University of Stavanger Business School
Stavanger, Spring 2021

New empirical evidence for popularity in company ESG data

Nikola Manojlovic and Kristian Tjessem

Supervisor: Kristoffer W. Eriksen

Master of Science in Business Administration

Major: Applied Finance

UNIVERSITY OF STAVANGER BUSINESS SCHOOL

Universitetet
i Stavanger

UNIVERSITY OF STAVANGER BUSINESS SCHOOL

MASTER'S THESIS

STUDY PROGRAMME: Master of Science in Business and Administration	THIS THESIS HAS BEEN WRITTEN WITHIN THE FOLLOWING FIELD OF SPECIALISATION: Applied Finance IS THE THESIS CONFIDENTIAL? **(NB!** Use the red form for confidential theses)

TITLE: New empirical evidence for popularity in company ESG data

AUTHOR(S)		SUPERVISOR:
Candidate number:	Name:	Kristoffer W. Eriksen
2066	Nikola Manojlovic	
2058	Kristian Tjessem	

Acknowledgements

This thesis was written in the spring of 2021 as a final step toward a master's degree in Business Administration at the University of Stavanger.

The process was both fun and challenging.

We would like to thank our supervisor Kristoffer W. Eriksen for advice along the way.

University of Stavanger Business School

Stavanger, June 2021

Nikola Manojlovic

Kristian Tjessem

Abstract

ESG awareness among both professional and retail investors has seen an increase in the last two decades. A large part of the literature finds that investing in companies with a high ESG score comes at a cost. We study to what extent company ESG data is consistent with the popularity framework, which claims that assets with preferences investors find popular demand a higher premium. 5917 global stocks were sorted into four quantiles based on their ESG scores. Backtesting was performed on these quantiles from 2003-2019. Analysis on global economic sectors and country specific ESG was also peformed. Overall we conclude that a company's ESG score is consistent with the popularity framework. The study shows that the least popular quantile, apart from a few minor cases, achieves superior returns when compared to the most popular quantile.

Keywords – Classical finance, behavioural finance, popularity, ESG

Contents

List of Figures

List of Tables

1 Introduction

In with the new and out with the old? A phrase that can be applied to many aspects of life. But does it have any merit in the financial world? The financial literature has to date been unable to come up with a single agreed upon explanation for the various market premiums and anomalies that exist. Broadly, there exists two fields of finance, classical and behavioral finance. Classical finance is based on the fundamental idea that investors are risk averse. This means that the market premiums encountered are referred to as risk premiums. Behavioral finance on the other hand is based on an idea that premiums observed are a result of either cognitive errors investors systematically make or preferences for company or security characteristics that might not be related to risks (Ibbotson et al., 2018). The intersection between classical finance and behavioral finance is an interesting place. This new intersection has seen much research in the past three decades. Ibbotson et al. presented their attempt to unite the thoughts of classic and behavioral finance, with the popularity hypothesis. The idea of popularity is that investors have preferences for certain characteristics of an asset. Some preferences are more popular than others. Some are almost universally liked or disliked. An asset with popular characteristics is going to have a higher demand, which results in higher pricing. Relative to the same cash flows, the asset is therefore going to have lower returns. This implies that an investor takes the contrary approach by buying the less popular and ignored asset at a lower price and ultimately achieves a higher return. Ibbotson et al. (2018) identified several new and previously unstudied company characteristics that could serve as proxies for dimensions of popularity. These characteristics were the company brand, a companies sustainable competitive advantage and a companies reputation. However, a company's ESG metrics have yet to be studied through the lens of popularity. Therefore the goal of this thesis is to explore the ESG factors of a company in order to determine whether they are consistent with popularity. The research question then aptly becomes:

Are a companies ESG metrics consistent with popularity?

This thesis consists of five parts. In the theory chapter we first go through the theory that leads up to the popularity framework and explain the idea and how it relates to the purpose of this thesis. We then go through the discussion in the literature on ESG, the profitability of ESG investing and if ESG investing works. The third chapter explains the methodology used in our analysis, its purpose, strengths and limitations. The fourth chapter show the result of our analysis, both with the full data set, by sector and countries to get the best insight as possible. In the fifth chapter we discuss our results in terms of the popularity framework and in context with literature on ESG investing, before finally concluding our thesis in the last chapter.

2 Theory

The ultimate goal of active investing is to beat the market in the long run. To outperform a benchmark index by skill rather than by random luck. Active investing is a broad umbrella term of investing where the investor takes active action in the selection, monitoring and evaluation of financial assets in their portfolio. It can take the form of different strategies such as value investing, growth investing or frequent trading. If one subscribes to the efficient market hypothesis and view the markets as efficiently priced, the costs of active investing should make it extremely hard (if not impossible) to outperform the market in the long run. Despite these observations, there are stories of legendary investors that have managed to successfully beat the market over time. CAPM, limited by it's single factor; market risk, is seemingly unable to explain these anomalies. Since the CAPM's inception in the 1960's, a great number of additional factors have been proposed. In their seminal paper from 1993, Eugene Fama and Kenneth French proposed two additional factors, the value and size factors (Fama and French, 1993). Carhart introduced momentum some five years later (Carhart, 1997). In 2015 Fama and French supplemented their initial model with two extra factors, profitability and investments (Fama and French, 2015). Over the years a great deal of factors have been proposed. So much that it has become a factor zoo (Feng et al., 2020). There have also been made attempts to adopt behavioral variables into asset pricing models to include how cognitive errors and limitations of market participants influence the dynamics of asset pricing (Shefrin and Statman, 1994).

Ibbotson et al. (2018) proposes their "popularity hypothesis" to explain risk premiums and anomalies that the other models fail to explain. They claim that the hypothesis serves as a link between classical and behavioral finance.

In this chapter we will go through the theory and the discussions in the literature, from classical finance to behavioral finance, the popularity hypothesis and ESG investing. Key focus points will be: (1) Why behavioral finance has emerged as an attempt to better understand what drives asset prices and returns, extending finance by adding concepts from psychology and sociology. (2) Explain the theory and empirical evidence of the popularity hypothesis and how it connects classical and behavioral finance. (3) How ESG investing applies to the popularity hypothesis and empirical evidence on ESG investing.

2.1 From classic to behavioral finance

Classic (or standard) finance, that is taught in most finance classes, is based on assumptions of risk averse and rational utility maximizing investors that operate in an efficiently priced market where an individual investor who outperforms the market can only be explained by luck or random variations around the market benchmark. The foundations of classical finance theory were developed in the 1950's and 1960's. Theories such as mean-variance portfolio theory (Markowitz, 1952), definitions of rational investors (Miller and Modigliani, 1961), asset pricing models such as CAPM (Sharpe, 1964; Lintner, 1965) and the efficient market hypothesis (Fama, 1970). These important works contributed strongly to what has been the leading paradigm of financial theory for the last 70 years. This paradigm consists of the perceived positive linear relationship between risk and return, benefits of portfolio diversification and the efficiency of the markets. These ideas suggest that all market participants should be passive investors that only invest in the broad market. If these theories hold, why are there active investors in the market?

While these seminal works are the pillars of classical finance, there exists a growing body of empirical evidence that challenge these theories and claim their assumptions have considerable shortcomings. The risk-return paradigm in finance has been challenged by several researchers in the past. Mehra and Prescott (1985) coined the term "the equity puzzle", which claimed that the equity premium had been too large compared to what academic models, such as CAPM, would predict. Vendrame et al. (2016) claim the simplicity and theoretical appeal of the CAPM model appears to have strongly outweighed the shortage of empirical evidence in its favour.

Market efficiency is also a pillar of classic finance that have been challenged. Ibbotson et al. (2018) argue that the markets are "beyond efficient", in which information irrelevant to fair value is reflected in security prices as a result of investors' behavioral preferences. In such a market, they consider prices to be biased, in contrast to fair prices.

Behavioral finance eases the assumption of rational market participants, which seem to be one of the major reason why the classical theories and models do not hold in practice.

2.2 Behavioral finance

Behavioral finance emerged as an attempt to explain anomalies that the neoclassical models failed to do. Here investors are treated as "normal" rather than strictly rational, have limits to their self control, affected by their own biases and make decisions influenced by cognitive errors. Tversky and Kahnemann (1974) pioneered what became the building blocks of behavioral economics and behavioral finance. They presented evidence that people in general are less rational than the literature previously had assumed, explained by heuristics and biases that we all tend to have and act on to varying degrees. Five years later they first presented their prospect theory, assessing decision making under risk. The value function in the prospect theory was a behavioral alternative to the utility function in the expected utility theory. The value function estimated the value that people assigned to their wins and losses. The article introduced the concept of "loss aversion", which suggested that people disliked losses more than they cared for equivalent gains. (Kahnemann and Tversky, 1979). They claimed that this loss aversion was a integral of the decision making process. In a finance perspective this can be translated to that investors fear and value the downside risk more than they value the upside potential of an investment. Apart from generally dismissing the market efficiency hypothesis (Grossman and Stiglitz, 1980; De Bondt and Thaler, 1985; Titman and Jegadeesh, 1993), the literature in behavioral finance include concepts such as overconfidence (Harrison and Kreps, 1978), overreaction (De Bondt and Thaler, 1985), confirmation bias, hindsight biases and herd behaviour.

While behavioral finance has progressed a lot through the years, it has yet to provide a full framework or theory to explain and understand asset prices (Ibbotson et al., 2018).

