Keep the Change:

Analyzing the Increase in ESG Ratings for U.S. Equities

April 2022

D E Shaw & Co

Introduction

Amidst an abundance of industry and academic research on environmental, social, and governance ("ESG") investing, we believe one crucial question merits further exploration: have companies improved their ESG-related behaviors over time?

The weighted average ESG rating for companies in the Russell 1000® Index has indeed increased in recent years, as can be seen in *Figure 1*, but that isn't necessarily indicative of changes in the behaviors of companies. This shift could simply be an artifact of the evolving composition of the Russell 1000®, or something akin to "grade inflation" given incremental changes in scoring rules.

After conducting an attribution analysis of Russell 1000® constituents and adjusting for three structural factors, we find that U.S. equities still experienced a meaningful increase in ESG ratings. We also present supplementary research on the distributional concentration of ratings changes and a case study on carbon emissions, each of which indicates that the aggregate ratings increase was driven at least in part by the behavior of individual companies.

We believe this multidimensional analysis offers insights into developments in ESG scoring methodology and how company-specific behaviors may have evolved.

Data Sources

We selected the ESG Ratings framework provided by MSCI Inc. ("MSCI") and the Russell 1000® Index ("R1000") as our key data sources, which in turn informed our decisions regarding three core parameters: (1) the sample period, (2) the individual pillar scores on which we conducted our attribution analysis, and (3) the constituent companies in the R1000 to which we applied the attribution.

By way of background, MSCI publishes an ESG score on a 0–10 scale for over 8,500 companies. Each of the "E," "S," and "G" pillars of the aggregate ESG score comprises a number of constituent "key issues," and pillar scores are derived from the weighted average of the scores assigned to each key issue in that pillar. MSCI uses the three pillar scores to compute an overall ESG score, as well as an AAA–CCC rating, for each company.[1]

With respect to the first parameter, sample period, we chose January 1, 2015 as our start date because MSCI implemented methodological changes as of that date that had material impacts on E, S, and G scores, and complete point-in-time data are not available for periods prior to 2015. The end date for all data presented is December 31, 2021.

Regarding the second parameter, individual pillar scores, we chose to focus our analysis on "E&S" scores. MSCI's G score methodology materially changed on two separate occasions after 2015, which noticeably contributed to large shifts in individual and weighted average G scores.

Figure 1 R1000 Aggregate Score Change
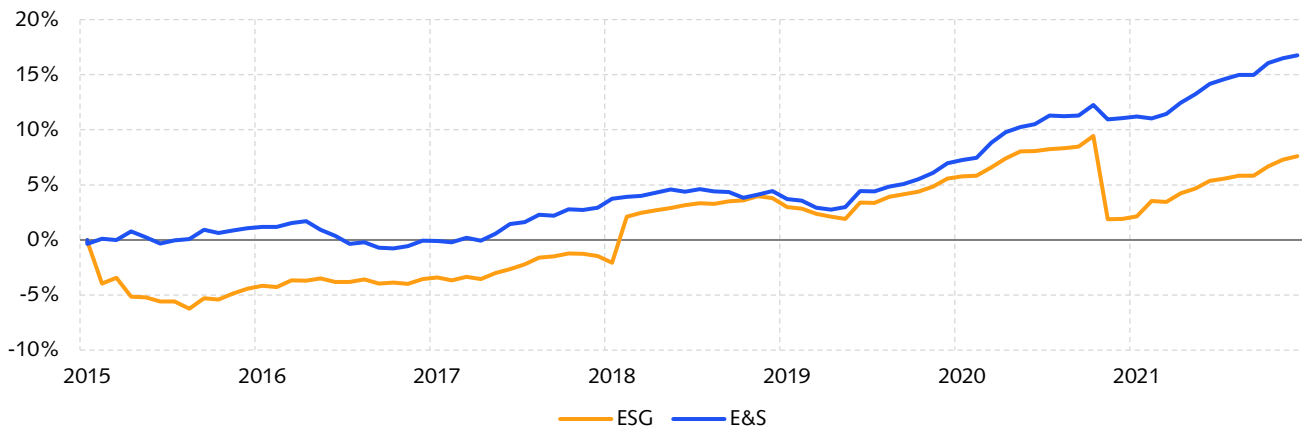
Sources: London Stock Exchange Group; MSCI; the D. E. Shaw group.

---

[1] MSCI provides a visual breakdown of its ESG, pillar, and key issue scores at https://www.msci.com/our-solutions/esg-investing/esg-ratings/esg-ratings-key-issue-framework.

Figure 1 compares the ESG and E&S scores for all R1000 constituents over the sample period. The two abrupt moves in the ESG line in 2018 and 2020 highlight the discontinuities in G scores. Because we concluded it was not feasible to adjust for those methodological changes in a robust manner, we chose to omit G scores from our attribution analysis.

With respect to the third parameter, company selection, for purposes of consistency we focused on the companies that were R1000 constituents over the entire sample period and for which the necessary ESG data were available.[2] For that smaller attribution sample, the total E&S score increase over the period is 18%, rather than the 17% computed for all R1000 constituents that is depicted by the blue line in *Figure 1*.

Structural Factors

Our attribution approach attempts to estimate how much of the raw E&S score increase can be attributed to structural factors that may have systematically biased the score, either positively or negatively, and carries out

sequential adjustments to arrive at an adjusted aggregate score improvement.

