Tradable Analytics - Trading ESG

Michel Crouhy

michel.crouhy-ext@natixis.com, NATIXIS

Dan Galai

dan.galai@huji.ac.il, The Hebrew University of Jerusalem

Aner Ravon

aner@zirra.com, Zirra Ltd., Israel

Zvi Wiener

zvi.wiener@huji.ac.il, The Hebrew University of Jerusalem (corresponding author)

Preliminary draft, please do not quote

January 2022

Abstract: ESG (Environment, Social and Governance) criteria become a relevant factor in the investment universe. We develop an AI-based algorithm that uses public data, mainly Web-based information, to assign E, S and G ratings to companies. Using our scoring procedure, we construct portfolios, comprising 50 firms each from the S&P 500 index, 50 firms with the highest scores and 50 with the lowest scores for 4 scoring categories: ESG, E, S and G, for the years 2018-2021. We find that, with the exception of 2021, high-score portfolios consistently outperform low score portfolios. We also observe that the shares of high G-score companies outperform low G-score portfolios. The data support the hypotheses that indicators of good corporate governance can identify better performing firms. We also note the outperformance of high S-rated portfolios in 2018-2020. In 2021, this changed, as the lowest S-rated portfolio outperformed both the index and the highest S-rated portfolio. We find that the E portfolios behave differently from the S and G portfolios and note that usually substantial investment is required to improve the E score, which is not the case for either the S or G scores.

Keywords: ESG, SRI, AI, investing, portfolio selection

JEL Codes: G11, G12, G14, G17, G19

Research support from the European Institute of Finance is gratefully acknowledged. We thank Philippe Jeanne from Natixis and June Dilevsky for helpful comments. We would like to thank the participants of the International Risk Management Conference 2021 in Cagliari and the Finance and Banking seminar at the Institute of Economics and Finance at the University of Szczecin for their comments. Wiener thanks the Krueger Center, and the Sanger Family Chair in Banking and Risk Management for support of this research. Galai thanks the Zagagi Center for support of this research. We thank Zirra Ltd., an Israeli FinTech company that developed the data analytics and ratings that were used in this study.

Tradable Analytics - Trading ESG

1. Introduction

The global Coronavirus pandemic that erupted in early 2020 has had major economic consequences, among them an immediate worldwide recession that has since lingered in most developed economies. (See, e.g., April 2020 IMF report.) The business models of numerous companies were affected, some irreversibly. Post-crisis market conditions will differ from pre-crisis times, and companies will be compelled to adapt to a new environment.

One consequence is that accounting data, including the annual reports for 2019 and quarterly reports for the end of March 2020, are virtually irrelevant as far as shaping projections for the future. Since accounting information is historical, ex-post, and economic conditions have changed drastically, in many cases, one cannot use past profitability data, or even sales trends, to extrapolate into the future. Hence, analysts and economists, who follow trends in the corporate world, urgently need to employ alternative sources of data. The major source for alternative data is undoubtedly the Internet, though data taken from the Web are usually unstructured, come from multiple sources within and outside the company, and usually require validation.

In parallel, over the past five years we have witnessed an increasing trend in the corporate world to follow emerging rules, standards and conventions related to ESG. ESG stands for environmental, social and governance best practices, which, while finding institutional expression, are still vaguely defined. The term ESG was officially coined with the publication of the 2004 UN Global Compact Initiative report "Who Cares Wins". It sets the ambitious goal of redefining and regrouping the ethical finance into three pillars: environmental, social and governance. Several mutual and hedge funds have since been established promising to select companies that adhere to ESG principles.

There is preliminary evidence that the ESG companies perform better than non-ESG firms, as echoed in the slogan "doing well by doing good". An increasing number of the investors worldwide integrate ESG considerations in their analysis of risks and opportunities. It is expected, therefore, that ESG equity funds investing in companies incorporating ethical business practices, positive treatment of employees, lower carbon emissions, and/or good corporate governance tend to outperform other funds. The academic literature on ESG funds however, does not necessarily confirm this expectation. As discussed below, numerous academic studies have been conducted, but do not provide a clear answer as to whether investment strategies built around ESG do in fact generate abnormal positive returns.

Aggregate assets under management (AUM) of funds with an explicit ESG policy is approaching $1 Trillion. According to Morningstar, in the first nine months of 2021, there was an inflow of approximately $ 55 billion into ESG funds, which is almost twice the amount during the respective period in 2020, and close to 10% greater than for the entire 2020 year. More than 70% of the ESG funds, across all asset classes, performed better than their counterparts since the beginning of 2020.[1]

The three ethical pillars of ESG (U.N. Global Compact Initiative, 2004) include:

The ***environmental pillar*** focuses on issues such as climate change, deforestation, air and water quality, land exploitation and biodiversity. In evaluating this pillar the efforts of a company in terms of energy efficiency, greenhouse gas emission, waste management, responsible water consumption and resource management should be considered.

The ***social pillar*** aims to measure the direct and indirect impact of a company's activity on stakeholders (customers, employees, stockholders, bondholders, local communities and society in general) and includes the rating of gender policies, human rights protection, labor standards, workplace and product safety, public health, income distribution, and other factors affecting employee satisfaction and welfare.

[1] Stankiewicz, Alyssa "The Number of New Sustainable Funds Hits an All-Time Record", *Morningstar* https://www.morningstar.com/articles/1062299/the-number-of-new-sustainable-funds-hits-an-all-time-record (accessed 6/1/2021).

The ***governance pillar*** relates to factors such as independence of the board, shareholder rights, executive compensation, control procedures, anti-competitive practices, business ethics, fraud, and respect for the law and regulations.

The ESG lens can help identify risks not detected by conventional risk analysis, e.g., environmental risks that could impact financial performance due to operational or litigation costs. Overall, ESG risks affect various industries differentially; corporate governance, however, is universal. For example, the water usage could determine whether a mining company remains profitable in the future. This is less relevant for the finance industry where cyber security is critical. In the past, companies that failed to manage ESG risks have historically experienced higher costs of capital, greater earnings volatility, and more frequent accounting irregularities.

The impact of ESG on portfolio performance is mixed. An early study by Kempf and Osthoff (2007) using data from the KLD rating agency demonstrates that the strategy of buying stocks with high socially responsible ratings and shorting stocks with low socially responsible ratings leads to high abnormal returns. In contrast, Kurtz and diBartolomeo (2011) show that the relative performance, positive and negative), of social investment with respect to the S&P 500 is explained by factor exposures. After adjusting for risk, the impact of social factors is negligible. (See also Bruno et al., 2021.)

Demers et al. (2020) find that ESG scores offer no positive explanatory power for returns during the initial period of the COVID-19 pandemic in 2020, after the firm's industry affiliation and market-based measures of risk have been accounted for and controlled. They conclude that ESG portfolios did not provide a good hedge against the market impact of COVID-19. They also show that cash position, debt level and the amount of "internally generated intangible assets (R&D, IT, employee training…) are more important determinants of stock price resilience.

