

## THE MATERIALITY CHALLENGE OF ESG RATINGS

**Zach Williams**

*Mercy College, Dobbs Ferry, New York, USA, E-mail address: [zwilliams1@mercy.edu](mailto:zwilliams1@mercy.edu), ORCID:  
<https://orcid.org/0000-0002-0854-7807>*

Received 21 November 2022; accepted 5 December 2022

### Abstract

**Research purpose.** The ESG literature supports a positive correlation between a firm's ESG performance and its financial performance; however, the details of this relationship are ambiguous, which has led to conflicting results in the literature. This article asserts that this ambiguity is largely related to the fact that ESG studies rarely consider the important issue of materiality. This article is an expanded replication of analysis by Williams & Apollonio (2022), a study that called for deeper analysis of both materiality and the causal link between ESG ratings and financial performance.

**Design / Methodology / Approach.** This methodology calculates Pearson Correlation coefficients between Bloomberg ESG scores and abnormal returns for S&P 500 firms from 2020 to 2022 after controlling for the material issue of GICS sectors.

**Findings.** The results show no relationship between ESG scores and abnormal returns, and the conclusion is that controlling for GICS sectors is not the methodology that will clarify the presumed positive correlation between ESG performance and financial performance.

**Originality / Value / Practical implications.** This study is one of the few that addresses materiality in ESG ratings. The finding that controlling for GICS sectors does not sufficiently control for materiality is a significant building block for future researchers. This article suggests that more granular categorization of sub-industries using a larger sample than the S&P 500 is likely to be useful future research.

**Keywords:** ESG Materiality; Abnormal Returns; Industry Classification.

**JEL codes:** G11.

### Introduction

Investing in a company based on an assessment of the environmental, social, and governance (ESG) risks that that company faces has become hugely popular in recent years. In 2020, there were \$35.3 Trillion of worldwide assets that were managed based on ESG requirements or factors; this was almost 40% of the world's managed assets in 2020, and it was an increase of 55% from 2016 (GSIA, 2021). Socially conscious investors have demanded these considerations (Zumente & Lace, 2021), and investment managers have found that considering these risks is consistent with traditional valuation (van Duuren et al., 2016); in other words, ESG investing is profitable.

However, there have been a number of concerns and criticisms in the ESG investing space. The lack of standardization (Whelan et al., 2021) of ESG ratings and the subjectivity (Wilhelmsen & Woods, 2021) involved are important concerns that have troubled empirical researchers. Different rating agencies might give the same company different ESG scores based on subjective agency criteria; furthermore, firms must voluntarily disclose ESG ratings in order to receive ESG ratings from an agency because there are no legal requirements for companies to provide standardized ESG information. This, of course, leads to a voluntary disclosure bias and has been associated with the practice of greenwashing, which involves, among other things, a company's attempt to benefit from the popularity of ESG investing by misleading consumers and investors about the environmental risks the company faces (Hayes, 2022a).

Another difficult challenge is the question of materiality. Which ESG factors are materially relevant to firms of different sizes, different products, different industries, or different operating models? In assessing the relationship between ESG ratings and financial performance, there is a two-fold challenge

of dealing with materiality issues: First, there is the question of which ESG metrics are financially relevant to a specific firm in a specific industry. Second, there is the question of exactly how that ESG metric translates into earnings and/or stock returns for a specific firm in a specific industry. The first question deals with the characteristics and risks of each industry, and the second question deals with how investors quantify the value associated with good or bad management of those risks (the second question could also deal with earnings and profitability, but researchers have been more likely to study valuation and stock returns). The research in this article attempts to address both questions by studying the relationship between abnormal returns and ESG ratings by controlling for the Global Industry Classification Standard (GICS) sector classification. Only a small number of researchers have attempted to address the problem of materiality, and this article will attempt to fill that gap. This article is an expanded replication of Williams & Apollonio (2022); that article noted the confusing issue of materiality and called for future research that might demonstrate how ESG performance drives financial performance and why confusion remains in the literature about the link between ESG performance and financial performance.

This article studies each of the 11 GICS sectors separately and runs regressions between abnormal returns and various ESG metrics. The ESG metrics are the individual current scores, the total current ESG score, the two-year changes in each individual score, and the two-year change in the total score. Materiality should vary by sector. For example, a company in the energy sector is likely to face different environmental risks than a company in the information technology sector. However, if both companies are successfully managing their different risks respectively, then both companies might receive the same high environmental ratings from ESG rating agencies. But the high environmental rating for the energy company and the high rating for the information technology company sends different signals to investors, so the valuation and stock returns could be different for the two companies with the same ESG ratings. Thus, academic researchers who run regressions of all firms with ESG scores are likely to find no correlation between environmental scores and abnormal returns, and these researchers might inaccurately conclude that environmental scores are not relevant to market returns. However, the researchers may have made a mistake by failing to consider that the high environmental score is more material for the energy company but less material for the information technology company. This article attempts to address this potential error in methodology by controlling for GICS sector classification. The risks that are financially material are likely to be similar within sectors, and the valuation implications of those risks are also likely to be similar within sectors.