(1) Index Composition

The first potential source of bias we explore is the composition of the R1000. The index weights of stocks in the R1000 shifted during the sample period, in some cases materially, and those changes together had a meaningful impact on the index's aggregate E&S score. For example, Microsoft Corp.—which consistently received high MSCI ESG scores over the period—grew significantly from 1.8% to 5.7% of the index's weight. Meanwhile, General Electric Co., one of the lowest-rated companies by ESG score, fell from 1.1% to 0.2% of the index's weight over the period.

With this in mind, we sought to quantify the effects of changing weights on the index's overall E&S score. As shown in *Figure 2*, by conducting a modified Brinson analysis,[3] we find that approximately 5% of the E&S score increase over the period resulted from changes in index weights.

Figure 2 R1000 Attribution Sample E&S Score Change: Index Composition and Residual Effects
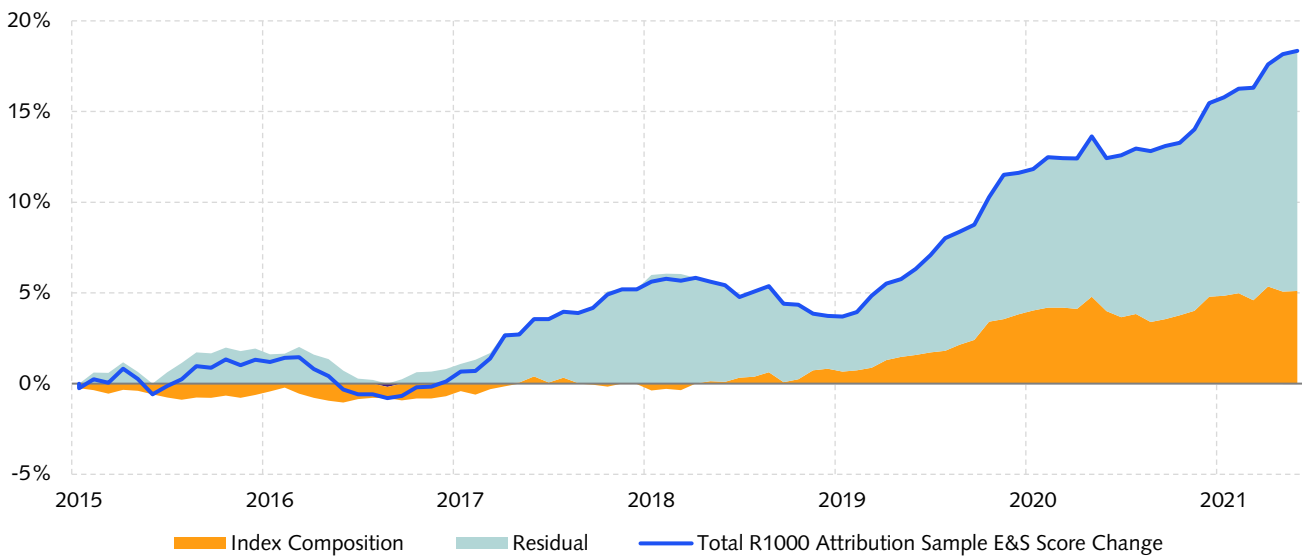
Sources: London Stock Exchange Group; MSCI; the D. E. Shaw group.

---

[2] Our attribution sample comprises 437 companies representing approximately 63% of the R1000 by weight as of January 1, 2015 and 67% as of December 31, 2021. We believe score changes in the sample and the full index track closely. See *Figure 11* in the Appendix for a comparison between the full R1000 and our attribution sample.

[3] We employ a modified version of the Brinson attribution described in Gary P. Brinson and Nimrod Fachler, "Measuring Non-U.S. Equity Portfolio Performance," *The Journal of Portfolio Management* (Spring 1985), Vol. 11, No. 3: 73–76. In *Figure 2*, "Residual" represents the E&S score change when holding index weights constant.
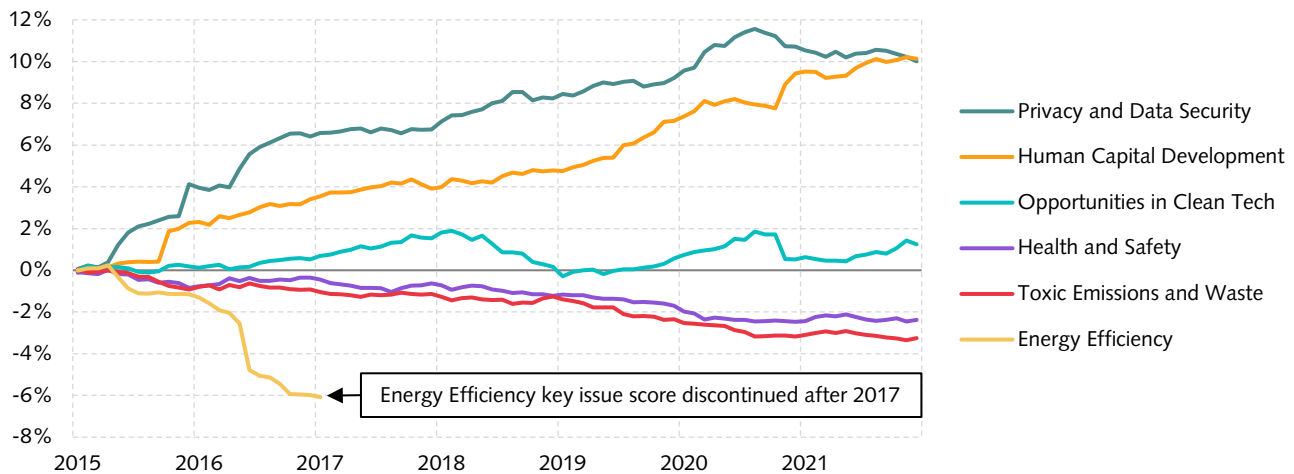
(2) Key Issue Scores

A second potential source of bias in the aggregate E&S score increase relates to changes in the weights assigned to the key issues that constitute each of MSCI's pillars. *Figure 3* plots the three E&S key issues with the largest total increases and the three with the largest total decreases in weight in the R1000 attribution sample over the period. As with the index weightings, some of these shifts during the sample period were clearly meaningful; in one case, for example, MSCI eliminated a key issue (Energy Efficiency) altogether. Shifts in key issue score

weights can result in substantial changes to overall pillar scores and the relative weightings of each pillar.