Also, Billio et al. (2020), show that non-ESG portfolios performed better than ESG portfolios (i.e., higher Sharpe and Sortino ratios) prior to 2005. Post 2005, following the introduction of the ESG concept, both portfolios performed similarly. Lo and Zhang (2021) derive conditions under which impact investing, such as ESG investing, detracts from or improves on performance relative to that of traditional mean-variance optimal portfolios.

Other studies offer various explanations of "green returns" as a separate risk factor (Pastor et. al., 2021) or within the framework equilibrium models Avramov et al. (2021). A series of recent papers by Lee et. al. (2020), Giese et.al. (2021) and Lee (2021), researchers from MSCI, deconstruct the performance of ESG rating by time horizon, sector and weighting and financial and social objective in assessing the performance of investment portfolios.

Our research focuses on employing alternative sources of financial and non-financial information to analyze the performance of ESG corporations. Our objective is to identify changes in trends and preferences in real time, rather than wait for new accounting reports, and to measure whether ESG firms perform better than non-ESG firms. Also, we empirically analyze which indicators are best suited to differentiate ESG from non-ESG firms. The E, S and G are analyzed separately to discover which, if any, weigh more in identifying better performing firms.

We use artificial intelligence (AI) and quantitative models to generate more precise corporate analysis for the S&P 500 companies. Accordingly, we develop consolidated and comprehensive data sets from complex structured and unstructured public sources, and through AI-powered event classifiers, we articulate powerful scoring algorithms. This provides an additional layer of deep and diverse alternative data sets for investment professionals seeking actionable insights and trading signals to complement their trading and investment strategies. We combine this AI analysis with advanced and innovative dynamic portfolio selection models to test the risk- adjusted performance of our portfolios relative to market benchmarks. We employ close to 200 variables related to E, S and G from web-based data as well as data provided by the company to score and rank each of the S&P 500 firms. The performance of the top-ranked 50 companies is compared to that of the bottom-ranked 50 companies, and the two portfolios are then compared to the performance of S&P 500 index. We expect the portfolio with the 50 best scores to outperform the Index, and expect the second portfolio of the worst 50 scores to have a lower performance.

2. The change of paradigm

The traditional neoclassic economic view holds that the maximization of shareholder profits is the primary objective of a company as posited by Milton Friedman from the University of Chicago. ESG was perceived as a potential burden and the source of unnecessary costs for the firm, which will likely impinge the shareholder interests. For the last twenty years, however, socially responsible investment (SRI) has taken hold and gained traction among investors[2]. Investor interest in SRI has since been followed by the development of ESG ratings, ESG indices and a growing inflow of investor money to ESG funds.

The shift in approach is also a reaction to events occurring in the last 20 years. Series of corporate scandals, including Enron, WorldCom, and Parmalat cast light on the perils of insufficient corporate governance. These scandals, led to huge financial losses and in some cases, bankruptcy (of the companies and their auditors). They also led to new and stricter corporate legislation in the US, which passed the 2002 Sarbanes-Oxley Act, a suite of governance measures, which requires, *inter alia*, that the CEO and CFO of publicly traded companies sign the annual financial report and affirm that it reflects all material information and that the financial numbers are correct.

The financial crisis of 2007-2009 raised public consciousness of social responsibility, and stressed the importance of good governance. Banks and other financial institutions invested in and issued complex and highly levered, x financial instruments, very often related to the real estate market, which had experienced accelerated growth, fuelled in part by questionable and at times, predatory mortgage lending practices. Many of these instruments, which were backed by low-grade debt, were highly rated by the rating agencies. Once the underlying real estate market collapsed, many of these instruments unravelled, and the financial institutions issuing and trading them went into bankruptcy.

The introduction of Basel II (especially Pillar II), in 2006, and Solvency II, requiring financial institutions to assess the risk associated with governance, also

[2] Socially responsible investment (SRI), sustainable and responsible investment (SRI), corporate social responsibility (CSR) and ESG are now used fairly interchangeably.

affected the move toward corporate ESG awareness and acceptance. The E in the ESG has also been enhanced by greater awareness of global warming, punctuated by what appears to be increasingly frequent extreme weather events, such as tsunamis, hurricanes, storms and wildfires, which in some cases devasted communities and demolished infrastructures.

Finally, activists emphasize the importance of environmental, social and governance issues in corporate strategy and the need to disclose progress in addressing them. The Sustainable Stock Exchanges Initiative reports that as of the end of 2021, 63 exchanges have issued written ESG guidance documents. The list of exchanges includes the New York Stock Exchange, Nasdaq, the London Stock Exchange, the Euronext exchanges, Deutsche Börse, and the Shanghai, Shenzhen and Hong Kong exchanges in China.[3] Larry Fink the chairman of Blackrock announced his firm's commitment to ESG compliant companies.

Interest in ESG is also reflected in the investment in private equity funds, mutual funds and ETFs specializing in ESG, leading to the proliferation of a new segment of the rating agencies. Unlike credit ratings, ESG measurement is somewhat nebulous given the lack of common definitions, reporting standards and shared characteristics among each ESG component and across rating providers. Consequently, we observe major disagreements in ESG ratings across ratings agencies and services. This is a key finding of several recent studies conducted by international institutions.

In 2020 the OECD issued a report "ESG Investing: Practices, Progress and Challenges", see Boffo and Patalano (2020). The key finding of their study is that ESG ratings vary strongly depending on the rating agency. This occurs for a number of reasons, such as different methods, measures, key indicators and metrics, data use, qualitative judgment, and the weighting of the subcategories. Returns have shown mixed results over the past decade, raising questions as to the true extent to which ESG drives positive performance. In November 2021, the International Organization of Securities Commissions (IOSCO) issued a report entitled "Environmental, Social and Governance (ESG) Ratings and Data Products Providers" which deals with the

[3] https://sseinitiative.org/esg-guidance-database/

use of ESG ratings by investors. Among other findings, the report states that "there is a lack of transparency about the methodologies underpinning these ratings or data products…", and "while there is wide divergence within the ESG ratings and data products industry, there is an uneven coverage of products offered…" and "there may be concerns about the management of conflicts of interest…" (IOSCO,2021, p. 1).

The leading ESG rating organizations include: KLD (MSCI Stats), Sustainalytics (Morningstar), Vigeo Eiris (Moody's), RobecoSAM (S&P Global), Asset4 (Refinitiv), ISS OEKOM (Germany), Bloomberg (US), FTSE RUSSELL (UK), and ECPI (Italy). Recently some exchanges started providing links to the ESG-related databases on public companies, see for example the ESG Data Hub at Nasdaq, https://www.nasdaq.com/solutions/nasdaq-esg-hub.

