ESG Ratings of ESG Index Providers*

Sonakshi Agrawal[1], Lisa Yao Liu[1], Shiva Rajgopal[1], Suhas A. Sridharan[2], Yifan Yan[2], and Teri Lombardi Yohn[2]

[1]Columbia Business School, Columbia University
[2]Goizueta Business School, Emory University

September 2023

*This paper integrates two earlier working papers: "ESG Ratings of ESG Index Providers" by Agrawal, Liu, and Rajgopal, and "ESG Rating Agency Incentives" by Sridharan, Yan, and Yohn. We appreciate helpful comments from Jonathan Berkovitch, Matthias Breuer, Dane Christensen, Jacquelyn Gillette, Jon Glover, Jeff Hales, Bianca He, Doron Israeli, Michael Jantzi, Sehwa Kim, Brian Williams, Leifu Zhang, Luigi Zingales, Christina Zhu, and workshop participants at CCSRG at the University of Chicago, Columbia Business School, Emory University, University of Miami, 2022 University of Chicago Stigler Center Affiliate Fellows Conference, 2023 Tel Aviv university conference, 2023 Southeastern Summer Accounting Research Conference, and 2023 Stanford Accounting Summer Camp. We acknowledge financial assistance from our respective schools. All errors are ours.

ESG Ratings of ESG Index Providers

September 2023

Abstract

Despite growing investor reliance on environmental, social, and governance (ESG) ratings, we know relatively little about how such ratings are constructed especially because widespread disagreement across ESG ratings raises concerns about their credibility. At the same time, several leading ESG raters not only construct ESG ratings but also market index products based on their ESG ratings. We examine whether the incentives associated with deriving revenue from ESG rating-based indices contribute to the variation in ESG ratings. Consistent with this notion, we find that raters with strong index licensing incentives issue higher ESG ratings for firms with better stock return performance and those added to their ESG indexes, compared to raters with weaker licensing incentives. The results hold after accounting for the firm's fundamental ESG performance and different rating methodologies. Overall, our findings suggest that index construction incentives affect the production of ESG ratings, highlighting the need for greater transparency in the production of ESG ratings.

Keywords: ESG, index providers, rating agencies, sustainability, disclosure
JEL Classifications: G24, M14, M40, M41, M48, Q56

1 Introduction

Investing with an environment, social, and governance (ESG) focus has experienced remarkable growth in recent years. ESG assets under management (AUM) surpassed $35 trillion in 2020 and are expected to comprise over half of all professionally managed assets globally by 2024 (Henze and Boyd, 2022; Taylor and Collins, 2022). Successful ESG investing requires reliable information about corporate ESG activities, but corporate disclosure in this arena is unregulated and often opaque. ESG ratings purport to offer a solution to this challenge by providing investors with a succinct representation of firms' complex ESG activities, and evidence suggests that investors rely heavily on ESG ratings (Berkovitch et al., 2022; Hartzmark and Sussman, 2019). However, there is growing awareness of divergence among ESG data providers, which raises concerns about the credibility of ESG ratings (Berg et al., 2022; Chatterji et al., 2016; Christensen et al., 2021).[1] To date, it is unclear how ESG data providers determine ESG ratings.

In this paper, we consider that while ESG raters face incentives to provide credible information, their alternative business lines – particularly those related to stock index construction and licensing - could also influence the production of ESG ratings. Several leading ESG ratings providers also use their own ESG ratings to construct ESG indices that are licensed to asset managers as benchmarks for mutual funds and exchange-traded funds. Asset managers typically pay index providers a fee that is a percentage of the fund's assets under management (AUM). Studies document a strong positive association between past equity return and future fund flows, which in turn leads to higher AUM and licensing revenues (An et al., 2023; Berk and Green, 2004). Thus, index licensing creates incentives to construct indices with high stock performance. We compare ESG ratings across raters who derive large and small portions of their revenue from ESG-based stock index licensing. If index licensing influences the production of ESG ratings, we expect that raters with strong index licensing incentives issue higher ESG ratings for firms with better stock return performance and those added to their ESG

[1] Media reports have also raised concerns about the potential conflict of interest in the ESG rating providers. As per an article in the *Financial Times*, the outperformance of MSCI ESG Indexes does not seem to be related to the underlying firms' ESG credentials (Boyde, 2022). Regulators in the European Union, the United Kingdom, and the United States started to consider regulating the ESG ratings industry (ESMA, 2022; FCA, 2022; IOSCO, 2021).

indexes, compared to raters with weaker licensing incentives.

We compare ESG ratings for a rater with high index incentives ($HighIndex$) and a rater with low index incentives ($LowIndex$). Specifically, we consider two of the leading providers of ESG ratings data: MSCI and Refinitiv. We consider MSCI as an example of an ESG rater with high index licensing incentives. Eight percent of MSCI's operating revenue derives from the direct sale of ESG ratings and other climate data to investors, while more than 60 percent of MSCI's operating revenue comes from sales of index products that are used as benchmarks by asset managers. We consider Refinitiv as an example of an ESG rater with low index licensing incentives. Like MSCI, Refinitiv is one of the largest ESG data providers of numeric ESG ratings; however, there is no significant ESG Index that relies on Refinitiv's ratings during our sample period.

We expect $HighIndex$ raters to face stronger incentives than $LowIndex$ raters to modify their ESG ratings for the purposes of index construction. We note, however, that an important component of data vendors' business model is the credibility they establish with clients. If investors perceive that a rater's ESG ratings do not accurately reflect the constructs they purport to measure, this perception could lead them to pursue alternative data providers and index products. Thus, even raters with strong index licensing incentives may also face incentives to provide ESG ratings that bolster their credibility. Moreover, $HighIndex$ raters might not view fund performance as critical to sustaining or increasing AUM, given recent evidence that fund flows are positively related to sustainability perceptions (Hartzmark and Sussman, 2019).

Considering these conflicting pressures, whether index construction and licensing incentives affect ESG ratings is an empirical question. We explore this question by testing if the difference between $HighIndex$ and $LowIndex$ ESG ratings for the same firm is associated with the firm's stock return performance. By comparing ESG ratings across $HighIndex$ and $LowIndex$ for the same firm, we control for the ESG activities of the firm. We further analyze whether $HighIndex$ ESG ratings are systematically higher or lower than those of $LowIndex$ for stocks added to or dropped from $HighIndex$ ESG indexes. We control for differences in such ratings across time, industry, the specific ESG index used, and different rating methodologies. Therefore, unobservable and omitted variables at the firm and the rater level are unlikely to explain the difference between the ratings.

To be in our sample, firms must receive ESG ratings from both $HighIndex$ and $LowIndex$ raters. The resulting sample comprises 7,214 firm-year observations for 1,691 unique US firms from 2012 to 2019. For each firm-year, we measure the difference between $HighIndex$ and $LowIndex$ ESG ratings (ESG_Diff). If index licensing incentives influence ESG ratings, we expect ESG_Diff to be significantly associated with measures of stock return performance: stock returns ($Return$) and book-to-market (BTM). Specifically, ESG_Diff will be positively associated with $Return$ and negatively associated with BTM. We also examine whether ESG_Diff is associated with firm profitability, namely, return on assets (ROA). We include ROA to consider the possibility that higher stock returns are associated with higher profitability, which may allow for greater ESG investment. As controls, we also include several variables shown to be associated with ESG rating divergence: firm size, analyst coverage, institutional ownership, and the level of ESG disclosure. To account for unobservable variation in ESG ratings, we include industry and year fixed effects in our estimation.

We find that firms with better stock return performance receive higher ESG ratings from $HighIndex$ raters relative to $LowIndex$ raters, suggesting that $HighIndex$ raters emphasize stock return performance when determining ESG ratings. To better understand whether the potential link between ESG ratings and stock return performance arises from the relative strength of raters' index licensing incentives, we also study whether stock return performance influences how $HighIndex$ raters change their index compositions. Specifically, we disaggregate the $HighIndex$ overall ESG ratings into a component that is related to stock return performance and a residual component that is unrelated to stock return performance. We then explore whether these components of $HighIndex$ ratings are associated with the likelihood of index inclusion as well as the intensity of index inclusion, measured as either the total number of indices in which the firm is included, or the cumulative weight assigned to the firm across the rater's indices. If index licensing incentives influence ESG ratings, we expect that index inclusion will be associated with the stock return performance component of ESG ratings. We find that the likelihood of index inclusion by $HighIndex$ raters is related to both the stock return and residual components of ESG ratings. However, changes in the intensity of *HighIndex* index inclusion are significantly associated with the stock return performance component of the ESG rating and not associated with the residual component. These findings suggest that both non-stock return performance factors and stock return performance impact index inclusion decisions, but only stock return performance impacts the intensity of index inclusion decisions.

To further corroborate the inference and strengthen the identification, we investigate how ESG_Diff changes with ESG index additions and deletions. To account for differences in rating methodology across rating agencies, we use the "stable" firms in our sample of ESG indices, those that were not added or deleted by the *HighIndex* rater, as a control group. Using a staggered difference-in-difference analysis where each index addition or deletion is a treatment event, we hold the firm's "fundamental" ESG performance relatively constant but vary the ESG rating agencies considered. This analysis controls, as best as we can, for various factors that could affect the ratings, such as differences in rating methodologies, firm fundamentals, methodology underlying the construction of a specific ESG index, and unobserved time-invariant factors.

We find that, relative to *LowIndex* raters, *HighIndex* raters assign systematically higher (lower) ESG ratings to firms being added to (deleted from) their ESG indices. By increasing subsequent ESG ratings for firms included in the index, *HighIndex* can potentially boost the marketability of its ESG index, which, in turn, suggests that its index business potentially influences its ESG ratings.[2] These results further support that index licensing incentives influence ESG ratings.

It is possible that our main results arise because *HighIndex* raters have access to private information about the firms' ESG fundamentals due to their index businesses. It is also possible that changes in the index composition itself prompt firms to improve their ESG performance. Again, it is unclear why such changes would lead to divergence in ratings from different ESG raters. Nonetheless, we examine the disclosed component scores that underlie

[2]For instance, MSCI promises investors that its ESG index "provides exposure to companies with high ESG performance" and is "designed to target companies with positive ESG characteristics."

the *HighIndex* and *LowIndex* ratings. We first consider the extent to which firms' overall ESG ratings relate to their ESG component scores for both *HighIndex* and *LowIndex* raters. We observe that variation in firms' *LowIndex* ESG component scores explains 97.1 percent of the variation in firms' overall *LowIndex* ESG ratings. In contrast, variation in firms' *HighIndex* ESG component scores explains only 74.1 percent of the variation in firms' overall *HighIndex* ESG ratings. After considering the individual ESG component scores, we also observe that stock returns are a strong incremental predictor of *HighIndex* overall ESG ratings but not *LowIndex* overall ESG ratings. The results suggest a contradiction to the notion that *HighIndex* raters have greater private information regarding the firms' ESG fundamentals.

Furthermore, we examine whether the addition of a firm to an ESG index or changes in the firm's ratings are correlated with firms' subsequent ESG outcomes (e.g., regulatory violations, gender diversity, racial diversity, and climate change exposure) following Heath et al. (2021). We do not find evidence supporting changes in subsequent fundamental ESG performance. We also visually plot rating changes prior to and following index composition and find that the *LowIndex* ESG ratings remain largely stable throughout the period. In contrast, *HighIndex* ESG ratings show an increase (a decrease) after stocks have been added to (dropped from) an ESG index. If *HighIndex* raters have private information on the ESG performance of added or removed firms, or if reconstituting the index causes a real change in firms' ESG behavior, we would expect the ratings of the other rating agencies to "catch up" with *HighIndex* rating changes. However, this is not what we find in the data. The weight of the evidence presented does not support the alternative explanations.

Our paper contributes to several streams of literature. First, we extend the nascent literature examining the production of corporate ESG ratings. Several recent studies highlight substantial disagreement among ESG data providers and attribute this disagreement to factors such as the number of ESG disclosures made by the firm and the measurement strategies used by ESG rating agencies (Berg et al., 2022; Christensen et al., 2021; Brandon et al., 2021). We investigate a new factor underlying disagreement across ESG rating agencies: the competing incentives that arise from index licensing business lines. While the factors identified by prior research could easily be interpreted as reflecting noise in ratings, we posit that rater's competing incentives induce more systematic differences in ratings.

A natural question that emerges when comparing multiple ESG ratings relates to the relative accuracy of the ratings. In contrast to credit ratings that can be validated by future default events (Piccolo and Shapiro, 2022), ESG ratings are less verifiable and entail less risk of incorrect ratings attribution compared to credit rating agencies. Thus, our analyses cannot speak to the relative accuracy of different ESG raters. However, our analyses do extend the existing literature by explaining *why* raters adopt different measurement strategies. This extends prior research that documents ESG rating disagreement but remains agnostic regarding its source (Larcker et al., 2022).

Second, in highlighting the incentives that ESG raters face and the potential consequences of those incentives, our study improves our understanding of the causes of divergence across ESG ratings. In doing so, we can help inform investors who rely on this information for their trading decisions. Our evidence can also inform regulators who are increasingly interested in the ESG investing landscape. For example, the Securities and Exchange Commission (SEC) is aware of the "widely different ratings" across ESG data providers and doubts if these ratings "can effectively guide investment decisions" (Pierce, 2019). Effective regulation requires a clear understanding of the existence and causes of different ratings. By linking ESG data providers' index licensing incentives with their ratings output, our findings can help guide regulators toward more effective disclosure policies.