To capture those potential distortions, we conduct a second Brinson analysis, as shown in *Figure 4*. We find that the E&S score increase, when adjusted for constant R1000 index weights as reflected in *Figure 2*, would have been 4% *larger* if the relative weights assigned to key issue scores had also remained constant over the period. The attributed value for "Key Issue Score Weight Allocation Effect" reflected in *Figure 4* illustrates the bias introduced by this structural element.

Figure 3 MSCI E&S Key Issue Score Weights in the R1000 Attribution Sample: Largest Positive and Negative Moves

Sources: MSCI; the D. E. Shaw group.

Figure 4 R1000 Attribution Sample E&S Score Change: Allocation to Key Issue Scores and Score Changes
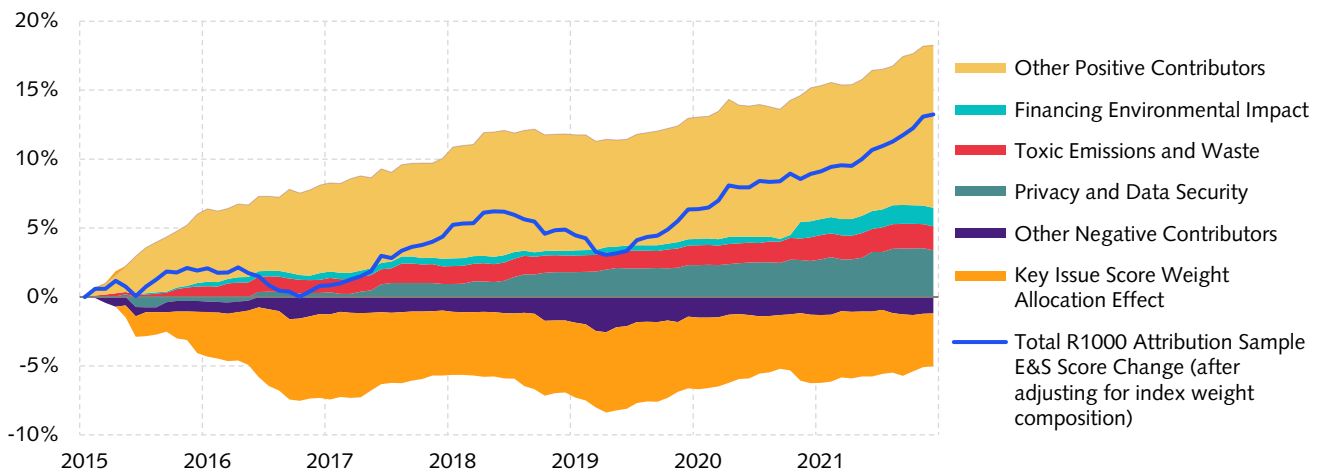
Sources: London Stock Exchange Group; MSCI; the D. E. Shaw group.
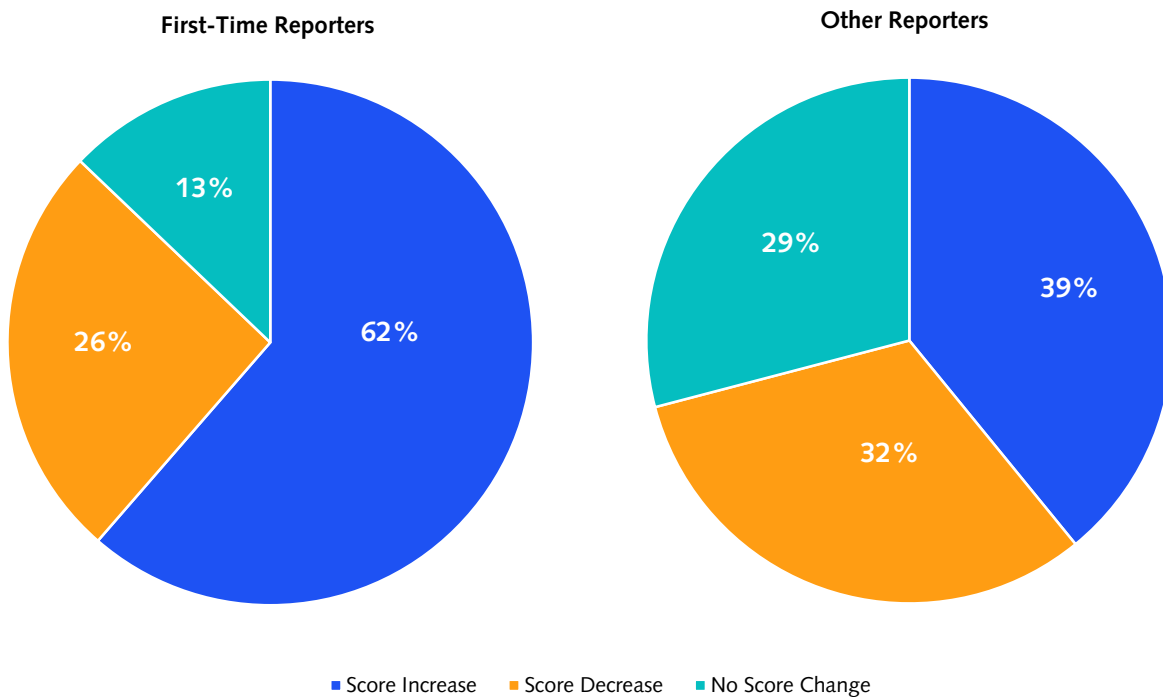
(3) The Effects of Disclosure