Ratings consist of three basic elements:

- ***Scope***: attributes, which together define the concept, e.g., carbon emission, water consumption, labor practices, gender equality, diversity, etc.
- ***Indicators*** or metrics used to measure the score of a given attribute, e.g., gender equality could be measured by the percentage of women on the board, or by the gender pay gap within the workforce.
- ***Weighting*** divergence, since rating agencies take different views on the relative importance of the specific attributes

The different rating systems apply various attributes, indicators and weights to create their list of ESG firms. It is not surprising that one finds very low correlations between ratings: it is 0.54 on average (ranging between 0.38-0.71). Sometimes opposite ratings are found for the same companies by different rating organizations. Berg et al. (2019) find that measurement divergence across rating systems is the most important reason that ratings diverge. Monk and Rook (2021) describe the barriers encountered in ESG measurement and argue that investors over-rely on third parties for ESG analysis.

Given the hype surrounding ESG, the lack of standards and the scope of public investment in ESG funds, the SEC is currently investigating possible mislabeling and "greenwashing" by securities issuers, investment advisors and fund managers in the U.S. Having established a Climate and ESG Task Force, in March 2021, the SEC has asked investment advisors to describe in detail the screening process they use to ensure assets are worthy of ESG designations. It is investigating the DWS Group, Deutsche Bank's asset-management firm following statements by the firms former head of sustainability that the firm overstated its use of sustainability criteria in its asset management practices.

Our proposed scoring system translates all the data related to a defined set of companies into relative and absolute scores. This scoring model is based on categories of signals, which comprise various data parameters, related to ESG, as defined in the data dictionary. The model is based on vector dynamics and the BERT Natural Language Processing (NLP) algorithm which categorizes articles by topics and determines whether the data is positive or negative. Weights are assigned to each data category.

The scoring is relative and it considers the change in the position of a company relative to its peers and its momentum. The relative scoring shows which companies rank highest within their respective universes. The relative scoring is used as an additional input for stock selection and portfolio construction. We measure each one of the factors E, S and G independently, and rank their relative strength for portfolio selection and identifying a change of market regime.

Different agencies may use various sources of information to support their ratings and the timing of changes in ratings. We use three major sources of information: first, the financial reports of traded companies, which are updated on a quarterly basis, usually within 2 months of the end of a fiscal quarter; second, public disclosures and published articles and blogs; third, web-based direct and indirect firm-specific information which may be relevant for ESG ratings.

3. On the need for robust automated portfolio selection approaches

The most classic and celebrated approach to portfolio selection is that of Markowitz (1952). Under this approach, investors find the combination of assets, which maximizes expected return under the condition that the portfolio's risk (assessed by the variance of returns) is below some fixed level set by the investor. This approach lies at the foundation of many portfolio management strategies.

Applying the portfolio optimization approach to highly rated ESG firms would lead to severe bias and possible dangerous underestimation of risk. To solve the Markowitz problem, one indeed needs to have access to the expected returns of the assets and their covariance matrix. In practice, these two fundamental variables are unknown and must be estimated from historical data.

As explained by Brodie et al (2009) and Laloux et al (2000), the estimation noise when inferring the covariance matrix leads to high instability in the resolution of Markowitz's program. In the case of ESG related assets, where availability of price data can be limited and subject to significant measurement error, the direct use of the sample covariance matrix would likely lead to unreliable results.

Our goal is to build an automated and robust recommendation procedure for an investor, based on his personal profile and on historical data concerning the ESG assets. Note that the same tools can be used in the situation where an investor is satisfied with the past performance of his portfolio but wishes to change its composition to invest in ESG assets. In this case, it means we should find a new portfolio constructed from ESG assets which mimics the performance of the investor's initial portfolio. This situation can also be identified as a regression problem with measurement errors. We use the approaches mentioned above to address it.

4. Data Collection and Identification

Our paper focuses on the set of S&P 500 companies. All stocks from the S&P 500 index were scanned for ESG-driven events (positive and negative) starting from the end of 2017.[4] Information sources such as Sustainalytics were used, as well as proprietary Zirra's identified 190 parameters. We aggregate and cluster data in accordance with the subject matter to which the data relate. These clusters are

[4] In general, historical ESG data is only available from 2014.

useful not only for targeted and vertical applications, but also for the creation of more compound scoring formulas and predictive tests.

We focus on data related to ESG, covering terminology signaling environmental, social and governance events, such as a sexual harassment event or a notable climate discussion or change that can be systematically picked up from our alternative data sources.

Our data processing flow is designed to be flexible, extensible and capable of identifying and processing data from thousands of data sources. We classify data sources into the following categories:

- *Structured Directories* - APIs, databases and events detected using natural language processing from the popular news (Google news), social media and networks, "database" websites, where the website is really just a thin presentation layer over structured data.
- *Restructured Time Series Data* - Website traffic, social media statistics such as Twitter followers, Facebook likes and Telegram subscribers.
- *Regulatory and Legal Filings* - Patent and trademark filings, incorporation and ownership filings, quarterly and annual financial reports and much more. Unstructured sources are internally extrapolated into manageable and scorable time series.
- *Unstructured Content* - Analyzed and classified with NLP and Machine Learning (ML) techniques. Examples of unstructured content include news articles, forum posts, social media posts, blogs and user comments and reviews.
- *Human Edits* - Human feedback is an invaluable source in cleaning and correcting data, and in (re)training AI classifiers.

All collected data undergo systematic processing steps to parse, normalize, identify, and aggregate the data into final entity representations. These steps are lengthy, complex, and proprietary, but it is worth highlighting some considerations and challenges:

- *Auditing and Compliance* - The process is built from the ground up to support advanced auditing and compliance specifications. Every piece of ingested data is preserved in its raw format, and every final datapoint can be tied to the raw data from which it is derived.

- *Unstructured Text Processing* - This employs many NLP techniques. To identify companies in unstructured text, we employ several well-known named-entity recognition (NER) tools together with a tool we constructed ourselves, in combination with our proprietary disambiguation tool. For event detection (discussed below), we use a custom feature extractor, which leverages, *inter alia*, grammar/syntax dependency trees, to construct feature vectors from every sentence. We then pass these feature vectors through a series of trained event classifiers.

- *General Company Identification* - We maintain an "ID authority service" that tracks identifiers for the companies we cover. An identifier is anything one can use to reliably point to a specific real-world company, such as a company's homepage URL, its CUSIP or ISIN number, its LinkedIn profile URL, its LEI (legal entity identifier), or its (internal) system ID are all examples of identifiers.

Time series generation can be divided into 3 sub-groups:

- Time series that are collected, filtered, normalized, and then linked to the respective companies. Examples would be "Number of positive/negative media mentions" or "Positive/negative employee/consumer reviews over a given period".

- Time series that are aggregated from several sources, such as "compound Web and mobile traffic score".

- Time series that are generated by using proprietary NLP and AI, such as trending sentiments.

Event Detection - As noted above all unstructured text is processed through the event detection system.

Our algorithm uses NLP sentiment scoring as well as specific event detection using the BERT algorithm (Bidirectional Encoder Representations from Transformers). The algorithm is first trained on historical data and then used for recommending stocks based on current news and social networks. The algorithm monitors news, then identifies relevant events and assigns them a score. This score is used for portfolio selection on a quarterly basis.

Sentences identified as events then undergo further post-processing, to categorize them properly (e.g., the same funding event mentioned in two different sources), and to extract additional structured data from the sentence (e.g., the cost of and date of a given acquisition). This structured data extraction is considered a distinct area of NLP, known as open information extraction (OIE).