2.3 Popularity hypothesis

Yale professor emiritus Roger Ibbotson and his co-authors from Morningstar presented their attempt to create a bridge between the ideas of classical and behavioral finance by presenting their "popularity framework", which they claim could give an explanation to all premiums and anomalies that the classical finance framework alone struggle to do (Ibbotson and Idzorek, 2014; Idzorek and Ibbotson, 2017; Ibbotson et al., 2018). The idea was first introduced in their 2014 article. In 2017 they studied factor-based quartile

portfolios through the lens of popularity and in 2018 a theory was presented with a fully developed mathematical framework. It draws insights from both the traditional views in finance and from behavioral research. The popularity hypothesis is a simple idea, yet it might seem counter-intuitive. The idea is that people and investors have preferences for certain characteristics of an asset. Some preferences are generally liked or disliked, or in other terms, certain types of characteristics are more popular than others. An asset with popular characteristics is going to be more demanded and therefore higher priced. When the asset is higher priced, relative to the same cash flows, the asset is going to have lower returns. This implies that an investor should buy the less popular and overlooked assets, which have a lower price and ultimately yield a higher return. Reasons for liking or disliking a certain characteristic or security can be both *rational* or *irrational*. The theory also acknowledges that an investor cares for non-risk factors that don't fit into the risk/return paradigm which classical finance often assumes away. Examples of such is liquidity, taxability, scalability, controllability and sustainability (ESG).
The thought of popularity in the security market is not new. Graham and Dodds recognized this, calling it votes, in their famous book *Security Analysis* over 80 years ago:

> "In other words, the market is not a weighing machine, on which the value of each issue is recorded by an exact and impersonal mechanism, in accordance with its specific qualities. Rather should we say that the market is a voting machine, whereon countless individuals register choices which are the product partly of reason and partly of emotion." (Graham and Dodd (1934, p.28), cited in Ibbotson and Idzorek (2014))

Ibbotson et al. (2018) showed empirically that most characteristics were consistent with the popularity framework. They tested their hypothesis on both traditional fundamental data and data that can be considered more behavioral in nature. This included both characteristics of the companies itself and characteristics of the stocks of the company. The authors tested return against previously unstudied characteristics of a company, which should almost universally be liked and disliked and thus serve as a proxy for the dimensions of popularity. The company specific features are powerful brand, a sustainable competitive advantage and good reputation, all three features thought to be popular characteristics of a company. The features that were tested that are linked to stocks of the company, were severe tail risk (low or negative coskewness) and lottery-like pay-offs. They argue that the former is nearly uniformly disliked, as opposed to the latter characteristic, which they claim to be a very popular characteristics. They ranked all these from least popular to most popular and divided them into four even quartiles. In their analysis they split the resulting quartiles into both cap-weighted and equally weighted portfolios. Their analysis showed that all 10 showed at least some consistency with the theory of popularity and that 5 were consistent with the more-risk/more-return paradigm .

They also took the same approach when testing traditional factors against return to test the popularity framework. These factors were the usual systematic risk factors such as CAPM beta, daily and monthly volatility, firm specific risk factors such as market cap, total assets, revenue, net income, B/M, E/P and the nonrisk factor illiquidity. Their analysis showed that 7 out of 10 factors were consistent with the theory of popularity and only two with the risk/return paradigm (Table 2.1)

Table 2.1: Consistency with popularity and risk/return from Ibbotson et al. (2018)

Characteristics	Consistent with popularity	Consistent with risk/return paradigm
Quality characteristics:		
Brand	Yes	No
Competitive advantage	Yes	Yes
Company reputation	Yes	Yes
Tail-risk	Yes	No
Lottery stocks	Mixed	Mixed
Factors:		
CAPM Beta	Yes	No
Daily Volatility	Yes	No
Monthly Volatility	Yes	No
Market Cap	Yes	Yes
Total Assets	No	No
Revenue	No	No
Net Income	No	No
Book/market	Yes	No
Earnings/Price	Yes	No
Illiquidity	Yes	Yes

2.4 ESG

While the primary goal of investments is to maximize the return for the investor, there is a growing segment of investors that are incorporating their social and environmental standards into the investment decision (Camilleri, 2020). How the particular firm operates in regard to sustainability and societal impact is often described through the three main factors environmental, social and corporate governance (ESG). In investment terms, ESG is by clients and professionals often used interchangeably with socially responsible investing (SRI) and impact investing. Langbein and Posner (1980) defined socially responsible investing as "excluding the securities of otherwise attractive companies from an investor's portfolio because the companies are judged to be socially irresponsible, and including the securities of certain otherwise unattractive companies because they are judged to be behaving in a socially laudable way". One way an investor can integrate sustainability into or evaluate the ESG performance of their portfolio, is to obtain ESG-scores for each asset in their portfolio combining into an overall ESG-score for the portfolio as a whole. Hence, the investor can select assets that results in a portfolio that suits the investors taste for ESG. The rationale behind a preference in investing in stocks with good ESG records is for the investor to promote and make a positive impact on social and environmental challenges (Blitz and Swinkels, 2020).

Ibbotson et al. (2018, p.2, p.39) suggest that investing in companies with strong ESG characteristics might be more in line with what the literature considers as "behavioral", than with classic finance.

2.4.1 The effect of ESG investing on portfolio performance

A heavily debated issue is the compatibility between the return-maximization objective of an investment and ESG. Many researchers have found that a focus on ESG leads to lower returns, while ESG proponents argue that ESG awareness give higher returns. Silva and Cortez (2016) claim that green funds tend under-perform, especially in times with absence of market turmoil. Adler and Kritzman (2008) point out that the investment manager need to be awere of the potential cost of socially responsible investing and that the investor suffers a financial loss if one "bad" company outperforms it's "good" substitute. They conclude that the cost of responsible investing is substantial.

Kempf and Osthoff (2007) on the other hand, showed results that responsible investing strategies worked. By taking long positions in stocks of socially responsible firms and selling short stocks of socially irresponsible firms, they claimed that their portfolio had abnormal returns of close to 9% per year. Their results were particularly convincing when only incorporating the firms with extreme scores (either positive or negative) on social responsibility. This is supported by a large meta-study by Friede et al. (2015) does not find evidence to support the common investor-perception of a mixed/neutral relationship between ESG and corporate financial performance. In contrast, Blitz and Fabozzi (2017) argues that there are compelling evidence that so called sin stocks provide abnormally high raw returns and that this can be explained with the two new factors in the 5-factor model from Fama and French (2015), profitability and investments. Blitz and Fabozzi further claim that a portfolio that excercise an exclusion policy of sin stocks will damanage it's return. A way to deal with with is if the manager of the portfolio can restore the exposure of the "good" stocks to these two quality factors, the expected return may be restored. Broadstock et al. (2021) examined the performance of high and low ESG portfolios during, prior and past peroids of financially distress, most notably during covid-19 crisis. They found that generally, high ESG scoring portfolios have outperformed low ESG portfolios, but that this difference was higher during times of financial distress. In normal times the low ESG portfolio did relatively better, but that the high ESG portfolio still outperformed. Other researchers have taken an approach that combine the returns of assets with ESG scores and ESG preference of the investor, thereby getting an ESG-adjusted return. Cooper et al. (2016) tries to quantify the impact the firm has and together with the usual financial performance, constructing a combined measure for evaluating the investment and therefore acknowledging that some investors care for more than monetary return. They produce a model where the ESG impact of the stock are paired with the degree of utility the investor has from investing in responsible stocks. They assign a "social coefficient" to the investor that is multiplied with the stocks impact, which produce a *social return* and a *social sharpe ratio* for the investment. Similarly, Pedersen et al. (2020) suggests an ESG-adjusted CAPM, where they divide investors into either ESG-unaware, -aware or -motivated from their preferences on the ESG factor. This results in an ESG-adjusted expected return for det individual stocks based on the the normal CAPM expected return in combination with the ESG score of the firm and the ESG preference of the investor.

2.4.2 Does ESG investing work?

Apart from the debate in the literature of potential cost or higher return of investing in companies with high ESG score, is whether investing in stocks with good ESG records promotes the goal of making a positive impact on social and environmental challenges as intended. The discussion in the literature on this topic is mainly centered around whether it works or not, whether to exclude firms with low ESG score from portfolio (ESG screening), if there are other, more effective ways to make an impact on these challenges and how to include non-financial performance when evaluating investments and maximizing utility for investors.

Blitz and Swinkels (2020) examine if decisions to exclude stocks with low ESG rating from the portfolio is meaningful if the goal is to make a positive impact. They argue that exclusion is effectively a transfer of ownership from concerned to less-concerned investor and the positive impact this have on the environment and society is less than obvious. They argue that investment managers like to promote sustainability issues because it is easier to claim success in this area than actually beating the market with higher returns. The article address and challenge four different ways to legitimate exclusion policies; increasing a firm's cost of capital, driving a firm out of business, improving investment performance, and signaling stakeholders to change behavior. They conclude that the justification of the various arguments in favor of excluding firms is doubtful and it is unclear if such a policy accomplish anything meaningful in the real world. Voting and engaging as active shareholders might serve the purpose in a more meaningful way. Camilleri (2020) points out that positive screening, as opposed exclusion policies (negative screening), are a better and more consistent approach to social and responsible investing. Positive screening does not exclude sin stocks, but rewards socially and environmentally responsible firms using some form of a point system. Adler and Kritzman (2008) notes that the managers selecting companies with good ESG score because the think they are going to outperform the low ESG scoring companies, are not engaging in responsible investing in itself, but rather in active management.

2.4.3 ESG in the perspective of the popularity hypothesis

Following the popularity hypothesis of Ibbotson et al. (2018), the increase in popularity of stocks with high ESG-scores should lead to lower returns due to the increased demand leads to higher prices and subsequently lower yields. As an increasing share of investors tend to avoid sin stocks and seek out responsible investments, this should give a opportunity to the investors that do no care for ESG and socially responsible investing to buy sin stocks at a discount and earn higher returns.

3 Methodology

In this chapter we lay out the methodological framework used in this thesis. We begin by explaining the data gathering and preprocessing phase. An introduction to the Refinitiv ESG scoring methodology follows were details on the four ESG variables are provided. We then move to an explanation of the portfolio sorting and backtesting procedure before we end the chapter with a brief discussion of some of the limitations of the methodological framework.

3.1 Data

The data was sourced from the Refinitiv Datastream platform[1]. Refinitiv is a global provider of financial market data and infrastructure. The Datastream platform is Refinitiv's historical financial database which features 70 years of data across 175 countries. The covered ESG universe currently consists of around 7300 companies.