Third, we extend the established literature on rating agencies as information intermediaries (Beatty et al., 2019; Jiang et al., 2012; Kedia et al., 2017). Most of this literature focuses on the credit rating industry. Like credit rating agencies, ESG data providers act as gatekeepers in financial markets (Christensen et al., 2021). However, there are several fundamental differences between ESG data providers and credit rating agencies. First, the payment models are different in the two industries. In the credit rating industry, the issuer-pay model is standard and has been shown to induce bias in ratings as raters cater to their clients (Jiang et al., 2012; Beatty et al., 2019). In contrast, most ESG data providers are paid by investors rather than firms. This can create greater credibility incentives for ESG data providers relative to credit rating agencies. As a result, the canonical catering bias that characterizes credit ratings need not apply to ESG ratings (Berg et al., 2022). However, Bonsall et al. (2023) notes the possibility for alternative conflicts of interest to emerge even in user-pay rating arrangements. Our paper suggests that the "investor-pays" model may be insufficient to address the conflict of interest concerns. Second, ESG rating agencies are less regulated and less transparent than credit agencies. Therefore, ESG rating agencies may have greater discretion in determining ESG ratings. Third, and the focus of our study, some ESG data providers also sell ESG indexes and index licensing incentives may lead to inflated ESG ratings for firms with better stock return performance. Our paper highlights these unique incentives of ESG data providers and investigates their consequences on reported ratings.

2 Institutional setting

ESG ratings are being increasingly used to construct ESG indexes. MSCI is perhaps the most prominent vendor that both produces ESG ratings and sells ESG indexes. It is also one of the largest ESG rating providers in the marketplace. The direct sale of ESG ratings and other climate data to investors comprises approximately eight percent of MSCI's operating revenue.

Formed when Morgan Stanley licensed the rights to indices published by Capital International, MSCI has grown to become a leading provider of index products across a variety of asset classes. MSCI's ESG indexes constitute the backbone of billions of dollars of assets managed by ESG ETFs and mutual funds. Growth in the ETF market, in general, reportedly helped ESG index providers earn \$5 billion in fee revenue in 2021 (Swink, 2022). Revenue attributed to ESG indexes reportedly grew by 123% in 2021, compared to 2020. As per the same report, around \$227 billion in equity ETF is linked to MSCI ESG and climate indexes as of 2021 end.

The market share of ESG indexes is highly concentrated, with MSCI being the index market leader with ESG ETFs. Specifically, MSCI ESG Indexes reportedly enjoyed a 73.3% market share in August 2021 (GIB, 2022). Furthermore, ETFGI (2021) claims that, in December 2020, 17 out of the largest 20 ESG ETFs tracked ESG indexes from the same index provider, MSCI, with combined assets of EUR 57 billion (Mazzacurrati, 2021). More than 60 percent of MSCI's operating revenue comes from sales of index products that are used as benchmarks by asset managers. We therefore use MSCI ratings as our measure of ratings generated under high index licensing incentives ($HighIndex$).

MSCI's index construction revenue is determined by the amount of AUM that is benchmarked to the indices. As index AUM is largely driven by the performance of the index (Berk and Green, 2004), MSCI has strong incentives to construct indices that generate positive returns. Moreover, since MSCI uses its own ESG ratings to select the constituent companies in its ESG indices, MSCI's reliance on index licensing revenues could influence the determination of ESG ratings. Specifically, MSCI earns revenue when the funds licensing its ESG indices increase their AUM. By issuing higher MSCI ESG ratings for firms with better stock return performance, MSCI can incorporate these firms into their ESG indices and improve their index performance with the goal of increasing the index AUM and related revenue. If firms' stock performance is not driven by ESG activities, this incentive could influence MSCI's ESG ratings.

Furthermore, index providers, in general, have strong incentives to minimize tracking errors and enhance the marketability of their ESG index, which is closely tied to the ESG performance of their constituents. For instance, MSCI ESG indices are "designed to represent the performance of companies that have high ESG ratings relative to their sector peers" (MSCI, 2020). MSCI assures investors that the ESG index focuses on firms that exhibit positive ESG attributes and strong ESG performance. Therefore, the incentive to inflate firms' ESG ratings after firms are included in the ESG index is high.

In contrast, we use Refinitiv ratings as our measure of ESG ratings generated under low index licensing incentives ($LowIndex$). Refinitiv is another leading financial data provider and, similar to MSCI, is one of the largest ESG data providers of numeric ESG ratings. Refinitiv's primary source of revenue is subscriptions to data platforms such as Datastream, Eikon, IBES, and Lipper. Through these platforms, Refinitiv sells numerous financial data products, including ESG ratings. Refinitiv's ratings are especially attractive for our setting because we could not identify any significant ESG Index that relies on Refinitiv's ratings. Hence, Refinitiv's ratings are relatively untainted by incentives that MSCI may be subject to because of its index business.

Disagreement or lack of correlation among ESG data providers poses several challenges to an ETF investor. Disagreement between data providers, on whose ratings the specific ESG index is based, will lead to investment portfolios that do not overlap by much. More relevant to our study, the index market leader, MSCI, produces the ESG ratings on which its own MSCI ESG index, is based.

On top of that, ETFs are primarily purchased by retail and unsophisticated investors. Professional investors typically use ESG ratings alongside their own internal assessments, and incorporate them into their valuation models, use them as indicators, or benchmark them against sector-wide ESG performance investors (Wong and Petroy, 2020). Many professional investors rely on in-house ESG expertise, but this is usually unavailable to average ESG ETF buyers. Therefore, an empirical investigation into the determinants of ESG rating disagreements used to construct their ESG index and its implications for investors is warranted.

3 Related literature and hypothesis development

In 2020, more than 90 percent of S&P 500 companies published an ESG report (Governance & Accountability Institute, 2022). However, there is no universal reporting standard for ESG disclosures, which makes it difficult for investors to reliably access and interpret ESG information directly from corporate ESG reports. In response to this challenge, many investors use ESG ratings provided by third-party data providers to compare the ESG activities of different companies and make investment decisions (Hartzmark and Sussman, 2019).

In theory, ESG ratings provide investors with a summary measure of a company's ESG performance (MSCI, 2022; Refinitiv, 2022). Thus, ESG data providers provide a similar service as credit rating agencies. However, differences between credit rating and ESG data providers limit our ability to draw inferences about ESG data providers from the conclusions of the credit ratings literature (Li et al., 2022). First, unlike credit ratings, ESG ratings focus on the issuer or company, rather than on a specific debt offering. Second, ESG ratings are usually organized under three distinct pillars ("E" for environmental, "S" for social, and "G" for governance), but these pillars are not uniformly defined by different ESG data providers. It is more difficult to verify the accuracy of ESG ratings ex post than it is for credit ratings, as there is no obvious ESG outcome variable to validate or a specified time horizon to check. Third, unlike most credit rating agencies, ESG data providers generally eschew the issuer-pay model in favor of a user-pay model. Several studies criticize the issuer-pay model for providing insufficient independence between the rater and the issuer. Because of this lack of independence, the issuer-pay model introduces conflicts of interest that often compromise the quality of ratings (Beatty et al., 2019; Cornaggia and Cornaggia, 2013; Jiang et al., 2012).

Whereas the issuer-pay model gives credit rating agencies incentives to provide high ratings to issuing firms, the user-pay model gives ESG data providers incentives to provide accurate and timely ESG ratings (Christensen et al., 2021). In a user-pay model, data providers generate revenue by selling data products directly to the users of the data (i.e., investors). Users' demand for data is largely determined by the extent to which it proves valuable for their decision-making. Because true rating quality can only be assessed in hindsight, perceptions of quality, or credibility, is often critical to users when contemplating a data purchase (White, 2001). Providing ratings that are perceived as credible and useful can help ESG data providers retain market share and attract new customers. Outside of the ESG data setting, research reveals that perceptions of credibility and investor reliance can be powerful forces in shaping the quality of the information provided by raters (DeHaan, 2017; Xia, 2014). Thus, there are likely strong incentives for ESG rating providers to consider the quality of their ratings.

However, recent evidence of pervasive and persistent ESG rating divergence raises questions about potential competing incentives in the production of ESG ratings. There are multiple ESG data providers in the marketplace, and the ratings from these different providers are highly inconsistent with one another (Berg et al., 2022; Billio et al., 2021; Boffo et al., 2020; Chatterji et al., 2016; Christensen et al., 2021; Brandon et al., 2021). Using a sample of six data providers (Sustainalytics, S&P Global, Moody's ESG, Refinitiv, KLD, and MSCI) in 2014, Berg et al. (2022) find that the correlations between ESG ratings range from 0.38 to 0.71. In addition, Christensen et al. (2021) show that ESG ratings from MSCI and TR (Sustainalytics) for a given firm-year differ by 19.7 (12.7) out of 100 on average.

There are two leading explanations for ESG rating divergence. On the one hand, ESG ratings may diverge because raters may interpret the same fundamental ESG data differently. Consistent with this view, Christensen et al. (2021) report greater rating disagreement when there is more ESG disclosure. Overall, Berg et al. (2022) estimate that this phenomenon explains 56 percent of the ESG rating divergence. That leaves room for an alternate explanation, namely that divergence across ESG ratings could indicate a failure of ESG ratings to capture firms' fundamental ESG activities. Consistent with this skeptical view, several studies find that ESG ratings are more correlated with the amount of ESG disclosure than the content of underlying ESG fundamentals (Yang, 2021; Raghunandan and Rajgopal, 2022). Yang (2021) further reports that ESG ratings do not have incremental predictive value for firms' future ESG problems or sustainability. Similarly, Serafeim and Yoon (2022) demonstrate that the usefulness of ESG ratings decreases in the presence of significant disagreement. Moreover, Raghunandan and Rajgopal (2022) find that ESG scores are unaffected by federal violations.

Despite mounting evidence of potential ESG ratings "greenwashing", the role of conflicting incentives in generating such behavior remains unexplored. We extend prior research by observing that ESG data providers face competing incentives arising from their multiple business lines. Several of the leading ESG ratings providers, such as MSCI and S&P, not only sell ESG ratings but also use ESG ratings to construct ESG indices. Index providers generate revenue by creating and calculating market indices and licensing them to asset managers who market their funds as passively tracking the licensed index (An et al., 2023). Asset managers typically pay index providers a fee that is a percentage of the fund's assets under management (AUM). This portion of the investment marketplace continues to grow rapidly; currently, more than 60 percent of MSCI's operating revenue comes from its index licensing fees.

As index licensing revenues are determined by AUM, index licensing businesses have incentives to improve the overall performance of their indices to attract investors. Similar incentives also exist in the ESG fund (Li et al., 2021). In the context of ESG ratings, this could generate an incentive to assign higher ESG ratings to firms with higher expected stock returns. Specifically, by assigning higher ESG ratings to firms with higher expected stock returns, one could justify incorporating such firms into more ESG indices and hence improve the overall return performance of these indices. This would, in turn, lead to increased index licensing revenue for the rater. However, ESG data providers may also have incentives to supply ratings that are perceived as credible and useful, aiming to boost the demand for these ratings. Ex-ante, it is unclear to what extent index licensing incentives affect ESG ratings. Therefore, we state our main hypothesis in the null form:

H1. Index licensing incentives do not influence ESG ratings.

4 Sample and Main Results

4.1 Sample construction

We conduct our analyses at the annual frequency since this is the frequency with which most firms disclose their ESG performance and most ESG data providers conduct in-depth reviews for ESG ratings. Prior research reveals that consistent ESG ratings data are not available prior to 2012. Therefore, our sample period spans 2012 to 2019.[3] In addition to ESG ratings data, we require sample firms to have data on financial performance and stock returns from COMPUSTAT and CRSP. We obtain data on institutional ownership from Thomson Reuters Institutional Holdings and analyst following data from the Institutional Brokers' Estimate System (I/B/E/S). We also access data on the quantity of ESG disclosures from Bloomberg. To ensure our inferences are not affected by firms with low stock prices, we exclude firm-year observations with a stock price of less than $1 (Israeli et al., 2021). The resulting sample comprises 7,214 firm-year observations from 1,691 unique US firms.

To ensure comparability across $HighIndex$ and $LowIndex$ ratings, we follow Christensen et al. (2021) and multiply MSCI ESG ratings (which range from 0 to 10) by 10 to be consistent with Refinitiv ESG ratings (which range from 0 to 100). As MSCI issues fractional ESG ratings, this scaling does not impact the relative granularity of the two ratings series. Figure 1 shows the average $HighIndex$ ratings and $LowIndex$ ratings by industry (based on the Fama-French 12-industry classification). The figure reveals a general consistency across the

[3]We end our sample in 2019 because the London Stock Exchange Group (LSEG) confirmed to acquire Refinitiv in 2019. LSEG is also the parent company of FTSE, another ESG data provider, which may impact Refinitiv ratings.

two rating providers. For example, both $HighIndex$ and $LowIndex$ raters provide high ESG ratings for the utility industry. At the same time, Figure 1 reveals non-trivial differences between $HighIndex$ and $LowIndex$ ESG ratings for each of the 12 Fama-French industries. While $HighIndex$ raters provide the energy industry with the lowest average ESG rating, $LowIndex$ raters provide the telephone industry with the lowest average ESG rating. This confirms the disagreement between ESG raters documented in prior research (Christensen et al., 2021).

Figure 2 focuses more closely on the differences between $HighIndex$ and $LowIndex$ ESG ratings by industry (in Panel A) and by year (in Panel B). Panel A reveals that average $HighIndex$ ratings exceed average $LowIndex$ ratings for seven of the Fama-French 12 industries: Business Equipment, Durables, Finance, Health, Manufacturing, Other, and Telephone industries. For the remaining five industries, average $LowIndex$ ratings exceed average $HighIndex$ ratings. Panel B of Figure 2 reveals an increasing time trend in ESG rating differences across $HighIndex$ and $LowIndex$ raters. This suggests that over time, $HighIndex$ raters provide increasingly higher ESG ratings for the same firms relative to $LowIndex$ raters. In addition, this is consistent with Christensen et al. (2021)'s findings that ESG disagreement has been increasing over time.

