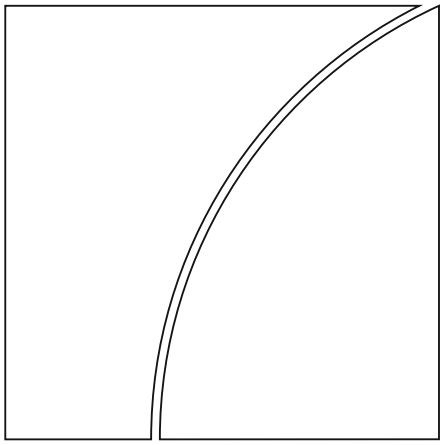
BANK FOR INTERNATIONAL SETTLEMENTS



BIS Working Papers

No 1008

Deconstructing ESG scores: How to invest with your own criteria

by Torsten Ehlers, Ulrike Elsenhuber, Anandakumar Jegarasasingam and Eric Jondeau

Monetary and Economic Department

March 2022

JEL classification: G11, G24, M14, Q01.

Keywords: sustainable investment, ESG ratings, ESG investing, negative screening, best-in-class screening.

BIS Working Papers are written by members of the Monetary and Economic Department of the Bank for International Settlements, and from time to time by other economists, and are published by the Bank. The papers are on subjects of topical interest and are technical in character. The views expressed in them are those of their authors and not necessarily the views of the BIS.

This publication is available on the BIS website (www.bis.org).

ISSN 1020-0959 (print)
ISSN 1682-7678 (online)

Deconstructing ESG Scores: How to Invest with Your own Criteria*

Torsten Ehlers, Ulrike Elsenhuber,

Anandakumar Jegarasasingam, and Eric Jondeau

February 2022

Abstract

Environmental, Social, and Governance (ESG) scores are becoming an increasingly important tool for asset managers to design and implement ESG investment strategies. They amalgamate a broad range of fundamentally different factors, creating ambiguity for investors as to the signals of higher or lower ESG scores. We explore the feasibility and performance of more targeted investment strategies based on specific categories by deconstructing ESG scores into their granular components. First, we investigate the characteristics of the various categories underlying ESG scores. Not all types of ESG categories lend themselves to more targeted strategies, which is related to both limits to ESG data disclosure and the fundamental challenge of translating qualitative characteristics into quantitative measures. Second, we consider an investment scheme based on the exclusion of firms with the lowest scores in each category of interest. In most cases, this targeted strategy still allows investors to substantially improve the portfolio headline ESG score, with only a marginal impact on financial performance relative to a broad stock market benchmark. The exclusion results in regional and sectoral biases relative to the benchmark, which may be undesirable for some investors. We then implement a "best-in-class" strategy, based on excluding firms with the lowest category scores and reinvesting the proceeds in firms with the highest scores maintaining the same regional and sectoral composition. This approach reduces the tracking error of the portfolio and slightly improves its risk-adjusted performance while still yielding a large gain in the headline ESG score.

JEL: G11, G24, M14, Q01

Keywords: Sustainable investment, ESG ratings, ESG investing, Negative screening, Best-in-class screening

*We would like to thank Raphael Auer, Claudio Borio, Pierre Cardon, Stijn Claessens, Ingo Fender, Enisse Kharoubbi, Luiz Pereira da Silva, Murphy Phan, and participants in a webinar held at the Bank for International Settlements (BIS) for useful comments and suggestions. T. Ehlers: Bank for International Settlements, Centralbahnplatz 2, 4051 Basel, Switzerland (e-mail: Torsten.Ehlers@bis.org). U. Elsenhuber: Bank for International Settlements, Aeschenplatz 1, 4051 Basel, Switzerland (e-mail: Ulrike.Elsenhuber@bis.org). A. Jegarasasingam: Bank for International Settlements, Centralbahnplatz 2, 4051 Basel, Switzerland (e-mail: Kumar.Jegarasasingam@bis.org). E. Jondeau: University of Lausanne, Enterprise for Society (E4S) Center, and Swiss Finance Institute, Faculty of Business and Economics (HEC Lausanne), Extranef 232, CH-1015 Lausanne, Switzerland (e-mail: Eric.Jondeau@unil.ch). The views expressed in this paper are those of the authors and should not be interpreted to reflect those of the BIS.

1 Introduction

Environmental, Social and Governance (ESG) investing has enjoyed rapid growth and by some measure has already reached $35 trillion – more than one third of global total assets under management (GSIA, 2020). Further rapid growth is expected. ESG funds' assets under management could exceed $50 trillion by 2025 (Bloomberg, 2021). This trend presents an opportunity for investment managers, and potentially for society as a whole.

ESG scores are a key tool for implementing investment managers' ESG strategies (Amel-Zadeh and Serafeim, 2021). Among the most popular strategies are ESG integration, which employs ESG factors alongside financial factors for portfolio selection, and negative screening, whereby assets with the worst (or "worst-in-class") ESG characteristics are excluded. Typically, investment managers rely on ESG scores by one or several data providers to measure ESG performance. ESG scores are therefore central to ESG investing and, by extension, to a substantial and rising share of investment allocations globally.

The use of ESG scores faces some well-known challenges, however. One key challenge for investment managers is the very low correlation of scores among the different major data providers. Investment managers may thus arrive at different portfolio selections using the same strategy but different ESG data providers.[1] While the correlation between credit ratings is usually close to 99%, Berg *et al.* (2019) find that on average, the correlation between ESG ratings is only slightly above 50%.

In this paper, we take an investment manager perspective and aim to circumvent the inconsistencies of ESG scores by deconstructing them and focusing on the underlying data points. Our main question is as follows: Can an investment manager construct a portfolio of equities (from a broad investable universe) that achieves a given financial performance, but with better underlying ESG characteristics?[2] Cutting straight through to the underlying characteristics not only shields against potential inconsistencies in ESG scoring and

[1] Perhaps the most famous example is the Japan Government Pension Investment Fund (GPIF), the largest investment fund in the world and one of the trailblazers in ESG investing. In looking to implement a positive screening strategy (selecting the assets with the best ESG scores) for its Japanese equity portfolio, the GPIF was presented with a starkly different set of assets based on the ESG scores from two major data providers (Bloomberg, 2017).

[2] Pedersen *et al.* (2021) propose the concept of an ESG-efficient frontier – the highest attainable Sharpe ratio for each ESG level. Conceptually, our approach attempts to maximize the ESG level, while keeping the performance of the portfolio as close as possible to that of a diversified benchmark, which could be viewed as the efficient frontier of a given universe of investable assets – including assets with both low and high ESG scores.

weighing methods, but also affords asset managers the flexibility to focus on specific aspects within the ESG sphere, mitigating the ambiguity that the amalgamation of a large set of diverse ESG factors inherently creates. Achieving given risk-return characteristics isolates the effect of implementing different degrees of ESG screening. Importantly, it also resembles the problem in implementing their ESG strategy that is faced by many large investors, who typically first screen for equity funds that match their desired financial performance. Indeed, for most investment funds, the primary mandate remains financial performance, while ESG considerations must either support or be neutral to the primary mandate.

To define the scope of the analysis, we naturally have to make a few fundamental choices. The first concerns the scope of underlying ESG data points. Our analysis uses Refinitiv (formerly Thomson Reuters Asset4) data, which to our knowledge is the only major ESG data provider that makes the underlying data points publicly available. We use all the ESG data points that Refinitiv uses to calculate its ESG scores – 186 comparable measures that Refinitiv uses to define ten ESG categories, which are then combined into the three E, S, and G pillars. The second choice regards the set of underlying ESG characteristics of interest. This choice is naturally subjective and depends on the desired ESG scheme of an investment manager. In our analysis, we consider the ESG categories in the Refinitiv dataset, which are likely representative of the thematic objectives that investment managers may want to implement. The third choice concerns the benchmark for financial performance, whose constituents span our universe of investable assets. Here we take a very broad equity index as our benchmark – the MSCI All Country World Index (ACWI) – to show the general geographical and sectoral applicability of our results. Choosing a broad benchmark index is possible given the broad coverage of listed firms in the Refinitiv ESG data.

Two challenges relating to how ESG data are disclosed and recorded are the treatment of missing data and the translation of qualitative information into a numeric value. As a general principle, numeric data points (e.g., annual CO_2 emissions in tons) are better suited for a screening investment strategy, as they enable a sharper distinction between firms. However, such distinction is not truly possible, when firms fail to disclose the related information, as the data point then assigned is the same value as if the firm scored poorly on the specific indicator. Boolean questions, on the other hand, only allow for a

limited differentiation between firms if many or all of the underlying ESG data points of interest are logical values.[3] As a result of both of these issues, screening strategies are only well suited for ESG strategies where the degree of acceptable exclusion is sufficiently high, as Boolean data points can be heavily skewed towards a missing or zero value. While these challenges are somewhat specific to the data that we use (in the Refinitiv ESG data, missing and zero values are indistinguishable), they nevertheless point to more general issues regarding the disclosure and numeric representation of ESG data.