Companies generally have increased their reporting of key ESG metrics over time, particularly in recent years, as stakeholder focus on ESG has intensified. For example, 346 companies in the R1000 reported carbon emissions data at the start of 2015; by the end of 2021, that number had risen to 507. Within our attribution sample of 437 index constituents, 193 reported emissions data at the start of the period compared to 294 at the end.

Focusing on this emissions example, we wanted to assess whether the relationship between disclosure of carbon emissions data and MSCI E scores was one that might have inflated the index's aggregate E&S score over the sample period. *Figure 5* shows that companies in the full R1000 that reported carbon emissions for the first time in any calendar year during the period were significantly more likely to experience an increase in their E score than companies that did not alter their reporting of carbon emissions or ceased reporting altogether.

To estimate the overall co-movement between E&S score improvement and increased disclosure, for each company in our sample, we ran a regression of changes in its E&S score (adjusted for changes in key issue score weights) on changes in its Bloomberg Finance L.P. ("Bloomberg") ESG disclosure score.[4]

Figure 5 Distribution of Year-Over-Year E Score Directional Changes: First-Time and Other Reporters of Carbon Emissions in the R1000 (2015–2021)

First-Time Reporters: Score Increase 62%, Score Decrease 26%, No Score Change 13%

Other Reporters: Score Increase 39%, Score Decrease 32%, No Score Change 29%

■ Score Increase   ■ Score Decrease   ■ No Score Change

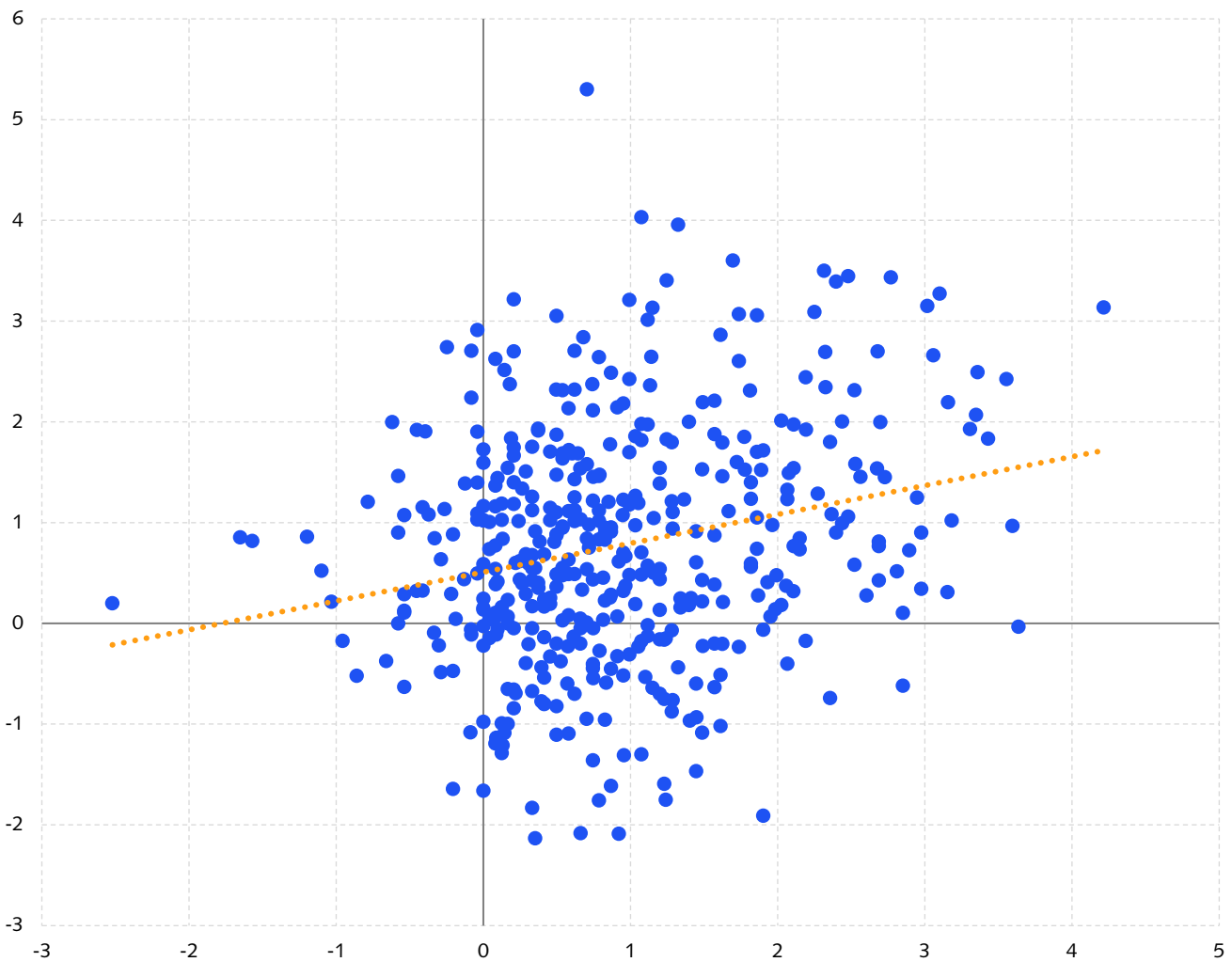*Sources: London Stock Exchange Group; MSCI; the D. E. Shaw group.*

[4] We relied on Bloomberg's ESG disclosure score data because MSCI does not report an ESG disclosure score or similar metric. Bloomberg reports these scores on a 0–100 scale, which we rescaled to 0–10 for interpretability relative to MSCI's ESG scores. We further lagged this disclosure data in an effort to more closely align the two data providers' reporting windows.