5. Results

Successful execution of this model enables smart stock selection based on a systematic synthesis and scoring of alternative data, both structured and unstructured. We construct portfolios based on the ESG scores of companies and check their performance relative to broader indices.

The proposed new algorithm provides a quantitatively tested AI model for analyzing and even predicting the performance of financial assets based on non-market data. This leads to a better understanding of the role corporate governance, social awareness, green energy policies and other sustainability metrics play in share price performance and on value for shareholders. With this understanding, this algorithm can be translated into a systematic scoring system and serve as a core element in portfolio construction. It is an innovative approach, which has a strong potential of becoming an industry standard.

Using our scoring procedure, we construct two portfolios of 50 firms from the S&P 500 index - 50 firms with the highest scores (denoted "High" or H), and 50 with the lowest scores (denoted "Low" or L) for 4 scoring categories: ESG, E, S and G. The portfolios are equally weighted and are reconstituted on a quarterly basis. The graphs below show the annual rates of return for each portfolio for the period 2018 - 2021, and compare them to the benchmark S&P 500 index (SPY).

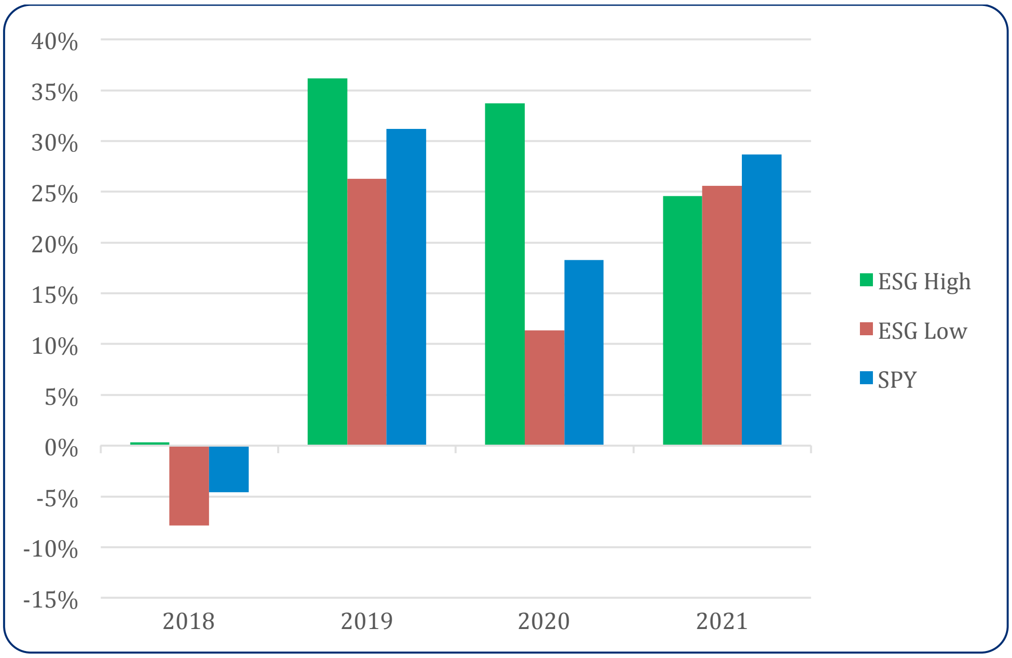
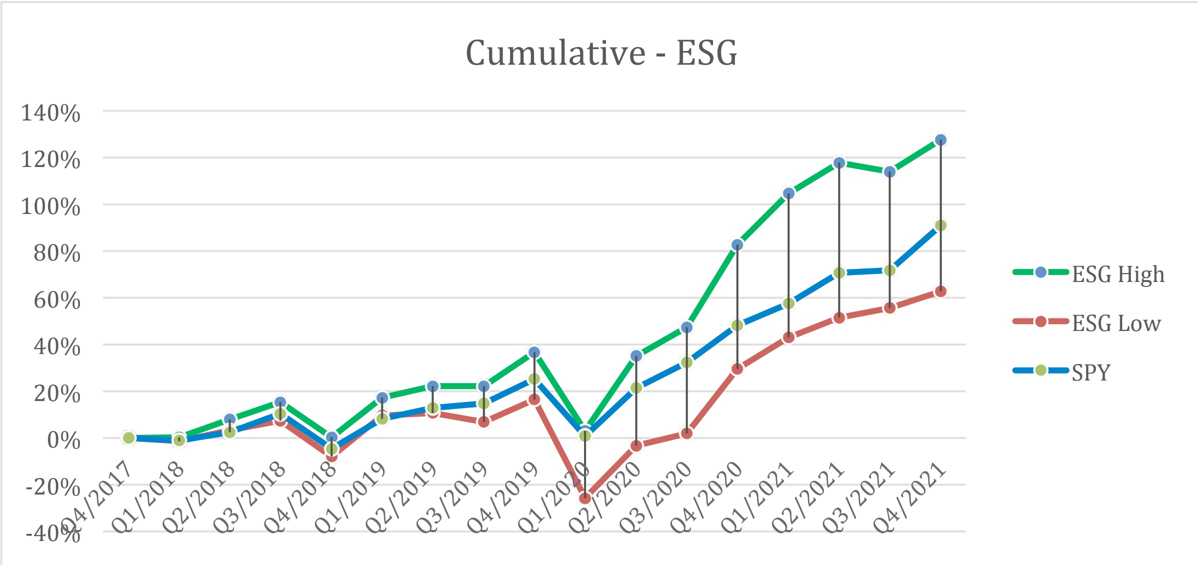


Figure 1: Yearly rates of return on ESG-based portfolios.

Figure 1 demonstrates that, with the exception of 2021, the High ESG-score portfolio consistently outperforms the Low ESG-score portfolio. In 2021, both portfolios performed roughly the same, neither beat the benchmark S&P 500 index. In most quarters, the S&P 500's performance is inferior to that of the high ESG-score companies, but superior to the lower ESG-score portfolio. In 2021, the S&P 500 rose primarily in the fourth quarter and the major contributors to its performance were the leading large-cap companies (Tesla (+36%), Apple (+ 26%), Microsoft (+21%) and others). Since our portfolios are equally weighted, and the S&P index is market cap weighted, we see a superior return for the market index especially in Q4 of 2021 and for the whole 2021.

The outperformance of the High-ESG portfolio is evident in the cumulative return graph below. Over the entire 4-year period, the H portfolio yielded more than the benchmark portfolio by almost 40%, and more than the L portfolio by over 60% (see also Table 1 below).



Graph 1. Cumulative returns for High-ESG portfolio,- Low-ESG portfolio and S&P 500 index for the period 2018-2021.

In order to better understand the ESG effect, we study each component separately by constructing portfolios based on each one of the E, S and G components. Again, we construct portfolios of 50 equally-weighted shares each with the highest and lowest scores for each category, update and rebalance them quarterly after adjusting for dividends). Figure 2 and Graph 2 show the results for the environment, E, scores.