The period of analysis in this thesis spans 17 years. From December 2002 to December 2019. Monthly observations for five variables were sourced. These variables were environment pillar score, social pillar score, corporate governance score, ESG combined score and return data. These five variables were collected for all companies in the Refinitiv ESG universe. The data preprocessing consisted of two main steps. First, OTC companies were removed as we found inconsistencies and large gaps in their pricing data. Second, companies with no available data and less than 24 months of observations were also removed. We found that by performing these two steps we got a more complete data set with fewer gaps and inconsistencies that could potentially lead to faulty results in the analysis. The final data set that was used in the analysis consisted of 5917 companies across 11 economic sectors and 81 countries. Refer to appendices A1.1 and A1.2 for a detailed breakdown of the economic sectors and countries.

[1] https://www.refinitiv.com/en/products/datastream-macroeconomic-analysis

3.2 Refinitiv ESG scoring methodology

A vast increase in research on ESG topics and integration of ESG ratings into financial management in recent years, have triggered a necessity to address the quality and reliability of the major ESG ratings providers. Thomson Reuters Refinitiv ESG is one of the key ESG rating providers in the financial industry. It has been used to provide ESG ratings in more than 1,200 published articles (Berg et al., 2021). It has also been identified by an OECD report as one of three major ESG data providers (Boffo and Patalano, 2020). Refinitiv use company-reported data in their ESG scoring methodology. This data is predominantly sourced from annual reports, company websites, CSR-reports and stock listing filings. The data is then standardized by research analysts and algorithms. The score model is compromised of two overall scores:

1. **ESG score**
   Measures the firm's ESG performance based on verifiable reported data publicly available
2. **ESG combined score**
   Overlays the ESG score with ESG controversies to deliver a more comprehensive evaluation of the company's sustainability impact and conduct over time

Roughly 500 ESG-metrics are grouped into 10 different categories that belong to either the environmental-, social- or governmental pillar of ESG. The environmental portion of the score are divided into the categories resource use, emissions and innovation. The social pillar consists of matters related to the workforce, human rights, the firms relation and behavior towards its community and product responsibility. The governmental part of the score consists of the categories management, shareholders and corporate social responsibility (CSR) reporting. These 10 different categories are given a weighting that is dependent on and differs from which industry they are part of and thus produce an industry adjusted ESG-score. A controversy overlay measure is added to form an ESG combined score, which discloses whether the firm has been involved in any controversies. The controversy score is adjusted for the magnitude and number of possible controversies and the size of the firm. Together with the ESG-score, this produces an ESG combined score between 0 and 1. Firms are then ranked by their score, forming percentiles which give an ESG combined rating between 0 and 100. For an overview of the descriptions of the different ESG scores, refer to appendix A1.3

3.3 Portfolio sorting and backtesting

A univariate portfolio sort and backtest for each of the four ESG variables was performed. For each month we divided all available stocks into four portfolios (quantiles). Quantile 1 (Q1) being stocks with the lowest ESG scores and quantile 4 (Q4) being stocks with the highest ESG scores. We paired each stocks ESG score for that month with its one month forward return[2]. The idea being that an investor invests at time t and then rebalances and measures the performance at $t + 1$. We iterated over all 204 months and obtained a measure of each quantiles out of sample performance. This methodology was performed with three different setups:

1. A portfolio sort and backtest on all 5918 global stocks to investigate the popularity hypothesis on a global scale. We ran this on all four ESG variables.
2. A portfolio sort and backtest with 5905 global stocks placed into their respective economic sectors to measure differences in popularity across sectors. The data set includes companies from 11 economic sectors but since the Academic & Educational Services sector only had 13 companies in it we chose to exclude it from this part of the analysis. This part was ran on only the ESG Combined Score variable.
3. Lastly, a portfolio sort and backtest with 3942 global stocks divided by country. There are a total of 81 countries in the data set. For this portfolio sort and backtest we chose to only include the top ten countries as they have the largest sample of companies. Table A1.2 in the appendix shows the countries included for this part of the analysis. As with the economic sectors, this part was also ran on only the ESG Combined Score variable.

One aspect of the whole data set that needs to be noted is that it includes missing values. This means that not every stock has a full 204 month history. Either ESG or return data might be missing. By doing it this way we have created a dynamically changing data set that more realistically reflects real life. The portfolio sorting and backtesting

[2]The first month of ESG data starts 31.12.2002 while the first month of the return data starts 31.01.2003. Likewise, the last month of the ESG data is 29.11.2019 while the last month of the return data is 31.12.2019

will only include stocks with complete data for that given month in its calculations[3]. An alternative way of doing it would have been to find a time frame were all stocks in the full ESG universe have complete data. This would lead to a smaller sample of stocks over a smaller time frame than what we have here. The data set would also be static, including the same stocks in the beginning as in the end.

3.4 Limitations

Currently there exists no standardized form of measuring company ESG data. As a result, the ESG score will heavily depend on the data provider used. Several articles that mark that the complexity of measuring ESG performance of firms can lead to significantly different results based on which data provider is used (Galema et al., 2008; Gibson et al., 2018; Berg et al., 2021). One way to deal with this is to use multiple data providers. When using ESG data in investing, it is important to address the materiality and validity of the data. Materiality is the degree to which the precise form of a metric reflects insight about investment-relevant questions the are useful. Validity is the degree to which the data is accurate in what they are measuring. Since ESG data to a large degree consists of somewhat subjective data that are self reported from the companies, ESG data often might have issues with validity, but still have a high degree of materiality. (CFA Institute, 2020). Further, it is important to keep in mind that Refinitiv have an ESG score update cycle of as low as one week if new company data became available (Refinitiv, 2020). As an extension of the last point, ESG scores remain unlocked going back five years. This means that ESG data from 2016 to 2019 in our data set can still be changed if new company information were to surface. This rather obviously has implications for the analysis as the time frame chosen becomes very important. We could have capped the data set from 2003 to 2016 but that would have excluded the last three years were the company ESG coverage has been the richest. This further raises reliability issues, questioning the results obtained. If major changes are made to the ESG data, any future attempt to reproduce the results of this thesis will surely fail.

[3]This means that early on in the data set, when ESG coverage was low, the quantiles calculated would have less stocks in them as opposed to later on in the data set.

4 Results

We find that all four ESG variables show consistency with the popularity framework. Table 4.1 summarizes our findings for the four ESG variables. The least popular quantile isn't always the best performing, as evident by the social and governance variables. However the most popular quantile is consistently the worst performing in terms of both raw returns and risk adjusted returns. Examining the standard deviation, a measure of total risk, we see that the most popular quantile is consistently lower than the rest of the quantiles across all variables, suggesting that stocks in the most popular quantile have lower overall volatility. In addition to all four variables showing consistency with popularity, figure 4.1 also illustrates that all four variables, to a large extent, show consistency with the risk and return paradigm.

Table 4.1: Annualized Performance Metrics (All global stocks, 2003-2019)

Characteristic	Performance	Least Popular Q1	Q2	Q3	Most Popular Q4	Consistent with Popularity
Environmental	Return	0.124	0.115	0.110	0.088	Yes
	Std Dev	0.150	0.148	0.146	0.143	
	Sharpe Ratio	0.827	0.778	0.755	0.617	
Social	Return	0.127	0.134	0.121	0.098	Yes
	Std Dev	0.151	0.154	0.147	0.141	
	Sharpe Ratio	0.837	0.866	0.822	0.690	
Governance	Return	0.131	0.132	0.113	0.103	Yes
	Std Dev	0.148	0.153	0.146	0.143	
	Sharpe Ratio	0.883	0.861	0.775	0.720	
ESG Combined	Return	0.137	0.134	0.109	0.099	Yes
	Std Dev	0.150	0.152	0.150	0.140	
	Sharpe Ratio	0.911	0.881	0.729	0.702	

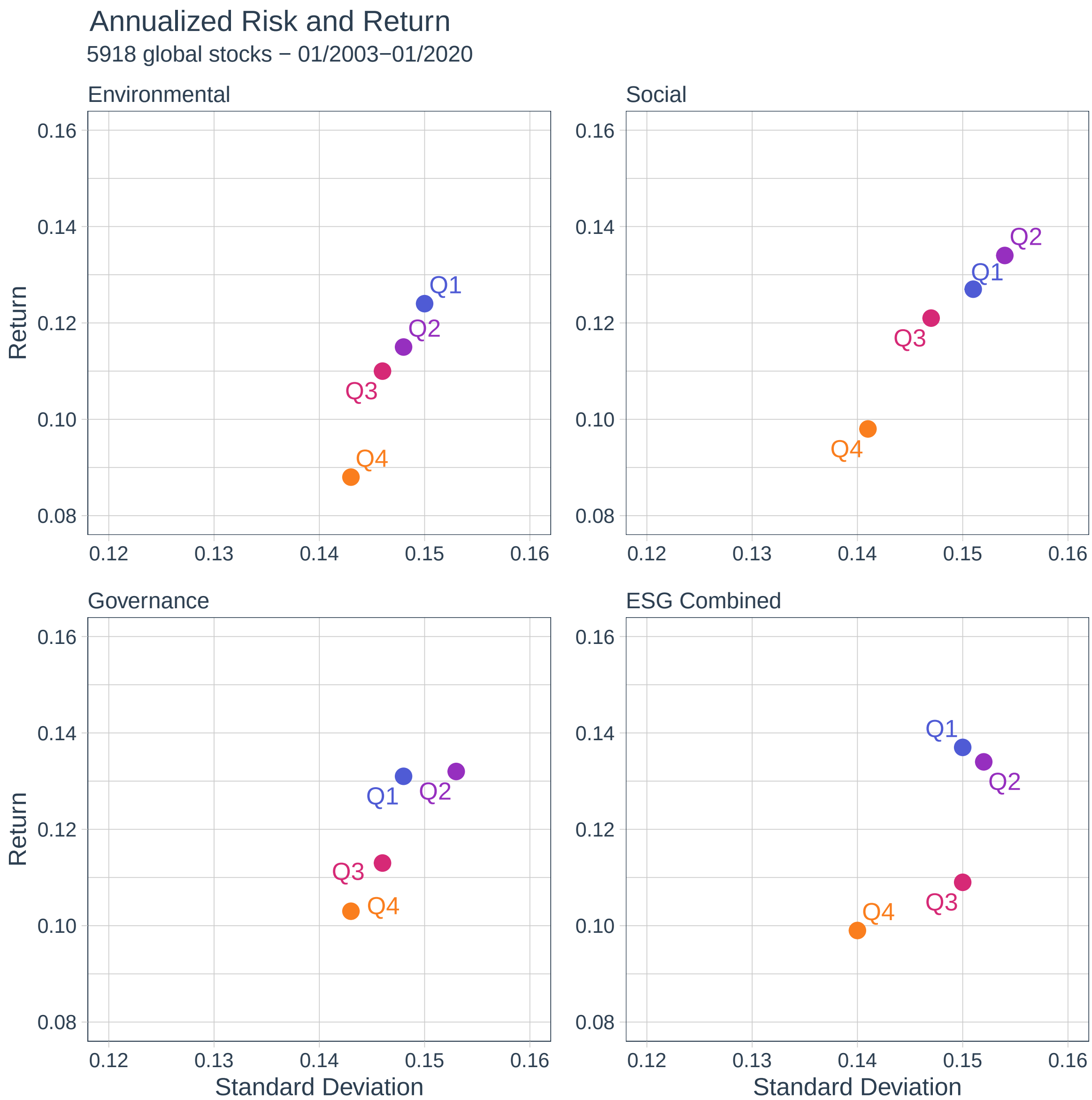

Figure 4.1: Annualized Risk and Return for all 5917 global stocks, 01/2003-01/2020