Figure 6 presents the results of that regression, which show a statistically significant (t = 5.3) relationship between E&S score and disclosure.

For every 1 point increase in its scaled Bloomberg disclosure score over the period, a company tended to experience a 0.3 point increase in its E&S score (*i.e.*, exhibited a "beta" of 0.3). The product of that beta and the weighted average percentage change in scaled Bloomberg disclosure scores for all companies in our attribution sample is 5%, meaning that the aggregate E&S score increase may be biased higher by up to 5% because of the disclosure effect associated with MSCI's ESG scores.

Figure 6 Changes in Adjusted E&S Scores (y-axis) Relative to Changes in Scaled Bloomberg ESG Disclosure Scores (x-axis)
(2015–2021)

Sources: Bloomberg; London Stock Exchange Group; MSCI; the D. E. Shaw group.

Sequentially Adjusting the Aggregate E&S Score Change

As we have noted, the aggregate E&S score for our attribution sample increased by approximately 18% on an unadjusted basis over the seven-year period. *Figure 7* shows the results of sequentially adjusting that raw score increase to account for the three structural effects we discuss above: changes in index composition, changes in key issue score weights, and estimated co-movement of score improvement with disclosure.

After making those adjustments, we arrive at a residual score increase of approximately 12%. While other structural effects may have influenced this residual in either direction, it seems plausible to us that changes in company-specific behaviors explain a meaningful amount of this residual increase. (The case study on carbon emissions in the penultimate section of this paper considers that possibility in greater depth.)

Figure 7 R1000 Attribution Sample E&S Score Change with Successive Adjustments
(2015–2021)
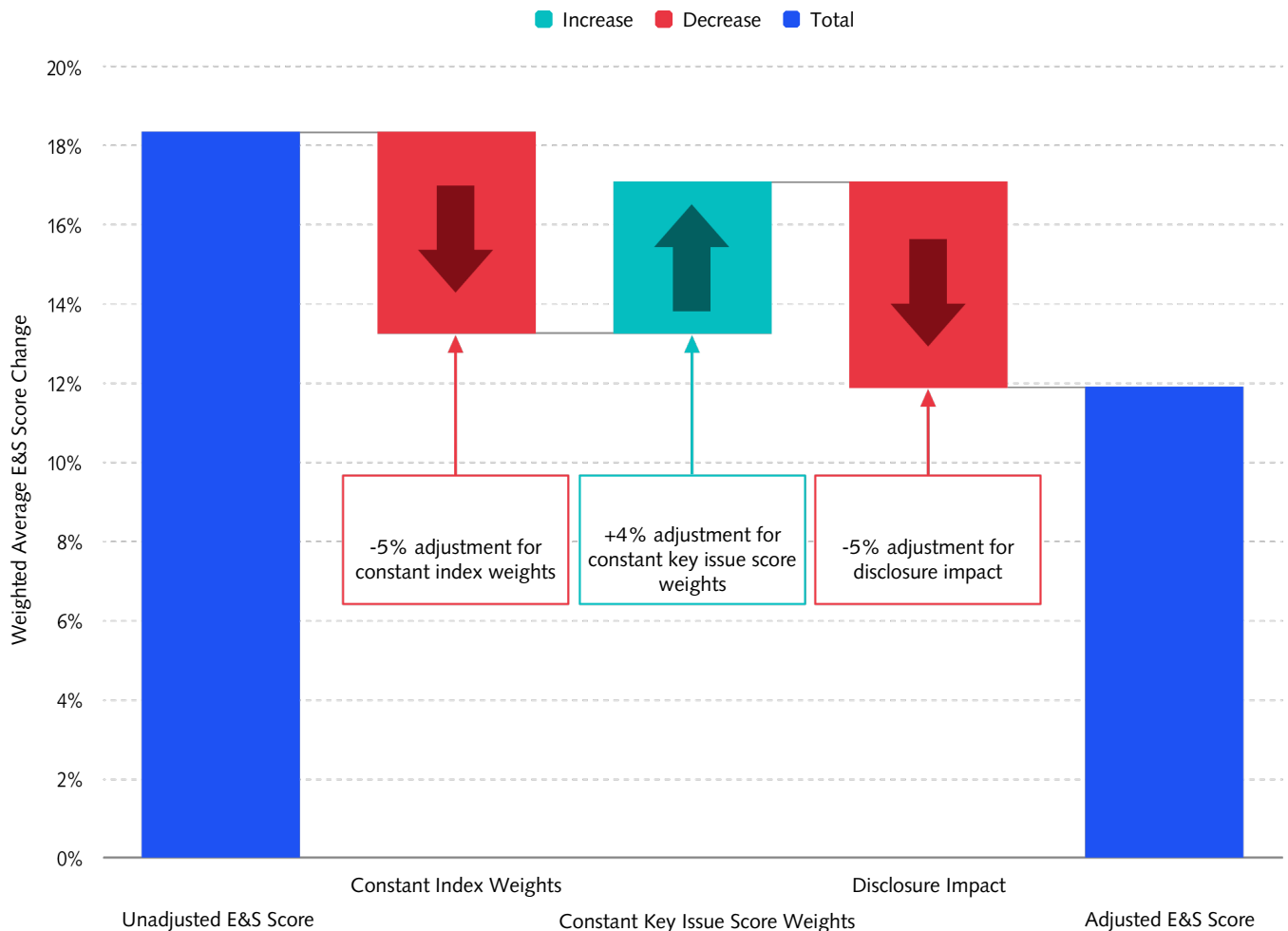
Sources: Bloomberg; London Stock Exchange Group; MSCI; the D. E. Shaw group.