Panel A of Table 1 displays descriptive statistics for the main variables in our analyses. It reports that the mean (rescaled) $HighIndex$ rating is 45.507 out of 100 and the mean $LowIndex$ rating is 43.019 out of 100. The mean difference between $HighIndex$ and $LowIndex$ ratings is 2.517. Together these statistics suggest that, on average, $HighIndex$ raters issue higher ESG ratings than $LowIndex$ raters for the same firm. For firms in our sample, the average annual equity return is -0.024 and the average ESG disclosure score is 38.795. The average firm in our sample is profitable (mean $ROA = 0.113$) and followed by approximately 12 analysts (mean $Analyst = 12.662$).

Panel B of Table 1 reports Pearson correlation coefficients between the main variables in our sample. The correlation between $HighIndex$ and $LowIndex$ ratings is low (0.268), consistent with Berg et al. (2022) that there is substantial disagreement between ESG data providers. The equity returns ($Return$) are positively correlated with the difference between $HighIndex$ and

$LowIndex$ ratings (ESG_Diff), which is consistent with stock return performance having a greater influence on $HighIndex$ ratings relative to $LowIndex$ ratings. The level of ESG disclosure ($ESG\_Disclosure$) and firm size ($Size$) are negatively correlated with the difference between $HighIndex$ and $LowIndex$ ratings (ESG_Diff), suggesting that $HighIndex$ raters provide higher ESG ratings than $LowIndex$ raters for smaller firms with less ESG disclosure.

For our index sample, as we do not have access to historical index constituencies, we use linked index fund constituents as a proxy. In keeping with our definition of MSCI as a $HighIndex$ rater, we focus our study on the construction of MSCI's USA ESG indices. First, we use MSCI's corporate website to obtain a list of their licensed ESG indices as well as the funds that license each of those indices. For example, MSCI USA Extended ESG Focus is an index that is licensed by the iShares ESG Aware MSCI USA ETF. For each of these funds, we obtain index holdings data from Refinitiv Eikon. Dyer and Guest (2022) note that index funds take varied approaches to tracking their benchmark indices; some opt for full replication while others use a more selective sampling approach. To ensure the holdings data from the linked funds accurately reflect the underlying index holdings, we read the prospectus and exclude funds that use sampling methods. Panel C of Table 1 provides information on the index sample included in the study. Many of these ESG indices hold significant economic importance, with the majority having total net assets (TNA) exceeding $1 billion and the largest one with TNA exceeding $20 billion.

4.2 ESG ratings and index incentives

We examine whether index construction incentives influence ESG ratings by estimating the following equation:

$$ESG\_Diff_{i,t+1} = \beta_1 StockPerformance_{i,t} + \gamma Controls_{i,t} + \phi_k + \tau_t + \epsilon_{i,t} \quad (1)$$

In Equation 1, the dependent variable (ESG_Diff_{t+1}) is the difference between $HighIndex$ and $LowIndex$ ESG ratings for firm i reported in January of year $t+1$.

If index construction incentives influence the development of ESG ratings, we expect

ESG_Diff_{t+1} to be positively associated with measures of stock return performance during year t: equity return ($Return_t$) and equity book-to-market ratio (BTM_t). We measure $Return_t$ as the change in stock price during calendar year t. We measure BTM_t as the ratio of equity book values to equity market values in year t. To ensure that both equity book and market values are observable to the rater prior to their development of year $t+1$ ESG ratings, we measure year t equity book values and market values as of the most recent fiscal quarter ending on or before September 30 of year t.

We expect index licensing incentives to induce a positive relation between $Return_t$ and ESG_Diff_{t+1}, since $Return_t$ measures stock return performance, which would affect the performance of the licensed index products. BTM_t is a measure of the market value of the firm's stock per dollar of the firm's accounting book value; therefore, it captures the market's sentiment toward the stock. We therefore expect index licensing incentives to induce a negative relation between BTM_t and ESG_Diff_{t+1}. We also include a measure of firm profitability in terms of return on assets (ROA_t), which is defined as operating income before depreciation during year t (summed over the four most recent quarter ending on or before September 30 of year t) divided by average total assets during year t. ROA_t may be related to both stock returns and ESG ratings if greater profitability leads to higher stock returns and greater investment in ESG.

As controls, we include several variables shown to be associated with the quality of a firm's information environment and thus the degree of ESG ratings divergence: firm size measured as the log of the firm's market capitalization as of the most recent fiscal quarter ending on or before September 30 of year t ($Size_t$), the number of analysts following the firm during year t ($Analyst_t$) which is defined as the number of estimates (average over the twelve months from September of year $t-1$ to September of year t), the percentage of shares outstanding held by institutional investors as of September 30 of year t (IOR_t), and the quantity of ESG disclosure made by the firm ($ESG\_Disclosure_t$) during year t. To further account for unobservable variation in ESG ratings, Equation 1 includes industry (ϕ_k) and year (τ_t) fixed effects. We define industries using the Fama-French 12-industry classifications. We base our inferences on t-statistics computed using robust standard errors clustered by industry.

Table 2 reports regression summary statistics from estimating Equation 1. The coefficients on *Return* are significantly positive in all columns, suggesting firms with higher (lower) annual equity returns have higher (lower) *HighIndex* ratings relative to *LowIndex* ratings. Moreover, the economic magnitude is nontrivial. A one-standard-deviation increase in *Return* leads to an increase in ESG_Diff of 0.768, which is 31% of the mean ESG_Diff.[4] The coefficients on equity book-to-market ratio (BTM) are significantly negative in all columns, indicating firms with higher (lower) BTM have lower (higher) *HighIndex* ratings than *LowIndex* ratings. This is consistent with *HighIndex* ratings being more closely aligned with investor sentiment than are *LowIndex* ratings. A one-standard-deviation decrease in BTM leads to an increase in ESG_Diff by 1.85, which is 73% of mean ESG_Diff. The coefficient of return on assets (ROA) is insignificant. This finding suggests that the difference in *HighIndex* versus *LowIndex* ESG ratings is driven by stock return performance but not fundamental firm performance. It reinforces a short-term perspective on the role of index licensing incentives in the production of ESG ratings. While returns themselves inform ratings when index licensing incentives are high, leading indicators of future return (like profitability) do not.

Table 2 reveals significantly negative coefficients on the quantity of ESG disclosure (*ESG_Disclosure*) and firm size (*Size*). Given that both *ESG_Disclosure* and *Size* are likely positively related to the quality of a firm's information environment, their negative associations with ESG_Diff suggest that *HighIndex* ratings are higher than *LowIndex* ratings for firms with a worse information environment. This finding is consistent with the cost of inflating ESG ratings for firms with a worse information environment being lower because the worse environment makes it harder to verify ESG ratings. Collectively, the results are consistent with index licensing incentives influencing ESG ratings and with *HighIndex* raters providing higher ESG ratings for firms with better stock return performance, especially for smaller firms and firms with less ESG disclosure.

[4] We construct ESG_Diff using *LowIndex* ratings from Refinitiv that are available at the time of our data collection (Fall 2022). Berg et al. (2020) report that Refinitiv applies retroactive changes to their reported ESG ratings. They do not examine whether the initial or revised data are better in capturing the ESG quality of firms. However, they find that the revised ratings exhibit stronger associations with equity returns. Since these revisions should increase the likelihood of our observing a positive association between *LowIndex* ratings and *Return*, they would also decrease the likelihood of an association between ESG_Diff and *Return*.

4.3 ESG ratings and index composition

Our descriptive analyses suggest that index licensing incentives influence ESG ratings by demonstrating that the divergence between $HighIndex$ and $LowIndex$ ESG ratings is related to stock return performance. However, the findings do not fully answer the question of whether this difference relates to index licensing incentives. To better understand whether the potential link between ESG ratings and stock return performance arises from raters' index licensing incentives, we also examine whether stock return performance influences the composition of ESG index products.

To do so, we construct a panel of ESG index constituents over time as shown in Panel C of Table 1. We merge this sample of index constituents to the set of firms with available $HighIndex$ ratings. The resulting sample comprises 7,866 firm-year observations for 1,669 unique US firms between 2012 and 2019.

Using this sample, we construct four measures of a firm's inclusion in ESG indices. First, we define the indicator variable $ESGIndexInclude_{i,t}$ equal to one when firm i is included in at least one index constructed by the $HighIndex$ rater in year t. Second, we define $ESGIndexNum_{i,t}$ as a count of the number of ESG indices in which MSCI includes firm i during year t. We then define $\Delta ESGIndexNum_{i,t}$ as the change in $ESGIndexNum$ from year t to year $t+1$. Third, we collapse $\Delta ESGIndexNum_{i,t}$ into a categorical variable, $\Delta ESGIndexNum\_UpOrDown_{i,t}$, that takes the value of -1, 0, or 1 if $\Delta ESGIndexNum_{i,t}$ is negative, zero, or positive, respectively. Fourth, we define $\Delta ESGIndexWeight_{i,t}$ as the change in cumulative weight assigned to firm i across all $HighIndex$ ESG indices during year t. Panel A of Table 3 displays descriptive statistics for these variables as well the controls in our index constituents sample. It reveals that the mean values of $ESGIndexNum$ and $ESGIndexWeight$ are positive, suggesting that a minority of firms with $HighIndex$ ratings are included in several related ESG indices.

Using these data, we study whether stock return performance is associated with $HighIndex$

rater' indices construction by first estimating the following equation:

$$ESGIndexInclude_{i,t+1} = \beta_1 HighIndex\_StockPerf_{i,t+1} + \beta_2 HighIndex\_NonStockPerf_{i,t+1} + \gamma Controls_{i,t} + \phi_k + \tau_t + \epsilon_{i,t} \quad (2)$$

In Equation 4, the dependent variable is $ESGIndexInclude_{i,t}$. The explanatory variables, $HighIndex\_StockPerf_{i,t}$ and $HighIndex\_NonStockPerf_{i,t}$, are the predicted and residual values of the change in firm i's $HighIndex$ ESG rating in year $t+1$, generated by estimating Equation 3 below:

$$HighIndex_{i,t+1} = \beta_1 StockPerformance_{i,t} + \epsilon_{i,t} \quad (3)$$

If stock performance influences ESG index inclusion, we expect $ESGIndexInclude_{i,t+1}$ to be associated with $HighIndex\_StockPerf_{i,t+1}$ in Equation 2. If $ESGIndexInclude_{i,t+1}$ is significantly associated with $HighIndex\_NonStockPerf_{i,t+1}$ as well, this would suggest that ESG factors beyond stock return performance affect a firm's likelihood of ESG index inclusion.

Panel B of Table 3 shows the results of estimating Equation 2. The coefficients of both $HighIndex\_StockPerf$ and $HighIndex\_NonStockPerf$ are significantly positive. This result indicates that $HighIndex$ raters do not solely focus on stock returns when constructing ESG indices. While higher stock performance increases the likelihood of a firm being included in an ESG index, so does higher non-stock ESG performance.

Next, we study how $HighIndex$ raters change their indices over time by estimating the following equation:

$$\Delta Y_{i,t+1} = \beta_1 \Delta HighIndex\_StockPerf_{i,t+1} + \beta_2 \Delta HighIndex\_NonStockPerf_{i,t+1} + \gamma Controls_{i,t} + \phi_k + \tau_t + \epsilon_{i,t} \quad (4)$$

In Equation 4, the dependent variable is either $\Delta ESGIndexNum\_UpOrDown_{i,t+1}$ or $\Delta ESGIndexNum_{i,t+1}$. The explanatory variables, $\Delta HighIndex\_StockPerf_{i,t+1}$ and

$\Delta HighIndex\_NonStockPerf_{i,t+1}$, are the predicted and residual values of the change in firm i's $HighIndex$ rating from year t to year $t+1$, generated by estimating Equation 3 using $\Delta HighIndex_{i,t+1}$ as the explanatory variable.

If index licensing incentives lead raters to consider stock returns in the construction of their indices, we expect $\Delta ESGIndexNum\_UpOrDown_{i,t+1}$ and $\Delta ESGIndexNum_{i,t+1}$ to be associated with $\Delta HighIndex\_StockPerf_{i,t+1}$. If these outcomes are also significantly associated with $\Delta HighIndex\_NonStockPerf_{i,t+1}$, it would suggest that index providers consider ESG factors beyond stock return performance in determining index inclusion. Because both $\Delta ESGIndexNum\_UpOrDown_{i,t+1}$ or $\Delta ESGIndexNum_{i,t+1}$ are discrete variables with limited support, we estimate Equation 4 using an ordered logistic regression.

Panel C of Table 3 reports results of estimating Equation 4 using $\Delta ESGIndexNum\_UpOrDown$ and $\Delta ESGIndexNum$ as the dependent variables. In column (1) where the dependent variable is $\Delta ESGIndexNum\_UpOrDown$, the coefficient of $\Delta HighIndex\_StockPerf$ is significantly positive. A one-standard-deviation increase in $\Delta HighIndex\_StockPerf$ raises the probability of being included in more MSCI ESG indices by 6.6%. This result indicates that higher *Return* is associated with a greater probability of increased index inclusion. Also, the coefficient of both $\Delta HighIndex\_NonStockPerf$ is significantly positive. Taken together, these findings suggest that the portion of ESG ratings that change index inclusion decisions from year-to-year relate to both stock return performance and variations in ESG ratings. Similarly, in column (2) where the dependent variable is $\Delta ESGIndexNum$, the coefficients of both $\Delta HighIndex\_StockPerf$ and $\Delta HighIndex\_NonStockPerf$ are significantly positive, suggesting that both the stock performance portion and non-stock performance portion of ESG ratings relate to the change in the number of MSCI ESG indices to which a firm belongs.

