ESG RATINGS

A COMPASS WITHOUT DIRECTION

BY DAVID F. LARCKER, LUKASZ POMORSKI, BRIAN TAYAN, AND EDWARD M. WATTS
AUGUST 2, 2022

INTRODUCTION

ESG ratings firms provide information to market participants (investors, analysts, and corporate managers) about the relation between corporations and non-investor stakeholders interests. They do so by sifting masses of data to extract insights into various elements of environmental, social, and governance performance and risk. Investors rely on this information to make investment decisions, while corporations use ratings to gain third-party feedback on the quality of their sustainability initiatives.

Recently, ESG ratings providers have come under scrutiny over concerns of the reliability of their assessments. In this Closer Look, we examine these concerns. We review the demand for ESG information, the stated objectives of ESG ratings providers, how ratings are determined, the evidence of what they achieve, and structural aspects of the industry that potentially influence ratings. Our purpose is to help companies, investors, and regulators better understand the use of ESG ratings and to highlight areas where they can improve. We find that while ESG ratings providers may convey important insights into the nonfinancial impact of companies, significant shortcomings exist in their objectives, methodologies, and incentives which detract from the informativeness of their assessments.

DEMAND FOR ESG INFORMATION

Demand for ESG information has exploded in recent years. Ten years ago, the term ESG—although in existence—was rarely used by the investment community or in corporate boardrooms.[1] Instead, public and professional interest were focused on the general concepts of corporate responsibility, sustainability, and impact investing. Only recently has the focus on ESG (environmental, social, and governance) as a unique concept come to the forefront and with it an explosion in the demand for information (see Exhibit 1).

Sources of this demand include:

- *Asset owners.* Investors concerned about the environmental and societal impact of the companies they invest in. These individuals generally do not want to invest in companies whose products (because of their sourcing, production, end use, or disposal) cause harm to society or otherwise represent practices deemed contrary to their personal values. These individuals use information on companies or funds as an ESG screen for their investments.
- *Institutional investors.* Institutional investors seek information about the environmental and societal performance of companies to create investment products and services to meet the needs of their clients. Additionally, some institutional investors have a view on the financial impact that societal and environmental forces can have on the short- and long-term performance of companies. Examples include passive funds (such as BlackRock, Vanguard, and StateStreet) who have increased their advocacy for environmental or social issues. They also include active managers (such as Parnassus Investments and Calvert Investments) who believe that mitigating ESG factors will improve the risk and performance characteristics of their fund. These funds seek information about the performance of specific companies along various dimensions of ESG and the potential risk that ESG factors pose to business.
- *Companies.* Companies want to demonstrate the extent to which they invest in stakeholder-facing initiatives and highlight their positive impact. On one hand, companies are a primary supplier of ESG information through voluntary disclosure, such as sustainability reports. These are used to highlight activities the company engages in and possibly counter public criticism of the alleged harm of their activities. On the other hand, companies also are consumers of third-party ESG information which they use to validate their claims of positive impact.
- *Regulators.* Regulators who are concerned that ESG information might be material to the financial performance of corporations, particularly human capital management practices and environmental impact. Also, regulators who

assess investment managers' claims about their incorporation of ESG-factors in the investment process.

- *Other stakeholders.* Stakeholders who are not direct beneficiaries or contributors to assets yet have an opinion about how those assets are invested, such as students concerned about a university endowment, consultants advising on the investment process, or local governments interested in pension assets.

Demand for ESG information has in many ways outstripped the ability of suppliers to supply the depth, detail, and accuracy of data required. This is perhaps due to the immense number of factors that plausibly fall under the heading of ESG, the difficulty in measuring ESG factors, and the daunting challenge of determining their impact. To this end, Amel-Zadeh and Serafeim (2018) find several informational impediments that hinder ESG integration in the investment process including lack of comparability across firms, lack of standards, the cost of gathering information, and a lack of quantifiable information.[2]

Commercially developed, third-party ESG ratings are one type of service provider that has evolved to meet the demand for ESG information. A 2020 survey by SustainAbility finds that ESG ratings are the most frequently referenced source of information that institutional investors rely on to gauge ESG performance (55 percent, tied with direct company engagement).[3] Another survey finds that 88 percent of investment professionals use third-party ESG ratings as a part of their investment process, with 92 percent expecting to do so in the future.[4]

The importance of ESG ratings to the asset management business is demonstrated by the flow of funds into ESG-labeled investment products. Bank of America calculates that over $200 billion was invested in ESG bond funds between 2019 and 2022.[5] Hartzmark and Sussman (2019) show that mutual funds with high ESG ratings (as measured by Morningtart) realized net inflows over the measurement period, compared with net outflows among firms with low ESG ratings.[6]

WHAT ARE ESG RATING SUPPOSED TO MEASURE?

ESG ratings are intended to measure "ESG quality." ESG quality itself, however, does not have a single agreed-upon definition. Two main views of ESG exist, and to some extent they work in directionally opposite ways.

One view of ESG is that it reflects the impact a company has on the welfare of its stakeholders, such as employees, suppliers, customers, local community, and the environment. Under this definition, a company can improve its ESG profile by withdrawing from activities that are harmful to stakeholders or improving business practices in affected areas to benefit these constituents. The cost of such investment, at least in the short run, is incurred by shareholders, while the long-term financial impact to the company is undetermined or unstated. This view of ESG ("doing good") is what most individual investors likely think of when they think about ESG quality.

A competing view is that ESG measures the impact societal and environmental factors have on the company, and that these factors are financially material. Under this definition, an ESG framework provides a set of risk factors that the company can plan for or mitigate through strategic planning, targeted investment, or a change in operating activity. Addressing ESG risk factors, even if costly in the short run, is expected to result in a long-term financial benefit to the corporation and its shareholders. This view of ESG (the impact of environmental and social risks on financial performance) is the one predominantly adopted by ESG ratings providers.

The tension between these viewpoints is demonstrated in a *Bloomberg BusinessWeek* article which takes a critical view of ESG ratings, with a focus on the ratings of MSCI. According to the article,

> *There's virtually no connection between MSCI's 'better world' marketing and its methodology. That's because the ratings don't measure a company's impact on the Earth and society. In fact, they gauge the opposite: the potential impact of the world on the company and its shareholders. MSCI doesn't dispute this characterization. It defends its methodology as the most financially relevant for the companies it rates.*

According to the article, MSCI's CEO

> *concedes ordinary investors piling into such funds have no idea that his ratings, and ESG overall, gauge the risk the world poses to a company not the other way around. 'No, they for sure don't understand that,' he said in an interview.*[7]

The authors of this piece make the assumption that ESG ratings are supposed to measure a company's impact on the environment and society and convey surprise that MSCI's ratings attempt to measure the opposite.[8]

WHO ARE THE PLAYERS?

The ESG ratings industry is highly fragmented with dozens of ratings agencies and data providers in existence. The backgrounds of these firms are not uniform, with many having entered the ESG ratings business from different areas of historical expertise. Some ESG ratings firms used to create ESG funds or referenced in the press include:

- *MSCI.* MSCI publishes ESG ratings on 8,500 companies (14,000

issuers) globally, and employs over 200 analysts. The data from MSCI ESG research analysts are also used to produce MSCI ESG Indexes.[9] MSCI was originally a subsidiary of Morgan Stanley (MSCI stands for Morgan Stanley Capital Investment) and its primary business is the compilation of stock-market indexes for license to investment management firms. In 2007, Morgan Stanley spun off MSCI as a separately traded public company. In 2010, MSCI acquired RiskMetrics, which owned KLD, one of the earliest providers of sustainability data in the U.S.[10] In 2014, MSCI purchased GMI Ratings, a provider of governance and accounting quality ratings.[11] In 2019, MSCI acquired a climate-change analytics company called Carbon Delta.[12]

- *ISS ESG.* ISS ESG publishes ratings on 11,800 issuers and 25,000 funds.[13] ISS ESG is a subsidiary of Institutional Shareholder Services, the largest proxy advisory firm that provides recommendations to investment management firms on how to vote various items on the annual proxy. ISS has historically provided governance ratings, and offers consulting services to companies on how they can improve governance quality. In 2014, MSCI sold ISS (which it had acquired through RiskMetrics) to a private-equity firm.[14] In 2017, ISS was sold to another private-equity firm.[15] In 2020, Deutsche Börse acquired majority ownership (80 percent) of ISS.[16]
- *Sustainalytics.* Sustainalytics publishes ESG ratings on over 13,000 companies, and employs 200 analysts.[17] Sustainalytics is owned by Morningstar (acquired in 2020), whose primary business is the rating of mutual funds and exchange-traded funds for use by individual investors.[18] Morningstar uses Sustainalytics ratings to provide sustainability ratings to the funds its rates. Funds are awarded "globes," with a high number of globes indicating lower ESG risk.
- *Refinitiv.* Refinitiv calculates ESG scores on 11,800 companies, and has 700 research analysts.[19] Refinitiv is the rebranded data provider ThomsonReuters, which owns the namesake database as well as newswire Reuters. Refinitiv was purchased by the London Stock Exchange Group (LSEG) in 2021.[20] Refinitiv ESG scores are included for purchase through the company's broader financial databases.
- *FTSE Russell.* FTSE Russell publishes ratings on 7,200 securities.[21] FTSE Russell's main business is the compilation of market indexes which, like MSCI, it licenses to investment management firms. FTSE Russell is also owned by the London Stock Exchange Group, which purchased Russell Indexes in 2015 and combined them with the FTSE Indexes, which it already owned and had jointly developed with the *Financial Times*.[22] LSEG sells FTSE Russell ESG ratings to investment managers for use in individual security selection, and also uses them to create customer benchmarks for mutual funds and exchange-traded funds.