Ultimately, this article does not find meaningful associations after controlling for sector classification. Thus, there are two important conclusions relevant to the future of ESG research. One, classifying by sector may not be the answer to understanding the materiality of ESG scoring as it relates to financial performance, and researchers should look elsewhere for control factors – such as company size or more granular classifications like the GICS industry or sub-industry. Two, the methodology in this article may not have been properly specified, and future researchers could adjust sample sizes, empirical tests, or other issues; nevertheless, addressing sector differences is an important step in ESG research, and this article will clearly articulate the steps in this process. The remainder of this article reviews the ESG literature, and explains the study methodology with the associated results, conclusions, limitations, and future research opportunities.

## **Literature Review**

This literature review shows that research in the ESG space decidedly points to a positive correlation between ESG ratings and financial performance, and these findings likely contributed to the explosion in ESG investing. However, this positive relationship is clouded with ambiguity and divergent study results. The key finding of Williams & Apollonio (2022) was that researchers, as a whole, do not agree on exactly how higher ESG ratings are causing better financial performance. Different research methodologies, voluntary ESG disclosures, and subjectively determined ESG ratings are some of the main drivers that have contributed to these divergent research results. The gap in the literature identified here is a lack of research to explain what controls are necessary for researchers to clearly identify and understand the relationship between ESG performance and financial performance.

## **ESG Ratings Drive Financial Performance**

A huge meta-study by Whelan et al. (2021) examines over 1,000 studies of ESG investing conducted between 2015 and 2020 and finds that most of the results show positive or neutral relationships between ESG ratings and financial performance. Whelan et al. (2021) also note that studies before 2015 pointed toward positive correlations between ESG scores and financial performance, and Whelan et al. (2021) credit those results with contributing to the rise in ESG investing. A highly cited individual study by Khan et al. (2016) demonstrated exciting ESG investing results by showing that investors could have earned high returns by investing in firms with strong, material ESG scores; King & Pucker (2022) acknowledge this study as feeding the ESG investing frenzy, and Berchicci & King (2022) note that this study is frequently cited by ESG researchers.

Williams & Apollonio (2022) reviewed the ESG literature, found an overall consensus of positive correlations between ESG scores and financial performance, and noted a number of important and supporting studies. Engelhardt et al. (2021), Ademi & Klungseth (2022), and Gregory (2022) all find positive correlations between ESG ratings and financial performance (as cited in Williams & Apollonio, 2022). Notably, Ademi & Klungseth (2022), and Gregory (2022) use S&P 500 data (as cited in Williams & Apollonio, 2022), which supports the assertion that there are positive correlations in S&P 500 data even though this article and Williams & Apollonio (2022) both found ambiguous results.

## **Unclear Causation between ESG Ratings and Financial Performance**

While the consensus in the literature points to a positive correlation between ESG scores and financial performance, several studies have found different results. A notable study by Alves et al. (2022) assesses more than 9,000 firms over 20 years in 46 countries using several control variables. The study finds no substantial evidence of a relationship between ESG scores and stock returns. Wilhelmsen & Woods (2021) found that ESG ratings and abnormal returns are negatively correlated; those researchers do not call the results a statistical anomaly; instead, they assert that this is consistent with risk-return asset management theory because riskier firms with lower ESG scores should yield higher returns. Hvidkjaer (2017) illustrates the ambiguity in the ESG literature by studying many different time periods; the study finds a positive correlation in one time period and no correlation in another period.

In addition to findings that refute the positive correlation, many studies point out problems with ESG metrics that make rigorous empirical testing difficult: inconsistent terminology, unstandardized and subjective ESG scores, and materiality problems are commonly mentioned challenges. Whelan et al. (2021) point to inconsistent terminology as a key problem in the ESG literature. Meuer et al. (2019) found 33 different definitions of corporate sustainability (as cited in Whelan et al., 2021), and the different definitions are often lumped together by different researchers, which creates confusion and noise in the literature (Douglas et al., 2017 as cited in Whelan et al., 2021). Whelan et al. (2021) comment on the unstandardized, subjective way that different ratings agencies issue ESG ratings, which would understandably lead to different conclusions. In addition to using unstandardized ESG data with inconsistent terminology, Whelan et al. (2021) note that there are different approaches to ESG investing (with different risks and return priorities), such as negative screening, socially responsible investing, ESG momentum, ESG integration, and decarbonizing. Whelan et al. (2021) note that comparing the results of different strategies without recognizing the difference leads to faulty conclusions. Other studies such as King & Pucker (2022), Berg et al. (2021), Gibson et al. (2021), and Christiansen et al. (2020) all note problems with unstandardized or subjective ESG ratings (as cited in Williams & Apollonio, 2022).

## **The Materiality Gap in the ESG Literature**

In addition to the challenges of inconsistent terminology and unstandardized ESG ratings, the challenge of materiality is mentioned by a number of studies, and that issue is the challenge addressed in this article. The study mentioned earlier by Khan et al. (2016) asserts that previous ESG studies did not find correlations between ESG ratings and stock performance because those studies did not consider materiality, and Khan et al. (2016) notes that that study is the first to incorporate materiality into ESG research. Khan et al. (2016) examine guidance on materiality from the Sustainability Accounting Standards Board (SASB) and develop materiality scores for each firm. Khan et al. (2016) claim that this

understanding of materiality is why the study results show a clear positive link between ESG ratings and financial performance. A later study by Khan (2019) also carefully examines materiality and again finds that investing in firms with high and material ESG scores yields high stock returns. In addition to Khan et al. (2016) and Khan (2019), Whelan et al. (2021) discussed above also note that researchers often fail to consider which ESG ratings are material to each company.