It is not only the inclusion but also the weighting of a firm in an index that affects index performance. We further consider how the cumulative weight of a firm across ESG indices changes with stock return performance-based ESG ratings by estimating the following equation:

$$\Delta ESGIndexWeight_{i,t+1} = \beta_1 \Delta HighIndex\_StockPerf_{i,t+1} + \beta_2 \Delta HighIndex\_NonStockPerf_{i,t+1} + \gamma Controls_{i,t} + \phi_k + \tau_t + \epsilon_{i,t} \quad (5)$$

In Equation 5, the dependent variable ($\Delta ESGIndexWeight_{i,t}$) is the change in cumulative weight assigned to firm i across all MSCI ESG indices from year t to year $t+1$. If ESG indices are tilted towards firms with higher stock performance, then we expect $\Delta ESGIndexWeight_{i,t+1}$ to be associated with $\Delta HighIndex\_StockPerf_{i,t+1}$. If other ESG factors influence ESG weights, then $\Delta HighIndex\_NonStockPerf_{i,t+1}$ will also be positive and significant. These variables, as well as the *Controls* vector, remain as previously defined.

Panel D of Table 3 reports the results of estimating Equation (5). The coefficient of $\Delta HighIndex\_StockPerf$ is significantly positive and suggests that a one-standard-deviation increase in $\Delta HighIndex\_StockPerf$ leads to a 3.7% increase in $\Delta ESGIndexWeight$ relative to the mean. In contrast, the coefficient of $\Delta HighIndex\_NonStockPerf$ is insignificant. This result suggests that the portion of ESG ratings that impact cumulative weighting changes only relates to a firms' stock return performance, and all other variation in ESG ratings is not significantly associated with changes in index composition. Specifically, higher *Return* is associated with higher cumulative weights. Collectively, Table 3 provides evidence that index licensing incentives influence ESG ratings and index composition.

4.4 Event study of index additions and deletions

To strengthen our identification of the link between the ESG indices and ESG ratings, we use an event-study design focused on the addition or deletion of firms from MSCI ESG indices.[5] We investigate whether firms added to (deleted from) these ESG indexes have higher (lower) *HighIndex* ESG ratings relative to the same firm's *LowIndex* ESG ratings (i.e., a higher

[5] The research design in the event study requires us to have a post-period. Therefore, we exclude the ESG indices that were launched in 2019, including MSCI USA ESG Enhanced Focus Index, MSCI USA SRI Select Reduced Fossil Fuel Index, MSCI USA Extended ESG Leaders TR USD, MSCI USA ESG Universal Select Business Screens Index, MSCI USA SRI S-Series PAB 5% Capped Index. This exclusion results in a total of 11 indices for the analyses related to the event study.

[lower] ESG_Diff value). We compare $HighIndex$ ESG ratings for firms added to (deleted from) $HighIndex$ ESG index funds in the pre-addition (deletion) period and the post-addition (deletion) period with the rating by the $LowIndex$ rater.

Specifically, we use the addition or deletion of a firm from any of the aforementioned ESG indices as a focal treatment event, and we study how ESG_Diff changes around these events. To further control for the fundamental differences in rating methodology used by different rating agencies, we define the control group as exclusively those "stable" firms that were not added or deleted from the index during the sample period. We compare changes in ESG_Diff between treatment and control groups to infer the impact of index inclusion on $HighIndex$ ESG rating decisions. In Table 4 Panel A, we provide summary statistics of our two treatment group companies (i.e., companies that are added to or deleted from the MSCI ESG index) relative to stable firms for the Refinitiv samples.

Our identification assumption is that any change in actual ESG performance should have a similar impact on $HighIndex$ and $LowIndex$ ESG raters. Also, any differences in rating methodology across different agencies will be similar for treatment firms and stable firms. As there is staggered variation in the timing of such changes, we follow a stacked difference-in-difference regression approach (Baker et al., 2022; Cengiz et al., 2019; Barrios, 2021). For each index-level treatment (i.e., addition or deletion) event, we create a cohort including all treated firms as well as control firms that are stable for that particular index. We generate 38 (28) addition (deletion) cohorts through this process. We then stack all the cohorts to create a sample of 14,110 (6,803) firm-index-year observations for the addition (deletion) events from 2012 to 2019. Using this sample, we estimate the following equation:

$$ESG\_Diff_{i,t} = \beta_1 Post_t \times Treat\_firm_{i,t} + \gamma Controls_{i,t} + Firm \times CohortFE + Year \times CohortFE + \epsilon_{i,t} \quad (6)$$

In Equation (6), $Treat\_firm$ is an indicator variable that takes the value of one if the observation is subject to a treatment event (i.e., addition or deletion); otherwise, that indicator is set to zero. $Firm \times Cohort$ fixed effects control for time-invariant firm-level characteristics within each cohort. $Year \times Cohort$ fixed effects control for time-varying shocks within each cohort. The $Controls$ vector includes several firm-specific time-variant factors likely to impact ESG ratings, such as leverage, tangibility of assets, cash holdings, institutional ownership, analyst following, and dividends paid (Li et al., 2022). We include a continuous measure of research and development (R&D) expenditure as well as an indicator variable for the existence of non-zero R&D expenditures because discretionary spending such as R&D arguably likely has a nonlinear impact on firms' ESG budgets. Furthermore, following prior literature (Christensen et al., 2021; Li et al., 2022), we control for sales growth, return on assets, equity market-to-book ratio, and Tobin's Q as firms with high growth opportunities may care more about sustainability. ESG ratings from different rating agencies may have different statistical distributions on account of idiosyncratic assignment of the proportion of the rated companies as "leaders" or "laggards." Therefore, we also test our results using standardized ESG ratings as an alternative dependent variable.

Table 4 Panel B reports the results of estimating equation (6). In all the specifications, the coefficients on $Post \times Treat\_firm$ are positive (negative) and significant at the 1% level for firms that are added to (deleted from) the index. Firms that are added to one of the $HighIndex$ rater's ESG indices experience an increase of 2.923 points in their $HighIndex$ ESG rating compared to their $LowIndex$ ESG rating, relative to the control sample of stable firms. The increase is equivalent to 41.8% of the average difference in ESG ratings between $HighIndex$ and $LowIndex$ raters for the control group. The results strengthen our inference that, relative to $LowIndex$ raters, $HighIndex$ raters increase ESG ratings of firms that are added to their ESG indices.

5 Additional Analyses

5.1 Distinguishing stock return performance from ESG performance

Our main analyses reveal a strong positive association between stock returns and the difference in ESG ratings from a rater with strong index licensing incentives and a rater with low licensing incentives. We interpret this as evidence that index licensing incentives lead raters to assign higher ESG ratings to firms with stronger expected future stock return performance. Alternatively, as ESG remains a nebulous concept that comprises a range of different activities, it is possible that different ESG raters issue different ratings because they focus on different components of ESG activity. Stated differently, it is possible that certain features of ESG performance lead to stronger stock performance, and certain ESG raters focus more heavily on these features. To explore this possibility, we take advantage of the fact that most ESG raters provide not only an overall ESG rating but also underlying component scores. For example, MSCI issues four layers of ESG ratings: indicator scores (i.e., carbon emission, labor management, etc.), category scores (i.e., climate change, human capital, etc.), pillar scores (for each of the Environment, Social, and Governance pillars), and an overall ESG rating. Raters provide little insight into how these component scores relate to the overall ESG rating they issue, but Berg et al. (2022) conjecture that raters aggregate category scores to arrive at their overall rating.

This disaggregation allows us to examine the possibility that our main findings arise because ESG raters with strong index licensing incentives focus on different components of corporate ESG activity. We study the extent to which firms' overall ESG ratings relate to their ESG component scores and whether stock returns are associated with the overall ESG ratings *after* controlling for firms' individual ESG component scores. We interpret the portion of ESG ratings unexplained by component scores as the most subjective portion of the ratings. If stock returns are related to this highly subjective portion, it will suggest that ESG raters consider stock returns on the margin when forming their ESG ratings. To test this idea, we estimate the following equation:

$$ESGRating_{i,t+1} = \beta_1 Return_{i,t} + \gamma Category\_Score_{i,t+1} + \phi_i + \tau_t + \epsilon_{i,t} \quad (7)$$

Tables 5 and 6 report regression results from estimating Equation (7) separately using *HighIndex* and *LowIndex* ratings as the dependent variable. The results reveal that variation in *HighIndex* ESG component scores explains only 74.1 percent of the variation in *HighIndex* overall ESG ratings. In contrast, variation in *LowIndex* ESG component scores explains 97.1 percent of the variation in *LowIndex* overall ESG ratings. This evidence suggests there is greater transparency in *LowIndex* ESG rating construction.

Turning our attention to the specific role of stock returns, we find that there is a significant positive association between *HighIndex* ESG ratings and stock returns after controlling for variation in all reported ESG component scores as determined by *HighIndex*. A one-standard deviation increase in *Return* leads to a 0.143 increase in *HighIndex* ESG ratings, which is 0.31% of the mean *HighIndex* ESG rating. In contrast, stock returns are not associated with *LowIndex* ESG ratings after controlling for variation in analogous component scores as determined by *LowIndex*. This finding is consistent with the notion that index licensing incentives lead raters to provide higher (lower) ESG ratings for firms with higher (lower) stock returns.

5.2 Does index licensing help raters generate more accurate or timely ESG ratings?

We interpret our results as evidence that *HighIndex* raters may have incentives to influence their ESG ratings because of their index licensing business. However, we acknowledge that alternative explanations may also be plausible: (1) index composition changes themselves might prompt firms to improve their ESG performance, or (2) *HighIndex* raters may have access to private information because of their index business. In this section, we conduct further analyses to investigate these alternative possibilities.

First, we examine firms' subsequent ESG outcomes (e.g., violations, gender diversity, racial diversity, climate change exposure) following Heath et al. (2021). We use the same specification as our main analysis to examine the changes in ESG outcomes after $HighIndex$ ESG index additions/deletions.

$$ESG\_Outcome_{i,t+1} = \beta_1 Post_t \times Treat\_firm_{i,t} + \gamma Controls_{i,t} + Firm \times CohortFE + Year \times CohortFE + \epsilon_{i,t} \quad (8)$$

The dependent variable $ESG\_Outcome$ is one of four measures: the number of penalties associated with regulatory violations, gender diversity, racial diversity, or climate change exposure, which are measured on an annual basis.

In Table 7 Columns (1) and (2), we do not find any significant association between $HighIndex$ raters' decisions to add or delete a firm from an ESG index and future ESG violations. In Columns (3) and (4), we find positive coefficients on $Treat\_Firm \times Post$ for gender diversity for firms both added to and deleted from ESG indices, suggesting a general trend instead of index inclusion or exclusion events being the main driving factor. For racial diversity, the coefficient on $Treat\_Firm \times Post$ is statistically insignificant for both the firms added or deleted from ESG indices in Columns (5) and (6). Finally, in Columns (7) and (8), our analysis suggests there is no statistically significant correlation between $HighIndex$ raters' inclusion or exclusion of a company from their ESG indices and changes in the firm's climate change exposure, as measured by management's references to climate in conference calls. In sum, we do not find evidence to support the claim that inclusion or exclusion in ESG indices leads to actual changes in any of the wide range of ESG outcomes that we observe. It may still be possible that ESG index inclusion decisions (and subsequent ratings changes) pertain to other unobservable dimensions of ESG that are not reflected in any of the measures that we study. However, this interpretation, when taken in conjunction with our findings, would still suggest that the features of ESG that determine index inclusion likely fall outside of the "mainstream" definitions of ESG.

Next, we plot the changes in ESG ratings and confidence intervals around these changes prior to and following additions to or deletions from ESG indices, in Figure 3. We find that

LowIndex ESG ratings remain largely stable over the sample period. In contrast, *HighIndex* ESG ratings exhibit an increase following index additions and a decrease following index deletions. If *HighIndex* raters have access to private information about the ESG performance of added or removed firms that is not immediately available to *LowIndex* raters, or if index composition events result in actual changes in firms' ESG outcomes related to ESG rating changes, we would expect other rating agencies to follow *HighIndex* rating changes (i.e., to "catch up" with MSCI's rating changes). However, we do not find evidence of such "catching up" in Figure 3.

In summary, along with the results in Table 7 documenting no significant changes in ESG outcomes, results suggesting the absence of "catching up" in *LowIndex* ratings casts doubt on the validity of the alternative explanations posited.

6 Conclusion

It is not clear how ESG data providers determine ESG ratings, and there is substantial disagreement in ratings across ESG data providers (Berg et al., 2022; Christensen et al., 2021). This raises concerns about the credibility of ESG ratings and underscores the need to understand the incentives that shape the production of ESG ratings. We examine whether raters with strong index licensing incentives issue higher ESG ratings for firms with better stock return performance and those added to their ESG indexes, compared to raters with weaker licensing incentives. We study MSCI as an example of an ESG rater with high index licensing incentives (*HighIndex*) and Refinitiv as an example of an ESG rater with low index licensing incentives (*LowIndex*). While most of Refinitiv's revenue is from selling data, more than 60 percent of MSCI's operating revenue is from index licensing fees. Using these raters, we also investigate the degree to which index licensing incentives influence ESG ratings by benchmarking *HighIndex* ESG ratings to *LowIndex* ESG ratings.

Our results offer several new insights. First, we report that firms with higher (lower) stock returns receive higher (lower) ratings from a rater with high index incentives relative to ratings from a rater with low index incentives. As our inferences are based on comparisons across ratings issued for the same firm, we effectively hold "fundamental" ESG performance constant in our analyses. Second, we find that ESG ratings from a rater with high index incentives are systematically higher (lower) than those of a rater with low index incentives for firms added (dropped) from the ESG indexes, even after controlling for rating methodology differences. Notably, these ESG ratings upgrades and downgrades, relative to peers, do not appear to be informative about "fundamental" ESG performance. Third, we show that ESG index inclusion decisions are associated with stock returns. Collectively, our findings suggest that ESG data providers' index licensing incentives influence their ESG ratings.