Our key result is that constructing a portfolio with an ESG objective based on underlying ESG data points comes at virtually no cost in terms of financial performance. Investment managers can construct portfolios with specific ESG benefits, without relying on potentially confounding ESG scores, while still being able to match a given desired financial performance. For instance, with the 33% screening threshold, the portfolio score improves by 18 percentage points (pp) on average across ESG categories (from 62 to 80, with a maximum possible score of 100), while the increase is equal to 11 pp only for the overall ESG score. The screening process has a substantial impact on regional and sectoral exposures of the portfolio, which an otherwise passive investor may not be able to accept. Therefore, we implement a "best-in-class" strategy by excluding firms with the lowest scores and reinvesting the proceeds in firms with the highest scores in the same region-sector. The scores associated with this strategy improve slightly, while the portfolios have the same regional and sectoral exposures as the benchmark. The main cost of the screening strategy is the tracking error relative to the MSCI ACWI, although most of this cost comes from the use of an intermediary benchmark (based on the constituents of the index with an ESG score): The tracking error of this benchmark relative to the MSCI ACWI is on average equal to 0.9% per year. The additional cost due to the screening is below 0.7% per year even with a 33% screening threshold.

Related Literature. Our paper is part of the broader literature on the disclosure of nonfinancial information and has implications for ongoing global efforts to improve ESG-type corporate disclosures. The impact of mandatory sustainability disclosure has been analyzed in several papers, including Khan *et al.* (2016) and Ioannou and Serafeim (2017). In particular, Khan *et al.* (2016) find that firms with high ratings on material sustainability

[3]For instance, the database contains 95% of zeros associated with the question "Does the company provide information about the total individual compensation of all executives and board members?"

topics significantly outperform firms with poor ratings on these topics in terms of risk-adjusted returns. In contrast, firms with good ratings on immaterial sustainability issues do not significantly outperform firms with poor ratings on the same issues. Jouvenot and Krueger (2020) evaluate the effect of a law in the United Kingdom that mandates that publicly listed firms disclose their greenhouse gas (GHG) emissions in a standardized way in their annual reports. The authors find that firms respond to the law by reducing GHG emissions by approximately 16% percent. Relatedly, Mésonnier and Nguyen (2020) investigate the impact of a French regulation that requires institutional investors (except banks) to report annually on both their climate-related exposure and climate change mitigation policy. They find that investors subject to the disclosure requirements curtail their financing of fossil energy companies by some 40% compared to investors in the control group.

In practice, in the United States, the Securities and Exchange Commission (SEC) requires companies to make disclosures that are material to investment decisions, including ESG-type disclosures relating to human capital management, key performance metrics, and climate risks. However, it currently does not mandate ESG disclosures. One important obstacle to mandating ESG disclosures is the need to precisely define what information is material to an investment decision.[4] Voluntary ESG disclosure can be based on a large set of frameworks, such as the Global Reporting Initiative (GRI), the Sustainability Accounting Standards Board (SASB), or the Task Force on Climate-related Financial Disclosures (TCFD).[5]

The European Union has introduced several instruments related to the disclosure of nonfinancial information (collectively called EU sustainable finance regulations). The Non-Financial Reporting Directive (NFRD) has required large corporations to report on ESG information since 2018. The EU Taxonomy Regulation, introduced in July 2020, defines six environmental objectives (such as climate change mitigation and the transition to a circular economy) and considers an economic activity sustainable if it contributes at least to one of these objectives without, at the same time, causing significant harm to any of the other objectives. The new Sustainable Finance Disclosure Regulation (SFDR), introduced in March 2021, imposes more stringent requirements for financial services

[4]See https://www.sec.gov/news/speech/can-the-sec-make-esg-rules-that-are-sustainable. See also Khan *et al.* (2016) and Esty *et al.* (2020).

[5]See https://iclg.com/practice-areas/environmental-social-and-governance-law/usa.

institutions on sustainability-related disclosures made by these institutions with regard to sustainability risks.

At the world level, the International Sustainability Standards Board was launched in November 2021. Its objective is to establish sustainability disclosure standards covering both climate-related and other sustainability disclosures.

Our paper is further related to the literature that investigates ESG data quality issues. Not surprisingly, as the development of ESG ratings is relatively recent, there are sizable discrepancies among ESG ratings produced by different data providers, raising issues about their reliability and comparability. Berg *et al.* (2019) identify three sources of divergence in ESG ratings (divergence in scope, in measurement, and in weights) and find that differences in measurement explain most of the differences among ESG ratings, meaning that the same ESG attribute is measured using different underlying indicators. Gibson *et al.* (2021), Billio *et al.* (2021), and Serafeim and Yoon (2021) analyze the disagreement across data providers and evaluate its impact on future stock returns. High disagreement regarding the ESG quality of a firm tends to be associated with lower subsequent stock returns.

Other issues have been identified for ESG data. Berg *et al.* (2020) document large and repeated changes in historical ESG scores. While they find a positive relation between ESG scores and stock returns when updated data are used, the authors do not observe such a relationship with the initial data. Sahin *et al.* (2021) document the large proportion of missing information, which cast doubt on the reliability of ESG scores.

Finally, we build on the literature on ESG investing and its financial performance, which has grown very rapidly, as demonstrated by Friede *et al.* (2015). For a long time, firms with low ESG scores and "sin stocks" were expected to enjoy superior performance (Fabozzi *et al.*, 2008 and Hong and Kacperczyk, 2009). Recent analysis has also been spurred by the financial outperformance of firms with high ESG scores during the great financial crisis (Lins *et al.*, 2017) and the COVID-19 shock (Garel and Petit-Romec, 2021), although there is some evidence to the contrary (Demers *et al.*, 2021, Pástor *et al.*, 2021b, Scatigna *et al.*, 2021). One possible explanation for these diverging research conclusions is the heterogeneity and inconsistency of data, including diverging imputation methods employed in ESG scoring to address data gaps (Kotsantonis and Serafeim, 2019). Deconstructing ESG scores into their individual elements is aimed at shedding light on this

question.

In equilibrium, ESG screening should result in lower expected returns for firms with high ESG scores if investors have a preference for firms with high ESG quality. Pedersen *et al.* (2021) describe a model in which the relation between the ESG score and financial performance of firms depends on the role of the ESG score in investors' decisions. If investors have ESG preferences, the *expected return* of high-score firms should be lower than that of low-score firms. Pástor *et al.* (2021a) also find that in an equilibrium model with ESG preferences, green assets have negative alphas and brown assets have positive alphas. However, as pointed out by Pástor *et al.* (2021b), green assets have delivered higher *realized returns* in the recent period because of the demand pressure driven by investors' climate concerns.

The remainder of the paper is structured as follows. Section 2 describes our data. In Section 3, we present the main results regarding firms' disclosure of ESG information. Section 4 summarizes the main results for portfolio screening. Section 5 concludes.

2 Data

2.1 Construction of the Scores

The methodology adopted by Refinitiv for scoring firms is relatively complex, as it combines a vast amount of different types of data and different aggregation schemes (see Refinitiv, 2021). At the same time, it is strongly data-driven and transparent – due not least to the disclosure of both the underlying methodology and data points. First, the database is based on 450 data points (or metrics), which can be Boolean indicators and numeric indicators, such as ratios and analytics. Of these 450 metrics, 186 comparable measures are actually used for ESG scoring. Other data points cover different topics of interest but are not directly used for ESG scoring. The 186 comparable measures are then aggregated, using different weightings, into 10 categories. The 10 categories, in turn, are aggregated further to compute the three (E, S, and G) pillars.

The definition and characteristics of the pillars and categories are summarized in the table below (Refinitiv, 2021). ESG pillar scores are obtained by multiplying category scores with their category weights. For the E and S pillars, category weights vary across

industries depending on the materiality of the associated indicators. Some indicators are material for some industries but are not included in the calculation of the scores for the other industries.[6] For the G pillar, the weights of the three categories are the same across all industries, as indicated in the table. The overall ESG score is based on combining the three pillars, with weights that are specific to the industry of the assessed firm. All the scores are between 0 and 100, with 100 being the best possible score.

Pillars and categories	Nb of comparable measures	Themes
Environmental		
(1) Emissions reduction	28	Emissions; Waste; Biodiversity; Environmental management systems
(2) Innovation	20	Product innovation; Green revenues, research and development and capital expenditures
(3) Resource use	20	Water; Energy; Sustainable packaging; Environmental supply chain
Social		
(1) Community	10	Public health; Business ethics
(2) Human rights	8	Respect for fundamental human rights conventions
(3) Product responsibility	30	Responsible marketing; Product quality; Data privacy
(4) Workforce	14	Diversity and inclusion; Career development and training; Working conditions; Health and safety
Governance (weight)		
(1) CSR strategy (0.13)	12	Corporate Social Responsibility strategy; ESG reporting and transparency
(2) Management (0.67)	9	Structure (independence, diversity, committees); Compensation
(3) Shareholders (0.20)	35	Shareholder rights; Takeover defenses

[6]For instance, for the coal industry, the weights to compute the E pillar score are equal to 0.20 for Emissions reduction, 0.19 for Resource use, and 0.02 for Innovation, reflecting the materiality of the first two categories for this industry. In contrast, for banking services, the weights are equal to 0.02, 0.02, and 0.10, respectively.