The problem with ignoring materiality is clear, but the solution is less evident. As mentioned, some ESG ratings are relevant to a firm, and some are not. For example, an environmental rating may be material for an oil company but immaterial for a technology company. But reliable materiality information can be challenging to find and problematic to use in analyses. Research by Berchicci & King (2022) criticizes the materiality calculation of Khan et al. (2016). Berchicci & King (2022) asserts that the flawed materiality calculation by Khan et al. (2016) erroneously led to a positive correlation, and in fact, there is not necessarily any added value from ESG investing. A related article by King & Pucker (2022) explains that the positive correlation in Khan et al. (2016) disappears if the faulty materiality calculation is removed.

The Bloomberg ESG ratings that will be used in this study consider materiality by rating different industries in different ways based on Bloomberg's assessment of what is financially material to each industry (Bloomberg Professional Services, 2022). But this does not solve the problem of correlating ESG scores with abnormal returns. As discussed in the introduction, the materiality problem is two-fold. The first issue seems to be addressed if Bloomberg rates firms based on financial material issues, but investors may see different financial valuation implications for two firms in different industries even though the firms have the same Bloomberg ESG scores, so further controls for materiality are necessary.

The need for studies that consider materiality represents a gap in the literature. The literature points to a positive correlation, but there have been numerous unexplained ambiguous results. Williams & Apollonio (2022) referred to this as the causation dilemma in ESG research; that study did not question the positive correlation, but it highlighted the confusion around how (and if) ESG actions and circumstances were causing strong financial performance. Very few studies have attempted to address materiality and those that have been subject to criticism. This article will attempt to fill this gap by testing the correlation among firms that are in the same sector. This analysis could go further to test firms that are similar in all ways except for ESG ratings, but this article will take an underpinning step in this direction to help researchers understand the importance of materiality in the correlation between ESG ratings and financial performance.

## **Research Methodology**

This article is an expanded replication of Williams & Apollonio (2022) which measures correlations between a firm's abnormal returns and its ESG scores. This article goes further by controlling for GICS sector classification. The firms used in this study are S&P 500 firms, and the abnormal return performance is measured from May 2020 to May 2022. Abnormal returns are calculated using the Capital Asset Pricing Model (CAPM), which calculates the excess return over the expected return based on the risk-free rate, the overall performance of the market, and the firm's correlation (beta) with the overall market (Kenton, 2022a).

The ESG ratings and abnormal return performance (including market rates and betas) are sourced from Bloomberg. Bloomberg is considered by investors to be one of the most dependable providers of global market data (Kolakowski, 2021). Bloomberg ESG scores are compiled from company-disclosed data; Bloomberg ensures that the self-reported metrics are relevant to 80% or more of a firm's workforce and operational imprint (Bloomberg Professional Services, n.d.b). The mechanics of Bloomberg ESG scores are based on elaborate, proprietary analyses that specifically consider the financial materiality of ESG risks for each industry (Bloomberg Professional Services, n.d.a). As discussed, while Bloomberg considers the financial materiality of risks, this materiality does not necessarily mean that the financial risks equate equally to abnormal returns in different industries. Bloomberg examines company information such as annual filings, sustainability reports, and company websites; Bloomberg attempts only to include data that are comparable across similar companies and comprehensive enough to illuminate the relevant financial risk (Bloomberg Professional Services, n.d.a). Bloomberg's

methodology identifies priorities based on the industry in order to formulate scoring that is industry-specific (Bloomberg Professional Services, n.d.a). Bloomberg recognizes the role of corporate governance in protecting shareholders from risk and identifying profitable opportunities; thus, Bloomberg governance scores evaluate issues such as shareholder rights, executive compensation, audit touch, and board composition (Bloomberg Professional Services, n.d.c).

This study calculates Pearson's Correlation coefficients, which measure the direction and strength of a linear relationship between two variables (Freed et al., 2014), and are the most commonly used correlation coefficients in scholarly research (Cozby & Bates, 2015). Pearson's Correlation coefficient calculates a p-value that measures the likelihood of randomly calculating the linear relationship from the sample data if there is truly no relationship in the population data (Freed et al., 2014). P-values below 5% or 1% (significance levels) are generally considered statistically significant results (Freed et al., 2014).

The Global Industry Classification Standard (GICS) defines firms by sector, industry group, industry, and sub-industry, and there are 11, 24, 69, and 158 of each, respectively (Hayes, 2022b). The GICS is a respected method to help investors identify a firm's competitors in the same line of business (Hayes, 2022b). For this research, the most granular distinction of sub-industries would be preferable; ideally, this study would compare firms that are similar in all ways except ESG ratings. However, the resulting small sample sizes of comparing 158 categories were not sufficient for sound empirical testing. Thus, this study used the 11 sector classifications with the goal of adequate sample sizes for each sector category; furthermore, the firms in each sector category were considered to be reasonably comparable because of the similarities of firms in the S&P 500. The Standard & Poor's 500 Index (S&P 500) is an index of leading US public companies that investors regard as an accurate gauge of global stocks (Kenton, 2022b).

The three hypotheses in this study are the same as the hypotheses of Williams & Apollonio (2022); the difference is that the data will first be categorized by GICS classification.