Our study makes several contributions to the literature. First, we extend the nascent literature exploring ESG rating divergence. Prior research attributes such divergence to the number of ESG disclosures (Christensen et al., 2021) and the different priorities of ESG data providers (Berg et al., 2022). In contrast, we show the unique index licensing incentives that shape the production of ESG ratings information. Second, we extend the existing rating agency literature that mainly focuses on credit rating agencies. Though ESG data providers also act as gatekeepers in financial markets, they are fundamentally different from credit rating agencies. Specifically, ESG data providers are paid by data users rather than firms and hence have different incentives from credit rating agencies. Our paper suggests that the "investor-pays" model may be insufficient to address the concerns regarding conflicts of interest.

ESG ratings have attracted significant attention given the continued growth of ESG investing. Investors, regulators, and the media have raised concerns about construct validity, accuracy, and divergent ratings provided by rating agencies for the same firm. We hope our findings kindle debate on what, if anything, needs to be done when an ESG rating agency also markets an ESG index based on such ratings. This concern is important as such an index is widely used by passive investors without the ability or the resources to do their own due diligence on whether the companies included in the index operate in a sustainable manner.

Appendix

A Variable Definitions

Panel A: ESG Rating Variables and ESG Index Variables	
Variables Name	Description
$HighIndex_{i,t}$	MSCI weighted average ESG score for firm i as of the first day of year t
$LowIndex_{i,t}$	Refinitiv ESG rating for firm i as of the first day of year t
$ESG\_Diff_{i,t}$	The rating difference between $HighIndex_{i,t}$ and $LowIndex_{i,t}$
$ESG\_Diff\_Std_{i,t}$	The difference between standardized $HighIndex_{i,t}$ and $LowIndex_{i,t}$ ratings
$ESGIndexInclude_{i,t}$	An indicator variable that equals 1 when firm i is included in MSCI ESG indices and equals 0 when the firm i is not included in any MSCI ESG indices during year t
$ESGIndexNum_{i,t}$	The number of MSCI ESG indices of which firm i is a constituent during year t
$ESGIndexWeight_{i,t}$	The cumulative weight assigned to firm i across all MSCI ESG indices during year t
$\Delta ESGIndexNum_{i,t}$	The change in number of MSCI ESG indices of which firm i is a constituent during year t
$\Delta ESGIndexNum\_UpOrDown_{i,t}$	A categorical variable that equals 1 when the $\Delta ESGIndexNum > 0$, equals 0 when $\Delta ESGIndexNum = 0$ and equals -1 when $\Delta ESGIndexNum < 0$.
$\Delta ESGIndexWeight_{i,t}$	The change in cumulative weight assigned to firm i across all MSCI ESG indices during year t

$\Delta HighIndex\_StockPerf_{i,t}$	The predicted change in firm i's $HighIndex$ ESG rating in year t from estimating the below equation: $$\Delta HighIndex_{i,t} = \beta_1 StockPerformance_{i,t-1} + \epsilon_{i,t}$$
$\Delta HighIndex\_NonStockPerf_{i,t}$	The residual change in firm i's $HighIndex$ ESG rating in year t from estimating the below equation: $$\Delta HighIndex_{i,t} = \beta_1 StockPerformance_{i,t-1} + \epsilon_{i,t}$$

Panel B: Firm Level Variables and Controls

$Analyst_{i,t}$	The number of analysts following the firm for firm i during year t (average over the twelve months from September of year $t-1$ to September of year t)
$BTM_{i,t}$	Book-to-market ratio for firm i in year t
$BVE_{i,t}$	Equity book value for firm i at the end of the most recent fiscal quarter ending on or before September 30 of year t
$Cash_{i,t}$	Cash and Cash equivalents (CHE) / Total Assets (AT) for firm i in year t
$Div_{i,t}$	Cash Dividends (DV) / Total Assets (AT) for firm i in year t
$ESG\_Disclosure_{i,t}$	The Bloomberg ESG disclosure score for firm i in year t
Has_RD	Indicator equals to one if RD >0
$IOR_{i,t}$	The percentage of the shares owned by institutional investors for firm i as of September 30 of year t
$Lev_{i,t}$	Total Debt (DT)/ Total Assets (AT) for firm i in year t
$MB_{i,t}$	Market value of common equity / Book value of common equity for firm i in year t
$MVE_{i,t}$	Market value for firm i at the end of the most recent fiscal quarter ending on or before September 30 of year t

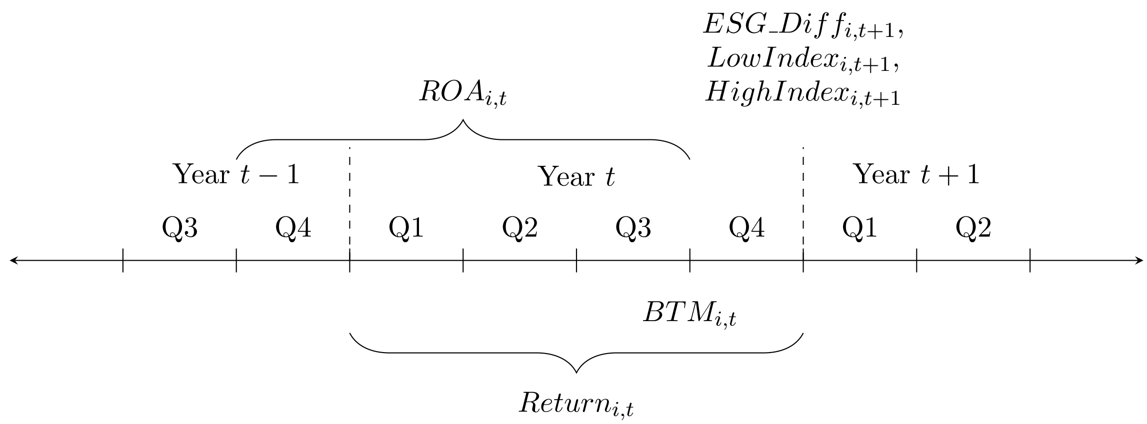
$RD_{i,t}$	Research & Development expenses (XRD) / Total Assets (AT) for firm i in year t
$Return_{i,t}$	The market-adjusted return for firm i in year t
$ROA_{i,t}$	Return on assets for firm i in year t, defined as operating income before depreciation (summed over the four most recent quarter ending on or before September 30 of year t) divided by average total assets during year t
$Sales\_growth_{i,t}$	[Current Period Sales - Past Period Sales]/Past Period Sales, where Sales is Compustat item SALE for firm i in year t
$Size_{i,t}$	The logarithm of market capitalization for firm i at the end of the most recent fiscal quarter ending on or before September 30 of year t
$Spread_{i,t}$	Average bid-ask spread based on daily close prices for firm i over year t
$Tangiblity_{i,t}$	Net Property, Plant & Equipment (PPENT) / Total Assets (AT) for firm i in year t
$TobinQ_{i,t}$	[Market value of common equity + Total assets - Book value of common equity]/ Total Assets Book value of common equity = SEQ (or CEQ+PSTK or AT - LT -MTB; in order of preference) - PSTKRV (or PSTKL or PSTK or 0 , in that order of prefernce)
Δ	Change operator
Panel C: Additional Variables for Index Addition/Deletion Tests	
$Post_t$	Indicator equal to one if month t is after a addition/deletion event
$Treat\_firm_{i,t}$	Indicator equal to one if firm i is added/deleted to an ESG index in month t, and 0 if the firm has been part of the ESG index since index inception

Panel D: Additional Variables for ESG Outcome Tests

$\#Penalty_{i,t+1}$	Number of Penalties imposed on firm i for violations in the calendar year $t+1$
$Gender\_diversity_{i,t+1}$	Ratio of women directors to total directors on the board of firm i in year $t+1$
$Racial\_diversity_{i,t+1}$	Ratio of non-Caucasian directors to total directors on the board of firm i in year $t+1$
$CCExposure_{i,t+1}$	Relative frequency with which bigrams related to climate change occur in the transcripts of earnings conference calls of firm i in year $t+1$. We measure relative frequency by dividing the number of such climate-change bigrams (numerator) by the total number of bigrams in the transcripts (denominator).

B Timeline

This figure provides a timeline of stock return performance and ESG ratings in a given firm-year. For each firm i, we measure potential determinants of ESG ratings ($Return$, BTM, ROA) such that they are observable to the rater prior to the release of year $t+1$ ESG ratings.



References

An, Yu, Matteo Benetton, and Yang Song, 2023, Index providers: Whales behind the scenes of etfs, *Journal of Financial Economics* .

Baker, Andrew C, David F Larcker, and Charles CY Wang, 2022, How much should we trust staggered difference-in-differences estimates?, *Journal of Financial Economics* 144, 370–395.

Barrios, John Manuel, 2021, Staggeringly problematic: A primer on staggered did for accounting researchers, *Available at SSRN 3794859* .

Beatty, Anne, Jacquelyn Gillette, Reining Petacchi, and Joseph Weber, 2019, Do rating agencies benefit from providing higher ratings? Evidence from the consequences of municipal bond ratings recalibration, *Journal of Accounting Research* 57, 323–354.

Berg, Florian, Kornelia Fabisik, and Zacharias Sautner, 2020, Rewriting history ii: The (un) predictable past of esg ratings, *European Corporate Governance Institute–Finance Working Paper* 708, 10–2139.

Berg, Florian, Julian F Koelbel, and Roberto Rigobon, 2022, Aggregate confusion: The divergence of ESG ratings, *Review of Finance* .

Berk, Jonathan B., and Richard C. Green, 2004, Mutual Fund Flows and Performance in Rational Markets, *Journal of Political Economy* 112, 1269–1295.

Berkovitch, Jonathan, Cassandra Estep, Doron Israeli, and Suhas A Sridharan, 2022, Does CSR engender investor trust? Evidence from investor reactions to corporate disclosures.

Billio, Monica, Michele Costola, Iva Hristova, Carmelo Latino, and Loriana Pelizzon, 2021, Inside the esg ratings:(dis) agreement and performance, *Corporate Social Responsibility and Environmental Management* 28, 1426–1445.

Boffo, Riccardo, Catriona Marshall, and Robert Patalano, 2020, Esg investing: Environmental pillar scoring and reporting, *Retrived* 14, 2021.

Bonsall, Samuel B, Jacquelyn R Gillette, Gabriel Pundrich, and Eric So, 2023, Conflicts of interest in subscriber-paid credit ratings, *Journal of Accounting and Economics* 101614.

Boyde, Emma, 2022, Msci esg indices' outperformance needs scrutiny, experts caution, https://www.ft.com/content/e437fe42-7483-47fa-83d6-983ae1bfd15d, 2022-2-9.

Brandon, Rajna Gibson, Philipp Krueger, and Peter Steffen Schmidt, 2021, Esg rating disagreement and stock returns, *Financial Analysts Journal* 77, 104–127.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, 2019, The effect of minimum wages on low-wage jobs, *The Quarterly Journal of Economics* 134, 1405–1454.

Chatterji, Aaron K, Rodolphe Durand, David I Levine, and Samuel Touboul, 2016, Do ratings of firms converge? Implications for managers, investors and strategy researchers, *Strategic Management Journal* 37, 1597–1614.

Christensen, Dane M, George Serafeim, and Anywhere Sikochi, 2021, Why is Corporate Virtue in the Eye of The Beholder? The Case of ESG Ratings, *The Accounting Review* 97, 147–175.

Cornaggia, Jess, and Kimberly J Cornaggia, 2013, Estimating the costs of issuer-paid credit ratings, *The Review of Financial Studies* 26, 2229–2269.

DeHaan, Ed, 2017, The financial crisis and corporate credit ratings, *The Accounting Review* 92, 161–189.

Dyer, Travis, and Nicholas Guest, 2022, A Tale of Two Index Funds: Full Replication vs. Representative Sampling .

ESMA, 2022, Esma publishes results of its call for evidence on esg ratings, https://www.esma.europa.eu/press-news/esma-news/esma-publishes-results-its-call-evidence-esg-ratings, 2022-06-27.

FCA, 2022, Esma publishes results of its call for evidence on esg ratings, https://www.fca.org.uk/publication/feedback/fs22-4.pdf, 2022-06.

GIB, Asset Management, 2022, Of msci and men, https://gibam.com/insights/of-msci-and-men, 2022-04.

Governance & Accountability Institute, 2022, Sustainability reporting in focus.

Hartzmark, Samuel M, and Abigail B Sussman, 2019, Do investors value sustainability? A natural experiment examining ranking and fund flows, *Journal of Finance* 74, 2789–2837.

Heath, Davidson, Daniele Macciocchi, Roni Michaely, and Matthew C Ringgenberg, 2021, Does socially responsible investing change firm behavior?, *Review of Finance-Forthcoming European Corporate Governance Institute–Finance Working Paper* .

Henze, Veronica, and Samantha Boyd, 2022, ESG Assets Rising to $50 Trillion Will Reshape $140.5 Trillion of Global AUM by 2025, Finds Bloomberg Intelligence .

IOSCO, 2021, Esma publishes results of its call for evidence on esg ratings, https://www.iosco.org/news/pdf/IOSCONEWS627.pdf, 2021-11-23.

Israeli, Doron, Ron Kasznik, and Suhas A Sridharan, 2021, Unexpected distractions and investor attention to corporate announcements, *Review of Accounting Studies* 27, 477–518.

Jiang, John Xuefeng, Mary Harris Stanford, and Yuan Xie, 2012, Does it matter who pays for bond ratings? Historical evidence, *Journal of Financial Economics* 105, 607–621.

Kedia, Simi, Shivaram Rajgopal, and Xing Alex Zhou, 2017, Large shareholders and credit ratings, *Journal of Financial Economics* 124, 632–653.

Larcker, David F, Lukasz Pomorski, Brian Tayan, and Edward M Watts, 2022, Esg ratings: A compass without direction, *Rock Center for Corporate Governance at Stanford University Working Paper* .