An important aspect of the Refinitiv database is the data collection process. For the list of firms covered by Refinitiv, analysts collect information on individual ESG measures using numerous publicly available sources (including annual reports, CSR reports, company websites, and news sources). Essentially, the data collected by Refinitiv reflect the disclosure policy of the firms, except for some particular situations such as controversies, which also reflect reports from global media.[7]

The Refinitiv database is unique and highly suitable for economic analyses for two reasons. First, it provides all the underlying data points used to build the scores, particularly the 186 comparable measures used to calculate the 10 category scores. These data points allow us to identify what generates the particular distribution of the category scores, as detailed in Section 3. Second, the methodology used to build the scores is transparent, which also allows us to precisely interpret the scores. One challenging aspect, which we explain in Section 3.2, is the methodological choice to assign a default value to a Boolean indicator when no relevant data are found in the public disclosure of a firm. The default value is 0 when answering 'yes' to the question is positive from a sustainability point of view (e.g., "Does the firm conduct corporate social responsibility reporting?") but 1 when answering 'yes' is 'negative' (e.g., "Is the structure of the company board classified?").

2.2 Data Coverage

We use all the data available in the Refinitiv database, from 2010 to 2019. Our analysis of ESG data ends in 2019 for two reasons. First, there is a substantial time lag for a complete update of the database for a given financial year. At the time of our last download (March 2021), some data were already available for 2020, but for a substantial number of firms, the data were still missing. Second, as we describe in Section 4, we evaluate the financial performance of a portfolio built at the end of year N using stock returns in year $N+1$, so that the performance of the portfolio built with 2019 data is based on financial returns at the end of 2020. Our sample is defined as the complete set of firms included in the

[7]The Refinitiv database is not without drawbacks. One particular problem is the widespread and repeated changes to the historical ESG scores. As data in the Refinitiv database are subject to backward changes of up to five years as new information is disclosed, changes to the historical data are relatively likely, as put forward by Berg *et al.* (2020). These authors find that the median ESG score in the rewritten data is 18% lower than that in the initial data, with a deviation of 44% for the E score. In early 2021, Refinitiv opened a portal for firms to submit their ESG data, although this does not apply to our sample period, which runs up to the fiscal year ending in 2019 for the firms covered by Refinitiv. See https://www.refinitiv.com/en/media-center/press-releases/2021/january/refinitiv-makes-esg-company-scores-free-rolls-out-esg-voice-app.

Refinitiv database for which market capitalization is available in a given year. At the time of our last download, the database contained 10′142 firms that have been evaluated at some point in time.

Table 1 reports summary statistics on the number of firms for which both market capitalization and ESG scores are available. All the numbers in the table are relative to the firms covered by Refinitiv ESG in 2020: The total number of firms worldwide in the database is 10′142 as of 2020 – the latest available data at the time of analysis. The proportion of firms with available data on market capitalization in 2010 is equal to 74.8%. Among these 7′590 firms, 3′911 (51.5%) also have a Refinitiv ESG score in the respective year. The proportion of firms with an ESG score remains fairly stable in our sample until 2014, at slightly above the 50% of covered firms in 2020. Starting in 2015, the coverage improves steadily, with a maximum in 2019 of 82.3% of the firms in our sample. In general, other scores (3 pillars and 10 categories) have coverage that is essentially the same as the aggregate ESG score.

In 2019, the regional coverage in terms of market capitalization is the following: 39% of firms are from North America, 20% from Europe, 14% from the Pacific, and 23% from emerging countries. The table also reveals that among the firms in the Refinitiv database, the proportion of firms with ESG scores varies substantially across regions. On average, it is relatively low in North America at the beginning of the sample, below 50% until 2014. In emerging countries, the proportion of firms with an available ESG score has been above 50% since 2011. In Europe and the Pacific, the ESG score coverage is relatively complete over the full sample.

[Insert Table 1 here]

3 Disclosure of ESG Information

In this section, we analyze the disclosure of ESG data by firms in the Refinitiv database. As discussed in the introduction, there is no generally agreed-upon regime for disclosure regulation: The United States does not currently mandate ESG disclosure by publicly listed companies, while the European Union has already introduced a set of instruments to regulate ESG disclosure. We present the two issues raised by ESG data (missing values in numeric indicators and proportion of zeros in Boolean indicators) and describe the

implications for category scores.

3.1 Missing Values in Numeric Indicators

For numeric indicators, a score (based on the relative percentile ranking) is calculated only if the firm has reported this information, and a missing value is assigned when Refinitiv cannot find the information in publicly available reports. To compute the proportion of valid (or nonmissing) values for a given numeric indicator in a given year, we start by calculating the number of firms for which a given indicator is available in that year; then, for this given indicator, we identify the industries for which the indicator is material. Finally, we calculate the number of firms with valid data and divide that by the total number of firms for which the indicator is material.

Table 2 reports summary statistics on the proportion of valid values for numeric indicators for each pillar and category, which we interpret as a proxy for the disclosure policy of the firms. Overall, the proportion of valid values is relatively low, close to 40% on average for all the numeric indicators over the sample. In fact, there is a large gap between the E and S pillars (20% and 30%) and the G pillar (80%). Within a given pillar, this proportion is usually homogeneous. Indicators related to Emissions reduction have a proportion of valid values equal to 22% worldwide on average. These results reveal the lack of disclosure, particularly regarding measures taken by firms to protect the environment (approximately 10% of firms provide data on their Renewable energy use ratio) or to reduce greenhouse gas emissions (25% of firms report data on their CO_2 equivalent Scope 3 indirect emissions). The proportion of valid values is particularly low for the Innovation category in the E pillar and Product responsibility in the S pillar (close to 10% on average worldwide).

We also consider the results for the different regions. The proportion of valid values is on average higher for Europe than for other regions. The average proportion is 48% for Europe but 35% for North America and the Pacific and 38% for emerging countries. The difference is larger for the E and S pillars: 32% and 38% for the indicators in the E and S pillars in Europe versus 10% and 21% in North America. Numeric indicators in the E pillar have a proportion of valid values close to 20% on average in Pacific and emerging countries.

[Insert Table 2 here]

In Table 3, we show the proportion of valid values across the main industries. We do not report results for the real estate sector because the number of firms is insufficient. The results are, broadly speaking, similar across industries, although there are some notable differences. In particular, the disclosure of information related to the E pillar differs somewhat across industries: Financial firms have a low record of numeric indicators (below 8% on average), while firms in basic materials and utilities have the highest reporting standards (23% and 20%, respectively). Importantly, firms in sectors with the highest carbon intensity (energy, utilities, and basic materials) disclose relatively more than those in other sectors for the numeric indicators in the Emissions reduction category (close to 20% on average). We also observe the same pattern for the Resource use category. One possible explanation for why firms in energy-related sectors disclose more information about environmental issues is that these issues are very material for their investors. For the S pillar, financial firms have the lowest level of disclosure for the aggregate pillar and for the three categories (except Human rights). Disclosure is particularly low for the Product responsibility and Workforce categories. In contrast, firms in the consumer noncyclical sector report a high level of disclosure on the S pillar. Regarding the G pillar, almost 90% of the firms report on the indicators related to governance on average at the end of the sample. We do not observe any significant difference across sectors.

[Insert Table 3 here]

3.2 Proportion of Zeros at Category Level

Boolean indicators usually do not have missing values in the Refinitiv database. When the information concerning a Boolean question is not available, Refinitiv assigns a value equal to 0 (corresponding to a 'no') when answering the question with 'yes' would be considered positive and a value equal to 1 (corresponding to a 'yes') when answering the question with 'yes' would be considered negative. This strategy of penalizing firms is relatively recent. As stated on the Refinitiv website, "the previous ESG scoring methodology allocated a score of 0.5 to companies which did not report on metrics, essentially giving them the 'benefit of the doubt'. However, as this may disincentivize companies to report on their ESG performance, the enhanced methodology assigns a score of zero to companies who don't report on metrics relevant to the industry. This new approach encourages company

disclosure and transparency."[8] We note that this methodological choice will incentivize companies to improve their nonfinancial information disclosure if they actually improve their policy.

Because of the choice to assign a value of 0 to missing Boolean indicators, the proportion of valid values is equal to 100% for most categories. However, one difficulty with assigning the same value of 0 to both negative answers and missing values is that the evolution of the indicator over time may be difficult to interpret. For instance, if we consider the question "Does the company have a policy to avoid the use of forced labor?", we find that the number of firms with a value of 0 has decreased from 95% in 2010 to 52% in 2019. However, we cannot identify whether this change is due to better reporting or to a real improvement in firms' policy.

An implication of this approach is that the proportion of firms with a value of 0 is very large for some Boolean indicators. For instance, in 2019 the proportion of 0 is equal to 74% for the reporting on firm's environmental expenditures and to 95% for the reporting about the total individual compensation of all executives and board members. When we turn to category scores, this attribution approach may have a considerable impact because some categories (Human rights and CSR strategy) are exclusively based on Boolean indicators. Consequently, for these categories, a substantial proportion of firms report a score equal to 0.

In Table 4, we report the proportion of firms with a category score equal to 0 for the three pillars and the ten categories for the various regions. The distribution of scores is also displayed in Figures 1 to 3. As the table reveals, the problem is particularly acute for the E pillar because the pick of scores equal to 0 also contaminates the E pillar score itself. In all regions, we find that the scores for the three E categories suffer from a high frequency of 0. Even if there are numeric indicators for these three categories, they also have a large proportion of missing values, so that the score of the categories is often based on Boolean indicators only and therefore may obtain a score equal to 0. This problem is substantial for Innovation, as 56% of firms worldwide have an Innovation score equal to 0 (71% in North America) in 2019. Emissions and Resource use scores are also affected by this issue but to a lesser extent, with a proportion of 0 equal to 28% and 29% worldwide in 2019.