These are just a few ESG ratings providers. Other well-known firms include S&P Global, Vigeo Eiris (owned by Moody's Investor Services), HIP, and TruValue Labs (owned by FactSet Research—See Exhibit 2).

WHAT DO THEY SAY THEY MEASURE?

ESG ratings firms aim to provide insight into ESG quality. However, the approaches they take are not the same. This can be seen in the variation in their stated objectives.

A common theme among ESG providers is investment risk reduction. The assumption is that ESG quality improves financial performance by reducing social and environmental factors that pose risk to the company's business model or operations. To this end, MSCI claims its ratings "support ESG risk mitigation and long-term value creation." Sustainalytics measures "the degree to which a company's economic value is at risk" because of ESG factors. If these providers are correct in their thesis and accurate in their measurement, we should be able to observe a correlation between ESG ratings and subsequent risk events (measured by such factors as financial performance or reduced likelihood of regulatory violations, litigation, or bankruptcy.

Risk reduction is not the only claim of ESG ratings providers. Some are explicit in designing their scores to predict returns. For example, HIP claims that its ratings "correlate with better returns for the same amount of risk."[23] Arabesque says its approach "is all about identifying companies that are better positioned to outperform over the long term. … When calculating the ESG score of a company, the algorithm will only use information that significantly helps explain future risk-adjusted performance."[24] These claims are also testable and can be verified by relating ESG ratings to subsequent stock or bond price changes.

In addition to these, some ESG ratings providers make additional claims, such as measuring a company's environmental or social impact (ISS), transparency and commitment to ESG (Refinitiv), or provide a screen for ESG selection in support of stewardship goals (FTSE Russell). The accuracy of these types of claims is somewhat harder to measure.

ESG ratings are generally reported on a letter or numeric basis to reflect the company's absolute or relative ESG risk or performance. Some companies (such as MSCI) use a 7-point scale from AAA to CCC, analogous to that used by major credit-rating agencies. Others use a 12-point scale from A+ to D-, similar to

an education system (ISS is an example). Another widely used approach is to publish scores on a percentile basis using a scale of 1 to 100, where 100 can either represent high ESG quality (positive) or high ESG risk (negative).

Many ratings providers claim to measure *industry-relative* ESG quality, while some claim to measure *absolute* quality. Industry-adjusted ratings allow investors to compare ESG risk or performance across firms within the same industry. In this way, an energy company that is more financially exposed to environmental risks can be identified against its peer group. However, industry-adjusted ratings do not allow for comparison of firms across industries, and a company's rating is highly dependent on the industry it is designated to. By contrast, ratings providers that claim to measure absolute ESG quality can be used for comparison across industries, although firms tend to receive systematically higher or lower ratings depending on their line of business.

WHAT ARE THE SUBCOMPONENTS?

To arrive at an overall ESG rating, ratings firms typically make separate assessments of the three components of ESG—E (environment), S (social), and G (governance)—which they then aggregate to compute an overall score. In measuring these, the firm must have a view of the major factors that contribute to each component. These might be derived using statistical analysis of historical data to identify drivers of E, S, and G, or they might be hypothesized based on a theoretical relation that is not tested.

For example, MSCI identifies the following subcomponents of E, S, and G:

Environment

Climate change. The company's contribution to climate change through emissions, or the company's exposure to harm due to climate change or climate-related regulatory action.

Natural capital. The degree to which the company relies on natural resources that might be at risk

Pollution and waste. The generation of waste (packaging, materials, or toxins) as part of the production or disposal of company goods.

Environmental opportunities. The potential to use environmental technology to improve operations or sales.

Social

Human capital. All aspects of human capital management including employment practices, talent development, safety, and the labor standards of suppliers.

Product liability. The potential for products to cause harm because of quality failures, safety failures, financial harm, privacy violations or data leaks, chemical harm, other health or demographic risk, and the potential benefits of responsible investment to improve product quality, safety, or impact.

Stakeholder opposition. Societal opposition to the company because of controversial sourcing techniques or locations, or other conflicts with local communities.

Social opportunities. The potential to benefit society by improving access to products.

Governance

Corporate governance. Factors relating to the quality of corporate oversight, including the structure and composition of the board of directors, shareholder ownership structure and control, CEO pay practices, and accounting quality.[25]

Corporate behavior. Evidence into the ethical behavior of the company, including anticompetitive practices, corruption, and tax shielding and transparency.

(See Exhibit 3 for examples of ESG frameworks).

Ratings providers may also leverage reporting frameworks developed by third-party organizations. Examples include the reporting standards developed by the Sustainability Accounting Standards Board (SASB), Task Force on Climate-Related Financial Disclosures, and the Global Reporting Initiative.[26] These frameworks offer the benefit of leveraging the work of independent organizations and are often similar to the proprietary frameworks developed by ESG ratings providers.

One observation is that the number of input variables is massively large. FTSE Russell claims its model uses 300 indicators. Refinitiv uses 630 ESG metrics. S&P Global uses 1,000 underlying data points.

Managing this number of variables requires the ratings provider to make important decisions or simplifying assumptions. One is assessing materiality. Not all variables are equally material across companies or industries. As a result, some variables might require larger or lesser weighting to reflect their relevance; some might be excluded entirely.[27] Another decision is how to deal with missing data. Even though a variable might be deemed material, this does not mean that the relevant data is available to measure that variable. (We discuss options for handling this decision below.) A related decision is how to standardize variables when they are reported differently and therefore are not directly comparable across companies. Finally, the ratings provider must decide how to weight both the variables in their importance to E, S, and G, and also the overall pillars of E, S, and G in relation to one another.

All of these choices will influence the reported ESG rating.

WHAT ARE THE SOURCES OF DATA?

The data sources used to populate ratings models include public, quasi-public, and private data. Public data includes company-reported filings with the SEC, company-produced sustainability reports, press releases, newswires, and media reports. Quasi-public information includes data captured in government, regulatory, and NGO datasets. Nonpublic information might be provided by the company in response to solicited questionnaires.

Working with data sets such as these brings inherent problems. Three major challenges are completeness of data, standardization, and consistency.

Completeness. A model that includes hundreds of material input variables requires data to support each variable. Much of this information is not publicly reported. As a result, the ratings firm will have to make decisions about how to handle missing data. One approach is to simply omit the data point, but this makes it difficult to compare scores across companies that report and do not report a value. Another is to make an assumption about what the data might be. For example, when information is not available to populate a data point, MSCI appears to assume that the company's performance is the industry average.[28] (In this case, the choice of industry peer group will influence how the data point is populated.) By contrast, FTSE assumes that the company's performance is the worst.[29] (This choice is intended to encourage transparency but is also likely punitive.) A third approach is to estimate the data using advanced statistical techniques to impute the missing value.[30]

Standardization. The problem of standardization occurs when companies report information on the same variable using scales that are not directly comparable. For example, one company might report workplace safety information using raw numbers (number of incidents), a time scale (injuries per unit of time worked), or a percentage scale (lost-time frequency).[31] The ratings provider must standardize these differences across companies in order to compute overall ESG performance.

Consistency. To improve the performance of models, a ratings provider might make retroactive adjustments to historical data. For example, the data included in a model five years ago might not be the same as the data in the model today for that same year. Data changes are made to improve the accuracy of models, as new or better data is made available. However, they have the effect of making a model look more predictive than it was. Revising past data based on observed subsequent outcomes can invalidate the results from back testing. This is an important concern when evaluating the predictability and validity of commercial ESG ratings.

The impact of routine methodological choices such as these can be seen in the example of Refinitiv. Berg, Fabisik, and Sautner (2021) show that methodological changes adopted by Refinitiv in 2020 resulted in major changes to both current and historical ratings. Median scores were 18 percent lower with rewritten changes, with 44 percent and 16 percent swings in E and S scores, respectively. These revisions also changed the predictive results of the ratings. Stocks with high ESG scores outperformed in the rewritten data but not in the original data. They observe that "data rewriting is an ongoing rather than a one-off phenomenon," no doubt reflective of firms working to improve the usefulness of their data.[32]

DO THEY DEMONSTRATE WHAT THEY SAY?