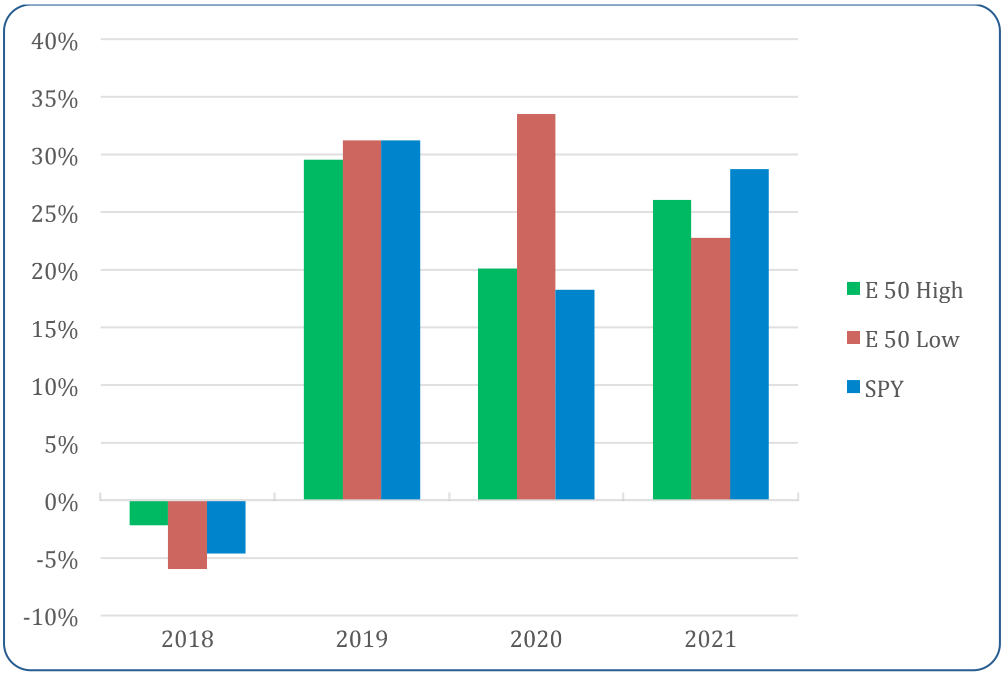
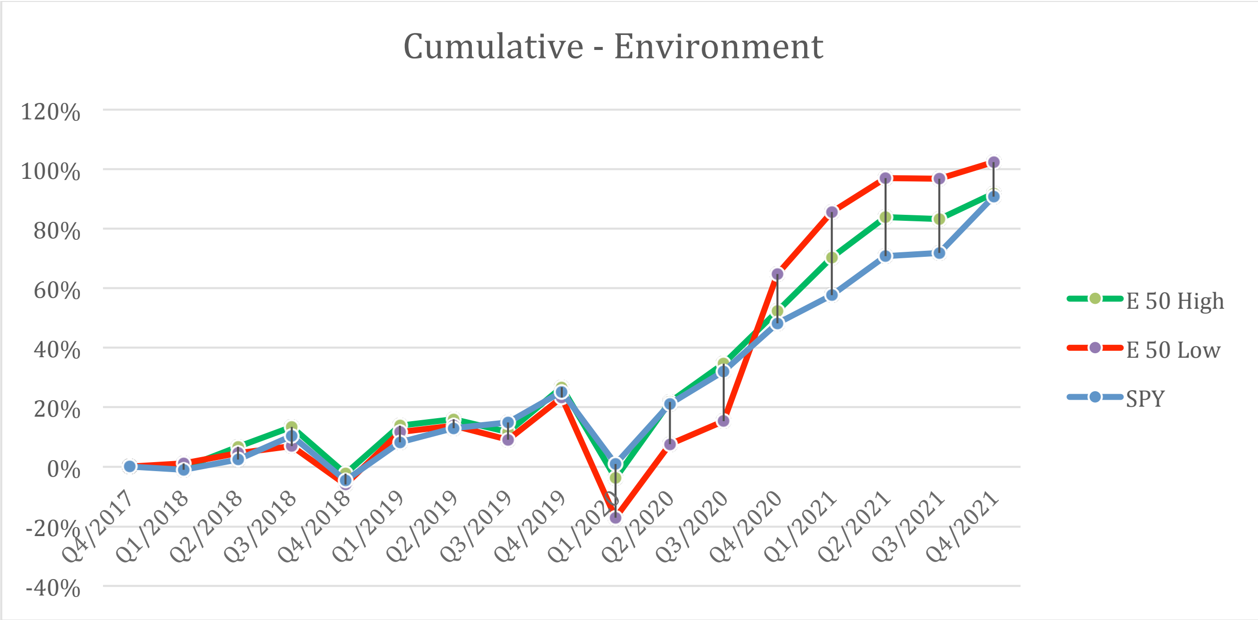


Figure 2: Yearly rates of return on E-based portfolios.



Graph 2. Cumulative returns for E-based high and low portfolios versus S&P 500 for the period 2018-2021.

Figure 2 and Graph 2 show that the E-component does not generate consistent over or under performance for either ESG-based portfolio. In 2020, the high E-score portfolio performed significantly worse than the Low E-score portfolio, although both outperformed the benchmark S&P 500 index. This was the year of the COVID-19 outbreak and many companies with low E-score (e.g. energy, cruise, mining, etc.) had very high returns, from the second half of 2020. It should be noted that in order to achieve a high E score, the firm must make heavy capital investments for which investors may not necessarily be rewarded. This point will be further discussed below.