[8] See https://www.refinitiv.com/en/media-center/press-releases/2020/april/refinitiv-enhances-esg-scoring-methodology-to-reflect-sustainable-industry-developments-and-market-changes.

Regarding the S categories, we note a substantial fraction of firms with a score equal to 0 for the Human rights and Product responsibility categories. The Human rights score is equal to 0 for 42% of firms worldwide (59% in North America) in 2019. The Product responsibility score is equal to 0 for 10% of firms (20% in emerging countries). Finally, as the CSR strategy category is based on Boolean indicators only, it reports approximately 34% of scores equal to 0 worldwide (61% in North America) in 2019.

[Insert Table 4 and Figures 1 to 3 here]

Table 5 indicates that the proportion of firms with a category score equal to 0 is also heterogeneous across industries. First, there are large differences regarding the proportion of scores equal to 0. For the E pillar, the proportion of scores equal to 0 is as high as 53% for health care and as low as 17% for utilities in 2019. Heterogeneity in categories is even more pronounced: 90% of health care firms fail to have one positive answer for the indicators underlying the Innovation score. In contrast, only 18% of utilities obtain an Innovation score equal to 0. We also find the same gap between health care firms and utilities for the Human rights score (S pillar) and the CSR strategy score (G pillar). Financial firms also report a substantial proportion of scores equal to 0 for some categories. Sectoral heterogeneity is not related to the (lack of) materiality of some indicators, because category scores are already based on indicators that are material at the sector level.

Second, the table reveals that the proportion of firms with category scores equal to 0 tends to decrease over time for most sectors, but this evidence is not universal. Health care reports a higher proportion of scores equal to 0 in 2019 than in 2010 for all E and G categories. Technology also demonstrates an increase in the frequency of 0 for the E categories.

[Insert Table 5 here]

3.3 Scores at the Category Level

The high frequency of scores equal to 0 for some categories may introduce some distortion in the resulting average score across categories and therefore across pillars. For this reason, we now consider the temporal evolution of scores across categories. Table 6 confirms the large differences in the average score across ESG categories. Categories based on Boolean indicators only (Human rights in the S pillar and CSR strategy in the G pillar)

or on a small proportion of numeric indicators with a large proportion of missing values (Innovation in the E pillar) are associated with low average scores. On average, scores are lower for the E pillar than for the S and G pillars.

The table also reveals substantial heterogeneity across regions. Overall, European firms have higher scores, particularly for the E and S categories. Firms in North America and emerging countries have lower E scores.

On average, scores tend to improve over time. Pacific and emerging countries benefit from large increases in ESG scores, particularly because of the E and S pillars. In contrast, ESG scores do not improve in North America, mainly because of the decrease in the E pillar.

[Insert Table 6 here]

As reported in Table 7, we find interesting and somewhat counterintuitive results across sectors. The Emissions score is much higher for firms in energy, utilities, and basic materials (45%, 47%, and 44%, respectively, in 2019), although these industries emit large quantities of greenhouse gases. In contrast, firms in the health care, financial, and technology sectors have very low Emissions scores (20%, 30%, and 32%, respectively), although they have low carbon intensity. This difference has two sources. First, a large fraction of energy and utilities companies report on their emissions policy, for instance whether they have environmental partnerships, a policy to improve emissions reduction, or targets or objectives to be achieved on emissions reduction. Therefore, high carbon emissions can be at least partly compensated, at the Emissions score level, by policy measures taken by the company.[9] In contrast, firms in health care or financial firms often do not report information on these topics, partly because they view themselves as less concerned about these issues and therefore receive low Emissions scores, even if they generate low carbon emissions. As a result, the average E score ranges between 17.6 for health care firms and 44.8 for utilities in 2019. This contrast due to reporting biases is less pronounced for the S and G pillars. The average S score ranges between 42.9 and 46.3 in 2019 across sectors. The average G score is between 40.6 and 53.5.

[Insert Table 7 here]

[9]This logic is similar to the "best-in-class" approach, in which firms with best practices in their sector can benefit from relatively high industry-adjusted scores.

In the Online Appendix, we also assess how the size of the firms affects their disclosure policy. We find that large firms (firms in the highest quartile of the market cap) usually tend to disclose more information about their activities; therefore, the proportion of numeric indicators with valid values is higher for large firms. Overall, for numeric indicators, this proportion is 31% for the lowest quartile and 45% for the highest quartile. In addition, the proportion of firms with category scores equal to 0 represents approximately 50% of small firms, whereas it represents approximately 10% of large firms. This heterogeneity in the proportion of missing numeric indicators and Boolean indicators equal to 0 is reflected in large differences in pillar and category scores across firms of different sizes. For small firms, the average ESG score is close to 30% over the sample. In contrast, for large firms, the average ESG score increases from 49% in 2010 to 58% in 2019. These results are consistent with the empirical evidence that large firms spend considerable resources reporting on ESG matters (Drempetic *et al.*, 2020).

4 ESG Screening at the Category Level

Our analysis identifies two issues with the implementation of an ESG-based screening investment strategy at the category level. First, the proportion of scores equal to 0 is substantial for 6 of 10 categories. Setting a low value for the screening threshold (for instance, excluding 1% or 5% of the firms with the lowest scores and reinvesting in the remaining firms proportionately) would result for these categories in the exclusion of some firms with a score equal to 0 while other firms with a score equal to 0 that would be kept in the portfolio. Therefore, screening at the category level is well suited for relatively high screening levels (e.g., 25% or 33%) as illustrated below.[10]

Second, given the large heterogeneity of scores across regions or sectors, the screening process will imply significant regional and sectoral biases in the ESG portfolio relative to market exposures. Such biases would be an issue for investors seeking to hold an otherwise passive portfolio. To address this issue, we proceed as follows. We assume a benchmark portfolio that reproduces the structure of the targeted market and provides representative weights for the companies. We construct an ESG portfolio based on excluding firms with the worst scores associated with a given ESG category. In the first strategy, the proceeds

[10] An ESG screening based on indicators instead of categories could not be designed because the vast majority of indicators (123 of 186) are Boolean indicators, which are not suited for a screening approach.

of the excluded firms are reinvested proportionately in the remaining firms. As this approach generates large regional and sectoral biases, we consider a second strategy, in which the screening is performed at the region-sector level: The proceeds of the exclusion of low-score firms in a given region-sector are reinvested in high-score firms in the same region-sector. This strategy is akin to what is often called a "best-in-class" approach, whereby investment managers select the firms with the highest scores within their sector and often also their region.

As a large worldwide stock market index, we use the MSCI ACWI, which covers developed and emerging markets. The list of constituents and the corresponding market weights, which we use to define the reference weights for regions and sectors, are available for this index. Hereafter, we consider the subset of firms in the Refinitiv database that are also in the MSCI ACWI.

4.1 Global Screening

The allocation exercise is performed over the period from 2010 to 2019. The screening is based on the scores available at the end of year N. For a screening threshold of, for example, 25%, we identify all firms with the worst scores until their cumulative market cap represents 25% of the market cap of the benchmark portfolio. The proceeds of the exclusion are reinvested in the remaining firms in proportion to their market weight. Stock market returns of the subsequent year are used to compute the financial performance of the portfolio, so a portfolio built at the end of year N is evaluated at the end of year $N+1$. We consider investors with a preference for some particular dimension of the ESG pillars (for instance, for Emissions reduction or Human rights). We may also imagine investors interested in combining two or more categories.

Table 8 reports summary statistics for screening portfolios based on the 2010–2019 sample. The MSCI ACWI represents the market index, including all firms, even those with no ESG score. The row labeled "Benchmark" represents the portfolio based on MSCI ACWI constituents for which Refinitiv ESG scores are available. As the table reveals, for the world index, we lose only 2.4% of the market cap on average due to the lack of Refinitiv scores among firms within the MSCI ACWI.

[Insert Table 8 here]

The first two columns represent the proportion of firms and the proportion of the market value with scores equal to zero, while the next two columns indicate for a given threshold how many firms are actually excluded and what fraction of the market cap is excluded. The comparison of these columns allows us to evaluate the impact of zero scores on the composition of the screening portfolio. First, we note that, as low-score firms also tend to have a low market cap, we in fact exclude a rather large fraction of small firms. For the 10% screening criterion (Panel A), we exclude 9.9% of the market cap but 26.2% of the firms with the lowest ESG scores. Similarly, for the 25% screening (Panel B), we exclude 24.7% of the market value but 50.4% of the firms. These proportions are equal to 33% and 60.2%, respectively, for the 33% screening (Panel C).

Second, we turn to categories with a large fraction of firms with a score equal to zero. For the Innovation category, we find that 40.3% of firms in the MSCI index (26.1% of the market cap) have a score equal to zero. Consequently, the lowest screening threshold that we can apply to build a screening portfolio is the 26.1% quantile (to avoid arbitrary selection of firms with a score equal to zero). Similarly, for the Human rights category, we cannot exclude less than 21.9% of the market cap (40% of the firms). Consequently, for these two categories, the impact of the screening process is much larger than for other scores because it actually corresponds to an approximately 25% screening. For the CSR strategy category, the lower bound for screening is 9.5% of the market cap. These results clearly illustrate the impact of the scoring methodology on the screening strategy. For these categories, because of the large proportion of firms with scores equal to zero, a screening strategy with a low screening threshold cannot be implemented.