Moving on to the sector specific results we find that almost all sectors show consistency with the popularity framework. Table 4.2 summarizes the results. One exception worth noting is the financial sector, which to some degree shows mixed consistency. The second least popular quantile, Q2, has a considerably higher risk and return compared to the other three quantiles. Adjusting for risk puts the least popular quantile on top. Overall, the most evident consistency with the popularity framework is found in the technology, energy and healthcare sectors. Also worth noting is that the consumer non-cyclical and utilities sector have the lowest volatility across all quantiles compared to the rest of the sectors. This is inline with empirical studies as these two sectors tend to be stable over time and less prone to follow economic trends. Figure 4.2 shows the annualized risk and return of all four quantiles by sector. We see a pattern in that Q4 are lower and further to the left than Q1.

Table 4.2: Annualized Performance Metrics by economic sector (2003-2019)

Economic Sector	Performance	Least Popular Q1	Q2	Q3	Most Popular Q4	Consistent with Popularity
Industrials	Return	0.130	0.128	0.127	0.113	Yes
	Std Dev	0.161	0.169	0.166	0.164	
	Sharpe Ratio	0.806	0.759	0.769	0.687	
Consumer Cyclicals	Return	0.146	0.132	0.116	0.104	Yes
	Std Dev	0.164	0.177	0.181	0.170	
	Sharpe Ratio	0.886	0.744	0.641	0.612	
Financials	Return	0.095	0.169	0.095	0.075	Mixed
	Std Dev	0.147	0.279	0.172	0.176	
	Sharpe Ratio	0.647	0.604	0.550	0.424	
Technology	Return	0.146	0.146	0.135	0.096	Yes
	Std Dev	0.164	0.169	0.167	0.149	
	Sharpe Ratio	0.894	0.862	0.812	0.643	
Basic Materials	Return	0.171	0.138	0.098	0.126	Yes
	Std Dev	0.222	0.196	0.197	0.183	
	Sharpe Ratio	0.772	0.705	0.499	0.692	
Consumer Non-Cyclicals	Return	0.128	0.109	0.107	0.089	Yes
	Std Dev	0.119	0.101	0.108	0.098	
	Sharpe Ratio	1.069	1.079	0.986	0.909	
Real Estate	Return	0.115	0.101	0.110	0.098	Yes
	Std Dev	0.177	0.167	0.181	0.177	
	Sharpe Ratio	0.648	0.604	0.606	0.555	
Energy	Return	0.181	0.129	0.129	0.088	Yes
	Std Dev	0.260	0.225	0.210	0.185	
	Sharpe Ratio	0.693	0.573	0.615	0.477	
Healthcare	Return	0.197	0.156	0.137	0.107	Yes
	Std Dev	0.154	0.131	0.121	0.116	
	Sharpe Ratio	1.278	1.191	1.140	0.924	
Utilities	Return	0.093	0.081	0.071	0.073	Yes
	Std Dev	0.117	0.107	0.100	0.101	
	Sharpe Ratio	0.798	0.752	0.711	0.729	

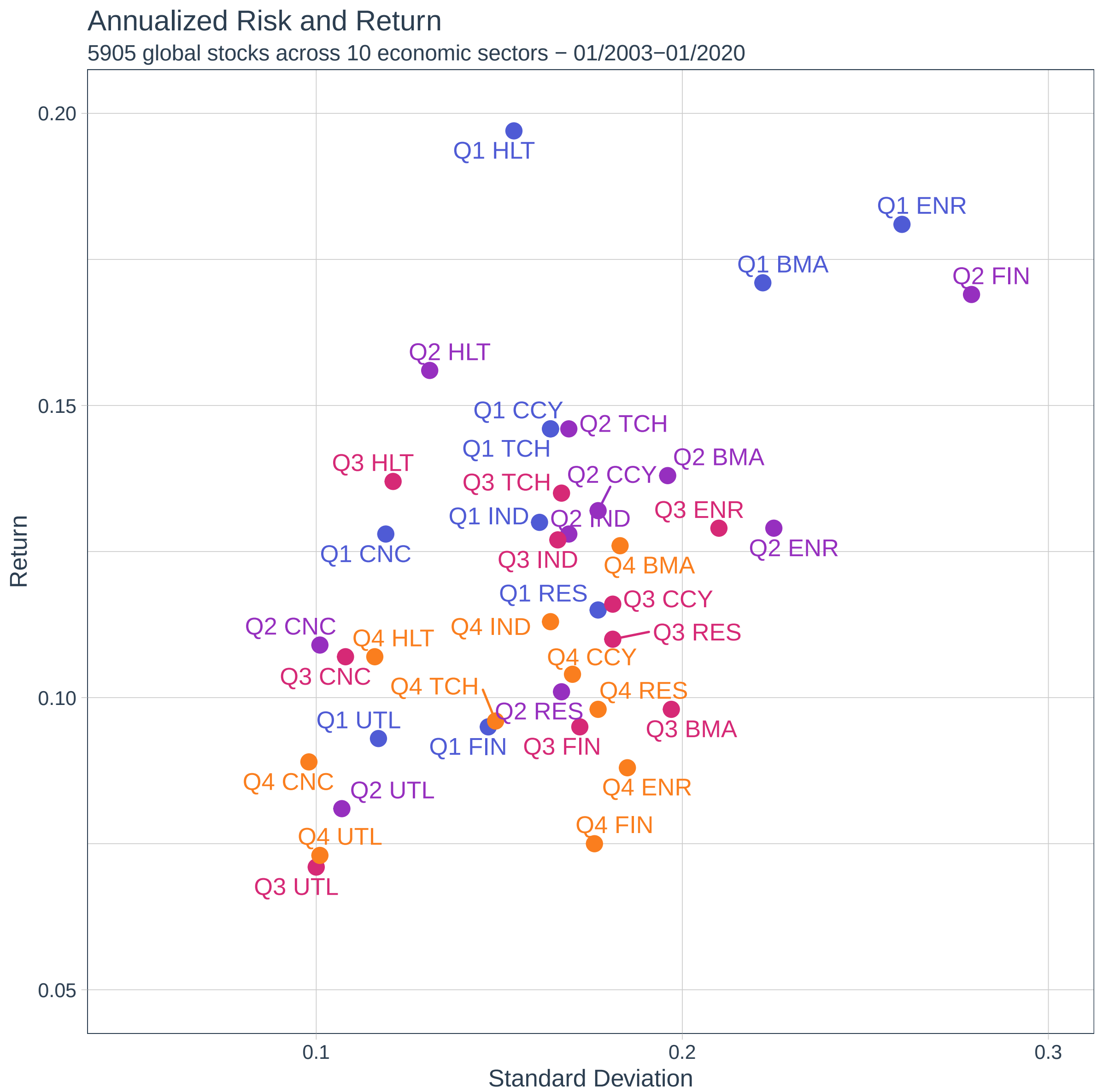

Figure 4.2: Annualized Risk and Return for 5905 global stocks across 10 economic sectors, 01/2003-01/2020

Lastly we have the country specific results. Table 4.3 summarizes the results. Some countries are more in line with the popularity framework than others. Most show similar numbers as we saw with the economic sectors where the least popular quantile outperforms the most popular. However there are a few exceptions that show mixed or no consistency with the popularity framework. The most notable deviation is China with an almost 7 percentage points higher absolute return in Q4 compared to Q1. The deviation is still significant after adjusting for risk. Figure 4.3 graphically shows a clustering of Q4 and Q3 in the low left corner, representing lower return and standard deviation.