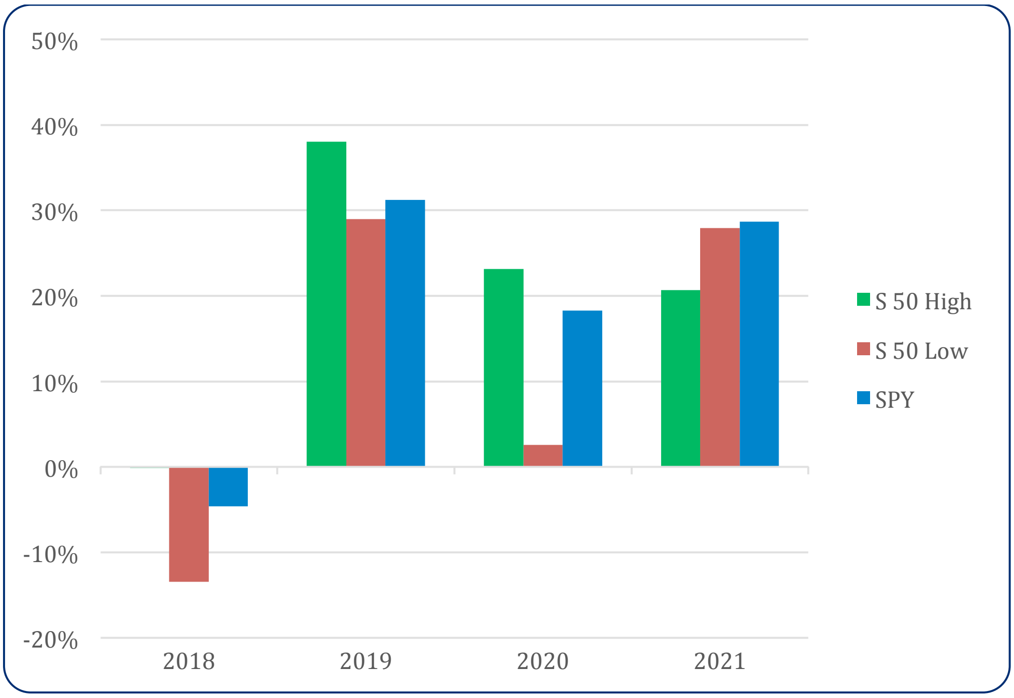
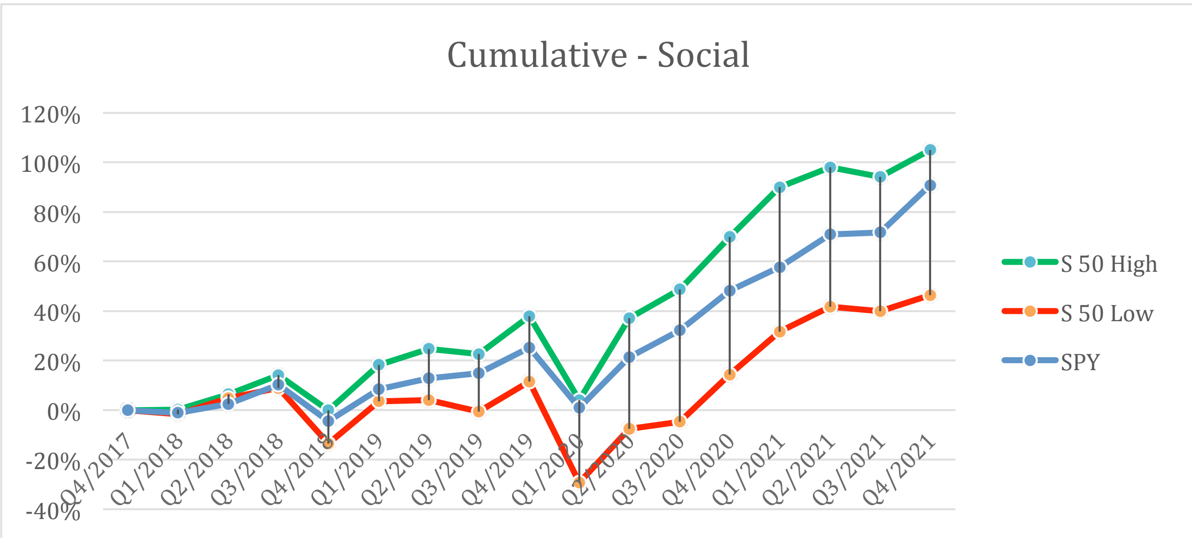


Figure 3: Yearly rates of return on S-based portfolios.



Graph 3. Cumulative returns for S-based High and Low portfolios versus S&P 500 for the period 2018-2021.

The High-S portfolio generated higher returns relative to the Low-S portfolio and the S&P 500 in each year other than 2021. The S&P 500 performed better than the Low-S portfolio in all 4 years. The cumulative difference is significant but not as strong as in the G-based portfolios presented in Figure 4 and Graph 4 below.

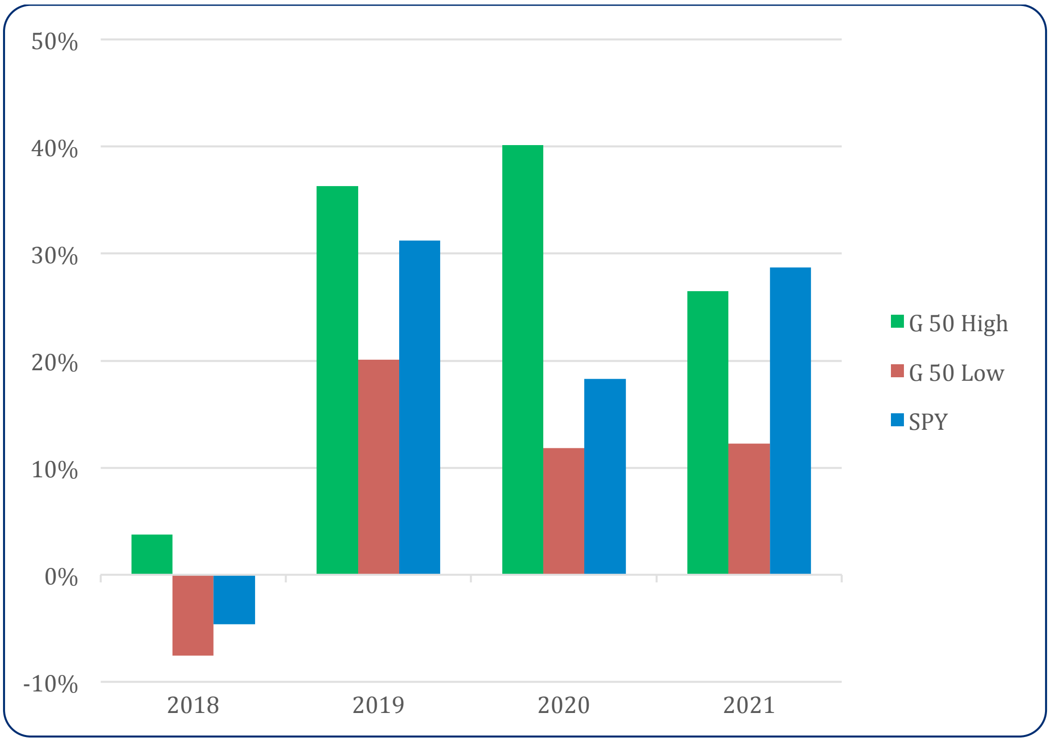
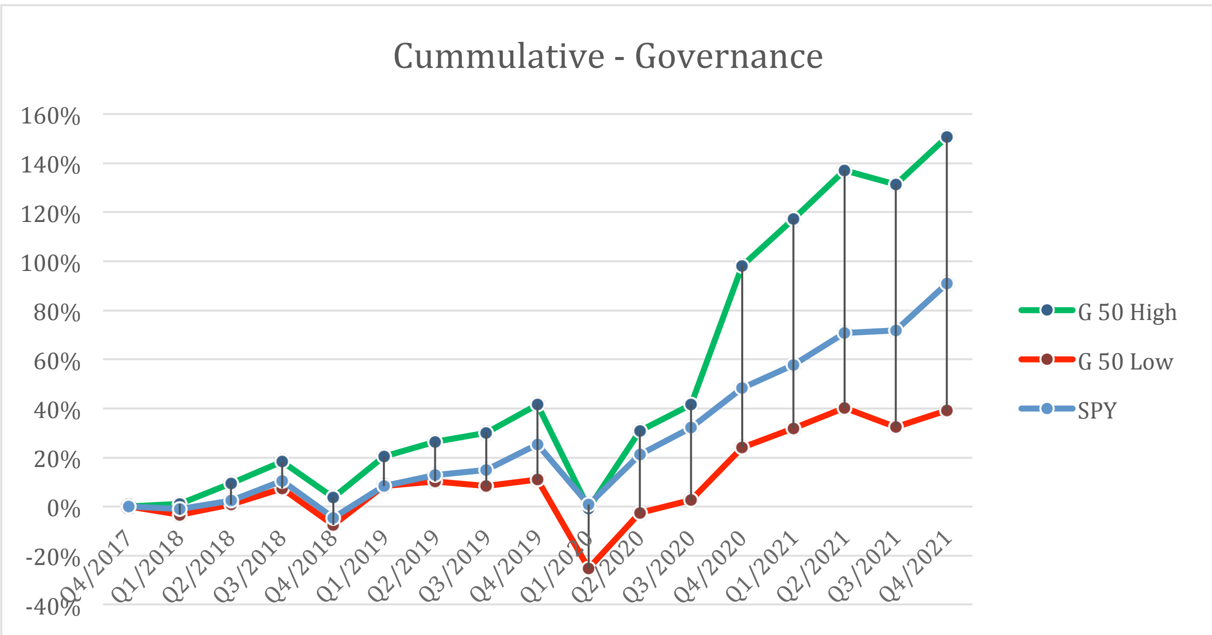