Hypothesis 1:

- $H_01$ : Correlation ( $R$ ) of Abnormal Returns vs. Current ESG Score = 0
- $H_a1$ : Correlation ( $R$ ) of Abnormal Returns vs. Current ESG Score  $\neq$  0

The abnormal returns are calculated as described above using the Capital Asset Pricing Model (CAPM) to determine the excess return beyond the expected return as calculated based on risk-free rates, market performance, and a given security's relationship with the overall market. The current ESG score includes three separate ESG scores (environmental, social, and governance) issued by Bloomberg as of May 2022 and one calculated summation score of all three scores added together. Hypothesis 1 tests the relationship between a firm's ESG performance and its financial performance.

Hypothesis 2:

- $H_02$ : Correlation ( $R$ ) of Abnormal Returns vs. Changes in ESG Scores = 0
- $H_a2$ : Correlation ( $R$ ) of Abnormal Returns vs. Changes in ESG Score  $\neq$  0

For hypothesis 2, the abnormal returns and the ESG scores are calculated as above in hypothesis 1. The change in ESG score is calculated from May 2020 to May 2022. ESG scores from both time periods were sourced from Bloomberg. Hypothesis 2 tests the relationship between changes in a firm's ESG performance and its financial performance.

Hypothesis 3:

- $H_03$ : Abnormal Returns of Firms with ESG Ratings are Equal to the  
Abnormal Returns of Firms Without Complete ESG Ratings
- $H_a3$ : Abnormal Returns of Firms with ESG Ratings are not Equal to the  
Abnormal Returns of Firms Without Complete ESG Ratings

For hypothesis 3, the abnormal returns are calculated as in hypotheses 1 & 2. Firms with complete ratings are defined as firms with all three Bloomberg ESG ratings: environmental, social, and governance. Hypothesis 3 tests the relationship between receiving full Bloomberg ESG ratings and financial performance.

## Research results

**Table 1. Descriptive statistics**

*Descriptive Statistics for ESG Study Data (S&P Companies 2020 to 2022)*

	Firms	Max	Min	Mean
Total S&P 500 Firms	501			
Environmental Rating	281	7.91	0	3.6931
Social Rating	281	9.57	0.69	3.7497
Governance Rating	484	8.69	4.08	6.6859
Firms with Complete Ratings	269			
Environmental Change	277	5.8	-3.06	0.7949
Social Change	277	7.68	-1.58	0.9704
Governance Change	468	1.79	-1.81	0.1048
Firms with Complete Change Ratings	259			
Abnormal Returns	501	4.27	-1.03	0.1638

Table 1 shows the available ESG data. Of the 501 S&P 500 firms, 269 have complete current ESG data and 259 have complete current and past ESG data, so the total change in ESG scores can be calculated for those 259 firms. Notably, there are 468 firms with current and past governance data, so analysis of changes in governance scores could be a rich area of analysis. The abnormal returns over the time period are given for all 501 firms, and the returns range from 427% to negative 103%.

**Table 2. Average abnormal return**

*Average Abnormal Return for Firms with Complete Current ESG Ratings and Firms without Complete Current ESG Ratings by Sector*

Sector	Total Firms in Sector	Percent of Total	With Complete Current Ratings			Without Complete Current Ratings			Mean Difference
			Mean Return	N	Percent of Total	Mean Return	N	Percent of Total	
Energy	21	4.2%	1.1775	20	7.4%	2.8584	1	0.4%	-1.6809
Material	28	5.6%	0.6765	24	8.9%	0.2487	4	1.7%	0.4278
Industrial	71	14.2%	0.1939	64	23.8%	0.0874	7	3.0%	0.1065
Consumer Discretionary	60	12.0%	0.0509	4	1.5%	0.1047	56	24.1%	-0.0538
Consumer Defensive	32	6.4%	-0.0326	29	10.8%	0.1976	3	1.3%	-0.2302
Health	64	12.8%	0.1445	1	0.4%	-0.0373	63	27.2%	0.1818
Financial	66	13.2%	-0.3076	5	1.9%	0.3031	61	26.3%	-0.6107
Information Technology	76	15.2%	0.0305	72	26.8%	-0.1496	4	1.7%	0.1801
Communications	25	5.0%	NA	0	0.0%	-0.1207	25	10.8%	NA
Utilities	28	5.6%	0.0332	25	9.3%	-0.0133	3	1.3%	0.0465
Real Estate	30	6.0%	0.1675	25	9.3%	0.2795	5	2.2%	-0.112
Total	501	100.0%	0.2129	269	100.0%	0.1069	232	100.0%	0.106

Table 2 shows the sector distribution of firms in the S&P 500, the details of firms in each sector with and without complete ESG ratings, the mean abnormal returns for the category, and the difference in mean returns for sector firms with complete ESG ratings and firms without complete ESG ratings. The percentage of total firms in the entire S&P 500 ranges from 4.2% for Energy firms to 15.2% for Information Technology firms. The percentages are notably lopsided for a few sectors with complete

ESG ratings; Industrials, Consumer Defensive, and Information Technology are notably larger percentages of the firms with complete ratings compared to the sector proportion in the S&P 500. These higher percentages may disproportionately influence the mean returns and correlations for firms with ratings. Similarly, the pool of firms without ratings may be disproportionally influenced by Consumer Discretionary, Health Care, Financials, and Communications. The other notable results in Table 2 are the mean difference of returns of firms with complete ratings and firms without complete ratings. Five of the sectors show higher mean abnormal returns for firms with complete ratings (Materials, Industrials, Health Care, Information Technology, and Utilities), and five of the sectors show higher mean returns for firms without complete ratings (Energy, Consumer Discretionary, Consumer Defensive, Financial, and Real Estate).