Having reviewed the objectives and methodological choices of ESG firms, we can better understand the research evidence regarding ESG ratings quality, consistency, and effectiveness. Unfortunately, it is rare for ratings providers to offer concrete, systematic evidence to back up claims about their ratings.

Investor Perception of ESG Ratings. Practitioners profess a lack of understanding about the methodologies and reliability of ESG ratings. The Alternative Investment Management Association (AIMA), which represents such firms globally, reports that its members "have experienced challenges in terms of understanding and validating the approaches used by different ratings providers."[33] The European Securities and Market Authority describes the market for ESG ratings as "immature," based on its structure and dispersion of methodologies.[34] A 2020 study of institutional investors uncovers widespread concerns, including inaccuracy and inconsistency of data, inexperienced research analysts, and a perception that ESG quality cannot be distilled to a score.[35]

Patterns in ESG Ratings. Systemic patterns are observed in ESG ratings. One pattern is related to company size: Large companies receive higher average ratings than smaller companies. This might be due to the more significant resources large firms are able to invest in ESG initiatives, or it might be due to the fact that large companies have greater disclosure of ESG data. A second pattern is industry-related: While some ESG ratings are industry-adjusted, those that are not may have higher average scores for certain industries (such as banks and wireless communications) than for others (such as tobacco and gaming). It is not clear if these patterns are due to fundamental differences in ESG quality across industries, or a result of the methodological choices and input variables that underpin ESG ratings models. A third pattern is country-related: European companies have higher average

ESG scores than U.S. companies, which might be due to political and regulatory differences across countries. Firms in emerging markets also have lower ratings than firms in more developed economies.[36]

Ratings Improvements. Research also demonstrates an upward drift in ESG ratings over time. D.E. Shaw (2022) analyzes the aggregate ESG scores for all Russell 1000 companies as calculated by MSCI between January 2015 and December 2021. They find an 18 percent aggregate improvement over the measurement period. Structural changes account for 6 percentage points of this improvement. These include:

1. Changes in the index composition, with higher-rated companies (such as Microsoft) growing over time to represent a larger percentage of the total index.
2. Changes in the weightings assigned to components in the MSCI model. For example, MSCI eliminated key issue scores for "energy efficiency," reduced weightings for "toxic emissions" and "health and safety," and increased weightings for "human capital development," leading to higher average scores than would have occurred if these weightings had not changed.
3. More disclosure by companies. Companies that increased disclosure (for example, by disclosing their carbon emissions) were significantly more likely to experience a subsequent upgrade without regard to the fact that the company's underlying performance in this area did not necessarily change.

Adjusting for these structural changes, D.E. Shaw still finds that MSCI ratings are subject to an aggregate 12 adjusted-improvement (which the report describes as "grade inflation"). They do not explain the reason for this improvement.[37]

Correlations Across Providers. Studies find low correlations across ESG ratings providers.[38] This is perhaps surprising if ESG ratings are supposed to measure the same construct.

CFA Institute (2021) finds correlations across the major providers ranging from 0.65 (between S&P Global and Sustainalytics) to 0.14 (between ISS and S&P Global).[39] Dimson, Marsh, and Staunton (2020) find not only that ESG ratings vary across providers but the individual components (E, S, and G) also vary widely. For example, assessments of the E, S, and G components as determined by MSCI and Sustainlytics exhibit correlations of only 0.11, 0.18, and -0.02, respectively. This suggests either they are measuring unrelated constructs or they have significant measurement error in measuring the same construct (see Exhibit 4).[40]

Berg, Kölbel, and Rigobon (2022) try to identify reasons why ESG ratings diverge across providers. They deconstruct ratings along three dimensions: scope (the attributes the ratings providers attempt to measure), measurement (the measures used to evaluate the same attributes), and weighting (the weights assigned to attributes in reflection of their relative importance). They find that differences in measurement (56 percent) and scope (38 percent) account for most of the divergence, with weighting differences accounting for just 6 percent of the variance.[41] This illustrates how fundamental the methodological differences are across firms.

Perhaps unexpected, Christensen, Serafeim, and Sikochi (2022) find that corporate disclosure does not reduce the divergence of ESG ratings but instead increases it.[42] They explain that "due to the subjective nature of ESG information … higher disclosure would be associated with higher disagreement, as disclosure expands opportunities for different interpretations of information." This suggests that greater corporate disclosure requirements of environmental and social data might not lead to more consistent ESG ratings. In this way, ESG ratings might be similar to equity analyst ratings, where the rating is ultimately dependent on the interpretation of information rather than its availability.

The divergence of ESG ratings has several implications. One is the potential to confuse investment decisions by giving unreliable information about the ESG quality of firms.[43] Another is that it confuses the disclosure that fund managers make to investors about the overall ESG quality of their portfolio. A third is that it reduces the incentive of companies to improve their ESG performance by sending unreliable signals about how their ESG initiatives are assessed by third-party observers.

Environmental and Social Outcomes. Studies find that ESG ratings have low associations with environmental and social outcomes.[44]

A review of MSCI ratings conducted by Bloomberg finds that most upgrades occur for what Bloomberg calls "rudimentary business practices" rather than substantive improvements. In justifying 155 upgrades, MSCI cited governance improvements almost half (42 percent) of the time—significantly more than social (32 percent) or environmental (26 percent) improvements. Upgrades were often driven by check-the-box practices, such as conducting an employee survey that might reduce turnover, and rarely for substantial practices, such as an actual reduction in carbon emissions. Half of companies were upgraded for doing nothing—the result of methodological changes.[45]

Raghunandan and Rajgopal (2022) find that companies in ESG portfolios (those with high Sustainalytics ratings) have worse records for compliance with labor and environmental laws relative to companies in non-ESG portfolios during the same period. Companies added to ESG portfolios also do not subsequently improve compliance with labor or environmental regulations.[46]

Gibson, Glossner, Krueger, Matos, and Steffen (2022) find that U.S. firms that join the Principles for Responsible Investment (PRI), which commit a company to incorporate ESG factors into their decision-making processes, earn worse ESG ratings (as assigned by MSCI, Refinitiv, and Sustainalytics) than U.S. firms that do not make this commitment.[47,48]

Stock Price Outcomes. The relation between financial performance and ESG ratings is uncertain.

Dunn, Fitzgibbons, and Pomorski (2018) study the risk characteristics of companies based on their ESG ratings (as provided by MSCI). They find that companies with the lowest ratings have volatility that is up to 15 percent higher and betas up to 3 percent higher than stocks with the highest ratings. They also find that ESG scores might be predictive of future risk, although the effects are modest. They conclude that "ESG information may play a role in investment portfolios that goes beyond the ethical considerations and may inform investors about the riskiness of the securities in a way that is complementary to what is captured by traditional statistical risk models."[49]

Hartzmark and Sussman (2019) examine the relation between fund sustainability and performance (using Sustainability fund ratings). They find that funds with low sustainability ratings perform better than those with high ratings.[50] Bansal, Wu, and Yaron (2022) find that companies with high ESG ratings (by MSCI) perform better during good economic times but worse during bad economic times.[51] Demers, Hendrikse, Joos, and Lev (2021) study the performance of companies at the onset of Covid-19 and find no evidence that ESG ratings predict performance during this unexpected risk event.[52] Lopez-de-Silanes, McCahery, and Pudschedl (2019) examine ESG ratings outside of the U.S.—primarily in European countries, Australia, and Japan. They find that ESG scores of companies domiciled in these countries are not associated with risk-adjusted performance.[53]

Schröder (2007) and Dimson, Marsh, and Staunton (2020) both find that ESG indexes created by ESG ratings firms (such as MSCI and FTSE Russell) exhibit outperformance during their prelaunch periods only to underperform after their launch dates. This suggests that ESG indexes are created through back-testing methods that do not result in a sustainable investment strategy.[54]

Atz, Liu, Bruno, and Van Holt (2021) provide a substantial literature review of over 1,100 primary peer-reviewed papers and 27 meta-analyses on ESG and sustainable investing published between 2015 and 2020. They conclude that "the financial performance of ESG investing has on average been indistinguishable from conventional investing."[55]

It might be the case that, while the ratings published by any single ratings provider are not predictive of performance, the assessments of multiple providers might be informative when considered in aggregate. To this end, Berg, Kölbel, Pavlova, and Rigobon (2021) attempt to combine the ratings of multiple providers to reduce the "noise" from conflicting assessments. They find some evidence that combining the scores from multiple firms leads to a stronger relationship between ESG and performance.[56]

STRUCTURAL CHARACTERISTICS OF THE INDUSTRY

Several structural features might influence the quality of ESG ratings. These include:

- The financial incentive for ratings to be adopted without regard to quality. Many advisory firms benefit from the use of ratings, including firms that advise companies on how to improve ESG disclosure and ratings, audit firms who are paid to attest to the accuracy of disclosure, and investment firms who market ESG-compliant products to the general public. These firms financially benefit from the use of ratings even if the ratings themselves ultimately do not provide reliable information to retail investors.
- Conflicts of interest due to the sale of consulting services to rated companies.[57] In general, the practice of offering paid services to rated companies to increase their ratings at least raises serious concerns about whether this compromises the independence of those ratings.
- Conflicts of interest when a ratings provider rates an affiliated company. Tang, Yan, and Yao (2022) show that "sister firms" of an ESG ratings provider receive higher ratings from the affiliated ratings firm than they do from independent firms.[58]
- Incentives to adopt aggressive methodological choices to gain market share or recognition. For example, a ratings agency might assign low ratings to a company to compel it to increase disclosure, even though that methodological choice is misleading to the investor. Or a ratings firm might assign artificially positive ratings to gain favor with, and recognition from, rated companies.