Table 4.3: Annualized Performance Metrics by country (2003-2019)

Country	Performance	Least Popular Q1	Q2	Q3	Most Popular Q4	Consistent with Popularity
United States	Return	0.157	0.140	0.125	0.118	Yes
	Std Dev	0.180	0.181	0.170	0.160	
	Sharpe Ratio	0.874	0.774	0.736	0.737	
Japan	Return	0.097	0.095	0.103	0.090	Mixed
	Std Dev	0.168	0.177	0.194	0.187	
	Sharpe Ratio	0.577	0.535	0.532	0.483	
United Kingdom	Return	0.176	0.109	0.102	0.102	Yes
	Std Dev	0.171	0.164	0.161	0.146	
	Sharpe Ratio	1.030	0.664	0.631	0.699	
China	Return	0.137	0.180	0.156	0.202	No
	Std Dev	0.269	0.313	0.286	0.282	
	Sharpe Ratio	0.510	0.575	0.545	0.716	
Australia	Return	0.205	0.113	0.098	0.073	Yes
	Std Dev	0.245	0.176	0.154	0.133	
	Sharpe Ratio	0.838	0.640	0.637	0.548	
Canada	Return	0.119	0.097	0.105	0.117	No
	Std Dev	0.185	0.164	0.145	0.146	
	Sharpe Ratio	0.644	0.587	0.729	0.800	
Germany	Return	0.144	0.182	0.079	0.115	Yes
	Std Dev	0.201	0.309	0.201	0.184	
	Sharpe Ratio	0.718	0.590	0.393	0.625	
France	Return	0.122	0.078	0.076	0.081	Yes
	Std Dev	0.178	0.179	0.179	0.188	
	Sharpe Ratio	0.684	0.434	0.426	0.429	
Hong Kong	Return	0.195	0.163	0.128	0.119	Yes
	Std Dev	0.246	0.261	0.253	0.196	
	Sharpe Ratio	0.790	0.623	0.506	0.606	
Sweden	Return	0.189	0.119	0.134	0.134	Yes
	Std Dev	0.215	0.190	0.176	0.190	
	Sharpe Ratio	0.883	0.623	0.761	0.703	

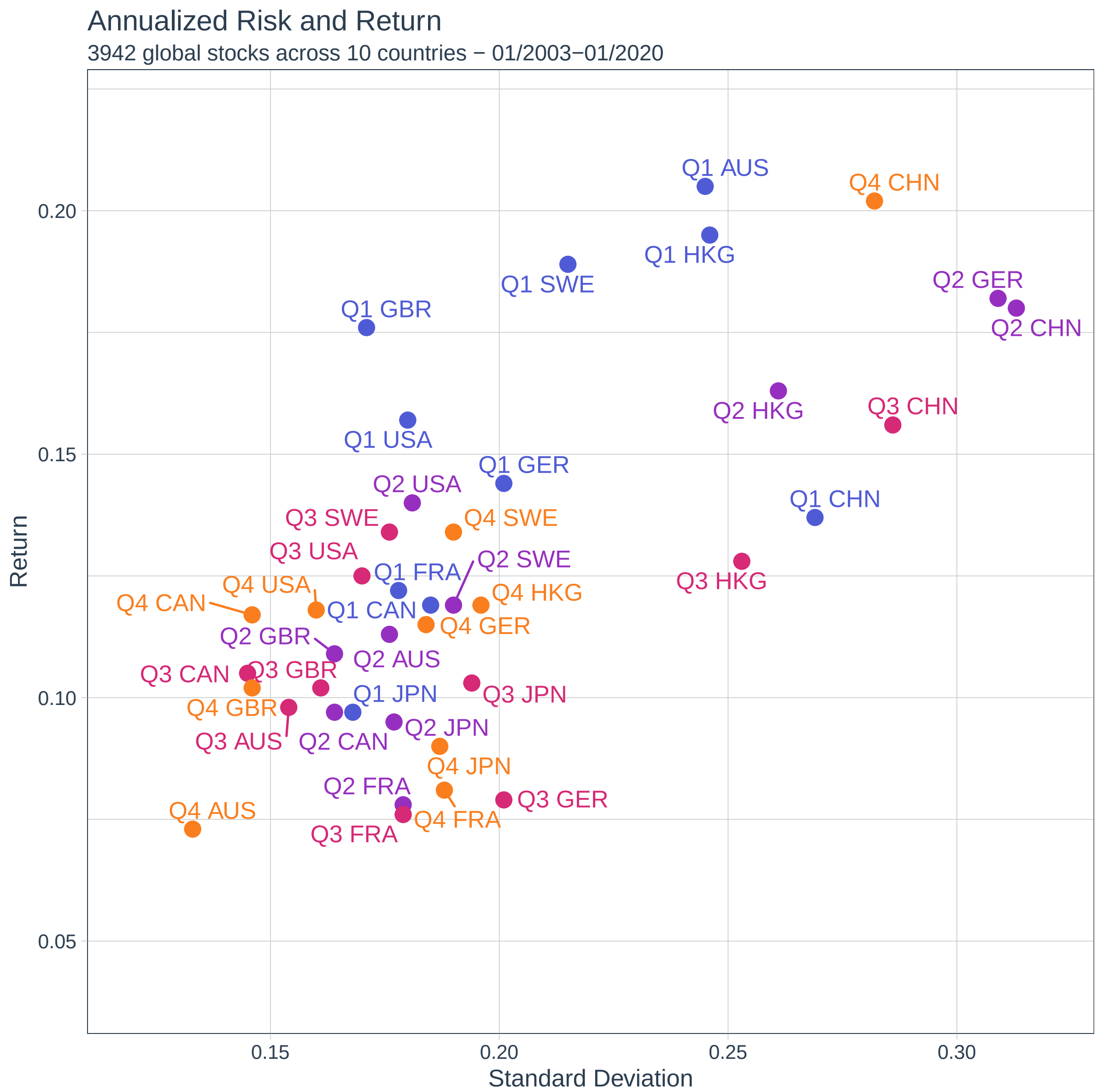

Figure 4.3: Annualized Risk and Return for 3942 global stocks across 10 countries, 01/2003-01/2020

5 Discussion

Ibbotson et al. (2018) demonstrated that the positive relationship between risk and return, put forth by classical finance, was flawed. More risk is not neccessarily linked to more return. The four main ESG variables in table 4.1 show a clear positive link between risk and return. However, looking at the economic sector results shown in table 4.2 we find a mix across sectors. The same can be observed if we look at the country specific results in table 4.3. This points to a potential diversification effect (Markowitz, 1952). Investing in a specific country or economic sector exposes the investor to specific country or sector risk. Investing in the whole set of global stocks spreads the risk out across multiple countries and economic sectors. This is evident if we compare the standard deviations found in table 4.2 and 4.3 with the standard deviation of the ESG Combined variable in table 4.1. A somewhat surprising realization was to see that the degree to which the spread between the returns across the four quantiles, was almost uniformly larger then the spread in standard deviation. The results show an almost uniform increase in returns from Q1 to Q4. Such a systematic pattern was unexpected, but strenghtens our results in confirming the consistency between avoiding the highest rated ESG companies and the popularity hypothesis. Camilleri (2015) indicates that, at least in Europe, ESG reporting and disclousure is primarily aimed more at larger businesses than small to medium businesses. This might suggest that some of the difference between the quantiles we see in our analysis, can be explained with the size factor described by Fama and French (1993). They found that stocks of smaller firms tended to outperform stocks of larger firms.

It may also be the case that some of the countries have companies with a high ESG score being a larger part of their economy. This could be an explanation why China shows no consistency with popularity at all. Perhaps ESG might be a less preferred characteristic here, than in western markets.

The results of our analysis are subject to limitations, mostly in the form of the reliability and validity of ESG data, as described in section 3.4. The applicability of our results comes from acknowledging that solely investing in stocks of high ESG-aware companies historically has come at a cost. It also serves the purpose of quantitatively confirming a consistency between the popularity hypothesis and ESG investing. As Ibbotson et al. (2018) states, their framework is still developing in terms of applicability to investing and investment strategies. Our result contribute to this, by validating this relationship. The difficulty of implementing popularity into an investment strategy can serve as a critique to the authors. There have also been critical voices that claim that their whole idea of popularity is obvious and too simple to work, but the authors answer this by referring to their empirical evidence which supports their view.

Addressing the popularity hypothesis in regards to ESG in the way we have done in our analysis and using it as a part of a investment strategy, also has limitations compared to the simplistic and convenient CAPM approach. First, data gathering is much more tedious and the data is much more subject to validity issues stemming from the subjective nature of ESG data. It demands more effort, time and resources than just calculating the historical beta from the market data, which is the CAPM way. There is also the decision on which data providers to use to gather the most reliable data possible. Averaging ESG scores from multiple sources could be a way to address this validity issue, but the cost of obtaining this data could potentially mean a negative net benefit overall.

Compared to previous research on the relationship between ESG and profitability, our results are in line with the part of the literature that claims that high scoring ESG stocks are outperformed by low scoring ESG stocks. If one where to think of our Q1 and Q4 (low scoring and high scoring, respectively) as portfolios, we have shown empirically that there exists a cost associated with investing in Q4 compared to Q1. The difference (cost) is substantial and in line with the results and conclusions from Adler and Kritzman (2008). Investing in the global ESG combined Q1 over our analysis horizon (2003-2019) would yield an annualized return of 4.4 percent more than investing in the Q4 (see table 4.1), with only a minor increase in volatility (from 14 to 15 percent). While dividing stocks into quantiles is meaningful for our analysis in testing the consistency with popularity and empirically showing an implicit cost of solely investing in high scoring ESG stock, one can reasonably argue that this approach should not be directly translated into an investment strategy. Ibbotson et al. (2018) warns that the "value trap" awaits those that follow the popularity hypothesis too blindly, i.e only investing in unpopular securities without any other supplemental analysis. In terms of our analysis, this would mean to blindly invest in the lowest scoring ESG stock with no regard for fundamental metrics. Examples of low ESG scoring companies include poorly govern companies and/or companies that exploit their employees, which one could plausibly argue is closely linked to profitability.

Such companies are naturally more subject to bankruptcy, which every investor would want to avoid being a part of. Blindly investing in low scoring ESG will not identify and exclude such securities in the portfolio and therefore posit a cost for the investor. This can explain why we see Q2 performing better in terms of raw return in both the social and governmental pillars and the higher Sharpe ratio in the social pillar in table 4.1. The rational for this being that this quantile avoids the most unpopular stocks (the most inadequately governed and socially ignorant firms) and thus lowers the overall default risk of the securities in the portfolio.

But why does a significant portion of the literature state that investing in securities of firms with good ESG records is profitable, in contrary to what the popularity predicts and what several articles show empirically? Kempf and Osthoff (2007) got positive results by going long in stocks of socially responsible firms and shorting the stocks of firms that acted socially irresponsible. This is another approach than in our analysis and what Ibbotson et al. (2018) did. While our results shows the connection between investing in low ESG scoring companies and superior returns in a descriptive way, we might get a different result if we had adopted this into an active strategy that both went long and short on the different securities. The extensive meta-study of Friede et al. (2015) found roughly 90 percent of the studies since the 1970's showed non-negative relationship between ESG and corporate financial performance (CFP). This is by far the largest study on ESG performance and company profitability done to date. They claim that there is a general perception of a neutral or mixed ESG/CFP relation, but that this view may be biased due to it being based on findings from portfolio studies. They find that there are clear evidence for the business case for ESG investing on the company level. The authors identify an ESG alpha, but a large group of investors fail to harvest this because of the costs of implementing ESG in their portfolios or the cost of fees for actively managed ESG funds. This somewhat contrast our results.

