# Financial Characteristics of Firms With High ESG Scores

Englich, Frans
fenglich@fastmail.fm

Gedda, Oscar
gusengwos@student.gu.se

June 1, 2020

Master Thesis Spring 2020
Master of Science in Finance

Supervisor: Conny Overland

UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

## Abstract

This study investigates the impact of Corporate Financial Performance (CFP) on Environmental, Social and Governance (ESG) measures. A large part of earlier literature concerns the effect of ESG on CFP. In later years, some studies have examined whether the causality may flow the other way which is the vantage point of this thesis. Out of a data set of more than 8000 firms globally, a random sample of 100 US firms is selected. The ESG data runs from 2004 to 2020 and is sampled monthly. To measure the effect of CFP on ESG, three different measures of CFP are chosen. These are size, measured by total assets and market capitalization, profitability, measured by return on assets (ROA) and return on equity (ROE), and net income. The results from running multiple OLS and GLS regressions while controlling for sector and time since listing on the stock market, indicate a generally positive but small effect of the CFP measures on ESG. These effects result in accepting the hypotheses that profitability and net income lead to higher ESG ratings. The regressions on the two measures of size have contradicting results, resulting in the hypothesis of a larger firm size leading to higher ESG not being accepted.

Keywords: CSR, ESG, CFP, CSP, sustainability

## Acknowledgements

During the writing of this master thesis we have learned about the topics and team work immensely. However, we stood upon the shoulders of many, in order to do this. We would like to express our deepest gratitude to Alexandra Pavlovskis and her team at Arabesque for handling our questions and data requests; Herman Andersson, Johannes Sivén, and Marcin Zamojski for guidance in the peculiarities of research and econometrics; and our supervisor Conny Overland for reviews. Omissions and errors are, of course, entirely our own.

# Contents

# List of Tables



# List of Figures

# Listings

# 1 Introduction

As thoroughly cemented in practice and documented by Berle and Means (1932), corporations should maximize shareholder value. Hence one would expect the Friedman doctrine (Smith, 2003), that "the business of business is business", to fully lay out the picture. In later decades, the tides have turned and more and more voices have risen in favor of the Stakeholder Theory (Freeman, 1984) which argues that a firm should attempt to maximize all its stakeholder groups' interests.

Following this rising trend, various measures of sustainability have begun to gain ground. Among others, the Triple Bottom Line (Elkington, 1994), Corporate Social Responsibility (CSR) (Bowen, 1953) and Corporate Social Performance (CSP) have been used as proxies for sustainability. One measure that has gained a lot of attention during later years is ESG (Environmental, Social, Governance). This, and the focus on quantifiable opportunities with ESG, is the reason this thesis discusses sustainability mainly with ESG perspective.

It can be noticed that the phrase sustainability has a dubious meaning. In some circumstances it refers to environmental issues solely. In other cases it is more broad, referring to a general ethical or "responsible" matters. One way the latter is quantified is through the three components found in ESG and this is how we define sustainability.

From main stream media, our networks and other channels, a general query formed which is whether sustainability measures require a virtuous stance or if they profitable as well. One common conception is the existence of a trade-off between ESG and Corporate Financial Performance (CFP). Although this certainly may be the case in specific circumstances, a large body of literature is now suggesting that the two have a positive relationship (Friede, Busch, and Bassen, 2015; Viehs, Clark, and Feiner, 2014). In the meta study done of about 2200 empirical studies, Friede, Busch, and Bassen (2015) find that roughly 90% report a non-negative relationship between ESG and CFP. Importantly, the large majority report positive results. Viehs, Clark, and Feiner (2014) report similar findings.

As such, there is a breadth of work in this area, and the inherent diversity of the literature acts as a strong robustness test. Therefore, we decided to not do yet another on this angle and hence changed perspective. Most literature has examined whether a firm's sustainability rating has an impact on its CFP or simply confirmed a positive relationship between the two. Another possible area of research is whether the opposite relation holds. That is, that CFP positively influences ESG ratings. Taking the results from the Friede, Busch, and Bassen (2015) paper as given, we investigate whether this opposite approach gives reasonable answers, and if so, what constitutes and characterizes firms with high ESG ratings.

We see this as a logical step from an academic perspective. We believe the earlier focus on whether ESG positively impacts CFP explains the sparse availability of literature on the opposite relation. We think this is an indicator to that we are investigating where it is needed, but also means it is an undertaking which isn't straight forward or which not necessarily leads to a streamlined, canned approach or result.

Using the findings of Carroll and Shabana (2010), sustainability and business success can be summarized in four aspects: Risk and cost reduction, better reputation and a stronger legitimacy, competitive advantages compared to competitors, and "synergistic value creation" which essentially means to consider several stakeholders, creating a symbiotic relationship. Empirically, our investigation has numerous usages. In essence, we are investigating which financial gears and mechanisms lead to ESG. One theme is a time perspective, that financial traits can be used as indicators for future, upcoming ESG ratings. Examples of applications are:

- Risk management. Since low ESG ratings are associated with negative risks, it is of interest to have a deeper insight into what leads to an increase in ESG rating so that perilous situations can be mitigated. Conversely, actions can be concluded that will increase ESG ratings, hence lowering risk.
- With an insight into what firm characteristics that lead to certain ESG ratings, it is possible to determine such developments in before hand. This is of interest for determining assets in portfolio management. For instance, a pension fund may be investing with an awareness of ESG, and therefore wants predictors for such future performance.
- With knowledge of what characteristics lead to ESG it can aid in long-term governance and management of companies, moving them in that direction.

The body of work by Friede, Busch, and Bassen (2015) investigate if sustainability affects financial performance, which had been the major focus of research up until then. Although the concrete usage or perspective of ESG stretches back to the 1950s and 1960s, when American pension funds started investing with a focus on sustainability (Roberts, 1958), ESG has recently gained more focus, which explains why Friede, Busch, and Bassen (2015) was able to summarize the direction of this. In essence, these papers investigated have some form of measure of CFP as the dependent variable, and establish a link from ESG indicators to it by having ESG as independent variable(s). Since the relationship between ESG and CFP has become widely established, this study attempts to take the discussion one step further. We now take this relationship as given, which enables us to have ESG as the dependent variable, using it as a proxy for CFP. The aim is then to research what specific factors, other than common variables that together constitute the ESG rating, lead to certain ESG ratings.

We met with representatives from Arabesque in spring 2019 at the conference *Sustainability in Finance*, Stockholm. From the presentations and convening discussions afterwards, it became clear that a collaboration for master thesis would be interesting. This is the reason we chose ratings produced by the London based company Arabesque as ESG measures. As independent variables we have chosen some traditional measures of CFP as well as firms' sector belonging in order to control for sector specific characteristics. With this we conceptually look at what characterizes companies which attain good ESG ratings. At first thought this might seem to have an obvious answer: it depends on however Arabesque's rating is determined. The ESG rating, which we elaborate more on further on, is in one sense shallow, it tells how a firm's actions are judged by external parties, such as stakeholders, media, non-governmental organizations, and so forth. What remains is an explanatory perspective: what are the underlying, driving factors and underlying characteristics of these ESG-performing companies? Using ESG as a dependent variable allows us to answer that.

We have formulated a general research question as: *Which financial characteristics lead to a high sustainability rating?* In order to properly investigate it we need to quantify both the financial characteristics and sustainability. As proxies for CFP we use two measures of size, one accounting-based and on market-based, two measures of profitability, ROE and ROA, and net income, which can be thought of as a combination of size and profitability. As sustainability rating we use ESG, computed by Arabesque.

Out of the Arabesque data set of more than 8000 firms globally, a random sample of 100 US firms is selected. The ESG data runs from 2004 to 2020 and is sampled monthly. This data set was provided by Arabesque and is of high quality, while we had some issues gathering relevant and correct data from Bloomberg. The independent variables of size, profitability, net income, and date of listing on the stock market, are heavily based on data extracted from Bloomberg. In the case of missing listing dates, this data is gathered from the respective firm's IR website.

One issue with the final data set has to do with sampling periods. We chose to go with monthly ESG sampling, daily sampling for market capitalization but only quarterly sampling of the remaining variables, due to their nature of only showing up in quarterly reports. This leads to a quite large decrease in the amount of observations for all but the market capitalization regressions.

The results from running multiple OLS and GLS regressions while controlling for sector and time since listing on the stock market indicate a generally positive but small effect of the CFP measures on ESG. These effects result in evidence of a positive correlation between ESG and profitability as well as net income. The two measures of size result in contradicting results, suggesting that a larger firm size does not necessarily lead to higher ESG.

With this thesis we intend to build upon the knowledge of how financial characteristics affect sustainability in firms. We consider the positive results to be interesting and mostly in line with earlier literature. We attribute the sometimes negligible impact of the results to a quite small sample size and a very large scope in the types of firms covered. An interesting way to build more depth in future research could be to narrow down the sample criteria and to maintain as large a portion as possible of the firms fulfilling these criteria.

# 2 Literature Review

Two decades ago, Global Compact (2004) published a study called "Who Cares Wins". In this report, the term ESG - an acronym consisting of the terms Environmental, Social, and Governance - was coined. Since then, the term and the general concept of sustainability have had upswings in popularity, which has resulted in developments in many areas. One area, which the Global Compact (2004) was addressed to, is the financial markets. Today the finance industry is full of green bonds, green funds and increasing requirements of sustainability reporting and regulations thereof (Hale, 2019).

Other similar terms such as the Triple Bottom Line (Elkington, 1994) and Corporate Social Responsibility (CSR) (Bowen, 1953) are sometimes used interchangeably with ESG. A fourth proxy for sustainability in earlier literature is Corporate Social Performance (CSP). As an example, Arouri, Gomes, and Pukthuanthong (2019) study CSR while using ESG databases. ESG has gained in popularity over CSR, partly due to the more quantified approach through the use of metrics describing a company's sustainability performance (NPower, 2019).

The business case for sustainability is largely discussed in Carroll and Shabana (2010). The main argument is simply that firms give value to the idea of a long-term interest to act sustainably. Reasons for this could be to avoid unfavorable regulatory changes and that it is more preferable to be proactive than reactive as a successful firm. Barnett (2007) argues that although CSR activities are not always profit maximizing, they may still be indirectly. He compares investments in sustainable practices to R&D and advertising, which have been given a value on financial statements and in valuations from developments in accounting over the years.

Firms with large measures of Corporate Financial Performance (CFP) may lead to stakeholders gaining the perception that shareholders profit at the expense of other stakeholders. It may therefore be in the interest of well-performing firms to focus on increasing measures such as ESG scores in order to maintain good relations with their stakeholders. Maximizing short term profits is likely to lead to underestimates of negative long-term effects (Carroll and Shabana, 2010).

## 2.1 ESG and Corporate Financial Performance

The literature on ESG in a financial setting has mainly focused on its relation to CFP. This area in the financial literature has been developed and discussed over a long period of time. There is a large body of papers written on the relation between ESG factors and financial success.

This thesis is based upon the findings of two relatively recent meta studies which have summarized the literature on the topic of sustainability and CFP. One of the studies is a meta study on the relation between ESG and CFP by Friede, Busch, and Bassen (2015) which evaluated more than 2200 studies on the topic. They found that most studies conclude the existence of a positive relation between the two and that only one in ten studies result in a negative ESG-CFP relation. The other meta study (Viehs, Clark, and Feiner, 2014) also finds positive relationships between sustainable investments and financial performance and is a corner stone in documenting current trends within sustainability and financial performance.

Even though these review articles generally show a significant relationship between ESG and corporate financial performance, and this study's whole approach is built on the generally widespread evidence toward a positive relation between ESG and CFP and that more ESG is better, some studies have attempted to nuance this notion. Barnett and Salomon (2012) find that firms with low CSP show greater CFP than companies with average CSP. At the same time, high CSP firms show higher CFP than the lower two tertiles. Other arguments include that large companies with high profitability and large sales can afford to invest in areas to improve ESG ratings, regardless of its impact on performance (Mishra, 2020).

Manescu (2011) discusses two arguments that can help explain high ESG firms' performance. The two arguments are called the economic argument and the discriminatory-tastes argument. The economic argument looks at the benefits as well as the costs of maintaining high ESG ratings and tries to compare the two. The discriminatory-tastes argument argues that the result of the economic argument is secondary. This is because some investors may find non-financial benefits of ESG investing. Therefore, these investors will impact security prices and thus the returns of certain stocks.

## 2.2 Causality of ESG and Corporate Financial Performance

The previous chapter mainly discusses the traditional view on the relation between sustainability measures and financial performance. This is largely related to stakeholder theory (Freeman, 1984) which suggests that taking all stakeholders into account will lead to better performance than just acknowledging shareholders. Despite the large interest of researching this field, using

ESG as the dependent variable, that is estimating whether CFP lead to positive effects on ESG measures, has not been a common approach. One explanation for the phenomenon that CFP lead to CSP, and not the other way around, has been named the slack resources hypothesis (Mcguire, Sundgren, and Schneeweis, 1988; Waddock and Graves, 1997). Their findings suggest that successful companies are better suited to comply with, and invest in, sustainability. Slack resources has been used to mean both its literal meaning, portrayed by debt/equity ratio and current assets/current liabilities, and a conceptual measure, using either accounting or market based measures (Melo, 2012).