The gain on the score (difference between the portfolio score and the benchmark score) is substantial, usually between 4 and 7 points for the 10% threshold. In relative terms (gain divided by benchmark score), the gain is between 6% and 10%. One factor that limits the score gain is that the portfolio is market cap-weighted. As mentioned above, large firms tend to have higher scores than small firms.[11] Therefore, the benchmark portfolio is already tilted in favor of firms with relatively high scores.

The score gain is the highest for the E category. For the same 10% proportion of excluded firms, the Resource use and Emissions scores deliver the highest score gains, above 7 pp. We note, however, that the gain on the E pillar score is much smaller than

[11] On average, the 25% smallest firms in the MSCI ACWI have an overall ESG score equal to 42.1, whereas the 25% largest firms have an overall ESG score equal to 58.6.

the gain in the E categories. The reason why the aggregate gain is well below the average of the gain on the categories is that category scores are summed at the firm level first, so that the large proportion of zeros observed for the Innovation score has a limited impact on the distribution of the E pillar score. As category scores are not perfectly correlated across firms, it is more difficult to improve the E pillar score than its components separately. We observe the same result for the other S and G pillar scores and the aggregate ESG score.

Figures 4–7 represent the temporal evolution of the E, S, and G scores of the screening portfolios based on various levels of screening. These figures demonstrate that, for the Innovation and Human rights categories, the increase in the score relative to the benchmark is very large for the 10% screening threshold because much more than 10% of firms are in fact excluded. For the other categories, the gains relative to the benchmark are similar across categories. The highest and lowest gains are obtained for the Product responsibility score and Workforce score, corresponding to gains equal to 14.2 and 11.5, respectively, for the 25% threshold. Importantly, with the 33% screening threshold, methodological issues related to the missing values in numeric and Boolean indicators no longer affect the portfolio construction. The gains in the score relative to the benchmark are substantial for all the categories, from 14.1 pp for Workforce to 21.4 pp for Human rights, while the overall ESG score only increases by 11.2 pp.

[Insert Figures 4 to 7 here]

As discussed in the literature review, the performance of ESG portfolios should be adjusted for the ESG risk factor (Pástor *et al.*, 2021b). Given the discrepancy between current ESG ratings produced by data providers, properly adjusting for the ESG risk factor is beyond the scope of our paper. For this reason, we use the Sharpe ratio as a measure of the risk-adjusted financial performance of the screening portfolios. We focus on the 33% threshold, as the screening strategy can be implemented for all the categories.

The last columns of Table 8 provide statistics on the financial performance of the screening portfolios. The Sharpe ratio remains in the same ballpark as that of the MSCI ACWI. The CSR strategy score generates the lowest Sharpe ratio (0.63 versus 0.69 for the benchmark), while the Community score improves the Sharpe ratio to 0.73. We also report the tracking error relative to the MSCI ACWI. As the table reveals, this is mostly due to the fraction of firms in the index with no ESG scores. On average, the annual tracking error of the benchmark (including all firms in the MSCI ACWI with ESG scores)

is equal to 0.9% relative to the MSCI ACWI. Even with the 25% and 33% thresholds, the tracking error of the screening portfolios is only increased to 1.4% and 1.5%, respectively.

One reason why the Sharpe ratio of screening portfolios differs from the benchmark Sharpe ratio may be that the screening process implies some changes in the regional and sectoral exposures. Average scores suggest that firms in North America are likely to be underweighted in favor of European firms and that health care firms are likely to be underweighted in favor of financial firms or utilities for almost all categories. As an illustration of the impact of screening on risk exposures, we consider the screening based on the E score with the 33% threshold. On average over the sample, the screening would imply an overweighting of 6.4 pp (from 23.7% to 28.7%) of European firms and an underweighting of 3.5 pp (from 11% to 7.5%) of firms in emerging countries. Similarly, the screening would imply an overweighting of 1.9 pp (from 11.7% to 13.6%) of financial firms and an underweighting of 1.7 pp (from 17.8% to 16.1%) of health care firms.

Such an impact on regional and sectoral exposures would be an issue for investors seeking to improve the ESG quality of their portfolio without altering their risk exposures. We address this issue in the next section.[12]

4.2 Screening at the Region-Sector Level

We now consider the same screening strategies but while maintaining the same sectoral and regional exposures as in the MSCI ACWI. Therefore, considering the 25% threshold, we exclude in each region-sector the firms with the lowest scores until their cumulative market cap represents 25% of the market cap of the region-sector in the benchmark portfolio. The proceeds are reinvested in the firms with the highest scores in the same region-sector until their cumulative market cap represents 25% of the market cap of the region-sector. This approach is therefore akin to a best-in-class strategy, in which investors reweight their portfolio with best-in-class instead of worst-in-class firms. For this reason, a region-sector approach is more likely to affect on the cost of financing of the reweighted firms: The cost of financing of excluded firms will tend to increase, while the cost of

[12]It is likely that ESG portfolios are exposed to other risk factors, such as size or value factors. Maintaining the same exposures to these risk factors as in the benchmark requires an optimization process in each period because the size and value characteristics of firms vary over time, which is not the case for regions and sectors (see Alessandrini and Jondeau, 2021).

financing of overweighted firms will tend to decrease.[13] In Tables 9 and 10, corresponding to the 25% and 33% thresholds, respectively, we consider the cases where reallocation is performed at the regional level, the sectoral level, and the region-sector level.

Starting with the 25% threshold (Table 9), we find that imposing regional exposures (Panel A) results in scores that are slightly lower than those obtained without exposure restrictions. The gain of reallocation is reduced by 0.8 pp (from 12.8 pp to 12 pp). With sectoral reallocation (Panel B), the impact of reallocation on the ESG scores is marginal. Finally, when reallocation is performed at the region-sector level, the gain in the overall ESG score is equal to 10.3 pp, while it is equal to 14, 11.3, and 11.7 pp for the E, S, and G scores, respectively. For ESG categories, the gain increases on average to 16.7 pp, so that gains are approximately 2.5 pp higher than with the global screening. This result suggests that reinvesting in the best-in-class firms at the region-sector level is more effective than reinvesting in all remaining firms proportionately.

For the 33% threshold (Table 10), even with constrained sectoral and regional exposures, reallocation strategies allow investors to benefit from substantial increases in ESG scores.[14] Gains are equal to 16.3, 13.3, and 14 pp for the E, S, and G pillars, respectively. For ESG categories, the gain increases to 19 pp.

Targeting some specific category usually results in larger gains than targeting ESG pillars, for the same reallocation threshold. For instance, in targeting the Emissions score, the 33% threshold would allow investors to improve their score from 66.9 to 84.4 in an otherwise passive portfolio. These findings could be particularly relevant for investors who wish to target certain ESG objectives, such as a lower portfolio carbon footprint compared to the benchmark. Without relying on the overall ESG score, these investors could exclude firms with the lowest Emissions scores (thus focusing on the particular underlying ESG category) and achieve similar results as the benchmark in terms of financial performance.

The Sharpe ratios of the reallocation portfolios based on category scores are in the same ballpark as the Sharpe ratio of the benchmark (0.68 vs. 0.69). In addition, the

[13] As pointed out by Pástor *et al.* (2021a) and Pástor *et al.* (2021b), in equilibrium investors with ESG preferences should expect lower returns, although realized returns may be higher in the short and medium terms because of the high demand pressure.

[14] We note that, for the 33% threshold, the proportion of market capitalization excluded is often slightly larger than 67%. This is because we reallocate firms at the region-sector level. Therefore, for each region-sector, we reallocate at most 33% of the market capitalization. As we do not have perfect granularity in the market cap, we often reduce slightly less than 33% of the market cap, which results in the discrepancy observed in the table.

tracking error relative benchmark is below 1.4% per year on average, which includes 0.9% due to the constituents of the MSCI ACWI with no ESG score.

[Insert Tables 9 and 10 here]

In Figure 8, we display how the increase in the portfolio score is affected by the reallocation threshold. We vary the reallocation threshold from 5% to 50% and consider the various pillars and categories. The Innovation, Human rights, Product responsibility, and CSR strategy categories achieve substantial increases in the category scores, even with a modest reallocation threshold, because all firms with zero scores are excluded simultaneously, resulting in large-than-expected reallocation. For other categories, the score gain increases steadily, up to 20 pp for the 50% threshold.

Figure 9 reports the Sharpe ratio of the various portfolios, and the Sharpe ratio of the MSCI ACWI (horizontal line). The negative impact of reallocation is marginal for the E categories and moderate for the S categories (for thresholds up to 40%). For the G pillar, only the Sharpe ratio of the CSR strategy category decreases significantly.

Finally, Figure 10 indicates that the annual tracking error usually increases as reallocation becomes more severe. However, it remains below 1.8%, even for reallocation thresholds as high as 50%, while the tracking error of the portfolio with no reallocation (but with only firms with an ESG score) is equal to 0.9%. Consequently, a reallocation strategy based on a rather high threshold (such as the 33% threshold) can be implemented at a relatively low financial cost.

Constraining regional and sectoral exposures not only results in lower gains in the ESG score but also allows investors to hold a portfolio that is otherwise passive, as it is not exposed to regional and sectoral risks relative to the market portfolio. This result suggests that investing in a portfolio with a higher ESG category score, with minimal regional and sectoral risk exposures relative to the market portfolio, can be attained at almost no cost in terms of financial performance.

