***)	(0.62)	(1.18)	(-1.85*)
FF5+1 alpha	0.04	0.1	0.02	-0.16	0.04	0.2	0.05	-0.03
	(0.76)	(1.01)	(0.17)	(-1.04)	(0.55)	(2.05**)	(0.39)	(-0.18)
MKT	0.98	1.09	1	1.09	1.02	1.1	1.03	1.15
	(66.26***)	(44.36***)	(29.01***)	(27.97***)	(54.84***)	(44.24***)	(31.45***)	(28.06***)
SMB	0.03	0.1	0.12	0.1	0.15	0.2	0.19	0.2
	(0.97)	(1.97**)	(1.71*)	(1.24)	(4.14***)	(4.03***)	(2.83***)	(2.48**)
HML	-0.15	0.43	-0.14	0.56	-0.11	0.43	-0.15	0.67
	(-5.17***)	(9.02***)	(-2.08**)	(7.34***)	(-2.95***)	(8.88***)	(-2.38**)	(8.32***)
RMW	0.21	-0.29	-0.02	-0.06	0.22	-0.24	0.04	-0.07
	(6.28***)	(-5.21***)	(-0.32)	(-0.71)	(5.29***)	(-4.34***)	(0.5)	(-0.75)
CMA	0.14	0.09	-0.14	-0.11	0.21	0.14	0.1	-0.14
	(3.39***)	(1.29)	(-1.41)	(-1.04)	(4.0***)	(2.04**)	(1.11)	(-1.26)
UMD	-0.04	-0.24	0.02	-0.15	-0.09	-0.25	0.02	-0.21
	(-2.64***)	(-10.01***)	(0.62)	(-3.99***)	(-5.13***)	(-10.12***)	(0.52)	(-5.16***)

This table shows performance metrics and coefficients from regressions of excess return on various sets of factor portfolios, for the Green and Brown parts of Quality and Junk portfolios, but initially sorted on size based on a median market equity breakpoint. These portfolios arise from a monthly 3x2 sort, first on quality and then on ESG score. The factor model specifications are as such: CAPM has only the market (MKT) return, "FF3" and "FF5" refers to the Fama and French (1992, 2015) models with added size (SMB), value (HML), profitability (RMW) and investment (CMA) factors, and "+1" to the addition of momentum (UMD). Treynor ratios are computed using CAPM betas. Results are reported under value-weights and equal-weights. This analysis is done on S&P500 stocks from 2003 to 2024 that have ESG scores in LSEG Workspace. Alphas are presented in monthly percentage terms. Sharpe, Treynor and Sortino ratios are annualized. T-statistics are shown in parentheses below coefficient estimates. 10% significance of a t-stat is denoted by *, 5% by ** and 1% by ***.

Table 9: Sample of our main panel data

	date	market equity	ret	ROE	Lev	assets	qmj	qmj_prof	qmj_growth	qmj_safety	matched_cusip	ESG_Score	Environmental_Pillar_Score	Social_Pillar_Score	Governance_Pillar_Score	in_sp500
10401	2003-04-30 00:00:00	13393.8656	0.052469037	0.056477626	3.241569409	55272	-0.701523898	0.708843635	-0.923948479	-0.974190851	001957505	38.77841006	17.05685619	32.0517719	68.92224788	1
10401	2003-05-30 00:00:00	15310.64218	0.143108547	0.056477626	3.241569409	55272	-0.71118658	0.703729393	-0.933685227	-0.973228154	001957505	38.77841006	17.05685619	32.0517719	68.92224788	1
10401	2003-06-30 00:00:00	15122.107	-0.002693677	0.056477626	3.241569409	55272	-0.716486335	0.698272111	-0.932814788	-0.974130522	001957505	38.77841006	17.05685619	32.0517719	68.92224788	1
10401	2003-07-31 00:00:00	16743.16436	0.104415596	0.059887782	2.886923165	50423	-0.580969382	0.933809641	-0.893168728	-0.942123683	001957505	38.77841006	17.05685619	32.0517719	68.92224788	1
10401	2003-08-29 00:00:00	17573.7374	0.048918109	0.059887782	2.886923165	50423	-0.56727184	0.937760044	-0.884022615	-0.940543689	001957505	38.77841006	17.05685619	32.0517719	68.92224788	1
10401	2003-09-30 00:00:00	16982.6924	-0.022982063	0.059887782	2.886923165	50423	-0.56749929	0.943672395	-0.894028381	-0.925770657	001957505	38.77841006	17.05685619	32.0517719	68.92224788	1

Jensen, Kelly and Pedersen (2023) Global Stock Returns and Characteristics database, all variables ending with "_Score" refer to ESG scores obtained from LSEG Workspace, "in_sp500", when equal to 1, indicates that the company was on the S&P500 for that date.