The literature on the direction of the causality is not unanimous. Waddock and Graves (1997) describe it as a chicken-or-egg problem where it remains uncertain as to which direction the causality goes. While the positive correlation between ESG and CFP is generally found to be statistically significant, the direction is not as certain. Orlitzky, Schmidt, and Rynes (2003) find evidence of a bidirectional and simultaneous relationship. This is in line with the findings of Waddock and Graves (1997) and Tuppura et al. (2016) that find this bidirectional causality in the clothing, energy and forest industries. Martinez-Ferrero and Frias-Aceituno (2015) find further evidence of this bidirectional causality and what they call a "synergistic circle" between sustainability and financial performance.

## 2.3 Financial Characteristics and ESG

The literature on which financial characteristics affect ESG is far from as well-developed as the literature on the general impact of ESG on CFP. By turning the question around and attempting to answer which typically non-ESG factors affect ESG scores positively this section is structured in a different way. We have focused on several characteristics individually and their impact on ESG. These are size, measured by total assets and market capitalization, profitability, measured by ROE and ROA, net income, and the controlling variable industry to control for sector specific factors.

### 2.3.1 Size of Firm

The size of a firm is a measure that in a way summarizes many measures of a firm. As size of a firm, we have decided to focus on one accounting-based measure and one market-based measure. Total assets is one of two characteristics in this study which act as a measure of firm size. This measure is accounting-based and as such, mostly backward-looking consisting of economic consequences of historic events. The other characteristic considered to act as a proxy for size of the firm is market capitalization. This measure is market-based and as such, mostly forward-looking and valued as a sum of the invested equity as well as some kind of financial measure of future events. Gregory, Tharyan, and Whittaker (2014) argue that when investigating CSR and performance, market-based measures may be more relevant.

The impact of market capitalization on stock performance was popularized by the Fama and French model (Fama and French, 1992) where the small minus big market capitalization factor was introduced. Brammer and Pavelin (2008) find that quality of environmental disclosure is primarily linked to large firms. Tuppura et al. (2016) find that market capitalization positively Granger-causes CSP in the clothing, energy and food industry while it also seem to negatively have an impact on CSP in the energy industry.

Drempetic, Klein, and Zwergel (2019) find a positive significant correlation between firm size and ESG performance where size is computed from number of employees, market capitalization, revenue and total assets. Despite finding significant results, the study problematizes the way ESG scores are computed. The main argument is that larger firms receive more attention, thus producing more data for the rating agencies, and that they have larger resources to produce data the ESG ratings consists of. The question then remains whether large firms are more sustainable than smaller ones or simply have better sustainability reporting. This discussion is in line with Baumann-Pauly et al. (2013) and Wickert, Scherer, and Spence (2016) who argue that smaller firms possess several characteristics which positively impact their work on sustainability, while lacking the ability to report and communicate these advances externally. Wickert, Scherer, and Spence (2016) call these two opposites the "large firm implementation gap" and "small firm communication gap" respectively. As such, the sustainability rating provider and its way to decide on materiality of certain KPIs impacts which firm size have better ESG scores.

### 2.3.2 Profitability

Profitability measures such as Return on Equity (ROE) (Beck, Frost, and Jones, 2018; Orlitzky, Schmidt, and Rynes, 2003) and Return on Assets (ROA) (Elsayed and Paton, 2005; Orlitzky, Schmidt, and Rynes, 2003; Tuppura et al., 2016) are proxies of CFP used in earlier studies.

There is evidence that suggest that higher economic and environmental performance can occur simultaneously. That is, companies can improve profitability measured by ROE and ROA while bettering their E score in ESG ratings (K.-H. Lee, Cin, and E. Y. Lee, 2016). Tuppura et al. (2016) report similar findings with a Granger causality between ROA and CSP and between CSP and ROA in certain industries, supporting a virtuous circle as discussed by Orlitzky, Schmidt, and Rynes (2003). In the food industry, ROA is shown to Granger-cause CSP while the opposite relation doesn't hold, signifying a direction of the causality (Tuppura et al., 2016).

### 2.3.3 Net Income

Net income could be argued to mainly be a consequence of the combination of firm size and profitability. As such, it has not been as investigated as the other measures. There is nonetheless some literature on the subject. Li and Thibodeau (2019) highlight an emerging trend of coupling executive compensation to sustainability measures. They find evidence of earnings management in firms where CSR ratings are low while they do not see it in high performing CSR-firms.

### 2.3.4 Industry/Sector

There is a large body of literature on the relation between ESG and CFP in different industries. One inherent difficulty when comparing these industries and their focus on ESG is the way ESG scores are computed. Arabesque's ESG scores take sectors into account and adjust the weights on certain criterion accordingly. Further important aspects are the general differences in size of total assets on the balance sheet as well as profitability measures which differ significantly between sectors. I.e. firms in the energy sector generally have higher total assets than consultancy firms.

Brammer and Pavelin (2008) find that quality of environmental disclosure is primarily linked to firms in sectors that are normally associated with environmental issues. These industries are also most likely to report environmental disclosures (Halme and Huse, 1997). Tuppura et al. (2016) find evidence of bidirectional causality between CSP and CFP in the clothing and energy sectors, which are prone to scrutiny of their sustainability practices. They further find that CFP Granger-causes CSP in the food industry and argue that this may be in accordance with the slack resources hypothesis.

# 3 Hypotheses

Having established that a high ESG rating leads to CFP, we now operationalize how to investigate what leads to ESG. Starting at as a broad perspective as possible, the immediate question is:

*What characteristics lead to a high ESG score?*

We believe this question to account for the wholeness of the true data generating process, which we speculate is affected by numerous conditions and aspects:

- Macroeconomic conditions. Acting like a form of systemic factor.
- Industry and sector specifics.
- Legislation and regulations.
- Cultural attitudes and norms.
- Demand and influence by the market. Intuitively this form of pressure could be a major factor that indirectly affects firms' decisions.
- Management. By this meaning strategic, operational decisions done by the firm.
- Corporate governance. Such as characteristics of the board of directors.
- (Other) financial aspects.

In order to make this feasible, we focus on only one: the financial aspect. That is, what financial characteristics lead to ESG? Note that in the above list the financial aspect spans several bullet points. Hence, to create a model for the true data generating process that not only includes finance but all true factors is an infeasible task but being aware of that reveals that we haven't captured the whole picture. In short, models that are used for regression analysis and which, as commonly found, are incomplete, often have low $R^2$-values and regression coefficients, even if, of course, considered relative to the dependent variable.

Our angle on this can be summarized with the following research question:

*What financial characteristics lead to a high ESG score?*

Consequently, we form the following hypotheses:

- $H_0$: No financial characteristics affect ESG ratings
- $H_1$: Larger firm size leads to higher ESG ratings
- $H_2$: Higher profitability leads to higher ESG ratings
- $H_3$: Higher net income leads to high ESG ratings

### 3.1 $H_1$: Larger firm size leads to higher ESG ratings

The size of a firm is a measure that in a way summarizes all the measures of a firm. In general, higher profitability and earnings should equal an increased size, both in terms of accounting and market based measures. As such, it is not unreasonable to expect that if profitability and earnings positively correlate with ESG scores, so should firm size. In this case we use total assets as an accounting based measure of size and market capitalization as a market based measure of size.

On the one hand, firm size has a positive impact on the amount of ESG information such as sustainability reports that firms put out. Medium and small size firms may not prioritize such exhaustive sustainability reports (Drempetic, Klein, and Zwergel, 2019). Many of them, however, might have products and services that fulfill the ESG criteria. Despite this contradiction, we believe a large firm size to be positively correlated with a high ESG score.

### 3.2 $H_2$: Higher profitability leads to higher ESG ratings

One issue of taking a randomized sample of an entire market is that several performance measures will depend largely on the size of the company and the industry in which it is active. To some degree, this is also the case with profitability ratios such as ROE or ROA. Adjusting for sector, we believe that higher profitability leads to higher ESG scores.

### 3.3 $H_3$: Higher net income leads to high ESG ratings

By net income we get an absolute measure of how well a firm manages to generate a surplus. We hypothesize that a high net income leads to a high ESG score because possibly managers and investors makes decisions towards sustainability when having an economic surplus. We hence suggest that it would reflect a virtuous approach where the sustainability investment is not done solely based on believing it's an investment leading to CFP. However, it might also be in a firm's long-term interest to act sustainably. With an economic surplus such in the case of high net income, long-term thinking is reasonably more prioritized.

## 4 Data

In order to approach our research question and the succeeding hypotheses, we gathered data points from two sources and merged them for succeeding analysis. The code achieving this is in appendix A. In short, our final data set is a time-series analysis of unbalanced panel data over 100 US listed equities, sampled monthly, stretching back to earliest 2004.

Both data sets, one from Arabesque and one from Bloomberg, can be considered proprietary.

While the classical data points from Bloomberg have transparent origins with straightforward interpretations and explanations, the score variables of Arabesque's need more attention. In short, they don't disclose the details of how the scores are produced beyond their general approach.

All programming and statistical work was done in Stata 16 and Matlab R2019b. We've used a number of user-written packages for Stata, see appendix A for details and source code.

## 4.1 Arabesque's Data

There exist a plethora of entities offering ESG ratings. Well-known data vendors are for instance Morningstar and Bloomberg while more novice are Arabesque's, Refinitiv's, or Vigeo Eiris' offerings, such as used in Ferrell, Liang, and Renneboog (2016). Arabesque was incorporated in 2013 (House, 2020), after a management buyout from Barclays Bank.

In addition to the various entities offering ratings, there is a diversity of ratings available. For instance, Arabesque offers in addition to the more common ESG score, a Temperature score as well as a rating on United Nations' Global Compact score. We chose the ESG score because it lies closer to existing literature and discourse. It can be argued that other ratings have their merits but for instance Temperature score is not of relevance of what we are investigating. The wealth available on this front serves well for robustness checks and cross-validation by performing the analyses against different data sets, and this is an area viable for future research.

### 4.1.1 How Arabesque's ESG Score is Computed

We will now account for the processes and procedures that lead to Arabesque's ESG scores, produced by their software service called S-Ray. The source of this information is from documents sent from, and correspondence with, Arabesque. Errors are entirely our own.

The overall ESG score measures how well firms perform on sustainability topics that are deemed material to their line of business. It ranges from 0 to 100, where higher is better. It is produced by considering the three ESG sub-scores (Environmental, Sustainability, Governance). These also rate firms from 0 to 100, higher is better. They are produced by materiality-weighting 22 so called *feature scores* and hence provide a granular perspective on their aspects. The ESG sub-scores and corresponding feature scores are summarized in Table 1.

| Environment | Social | Governance |
|---|---|---|
| Emissions | Diversity | Business Ethics |
| Environmental Stewardship | Occupational Health and Safety | Corporate Governance |
| Resource Use | Training and Development | Transparency |
| Environmental Solutions | Product Access | Forensic Accounting |
| Waste | Community Relations | Capital Structure |
| Water | Product Quality and Safety | |
| Environmental Management | Human Rights | |
| | Labour Rights | |
| | Compensation | |
| | Employment Quality | |

Table 1: Overview of S-Ray feature scores.

In contrast to other sustainability ratings, Arabesque's approach is quantitative and technology driven. It is possible to have a general discussion of pros and cons of quantitative and qualitative methods. The part called the *Input Layer* scans reports, news and NGO-based activity. For instance, this involves natural language processing[1] of about 3,000 news sources. In later steps this data is used to compute feature scores that in the end manifest in the various ratings. For more information, see Arabesque (2020).

Table 2 shows rudimentary statistics of the ratings as well as the other variables. As can be seen in the table, the environment rating has the highest variance. The variables chosen to investigate are return on assets (ROA) and return on equity (ROE) as measures of profitability, total assets and market capitalization as size of the firm, and net income. ROA and ROE are measured in

[1]Natural language processing is computer science and means the parsing, understanding and processing of natural languages, such as Arabic or French, by software. It is a complicated and generally considered a difficult and inexact undertaking because these languages were not designed by a mathematical rigour that suits software.

percentages with a mean of 2.5 % and 4.0 % respectively. Total assets and market capitalization are presented in millions of dollars with a mean of 28.8 billion and 31.9 billion respectively. Net income is measured in millions of dollars with a mean of 188.2 million. The data for the five variables is less extensive than for ESG, ESG-E, ESG-S and ESG-G. The reason is mainly that the sustainability measures are taken on the first day of each month. If this day happen to occur at a date where the markets are closed, there will be no measure of the market capitalization of that day. Furthermore; Net income, total assets, ROA and ROE are based on quarterly reports, thus further decreasing the number of data points by two thirds.