Li, Tao, S Lakshmi Naaraayanan, and Kunal Sachdeva, 2021, Conflicting objectives of esg funds: Evidence from proxy voting, *Available at SSRN 3760753* .

Li, Xuanbo, Yun Lou, and Liandong Zhang, 2022, Do Commercial Ties Influence ESG Ratings? Evidence from Moody's and S&P.

Mazzacurrati, J, 2021, Esg ratings: Status and key issues ahead, *ESMA report on Trends, Risks, and Vulnerabilities* 105–115.

MSCI, 2020, MSCI ESG Leaders Indexes Methodology,.

MSCI, 2022, MSCI ESG ratings methodology.

Piccolo, Alessio, and Joel Shapiro, 2022, Credit ratings and market information, *The Review of Financial Studies* 35, 4425–4473.

Pierce, Hester, 2019, Scarlet letters: Remarks before the american enterprise institute, https://www.sec.gov/news/speech/speech-peirce-061819, 2019-06-18.

Raghunandan, Aneesh, and Shiva Rajgopal, 2022, Do ESG funds make stakeholder-friendly investments?, *Review of Accounting Studies* 27, 822–863.

Refinitiv, 2022, Environmental, social and governance (ESG) scores from Refinitiv.

Serafeim, George, and Aaron Yoon, 2022, Stock price reactions to ESG news: The role of ESG ratings and disagreement, *Review of accounting studies* 1–31.

Swink, Sonya, 2022, Index providers take record 5*bninrevenuein*2021, , 2022-05-24.

Taylor, Tania Lynn, and Sean Collins, 2022, Ingraining sustainability in the next era of ESG investing, *Deloitte Insights* .

White, Lawrence J., 2001, The Credit Rating Industry: An Industrial Organization Analysis, Technical report.

Wong, C, and E Petroy, 2020, Rate the raters 2020: Investor survey and interview results. sustainability.

Xia, Han, 2014, Can investor-paid credit rating agencies improve the information quality of issuer-paid rating agencies?, *Journal of Financial Economics* 111, 450–468.

Yang, Ruoke, 2021, What do we learn from ratings about corporate social responsibility (CSR)?, *Columbia Business School Research Paper* .

Figure 1: ESG ratings by industry and provider

This figure shows the average ESG ratings provided by *HighIndex* (solid bar) and *LowIndex* (dotted bar) raters for each of the Fama-French 12 industry classifications.

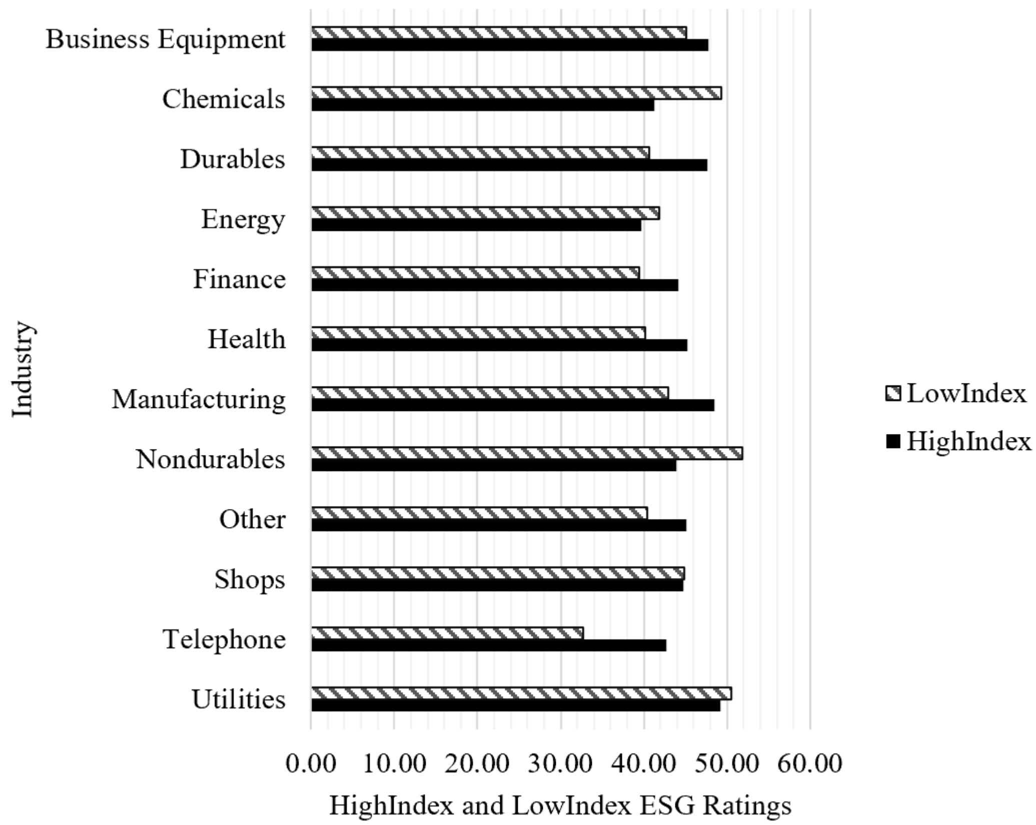
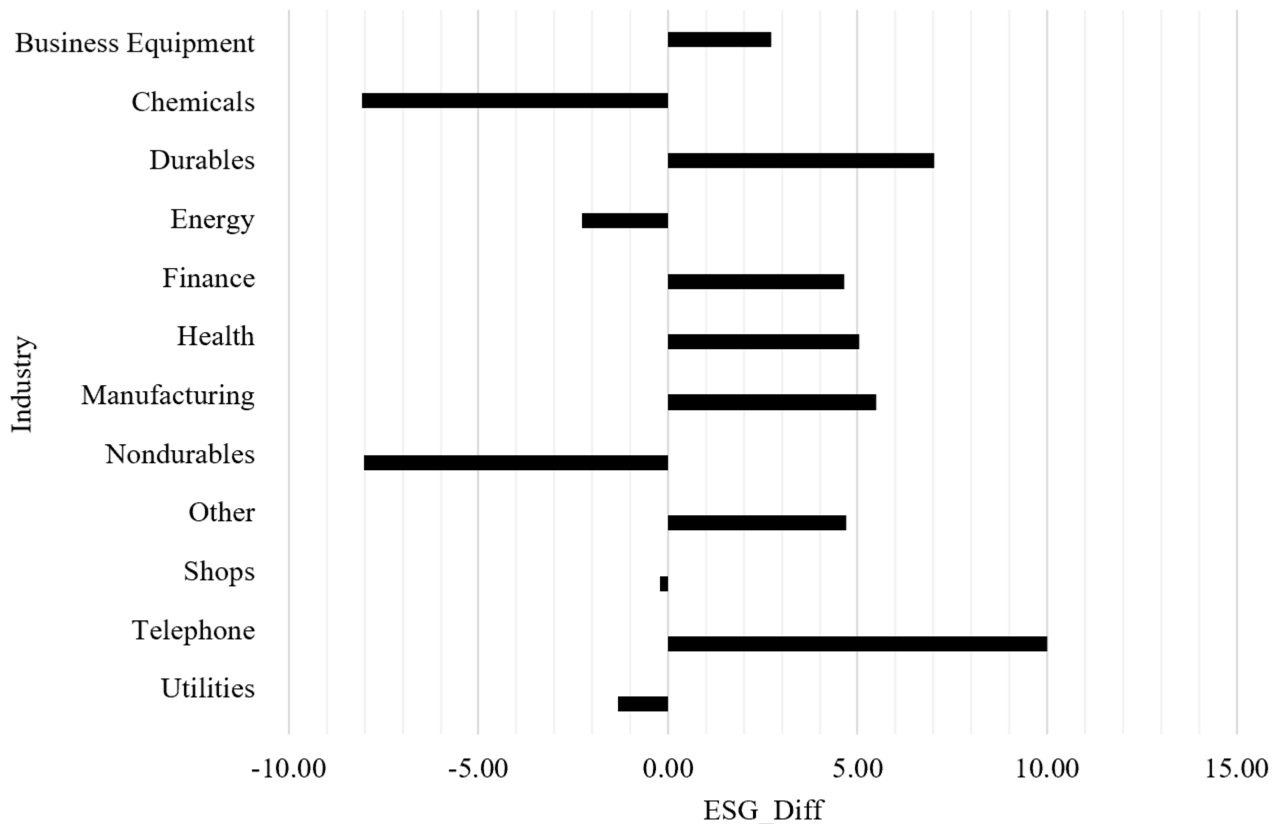


Figure 2: ESG rating differences by industry and year

This figure shows the mean difference between *HighIndex* and *LowIndex* ESG ratings for each of the Fama-French 12 industry classifications (Panel A) and each year from 2012 to 2019 (Panel B).

Panel A: Differences in ESG ratings by industry



Panel B: Differences in ESG ratings over time

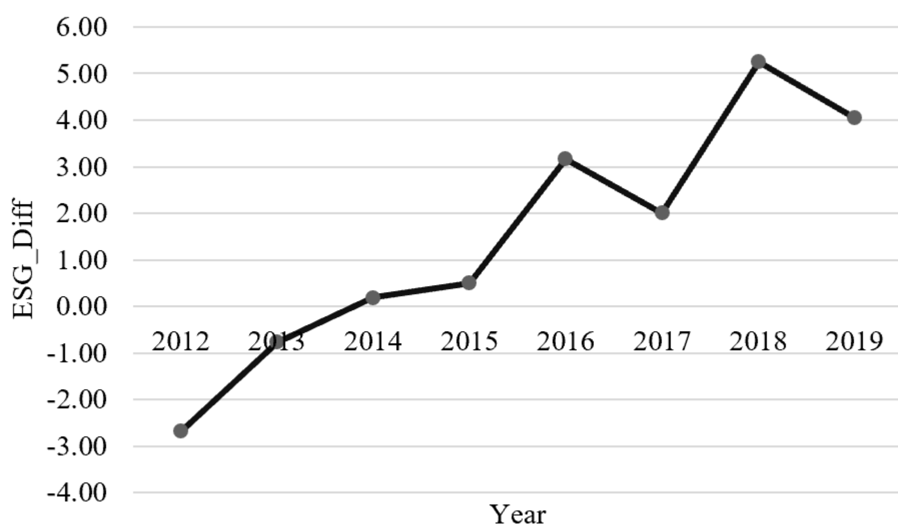


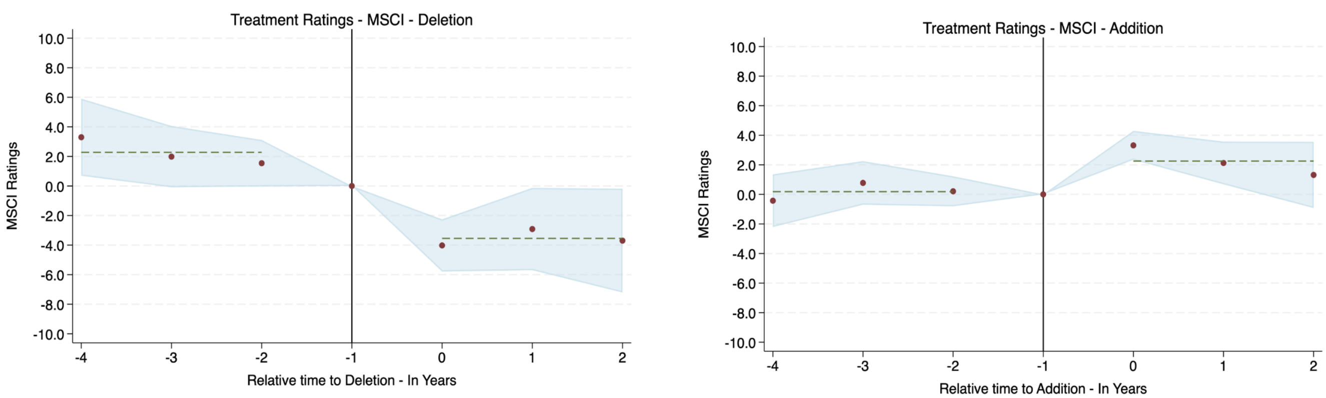
Figure 3: ESG ratings changes around index addition/deletion events

This figure plots the time coefficient from the following equation, separately for $HighIndex$ and $LowIndex$ rating agencies.

$$ESG\_Diff_{i,t} = \beta_1 Post_t \times Treat\_firm_{i,t} + \gamma Controls_{i,t} + Firm \times CohortFE + Year \times CohortFE + \epsilon_{i,t}$$

Panel A plots the trend of the coefficient on $HighIndex$ ratings. Panel B plots the trend of the coefficient on $LowIndex$ ratings.

Panel A: Trends in $HighIndex$ ratings around deletion/addition



Panel B: Trends in $LowIndex$ ratings around deletion/addition

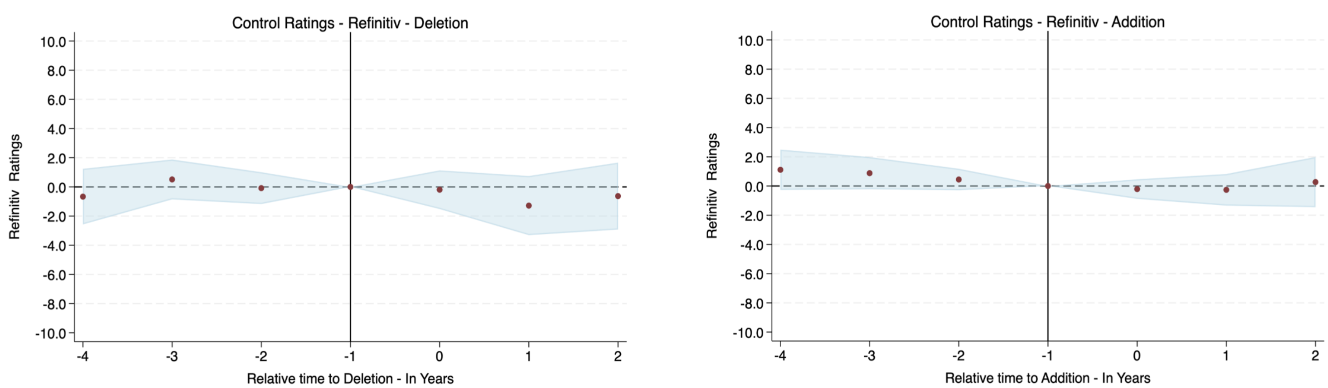


Table 1: Descriptive Statistics

This table presents descriptive statistics for key variables used in our analyses. Panel A presents mean, standard deviation, and distributional statistics. Panel B presents Pearson correlation coefficients. Panel C presents the ESG indexes we use for our study. All variable definitions appear in Appendix A.