Improvements Cluster among Lower Scorers

Having adjusted for those three distorting effects, we can examine how the adjusted E&S score increase is distributed across our sample. We find that companies that had relatively low E&S scores at the start of the period are more likely to have demonstrated score improvement.

Figure 8 assigns each company in our attribution sample to one of five segments based on its E&S score at the start of the period: AAA–AA, A, BBB, BB, or B–CCC. We compute the percentage of companies that transitioned from one segment to another, as well as the percentage that remained within the same segment at the end of the sample period. The observed score increase skews towards those companies in initially lower-rated cohorts of the sample.

Having established that companies with initially lower E&S scores are more likely to have shown score improvement, we now examine the relationship between initial E&S score and the relative size of the score increase. To do so, we assigned each company in the attribution sample to a decile by its starting E&S 0–10 numerical score, and then plotted the weighted average score change per decile (after making the same adjustments described in prior sections). *Figure 9* shows that companies in the lower deciles (1–3) exhibit a significantly larger score increase than those in the middle (4–7) or higher deciles (8–10). Although the higher deciles (highest initial scores) decrease on average, the weighted magnitude of these decreases is less than the corresponding increases experienced by the lowest scorers.

Figure 8 R1000 Attribution Sample E&S Score Transitions
(2015–2021)
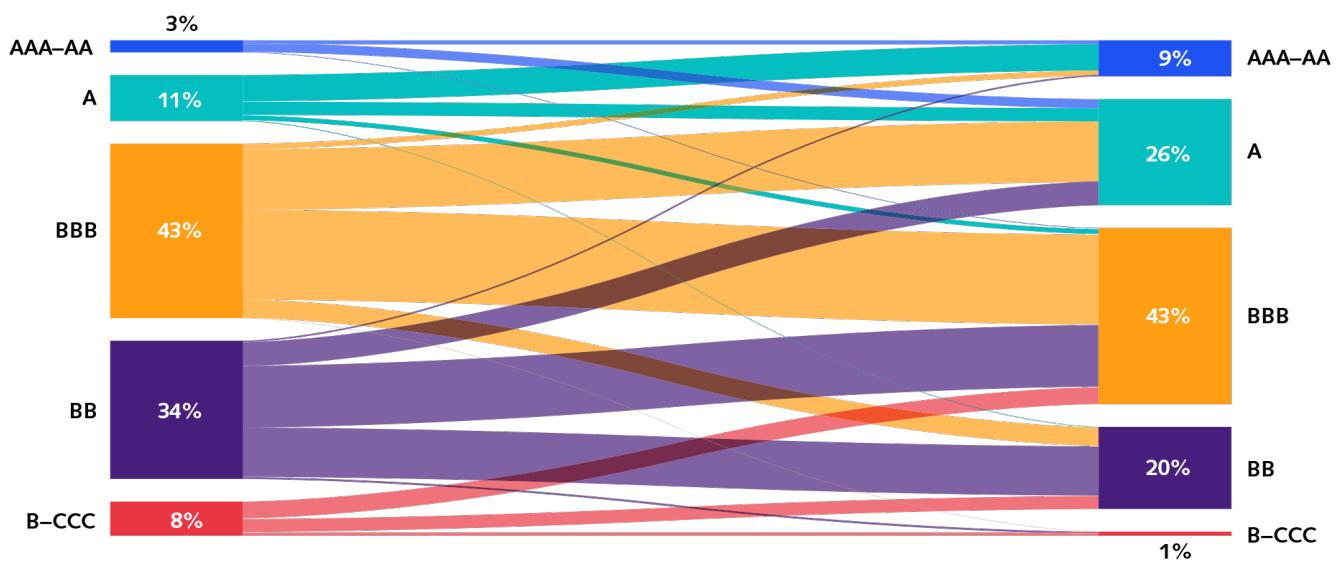
Sources: Bloomberg; London Stock Exchange Group; MSCI; the D. E. Shaw group.

Figure 9 R1000 Attribution Sample E&S Score Change by Starting Score Decile
(2015–2021)
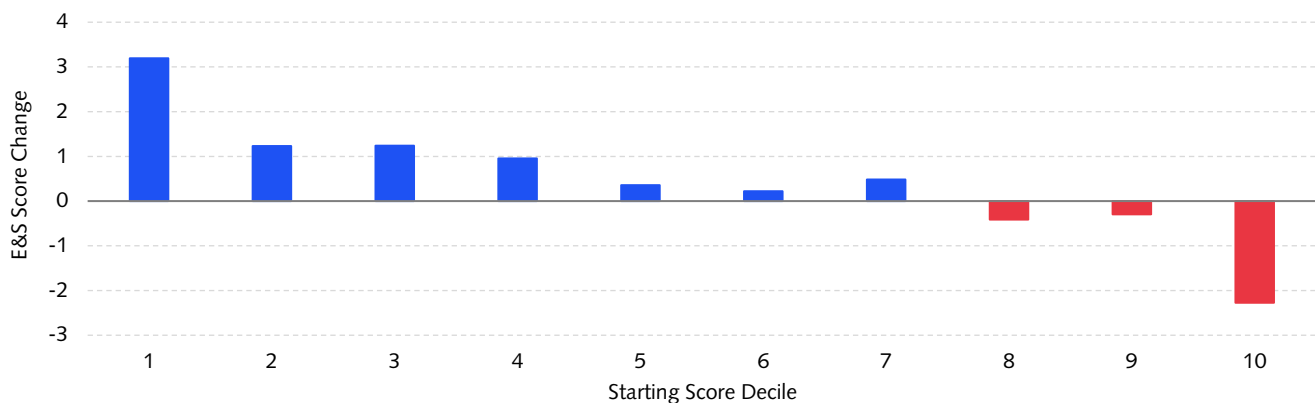
Sources: Bloomberg; London Stock Exchange Group; MSCI; the D. E. Shaw group.