WHY THIS MATTERS

1. The purpose of ESG ratings is to provide information to market participants about the quality of a company's ESG program and potential risks that might arise due to societal or environmental exposure. However, current evidence is mixed on whether these models, which rely on a large number of input variables, predict investment risk or return. It is also increasingly unclear whether they capture or predict improvements in stakeholder outcomes. What is the source

of this failure? Is it due to methodological choices these firms make? Or is it due to the sheer challenge of measuring a concept as broad and all-encompassing as "ESG?"

2. ESG ratings are relied on by institutional investors to develop portfolios and attract investment dollars from retail investors. These funds often charge higher fees than non-ESG funds.[59] Are institutional fund managers properly motivated to ensure that the ESG ratings they rely on to create these funds are reliable in predicting risk or performance? What steps do they take to validate ratings before using them?
3. Many retail investors purchase ESG funds in order to ensure their investments reflect certain societal values or environmental standards. Do they know that the ESG ratings used to create these portfolios do not necessarily attempt to measure a company's commitment to those values or standards? Should ESG fund managers disclose this?
4. Given the substantial research evidence that ESG ratings are unreliable in predicting outcomes, why do individual and institutional investors rely so heavily on them? Despite these weaknesses, do ESG ratings still have a role to play as a trusted third-party opinion of ESG risk, or as a common language for use in reporting in compliance purposes?
5. A fundamental challenge for ESG ratings providers is access to quality data to use in their models. Would more expansive corporate disclosure improve the reliability of ESG ratings, or would it add noise to already extensive disclosure requirements? Is it possible for companies to effectively report on the vast number of potential stakeholder-related metrics that would be required (carbon emissions, pollution and waste, human capital management, supply chain practices, product use and safety, etc.)?[60]
6. The major credit rating agencies Moody's, Standard & Poor's, and Fitch are subject to regulation by the Securities and Exchange Commission which requires covered firms to adhere to certain policies, procedures, and protections to reduce conflicts of interest and improve market confidence in their quality. Should ESG ratings be subject to similar requirements?[61] ■

[1] According to Wikipedia, the term ESG was first used in a 2004 report to the United Nations titled, "Who Cares Wins." A search of the research literature does not contain a notable number of papers using the term ESG until 2016, with a significant increase in the use of the term only occurring in the titles of peer-reviewed journal articles in 2019 and beyond. See Wikipedia, "environmental, social and corporate governance," (accessed June 20, 2022). See also, International Finance Corporation, "Who Cares Wins: Connecting Financial Markets to a Changing World," (2004).

[2] Amir Amel-Zadeh and George Serafeim, "Why and How Investors Use ESG Information: Evidence from a Global Survey," *Financial Analysts Journal* (2018).

[3] The SustainAbility survey finds that investors rely on ESG ratings and direct engagement the most (55 percent), followed by sustainability reports (50 percent), in-house research (41 percent), corporate ESG rankings (23 percent), corporate filings (18 percent), the media (14 percent), and other sources (9 percent). See SustainAbility, "Rate the Raters 2020: Investor Survey and Interview Results," (March 2020).

[4] Survey of 130 fund management professionals conducted by Ninety One, cited in: Natalie Kenway, "Investors 'Overly Reliant' on ESG ratings," *ESG Clarity* (May 23, 2022).

[5] Kay Hope, "ESG Matters: Skinny but Positive YTD Inflows to ESG," BofA Global Research (July 11, 2022).

[6] Samuel M. Hartzmark and Abigail B. Sussman, "Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows," *Journal of Finance* (2019).

[7] Cam Simpson, Akshat Rathi, and Saijel Kishan, "The ESG Mirage," *Bloomberg BusinessWeek* (December 10, 2021).

[8] The Securities and Exchange Commission also mixes these views of ESG. The preface to its draft rule on climate-related disclosure states that "climate risk can pose significant financial risks to companies, and investors need reliable information about climate risks to make informed investment decisions." The preponderance of the draft rule, however, specifies disclosure on the size of corporate emissions, which reflect the company's impact on the environment rather than the impact of the environment on the company. See Securities and Exchange Commission, "SEC Proposes Rules to Enhance and Standardize Climate-Related Disclosures for Investors," press release (March 21, 2022).

[9] MSCI ESG Ratings, available at: https://www.msci.com/our-solutions/esg-investing/esg-ratings. (Accessed June 2022). Subsequent explanations of the firm's objectives and methodologies are from this source, unless otherwise referenced.

[10] MSCI, "MSCI Inc. to Acquire RiskMetrics Group, Inc," press release (March 1, 2010).

[11] GMI Ratings was formed in 2010 through a three-way merger of GovernanceMetrics International, The Corporate Library, and Audit Integrity, which each operated in the governance rating, proxy advisory, and accounting quality space. MSCI, "MSCI to Acquire GMI Ratings," press release (June 27, 2014).

[12] MSCI, "MSCI to Strengthen Climate Risk Capability with Acquisition of Carbon Delta," press release (September 9, 2019).

[13] ISS ESG, available at: https://www.issgovernance.com/esg/. (Accessed June 2022). Subsequent explanations of the firm's objectives and methodologies are from this source, unless otherwise referenced.

[14] MSCI, "MSCI Completes Sale of ISS," press release (April 30 2014).

[15] Genstar Capital, "Genstar Capital, in Partnership with Management, Announces Acquisition of Institutional Shareholder Services," press release (September 7, 2017).

[16] Genstar Capital, "Deutsche Börse Acquires Leading Governance, ESG Data and Analytics Provider ISS," press release (November 17, 2020).

[17] Sustainalytics, ESG Risk Ratings, available at: https://www.sustainalytics.com/esg-data. (Accessed June 2022). Subsequent explanations of the firm's objectives and methodologies are from this source, unless otherwise referenced.

[18] Morningstar, "Morningstar to Acquire Sustainalytics and Expand Access to ESG Research, Data, and Analytics for Investors Worldwide," press release (April 21, 2020).

[19] Refinitiv ESG Company Scores, May 2022, available at: https://www.refinitiv.com/en/sustainable-finance/esg-scores. (Accessed June 2022). Subsequent explanations of the firm's objectives and methodologies are from this source, unless otherwise referenced.

[20] Refinitiv, "Proposed All Share Acquisition of Refinitiv by London Stock Exchange Group (LSEG)," press release (August 1, 2019).

[21] FTSE Russell ESG Ratings, available at: https://www.ftserussell.com/data/sustainability-and-esg-data/esg-ratings. (Accessed June 2022). Subsequent explanations of the firm's objectives and methodologies are from this source, unless otherwise referenced.

[22] LSEG, "Proposed Acquisition of Frank Russell Company and Fully Underwritten Rights Issue," press release (June 26, 2014).

[23] HIP Human Impact + Profit, Corporate ESG Ratings, available at: https://hipinvestor.com/corporate-ratings/. (Accessed June 2022). Subsequent explanations of the firm's objectives and methodologies are from this source, unless otherwise referenced.

[24] Arabesque S-Ray, available at: http://arabesque.com/docs/sray/S-Ray%20Methodology%20v260.pdf. (Accessed July 2022).

[25] The research literature suggests that corporate governance quality cannot be easily measured through outside observation and that the variables used by most ESG ratings providers in this area are not supported by peer-review research. For a discussion of why corporate governance does not belong in ESG, see David F. Larcker and Brian Tayan, "The Case for Taking the 'G' Out of ESG," *The Wall Street Journal* (April 28, 2022).

[26] Efforts are underway to harmonize the frameworks developed by these third-party organizations. Any changes made to them will of course flow through to impact the ratings of rating providers that rely on them. See Sustainability Accounting Standards Board (SASB), Materiality Map, available at: https://www.sasb.org/standards/materiality-map/; Task Force on Climate-related Financial Disclosures, available at: https://www.fsb-tcfd.org/; the Global Reporting Initiative, available at: https://www.globalreporting.org/.