If the results from our analysis and the large body of the literature that claim that ESG investing comes at a cost is true, then those who are incorporating their environmental and social standards into their investing by seeking out the stocks of high scoring ESG companies, are either what the literature call the "willing"- or "unknowing losers". Some may prioritize high ESG stocks and avoid sin stocks because they mean that this will perform better financially in the future. As stated by Adler and Kritzman (2008) these are not engaging in impact investing, but active management. The portion of both professional and retail investors that screen for ESG are primarily driven by the notion of making an impact and recognize that this might come at a cost by lower returns, are those who do impact investing. This is not in line with the classical idea of a rational investor that only cares about risk and return. This obviously has more "behavioral" aspects to it. Where the utility in classical finance is strictly in the return/risk dimension, we argue utility in a microeconomic sense could and maybe should include all the utility an investor gets from their investment. From this perspective it would be rational for the ESG minded investor to forgo some financial return if this maximizes his or hers utility. These investors get non-monetary return by avoiding sin stocks and promoting ESG stocks. One could argue that these investors could get higher net return on their investments if they would evaluate the performance by quantifying the non-monetary return such as proposed by Cooper et al. (2016) and Pedersen et al. (2020).

As environmental and social challenges keep dominating the agenda today and in the future, one reason for the popularity in investing in high ESG stocks might stem from the loss aversion concept from Kahnemann and Tversky (1979). As the global community tries to find solutions to these challenges, investors may fear bountiful losses if when investing in low ESG stocks, as regulations, technology advances and declining demand for the products of these companies may drive low ESG companies out of business. In terms of loss aversion, they would prefer to forgo some return and avoid large losses. This contributes to the notion of ESG investing being popular, and as our results show and the popularity framework suggest, there might be gains to be made for investors that do not care for such characteristics. The ethical side of this approach can although be questioned.

As the world moves forward, so will probably ESG-awareness by both companies and investors. What is excellent ESG practice now, could well be the standard in a few years. One could argue that there always will be companies that lags behind regarding ESG, and if ESG preferences continue to be popular in the future, we argue that implications of the results in our thesis will hold. We have showed that the popularity framework is consistent with ESG investing, in that only investing in the highest scoring ESG companies comes at a cost because of the high demand for this characteristic and the subsequent lower yield on the investment. Whether the investor chooses to take advantage of this and rather invest in the firms with a "lagging" ESG records or not, comes down to the degree of rationality of their investment behaviour, their utility function and ability to maximize it.

6 Conclusion

Our thesis shows that investing in ESG comes at a premium that is not directly related to risk, and therefore confirming our research question that investing in securities with a low ESG score is consistent with the popularity hypothesis put forth by Ibbotson et al. (2018). The results were systematic on the whole data set, but showed some varying consistency when comparing sectors and countries. The implications of these findings is that investors should be aware of the possibility of overpaying for securities of the companies with the highest ESG records. Our analysis shows notable differences in degree of consistency over the different sectors and countries.

Utilizing ESG data in our analysis, our findings are somewhat sensitive to the qualitative nature of these types of subjective company self-reported data, but this is the best way to get ESG data into a quantitative analysis and is the industry standard when evaluating the ESG performance of both companies, portfolios and funds.

Our empirical evidence is in line with what a large body of the literature suggests, that investing in the securities of high ESG rated firms comes at a cost in the form of lower returns. We also put this in context on the motivation and preference that both professional and retail investors have for investing in ESG, which ultimately forms the popularity, high demand and lower yields for these types of securities.

We have contributed to validate the consistency of ESG investing and the popularity framework and possibly answered why this consistency exists. Future research should address how to best apply these insights into real world scenarios so that one may achieve superior returns.

References

Adler, T. and Kritzman, M. (2008). The cost of socially responsible investing. *Journal of Portfolio Management*, 35(1):52–56.

Berg, F., Fabisik, K., and Sautner, Z. (2021). Rewriting History II: The (Un)Predictable Past of ESG Ratings.

Blitz, D. and Fabozzi, F. (2017). Sin Stocks Revisited: Resolving the Sin Stock Anomaly. *Journal of Portfolio Management*, pages 105–111.

Blitz, D. and Swinkels, L. (2020). Is exclusion effective? *Journal of Portfolio Management*, 46(3):42–48.

Boffo, R. and Patalano, R. (2020). ESG Investing: Practices, Progress and Challenges. Technical report, OECD Paris.

Broadstock, D. C., Chan, K., Cheng, L. T., and Wang, X. (2021). The role of ESG performance during times of financial crisis: Evidence from COVID-19 in China. *Finance Research Letters*, 38:101716.

Camilleri, M. A. (2015). Environmental, social and governance disclosures in Europe. *Sustainability Accounting, Management and Policy Journal*, 6(2):224–242.

Camilleri, M. A. (2020). The market for socially responsible investing: a review of the developments. *Social Responsibility Journal*, 17(3):412–428.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1):57–82.

CFA Institute (2020). Future of Sustainability in Investment Management. Technical report, CFA Institute.

Cooper, L., Evnine, J., Finkelman, J., Huntington, K., and Lynch, D. (2016). Social Finance and the postmodern portfolio: Theory and practice. *Journal of Wealth Management*, 18(4):9–21.

De Bondt, W. F. M. and Thaler, R. (1985). Does the Stock Market Overreact? *The Journal of Finance*, 40(3):793–805.

Fama, E. F. (1970). Efficient Market Hypothesis: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2):383–417.

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

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

Feng, G., Giglio, S., and Xiu, D. (2020). Taming the Factor Zoo: A Test of New Factors. *Journal of Finance*, 75(3):1327–1370.

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

Galema, R., Plantinga, A., and Scholtens, B. (2008). The stocks at stake: Return and risk in socially responsible investment. *Journal of Banking and Finance*, 32(12):2646–2654.

Gibson, R., Krueger, P., and Schmidt, P. S. (2018). ESG Rating Disagreement and Stock Returns. *Swiss Finance Institute Research Paper Series*, 19(67).

Graham, B. and Dodd, D. L. (1934). *Security Analysis*. McGraw-Hill, New York, 2nd editio edition.

Grossman, S. J. and Stiglitz, J. E. (1980). On the Impossibility of Informationally Efficient Markets. *The American Economic Review*, 70(3):393–408.

Harrison, J. M. and Kreps, D. M. (1978). Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations. *The Quarterly Journal of Economics*, 92(2):323–336.

Ibbotson, R. G. and Idzorek, T. M. (2014). Dimensions of Popularity. *Journal of Portfolio Management*, 40(5):68–74.

Ibbotson, R. G., Idzorek, T. M., Kaplan, P. D., and Qiong, J. X. (2018). *Popularity - A bridge between Classical and Behavioral Finance*. CFA Institute Research Foundation.

Idzorek, T. M. and Ibbotson, R. G. (2017). Popularity and Asset Pricing. *The Journal of Investing*, 26(1):46–56.

Kahnemann, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263–292.

Kempf, A. and Osthoff, P. W. (2007). The effect of socially responsible investing on portfolio performance.

Langbein, J. H. and Posner, R. A. (1980). Social Investing and the Law of Trust. *Michigan Law Review*, 79(1):72–112.

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, 47(1):13–37.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1):77–91.

Mehra, R. and Prescott, E. C. (1985). The Equity Premium: A Puzzle. *Journal of Monetary Economics*, 15(2):145–161.

Miller, M. H. and Modigliani, F. (1961). Dividend Policy, Growth and the Valuaiton of Shares. *The Journal of Business*, 34(4):411–433.

Pedersen, L. H., Fitzgibbons, S., and Pomorski, L. (2020). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics*, xxx(xxxx).

Refinitiv (2020). Environmental, Social and Governance (ESG) Scores from Refinitiv.

Sharpe, W. (1964). Capital Asset Prices: A Theory Of Market Equilibrium Under Conditions Of Risk. *The Journal of Finance*, 19(3):425–442.

Shefrin, H. and Statman, M. (1994). Behavioral Capital Asset Pricing Theory. *The Journal of Financial and Quantitative Analysis*, 29(3):323–349.

Silva, F. and Cortez, M. C. (2016). The performance of US and European green funds in different market conditions. *Journal of Cleaner Production*, 135:558–566.

Titman, S. and Jegadeesh, N. (1993). Returns to Buying Winners and Selling Losers : Implications for Stock Market Efficiency. *The Journal of Finance*, 48(1):65–91.

Tversky, A. and Kahnemann, D. (1974). Judgement Under Uncertainty: Heuristics and Biases. *Science*, 185(4157):1124–1131.

Vendrame, V., Tucker, J., and Guermat, C. (2016). Some extensions of the CAPM for individual assets. *International Review of Financial Analysis*, 44:78–85.

Appendix

A1

Table A1.1: Companies by economic sector

Economic Sector	Companies
Industrials	963
Consumer Cyclicals	849
Financials	841
Technology	674
Basic Materials	618
Consumer Non-Cyclicals	472
Real Estate	463
Energy	391
Healthcare	360
Utilities	273
Academic & Educational Services	13
Total	5917

Table A1.2: Top ten countries and the number of companies

Country	Companies
United States	1631
Japan	414
United Kingdom	393
China	378
Australia	297
Canada	265
Germany	164
France	148
Hong Kong	129
Sweden	123
Total	3942

Table A1.3: ESG score card (adapted from Refinitiv (2020))

Score range	Grade	Description
$0.000 \leq \text{score} \leq 0.083$	D-	Poor relative ESG performance and insufficient
$0.083 < \text{score} \leq 0.166$	D	degree of transparency in reporting material ESG
$0.166 < \text{score} \leq 0.250$	D+	data publicly.
$0.250 < \text{score} \leq 0.333$	C-	Satisfactory relative ESG performance and
$0.333 < \text{score} \leq 0.416$	C	moderate degree of transparency in reporting
$0.416 < \text{score} \leq 0.500$	C+	material ESG data publicly.
$0.500 < \text{score} \leq 0.583$	B-	Good relative ESG performance and above-
$0.583 < \text{score} \leq 0.666$	B	average degree of transparency in reporting
$0.666 < \text{score} \leq 0.750$	B+	material ESG data publicly.
$0.750 < \text{score} \leq 0.833$	A-	Excellent relative ESG performance and high
$0.833 < \text{score} \leq 0.916$	A	degree of transparency in reporting material ESG
$0.916 < \text{score} \leq 1.000$	A+	data publicly.