Case Study: Carbon Emissions

In the attribution analysis described above, we adjust aggregate E&S scores for three distinct structural factors. However, it's entirely possible that some of the remaining score increase is driven by structural factors not captured in our analysis. We conducted a case study on carbon emissions to evaluate whether companies have indeed changed their behaviors in ways that lead to higher scores.

We employed two emissions metrics: carbon intensity ("CI"), as measured by tons of MSCI Scope 1 + 2 emissions[5] per $1 million of sales, and carbon footprint ("CF"), as measured by tons of Scope 1 + 2 emissions per $1 million of market capitalization. We applied those two metrics to a sample of the R1000 for which carbon emissions data were available over the 2015–2021 period.[6] *Figure 10* shows that, on a weighted average basis, both emissions metrics declined over the sample period, whether or not index weights are held constant. These results suggest that, at least for carbon emissions, company-specific behavioral changes are captured by the overall E&S score increase. Although that may not be true for all key issues driving ESG scores, it is encouraging to see scores reflect at least certain "real world" ESG changes.

Conclusion

This paper applies statistical techniques to shed light on the observed increase in aggregate ESG scores since 2015. The results of our sequential adjustments indicate that, of the approximately 18% unadjusted aggregate E&S score improvement observed in our attribution sample, roughly 12% remains after accounting for the structural effects described above. Additionally, we find that score increases are generally concentrated in issuers that began the period with relatively low E&S scores, suggesting that companies at the bottom of the score distribution drove the improvement.

We believe that a multi-dimensional attribution approach is valuable not only for assessing the overall trend in ESG scores but also in evaluating how such scores map to companies' adoption of ESG-friendly practices. As ESG investing continues to grow and evolve, it will be critical for investors and analysts to apply rigorous approaches to ESG data, trends, and observations—not only to inform their own investment decisions, but also to help them understand the connections between ESG ratings and ESG behaviors at the company level.

Figure 10 R1000 Case Study Sample Carbon Intensity and Carbon Footprint
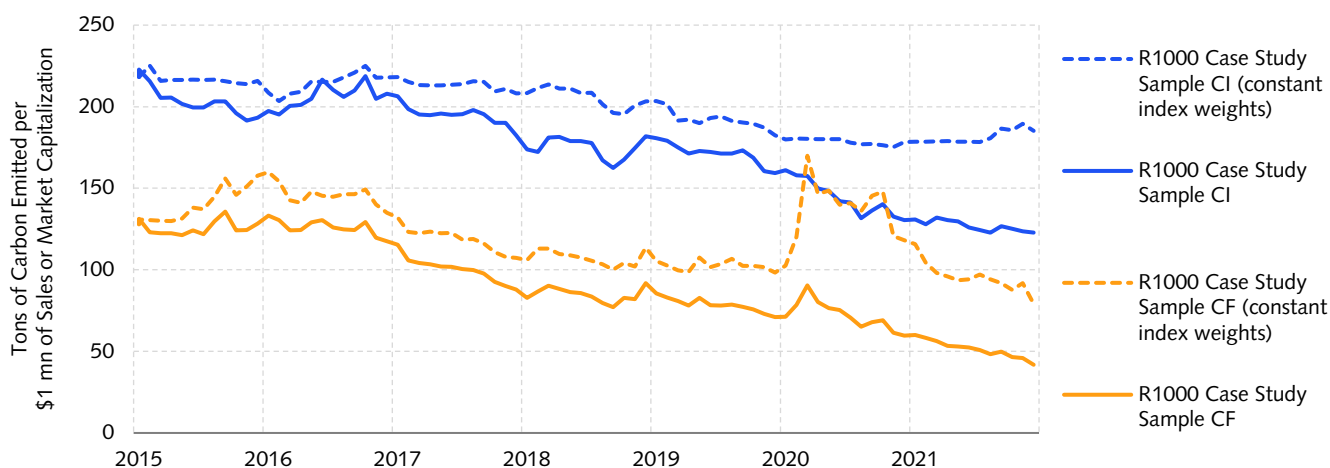
Sources: MSCI; Reuters DataScope; the D. E. Shaw group.

---

[5] Scope 1 emissions cover direct emissions from a company's owned or controlled sources. Scope 2 emissions cover indirect emissions from a company's energy consumption (*e.g.*, electricity purchased).

[6] Because of differences in data availability, the 581 companies in this case study sample are not the same as those included in the attribution sample. We believe this case study sample is representative of the full R1000 over this period, as supported by *Figure 12* in the Appendix.