[27] For example, under the SASB framework, human rights and community relations are not deemed material for auto parts manufacturers but they are for healthcare delivery firms. This means that human rights issues will not factor into the ESG ratings of an auto parts firm, presumably unless they are egregious enough violation to merit overriding the framework. See SASB Materiality Finder, available at: https://www.sasb.org/standards/materiality-finder/?lang=en-us.

[28] James Mackintosh, "Social, Environmental Investment Scores Diverge," *The Wall Street Journal* (September 18, 2018).

[29] Ibid.

[30] For a more detailed discussion, see Sakis Kotsantonis and George Serafeim, "Four Things No One Will Tell You About ESG Data," *Journal of Applied Corporate Finance* (2019).

[31] Ibid.

[32] Florian Berg, Kornelia Fabisik, and Zacharias Sautner, "Is History Repeating Itself? The (Un)predictable Past of ESG Ratings," *Social Science Research Network* (August 2021), available at: https://ssrn.com/abstract=3722087.

[33] AIMA, "AIMA Comments on the European Commission's Targeted Consultation on the Functioning of the ESG Ratings Market in the European Union and on the Consideration of ESG factors in Credit Ratings," (June 7, 2022).

[34] European Securities and Markets Authority, "Outcome of ESMA Call for Evidence on Market Characteristics of ESG Rating and Data Providers in the EU," (June 24, 2022).

[35] SustainAbility, "Rate the Raters 2020," op. cit.

[36] Cited in Osman T. Akgun, Thomas J. Mudge III, and Blaine Townsend, "How Company Size Bias in ESG Scores Impacts the Small Cap Investor," *The Journal of Impact and ESG Investing* (2021).

[37] The improvement might be due to the relative ratings performance of U.S. companies (which are the focus on the study) compared to non-U.S. companies. See D.E. Shaw & Co, "Keep the Change: Analyzing the Increase in ESG Ratings for U.S. Equities," (April 2022).

[38] Correlations run from 1.0 for perfectly positive correlation to 0 for no correlation to -1.0 for perfectly negative correlation. Perfect positive correlation means that the results of one ratings provider exactly match the ratings of the other provider. No correlation means that there is no statistical relation (positive or negative) between the providers. A perfectly negative correlation means that the results of one ratings provider are exactly the opposite of the ratings of the other provider.

[39] Kevin Prall, "ESG Ratings: Navigating Through the Haze," blog posting at CFA Institute (August 10, 2021).

[40] Elroy Dimson, Paul Marsh, and Mike Staunton, "Divergent ESG Ratings," *The Journal of Portfolio Management* (2020).

[41] Florian Berg, Julian F. Kölbel, and Roberto Rigobon, "Aggregate Confusion: The Divergence of ESG Ratings," *Review of Finance* (2022).

[42] Dane Christensen, George Serafeim, and Anywhere Sikochi, "Why Is Corporate Virtue in the Eye of the Beholder? The Case of ESG Ratings," *The Accounting Review* (2022).

[43] Avramov, Cheng, Lioui, and Tarelli (2021) find that market premiums increase and demand for stocks declines when there is greater uncertainty about ESG. See Doron Avramov, Si Cheng, Abraham Lioui, and Andrea Tarelli, "Sustainable Investing with ESG Rating Uncertainty," *Journal of Financial Economics* (2022).

[44] While this might not be surprising to institutional investors who understand that ESG ratings are intended to measure ESG risk exposure, it would likely surprise retail investors who believe they are investing in companies that reflect certain values.

[45] Cam Simpson, Akshat Rathi, and Saijel Kishan, "The ESG Mirage," op. cit.

[46] Aneesh Raghunandan and Shiva Rajgopal, "Do ESG Funds Make Stakeholder-Friendly Investments?" *Review of Accounting Studies* (2022).

[47] The Principles for Responsible Investment were developed by global institutional investors at the request of the United Nations. See Principles for Responsible Investment, "What Are the Principles for Responsible Investment?" available at: https://www.unpri.org/about-us/what-are-the-principles-for-responsible-investment.

[48] Rajna Gibson Brandon, Simon Glossner, Philipp Krueger, Pedro Matos, and Tom Steffen, "Do Responsible Investors Invest Responsibly?" *Social Science Research Network* (February 2022), available at: https://ssrn.com/abstract=3525530.

[49] Jeff Dunn, Shaun Fitzgibbons, and Lukasz Pomorski, "Assessing Risk Through Environmental, Social and Governance Exposures," *Journal of Investment Management* (2018).

[50] Hartzmark and Sussman, op. cit.

[51] Ravi Bansal, Di (Andrew) Wu, and Amir Yaron, "Socially Responsible Investing in Good and Bad Times," *The Review of Financial Studies* (2022).

[52] Elizabeth Demers, Jurian Hendrikse, Philip Joos, and Baruch Lev, "ESG Didn't Immunize Stocks During the COVID-19 Crisis, But Investments in Intangible Assets Did," *Journal of Business Finance & Accounting* (2021).

[53] Florencio Lopez-de-Silanes, Joseph A. McCahery, and Paul C. Pudschedl, "ESG Performance and Disclosure: A Cross-Country Analysis," *Social Science Research Network* (2019), available at: https://ssrn.com/

abstract=3506084.

[54] Michael Schröder, "Is There a Difference? The Performance Characteristics of SRI Equity Indexes," *Journal of Business Finance & Accounting* (2007); and Dimson, Marsh, and Staunton, op. cit.

[55] The papers reviewed were originally published in the 2015 to 2020 period. See Ulrich Atz, Zongyuan (Zoe) Liu, Christopher C. Bruno, and Tracy Van Holt, "Does Sustainability Generate Better Financial Performance? Review, Meta-analysis, and Propositions," *Social Science Research Network* (September 9, 2021), available at: https://ssrn.com/abstract=3708495.

[56] Florian Berg, Julian F. Kölbel, Anna Pavlova, and Roberto Rigobon, "ESG Confusion and Stock Returns: Tackling the Problem of Noise," *Social Science Research Network* (2021), available at: https://ssrn.com/abstract=3941514.

[57] For an example of a rating company offering consulting services related to the improvement of their rating, see ISS Corporate Solutions, "Sustainability Solutions," (accessed July 2022), available at: https://www.isscorporatesolutions.com/improve-esg-ratings/.

[58] Dragon Yongjun Tang, Jiali Yan, Chelsea Yaqiong Yao, "The Determinants of ESG Ratings: Rater Ownership Matters," *Social Science Research Network* (2022), available at: https://ssrn.com/abstract=3889395.

[59] Andy Kessler, "The Many Reasons ESG Is a Loser," *The Wall Street Journal* (July 10, 2022).

[60] To this end, one recent survey finds that a large majority of executives have reservations about the quality of the ESG data they currently report. See Mark Segal, "Over 70% of Execs Lack Confidence in Their Own ESG Data Reported to Stakeholders," *ESG Today* (June 23, 2022).

[61] In the United Kingdom, the Financial Conduct Authority (FCA) has outlines a rationale for regulating ESG ratings providers in that country. See Patrick Temple-West, "FCA to Regulate ESG Ratings Businesses," *The Financial Times* (July 1, 2022).

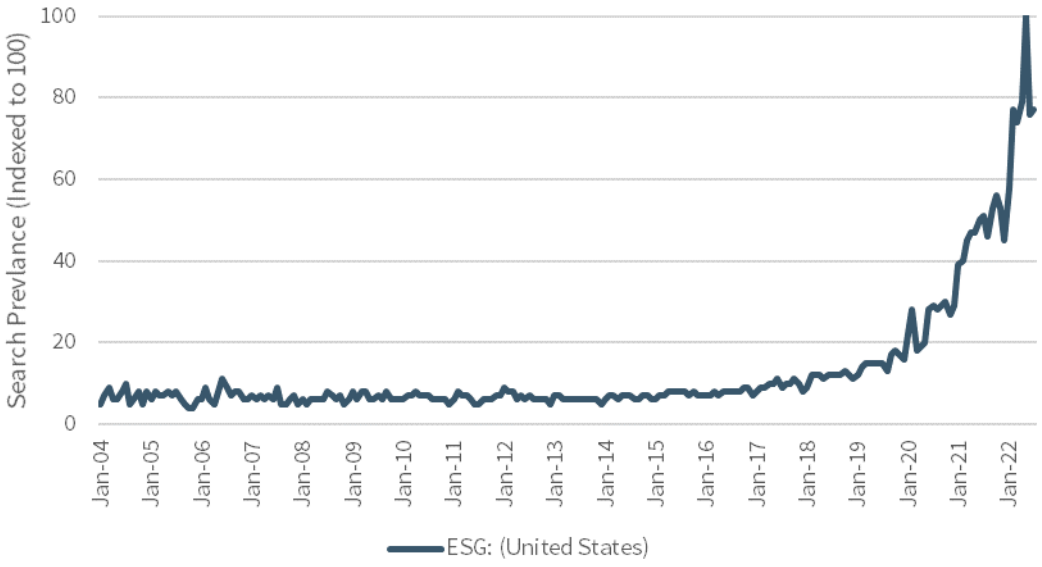
David Larcker is Director of the Corporate Governance Research Initiative at Stanford Graduate School of Business and senior faculty member at the Rock Center for Corporate Governance at Stanford University. Lukasz Pomorski is Head of ESG Research and a Managing Director at AQR Capital Management. Brian Tayan is a researcher with Stanford's Corporate Governance Research Initiative. Edward Watts is Assistant Professor of Accounting at Yale School of Management. Larcker and Tayan are coauthors of the books Corporate Governance Matters *and* A Real Look at Real World Corporate Governance*. AQR Capital Management is a global investment management firm, which may or may not apply similar investment techniques or methods of analysis as described herein. The views expressed here are those of the authors and not necessarily those of AQR.*

The Stanford Closer Look Series is dedicated to the memory of our colleague Nicholas Donatiello and to the retirement of our colleague Michelle E. Gutman.