A2

R Package References

Dickerson, Alex; Spohnholtz, Jonathan (2018). portsort: Factor-Based Portfolio Sorts. R package version 0.1.0. https://cran.r-project.org/package=portsort

Peterson, Brian G.; Carl, Peter; Boudt, Kris; Bennett, Ross; Ulrich, Joshua; Zivot, Eric; Cornilly, Dries; Hung, Eric; Lestel, Matthieu; Balkissoon, Kyle; Wuertz, Diethelm; Christidis, Anthony Alexander; Martin, R. Douglas; Zhou, Zeheng 'Zenith'; Shea, Justin M. (2020). PerformanceAnalytics: Econometric Tools for Performance and Risk Analysis. R package 2.0.4. https://cran.r-project.org/package=PerformanceAnalytics

Ryan, Jeffrey A.; Ulrich, Joshua M.; Bennett, Ross; Joy, Corwin (2020). xts: eXtensible Time Series. R package version 0.12.1. https://cran.r-project.org/package=xts

A3

Table A3.1: Summary Statistics for ESG variables, monthly observations

	Environmental				Social			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Measure								
Observations	204	204	204	204	204	204	204	204
NAs	0	0	0	0	0	0	0	0
Minimum	-0,221	-0,212	-0,214	-0,194	-0,225	-0,22	-0,203	-0,195
Quartile 1	-0,007	-0,009	-0,009	-0,012	-0,008	-0,009	-0,01	-0,011
Median	0,012	0,014	0,012	0,011	0,013	0,015	0,013	0,011
Arithmetic Mean	0,01	0,01	0,009	0,007	0,011	0,011	0,01	0,008
Geometric Mean	0,009	0,009	0,008	0,007	0,01	0,01	0,009	0,007
Quartile 3	0,037	0,036	0,033	0,03	0,036	0,037	0,032	0,03
Maximum	0,168	0,197	0,181	0,169	0,17	0,189	0,183	0,173
SE Mean	0,003	0,003	0,003	0,003	0,003	0,003	0,003	0,003
LCL Mean (0,95)	0,004	0,004	0,003	0,002	0,005	0,005	0,004	0,002
UCL Mean (0,95)	0,016	0,016	0,015	0,013	0,017	0,017	0,016	0,014
Variance	0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,002
Stdev	0,043	0,043	0,042	0,041	0,044	0,045	0,042	0,041
Skewness	-0,799	-0,653	-0,771	-0,663	-0,878	-0,617	-0,624	-0,64
Kurtosis	4,433	4,827	4,561	3,26	4,586	4,393	3,859	3,657
	Governance				**ESG Combined**			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Observations	204	204	204	204	204	204	204	204
NAs	0	0	0	0	0	0	0	0
Minimum	-0,209	-0,219	-0,209	-0,206	-0,22	-0,21	-0,216	-0,197
Quartile 1	-0,009	-0,01	-0,009	-0,008	-0,007	-0,009	-0,01	-0,01
Median	0,014	0,013	0,012	0,013	0,014	0,014	0,013	0,012
Arithmetic Mean	0,011	0,011	0,009	0,009	0,011	0,011	0,009	0,008
Geometric Mean	0,01	0,01	0,009	0,008	0,01	0,01	0,008	0,007
Quartile 3	0,036	0,034	0,033	0,03	0,038	0,035	0,032	0,031
Maximum	0,187	0,181	0,17	0,178	0,174	0,185	0,194	0,162
SE Mean	0,003	0,003	0,003	0,003	0,003	0,003	0,003	0,003
LCL Mean (0,95)	0,005	0,005	0,004	0,003	0,005	0,005	0,003	0,003
UCL Mean (0,95)	0,017	0,017	0,015	0,014	0,017	0,017	0,015	0,014
Variance	0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,002
Stdev	0,043	0,044	0,042	0,041	0,043	0,044	0,043	0,04
Skewness	-0,702	-0,651	-0,769	-0,687	-0,842	-0,547	-0,674	-0,712
Kurtosis	4,24	4,256	4,068	4,302	4,334	4,05	4,574	3,648

Table A3.2: Summary Statistics for ESG Combined by economic sector

	Industrials				Consumer Cyclicals			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Measure								
Observations	204	204	204	204	204	204	204	204
NAs	0	0	0	0	0	0	0	0
Minimum	-0.223	-0.220	-0.218	-0.217	-0.216	-0.219	-0.235	-0.206
Quartile 1	-0.011	-0.012	-0.010	-0.011	-0.011	-0.014	-0.016	-0.014
Median	0.014	0.014	0.013	0.013	0.015	0.013	0.014	0.012
Arithmetic Mean	0.011	0.011	0.011	0.009	0.012	0.011	0.010	0.009
Geometric Mean	0.010	0.009	0.009	0.008	0.011	0.010	0.008	0.007
Quartile 3	0.039	0.038	0.038	0.036	0.038	0.038	0.037	0.034
Maximum	0.182	0.210	0.199	0.163	0.242	0.309	0.284	0.263
SE Mean	0.003	0.003	0.003	0.003	0.003	0.004	0.004	0.003
LCL Mean (0.95)	0.004	0.004	0.004	0.003	0.006	0.004	0.002	0.002
UCL Mean (0.95)	0.017	0.017	0.017	0.016	0.019	0.018	0.017	0.015
Variance	0.002	0.002	0.002	0.002	0.002	0.003	0.003	0.002
Stdev	0.046	0.049	0.048	0.047	0.047	0.051	0.052	0.049
Skewness	-0.749	-0.683	-0.565	-0.804	-0.319	0.381	0.082	0.106
Kurtosis	3.413	3.600	3.146	2.758	4.772	6.991	5.716	4.551

Table A3.2: (continued)

	Financials				Technology			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Observations	204	204	204	204	204	204	204	204
NAs	0	0	0	0	0	0	0	0
Minimum	-0.186	-0.196	-0.197	-0.223	-0.192	-0.192	-0.198	-0.185
Quartile 1	-0.013	-0.016	-0.020	-0.023	-0.011	-0.018	-0.012	-0.016
Median	0.012	0.015	0.012	0.010	0.015	0.018	0.015	0.010
Arithmetic Mean	0.008	0.014	0.008	0.006	0.012	0.012	0.011	0.008
Geometric Mean	0.007	0.012	0.007	0.005	0.011	0.011	0.010	0.007
Quartile 3	0.033	0.035	0.037	0.036	0.042	0.040	0.037	0.035
Maximum	0.170	0.939	0.211	0.237	0.152	0.148	0.204	0.112
SE Mean	0.003	0.006	0.003	0.004	0.003	0.003	0.003	0.003
LCL Mean (0.95)	0.002	0.003	0.001	-0.001	0.006	0.005	0.005	0.002
UCL Mean (0.95)	0.014	0.025	0.015	0.013	0.019	0.019	0.018	0.014
Variance	0.002	0.006	0.002	0.003	0.002	0.002	0.002	0.002
Stdev	0.042	0.081	0.050	0.051	0.047	0.049	0.048	0.043
Skewness	-0.359	7.309	-0.363	-0.135	-0.718	-0.582	-0.182	-0.620
Kurtosis	2.920	83.337	2.977	3.787	2.370	1.817	3.011	1.824

Table A3.2: (continued)

	Basic Materials				Consumer Non-Cyclicals			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Observations	204	204	204	204	204	204	204	204
NAs	0	0	0	0	0	0	0	0
Minimum	-0.338	-0.303	-0.285	-0.236	-0.170	-0.155	-0.153	-0.115
Quartile 1	-0.022	-0.023	-0.022	-0.019	-0.003	-0.003	-0.006	-0.006
Median	0.015	0.013	0.012	0.009	0.013	0.012	0.011	0.009
Arithmetic Mean	0.014	0.012	0.008	0.011	0.011	0.009	0.009	0.007
Geometric Mean	0.012	0.010	0.007	0.009	0.010	0.009	0.008	0.007
Quartile 3	0.056	0.043	0.042	0.044	0.030	0.024	0.027	0.027
Maximum	0.161	0.178	0.217	0.175	0.107	0.122	0.099	0.066
SE Mean	0.004	0.004	0.004	0.004	0.002	0.002	0.002	0.002
LCL Mean (0.95)	0.005	0.004	0.000	0.003	0.006	0.005	0.005	0.004
UCL Mean (0.95)	0.023	0.019	0.016	0.018	0.015	0.013	0.013	0.011
Variance	0.004	0.003	0.003	0.003	0.001	0.001	0.001	0.001
Stdev	0.064	0.057	0.057	0.053	0.034	0.029	0.031	0.028
Skewness	-0.853	-0.721	-0.755	-0.441	-0.994	-0.980	-0.940	-1.138
Kurtosis	3.898	4.345	4.281	2.321	3.524	5.438	3.486	2.695

Table A3.2: (continued)