Table 2: Descriptive Statistics of all variables

| | N | mean | sd | variance | min | max |
|---|---|---|---|---|---|---|
| ESG | 7147 | 50.901466 | 7.5972377 | 57.71802 | 28.719999 | 71.769997 |
| ESG-E | 7147 | 46.010024 | 14.389277 | 207.05128 | 26.200001 | 85.129997 |
| ESG-S | 7147 | 51.173387 | 8.0380898 | 64.610888 | 36.25 | 78.57 |
| ESG-G | 7147 | 53.169138 | 11.354861 | 128.93287 | 17.58 | 81.839996 |
| Net Income | 1893 | 188.1633 | 426.6013 | 181988.67 | -5389 | 3134.1699 |
| Total Assets | 1888 | 28198.905 | 62828.068 | 3.947e+09 | 1.71 | 410295 |
| Market Cap | 6132 | 31929.542 | 58534.19 | 3.426e+09 | 19.094601 | 422741.16 |
| ROA | 1882 | 2.4977665 | 21.209553 | 449.84515 | -315.69839 | 45.882401 |
| ROE | 1799 | 4.0007441 | 176.52475 | 31160.988 | -6822.2622 | 176.02341 |

To get a sense of how the scores evolve over time and for consideration of possible volatility modelling, we looked at the score rating in log format, like log returns. We picked a firm, Wells Fargo (ticker WFC), that in 2018 had a severe negative governance-related event (Shapiro, 2018). We found no indications of that this exogenous shock affected the score. The result is plotted in Figure 1. For the shock in 2010, evident in the figure, we couldn't find an explanation.


Figure 1: Log returns of ESG scores for Wells Fargo.

#### 4.1.2 Arabesque's Universe

Arabesque S-Ray provides data on a global basis, covering about 8400 firms, stretching back to 2004, available in up to a daily sampling. Working on such a large data set, roughly 50 million observations times variables, proved infeasible for a number of reasons. For that we chose a sub-set, as will be discussed.

At first we attempted working on this global data set whose result would be useful because most investment strategies typically involve investing globally due to global diversification strategies. However, while we combined the various data sets we encountered many challenges related to this. The main problem was analyzing companies that are listed on different continents and exchanges. One problem is related to calendars, that time-series analysis is greatly complicated by that availability differs between the different exchanges. Another aspect is conversion and normalization of financial data, such as pricing. In general it involves bias-creating policy decisions, and some press ahead on that using for instance MSCI's indices. Another question is whether firms can be globally compared for this kind of analysis, due to underlying assumptions such as culture and legislations. Due to the constraints for this paper, we narrowed down to US firms.

Generally during our excavations the data set provided by Arabesque is of high quality, and we've found a few faulty data points that we had to circumvent, no systematic errors. Some tickers[2] seemed to be dubious by having been reused, probably caused by acquisitions, mergers, delistings and similar. In those relatively few cases we chose other equities. While this practice has its advantages, it can also be argued to cause a bias, albeit likely small.

Still, Arabesque doesn't index all firms in universe. Hence their selection poses one source of bias for this investigation. One determinant of whether they index a firm is whether data is available for the firm. Hence here is a bias too, firms that have been mentioned in the data sources that Arabesque mines will be indexed. We don't know the extent of not-indexed firms, but this is in other words a survivorship bias.

## 4.2 Bloomberg Fields

As our source for financial data points of general character, we chose Bloomberg, partly because of familiarity. Our data handling has been a limiting factor and consumed significant amount of time. Our hypotheses, scope and directions have been changed and shaped by what turned out feasible given our data set, hence going back and forth between data gathering and other areas. During this process we gathered several insights in Bloomberg, where primarily unavailability forced us to take turns.

One limiting factor was the university's Bloomberg license whose limit we reached through Bloomberg's Excel plugin, and which the help desk couldn't levitate. Luckily we managed to circumvent this through a helping hand, outside the university. Another limiting factor is the availability of Bloomberg fields, which are advertised as being extensive (Bloomberg, 2020). Although there are many nontraditional fields declared, and would promise interesting research, we discovered that few data points are actually available for these fields. In general Bloomberg's

[2] Arabesque's data set identifies equities using SEDOL identifiers, as well as regions and the tickers they are identified by.

availability seems to be in the order of 1) financial data; 2) accounting data; and 3) more unconventional, such as within corporate governance. In some cases the sample rates have natural explanations, such as that quarterly reports (also called 10-Q, from the SEC) are available at quarterly intervals.

For the unconventional fields, the best explanation we have is that only some, mainly large capitalization, firms are covered by the unconventional fields. This means two things for this paper: we have reduced the aspects that lead to a large amount of data points; and some hypotheses have been scrapped due to that we haven't found data points for them. That is, we considered the hypotheses valid in themselves, but they were infeasible to carry out. We've also encountered data points for some fields being available at inexplicable, random dates. In short, data availability is a concern.

The implication is that for some fields missing values in the data set have valid, natural explanations and aren't at fault. Stata, generally has as practice to drop observations with missing values for instructions such as `regress`, hence reducing asymptotic behavior.

### 4.3 Final Data Set

That we are reasoning around essentially a software service is interesting in itself, and could be argued to reflect how times have changed. Software through its algorithms reflects assertions, and is one way to take stances. Writing a paper is another. While Bloomberg acts as a more traditional data vendor and is an intermediary of data from other sources, such as 10-K (annual) reports to the SEC, Arabesque make contributions through its business logic and design.

One way to see this is that what Arabesque now automates through its, for instance, natural language processing, researchers previously did manually. This was probably a tedious and error prone process, and such massive manual endeavours were carried out in for instance Klassen and Mclaughlin (1996).

The subset of Arabesque's data set, which essentially is Arabesque's population data set, is 100 US equities. The count was determined in order to not create a too large data set for the data point requests via Bloomberg. The selection was done by a randomized draw from the whole population of US equities. This can be argued to be one source of bias. Though no matter what, this bias would likely enter in some way another, since a subset of the population would have had to been chosen. One approach could have been to select a stratified sample based on the ESG rating, but a source of randomness would occur in the other variables then. Figure 2 shows the draw related to the population, which shows asymptotic movement toward the same distribution, Gaussian. The vertical axes are number of observations (density), while the horizontal axes are the ratings (intrinsic values).

The data set stretches back to 2004 and is sampled monthly. This is as far back as Arabesque offers, and we chose that as data set in order to get as good asymptotic statistical and mathematical properties as possible.

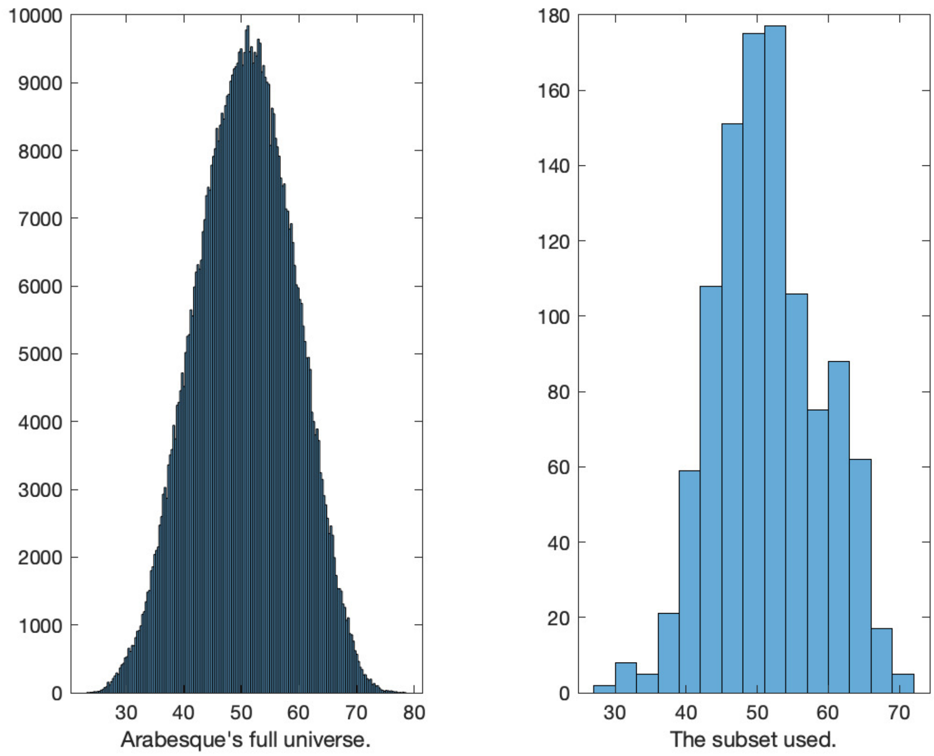


Figure 2: Histograms over ESG rating.

Another aspect was matching various dates. Essentially, the two data sets, from Bloomberg and Arabesque, were combined by custom joins using tickers and dates as keys. The latter generally didn't match. For that we chose the closest nearby date. This affects our data quality unfortunately, and means that exogenous events increase the endogeneity in our models.

The data set is unbalanced. This is natural due to that firms are listed (and delisted) at different moments in time. It affects our regressions. That the data is unbalanced could have been avoided by selecting a sub-set that doesn't have this trait, but this has its own set of complications, such as that likely a small sample would be available, as well as having a bias.

In total the data set has 7,147 observations, adjusting for companies without ESG scores. However, as discussed before, for natural reasons some observations drop out due to missing fields, leading to that in regressions the actual count can be lower. This is also the explanation to why the firms participating in the regressions sometimes drop below the data set's full 100 – there are no complete observations for those firms. This reduces efficiency, but the the worst case is at 84, and therefore we consider that the approach remains intact.

In summary, the data sources and situation of the university's license have put upon us these limitations:

- What hypotheses we are able to answer.
- The amount of data points we can have. Most notably this has been the sample rate and count of equities.

#### 4.3.1 Sample Rate

Allegedly, S-Ray's database over its covered universe is updated on a daily basis. However, from our understanding the whole universe is not updated in unison, but incrementally. This means that the integrity is diminished. A sample at a given point in time will involve firms that are updated at different times during that window. Hence the larger time window, determined by the sample rate, the more firms that are contained within it. On the other hand, a lower sample rate, hence larger window, means the time-series nature is less present. Still, the continuous updating scheme, which we speculate is an unfortunate effect of that there simply isn't sufficient resources, such as bandwidth and processing power, for updating the whole universe at once, meaning exogenous shocks aren't captured in unison. We chose a monthly sampling because it posed a trade-off between being informational, but at the same time not causing an unwieldy data set.

# 5 Methodology

Our methodological approach is, introductorily briefly, to run multiple Ordinary Least Squares (OLS) and Generalized Least Squares (GLS) regressions on the panel data set, which is unbalanced. Panel data is also called longitudinal data or cross-sectional time-series. The cross-sectional aspect consists of 100 firms (often called individuals), and the time-series aspect is covered with monthly sampling retrospectively, from year 2004 to 2020.

We use linear models because according to our judgment the (eventual) causal relationships are linear; a relatively simple, direct effect. That is, we see no justification for a more complex or non-linear model. Another area where non-linearity may enter the models is through transformations, such as log or squared, of the independent variables. We neither saw such alterations relevant in any of the cases.

The reason we use multiple methods is to cast as robustness tests and model the data from multiple angles. These can be described as that we as researchers examine how intuitively speaking core regressions behave when viewing the problem from different angles by controlling, removing or adding independent variables, or changing methods.

One way that points in the direction of that GLS is applicable is as follows (Lazzaro, 2020). Let T be the count of observations in the time series dimension, and let N be the count of observations in the cross-sectional dimension. The question is if the number of firms (individuals) is larger than the (time) observations. If N $>$ T, called a short panel, then OLS and hence `xtreg` in Stata is suitable. If T $>$ N, also called a long panel, then GLS and hence `xtgls` is suitable. In our case

T is in general magnitudes larger than N, which is 100 (or slightly less). Therefore we decide to in addition use GLS, which has as advantage over OLS and Weighted Least Squares that it is statistically more efficient in some cases, where the latter two also can lead to misleading inferences, that is, statistical inconsistency.

Another question concerning the regressions and data is auto-correlation (also called serial correlation), in particular whether auto-correlation is present in the error term. When the error terms are correlated with the response, we can not simply omit these errors since they affect the result. For the scope of this thesis, we didn't manage to formally investigate auto-correlation, although there are formal tests for this, such as Durbin-Watson.