In addition to the noted instances of disproportionate representation noted above, there are a number of sectors that seem to be almost entirely with ESG scores or entirely without ESG scores. All or most of the firms in Energy, Industrials, Consumer Defensive, Information Technology, Utilities, and Real Estate firms have complete ESG information. Whereas all or most of the firms in Consumer Discretionary, Health Care, Financials, and Communications do not have complete ESG information. These sectors will likely influence the overall returns and correlations of firms with and without ESG information, respectively.

**Table 3. Correlation between abnormal return and ESG ratings**

*Abnormal Return Correlated with Current ESG Ratings - Firms with Complete Data*

	<b>Correlation</b>	<b>P-Value</b>	<b>N</b>
<b>Environmental Rating</b>	-0.051	0.401	269
<b>Social Rating</b>	0.134*	0.028	269
<b>Governance Rating</b>	0.072	0.239	269
<b>Total Current ESG Rating</b>	0.074	0.229	269

\*Statistically significant at the 5% level

\*\*Statistically significant at the 1% level

Table 3 shows the results of overall correlations between abnormal returns and ESG ratings. The results indicate that the social rating is the only rating correlated with abnormal returns, with a correlation coefficient of 0.134 and a p-value of 0.028. The environmental rating, the governance rating, and the total current rating are all uncorrelated with abnormal returns. The results in Table 3 are based on regressions of only firms with complete data. Williams & Apollonio (2022) found that including firms with incomplete data in the individual regressions had a negligible effect on the results, so those results are not included here. Additionally, including firms with incomplete data in the total current ESG rating correlations would not be meaningful because firms with incomplete data would arbitrarily be lower because of the missing data.

Table 4 and Table 6 are shown in Appendix A and Appendix B, respectively. Table 4 and Table 6 illustrate the key results of this study. Table 4 shows the correlation between abnormal returns and ESG scores when firms are separated by sector classification. Firms are further classified into firms with complete ratings and firms without complete ratings. As Table 4 illustrates, very few of the correlations are statistically significant. For firms with complete ratings, only two correlations are significant – the consumer defensive governance rating, and the information technology social rating. Both are moderately strong correlations (-0.381 & 0.274, respectively), and both are based on reasonable sample sizes (29 & 72, respectively). For the firms without complete ratings, there are three statistically significant correlations – industrials governance, real estate social, and real estate total. However, all three correlations are based on small sample sizes (6, 3, & 5, respectively).

**Table 5. Correlation between abnormal return and ESG ratings (change)***Abnormal Return Correlated with ESG Ratings Change - 2020 to 2022*

	Correlation	P-Value	N
<b>Environmental Rating Change</b>	-0.112	0.072	259
<b>Social Rating Change</b>	-0.094	0.133	259
<b>Governance Rating Change</b>	-0.067	0.285	259
<b>Total Current ESG Rating Change</b>	-0.136*	0.028	259

\*Statistically significant at the 5% level

\*\*Statistically significant at the 1% level

Table 5 shows the results of overall correlations between abnormal returns and ESG rating changes. The total current ESG rating change is negatively correlated with abnormal returns, with a Pearson Correlation coefficient of -0.136. There are no other significant correlations between the change in ESG ratings and abnormal returns.

Table 6 (shown in Appendix B) shows the correlations between abnormal returns and changes in ESG ratings by industry and the availability of Bloomberg ESG ratings. For firms with complete ratings, there are three statistically significant correlations – consumer defensive governance change, information technology environmental change, and real estate governance change. The correlation coefficients are -0.431, -0.577, and -0.184, respectively. Moreover, the sample sizes are 29, 71, and 25, respectively. For the firms without complete ratings, none of the correlations are statistically significant.

## Conclusions

In this research, none of the three null hypotheses is rejected. The overall results of Hypothesis 1 can be seen in Table 3: The social score is positively correlated with abnormal returns. However, the important tests in this article are those that control sector classification. Table 4 shows that when individual sectors are assessed, there are very few statistically significant correlations. There are only two significant correlations for firms with complete data and only three significant correlations for firms without complete data. Given a 5% significance level, it is presumable that 5% of the correlations could have been Type I Errors (rejecting the null hypothesis when no actual correlation exists (Freed et al., 2014)). Table 5 and Table 6 tell the same story for Hypothesis 2. Table 5 shows that there is an overall negative correlation between abnormal returns and the total change in ESG ratings from 2020 to 2022. However, Table 6 only shows three statistically significant correlations, which could very conceivably be Type I Errors. Table 2 demonstrates the results of Hypothesis 3. The mean difference in the last column of Table 2 indicates whether firms with or without ratings have higher returns. There is a difference seen in the total row of 10.6% higher for firms with complete ESG ratings, and this conclusion was explored in Williams & Apollonio (2022), but the difference disappears after controlling for sector classification. Five of the sectors have higher returns for firms with complete ESG ratings, five of the sectors have lower returns for firms with complete ESG ratings, and one sector is not comparable because there are no communications firms in the sample with complete ESG ratings. The even five-five split could not be any clearer of an indication that firms with ESG ratings do not have higher abnormal returns.