	Real Estate				Energy			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Observations	204	204	204	204	204	204	204	204
NAs	0	0	0	0	0	0	0	0
Minimum	-0.254	-0.258	-0.297	-0.262	-0.235	-0.227	-0.235	-0.245
Quartile 1	-0.010	-0.012	-0.012	-0.010	-0.029	-0.027	-0.023	-0.023
Median	0.011	0.009	0.013	0.010	0.014	0.008	0.008	0.011
Arithmetic Mean	0.010	0.008	0.009	0.008	0.015	0.011	0.011	0.007
Geometric Mean	0.008	0.007	0.008	0.007	0.012	0.009	0.009	0.006
Quartile 3	0.034	0.035	0.032	0.033	0.070	0.054	0.051	0.042
Maximum	0.219	0.207	0.299	0.257	0.245	0.165	0.166	0.132
SE Mean	0.004	0.003	0.004	0.004	0.005	0.005	0.004	0.004
LCL Mean (0.95)	0.003	0.002	0.002	0.001	0.005	0.002	0.002	0.000
UCL Mean (0.95)	0.017	0.015	0.016	0.015	0.025	0.020	0.019	0.015
Variance	0.003	0.002	0.003	0.003	0.006	0.004	0.004	0.003
Stdev	0.051	0.048	0.052	0.051	0.075	0.065	0.061	0.053
Skewness	-0.437	-0.890	-0.500	-0.371	-0.092	-0.310	-0.442	-0.759
Kurtosis	5.429	5.787	9.965	6.132	0.622	0.716	1.129	2.313

Table A3.2: (continued)

	Healthcare				Utilities			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Observations	204	204	204	204	204	204	204	204
NAs	0	0	0	0	0	0	0	0
Minimum	-0.199	-0.156	-0.125	-0.154	-0.101	-0.118	-0.112	-0.088
Quartile 1	-0.004	-0.005	-0.005	-0.011	-0.008	-0.009	-0.009	-0.010
Median	0.022	0.017	0.016	0.015	0.010	0.009	0.009	0.011
Arithmetic Mean	0.016	0.013	0.011	0.009	0.008	0.007	0.006	0.006
Geometric Mean	0.015	0.012	0.011	0.008	0.007	0.006	0.006	0.006
Quartile 3	0.041	0.038	0.033	0.030	0.028	0.028	0.025	0.027
Maximum	0.155	0.117	0.096	0.080	0.116	0.087	0.083	0.080
SE Mean	0.003	0.003	0.002	0.002	0.002	0.002	0.002	0.002
LCL Mean (0.95)	0.010	0.008	0.007	0.004	0.003	0.002	0.002	0.002
UCL Mean (0.95)	0.023	0.018	0.016	0.014	0.012	0.011	0.010	0.010
Variance	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Stdev	0.044	0.038	0.035	0.033	0.034	0.031	0.029	0.029
Skewness	-0.920	-0.881	-0.811	-1.203	-0.377	-0.678	-0.570	-0.706
Kurtosis	3.091	2.422	1.060	3.030	1.295	1.683	1.485	0.821

Table A3.3: Summary Statistics for ESG Combined by country

	United States				Japan			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Measure								
Observations	204	204	204	204	204	204	204	204
NAs	0	0	0	0	0	0	0	0
Minimum	-0.245	-0.218	-0.211	-0.204	-0.161	-0.210	-0.247	-0.232
Quartile 1	-0.012	-0.015	-0.013	-0.011	-0.021	-0.022	-0.018	-0.014
Median	0.015	0.015	0.014	0.013	0.010	0.010	0.010	0.011
Arithmetic Mean	0.013	0.012	0.010	0.010	0.008	0.008	0.009	0.008
Geometric Mean	0.012	0.010	0.009	0.009	0.007	0.007	0.007	0.006
Quartile 3	0.046	0.043	0.036	0.035	0.041	0.039	0.043	0.041
Maximum	0.220	0.250	0.222	0.186	0.132	0.133	0.140	0.126
SE Mean	0.004	0.004	0.003	0.003	0.003	0.004	0.004	0.004
LCL Mean (0.95)	0.006	0.004	0.004	0.003	0.001	0.001	0.001	0.000
UCL Mean (0.95)	0.020	0.019	0.017	0.016	0.015	0.015	0.016	0.015
Variance	0.003	0.003	0.002	0.002	0.002	0.003	0.003	0.003
Stdev	0.052	0.052	0.049	0.046	0.049	0.051	0.056	0.054
Skewness	-0.538	-0.199	-0.429	-0.456	-0.242	-0.487	-0.553	-0.586
Kurtosis	3.387	3.422	3.405	3.089	0.133	1.082	1.662	1.529

Table A3.3: (continued)

	United Kingdom				China			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Observations	204	204	204	204	191	191	179	204
NAs	0	0	0	0	13	13	25	0
Minimum	-0.179	-0.202	-0.208	-0.172	-0.226	-0.304	-0.279	-0.268
Quartile 1	-0.013	-0.017	-0.018	-0.014	-0.026	-0.032	-0.025	-0.026
Median	0.014	0.013	0.012	0.011	0.015	0.011	0.018	0.014
Arithmetic Mean	0.015	0.009	0.008	0.009	0.011	0.015	0.013	0.017
Geometric Mean	0.013	0.008	0.007	0.008	0.008	0.011	0.010	0.014
Quartile 3	0.047	0.038	0.036	0.033	0.061	0.063	0.059	0.060
Maximum	0.256	0.220	0.212	0.206	0.229	0.359	0.248	0.286
SE Mean	0.003	0.003	0.003	0.003	0.006	0.007	0.006	0.006
LCL Mean (0.95)	0.008	0.003	0.002	0.003	0.000	0.002	0.001	0.006
UCL Mean (0.95)	0.021	0.016	0.015	0.014	0.023	0.028	0.025	0.028
Variance	0.002	0.002	0.002	0.002	0.006	0.008	0.007	0.007
Stdev	0.049	0.047	0.046	0.042	0.078	0.090	0.083	0.081
Skewness	0.105	-0.453	-0.390	-0.009	-0.267	0.099	-0.291	-0.027
Kurtosis	3.412	3.620	3.631	3.293	0.588	1.658	1.240	0.910

Table A3.3: (continued)

	Australia				Canada			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Observations	204	204	204	204	204	204	204	204
NAs	0	0	0	0	0	0	0	0
Minimum	-0.310	-0.269	-0.201	-0.152	-0.197	-0.217	-0.186	-0.188
Quartile 1	-0.023	-0.015	-0.017	-0.016	-0.022	-0.019	-0.016	-0.013
Median	0.017	0.011	0.015	0.012	0.010	0.008	0.012	0.009
Arithmetic Mean	0.017	0.009	0.008	0.006	0.010	0.008	0.009	0.010
Geometric Mean	0.015	0.008	0.007	0.005	0.009	0.007	0.008	0.009
Quartile 3	0.056	0.039	0.036	0.031	0.046	0.040	0.037	0.037
Maximum	0.266	0.144	0.106	0.086	0.166	0.166	0.147	0.131
SE Mean	0.005	0.004	0.003	0.003	0.004	0.003	0.003	0.003
LCL Mean (0.95)	0.007	0.002	0.002	0.001	0.003	0.001	0.003	0.004
UCL Mean (0.95)	0.027	0.016	0.014	0.011	0.017	0.015	0.015	0.016
Variance	0.005	0.003	0.002	0.001	0.003	0.002	0.002	0.002
Stdev	0.071	0.051	0.045	0.038	0.054	0.047	0.042	0.042
Skewness	-0.123	-0.871	-0.827	-0.725	-0.324	-0.545	-0.923	-0.483
Kurtosis	2.324	3.903	1.808	1.103	1.050	2.888	3.538	2.470

Table A3.3: (continued)

	Germany				France			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Observations	204	204	204	204	204	204	204	204
NAs	0	0	0	0	0	0	0	0
Minimum	-0.219	-0.239	-0.225	-0.191	-0.177	-0.207	-0.177	-0.175
Quartile 1	-0.018	-0.022	-0.025	-0.020	-0.017	-0.022	-0.023	-0.024
Median	0.017	0.013	0.008	0.012	0.012	0.008	0.007	0.007
Arithmetic Mean	0.012	0.015	0.007	0.010	0.010	0.006	0.006	0.007
Geometric Mean	0.010	0.012	0.005	0.008	0.009	0.005	0.005	0.005
Quartile 3	0.044	0.050	0.044	0.042	0.049	0.038	0.035	0.042
Maximum	0.273	0.871	0.268	0.213	0.170	0.182	0.196	0.198
SE Mean	0.004	0.006	0.004	0.004	0.004	0.004	0.004	0.004
LCL Mean (0.95)	0.004	0.003	-0.001	0.002	0.003	-0.001	-0.001	-0.001
UCL Mean (0.95)	0.020	0.028	0.015	0.017	0.017	0.014	0.013	0.014
Variance	0.003	0.008	0.003	0.003	0.003	0.003	0.003	0.003
Stdev	0.058	0.089	0.058	0.053	0.051	0.052	0.052	0.054
Skewness	-0.009	4.358	-0.070	-0.278	-0.386	-0.338	0.107	-0.202
Kurtosis	2.716	41.062	2.715	1.862	1.047	1.903	1.521	1.252

Table A3.3: (continued)

	Hong Kong				Sweden			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Observations	204	204	204	204	204	204	204	204
NAs	0	0	0	0	0	0	0	0
Minimum	-0.231	-0.317	-0.313	-0.194	-0.194	-0.178	-0.152	-0.260
Quartile 1	-0.021	-0.021	-0.022	-0.018	-0.021	-0.022	-0.017	-0.011
Median	0.017	0.011	0.009	0.012	0.017	0.010	0.007	0.013
Arithmetic Mean	0.016	0.014	0.011	0.010	0.016	0.010	0.011	0.011
Geometric Mean	0.014	0.011	0.008	0.008	0.014	0.008	0.010	0.010
Quartile 3	0.046	0.052	0.043	0.038	0.052	0.038	0.043	0.039
Maximum	0.309	0.241	0.282	0.179	0.419	0.292	0.218	0.283
SE Mean	0.005	0.005	0.005	0.004	0.004	0.004	0.004	0.004
LCL Mean (0.95)	0.006	0.003	0.001	0.002	0.007	0.002	0.004	0.004
UCL Mean (0.95)	0.026	0.024	0.021	0.018	0.024	0.017	0.018	0.019
Variance	0.005	0.006	0.005	0.003	0.004	0.003	0.003	0.003
Stdev	0.071	0.075	0.073	0.056	0.062	0.055	0.051	0.055
Skewness	0.171	-0.038	-0.256	-0.274	0.971	0.511	0.147	-0.286
Kurtosis	2.513	2.525	3.233	1.418	8.300	3.976	1.558	5.557