[Insert Figures 8 to 10 here]

5 Conclusion

While ESG investing has gained popularity in recent decades, many institutional investors struggle with the inherent limitations of ESG scores. These limitations include the lack

of transparency in the methodologies, the wide divergence among ESG ratings, and potential conflicts of interest in the ESG rating business (Kotsantonis and Serafeim, 2019, Berg *et al.*, 2019, IOSCO, 2021). Moreover, devising investment strategies based on an amalgamation of three fundamentally different topics underpinning ESG investing has been a practical hurdle, especially given the potential for weak scores in one pillar to be offset by strong scores in another pillar. We use the Refinitiv ESG database, which allows us to deconstruct ESG scores and analyze indicators in depth. On this basis, we address two important questions regarding ESG investment in equity markets. First, we investigate the characteristics of the various categories of ESG factors and find that they do not all contain the same quality of underlying information – and hence may not have the same desired positive impact from an investment perspective. Specifically, several category scores are essentially based on Boolean indicators. Given the methodological choice of Refinitiv to assign a negative score when firms fail to disclose 'yes' or 'no' information, these categories suffer from a high proportion of scores equal to 0, which makes it difficult to differentiate among firms. Consequently, an investment strategy based on excluding firms with low category scores may not be implementable. For other categories, implementing an exclusion strategy is feasible and allows investors to substantially improve the score of their portfolio after exclusion.

Second, regarding the financial characteristics of the screening portfolios, we find that they do not suffer from a lower risk-adjusted performance compared to a wide stock market benchmark. However, the screening process also results in significant regional and sectoral biases relative to this benchmark. Such biases may be undesirable for investors seeking to hold an otherwise passive portfolio. We demonstrate that a best-in-class approach imposing the same regional and sectoral exposures as the benchmark slightly increases the gain in the targeted score with no material impact on the risk-adjusted performance and minimal increase in the tracking error of the portfolio.

In addition, shifting focus from aggregate ESG pillar scores and ratings to more granular characteristics (to the extent that they may be available from the various ESG data vendors) has three key nonfinancial advantages. First, a focus on specific categories would enable investors to overcome the "aggregate confusion" created by consolidated ESG scores or ratings and directly focus on factors that are most relevant to their investment mandates. For example, an investor seeking to protect the environment and universal human

values could target themes such as emissions reduction or human rights. Second, focusing on specific themes would help them better track the sustainability performance trajectory of their investments vis-à-vis their stated sustainable investment objectives. This would also help initiate divestments or reweight investments within the portfolio when there is a notable development. Finally, over time, the focus on themes would also enable investors to develop their own ESG assessment models using actual and observed third-party vendor data, thereby overcoming vendor-specific concerns.

References

Alessandrini, F. and Jondeau, E. (2021). Optimal strategies for ESG portfolios. *Journal of Portfolio Management*, **47**(6), 114–138.

Amel-Zadeh, A. and Serafeim, G. (2021). Why and how investors use ESG information: Evidence from a global survey. *Financial Analysts Journal*, **74**(3), 87–103.

Berg, F., Koelbel, J., and Rigobon, R. (2019). Aggregate confusion: The divergence of ESG ratings. MIT Sloan School, Working Paper No. 5822-19.

Berg, F., Fabisik, K., and Sautner, Z. (2020). Rewriting history II: The (un)predictable past of ESG ratings. ECGI Finance Working Paper No. 708/2020.

Billio, M., Costola, M., Histova, I., Latino, C., and Pelizzon, L. (2021). Inside the ESG ratings: (dis)agreement and performance. *Corporate Social Responsibility and Environmental Management*, **28**(5), 1426–1445.

Bloomberg (2017). GPIF selected ESG indices. Government Pension Investment Fund of Japan. Available at www.gpif.go.jp/en/investment/pdf/ESG_indices_selected.pdf.

Bloomberg (2021). ESG assets rising to $50 trillion will reshape $140.5 trillion of global AUM by 2025, finds Bloomberg Intelligence. Bloomberg, 21 July. Available at https://www.bloomberg.com/company/press/esg-assets-rising-to-50-trillion-will-reshape-140-5-trillion-of-global-aum-by-2025-finds-bloomberg-intelligence/.

Demers, E., Hendrikse, J., Joos, P., and Lev, B. (2021). ESG did not immunize stocks during the COVID-19 crisis, but investments in intangible assets did. *Journal of Business Finance and Accounting*, **48(3–4)**, 433–462.

Drempetic, S., Klein, C., and Zwergel, B. (2020). The influence of firm size on the ESG score: Corporate sustainability ratings under review. *Journal of Business Ethics*, **167**, 333–360.

Esty, D., Cort, T., Strauss, D., Wyatt, K., and Yeargain, T. (2020). Toward enhanced sustainability disclosure: Identifying obstacles to broader and more actionable ESG reporting. Yale Initiative on Sustainable Finance White Paper.

Fabozzi, F. J., Ma, K. C., and Oliphant, B. J. (2008). Sin stock returns. *Journal of Portfolio Management*, **35**(1), 82–94.

Friede, G., Busch, T., and Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance and Investment*, **5**(4), 210–233.

Garel, A. and Petit-Romec, A. (2021). Investors rewards to environmental responsibility: Evidence from the COVID-19 crisis. *Journal of Corporate Finance*, **68**, 101948.

Gibson, R., Krueger, P., Riand, N., and Schmidt, P. S. (2021). ESG rating disagreement and stock returns. *Financial Analyst Journal*, **77(4)**, 104–127.

GSIA, Global Sustainable Investment Alliance (2020). Global sustainable investment review. Biennial Report.

Hong, H. and Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, **93**(1), 15–36.

Ioannou, I. and Serafeim, G. (2017). The consequences of mandatory corporate sustainability reporting. Harvard Business School Working Papers 11-100, Harvard Business School, available at SSRN: https://ssrn.com/abstract=1799589.

IOSCO, International Organization of Securities Commissions (2021). Environmental, Social and Governance (ESG) ratings and data products providers. The Board of the International Organization of Securities Commissions, Final Report.

Jouvenot, V. and Krueger, P. (2020). Mandatory corporate carbon disclosure: Evidence from a natural experiment. University of Geneva Working Paper, available at SSRN: https://ssrn.com/abstract=3434490.

Khan, M., Serafeim, G., and Yoon, A. (2016). Corporate sustainability: First evidence on materiality. *Accounting Review*, **91**(6), 1697–1724.

Kotsantonis, S. and Serafeim, G. (2019). Four things no one will tell you about ESG data. *Journal of Applied Corporate Finance*, **31**(2), 50–58.

Lins, K. V., Servaes, H., and Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *Journal of Finance*, **72**(4), 1785–1824.

Mésonnier, J.-S. and Nguyen, B. (2020). Showing off cleaner hands: Mandatory climate-related disclosure by financial institutions and the financing of fossil energy. Banque de France Working Paper, available at SSRN: https://ssrn.com/abstract=3733781.

Pástor, L., Stambaugh, R. F., and Taylor, L. A. (2021a). Sustainable investing in equilibrium. *Journal of Financial Economics*, **142(2)**, 550–571.

Pástor, L., Stambaugh, R. F., and Taylor, L. A. (2021b). Dissecting green returns. University of Chicago, Becker Friedman Institute for Economics Working Paper No. 2021-70, available at SSRN: https://ssrn.com/abstract=3869822.

Pedersen, L. H., Fitzgibbons, S., and Pomorski, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics*, **142(2)**, 572–597.

Refinitiv (2021). Environmental, Social, and Governance (ESG) scores from Refinitiv. Technical report.

Sahin, O., Bax, K., Paterlini, S., and Czado, C. (2021). ESGM: ESG scores and the Missing pillar. Available at SSRN: https://ssrn.com/abstract=3890696.

Scatigna, M., Xia, D., Zabai, A., and Zulaica, O. (2021). Achievements and challenges in ESG markets. BIS Quarterly Review, December 2021.

Serafeim, G. and Yoon, A. (2021). Stock price reactions to ESG news: The role of ESG ratings and disagreement. Forthcoming in Review of Accounting Studies.

Table 1. Global Coverage of Refinitiv Database

	Firms with valid market cap		Firms with valid ESG score		Firms with valid market cap		Firms with valid ESG score	
	Number	(%)	Number	(%)	Number	(%)	Number	(%)
	World				**North America**			
2010	7590	74.8	3911	51.5	2762	69.6	1264	45.8
2011	7739	76.3	4017	51.9	2818	71.0	1250	44.4
2012	7979	78.7	4096	51.3	2914	73.4	1232	42.3
2013	8236	81.2	4212	51.1	3053	76.9	1246	40.8
2014	8640	85.2	4342	50.3	3231	81.4	1256	38.9
2015	8959	88.3	5231	58.4	3383	85.2	1908	56.4
2016	9118	89.9	6110	67.0	3451	86.9	2622	76.0
2017	9381	92.5	6696	71.4	3576	90.1	2797	78.2
2018	9761	96.2	7433	76.1	3780	95.2	2875	76.1
2019	9944	98.0	8182	82.3	3884	97.9	3249	83.7
	Europe				**Pacific**			
2010	1538	76.6	918	59.7	1159	82.8	832	71.8
2011	1552	77.3	929	59.9	1184	84.6	851	71.9
2012	1586	79.0	930	58.6	1196	85.5	863	72.2
2013	1627	81.1	937	57.6	1218	87.1	923	75.8
2014	1715	85.5	957	55.8	1277	91.3	958	75.0
2015	1792	89.3	1059	59.1	1306	93.4	995	76.2
2016	1822	90.8	1072	58.8	1339	95.7	1029	76.8
2017	1876	93.5	1152	61.4	1357	97.0	1053	77.6
2018	1941	96.7	1599	82.4	1370	97.9	1068	78.0
2019	1972	98.3	1630	82.7	1380	98.6	1117	80.9
	Emerging countries							
2010	1807	79.1	840	46.5				
2011	1846	80.8	924	50.1				
2012	1926	84.3	1004	52.1				
2013	1967	86.1	1040	52.9				
2014	2008	87.9	1105	55.0				
2015	2056	90.0	1158	56.3				
2016	2076	90.9	1253	60.4				
2017	2130	93.2	1542	72.4				
2018	2212	96.8	1718	77.7				
2019	2246	98.3	2003	89.2				

Note: This table reports the global coverage of the Refinitiv database: it reports the number of firms with a valid market capitalization and its proportion relative to the total number of firms in the database in the given region (as of 2020); it also reports the number of firms with a valid ESG score and its proportion relative to the total number of firms with a valid market capitalization in the given region. The sample covers the period from 2010 to 2019.