Panel A: Summary statistics

	N	Mean	Std. Dev.	25%	Median	75%
Analyst	7,214	12.662	8.133	5.833	11.167	18.083
BTM	7,214	0.422	0.322	0.194	0.354	0.593
ESG_Diff	7,214	2.517	18.453	-10.353	4.809	16.372
ESG_Disclosure	7,214	38.795	10.585	31.496	33.764	44.653
HighIndex	7,214	45.507	9.347	40.000	45.000	51.000
IOR	7,214	0.810	0.189	0.721	0.858	0.953
LowIndex	7,214	43.019	18.959	28.087	40.116	56.614
Return	7,214	-0.024	0.333	-0.221	-0.048	0.134
ROA	7,214	0.113	0.121	0.067	0.118	0.170
Size	7,214	8.550	1.470	7.477	8.399	9.511

Table 1: Descriptive Statistics (cont.)

Panel B: Correlation matrix

Variables	*Analyst*	*BTM*	*ESG_Diff*	*ESG_Disclosure*	*IOR*	*HighIndex*	*LowIndex*	*Return*	*ROA*	*Size*
Analyst	1.000									
BTM	-0.174	1.000								
ESG_Diff	-0.446	-0.003	1.000							
ESG_Disclosure	0.472	-0.097	-0.617	1.000						
IOR	0.124	-0.078	0.017	0.066	1.000					
HighIndex	0.108	-0.141	0.220	0.242	0.118	1.000				
LowIndex	0.480	-0.058	-0.880	0.733	0.028	0.268	1.000			
Return	-0.016	-0.199	0.041	-0.037	0.007	0.016	-0.024	1.000		
ROA	0.180	-0.215	-0.149	0.128	0.141	0.047	0.171	0.007	1.000	
Size	0.746	-0.298	-0.551	0.610	0.053	0.164	0.612	0.063	0.247	1.000

Table 1: Descriptive Statistics (cont.)

Panel C: Index Sample

S. No	*Index*	*Index Launch Date*	*Linked ETF/MF*	*ETF/MF Launch Date*	*TNA as of Sep 2022*	*Remarks*
1	*MSCI USA Extended ESG Focus*	Mar-27-18	iShares ESG Aware MSCI USA ETF	Dec-1-16	20,038	Previously called iShares MSCI USA ESG Optimized ETF. Prior to Mar 2018 the ETF was tracking MSCI USA ESG Focus Index.
2	*MSCI USA SRI Select Reduced Fossil Fuel Index*	Oct-4-19	iShares MSCI USA SRI UCITS ETF USD (Acc)	Jul-11-16	7,347	Prior to Nov-2019, the Fund used a different benchmark
3	*MSCI USA ESG Enhanced Focus Index*	Jan-15-19	iShares MSCI USA ESG Enhanced UCITS ETF	Mar-6-19	5,866	
4	*MSCI USA Low Carbon SRI Leaders Index*	Feb-27-18	Xtrackers MSCI USA ESG UCITS ETF 1C	May-8-18	4,549	ETF is domiciled outside USA
5	*MSCI KLD 400 Social*	Sep-1-10	iShares MSCI KLD 400 Social ETF	Nov-14-06	3,285	The index was operated by FTSE KLD prior to Sep-2010.
6	*MSCI USA Extended ESG Select*	Mar-27-18	iShares MSCI USA ESG Select ETF	Jan-24-05	2,975	ETF was previously linked to MSCI USA ESG Select Index. ETF filled the Change document on 16-Mar-2018
7	*MSCI USA Extended ESG Leaders TR USD*	Feb-27-19	iShares ESG MSCI USA Leaders ETF	May-7-19	2,830	
8	*TIAA ESG USA Large-Cap Value*	Nov-7-16	Nuveen ESG Large-Cap Value ETF	Dec-13-16	1,385	Customised Index by MSCI for TIAA (parent company of Nuveen)
9	*MSCI USA Small Cap Extended ESG Focus*	Dec-12-17	iShares ESG Aware MSCI USA Small-Cap ETF	Apr-10-18	1,336	
10	*MSCI USA ESG Universal Select Business Screens Index*	Apr-25-19	Invesco MSCI USA ESG Universal Screened UCITS ETF	Jun-13-19	978	ETF is domiciled outside USA
11	*TIAA ESG USA Small-Cap*	Nov-7-16	Nuveen ESG Small-Cap ETF	Dec-13-16	836	Customised Index by MSCI for TIAA (parent company of Nuveen)
12	*MSCI ACWI Low Carbon Target*	Sep-23-14	iShares MSCI ACWI Low Carbon Target ETF	Dec-8-14	761	Includes 21 developed (including USA) and 25 emerging market countries
13	*MSCI USA SRI S-Series PAB 5% Capped Index*	Apr-25-19	BNP Paribas Easy MSCI USA SRI S-Series 5% Capped UCITS ETF - USD (Cap)	Oct-18-17	523	Prior to Apr-2019, ETF was linked to some other Index.
14	*MSCI USA ESG Screened Index*	Oct-22-18	iShares MSCI USA ESG Screened UCITS ETF	Oct-19-18	318	ETF is domiciled outside USA
15	*MSCI USA Select ESG Rating and Trend Leaders Index*	Jan-10-18	Lyxor MSCI USA ESG Trend Leaders UCITS ETF	Mar-3-18	300	ETF is domiciled outside USA
16	*MSCI USA Islamic Index*	Jul-26-07	iShares MSCI USA Islamic UCITS ETF	Dec-7-07	158	ETF is domiciled outside USA

Table 2: ESG ratings differences and stock return performance

This table presents regression summary statistics from estimating the following equation:

$$ESG\_Diff_{i,t+1} = \beta_1 StockPerformance_{i,t} + \gamma Controls_{i,t} + \phi_k + \tau_t + \epsilon_{i,t}$$

The dependent variable is the difference between *HighIndex* ESG ratings and *LowIndex* ESG ratings. The independent variables are stock return performance (*Return*, *BTM*) and firm profitability (*ROA*). *, **, and *** denote statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are clustered by industry. T-statistics are reported in parentheses. All variable definitions appear in Appendix A.

	Dependent variable: ESG_Diff			
	(1)	(2)	(3)	(4)
Return	2.306***			1.561**
	(3.38)			(2.28)
BTM		-5.740***		-5.550***
		(-5.27)		(-5.97)
ROA			1.236	-1.381
			(0.41)	(-0.44)
Analyst	-0.084	-0.074	-0.096	-0.068
	(-0.55)	(-0.47)	(-0.61)	(-0.42)
IOR	0.555	0.767	0.533	0.813
	(0.19)	(0.26)	(0.18)	(0.28)
ESG_Disclosure	-0.874***	-0.843***	-0.881***	-0.840***
	(-20.38)	(-19.86)	(-20.75)	(-19.74)
Size	-2.889***	-3.347***	-2.793***	-3.378***
	(-7.17)	(-7.65)	(-6.45)	(-6.78)
Year F.E.	Yes	Yes	Yes	Yes
Industry F.E.	Yes	Yes	Yes	Yes
R^2	0.482	0.488	0.480	0.488
N	7214	7214	7214	7214

Table 3: ESG ratings, stock return performance, and index inclusion

Panel A presents descriptive statistics for key variables used in our index constituent sample. Panels B, C, and D present regression summary statistics from estimating the following equation: estimating the following equation:

$$Y_{i,t+1} = \beta_1(\Delta)HighIndex\_StockPerf_{i,t} + \beta_2(\Delta)HighIndex\_NonStockPerf_{i,t} + \gamma Controls_{i,t} + \phi_k + \tau_t + \epsilon_{i,t}$$

In Panel B, the dependent variable ($ESGIndexInclude$) is an indicator variable that equals 1 when the firms are included in MSCI ESG indices and equals 0 when the firms are not included in any MSCI ESG indices. In Panel C, the dependent variable ($\Delta ESGIndexNum\_UpOrDown$) indicates whether the number of MSCI ESG indices including the firm increased, decreased, or did not change. In Panel D, the dependent variables are: $\Delta ESGIndexNum$ is the change in the number of MSCI ESG indices that include the firm, and $\Delta ESGIndexWeight$ is the change in the cumulative weight assigned to the firm across all MSCI ESG indices. *, **, and *** denote statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are clustered by industry. T-statistics are reported in parentheses. All variable definitions appear in Appendix A.

Panel A: Summary statistics

	N	Mean	Std. Dev.	25%	Median	75%
Dependent variables:						
$ESGIndexInclude$	7,866	0.509	0.500	0.000	1.000	1.000
$ESGIndexNum$	7,866	1.272	2.124	0.000	1.000	2.000
$ESGIndexWeight$	7,866	0.372	1.608	0.000	0.014	0.210
$\Delta ESGIndexNum$	7,866	0.460	1.112	0.000	0.000	1.000
$\Delta ESGIndexWeight$	7,866	0.137	0.870	0.000	0.000	0.051
Independent variables:						
$Analyst$	7,866	11.400	8.060	5.000	9.167	16.250
BTM	7,866	0.444	0.444	0.213	0.382	0.621
$ESG\_Disclosure$	7,866	37.281	10.700	30.520	32.673	42.108
$HighIndex\_StockPerf$	7,866	44.824	0.022	44.818	44.825	44.832
$HighIndex\_NonStockPerf$	7,866	0.540	9.116	-4.839	0.179	6.175
IOR	7,866	0.809	0.191	0.721	0.860	0.954
ROA	7,866	0.115	0.125	0.065	0.117	0.168
$Size$	7,866	8.234	1.544	7.087	8.040	9.238
$Spread$	7,866	0.079	0.169	0.024	0.045	0.085
ΔBTM	7,866	-0.000	0.315	-0.068	-0.010	0.054
$\Delta ESG\_Disclosure$	7,866	1.477	3.533	0.000	0.211	2.133
$\Delta HighIndex\_StockPerf$	7,866	0.886	0.008	0.884	0.886	0.889
$\Delta HighIndex\_NonStockPerf$	7,866	-0.000	5.651	-2.886	0.110	3.109
ΔROA	7,866	-0.002	0.064	-0.014	-0.000	0.010

Table 3: ESG ratings, stock return performance, and index inclusion (cont.)

Panel B: The inclusion of ESG indices

	Dependent variable: $ESGIndexInclude$		
	(1)	(2)	(3)
$HighIndex\_StockPerf$	7.202**		8.288**
	(2.05)		(2.15)
$HighIndex\_NonStockPerf$		0.089***	0.089***
		(13.78)	(14.05)
BTM	0.085	0.388*	0.317
	(0.44)	(1.80)	(1.46)
ROA	0.974**	0.862	0.803
	(2.09)	(1.59)	(1.53)
$Spread$	-0.611	-0.416	-0.631
	(-1.07)	(-0.60)	(-0.86)
$Analyst$	0.032***	0.046***	0.044***
	(3.15)	(6.73)	(6.25)
IOR	0.669**	0.358	0.351
	(2.56)	(1.29)	(1.29)
$ESG\_Disclosure$	0.030***	0.014	0.014
	(3.16)	(1.41)	(1.38)
$Size$	0.686***	0.738***	0.742***
	(16.60)	(16.01)	(16.15)
Year F.E.	Yes	Yes	Yes
Industry F.E.	Yes	Yes	Yes
N	7866	7866	7866

Table 3: ESG ratings, stock return performance, and index inclusion (cont.)

Panel C: Changes in ESG index inclusion

	Dependent variable:	
	$\Delta ESGIndexNum\_UpOrDown$	$\Delta ESGIndexNum$
	(1)	(2)
$\Delta HighIndex\_StockPerf$	8.016*	7.718**
	(1.68)	(1.99)
$\Delta HighIndex\_NonStockPerf$	0.022***	0.023***
	(4.47)	(4.71)
ΔBTM	-0.626***	-0.716***
	(-2.78)	(-4.04)
ΔROA	0.126	0.085
	(0.25)	(0.14)
$Spread$	-0.407	-0.461
	(-1.47)	(-1.64)
$Analyst$	0.008	0.011
	(1.23)	(1.47)
IOR	-0.188	-0.188
	(-0.44)	(-0.44)
$\Delta ESG\_Disclosure$	-0.003	0.001
	(-0.27)	(0.11)
$Size$	0.431***	0.468***
	(12.73)	(11.51)
Year F.E.	Yes	Yes
Industry F.E.	Yes	Yes
N	7866	7866

Table 3: ESG ratings, stock return performance, and index inclusion (cont.)