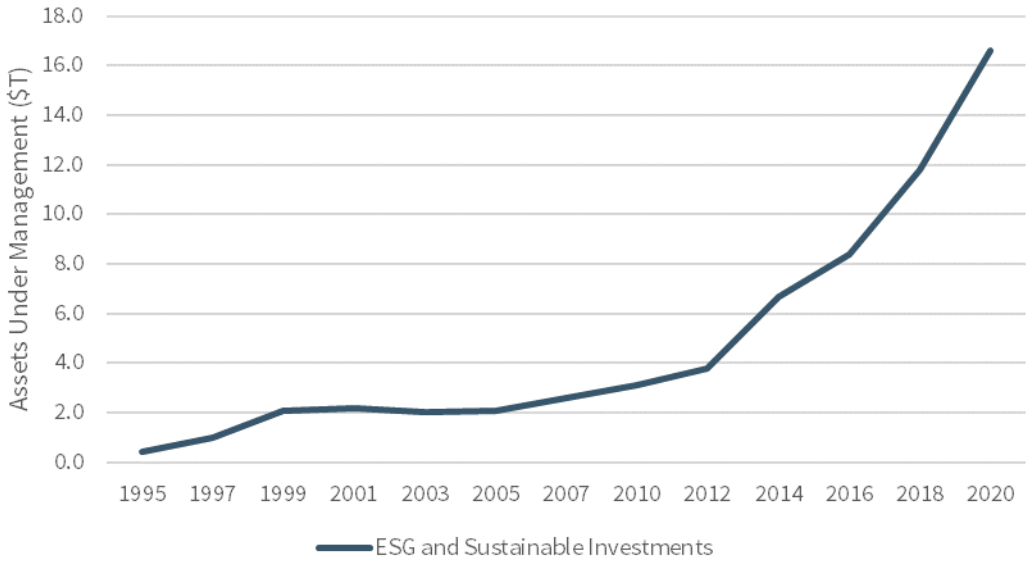
The Stanford Closer Look Series is a collection of short case studies that explore topics, issues, and controversies in corporate governance and leadership. It is published by the Corporate Governance Research Initiative at the Stanford Graduate School of Business, The Hoover Institution at Stanford University, and the Rock Center for Corporate Governance at Stanford University. For more information, visit: http:/www.gsb.stanford.edu/cgri.

EXHIBIT 1 — INTEREST IN ESG OVER TIME

GOOGLE TRENDS: SEARCHES FOR ESG (2004-2022)



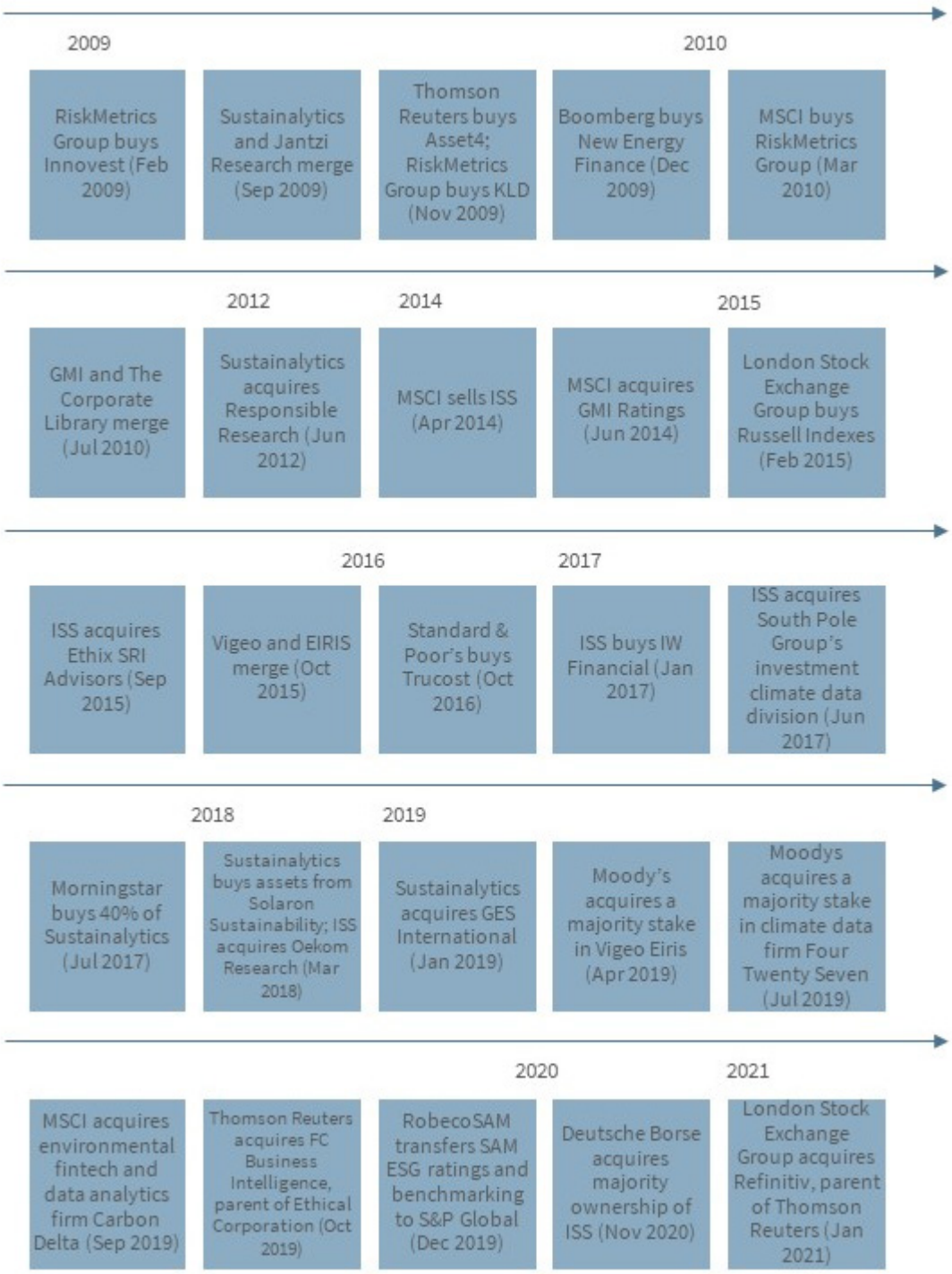
ESG MUTUAL FUNDS: ASSETS UNDER MANAGEMENT (1995-2020)



Sources: Google Trends, "ESG," (2004 to July 2022); U.S. SIF Foundation, "Sustainable and Impact Investing—Money Managers," (2020).