Appendix

For various portions of our analysis, we relied on subsets of R1000 constituents, both for methodological reasons, such as the desire to maintain a constant sample throughout the period, and because of data availability. As can be seen in the two figures below, the relevant metrics for those index subsets generally track those of the full R1000, indicating that our conclusions are likely to be applicable to the index itself.

As can be seen in *Figure 11*, the aggregate E&S score of our attribution sample increased by approximately 18%, compared to the index's 17%. At any given point in time, the two data series differ by no more than 1.9%, and the monthly score changes within the two series exhibit a correlation of 0.92. That said, it's still important to note that the attribution sample's slightly higher increase in weighted average score may suggest that the magnitude of the attributed effects could be somewhat larger for the attribution sample than for the full index.

Figure 12 plots the weighted average CI and CF for the emissions-specific case study sample and the full R1000, indicating a close relationship between those two data series. For both carbon metrics, the full index and case study sample converge over time, and the two series' monthly percentage changes demonstrate correlations of approximately 0.87 (CI) and 0.98 (CF). In this case, the observed decrease in carbon emissions is slightly larger for the full R1000 than for the case study sample, which suggests that changes in company behavior may have had a greater impact in reducing carbon emissions in the context of the full index than in the case study sample.

Figure 11 E&S Score Change by Index Sample
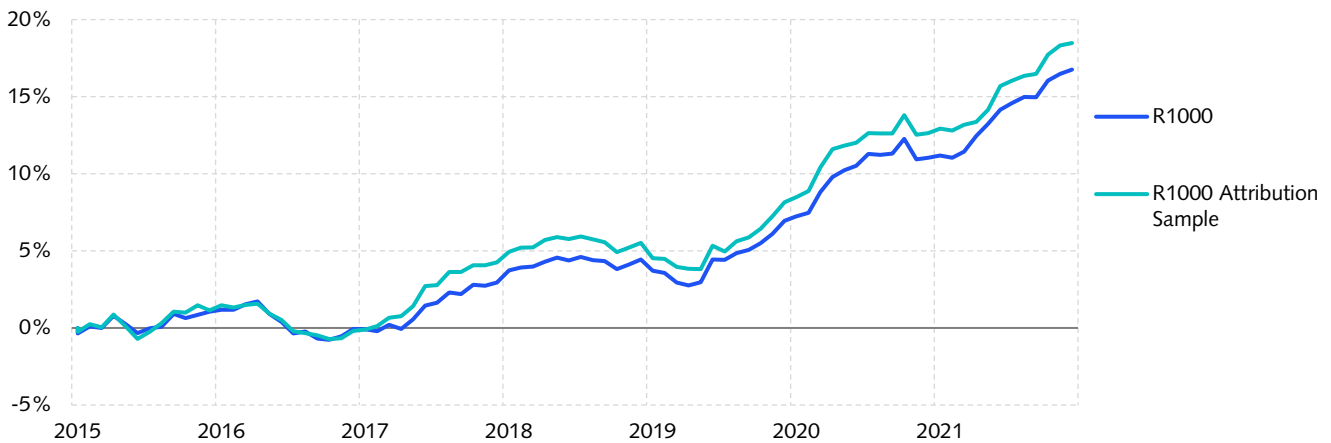
Sources: London Stock Exchange Group; MSCI; the D. E. Shaw group.

Figure 12 Carbon Intensity and Carbon Footprint by Index Sample
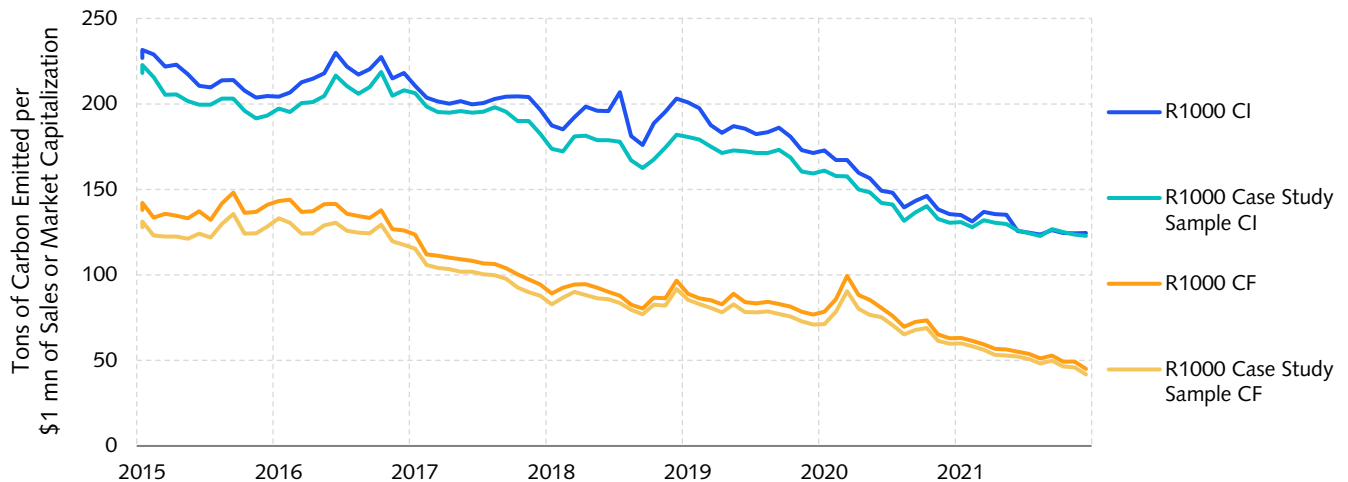
Sources: London Stock Exchange Group; MSCI; Reuters DataScope; the D. E. Shaw group.

Note Regarding the Use of Certain MSCI Inc. Data