The error term is always a central piece, and it's of interest to minimize it. Hence, a thorough understanding of its nature is relevant. The classic assumption is that the error term is normally distributed

$$\varepsilon | X \sim N(0, \sigma^2)$$

where we have zero mean and variance $\sigma^2$. However, in our case we question whether the data is generally normally distributed and also, due to that we have time-series, auto-correlation may be present. The problem with auto-correlation in linear panel data is that it biases the standard errors and create a less efficient result (Drukker, 2003). For the auto-correlation there are tests available, which various disadvantages, advantages and areas of application. Ljung Box is one, through Stata's `wntestq`, but performs poorly in our case due to requirements of uni-variate time-series as well as strictly exogenous independent variables. They are first-order auto-regressive (F Baum and Schaffer, 2013). However, they are not panel data-aware.[3] Still, we judge that the data has auto-correlation makes sense, from the perspective of economic theory. Also, visually inspecting the graph in Figure 1 indicates serial correlation. Due to this we choose to also view the data in the light of Generalized Least Squares (GLS).

Auto-correlation, at for instance lag 1, means that for the generic regression

$$y_t = \beta_0 + \beta_1 X_t + \cdots + \beta_n X_{n,t} + u_t$$

[3]There is research in this area, though we failed at managing to get it to work. See Drukker (2003) and Wooldridge (2002).

the error term is modelled

$$u_t = \phi u_{t-1} + \widetilde{u}_t$$

for which (regular) OLS can give misleading estimates. For this, GLS steps in. Auto-correlation can also be caused by misspecification (Verbeek, 2017).

The GLS estimator (ibid., p. 100) is defined as

$$\hat{\beta} = (X'\psi^{-1}X)^{-1}X'\psi^{-1}y$$

where $\psi$ is a positive definite matrix. It is easily seen that if $\psi = I$, the identity matrix, the estimator reduces to the OLS estimator. The estimator can only be computed given the data if $\psi$ is known, which it typically is not. The same holds for our case. The solution to this is to first estimate $\psi$, which then yields the Feasible Generalized Least Squares (FGLS). This is what Stata's command `xtgls` in effect does. What most of our OLS regressions have in common with GLS is that they are random effects, reflecting the nature of the data.

## 5.1 Heteroscedasticity

For linear panel data regression there is standard assumptions[4], of which one the following assumption (Brooks, 2014, p. 181):

$$Var(u_t) = \sigma^2 < \infty.$$

One way to formally test this is using the test White (ibid., p. 183). Based on visual inspection of the plots in Figure 3, we conclude that homoscedasticity does not hold, and hence the above assumption does not hold. One remedy for this is to do robustness checks with GLS.

The panel data is two-dimensional, a combination between cross-section and time-series. The advantage of longitudinal data or study is that it tracks the same entity (in our case firms) across time and hence the differences between those entities, considered as whole, are less likely to be time-dependent. A longitudinal study is hence more accurate and reduces systematic bias or inefficiency. Another way to see this efficiency is that long-term effects are found as opposed to effects dependent on the short-term.

[4]Ideally all would be investigated, but due to size limitations we do what we consider the most important.

In Figure 3 we have plotted the various independent variables against the final ESG score. In intuition these mimic single OLS regressions; however, they don't have the full precision of the regressions, that also include controlling variables. As can be seen, many variables are heteroscedastic. Therefore all OLS regressions we run in Stata are with the option `robust`, leading to robust standard errors.

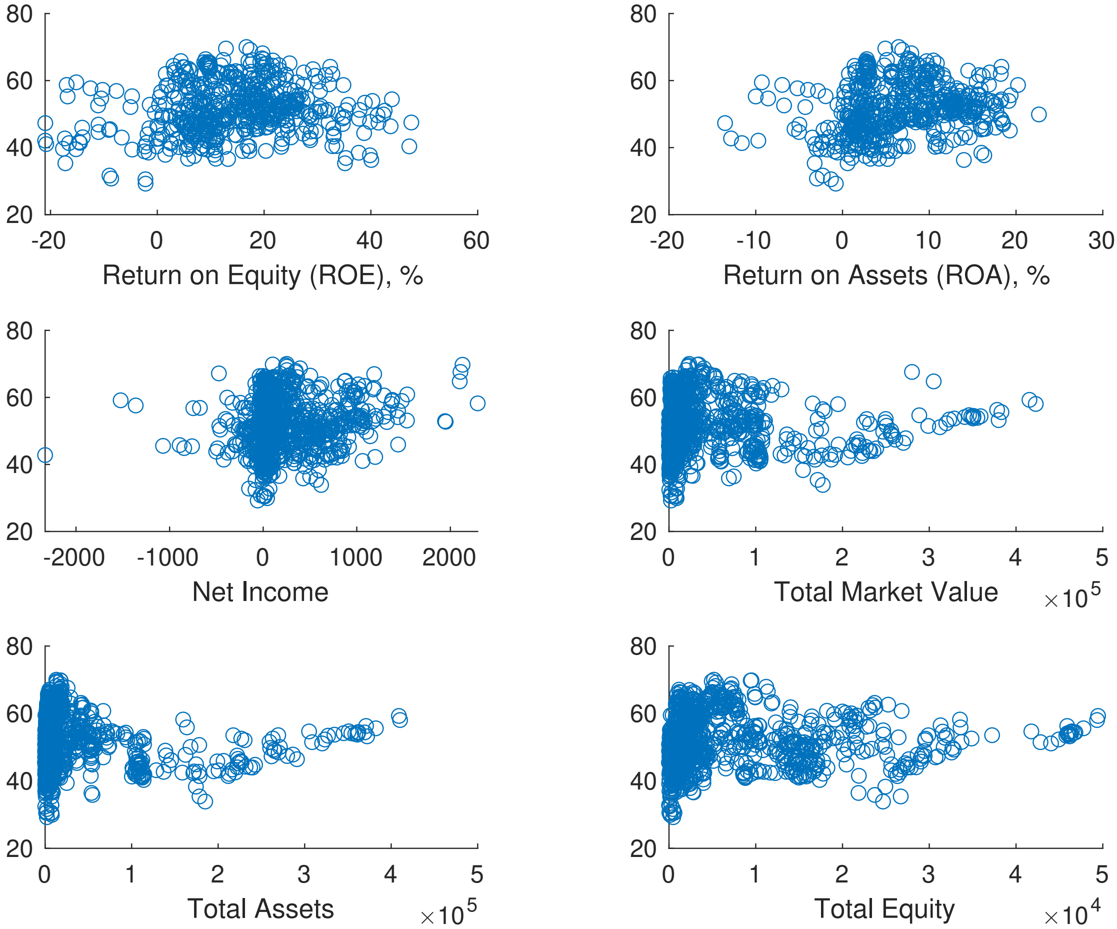


Figure 3: Scatter plot of key variables. All vertical axes are ESG scores.

In the data for ROE and ROA, as seen in Figure 3, we removed the outliers in the scatter plots because they concealed the view of the data and we want to improve the robustness, although from an extreme value perspective this practice can be questioned. Our definition of outlier is three scaled median absolute deviations, as practiced by Matlab's function `rmoutliers()`.

## 5.2 Controlling Variables

For all regressions related to the hypotheses we included controlling independent variables in order to increase robustness. These are:

- Economic sector. This is the category supplied and defined by Arabesque's data set.
- Industry. This is again defined and supplied by Arabesque's data set.
- Age, in years. The definition of this is Bloomberg's field `EQY_INIT_PO_DT` and hence age is defined as the amount of years the firm has been listed on the exchange, and can therefore be considered a proxy with characteristics that is suitable or the contrary. As a side note, the date that the firm was listed could be earlier than the year that data is available, which is capped by Arabesque's data by year 2004. We cannot find any negative impact of this or change in intuition or interpretation. The code for this can be found in appendix A. Bloomberg didn't have data points for some firms. For most of these firms, the IPO dates have been extracted from the firms' respective investor relations pages. This has diminished the problem of a smaller amount of observations.

The categorical variables for economic sector and industry are not subject to the dummy variable trap, multicollinearity. Assuming the underlying model specification is valid, there is no biased result caused by multicollinearity, it just produces larger standard errors for the independent variables in question (Maddala and Lahiri, 2009). The best model specification for the fitted data is where the independent variables correlate with the dependent variable, but minimally with each other.

Controlling might seem like an unanimously good idea. However, the more dummy variables that are added, the more noise in the model is controlled for. This could lead to a kind of overfit, an over-dampening of the model, reducing the useless information as well as the useful.

Another area we considered was detrending the data[5] but we were unable to intuitively identify any such trends. Possibly, one relevant adjustment is to normalize for inflation, but that turned out to be outside the scope of this thesis.

[5] See Matlab's function `detrend()` for a definition

## 5.3 Endogeneity

Endogeneity is a concern that needs to be addressed. Endogeneity is that the explanatory variables are correlated with the error term. Above discussion, where we have investigated model specification and endogeneity, shows that this specification most likely is endogenous. The technical parry for this is typically instrumental variable (IV) regression and to our best ability, as well as the competence in our vicinity available for conference, we haven't been able to identify a variable appropriate for the first-stage for IV regression. Identifying instrument variables truly is an inventive craft, and often picks from creative, unconventional fields in order to namely not suffer from the same problems as that it is set out to solve – be correlated with the error term. Sometimes it is for instance financial policies, such as found in Ferrell, Liang, and Renneboog (2016). However, we haven't been able to find appropriate instruments in earlier literature. In effect this means that we haven't been able to rule out reverse or simultaneous causality in our regressions on a model-basis. However, see section 6 for interpretation on this when taking the results into account.[6]

## 5.4 Robustness & Sensitivity

Panel data regressions can be carried out as fixed effects and random effects. Regarding independent individual variables, these attributes may or may not be correlated with them. In random effects, on the other hand, there are time constant, unique properties that are not correlated with the independent variables. In order to determine whether to run the panel data regression as a random effects or a fixed effects regression, we conducted a (Durbin-Wu-)Hausman test. The test is be expressed as

[6]We found, unfortunately too late, out about the method fronted by Tuppura et al. (2016), of using Granger-causality. In short, it involves an equation system of two Vector Auto-regressive (VAR) equations, where they have formulated corporate social performance and corporate financial performance, and subsequently see if there is Granger-causality. This is not causality in the traditional sense, but whether one system improves the forecast of the other. This is to our knowledge the best approach at addressing reverse or simultaneous causality. However, carrying out this approach at this time of writing would require a fundamental, infeasible rewrite of this paper.

$$W = \frac{(\hat{\beta}^*_{FE} - \hat{\beta}^*_{RE})^2}{Var(\hat{\beta}_{FE}) - Var(\hat{\beta}_{RE})} \overset{H_0}{\sim} {\chi^2}_1 \quad (1)$$

where $\hat{\beta}^*_{FE}$ is the coefficient of the regression – signified by the star, not the regression function, and it is distributed over a chi-squared distribution with one degree of freedom assuming the null hypothesis holds. In fixed effects there are unique attributes of the firms (individuals) that do not vary across time – $\alpha_i$. Note that Hausman test has two meanings: testing whether a variable can be treated as exogenous or if the research has to supply a structural equation for it; and, as in our case, whether random effects are valid or whether a fixed effect model is required (Brooks, 2014).

Random effects assumes that $Cov(\alpha_i, x_{it}) = 0$, which is the null hypothesis for this test. The alternative hypothesis is that $Cov(\alpha_i, x_{it}) \neq 0$. The null means that both random effects and fixed effects estimators are consistent. However, it also means that random effects are more efficient than fixed effects, that is $SE(\hat{\beta}_{RE}) < SE(\hat{\beta}_{FE})$. The Stata commands are `estimate` and `hausman`, code is available in appendix A. What estimator(s) to use is essentially based on economic intuition about the data and distributional assumptions, in particular related to the error term.

## 5.5 Model Specification

A number of formal tests exist for assessing correction of model specification. One is to test for using Ramsey RESET for regular regressions, using the test `ovtest` in Stata. RESET stands for *Ramsey Regression Equation Specification Error Test* and was developed in 1968 (Ramsey, 1969). It is related to the topic of $R^2$ and other goodness of fit tests. It tests that any possible omitted variable bias isn't caused by model misspecification (Verbeek, 2017, p. 74). It is applicable for linear models (Brooks, 2014, p. 220).

In our case we have panel data, for which `ovtest` does not work. For this the user-supplied extension `resetxt` is available (Shehata and Mickaiel, 2015), though is pending correction of faults (bugs), see Englich (2020). The purpose of this test is to improve the fit of the model.

# 6 Results

We now present our results, using the methodological tools presented in the previous chapter.

By sporadic investigation we didn't find high $R^2$-values which tells about the nature of these relationships. That is, we find no simple relationship for what leads to high ESG. It is a wealth of many small factors, and the most important factors are not in our models. For the fixed effects and GLS regressions, $R^2$-values are unavailable. In those cases we have provided the Wald Chi-square statistic.

Except for market capitalization, we decided to use accounting-based measures of CFP since they would be more suitable to evaluate performance. However, accounting has the inherent disadvantage of being backward looking. With this reasoning as an argument, Gregory, Tharyan, and Whittaker (2014) suggest that market-based measures are more relevant when investigating the effect of CSR on performance. When comparing the results of the regressions of firm size on ESG both the accounting based and market based measure give interesting results. Both GLS regressions are significant with the respective $\beta$ values turning in opposite directions as shown in Table 5 and Table 15.