EXHIBIT 2 — EVOLUTION OF ESG RATINGS INDUSTRY

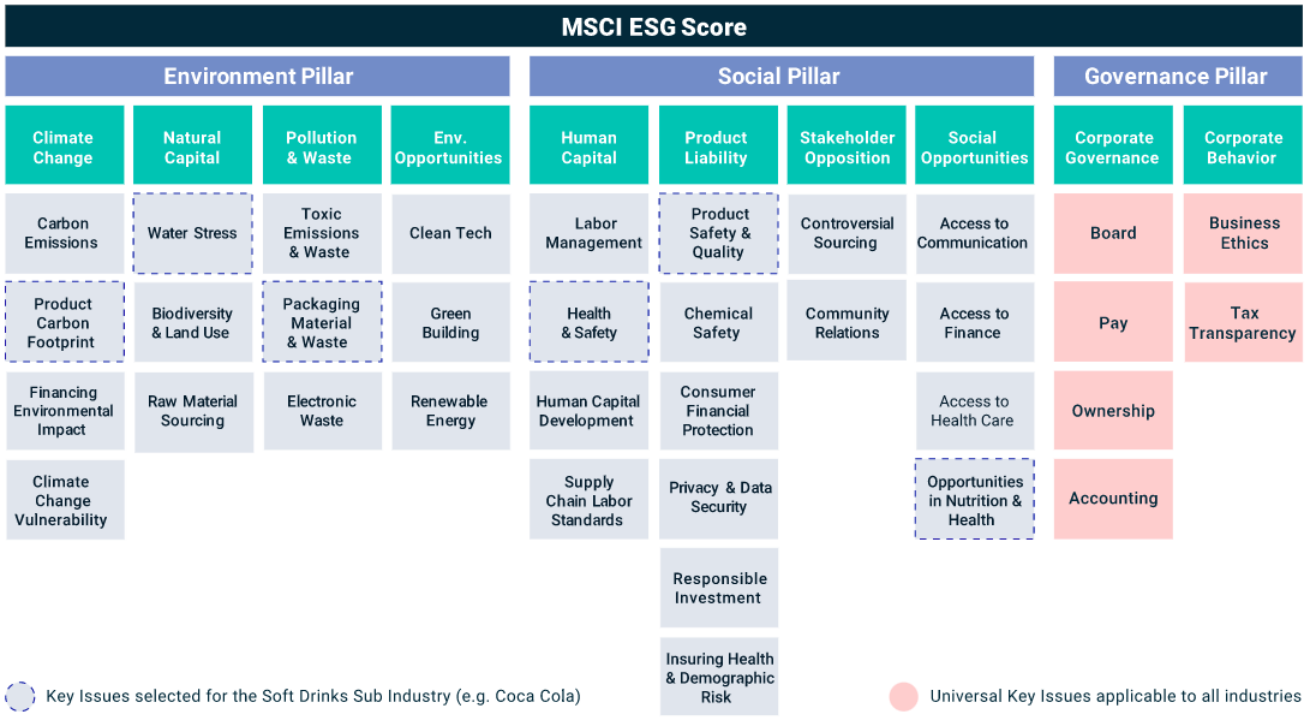
ESG MERGER AND ACQUISITION ACTIVITY (SELECTED)



Source: SustainAbility, Rate the Raters, 2020; research by the authors.

EXHIBIT 3 — ESG RATINGS FRAMEWORKS

MSCI ESG RATINGS: KEY ISSUES FRAMEWORK



FTSE ESG RATINGS MODEL

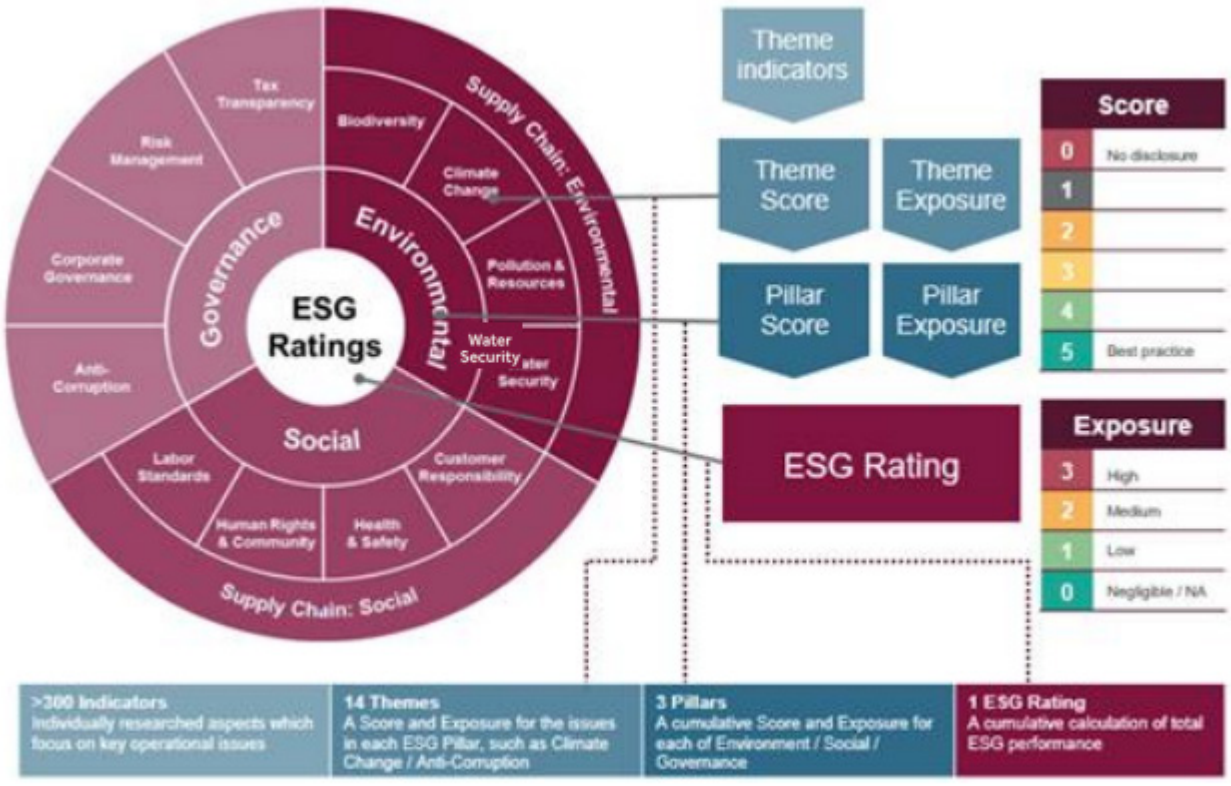
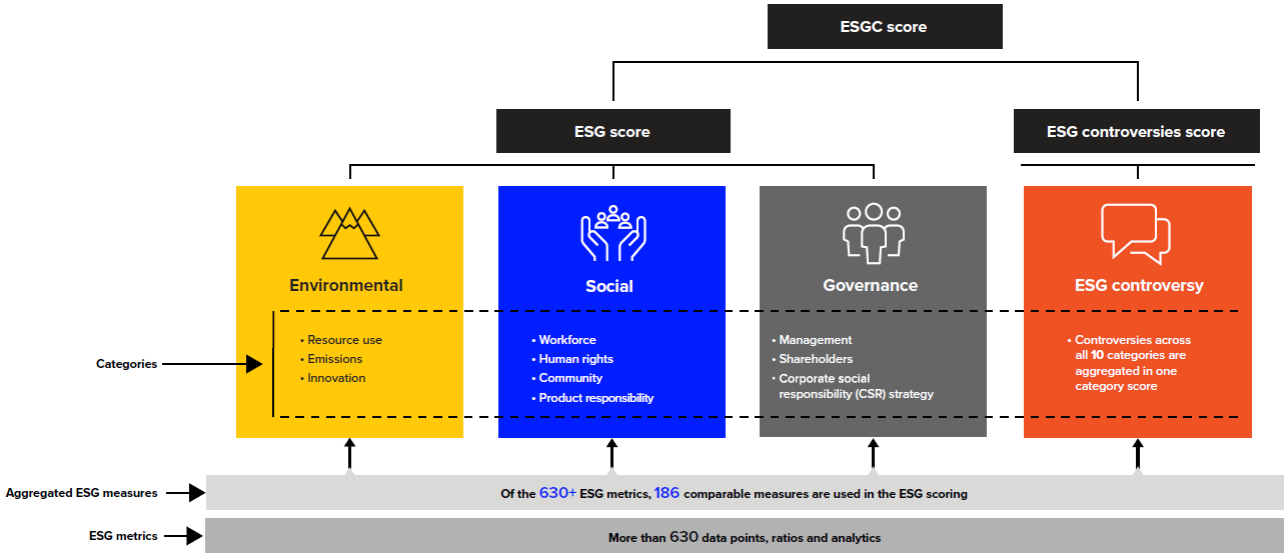