Figure 4: Yearly yields of G-based portfolios



Graph 4. Cumulative returns for G-based High and Low portfolios versus S&P 500 for the period 2018-2021.

The G component generates the best stock performance. The High G-score portfolio consistently and significantly outperformed the Low G-score portfolio, in each of the 16 quarters. This leads to the hypothesis that a higher G-score helps in identifying the better performing stocks. In all four years the high G-score portfolio performed better than the G 50 Low-score portfolio. In three out of the four years, it also outperformed the benchmark index and generated a cumulative 4-year return of more than 150%, as opposed to the S&P 500's return of 91%. and the Low G-score portfolio, in contrast gained only 39% over this period. Again, the superior rate of return of the S&P 500 in 2021, is due primarily to the different weighting methodologies used for the S&P 500 and our constructed portfolios, market cap and equal weighting, respectively.

In order to calculate the relative risk-adjusted performance of the various portfolios we shift from discrete to continuous compounding, and calculate risk measures and risk- adjusted performance, as shown in the following table.

Portfolio	Cumulative yield discrete comp.	Cumulative yield cont. comp.	Volatility annualized	Sharpe ratio
ESG 50 High	127.6%	82.2%	26%	0.79
ESG 50 Low	62.6%	48.6%	33%	0.37
E 50 High	91.9%	65.2%	24%	0.67
E 50 Low	102.3%	70.4%	33%	0.53
S 50 High	104.9%	71.7%	25%	0.71
S 50 Low	46.5%	38.2%	34%	0.28
G 50 High	150.7%	91.9%	31%	0.75
G 50 Low	39.3%	33.1%	29%	0.28
S&P 500	90.9%	64.6%	20%	0.81

Table 1. Relative performance of ESG-based and benchmark portfolios.

Our eight ESG-based portfolios are riskier in terms of standard deviation and betas compared to the benchmark S&P 500. The S&P 500 index has the highest return/risk profile, with a Sharpe ratio of 0.81, followed by the High ESG-score portfolio with 0.79 and the High G-score portfolio with 0.74. The worst performing portfolios as measured by the Sharpe ratio are the Low S-score and low G-score portfolios with a Sharpe ratio of only 0.28. Note that the Low-score portfolios have consistently a much lower Sharpe ratio than their High-score counterparts.

To compare the ESG, E, S, and G portfolios with the S&P 500 index we used a CAPM linear regression, based on the quarterly observations.

Portfolio	alpha	beta	R square
ESG 50 High	0.1%	1.25	0.922
ESG 50 Low	-3.1%	1.53	0.882
E 50 High	-0.7%	1.18	0.930
E 50 Low	-1.6%	1.48	0.796
S 50 High	-0.4%	1.21	0.919
S 50 Low	-4.3%	1.65	0.922
G 50 High	0%	1.42	0.855
G 50 Low	-3.6%	1.41	0.912

Table 2. Mapping of the ESG portfolios on the index.

One can see that alphas (intercept) of high-score portfolios are close to zero while they are consistently negative for the low-score portfolios. Betas for all portfolios are in the 1.18-1.61 range and R-square are relatively high, corresponding to a correlation with the index ranging between 90% and 96%.

Part of the outperformance of the high-score portfolios can be attributed to the high beta of these portfolios in a bull market environment. The inferior performance of the low-score portfolios in 2018-2020, on the other hand, stems mainly from their high negative alpha. This is especially true for the G, S and ESG portfolios.

6. Summary and Conclusions

The topic of ESG has become a hot topic in the investment world, attracting billions of dollars world-wide to specialized mutual and hedge funds and ETFs tracking ESG indices. The interest in ESG, a concept that is not precisely defined, attracted many advisory services to provide ESG ratings as part of their product line of analytic tools. There is a wide disagreement across different rating systems.

We propose a proprietary AI-based algorithm to rate the firms in the S&P 500 index according to ESG as well as its E, S, and G components separately. The AI algorithm is applied to various data including the reports issued by the firms and other public (primarily Web-based) information that is mined using 190 indicators of different metrics. We then construct high-score and Low-score portfolios reconstituted quarterly over the period 2018-2021. We then analyze the performance of the 8 portfolios based on their E, S, G and ESG-scores against the performance of the benchmark S&P 500 index.

Our major result is that both S and G-scores perform well in identifying superior performing companies within the S&P 500 index, while the G-score performs better. We also find the E indicator is problematic in terms of predicting future performance. This result is not necessarily surprising. The E, S and G components indicate different aspects of the firms' characteristics and behavior. Creating a unique ESG score is very arbitrary as explained in the paper. Specifically, E often requires massive capital investment by the company, which is not the case for either S and G. Also, E is very much dependent on the industry affiliation of the firm. Future research should concentrate on the components of ESG and especially on separating E from S and G.

We find that the high-score portfolios for ESG, S and G perform consistently better than the low-score portfolios, particularly after adjusting for risk. Based on data for 2018-2021, the results indicate that investors should reject investments in low-score firms for ESG, S and G.

References

Abramson, L. and D. Chung, (2000). Socially responsible investing: Viable for value investors? *Journal of Investing*, 9(3), 73–80.

Amel-Zadeh, A. and G. Serafeim, (2018), Why and how investors use ESG information: Evidence from a global survey, *Financial Analysts Journal*, 74:3, 87-103.

Auer, B.R. and F. Schuhmacher , (2016), Do socially (ir) responsible investments pay? New evidence from international ESG data, *The Quarterly Review of Economics and Finance*, Volume 59, Pages 51-62.