Table 2. Proportion of Valid Values in Numeric Indicators by Region

Category	Nb of numeric items	2010	2015	2019	2010	2015	2019	2010	2015	2019
		World			**North America**			**Europe**		
ESG	63	37.2	39.1	39.7	35.1	34.8	34.7	47.1	49.3	48.4
Environment	27	18.9	19.6	20.8	13.3	12.7	13.8	33.2	34.0	32.6
Emissions	14	21.8	22.4	21.3	13.9	12.8	11.7	37.4	38.0	32.9
Innovation	7	12.1	11.1	16.1	13.6	11.9	19.7	24.6	24.7	26.9
Resource use	6	20.3	23.2	24.8	11.5	13.2	13.5	31.9	33.9	37.4
Social	19	24.1	29.2	30.5	21.4	23.1	22.0	32.8	40.0	41.1
Community	1	39.3	37.0	35.2	27.9	22.1	17.4	55.0	45.9	34.4
Human rights	0	–	–	–	–	–	–	–	–	–
Product resp.	2	4.9	13.1	12.9	2.5	5.9	5.0	8.9	21.4	19.9
Workforce	16	25.5	30.7	32.4	23.3	25.3	24.4	34.5	42.0	44.1
Governance	17	80.9	81.2	80.1	87.9	85.7	85.0	86.9	85.9	83.7
CSR strategy	0	–	–	–	–	–	–	–	–	–
Management	14	80.8	80.0	78.6	87.1	83.9	82.9	85.9	84.7	82.0
Shareholders	3	81.4	86.9	87.3	91.7	94.2	94.9	91.9	91.8	91.7
		Pacific			**Emerging countries**					
ESG	63	32.3	35.9	39.8	33.9	38.8	39.4			
Environment	27	19.0	19.8	25.4	16.6	22.1	23.7			
Emissions	14	20.1	20.5	24.7	18.8	24.7	23.9			
Innovation	7	16.7	16.2	23.7	6.0	7.4	17.3			
Resource use	6	18.3	21.3	28.6	20.3	28.5	28.5			
Social	19	16.4	25.0	29.5	25.9	32.9	35.7			
Community	1	23.1	30.9	36.5	55.3	59.6	64.8			
Human rights	0	–	–	–	–	–	–			
Product resp.	2	1.3	11.7	13.5	7.5	17.7	17.5			
Workforce	16	17.9	26.3	31.1	26.3	33.2	36.2			
Governance	17	73.0	75.7	76.4	71.2	74.2	71.2			
CSR strategy	0	–	–	–	–	–	–			
Management	14	73.3	74.3	74.6	72.7	74.0	70.8			
Shareholders	3	71.6	82.2	84.8	64.4	75.2	73.1			

Note: This table reports the proportion of valid values in numeric indicators for each region in the Refinitiv database. The sample covers the period from 2010 to 2019.

Table 3. Proportion of Valid Values in Numeric Indicators by Sectors

Category	Nb of numeric items	2010	2015	2019	2010	2015	2019	2010	2015	2019
		Energy			**Basic Materials**			**Industrials**		
ESG	63	35.2	37.8	39.6	37.5	41.3	42.1	33.6	35.9	36.6
Environment	27	18.1	21.7	26.3	24.8	27.9	30.2	17.9	19.1	20.5
Emissions	14	19.2	22.4	25.5	24.9	28.4	29.8	18.8	19.7	20.0
Innovation	7	–	–	–	21.4	19.0	18.5	12.4	10.0	6.8
Resource use	6	15.4	19.9	28.2	25.8	29.8	35.0	16.9	19.6	24.4
Social	19	23.9	30.5	32.8	25.1	33.0	34.8	23.4	29.1	31.0
Community	1	32.7	41.0	39.6	39.8	45.0	50.1	39.5	33.8	34.7
Human rights	0	–	–	–	–	–	–	–	–	–
Product resp.	2	3.2	8.1	10.0	2.6	13.4	14.2	5.1	11.9	11.3
Workforce	16	26.0	32.7	35.2	27.0	34.7	36.4	24.7	31.0	33.2
Governance	17	83.8	83.1	81.8	80.4	81.6	79.8	79.6	80.1	79.3
CSR strategy	0	–	–	–	–	–	–	–	–	–
Management	14	83.8	82.0	80.2	81.0	80.8	78.4	79.0	78.5	77.5
Shareholders	3	83.5	88.0	89.2	77.3	85.6	86.2	82.2	87.5	87.9
		Consumer Cyclical			**Consumer Non Cycl.**			**Financials**		
ESG	63	33.4	35.1	35.8	34.3	35.8	37.0	32.3	33.6	34.0
Environment	27	16.7	16.8	17.4	21.3	21.9	25.2	19.0	21.2	19.6
Emissions	14	18.7	18.6	19.0	23.5	22.5	24.6	21.2	23.3	19.7
Innovation	7	12.5	10.4	5.0	7.5	11.3	10.0	–	–	–
Resource use	6	14.8	16.9	21.2	19.7	22.9	29.5	16.3	18.5	19.4
Social	19	22.1	26.9	29.5	24.0	29.1	32.6	27.5	31.0	32.3
Community	1	38.7	34.4	31.1	42.9	37.1	42.8	45.0	42.5	38.7
Human rights	0	–	–	–	–	–	–	–	–	–
Product resp.	2	5.2	12.7	13.8	1.9	10.1	11.3	11.6	12.5	10.0
Workforce	16	23.1	28.2	31.3	25.5	31.0	34.6	27.3	31.5	33.3
Governance	17	80.5	81.0	80.3	80.3	79.6	77.2	81.0	81.3	80.6
CSR strategy	0	–	–	–	–	–	–	–	–	–
Management	14	80.2	79.7	78.6	80.5	78.8	75.8	80.8	80.1	79.3
Shareholders	3	82.1	87.0	88.4	79.4	83.4	83.5	82.0	86.9	86.4
		Health Care			**Technology**			**Utilities**		
ESG	63	35.0	35.5	33.8	33.0	35.3	34.8	39.3	40.1	43.3
Environment	27	23.2	20.1	15.7	18.6	20.7	17.8	23.5	23.7	30.4
Emissions	14	25.4	20.2	14.6	20.3	21.0	17.6	28.6	28.2	32.4
Innovation	7	–	–	–	9.2	10.1	5.7	10.5	10.1	21.6
Resource use	6	18.3	19.9	18.1	16.8	22.1	20.8	21.7	23.8	32.6
Social	19	22.2	26.0	25.0	21.2	25.8	27.1	31.9	36.1	42.1
Community	1	30.8	27.4	19.0	26.9	22.1	25.1	60.3	56.8	60.3
Human rights	0	–	–	–	–	–	–	–	–	–
Product resp.	2	3.1	11.8	11.4	3.1	18.7	16.2	10.1	15.1	14.8
Workforce	16	24.1	27.6	27.0	23.1	26.9	28.6	32.9	37.4	44.4
Governance	17	83.0	83.8	82.5	79.9	81.5	80.9	80.7	78.9	75.7
CSR strategy	0	–	–	–	–	–	–	–	–	–
Management	14	82.8	82.3	81.0	79.9	80.1	79.1	80.7	78.1	74.6
Shareholders	3	84.1	90.6	89.5	80.0	87.8	89.1	81.0	82.6	81.3

Note: This table reports the proportion of valid values in numeric indicators for each sector in the Refinitiv database. The sample covers the period from 2010 to 2019.