## 6.1 $H_1$: Larger firm size leads to higher ESG ratings

We investigate whether the size of firms determines how they will perform on ESG. We will conduct this using multiple regressions in order to arrive at a more robust result.

The regression formula is expressed as

$$Y_{ESG,it} = \beta_0 + \beta_1 X_{it} + C + \varepsilon_{it} \tag{2}$$

where $\varepsilon$ signifies the error term. The constant $C$ is a general place holder for control variables. $\beta_0$ is labeled Constant in Table 3 and $\beta_1$ is the coefficient.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | MC on ESGE | MC on ESGS | MC on ESGG | Market Cap on ESG |
| mktcap | 1.54e-05 | -1.20e-05 | -7.92e-06 | -6.48e-06 |
| | (2.03e-05) | (7.73e-06) | (1.12e-05) | (7.52e-06) |
| Constant | 14.59 | 33.56*** | 43.57*** | 32.51*** |
| | (9.537) | (5.353) | (6.379) | (4.515) |
| Observations | 6,020 | 6,020 | 6,020 | 6,020 |
| Number of id | 96 | 96 | 96 | 96 |
| R-Squared (O) | 0.462 | 0.556 | 0.226 | 0.242 |

Robust standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3: Market Capitalization (MC) on ESG scores.

As can be seen, the results are not statistically significant. However, when for robustness regressing *without* control for age, our coefficients for the main variable became significant. However, though enticing as it is, we present the fully controlled models. This is because anything else would constitute p-hacking, also called data dredging.

Hausman test failed to reject its null hypothesis, meaning random effects regression is appropriate for both OLS regressions, as seen in Table 3 and Table 4. The result of running equation (2) is shown in Table 3.

Firm size can be measured or considered in multiple ways. Either through market valuation or book value. Table 4 is similar to Table 3, that investigates market capitalization, in approach and controls, but looks at the book value of total assets. In both OLS regressions of size on ESG, the intercepts are very low compared to the mean values of ESG in Table 2.

| VARIABLES | (1)<br>TA on ESGE | (2)<br>TA on ESGS | (3)<br>TA on ESGG | (4)<br>TA on ESG |
|---|---|---|---|---|
| totassets | 4.93e-08 | -1.40e-05** | -1.32e-05 | -1.20e-05 |
| | (1.31e-05) | (6.47e-06) | (1.46e-05) | (9.47e-06) |
| Constant | 23.26** | 35.15*** | 47.06*** | 36.45*** |
| | (9.842) | (4.130) | (6.957) | (4.951) |
| Observations | 1,862 | 1,862 | 1,862 | 1,862 |
| Number of id | 93 | 93 | 93 | 93 |
| R-Squared (O) | 0.498 | 0.600 | 0.330 | 0.290 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 4: Total Assets (TA) regressed on ESG.

For robustness, we in Table 5 are using GLS on the aforementioned variables. Here, statistical significance has comparatively improved. Regression (2) is also statistically significant, but with small coefficients for both the market capitalization and total assets. The small coefficients can be partly explained by the large regressors. As shown in Table 2 the mean in market capitalization is 31 900 million and in total assets 28 200 million. This suggests that in average, market capitalization is associated with ESG by an increased score of 0.43 and total assets with a decreased score of 0.32. This is not in line with Drempetic, Klein, and Zwergel (2019) which found a positive significant correlation for ESG and both market capitalization and total assets. It is noteworthy that there are roughly one third of the amount of observations in total assets on ESG. This is due to the monthly ESG and market capitalization sampling while total assets are taken from quarterly reports.

| VARIABLES | (1)<br>TA on ESG | (2)<br>Market Cap on ESG |
|---|---|---|
| totassets | -1.13e-05*** | |
| | (3.74e-06) | |
| mktcap | | 1.36e-05*** |
| | | (2.15e-06) |
| Constant | 54.66*** | 54.10*** |
| | (0.857) | (0.519) |
| Observations | 1,862 | 6,020 |
| Number of id | 93 | 96 |
| Wald chi2 | 1842 | 5291 |

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5: ESG regressed on size variables using GLS.

## 6.2 $H_2$: Higher profitability leads to higher ESG ratings

For measuring profitability, we use two traditional measures: ROA and ROE. This has both advantages and disadvantages. The data on the two measures are widely available, are simple to interpret and compute, and their use are well established. A drawback is that the indications of these metrics can be tampered or be misleading by themselves. For that, measures such as Return on Net Operating Assets (RNOA) is used. However, this involves semantic judgment of the balance sheet in order to assess what should be considered operating assets. This is typically done by analysts, and hence RNOA also incorporates an element of judgment. Hence, ROE and ROA are imperfect, but knowledge of their perils and drawbacks can accommodate the appropriate use and treatment. Due to the size of our data set, we could not do an analysis of RNOA on all the firms. Market-based profitability measures such as Tobin's q have been used in earlier studies (Elsayed and Paton, 2005) but have not been taken into account in this thesis.

The regression formula is expressed as

$$Y_{ESG,it} = \beta_0 + \beta_1 X_{it} + C + \alpha_i + \varepsilon_{it} \tag{3}$$

where $\varepsilon$ signifies the error term. The constant $C$ is a general place holder for control variables. $\beta_0$ is labeled Constant in Table 6 and Table 7 and $\beta_1$ is the coefficient for ROA or ROE, respectively. The constant $\alpha_i$ is fixed effects, if any.

In Table 6, ROA is regressed on the various ESG scores. As according to a Hausman test, the regression is fixed effects.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | ROA on ESGE | ROA on ESGS | ROA on ESGG | ROA on ESG |
| roa | -0.0166 | -0.0142 | 0.0517 | 0.0137 |
| | (0.0259) | (0.0119) | (0.0499) | (0.0251) |
| Constant | 23.59** | 36.70*** | 48.43*** | 38.82*** |
| | (9.171) | (3.884) | (6.416) | (4.506) |
| Observations | 1,857 | 1,857 | 1,857 | 1,857 |
| Number of id | 93 | 93 | 93 | 93 |
| R-Squared (O) | 0.500 | 0.611 | 0.352 | 0.311 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6: ESG regressed on Return on Assets (ROA).

As can be seen, the impact of ROA on the various scores ranges from roughly -0.017 to 0.052, but more importantly, the results are not statistically significant. We do not find evidence that ROA affects ESG based on an OLS regression. One reason could be the relatively low amount of observations.

In Table 7, ROE is regressed on the various ESG scores. As according to a Hausman test, it is a random effects regression. Compared to Table 6, the regression of ROA in Table 7 has stronger statistical validity. Here we can see positive correlations between ROE and ESG-Governance as well as ROE on ESG that are statistically significant. Still, the coefficients are small, close to zero. Since the regressor itself is quite small as seen by its mean value in Table 2, the magnitude of these findings are small. In Table 7, column 4, we can see an intercept of less than 40, compared to the mean value of ESG of more than 50. This suggests that something else mainly explains this difference.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | ROE on ESGE | ROE on ESGS | ROE on ESGG | ROE on ESG |
| roe | 0.000106 | 0.000199 | 0.00398*** | 0.00184*** |
| | (0.000866) | (0.000472) | (0.00101) | (0.000570) |
| Constant | 24.28*** | 37.99*** | 49.01*** | 39.42*** |
| | (7.373) | (4.583) | (9.160) | (5.564) |
| Observations | 1,775 | 1,775 | 1,775 | 1,775 |
| Number of id | 90 | 90 | 90 | 90 |
| R-Squared (O) | 0.528 | 0.641 | 0.356 | 0.323 |

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 7: ESG regressed on Return on Equity (ROE).

For robustness, we in Table 8 are using GLS. As can be seen in the table, both regressions are statistically significant at $p < 0.01$. However, both report small coefficients, with ROE being very small.

Although we find statistically significant values for ROE and ROA on ESG scores while using the GLS approach, the effects are still somewhat small. Due to the fairly small amount of observations, resulting from profitability measures based upon quarterly earnings, a larger sample size would be interesting to use to investigate this relationship.

| | (1) | (2) |
|---|---|---|
| VARIABLES | ROA on ESG | ROE on ESG |
| roa | 0.0892*** | |
| | (0.0103) | |
| roe | | 0.00286*** |
| | | (0.000726) |
| Constant | 54.67*** | 55.25*** |
| | (0.840) | (0.835) |
| Observations | 1,857 | 1,775 |
| Number of id | 93 | 90 |
| Wald chi2 | 1951 | 1901 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8: ESG regressed on size variables using GLS.

### 6.3 $H_3$: Higher net income leads to high ESG ratings

In regression (4) in Table 9 we see a statistically insignificant result, when regressing net income on ESG score. The $R^2$ value is very low as well, confirming that little is explained by the model. Even in the case of taking the insignificant value seriously at that significance level, the coefficient is close to zero, essentially showing no affect of net income on ESG score. Further, the intercepts are considerably lower than the sample means which suggest that something else is affecting the results. This is especially apparent in column (1) in Table 9. The regression formula is:

$$Y_{ESG,it} = \beta_0 + \beta_1 X_{it} + C + \alpha_i + \varepsilon_{it} \tag{4}$$

where $\varepsilon$ signifies the error term. $C$ is a general place holder for control variables. $\beta_0$ is labeled Constant in Table 9 and $\beta_1$ is the coefficient for net income. $\alpha_i$ is the constant for fixed effects, if any.

| VARIABLES | (1)<br>NI on ESGE | (2)<br>NI on ESGS | (3)<br>NI on ESGG | (4)<br>NI on ESG |
|---|---|---|---|---|
| netincome | 0.000540 | 0.000195 | 0.000240 | 0.000231 |
| | (0.000823) | (0.000522) | (0.000633) | (0.000445) |
| Constant | 17.26*** | 38.46*** | 51.56*** | 39.24*** |
| | (5.122) | (2.268) | (3.616) | (2.345) |
| Observations | 1,867 | 1,867 | 1,867 | 1,867 |
| R-squared | 0.296 | 0.220 | 0.002 | 0.147 |
| Number of id | 93 | 93 | 93 | 93 |
| R-Squared (O) | 0.178 | 0.335 | 0.0246 | 0.0812 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 9: Net Income on ESG scores.

For robustness, we in Table 10 are using GLS. The results suggest that all regressions are statistically significant at a 0.01 confidence level. However, for all very small coefficients are reported. Still, the approach of GLS improved the statistical significance.

Conclusively we say that although net income does not seem to have a large effect on ESG scores, it does seem to have a somewhat positive effect on ESG scores. These results are in line with our hypothesis $H_3$: Higher net income leads to high ESG ratings. As net income is a function of total assets, which we have found a small but negative correlation with ESG for, and ROA, which we have found a small and positive correlation with ESG for, the GLS net income regression returns a rather non-surprising result. That is, a small, positive effect which is in line with the slack resources hypothesis (Mcguire, Sundgren, and Schneeweis, 1988; Waddock and Graves, 1997).

| VARIABLES | (1)<br>NI on ESGE | (2)<br>NI on ESGS | (3)<br>NI on ESGG | (4)<br>NI on ESG |
|---|---|---|---|---|
| netincome | 0.00338*** | 0.00158*** | 0.00183*** | 0.00208*** |
| | (0.000517) | (0.000273) | (0.000539) | (0.000338) |
| Constant | 55.36*** | 49.74*** | 57.88*** | 55.23*** |
| | (1.298) | (0.685) | (1.352) | (0.849) |
| Observations | 1,867 | 1,867 | 1,867 | 1,867 |
| Number of id | 93 | 93 | 93 | 93 |
| Wald chi2 | 4245 | 5083 | 1285 | 1899 |

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 10: ESG regressed on Net Income using GLS.

## 6.4 Concluding Remarks

Summarily, we combine the regressors in one regression, as seen in Table 15. The regression formula for this is expressed as

$$Y_{ESG,it} = \beta_0 + \beta_1 X_{mktcap,it} + \beta_2 X_{netincome,it} + \beta_3 X_{totassets,it} + \beta_4 X_{roa,it} + \beta_5 X_{roe,it} + C + \alpha_i + \varepsilon_{it} \quad (5)$$

where $\varepsilon$ signifies the error term. $C$ is a general place holder for control variables. $\beta_0$ is labeled Constant in Table 15 (in the appendix) and $\beta_N$ is the coefficient for the independent variables. $\alpha_i$ is the constant for fixed effects, if any.

Aggregating all the previously investigated independent variables into one GLS regression seen in Table 15 (in the appendix) highlights some of the characteristics shown earlier. The effects of market capitalization and total assets on ESG seem to almost cancel each other out. Both regressors show roughly a sevenfold increase in their effects, compared to Table 5. When congregating all independent variables, the number of observations decrease dramatically. This is mainly due to the quarterly sampling of some regressors, due to quarterly reporting, and to different missing variables for different firms. This means that the observations solely result from firms and periods for which the data set contains complete data. It is worth noting that when regressing all independent variables, net income and ROA are no longer significant. It is also apparent that a large degree of the ESG scores are determined outside of the model, considering the large intercept or $\beta_0$ value.