The key conclusion in this study is that after controlling for industry classification, there is not a statistically significant relationship between abnormal returns and ESG ratings. This study does not support the assertion that ESG scores are related to increased financial performance. However, this study is a valuable contribution to the literature as researchers attempt to understand how ESG ratings are related to financial performance. As demonstrated in the literature review, the clear consensus of the research points to a positive correlation between high ESG ratings and strong financial performance. Research by Williams & Apollonio (2022) supported the positive correlation but found numerous instances of ambiguity that warranted further research. This article takes an important step in that further research by suggesting that sector classification is not the answer to resolving the ambiguous relationship between financial performance and ESG ratings.



The secondary conclusion from this research is that firms with ESG information do not have higher abnormal returns when viewed by the industry. Williams & Apollonio (2022) found that firms with complete Bloomberg ESG ratings had higher abnormal returns than those without complete ESG ratings, but this study shows that the higher returns are not observed within industries. The findings of Williams & Apollonio (2022) could have been attributable to industry biases; for example, some industries are more likely to disclose ESG information, and those industries might have coincidentally performed better from 2020 to 2022. This is an inspiring finding that strengthens the power of ESG ratings. This study shows that ESG ratings are not meaningful just because they are given for a firm; investors did not reward companies with higher returns just because they disclosed ESG information and were issued ESG ratings. Instead, it seems more likely that investors carefully evaluate ESG scores and associate higher scores with higher company valuations (although this particular finding was not proven in this study).

This study shows a piece of the mosaic that illustrates the relationship between ESG scores and financial performance. The ESG literature strongly suggests that there is a positive correlation between high ESG ratings and strong financial performance; however, finding that relationship has been difficult for researchers who have struggled to understand the underlying cause of stronger financial performance. Is it that firms that disclose ESG information are stronger firms? Are ESG ratings a positive signal to investors? Do high ESG ratings indicate a fundamentally stronger company? Future research is necessary to answer these questions. This research showed that the answer is not any clearer after controlling for sector classification. A plethora of other controls could be assessed in future research, such as company size, investor base, company leverage, etc. The most logical extension of this article would be to control for the more granular industry group, industry, or sub-industry. This future research could be successful with large, homogenous sample sizes. The sample sizes in this article would not have been large enough to see correlations within 158 sub-industry classifications, so pools of data larger than the S&P 500 would be necessary. However, these larger samples must either be companies that are similar (except for ESG ratings), or the researchers must adequately control for these differences in order to determine the isolated effect of different ESG scores.

## References

- Alves, R., Krueger, P., & van Dijk, M. (2022). Drawing Up the Bill: Does Sustainable Investing Affect Stock Returns Around the World? *Network for Studies on Pension, Aging and Retirement*.
- Berchicci, Luca and King, Andrew A., Corporate Sustainability: A Model Uncertainty Analysis of Materiality (May 18, 2021). Available at SSRN: <https://ssrn.com/abstract=3848664> or <http://dx.doi.org/10.2139/ssrn.3848664>
- Bloomberg Professional Services. (2022). *Environmental & Social Scores*. Bloomberg. Retrieved May 20, 2022, from <https://assets.bbhub.io/professional/sites/10/Environmental-Social-Scores-Fact-Sheet1.pdf>
- Bloomberg Professional Services. (n.d.a) *Environmental and Social Scores. Methodology and Field Information*. Retrieved December 2, 2022, from Bloomberg Terminal.
- Bloomberg Professional Services. (n.d.b) *Global Environmental, Social & Governance – ESG Data*. Retrieved June 5, 2022, from <https://www.bloomberg.com/professional/dataset/global-environmental-social-governance-data/>
- Bloomberg Professional Services. (n.d.c) *Governance Scores. Methodology and Field Information*. Retrieved December 2, 2022, from Bloomberg Terminal.
- Cozby, P., & Bates, S. (2015). *Methods in behavioral research*. New York, NY: McGraw-Hill Education.
- van Duuren, E., Plantinga, A. & Scholtens, B. ESG Integration and the Investment Management Process: Fundamental Investing Reinvented. *J Bus Ethics* 138, 525–533 (2016). <https://doi.org/10.1007/s10551-015-2610-8>
- Freed, N., Jones, S., & Bergquist, T. (2014). *Understanding business statistics*. Hoboken, NJ: John Wiley & Sons.