EXHIBIT 3 — CONTINUED

REFINITIV ESG SCORE



S&P GLOBAL ESG SCORES

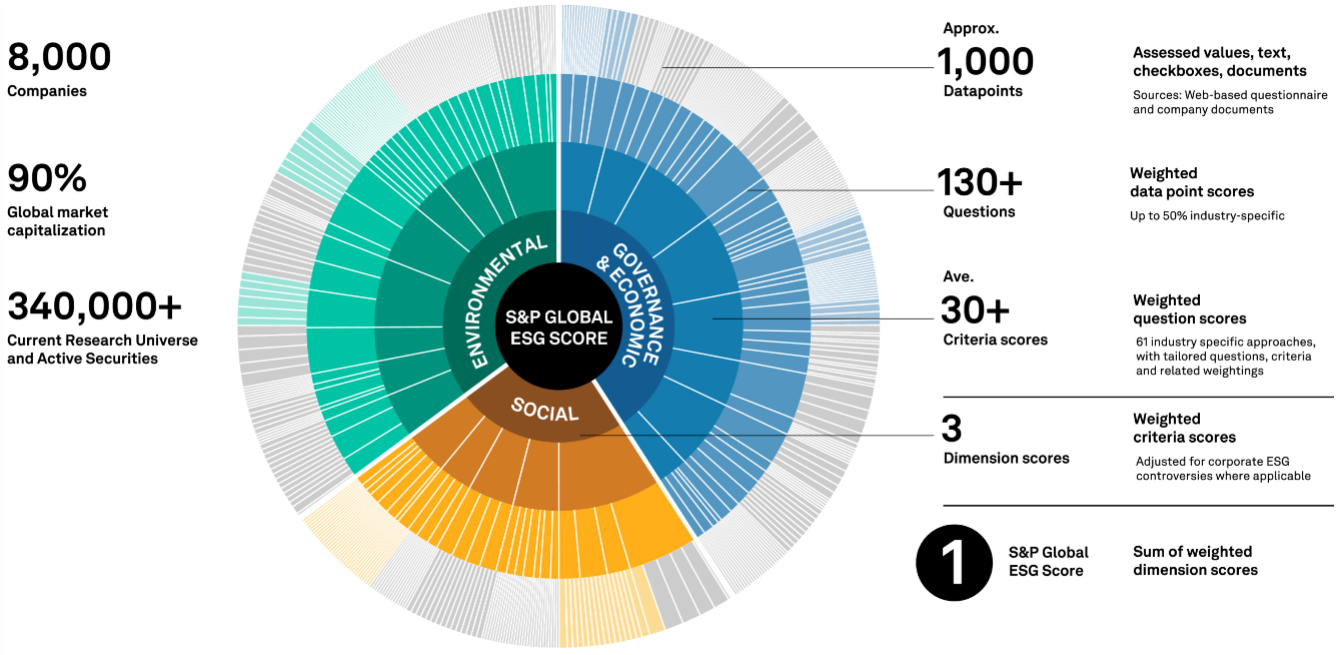
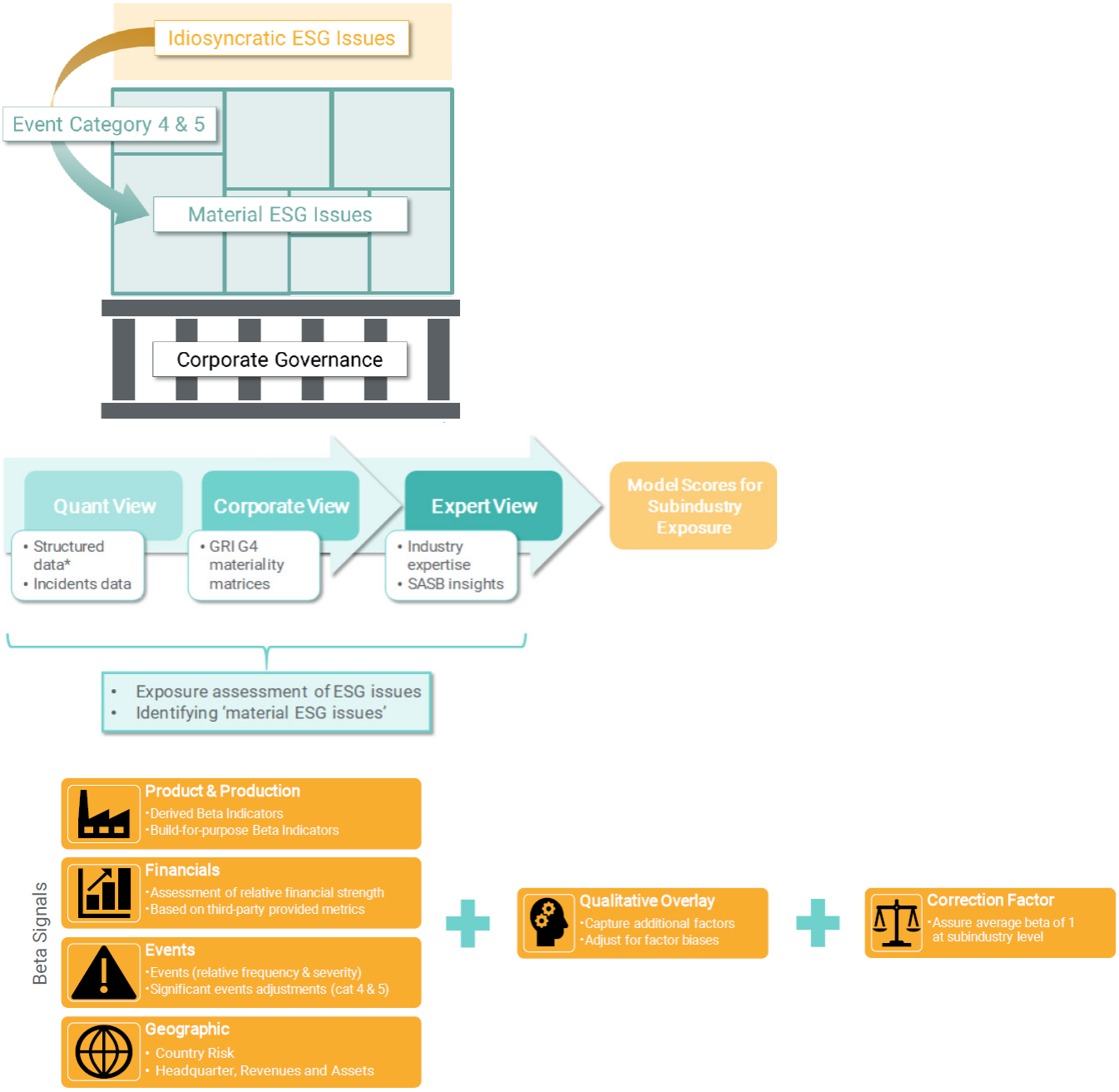


EXHIBIT 3 — CONTINUED

SUSTAINALYTICS ESG RISK RATINGS



Sources: MSCI Key Issue Framework (as of July 2022), available at: https://www.msci.com/our-solutions/esg-investing/esg-ratings/esg-ratings-key-issue-framework; FTSE ESG Ratings Model (as of June 2021), available at: https://research.ftserussell.com/products/downloads/Guide_to_FTSE_Sustainable_Investment_Data_used_in_FTSE_Russell_Indices.pdf; Refinitiv ESG Scores (as of May 2022), available at: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf; S&P Global ESG Ratings (as of July 2022), available at: https://www.spglobal.com/esg/solutions/data-intelligence-esg-scores; Sustainalytics ESG Risk Ratings (as of January 2021), available for download at: https://www.sustainalytics.com/esg-data.

EXHIBIT 4 — CORRELATIONS ACROSS ESG RATINGS PROVIDERS

CFA INSTITUTE (2021)

	MSCI	S&P	Sustainalytics	CDP	ISS	Bloomberg
MSCI	x	36%	35%	16%	33%	37%
S&P	36%	x	65%	35%	14%	74%
Sustainalytics	35%	65%	x	29%	22%	58%
CDP	16%	35%	29%	x	7%	44%
ISS	33%	14%	22%	7%	x	21%
Bloomberg	37%	74%	58%	44%	21%	x

Note: CDP is the Carbon Disclosure Project Score

BERG, KOLBEL, AND RIGOBON (2022)

	KL SA	SL VI	KL RS	KL A4	KL MS	SA VI	SA RS	SA A4	SA MS	VI RS	VI A4	VI MS	RS A4	RS MS	A4 MS	Average
ESG	0.53	0.49	0.44	0.42	0.53	0.71	0.67	0.67	0.46	0.70	0.69	0.42	0.62	0.38	0.38	0.54
E	0.59	0.55	0.54	0.54	0.37	0.68	0.66	0.64	0.37	0.73	0.66	0.35	0.70	0.29	0.23	0.53
S	0.31	0.33	0.21	0.22	0.41	0.58	0.55	0.55	0.27	0.68	0.66	0.28	0.65	0.26	0.27	0.42
G	0.02	0.01	-0.01	-0.05	0.16	0.54	0.51	0.49	0.16	0.76	0.76	0.14	0.79	0.11	0.07	0.30

Note: Note: SA is Sustainalytics, RS is RobecoSAM, VI is Vigeo Eiris, A4 is Asset4, KL is KLD, and MS is MSCI.

Sources: Kevin Prall, "ESG Ratings: Navigating Through the Haze," blog posting at CFA Institute (August 10, 2021); Florian Berg, Julian F. Kölbel, and Roberto Rigobon, "Aggregate Confusion: The Divergence of ESG Ratings," *Review of Finance* (2022).