### 6.4.1 Effect Size

Statistical significance is rudimentary, but what must also be considered is the magnitude of the independent variable, known as effect size (Grace-Martin, 2020). One approach is normalizing into standard deviations; we decide to present it as an average effect of the regressor on ESG. This is presented in Table 11, below. It is calculated using the coefficients from GLS regressions, which are all statistically significant at $p < 0.01$. For instance, a one billion increase in total assets corresponds to a decrease in ESG with 0.0113.

| Variable | Average change | Change per unit | Unit |
|---|---|---|---|
| Total Assets | -0.3186 | -0.0113 | billions |
| Market Capitalization | 0.4342 | 0.0136 | billions |
| ROA | 0.2228 | 0.0892 | percentage points |
| ROE | 0.0114 | 0.0029 | percentage points |
| Net Income | 0.3914 | 0.0021 | millions |

Table 11: Effect sizes of the independent variables.

In Table 13 in appendix B a correlation matrix, also known as a kind of contingency table, for the variables is presented. The correlation between total assets and market capitalization is interesting with regards to their respective regressions, and especially the joint regression in Table 15 where they almost cancel each other out. This suggests that combined, size will not have much of an impact on ESG.

# 7 Conclusion

Although there are contradicting evidence and beliefs on the impact of sustainability on CFP, we early decided to give significant weight to the meta studies of Friede, Busch, and Bassen (2015) and Viehs, Clark, and Feiner (2014). Instead of investigating this thoroughly discussed topic, the direction of the causality was inverted which lead to the question of which financial characteristics positively affect ESG ratings. The main argument for why CFP would have an impact on sustainability is the slack resources hypothesis (Mcguire, Sundgren, and Schneeweis, 1988) which argue that firms with larger financial resources would be able to more heavily invest in sustainability enhancing activities. The most interesting characteristics were mainly deemed to be firm size and profitability, with total assets and market capitalization as proxies for size and ROE and ROA as profitability measures. We also looked at net income as a way to combine the two.

In retrospect, there are several opportunities for improvement with this thesis' approach. Due to limitations on data gathering we early decided to solely focus on the US firms in the Arabesque universe, representing roughly half the population. This limitation was not enough and in order to gather sufficient data for all characteristics examined, a dramatic decrease to 100 randomly selected US equities represented the final sample. Due to these limitations, a much more narrowly defined approach would have given the opportunity for more depth.

The formulas behind the utilized ESG ratings give its constituents different weight depending on sector which is why dissimilar characteristics in different sectors are rewarded. Furthermore, variables such as total assets and ROE are highly characterized by a firm's sector belonging. These reasons combined make a case for replicating the investigation while singling in on one or a limited number of sectors. This would enable a more thorough analysis. These issues were nevertheless addressed to some degree by controlling for sector affiliation.

As discussed by Tuppura et al. (2016), the decision as to which measure of CFP to use may vary among sectors investigated. This supports our decision to incorporate different measures of profitability as well as different measures of size. It further underlines the importance of controlling for sector when regressing CFP on ESG.

What is considered relevant effect size is highly circumstantial, subjective. However, broadly speaking we argue that the effects, as reported in Table 11, are sufficient for deciding about our hypotheses. One exception is our firm size hypothesis where we found contradicting results in the two measures. Due to these contradictions and the fact that they more or less cancel each other out, we are reluctant to confidently accept this hypothesis. From our gathering we cannot see indicators to a Type I or Type II error. Generally apart from the OLS regressions, the relevant regressions show statistical significance in several of the cases to the degree that we conclude these robustness tests have achieved a reliable answer.

| Hypothesis | Status |
|---|---|
| $H_0$: No financial characteristics affect ESG ratings | Rejected |
| $H_1$: Larger firm size leads to higher ESG ratings | Fail to accept |
| $H_2$: Higher profitability leads to higher ESG ratings | Accepted |
| $H_3$: Higher net income leads to high ESG ratings | Accepted |

Table 12: Summary of hypotheses.

Though controlling for reverse and simultaneous causality turned out to be beyond the scope of this thesis, we note that the longitudinal aspect, which is monthly stretching back to 2004, may improve the reasoning in this area. Though one academic said to us that no one cared about ESG about twenty years ago; precisely this fact might help rule out fads and trends, meaning that management possibly didn't reason about ESG for the purpose of opportunistic gains related to this. Another point is to disregard the theoretical correctness and have a pragmatic approach: maybe any possible simultaneous causality doesn't matter for a firm's management decisions.

Like we noted in Section 4, the multitude of data vendors as well as the various products offered by these firms, serve well for robustness checks and cross-validation by performing the analyses against different data sets. A take that centrally addresses simultaneous/reverse causality may also serve as an area of future research. Another area is to pick other independent variables that are innovative and revealing. This will, as in our case, largely be constrained by the availability from the data vendor. Other interesting areas for future research would be to narrow down the sample to only include firms that possess certain characteristics. Examples include but are not limited to size of firms, sectors or other geographies.

## References


Arabesque (2020). *S-Ray FAQs*. URL: https://www.arabesque.com/s-ray/faq/ (visited on 04/01/2020).

Arouri, Mohamed, Mathieu Gomes, and Kuntara Pukthuanthong (2019). "Corporate social responsibility and M&A uncertainty". In: *Journal of Corporate Finance* 56, pp. 176–198. ISSN: 0929-1199. DOI: https://doi.org/10.1016/j.jcorpfin.2019.02.002.

Barnett, Michael L. (2007). "Stakeholder Influence Capacity and the Variability of Financial Returns to Corporate Social Responsibility". eng. In: *The Academy of Management Review* 32.3, pp. 794–816. ISSN: 03637425.

Barnett, Michael L. and Robert M. Salomon (2012). "Does it pay to be really good? addressing the shape of the relationship between social and financial performance". eng. In: *Strategic Management Journal* 33.11, pp. 1304–1320. ISSN: 0143-2095.

Baumann-Pauly, Dorothée et al. (2013). "Organizing Corporate Social Responsibility in Small and Large Firms: Size Matters". eng. In: *Journal of Business Ethics* 115.4, pp. 693–705. ISSN: 0167-4544.

Beck, Cornelia, Geoffrey Frost, and Stewart Jones (2018). "CSR disclosure and financial performance revisited: A cross-country analysis". eng. In: *Australian Journal of Management* 43.4, pp. 517–537. ISSN: 0312-8962.

Berle, A. and G. Means (1932). *The Modern Corporation and Private Property*. Harcourt, Brace & World, New York, NY.

Bloomberg (2020). *Reference Data — Bloomberg Professional Services*. URL: https://www.bloomberg.com/professional/product/reference-data/ (visited on 04/23/2020).

Bowen, Howard (1953). *Social responsibilities of the businessman*. Harper.

Brammer, Stephen and Stephen Pavelin (2008). "Factors influencing the quality of corporate environmental disclosure". eng. In: *Business Strategy and the Environment* 17.2, pp. 120–136. ISSN: 0964-4733.

Brooks, Chris (2014). *Introductory Econometrics for Finance*. Cambridge University Press.

Carroll, Archie B. and Kareem M. Shabana (2010). "The Business Case for Corporate Social Responsibility: A Review of Concepts, Research and Practice". In: *International Journal of Management Reviews* 12.1, pp. 85–105. ISSN: 1460-8545.

Drempetic, S., C. Klein, and B. Zwergel (2019). "The Influence of Firm Size on the ESG Score: Corporate Sustainability Ratings Under Review". In: *Journal of Business Ethics*. ISSN: 01674544.

Drukker, David M. (2003). "Testing for Serial Correlation in Linear Panel-data Models". In: *The Stata Journal* 3.2, pp. 168–177. DOI: 10.1177/1536867X0300300206.

Elkington, John (1994). "Towards the sustainable corporation: Win-win-win business strategies for sustainable development". eng. In: *California Management Review* 36.2, p. 90. ISSN: 00081256.

Elsayed, Khaled and David Paton (2005). "The impact of environmental performance on firm performance: static and dynamic panel data evidence". eng. In: *Structural Change and Economic Dynamics* 16.3, pp. 395–412. ISSN: 0954-349X.

Englich, Frans (2020). *resetxt locks on small data set*. URL: https://www.statalist.org/forums/forum/general-stata-discussion/general/1545709-resetxt-locks-on-small-data-set (visited on 04/09/2020).

F Baum, Christopher and Mark E Schaffer (2013). *A general approach to testing for autocorrelation*. URL: https://www.stata.com/meeting/new-orleans13/abstracts/materials/nola13-baum.pdf (visited on 04/27/2020).

Fama, Eugene F. and Kenneth R. French (1992). "The Cross-Section of Expected Stock Returns". In: *Journal of Finance* 47.2, pp. 427–465. ISSN: 0022-1082.

Ferrell, Allen, Hao Liang, and Luc Renneboog (2016). "Socially responsible firms". In: *Journal of Financial Economics* 122.3, pp. 585–606.

Freeman, R Edward (1984). *Strategic management : a stakeholder approach*. Pitman.

Friede, Gunnar, Timo Busch, and Alexander Bassen (2015). "ESG and financial performance: aggregated evidence from more than 2000 empirical studies". In: *Journal of Sustainable Finance & Investment* 5.4, pp. 210–233. DOI: 10.1080/20430795.2015.1118917.

Global Compact, The (2004). *Who cares wins*. Tech. rep. The Global Compact.

Grace-Martin, Karen (2020). *A Comparison of Effect Size Statistics*. URL: https://www.theanalysisfactor.com/effect-size/ (visited on 05/05/2020).

Gregory, Alan, Rajesh Tharyan, and Julie Whittaker (2014). "Corporate Social Responsibility and Firm Value: Disaggregating the Effects on Cash Flow, Risk and Growth". eng. In: *Journal of Business Ethics* 124.4, pp. 633–657. ISSN: 0167-4544.

Hale, Jon (2019). *Sustainable Investing Interest Translating Into Actual Investments.* URL: https://www.morningstar.com/articles/952254/sustainable-investing-interest-translating-into-actual-investments (visited on 04/06/2020).

Halme, Minna and Morten Huse (1997). "The influence of corporate governance, industry and country factors on environmental reporting". eng. In: *Scandinavian Journal of Management* 13.2, pp. 137–157. ISSN: 0956-5221.

House, Companies (2020). *ARABESQUE ASSET MANAGEMENT HOLDING LIMITED.* URL: https://beta.companieshouse.gov.uk/company/08616956 (visited on 04/09/2020).

Klassen, Robert and Curtis Mclaughlin (Nov. 1996). "The Impact of Environmental Management on Firm Performance". In: *Management Science* 42, pp. 1199–1214. DOI: 10.1287/mnsc.42.8.1199.

Lazzaro, Carlo (2020). *Comment no. 2.* URL: https://www.statalist.org/forums/forum/general-stata-discussion/general/1547290-xtreg-re-vs-xtgls?p=1547295#post1547295 (visited on 04/27/2020).

Lee, Ki-Hoon, Beom Cheol Cin, and Eui Young Lee (2016). "Environmental Responsibility and Firm Performance: The Application of an Environmental, Social and Governance Model". In: *Business Strategy and the Environment* 25.1, pp. 40–53. ISSN: 0964-4733.

Li, Zhichuan and Caleb Thibodeau (2019). "CSR-Contingent Executive Compensation Incentive and Earnings Management". eng. In: *Sustainability* 11.12. ISSN: 20711050.

Maddala, G. S. and Kajal Lahiri (2009). *Introduction to Econometrics (Fourth ed.)* Chichester: Wiley, pp. 279–312. ISBN: 978-0-470-01512-4.

Manescu, Cristiana (2011). "Stock Returns in Relation to Environmental, Social and Governance Performance: Mispricing or Compensation for Risk?" eng. In: *SUSTAINABLE DEVELOPMENT, 2011, Vol. 19, Iss. 2, pp. 95-.118* 19.2, pp. 95–118.

Martinez-Ferrero, Jennifer and Jose Valeriano Frias-Aceituno (2015). "Relationship Between Sustainable Development and Financial Performance: International Empirical Research". In: *Business Strategy and the Environment* 24.1, pp. 20–39. ISSN: 0964-4733.

Mcguire, Jean, Alison Sundgren, and Thomas Schneeweis (1988). "Corporate Social Responsibility and Firm Financial Performance". eng. In: *Academy of Management Journal* 31.4, p. 854. ISSN: 0001-4273.

Melo, Tiago (2012). "Slack-resources hypothesis: a critical analysis under a multidimensional approach to corporate social performance". eng. In: *Social Responsibility Journal* 8.2, pp. 257–269. ISSN: 1747-1117.

Mishra, Subodh (2020). *ESG Matters*. URL: https://corpgov.law.harvard.edu/2020/01/14/esg-matters/ (visited on 04/06/2020).