- Global Sustainable Investment Review 2020* - [gsi-alliance.org](http://www.gsi-alliance.org). (2021). Retrieved from <http://www.gsi-alliance.org/wp-content/uploads/2021/08/GSIR-20201.pdf>
- Hayes, A. (2022a). *What is greenwashing? how it works, examples, and Statistics*. Investopedia. Retrieved November 19, 2022, from <https://www.investopedia.com/terms/g/greenwashing.asp>
- Hayes, A. (2022b). *What is the Global Industry Classification Standard (GICS)?* Investopedia. Retrieved November 19, 2022, from <https://www.investopedia.com/terms/g/gics.asp>
- Hvidkjær, S. (2017). *ESG Investing: A literature Review. Report Prepared for Dansif*.
- Kenton, W. (2022a). *Capital Asset Pricing Model (CAPM) and assumptions explained*. Investopedia. Retrieved November 19, 2022, from <https://www.investopedia.com/terms/c/capm.asp>
- Kenton, W. (2022b). *S&P 500 index: What it's for and why it's important in investing*. Investopedia. Retrieved November 19, 2022, from <https://www.investopedia.com/terms/s/sp500.asp>
- Khan, Mozaffar N., George Serafeim, and Aaron Yoon. "Corporate Sustainability: First Evidence on Materiality." Harvard Business School Working Paper, No. 15-073, March 2015.
- Khan, M. (2019). Corporate governance, ESG, and stock returns around the world. *Financial Analysts Journal*, 75(4), 103–123. <https://doi.org/10.1080/0015198x.2019.1654299>
- King, A., & Pucker, K. (2022, March 3). *ESG and alpha: Sales or substance?* Institutional Investor. Retrieved from <https://www.institutionalinvestor.com/article/b1wxqznltqnyzj/ESG-and-Alpha-Sales-or-Substance>
- Kolakowski, M. (2022). *Bloomberg vs. Reuters: What's the difference?* Investopedia. Retrieved from <https://www.investopedia.com/articles/investing/052815/financial-news-comparison-bloomberg-vs-reuters.asp>
- Whelan, T., Atz, U., Van Holt, T., & Clark, C. (2021). *ESG and financial performance. Uncovering the Relationship by Aggregating Evidence from, 1,000 Plus Studies Published between 2015-2020*. NYU Stern Center for Sustainable Business. Retrieved May 22, 2022 from [https://www.stern.nyu.edu/sites/default/files/assets/documents/NYU-RAM\\_ESG-Paper\\_2021.pdf](https://www.stern.nyu.edu/sites/default/files/assets/documents/NYU-RAM_ESG-Paper_2021.pdf)
- Wilhelmsen, E., & Woods, E. (2021). *Esg Ratings and Stock Performance* (thesis).
- Williams, Z., & Apollonio, H., (2022). The causation dilemma in ESG research [Manuscript submitted for publication].
- Zumente, I., & Lāce, N. (2021). ESG rating—necessity for the investor or the company? *Sustainability*, 13(16), 8940. <https://doi.org/10.3390/su13168940>

## Appendix A

Table 4 Correlations between Abnormal Returns and Current ESG Ratings

	Firms with Complete Ratings			Firms without Complete Ratings			Correlation Difference Absolute Value
	Pearson Correlation	P-Value	N	Pearson Correlation	P-Value	N	
Energy Environmental Rating	0.068	0.777	20	NA	NA	1	NA
Energy Social Rating	-0.079	0.741	20	NA	NA	1	NA
Energy Governance Rating	0.388	0.091	20	NA	NA	1	NA
Energy Total Current ESG Rating	0.8	0.739	20	NA	NA	1	NA
Materials Environmental Rating	-0.206	0.333	24	NA	NA	1	NA
Materials Social Rating	0.113	0.599	24	NA	NA	1	NA
Materials Governance Rating	0.316	0.132	24	0.458	0.697	3	-0.142
Materials Total Current ESG Rating	0.05	0.816	24	0.069	0.931	4	-0.019
Industrials Environmental Rating	-0.178	0.158	64	NA	NA	1	NA
Industrials Social Rating	0.023	0.857	64	NA	NA	1	NA
Industrials Governance Rating	-0.042	0.74	64	.851*	0.032	6	-0.809
Industrials Total Current ESG Rating	-0.088	0.489	64	0.659	0.108	7	-0.571
Consumer Discretionary Environmental Rating	0.251	0.749	4	NA	NA	0	NA
Consumer Discretionary Social Rating	-0.305	0.695	4	NA	NA	0	NA
Consumer Discretionary Governance Rating	0.106	0.894	4	-0.152	0.267	55	-0.046
Consumer Discretionary Total Current ESG Rating	0.168	0.832	4	-0.123	0.365	56	0.045
Consumer Defensive Environmental Rating	-0.215	0.262	29	NA	NA	1	NA
Consumer Defensive Social Rating	0.105	0.587	29	NA	NA	1	NA
Consumer Defensive Governance Rating	-.381*	0.042	29	NA	NA	2	NA
Consumer Defensive Total Current ESG Rating	-0.194	0.313	29	0.983	0.117	3	-0.789
Health Care Environmental Rating	NA	NA	1	NA	NA	0	NA
Health Care Social Rating	NA	NA	1	NA	NA	0	NA
Health Care Governance Rating	NA	NA	1	0.243	0.064	59	NA
Health Care Total Current ESG Rating	NA	NA	1	0.061	0.635	63	NA
Financial Environmental Rating	0.686	0.201	5	NA	NA	0	NA
Financial Social Rating	0.025	0.968	5	NA	NA	0	NA
Financial Governance Rating	-0.409	0.495	5	-0.138	0.288	61	0.271
Financial Total Current ESG Rating	0.355	0.558	5	-0.138	0.288	61	0.217
IT Environmental Rating	-0.146	0.22	72	-0.843	0.361	3	-0.697
IT Social Rating	.274*	0.02	72	-0.295	0.809	3	NA
IT Governance Rating	0.09	0.451	72	NA	NA	1	NA
IT Total Current ESG Rating	0.069	0.562	72	-0.813	0.187	4	-0.744
Communications Environmental Rating	NA	NA	0	NA	NA	0	NA
Communications Social Rating	NA	NA	0	NA	NA	0	NA
Communications Governance Rating	NA	NA	0	0.083	0.693	25	NA
Communications Total Current ESG Rating	NA	NA	0	0.083	0.693	25	NA
Utilities Environmental Rating	0.163	0.437	25	NA	NA	2	NA
Utilities Social Rating	-0.247	0.234	25	NA	NA	2	NA
Utilities Governance Rating	-0.216	0.299	25	NA	NA	1	NA
Utilities Total Current ESG Rating	-0.123	0.557	25	-0.776	0.435	3	-0.653
Real Estate Environmental Rating	-0.188	0.367	25	-0.31	0.799	3	-0.122
Real Estate Social Rating	0.154	0.463	25	-1.000*	0.02	3	NA
Real Estate Governance Rating	0.143	0.495	25	NA	NA	2	NA
Real Estate Total Current ESG Rating	0.071	0.736	25	-.925*	0.025	5	NA