Avramov, D., A. Lioui, Y. Liu and A. Tarelli, 2021, Dynamic ESG Equilibrium, working paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3935174

Bhagat, S. and B. Bolton, (2008), Corporate governance and firm performance, *Journal of Corporate Finance*, Volume 14, Issue 3, Pages 257-273.

Berg, F., J. F. Koelbel and R. Rigobon, 2019, Aggregate Confusion: The Divergence of ESG Ratings, MIT Sloan School Working Paper 5822-19.

Billio, M., M. Costola, I. Kristova, C. Latino and L. Pelizzon, 2020, Inside ESG Ratings: (Dis)agreement and Performance, Leibniz Institute for Financial Research SAFE, Working Paper No. 284.

Boffo, R. and R. Patalano (2020), "ESG Investing: Practices, Progress and Challenges", OECD Paris, www.oecd.org/finance/ESG-Investing-Practices-Progress-and-Challenges.pdf

Boudoukh, J., R. Feldman, S. Kogan and M. Richardson, (2013), Which News Moves Stock Prices? A Textual Analysis NBER Working Paper No. 18725.

Brav, A., W. Jiang, F. Partnoy and R. Thomas, (2008), Hedge Fund Activism, Corporate Governance, and Firm Performance. *The Journal of Finance*, 63: 1729-1775.

Breedt, A., S. Ciliberti, S. Gualdi and P. Seager, (2019), Is ESG an Equity Factor or Just an Investment Guide? *The Journal of Investing*, 28 (2) 32-42.

Brodie, J., I. Daubechies, C. De Molc, D. Giannone and I. Loris, (2009), Sparse and stable Markowitz portfolios. PNAS, 106, 12267–12272.

Bruno, G., M. Esakia and F. Goltz. (2021), "Honey, I Shrunk the ESG Alpha": Risk-Adjusting ESG Portfolio Returns, EDHEC Scientific Beta.

Demers, E., J. Hendrikse, P. Joos and B. Lev, 2020, ESG Didn't Immunize Stocks Against the COVID-19 Market Crash, Working Paper.

Eccles, N.S. and Viviers, S. (2011), The Origins and Meanings of Names Describing Investment Practices that Integrate a Consideration of ESG Issues in the Academic Literature. *J Bus. Ethics*, 104, 389–402.

Eccles, R.G., Kastrapeli, M.D. and Potter, S.J. (2017), How to Integrate ESG into Investment Decision-Making: Results of a Global Survey of Institutional Investors, *Journal of Applied Corporate Finance*, 29, 125-133.

Environmental, Social and Governance (ESG) Ratings and Data Products Providers. Final Report. IOSCO. The board of the International Organization of Securities Commissions. November 2021. www.iosco.org

Feldman, R. and J. Sanger The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data, Cambridge University Press.

Feldman, R., (2013), Techniques and applications for sentiment analysis. *Communications of the ACM*, 56:4, 82-89.

Feldman, R., S. Govindaraj and J. Livnat, (2010), Management's tone change, post earnings announcement drift and accruals. *Review of Accounting Studies*, 15, 915–953.

Giese, G., Z. Nagy and L-E. Lee, (2021), Deconstructing ESG Ratings Performance: Risk and Return for E, S, and G by Time Horizon, Sector, and Weighting. *The Journal of Portfolio Management*, 47:3, 1-18.

IMF, *World Economic Outlook*, Chapter 1 – The Great Lockdown, April 2020.

IOSCO Board, Environmental, Social and Governance (ESG) Ratings and Data Products Providers Final Report, Nov. 23, 2020.

Kempf, A. and P. Osthoff, 2007, The Effect of Socially Responsible Investing on Portfolio Performance. *European Financial Management*, 13, 908-922.

Kotsantonis, S., C. Pinney and G. Serafeim, (2016), ESG Integration in Investment Management: Myths and Realities. *Journal of Applied Corporate Finance*, 28 (2).

Kumar, N., C. Ashwin, C. Smith, L. Badis, N. Wang, P. Ambrosy and R, Tavares, (2016), ESG factors and risk-adjusted performance: a new quantitative model, *Journal of Sustainable Finance & Investment*, 6:4, 292-300.

Kurtz L. and D. diBartolomeo, (2011), The Long-Term performance of a Social Investment Universe. *The Journal of Investing*, Fall, pp. 95-102.

Laloux, L., P. Cizeau, M. Potters and J. P. Bouchaud, (2000), Random matrix theory and financial correlations. *International Journal of Theoretical and Applied Finance*, 3(03), 391-397.

Lee, L-E., What Does ESG Investing Really Mean? Implications for Investors of Separating Financial Materiality and Social Objectives, working paper 9-2021, Wharton Pension Research Council.

Lee, L-E., G. Giese and Z. Nagy, (2020), Combining E, S, and G Scores: An Exploration of Alternative Weighting Schemes. *The Journal of Impact and ESG Investing*, 1:1, 1-9.

Lo, A. and R. Zhang, 2021, Quantifying the Impact of Impact Investing, working paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3944367

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

Monk, A. and D. Rook, 2021, Resilience as an Analytical Filter for ESG Data, working paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3968081

Nagy, Z., A. Kassam and L. E. Lee, (2016), Can ESG Add Alpha? An Analysis of ESG Tilt and Momentum Strategies. *The Journal of Investing*, 25 (2), 113-124.

Pastor, L., R. F. Stambaugh and L. A. Taylor, 2021, Dissecting Green Returns, working paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3864502

Patel, K., (2018), ESG Investing Moves to the Mainstream. *Financial Analysts Journal*, 74 (3), 39-41.

Shleifer, A. and R. W. Vishny, (1997), A Survey of Corporate Governance. *The Journal of Finance*, 52: 737-783.

Starks, L. T., P. Venkat and Q. Zhu, (2019), Corporate ESG Profiles and Investor Horizons, *The Journal of Investing*, 28:2 ESG.

Tibshirani, R. (1996), Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society (B)*, 58(1), 267-288.

U.N., (2004), Who Cares Wins: Connecting Financial Markets to a Changing World, United Nations Global Impact.

Appendix A – ESG Driven Detected Events

Specifically, we identity, track and quantitatively research a number of ESG-driven events. These events are detected using various NLP techniques, scored for sentiment strength and direction (positive vs. negative) and then quantitatively measured against the market performance of the company's stock during the following periods.

Example of relevant detected events include (a) Scandals/Improper Corporate and Management behavior (b) Gender & Minorities Inclusion/Exclusion (c) Adoption/Rejection of Environmental Protocols/Standards.

These events are picked in accordance with accepted ESG scoring practices as employed by ESG agencies such as Vigeo AIRES, Sustainalytics, Owl, etc.

Each of these events is detected from the structured as well as unstructured data found for each company. The event is scored for (a) strength of association with its ESG context, (b) direction of sentiment, positive or negative, on a linear scale from extremely positive to extremely negative, (c) strength of signal on linear scale from very weak to very strong.

These events are then clustered and quantitatively analyzed as information coefficients with future performance of the company's traded shares. The result is a quantitative, predictive stock scoring model that uses open data to detect ESG sentiment.