NPower (2019). *Why ESG is replacing CSR and what this means to your business*. URL: https://www.npower.com/business-solutions/blog/2019/04/26/why-esg-is-replacing-csr-and-what-this-means-to-your-business/ (visited on 04/06/2020).

Orlitzky, Marc, Frank L Schmidt, and Sara L Rynes (2003). "Corporate Social and Financial Performance: A Meta-Analysis". eng. In: *Organization Studies* 24.3, pp. 403–441. ISSN: 0170-8406.

Ramsey, J. B. (1969). "Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 31.2, pp. 350–371. ISSN: 00359246.

Roberts, B.C. (1958). *Trade Union Government and Administration in Great Britain*. Harvard University Press.

Shapiro, Attorney General (2018). *Attorney General Shapiro Announces $575 Million 50-State Settlement with Wells Fargo Bank for Opening Unauthorized Accounts and Charging Consumers for Unnecessary Auto Insurance, Mortgage Fees*. URL: https://www.attorneygeneral.gov/taking-action/press-releases/attorney-general-shapiro-announces-575-million-50-state-settlement-with-wells-fargo-bank-for-opening-unauthorized-accounts-and-charging-consumers-for-unnecessary-auto-insurance-mortgage-fees/ (visited on 12/18/2018).

Shehata, Emad and Sahra Mickaiel (2015). *RESETXT: Stata Module to Compute Panel Data REgression Specification Error Tests (RESET)*.

Smith, H. Jeff (2003). In: *MIT Sloan Management Review*.

Tuppura, Anni et al. (2016). "Corporate social and financial performance in different industry contexts: the chicken or the egg?" eng. In: *Social Responsibility Journal* 12.4, pp. 672–686. ISSN: 17471117.

Verbeek, Marno (2017). *A Guide to Modern Eceonometric*. Wiley Custom.

Viehs, Michael, Gordon Clark, and Andreas Feiner (Sept. 2014). "From The Stockholder To The Stakeholder - How Sustainability Can Drive Financial Outperformance". In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.2508281.

Waddock, Sandra A. and Samuel B. Graves (1997). "THE CORPORATE SOCIAL PERFORMANCE, FINANCIAL PERFORMANCE LINK". eng. In: *Strategic Management Journal* 18.4, pp. 303–319. ISSN: 0143-2095.

Wickert, Christopher, Andreas Georg Scherer, and Laura J. Spence (2016). "Walking and Talking Corporate Social Responsibility: Implications of Firm Size and Organizational Cost". In: *Journal of Management Studies* 53.7, pp. 1169–1196. ISSN: 0022-2380.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press.

# A Code

```
%% Read fields.txt, one field per line, into ‘fields’.
ffields=fopen('fields.txt');
tline = fgetl(ffields);

fields = {};

while ischar(tline)
    fields = [fields; tline];
    tline = fgetl(ffields);
end

%% Read tickers.txt, one ticker per line, into ‘tickers’.
ftickers=fopen('tickers.txt');
tline = fgetl(ftickers);

tickers = {};

while ischar(tline)
    tickers = [tickers; tline];
    tline = fgetl(ftickers);
end

%% pull data
startDate   = datenum('01-Jan-2004');
endDate     = floor(now);

out_dir = 'H:/Frans/';

str_format = '%f,%f,%f,%f,%f,%f,%f,%f,%f,%f';

for j = 2:numel(tickers)
    tic = [tickers{j} ' US Equity'];
    disp(tic)
    [dt,val] = get_historical_data(tic,fields,startDate,endDate);

    fid = fopen([out_dir tic '.csv'],'w');
    fprintf(fid,'date');
    for k = 1:numel(fields)
        fprintf(fid,',%s',fields{k});
```

```
    end
    fprintf(fid,'\n');
    for k = 1:numel(dt)
        fprintf(fid,'%s,%f,%f,%f,%f,%f,%f,%f,%f,%f,%f\n',datestr(dt(k)),val(k,:));
    end
    fclose(fid);
end

%%

%
% fid = fopen([out_dir tic '.csv'],'r');
% data = textscan(fid,'%s%f%f%f%f%f%f%f%f%f%f','headerlines',1,'delimiter',',')
% fclose(fid);
```

Listing 1: Fetch data from Bloomberg (Matlab)

```

% This script requires Matlab 2019.

% This can likely be done in a better way, see Matlab's join(), outerjoin()
% and innerjoin(). Discussions at:
% https://se.mathworks.com/matlabcentral/answers/505997-searching-tables-for-modification

clear all;

from_Arabesque  = "../Data/us_only_monthly.csv";
tickers_CSV_dir = '../Data/Tickers/';
out_file        = "../Data/us_only_monthly_with_bb.csv";

% Debugging:
%from_Arabesque = "test.csv";
%tickers_CSV_dir = 'Test/';
%out_file        = "master_test.csv";

final_table = readtable(from_Arabesque);

%% First we need to add our columns to us_only. We do this by extracting
% the column names from an arbitrary CSV file, construct empty columns, and
% add them to us_only.

```

```
CSV_files = dir(strcat(tickers_CSV_dir, '*.csv'));

first_CSV_filename = CSV_files(1); % Arbitrary file.
first_CSV = readtable(strcat(tickers_CSV_dir, first_CSV_filename.name));

empty_column = zeros(height(final_table), 1);
empty_column(:) = NaN;

final_column_names = final_table.Properties.VariableNames;
first_CSV_column_names = first_CSV.Properties.VariableNames;

final_column_names_size = size(final_column_names);
size_before_append = final_column_names_size;
size_before_append = size_before_append(2);

% From 2 and - 1: we skip the first, which is the date.
for i = 2:length(first_CSV_column_names)
    final_column_names{1, final_column_names_size(2) + i - 1} ...
        = string(first_CSV_column_names(i));
    final_table = addvars(final_table, empty_column);
end

final_table.Properties.VariableNames = cellstr(final_column_names);

% We now have final_table, which has the dimensions and column names of the
% the final table we want. We now need to populate the cells with the
% actual data.

for i = 1:length(CSV_files)
    base_filename = CSV_files(i).name;
    full_filename = fullfile(tickers_CSV_dir, base_filename);

    % Strip away the ' US Equity.csv'
    ticker = string(extractBetween(base_filename, 1, ...
        length(base_filename) - 14));

    ticker_table = readtable(full_filename);

    locations = find(final_table.ticker == ticker);

    if isempty(locations)
```

```
        error("Don't have ticker %s in Arabesque's data set.", ticker);
    end

    for l = (locations')
        % We need to:
        % 1) Figure out the date for this row
        % 2) Find the row in ticker_table that has the data. Lookup based
        %    on date.
        % 3) Fetch the corresponding columns in ticker_table
        % 4) Insert the columns into final_table

        lookup_date = datetime(string(table2cell(final_table(l, 1))));

        % The location in ticker_table that has the data we want.
        ticker_row = [];

        % Since we sometimes don't have matching dates, we need to find the
        % closest one. We do this by looping over up to five upcoming days.
        for i = 0:5

            candidate = lookup_date + days(i - 1);
            ticker_row = find(datetime(ticker_table.date) == candidate);

            if ~isempty(ticker_row) % We found a matching date.
                break;
            end
        end

        if isempty(ticker_row)
            break; % We didn't find any data for this date.
        end

        % At location l, in final_table, we need to append the data after
        % size_before_append using the data found in ticker_table at
        % row row after the first column, which is the date.

        for i = 2:(length(first_CSV_column_names))
            cell = (table2cell(ticker_table(ticker_row, i)));
            val = double(string(cell));
            final_table{l, size_before_append + i - 1} = val;
        end
```

```
    end
end

writetable(final_table, out_file);
```

Listing 2: Merge data sources (Matlab)

```

clear all;

tic;

in_data_file = "../Data/us_only_monthly_with_bb.csv";
in_ipo_date_file = "../Data/IPO_date.csv";
out_file = "../Data/us_only_monthly_with_bb_with_age.csv";

final_table = readtable(in_data_file);

%% Add an empty column named 'age'.
empty_column = zeros(height(final_table), 1);
empty_column(:) = NaN;
final_table = addvars(final_table, empty_column);

column_names = final_table.Properties.VariableNames;
age_col_pos = size(column_names, 2);
column_names{1, age_col_pos} = 'age';
final_table.Properties.VariableNames = cellstr(column_names);

%%

ipo_dates = readtable(in_ipo_date_file);

%for i = 1:height(final_table)
i = 1;
while i <= height(final_table)

    current_ticker = strcat(string(final_table{i, 5}), ' US Equity');

    tbl_listing_date = ipo_dates(ipo_dates.ticker == current_ticker, 3);

    if isempty(tbl_listing_date)
        continue;
```

```
    end

    listing_date = datetime(tbl_listing_date{1, 1});
    % However, still might listing_date be NaT due to no entry in
    % ipo_dates. For those we will get NaN below.

    % Note, we let l = i, and update i afterwards.
    for l = i:height(final_table)

        timestamp = final_table{l, 1};
        duration = timestamp - listing_date;
        yrs = round(years(duration));

        new_ticker = strcat(string(final_table{l, 5}), ' US Equity');

        if new_ticker ~= current_ticker
            % It's a new one, we don't want to skip this entry, so do it
            % from the start.
            i = l;
            break;
        else
            i = i + 1;
            final_table{l, age_col_pos} = yrs;
        end
    end
end

writetable(final_table, out_file);
toc;
```

Listing 3: Control for firm age (Matlab)

```
close all;
clear all;

% master_monthly.csv is US only merged with tickers.
in_file = "../Data/master_monthly.csv";

table = readtable(in_file);
table = rmmissing(table);
esg = table.esg;

```

```


%%
% ESG histogram. Takes a long time, so let's easily disable it.
doHistograms = true;
if doHistograms
    tiledlayout(1, 2);

    nexttile;
    global_csv = "../../../../../Google Drive/Google Drive Master Thesis/Data from
     Arabesque/sray_260_historical_20200201.csv";
    global_table = readtable(global_csv);
    esg_global = global_table.esg;
    histogram(esg_global);
    xlabel("Arabesque's full universe.");

    nexttile;
    histogram(esg);
    xlabel("The subset used.");

    saveas(gcf, "../Figures/esg_histograms.eps", 'epsc')

    % We want a new figure now.
    hold off;
end


%%
% Scatter plot of overview.
tiledlayout(3,2);

nexttile;
% We have two enormous negative outliers, we delete those.
rtn_table = table;

% Remove the string and date variables because rmoutliers() cough on those.
rtn_table(:, 1:9) = [];

rtn_table = rmoutliers(rtn_table);
scatter(rtn_table.RETURN_COM_EQY, rtn_table.esg);
xlabel("Return on Equity (ROE), %");
```

```
nexttile;
scatter(rtn_table.RETURN_ON_ASSET, rtn_table.esg);
xlabel("Return on Assets (ROA), %");


nexttile;
scatter(table.NET_INCOME, esg);
xlabel("Net Income");

nexttile;
scatter(table.TOT_MKT_VAL, esg);
xlabel("Total Market Value");

nexttile;
scatter(table.BS_TOT_ASSET, esg);
xlabel("Total Assets");

nexttile;
scatter(table.TOTAL_EQUITY, esg);
xlabel("Total Equity");

saveas(gcf, "../Figures/scatter_overview.eps", 'epsc');



%%
% Produce log returns of WFC.

in_file = "../Data/wfc_daily.csv";

table = readtable(in_file);

dates = table.date;

esg     = str2double(table.esg);
esge    = str2double(table.esg_e);
esgs    = str2double(table.esg_s);
esgg    = str2double(table.esg_g);

dates = table.date(2:size(table.date));
```

```

ret_esg      = diff(log(esg));
ret_esge     = diff(log(esge));
ret_esgs     = diff(log(esgs));
ret_esgg     = diff(log(esgg));

figure;
plot(dates, ret_esg);
hold on;

plot(dates, ret_esge);
plot(dates, ret_esgs);
plot(dates, ret_esgg);

legend("ESG", "ESG Environment", "ESG Social", "ESG Governance");
hold off;

saveas(gcf, "../Figures/wfc_esg_log_returns.eps", 'epsc')
```