Panel D: Changes in cumulative weight across ESG indices

	Dependent variable: $\Delta ESGIndexWeight$		
	(1)	(2)	(3)
$\Delta HighIndex\_StockPerf$	0.624***		0.626***
	(3.82)		(3.89)
$\Delta HighIndex\_NonStockPerf$		0.002	0.003
		(1.62)	(1.66)
ΔBTM	-0.013	-0.088***	-0.010
	(-0.47)	(-3.12)	(-0.37)
ΔROA	-0.145	-0.021	-0.147
	(-1.39)	(-0.19)	(-1.43)
$Spread$	0.092	0.083	0.092
	(1.20)	(1.12)	(1.20)
$Analyst$	0.006*	0.006*	0.006*
	(2.10)	(2.01)	(2.11)
IOR	-0.193**	-0.193**	-0.193**
	(-2.61)	(-2.55)	(-2.61)
$\Delta ESG\_Disclosure$	-0.004	-0.004	-0.004
	(-0.75)	(-0.78)	(-0.77)
$Size$	0.133***	0.136***	0.134***
	(6.16)	(6.11)	(6.13)
Year F.E.	Yes	Yes	Yes
Industry F.E.	Yes	Yes	Yes
N	7866	7866	7866

Table 4: Index additions and deletions

Panel A presents descriptive statistics for variables we used in our regression. Panel B presents regression summary statistics from estimating the following equation:

$$ESG\_Diff_{i,t} = \beta_1 Post_t \times Treat\_firm_{i,t} + \gamma Controls_{i,t} + Firm \times CohortFE + Year \times CohortFE + \epsilon_{i,t}$$

The dependent variable ESG_Diff is the difference between $HighIndex$ and $LowIndex$ ratings. $Post$ is an indicator variable that equals 1 for the period after the firm is added/deleted from the index and zero otherwise. $Treat\_firm$ is an indicator variable that equals 1 if the firm is added/deleted to the index after initial index creation and 0 if the firm has been part of the index since the inception of the index. Additions consists of observations corresponding to firms that are added to an index post its creation. Deletion comprises of observations corresponding to firms that are deleted from an index. *, **, and *** denote statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. See Appendix A for variable definitions.

Panel A: Summary statistics

	Additions (n=1,171)			*Deletions (n=281)*			*Diff in Mean*	*Stable (n=1,031)*		
	Mean	*Median*	*Std. Dev.*	*Mean*	*Median*	*Std. Dev.*	*Add - Del*	*Mean*	*Median*	*Std. Dev.*
HighIndex	53.188	53.000	18.357	40.004	39.000	19.813	13.184***	59.171	61.000	19.387
LowIndex	51.952	51.658	18.908	50.320	49.188	18.114	1.632	52.177	52.519	19.899
ESG_Diff	1.236	0.366	21.566	-10.316	-9.077	23.474	11.552***	6.994	9.260	22.630
ESG_Diff_Std	0.014	-0.062	1.100	-0.557	-0.498	1.193	0.570***	0.299	0.393	1.155
Analyst	17.838	17.000	8.643	18.459	17.000	9.744	-0.621	15.618	14.000	9.882
Cash	0.145	0.091	0.149	0.126	0.073	0.135	0.019**	0.152	0.100	0.157
Div	0.019	0.012	0.022	0.019	0.013	0.021	0.000	0.020	0.012	0.022
Has_RD	0.478	0.000	0.500	0.509	1.000	0.501	-0.031	0.404	0.000	0.491
IOR	0.828	0.846	0.130	0.859	0.872	0.130	-0.031***	0.843	0.861	0.131
Lev	0.261	0.249	0.162	0.295	0.296	0.162	-0.033***	0.258	0.247	0.154
MB	5.497	3.426	6.506	4.777	2.798	7.295	0.720	4.388	3.344	4.162
RD	0.030	0.002	0.055	0.024	0.000	0.046	0.007**	0.033	0.010	0.051
ROA	0.101	0.095	0.083	0.087	0.084	0.079	0.014***	0.087	0.088	0.065
Sales_Growth	0.116	0.081	0.172	0.053	0.045	0.152	0.063***	0.091	0.066	0.142
Size	8.950	8.946	1.470	8.954	8.855	1.431	-0.004	8.865	8.509	1.735
Tangibility	0.192	0.114	0.195	0.242	0.178	0.214	-0.050***	0.184	0.132	0.174
TobinQ	2.591	2.039	1.633	2.101	1.693	1.168	0.490***	2.332	1.969	1.272
Total assets	22,449	7,679	44,567	22,057	7,007	46,523	392	30,660	4,958	58,932

Table 4: Index additions and deletions (cont.)

Panel B: Difference between *HighIndex* and *LowIndex* Ratings

	Deletions			*Additions*		
Dependent Variable	*ESG_Diff* (1)	*ESG_Diff* (2)	*ESG_Diff_Std* (3)	*ESG_Diff* (4)	*ESG_Diff* (5)	*ESG_Diff_Std* (6)
Post × *Treat*	-4.207***	-4.697***	-0.232***	2.783***	2.928***	0.147***
	(-2.96)	(-3.45)	(-3.38)	(3.68)	(3.85)	(3.82)
Lev		0.441	0.035		1.543	0.085
		(0.11)	(0.18)		(0.59)	(0.65)
Tangibility		3.761	0.209		-0.379	-0.008
		(0.47)	(0.52)		(-0.07)	(-0.03)
TobinQ		-0.890*	-0.044		0.243	0.012
		(-1.67)	(-1.62)		(0.65)	(0.66)
ROA		-1.739	-0.074		2.814	0.141
		(-0.26)	(-0.22)		(0.56)	(0.56)
Cash		3.934	0.173		-5.594**	-0.300**
		(1.21)	(1.05)		(-2.28)	(-2.42)
RD		-71.963***	-3.647***		-25.806**	-1.337**
		(-3.65)	(-3.67)		(-2.18)	(-2.25)
Has_RD		-7.886**	-0.395**		-3.580	-0.188
		(-2.23)	(-2.21)		(-1.32)	(-1.40)
MB		0.020	0.001		-0.106	-0.005
		(0.16)	(0.18)		(-1.46)	(-1.44)
Sales_growth		1.507	0.085		2.300**	0.123**
		(0.93)	(1.05)		(2.05)	(2.19)
Size		-7.299***	-0.384***		-6.131***	-0.322***
		(-4.67)	(-4.84)		(-6.04)	(-6.26)
Div		-51.937	-2.797		15.642	0.603
		(-1.32)	(-1.39)		(0.64)	(0.49)
IOR		0.847	0.047		-0.581	-0.018
		(0.30)	(0.33)		(-0.33)	(-0.20)
Analyst		1.002	0.052		2.959**	0.149**
		(0.54)	(0.56)		(2.26)	(2.24)
Firm×Cohort F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Period×Cohort F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.766	0.769	0.771	0.736	0.739	0.742
N	6803	6803	6803	14110	14110	14110

Table 5: *HighIndex* overall ESG ratings and category scores

This table presents regression summary statistics from estimating the following model:

$$HighIndex_{i,t+1} = \beta_1 Return_{i,t} + \gamma Category\_score_{i,t+1} + \phi_k + \tau_t + \epsilon_{i,t}$$

The dependent variable is the *HighIndex* ESG ratings. The independent variables are stock performance (*Return*) and scores for each of the 10 component categories outlined in the MSCI ESG ratings methodology. These categories are Climate Change, Natural Resource Use, Waste Management, Environmental Opportunities, Human Capital, Product Safety, Social Opportunities, Corporate Governance, Business Ethics, and Stakeholder Opposition. The missing values of category scores are replaced with the mean. *, **, and *** denote statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are clustered by industry. T-statistics are reported in parentheses.

	Dependent variable: *HighIndex*
Return	0.428***
	(3.16)
Climate_Change	1.061***
	(9.28)
Natural_Resource_Use	0.402
	(1.73)
Waste_Management	0.844**
	(2.93)
Environmental_Opportunities	2.261***
	(18.43)
Human_Capital	1.999***
	(11.65)
Product_Safety	2.320***
	(12.60)
Social_Opportunities	0.798***
	(4.28)
Corporate_Governance	1.146***
	(13.67)
Business_Ethics	2.006***
	(14.49)
Stakeholder_Opposition	1.575***
	(5.19)
Year F.E.	Yes
Industry F.E.	Yes
R^2	0.741
N	7214

Table 6: *LowIndex* overall ESG ratings and category scores

This table presents regression summary statistics from estimating the following equation:

$$LowIndex_{i,t+1} = \beta_1 Return_{i,t} + \gamma Category\_score_{i,t+1} + \phi_k + \tau_t + \epsilon_{i,t}$$

The dependent variable is the *LowIndex* ESG rating. The independent variables are stock performance (*Return*) and scores for each of the 10 component categories outlined in the Refinitiv ESG ratings methodology. These categories are Resource Use, Emissions, Environmental Innovation, Workforce, Human Rights, Community, Product Responsibility, Management, Shareholder, and CSR strategy. The missing values of category scores are replaced with the mean. *, **, and *** denote statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. Standard errors are clustered by industry. T-statistics are reported in parentheses.

	Dependent variable: *LowIndex*
Return	0.048
	(0.40)
Resource_Use	0.091***
	(10.17)
Emissions	0.083***
	(7.34)
Environmental_Innovation	0.078***
	(6.60)
Workforce	0.138***
	(12.85)
Human_Rights	0.099***
	(14.41)
Community	0.118***
	(15.13)
Product_Responsibility	0.102***
	(10.49)
Management	0.211***
	(20.81)
Shareholder	0.060***
	(14.16)
CSR_Strategy	0.025***
	(6.33)
Year F.E.	Yes
Industry F.E.	Yes
R^2	0.971
N	7214

Table 7: ESG ratings changes and future ESG outcomes

This table presents regression summary statistics from estimating the following equation:

$$ESG\_Outcome_{i,t+1} = \beta_1 Post_t \times Treat\_firm_{i,t} + \gamma Controls_{i,t} + Firm \times CohortFE + Year \times CohortFE + \epsilon_{i,t}$$

The dependent variables are *#Penalty* (the number of penalties imposed on the firm for violations in the following calendar year), *Gender_Diversity* (Ratio of women directors to total directors on the board in the following calendar year), *Racial_Diversity* (Ratio of non-Caucasian directors to total directors on the board in the following calendar year), and *CCExposure* (Relative frequency with which bigrams related to climate change occur in the transcripts of earnings conference calls in the following calendar year). *, **, and *** denote statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively. All variable definitions appear in Appendix A.

	# Penalty		*Gender Diversity*		*Racial Diversity*		*CCExposure*	
Dependent Variable	*Deletion* (1)	*Addition* (2)	*Deletion* (3)	*Addition* (4)	*Deletion* (5)	*Addition* (6)	*Deletion* (7)	*Addition* (8)
Post×Treat_firm	0.350	-0.105	0.008	0.008*	-0.005	0.000	0.014	-0.010
	(1.45)	(-0.98)	(1.24)	(1.85)	(-0.53)	(0.02)	(0.20)	(-0.28)
Lev	0.475*	0.315	-0.040**	-0.009	0.045**	0.025*	-0.063	-0.035
	(1.90)	(1.41)	(-2.52)	(-0.67)	(2.51)	(1.85)	(-0.42)	(-0.37)
Tangibility	1.463	-0.231	-0.009	-0.050	-0.013	0.048	-0.324	-0.435*
	(1.55)	(-0.35)	(-0.24)	(-1.49)	(-0.30)	(1.32)	(-0.74)	(-1.69)
TobinQ	0.086**	0.068**	0.004*	0.004***	-0.007**	-0.001	0.002	0.003
	(2.11)	(2.54)	(1.78)	(3.01)	(-2.16)	(-0.71)	(0.14)	(0.31)
ROA	0.333	0.831	-0.098***	0.036	-0.053	-0.046*	0.124	0.198
	(0.37)	(1.22)	(-3.05)	(1.57)	(-1.41)	(-1.71)	(0.47)	(1.06)
Cash	-0.533*	-0.873***	0.042**	0.057***	0.091***	0.075***	-0.425**	-0.117
	(-1.67)	(-4.18)	(2.12)	(4.11)	(4.98)	(5.58)	(-2.51)	(-1.06)
RD	-2.412	1.417	0.190*	0.113	-0.521***	-0.230*	0.454	0.626
	(-1.47)	(1.58)	(1.94)	(1.23)	(-3.28)	(-1.87)	(0.57)	(1.08)
Has_RD	0.206	0.383**	-0.026*	0.009	0.116***	0.062**	-0.057	-0.077
	(0.77)	(2.13)	(-1.76)	(0.77)	(3.54)	(2.09)	(-0.38)	(-0.70)
MB	-0.016	-0.015	-0.000	0.000	-0.000	-0.000	-0.004	-0.005***
	(-1.31)	(-1.49)	(-0.52)	(0.06)	(-0.61)	(-0.40)	(-1.56)	(-2.81)
Sales_growth	-0.070	-0.208	-0.018***	-0.028***	0.025**	0.019***	0.009	0.065
	(-0.40)	(-1.33)	(-2.76)	(-5.37)	(2.42)	(2.90)	(0.12)	(1.39)
Size	0.215	0.557***	0.029***	0.024***	-0.035***	-0.017***	0.079	-0.018
	(1.41)	(4.71)	(4.51)	(5.01)	(-4.73)	(-3.45)	(1.47)	(-0.56)
Div	-6.703**	-2.872	-0.139	0.020	-0.060	-0.163	-0.797	-0.672
	(-2.09)	(-1.37)	(-1.05)	(0.18)	(-0.42)	(-1.54)	(-0.80)	(-0.89)
IOR	0.174	-0.076	0.020	0.008	0.052***	0.028***	0.171	0.008
	(0.51)	(-0.37)	(1.39)	(0.86)	(3.41)	(2.93)	(1.21)	(0.11)
Analyst	-0.075	-0.075	0.010	0.008	-0.012	-0.022***	-0.108	0.076
	(-0.58)	(-0.86)	(1.25)	(1.48)	(-1.52)	(-3.90)	(-1.27)	(1.45)
Firm×Cohort F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period×Cohort F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.735	0.698	0.745	0.756	0.796	0.800	0.772	0.750
N	6353	13357	7325	12884	5214	10766	6165	12906