\*Statistically significant at the 5% level

\*\*Statistically significant at the 1% level

## Appendix B

Table 6 Correlations between Abnormal Returns and ESG Changes

	Firms with Complete Ratings			Firms without Complete Ratings			Correlation
	Pearson Correlation	P-Value	N	Pearson Correlation	P-Value	N	Difference Absolute Value
Energy Environmental Change	0.101	0.682	19	NA	NA	2	NA
Energy Social Change	-0.2	0.412	19	NA	NA	2	NA
Energy Governance Change	-0.045	0.853	19	NA	NA	0	NA
Energy Total Current ESG Change	-0.03	0.904	19	NA	NA	2	NA
Materials Environmental Change	-0.334	0.139	21	NA	NA	2	NA
Materials Social Change	-0.184	0.425	21	NA	NA	2	NA
Materials Governance Change	0.285	0.211	21	NA	NA	0	NA
Materials Total Current ESG Change	-0.276	0.225	21	NA	NA	2	NA
Industrials Environmental Change	-0.131	0.325	59	-0.604	0.396	4	-0.473
Industrials Social Change	0.006	0.966	59	0.642	0.358	4	-0.636
Industrials Governance Change	0.078	0.557	59	0.329	0.525	6	-0.251
Industrials Total Current ESG Change	-0.057	0.666	59	0.337	0.284	12	-0.28
Consumer Discretionary Environmental Change	0.62	0.38	4	NA	NA	0	NA
Consumer Discretionary Social Change	-0.04	0.96	4	NA	NA	0	NA
Consumer Discretionary Governance Change	0.85	0.15	4	-0.046	0.739	54	0.804
Consumer Discretionary Total Current ESG Change	0.567	0.433	4	-0.048	0.727	56	0.519
Consumer Defensive Environmental Change	-0.204	0.289	29	NA	NA	1	NA
Consumer Defensive Social Change	-0.06	0.758	29	NA	NA	1	NA
Consumer Defensive Governance Change	-.431*	0.019	29	NA	NA	2	NA
Consumer Defensive Total Current ESG Change	-0.285	0.133	29	-0.032	0.98	3	0.253
Health Care Environmental Change	NA	NA	1	NA	NA	0	NA
Health Care Social Change	NA	NA	1	NA	NA	0	NA
Health Care Governance Change	NA	NA	1	0.24	0.07	58	NA
Health Care Total Current ESG Change	NA	NA	1	0.231	0.068	63	NA
Financial Environmental Change	0.139	0.823	5	NA	NA	0	NA
Financial Social Change	-0.312	0.61	5	NA	NA	0	NA
Financial Governance Change	-0.393	0.513	5	0.054	0.682	61	0.339
Financial Total Current ESG Change	-0.156	0.803	5	0.054	0.682	61	0.102
IT Environmental Change	-.242*	0.042	71	0.428	0.572	4	NA
IT Social Change	-0.013	0.915	71	0.062	0.938	4	-0.049
IT Governance Change	0.082	0.498	71	NA	NA	1	NA
IT Total Current ESG Change	-0.149	0.215	71	0.552	0.335	5	-0.403
Communications Environmental Change	NA	NA	0	NA	NA	0	NA
Communications Social Change	NA	NA	0	NA	NA	0	NA
Communications Governance Change	NA	NA	0	0.175	0.436	22	NA
Communications Total Current ESG Change	NA	NA	0	0.159	0.446	25	NA
Utilities Environmental Change	0.128	0.542	25	NA	NA	2	NA
Utilities Social Change	-0.276	0.182	25	NA	NA	2	NA
Utilities Governance Change	-0.152	0.468	25	NA	NA	1	NA
Utilities Total Current ESG Change	-0.129	0.54	25	-0.444	0.707	3	-0.315
Real Estate Environmental Change	-0.104	0.62	25	0.363	0.763	3	-0.259
Real Estate Social Change	-0.12	0.567	25	-0.709	0.499	3	-0.589
Real Estate Governance Change	-.577**	0.003	25	NA	NA	2	NA
Real Estate Total Current ESG Change	-0.184	0.378	25	-0.832	0.081	5	-0.648

\*Statistically significant at the 5% level

\*\*Statistically significant at the 1% level