Listing 4: Produce figures (Matlab)

```
clear all

* This script imports the CSV file of Arabesque's data, and finalizes it.

* Note:
* 1. You need to change the working directory (see the File menu) to where the
*    the CSV-file is.
**
* 2. Stata got a 32-characters limit for variable names.


* NOTE: The working directory (see the File menu) needs to be of where the data
*       file is.

import delimited "us_only_monthly_with_bb_with_age.csv"


* Convert to numbers.
destring esg esg_e esg_s esg_g gc gc_hr gc_lr gc_en gc_ac, replace
destring wacc_cost_equity return_com_eqy return_on_asset net_income
    tot_return_index_net_dvds tot_mkt_val bs_tot_asset total_equity, replace
```

```
* We need a unique ID but the SEDOLs contains alphanumerics, so we generate an
* id based on that.
egen id = group(sedol)

* For controlling.
tabulate economic_sector, gen(g_economic_sector)
tabulate industry, gen(g_industry)

* We drop in order to avoid the dummy variable trap.
drop g_economic_sector18
drop g_industry52

* Change datatype for our date.
gen date2 = date(date, "YMD")
drop date
rename date2 date
format date %td

* This is for the file as a whole.
label data "Data set of ESG ratings and related data."

* Single underlines doesn't play well with Latex, so avoid that.
rename esg_e esge
rename esg_s esgs
rename esg_g esgg
rename net_income netincome
rename tot_mkt_val mktcap
rename bs_tot_asset totassets
rename return_on_asset roa
rename return_com_eqy roe

* Set descriptions for our variables.
label variable esg        "ESG"
label variable esge       "ESG-E"
label variable esgs       "ESG-S"
label variable esgg       "ESG-G"
label variable netincome    "Net Income"
label variable totassets    "Total Assets"
label variable roe        "ROE"
```

```
label variable mktcap      "Market Cap"
label variable roa         "ROA"

* Set up our panel data.
xtset id date
tset id date

* Write out.
save us_only_monthly_with_bb_with_age.dta, replace
```

Listing 5: Export to Stata (Stata)

```
clear all

* NOTE: The working directory (see the File menu) needs to be of where the data
*       file is.


* ssc install outreg2
* ssc install outtable
* ssc install corrtex

local esg_variables "esg esge esgs esgg"

* The other variables that we have, the rest.
local rest_variables "netincome totassets mktcap roa roe"

* OLD: local controls "i.grp_economic_sector i.grp_industry"
local controls "g_economic_sector* g_industry* age"

* Open our data set.
use "Data/us_only_monthly_with_bb_with_age.dta"


* --------------------- Co-variance ------------------------
corrtex `esg_variables' `rest_variables', landscape file(Tables/corrtable.tex)
    replace



* --------------------- All Summaries ------------------------
* From https://roastata.wordpress.com/2008/09/30/export-summary-statistics-to-
```

```
latex/
* quietly: tabstat VARLIST, c(s) stat(mean) save
tabstat `esg_variables' `rest_variables', c(s) stat(count mean sd var min max) save

matrix output = r(StatTotal)'

outtable using Tables/descriptive_all_stats, mat(output) center replace ///
caption("Descriptive Statistics of all variables") nobox label ///
clabel(table_descriptive_all_stats)


* --------------------- Regressions ------------------------

* ------------ Base regression ------------------
local all_summarized "mktcap netincome totassets roa roe"
xtgls esg `all_summarized' `controls', i(id)
outreg2 using Tables/base_regression.tex, keep(`all_summarized') replace ctitle(ESG) tex(fragment) addstat(Wald chi2, e(chi2))

xtreg esg `all_summarized', fe
estimate store fe

xtreg esg `all_summarized', re
estimate store re

hausman fe re


* ------------ Size of Firm ------------------
* ---------- mktcap ---------------
xtreg esge mktcap `controls', i(id) re robust
outreg2 using Tables/esg_mktcap.tex, replace keep(mktcap) ctitle(MC on ESGE) tex(fragment) addstat(R-Squared (O), e(r2_o))

xtreg esgs mktcap `controls', i(id) re robust
outreg2 using Tables/esg_mktcap.tex, append keep(mktcap) ctitle(MC on ESGS) tex(fragment) addstat(R-Squared (O), e(r2_o))

xtreg esgg mktcap `controls', i(id) re robust
outreg2 using Tables/esg_mktcap.tex, append keep(mktcap) ctitle(MC on ESGG) tex(
```

```
    fragment) addstat(R-Squared (O), e(r2_o))

* Copied to below
xtreg esg mktcap `controls', i(id) re robust
outreg2 using Tables/esg_mktcap.tex, append keep(mktcap) ctitle(Market Cap on ESG)
     tex(fragment) addstat(R-Squared (O), e(r2_o))

* Hausman test. If null is not rejected, random effects is appropriate.
xtreg esg mktcap, fe
estimate store fe

xtreg esg mktcap, re
estimate store re

hausman fe re

* ---------- totassets ---------------
xtreg esge totassets `controls', i(id) re robust
outreg2 using Tables/esg_totassets.tex, replace keep(totassets) ctitle(TA on ESGE)
     tex(fragment) addstat(R-Squared (O), e(r2_o))

xtreg esgs totassets `controls', i(id) re robust
outreg2 using Tables/esg_totassets.tex, append keep(totassets) ctitle(TA on ESGS)
    tex(fragment) addstat(R-Squared (O), e(r2_o))

xtreg esgg totassets `controls', i(id) re robust
outreg2 using Tables/esg_totassets.tex, append keep(totassets) ctitle(TA on ESGG)
    tex(fragment) addstat(R-Squared (O), e(r2_o))

* Copied to below
xtreg esg totassets `controls', i(id) re robust
outreg2 using Tables/esg_totassets.tex, append keep(totassets) ctitle(TA on ESG)
    tex(fragment) addstat(R-Squared (O), e(r2_o))

* Hausman test. If null is not rejected, random effects is appropriate.
xtreg esg totassets, fe
estimate store fe

xtreg esg totassets, re
estimate store re

```

```
hausman fe re

* ---------------------- START size GLS ---------------------------
xtgls esg totassets `controls', i(id)
outreg2 using Tables/esg_gls_size.tex, replace keep(totassets) ctitle(TA on ESG)
    tex(fragment) addstat(Wald chi2, e(chi2))

xtgls esg mktcap `controls', i(id)
outreg2 using Tables/esg_gls_size.tex, append keep(mktcap) ctitle(Market Cap on
    ESG) tex(fragment) addstat(Wald chi2, e(chi2))


* ---------------------- END size GLS ---------------------------


* ---------- roa ---------------
xtreg esge roa `controls', i(id) re robust
outreg2 using Tables/esg_roa.tex, replace keep(roa) ctitle(ROA on ESGE) tex(
    fragment) addstat(R-Squared (O), e(r2_o))

xtreg esgs roa `controls', i(id) re robust
outreg2 using Tables/esg_roa.tex, append keep(roa) ctitle(ROA on ESGS) tex(
    fragment) addstat(R-Squared (O), e(r2_o))

xtreg esgg roa `controls', i(id) re robust
outreg2 using Tables/esg_roa.tex, append keep(roa) ctitle(ROA on ESGG) tex(
    fragment) addstat(R-Squared (O), e(r2_o))

* Copied to below
xtreg esg roa `controls', i(id) re robust
outreg2 using Tables/esg_roa.tex, append keep(roa) ctitle(ROA on ESG) tex(fragment
    ) addstat(R-Squared (O), e(r2_o))

* Hausman test. If null is not rejected, random effects is appropriate.
xtreg esg roa, fe
estimate store fe

xtreg esg roa, re
estimate store re

hausman fe re
```

```
* Result: Prob>chi2 =      0.3353

* ---------- roe ---------------
xtreg esge roe `controls', i(id)
outreg2 using Tables/esg_roe.tex, replace keep(roe) ctitle(ROE on ESGE) tex(
    fragment) addstat(R-Squared (O), e(r2_o))

xtreg esgs roe `controls', i(id)
outreg2 using Tables/esg_roe.tex, append keep(roe) ctitle(ROE on ESGS) tex(
    fragment) addstat(R-Squared (O), e(r2_o))

xtreg esgg roe `controls', i(id)
outreg2 using Tables/esg_roe.tex, append keep(roe) ctitle(ROE on ESGG) tex(
    fragment) addstat(R-Squared (O), e(r2_o))

* Copied to below
xtreg esg roe `controls', i(id)
outreg2 using Tables/esg_roe.tex, append keep(roe) ctitle(ROE on ESG) tex(fragment
    ) addstat(R-Squared (O), e(r2_o))

* Hausman test. If null is not rejected, random effects is appropriate.
xtreg esg roe, fe
estimate store fe

xtreg esg roe, re
estimate store re

hausman fe re
* Result: Prob>chi2 =      0.0000


* ---------------------- START profit GLS ---------------------------
xtgls esg roa `controls', i(id)
outreg2 using Tables/esg_gls_profit.tex, replace keep(roa) ctitle(ROA on ESG) tex(
    fragment) addstat(Wald chi2, e(chi2))

xtgls esg roe `controls', i(id)
outreg2 using Tables/esg_gls_profit.tex, append keep(roe) ctitle(ROE on ESG) tex(
    fragment) addstat(Wald chi2, e(chi2))


```

```
* ---------------------- END profit GLS --------------------------


* ----------- Net Income -----------------
xtreg esge netincome `controls', i(id) fe robust
outreg2 using Tables/esg_netincome.tex, replace keep(netincome) ctitle(NI on ESGE) tex(fragment) addstat(R-Squared (O), e(r2_o))

xtreg esgs netincome `controls', i(id) fe robust
outreg2 using Tables/esg_netincome.tex, append keep(netincome) ctitle(NI on ESGS) tex(fragment) addstat(R-Squared (O), e(r2_o))

xtreg esgg netincome `controls', i(id) fe robust
outreg2 using Tables/esg_netincome.tex, append keep(netincome) ctitle(NI on ESGG) tex(fragment) addstat(R-Squared (O), e(r2_o))

xtreg esg netincome `controls', i(id) fe robust
outreg2 using Tables/esg_netincome.tex, append keep(netincome) ctitle(NI on ESG) tex(fragment) addstat(R-Squared (O), e(r2_o))


* ---------------------- START NI GLS --------------------------
xtgls esge netincome `controls', i(id)
outreg2 using Tables/esg_gls_ni.tex, replace keep(netincome) ctitle(NI on ESGE) tex(fragment) addstat(Wald chi2, e(chi2))

xtgls esgs netincome `controls', i(id)
outreg2 using Tables/esg_gls_ni.tex, append keep(netincome) ctitle(NI on ESGS) tex(fragment) addstat(Wald chi2, e(chi2))

xtgls esgg netincome `controls', i(id)
outreg2 using Tables/esg_gls_ni.tex, append keep(netincome) ctitle(NI on ESGG) tex(fragment) addstat(Wald chi2, e(chi2))

xtgls esg netincome `controls', i(id)
outreg2 using Tables/esg_gls_ni.tex, append keep(netincome) ctitle(NI on ESG) tex(fragment) addstat(Wald chi2, e(chi2))


* ---------------------- END NI GLS --------------------------

```

```
* ------------ ESG sub-scores ------------------
xtreg esg esge esgg esgs, i(id) fe robust
outreg2 using Tables/esg_subsets.tex, replace ctitle(ESG and sub-scores) tex(
    fragment)
```

Listing 6: Export tables (Stata)

# B Additional Tables

| | (1) |
|---|---|
| VARIABLES | ESG and sub-scores |
| | |
| esge | 0.238*** |
| | (0.0101) |
| esgg | 0.450*** |
| | (0.00780) |
| esgs | 0.314*** |
| | (0.0135) |
| Constant | -0.0824 |
| | (0.629) |
| | |
| Observations | 7,147 |
| Number of id | 100 |
| R-squared | 0.985 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 14: ESG score and its relation to subsets.

| | **ESG** | **ESG-E** | **ESG-S** | **ESG-G** | **Net Income** | **Total Assets** | **Market Cap** | **ROE** | **ROA** |
|---|---|---|---|---|---|---|---|---|---|
| **ESG** | 1.000 | | | | | | | | |
| **ESG-E** | 0.713 | 1.000 | | | | | | | |
| **ESG-S** | 0.671 | 0.789 | 1.000 | | | | | | |
| **ESG-G** | 0.660 | -0.007 | -0.023 | 1.000 | | | | | |
| **Net Income** | 0.190 | 0.167 | 0.282 | 0.030 | 1.000 | | | | |
| **Total Assets** | -0.025 | 0.080 | 0.208 | -0.176 | 0.403 | 1.000 | | | |
| **Market Cap** | 0.071 | 0.132 | 0.261 | -0.103 | 0.539 | 0.952 | 1.000 | | |
| **ROA** | 0.244 | 0.109 | 0.137 | 0.216 | 0.238 | 0.005 | 0.055 | 1.000 | |
| **ROE** | 0.132 | 0.053 | 0.070 | 0.124 | 0.084 | 0.016 | 0.033 | 0.238 | 1.000 |

Table 13: Correlation matrix of variables.

| | (1) |
|---|---|
| VARIABLES | ESG |
| | |
| mktcap | 8.30e-05*** |
| | (1.06e-05) |
| netincome | -0.000130 |
| | (0.000597) |
| totassets | -9.28e-05*** |
| | (1.03e-05) |
| roa | 0.0268 |
| | (0.0198) |
| roe | 0.00178*** |
| | (0.000681) |
| Constant | 55.58*** |
| | (0.961) |
| | |
| Observations | 1,046 |
| Number of id | 84 |
| Wald chi2 | 1560 |

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 15: Result of base regression using GLS.