Table 4. Proportion of Scores with Zero Value by Region

Category	2010	2015	2019	2010	2015	2019	2010	2015	2019
	World			**North America**			**Europe**		
Environment	17.9	21.0	21.2	24.4	32.9	37.5	5.1	5.7	6.0
Emissions	26.6	29.8	28.2	34.6	43.9	46.4	9.8	9.1	9.9
Innovation	56.1	56.6	56.5	62.3	66.0	71.0	40.5	40.2	39.2
Resource use	27.6	29.1	29.4	37.0	43.0	49.8	9.9	11.9	11.2
Social	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Community	0.6	0.0	0.0	0.1	0.0	0.0	1.1	0.0	0.0
Human rights	66.6	55.4	41.9	75.4	64.5	59.4	41.7	30.2	13.7
Product resp.	24.8	17.2	10.0	19.0	11.6	5.9	20.5	13.6	6.6
Workforce	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Governance	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CSR Strategy	35.1	36.4	33.9	45.4	57.0	61.1	16.0	12.7	10.0
Management	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Shareholders	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Pacific			**Emerging countries**					
Environment	19.5	19.1	14.1	19.2	15.1	11.0			
Emissions	29.0	28.8	20.7	29.4	24.3	17.9			
Innovation	55.7	55.0	48.3	62.6	55.0	50.7			
Resource use	29.0	27.3	19.1	30.0	21.3	16.4			
Social	0.0	0.0	0.0	0.0	0.0	0.0			
Community	1.2	0.0	0.0	0.5	0.1	0.0			
Human rights	73.9	64.8	38.9	72.3	53.0	37.6			
Product resp.	33.9	22.9	9.3	28.2	23.3	19.7			
Workforce	0.0	0.0	0.0	0.0	0.0	0.0			
Governance	0.0	0.0	0.0	0.0	0.0	0.0			
CSR Strategy	40.1	34.0	21.0	34.7	24.8	16.4			
Management	0.0	0.0	0.0	0.0	0.0	0.0			
Shareholders	0.0	0.0	0.0	0.0	0.0	0.0			

Note: This table reports the proportion of scores with zero values for each region in the Refinitiv database. The sample covers the period from 2010 to 2019.

Table 5. Proportion of Scores with Zero Value by Sector

Category	2010	2015	2019	2010	2015	2019	2010	2015	2019
	Energy			**Basic Materials**			**Industrials**		
Environment	14.7	12.1	6.0	10.7	7.6	7.0	9.9	13.5	11.6
Emissions	22.8	20.2	10.7	20.0	14.1	12.1	19.5	24.4	17.5
Innovation	76.4	75.1	68.7	61.0	56.9	49.7	46.9	44.4	44.0
Resource use	29.6	26.4	19.7	19.6	15.7	14.6	21.8	21.3	18.3
Social	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Community	0.3	0.0	0.0	1.2	0.0	0.0	1.5	0.0	0.0
Human rights	70.4	60.9	39.6	60.8	45.6	29.1	64.0	46.2	29.5
Product resp.	34.4	23.2	13.4	31.8	21.3	13.7	15.8	12.7	7.9
Workforce	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Governance	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CSR strategy	32.7	30.2	21.7	25.1	19.5	13.7	28.9	31.4	26.1
Management	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Shareholders	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Consumer Cyclical			**Consumer Non Cycl.**			**Financials**		
Environment	20.6	20.7	14.5	11.5	14.6	9.4	28.9	34.7	33.6
Emissions	31.8	29.7	25.0	19.0	20.5	14.8	35.3	41.9	40.1
Innovation	58.7	58.1	55.6	44.4	49.6	47.7	59.4	60.3	58.2
Resource use	27.5	27.8	23.9	18.0	18.9	14.4	40.6	43.8	45.1
Social	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Community	0.5	0.1	0.0	0.0	0.0	0.0	0.2	0.0	0.0
Human rights	60.0	49.9	32.8	56.3	41.9	30.4	78.4	71.8	57.3
Product resp.	27.6	18.6	8.4	10.5	8.7	7.5	30.0	20.5	10.0
Workforce	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Governance	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CSR strategy	40.7	38.3	33.9	26.0	24.8	21.4	42.3	43.4	41.9
Management	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Shareholders	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Health Care			**Technology**			**Utilities**		
Environment	34.6	41.5	53.1	17.5	23.6	26.8	2.7	3.8	1.7
Emissions	40.9	51.0	59.0	28.5	34.1	35.4	5.3	7.7	5.2
Innovation	80.3	87.9	90.3	39.7	49.6	58.9	41.0	30.3	17.6
Resource use	38.5	45.5	58.8	26.1	30.7	32.4	13.3	15.0	8.3
Social	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Community	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.0
Human rights	69.2	62.5	64.8	62.3	45.4	38.9	62.2	56.8	26.9
Product resp.	18.3	13.0	8.5	14.7	10.3	9.7	17.0	14.1	12.8
Workforce	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Governance	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CSR strategy	44.2	54.2	62.7	41.2	45.1	41.6	14.8	18.4	8.6
Management	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Shareholders	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Note: This table reports the proportion of scores with zero values for each sector in the Refinitiv database. The sample covers the period from 2010 to 2019.

Table 6. Average Scores by Region

Category	2010	2015	2019	2010	2015	2019	2010	2015	2019
	World			**North America**			**Europe**		
ESG	40.4	41.4	42.8	38.9	37.5	37.1	49.9	51.9	52.5
Environment	32.8	31.9	32.1	26.9	23.3	20.5	48.5	47.9	46.3
Emissions	36.7	35.2	36.1	29.5	24.5	22.3	54.8	54.1	52.5
Innovation	21.9	21.7	21.8	18.1	15.7	13.7	32.0	32.5	32.3
Resource use	36.3	35.5	35.5	29.5	26.4	23.0	54.8	53.7	51.5
Social	39.3	42.0	44.8	40.7	39.5	39.9	50.3	55.3	57.7
Community	49.9	50.0	50.2	64.8	59.3	56.7	50.3	53.1	51.3
Human rights	16.7	22.3	29.3	12.4	16.5	19.4	31.9	41.1	51.9
Product resp.	37.6	41.4	45.0	37.1	37.1	38.8	46.2	52.5	55.7
Workforce	50.3	50.1	50.4	46.1	40.4	37.6	68.1	68.9	67.8
Governance	47.9	47.6	47.9	47.1	46.3	45.9	49.4	49.5	50.1
CSR Strategy	32.6	31.8	33.3	27.3	21.4	19.4	42.3	43.8	45.7
Management	50.3	50.1	50.1	50.1	50.2	49.9	50.5	50.3	50.8
Shareholders	50.2	50.1	50.1	50.2	50.1	50.0	50.4	50.3	50.6
	Pacific			**Emerging countries**					
ESG	36.1	38.4	45.0	36.6	41.4	43.3			
Environment	35.5	31.5	39.7	27.0	33.0	35.8			
Emissions	31.3	34.4	44.4	31.2	37.2	41.0			
Innovation	34.4	23.7	28.0	16.2	21.3	23.7			
Resource use	23.4	33.7	42.3	30.5	36.8	39.6			
Social	33.3	35.3	44.4	34.5	40.8	43.0			
Community	30.5	37.7	43.1	40.3	43.4	43.2			
Human rights	36.7	13.4	26.3	13.3	23.5	28.7			
Product resp.	10.5	39.8	48.7	33.9	41.1	44.9			
Workforce	33.5	47.7	56.8	48.1	52.3	54.0			
Governance	39.4	47.6	48.8	47.5	48.2	48.9			
CSR Strategy	47.2	33.2	40.4	32.6	37.5	42.1			
Management	30.2	49.8	50.3	49.7	49.8	50.0			
Shareholders	50.0	49.7	49.3	50.1	49.9	50.0			

Note: This table reports the average scores for each region in the Refinitiv database. The sample covers the period from 2010 to 2019.

Table 7. Average Scores by Sector

Category	2010	2015	2019	2010	2015	2019	2010	2015	2019
	Energy			**Basic Materials**			**Industrials**		
ESG	36.7	38.9	42.8	39.1	42.7	45.1	41.0	41.9	44.2
Environment	31.1	32.5	37.0	37.4	39.5	41.4	36.5	35.8	37.3
Emissions	38.6	40.0	45.3	40.1	43.1	44.5	40.3	37.8	41.4
Innovation	11.7	12.5	15.8	19.5	21.5	25.4	26.6	27.8	27.9
Resource use	35.1	37.0	40.6	40.3	42.3	43.5	39.1	39.4	40.9
Social	35.7	37.9	42.9	36.3	40.0	44.3	38.9	42.2	45.0
Community	50.0	50.2	50.7	49.6	50.2	51.2	49.2	50.1	50.1
Human rights	14.8	19.6	30.4	19.6	27.2	36.1	18.0	26.9	35.4
Product resp.	32.7	38.5	43.0	34.2	39.4	43.2	42.1	43.7	45.8
Workforce	50.3	50.2	50.8	50.2	50.2	51.0	50.0	50.1	50.1
Governance	47.7	49.2	50.9	46.8	52.0	52.1	48.1	47.3	49.9
CSR strategy	35.0	36.3	41.6	41.3	44.5	49.6	34.2	33.2	36.5
Management	49.4	51.2	51.8	46.9	52.4	52.1	49.6	49.4	52.2
Shareholders	50.6	51.4	54.2	50.1	55.5	53.6	52.5	49.5	51.0
	Consumer Cyclical			**Consumer Non Cycl.**			**Financials**		
ESG	39.2	41.0	43.8	42.8	43.3	45.1	41.4	41.6	42.3
Environment	31.9	32.7	34.5	38.9	38.0	40.4	24.6	23.4	23.9
Emissions	34.1	35.1	37.8	40.6	39.8	43.3	32.5	29.2	30.1
Innovation	20.6	21.0	22.3	27.9	25.2	26.0	20.3	19.9	21.1
Resource use	36.4	36.1	38.1	41.0	40.5	42.9	29.8	28.2	27.6
Social	39.0	41.5	45.4	40.8	43.1	45.3	40.8	42.4	44.4
Community	49.8	49.9	50.2	49.8	50.0	50.6	50.1	50.1	49.8
Human rights	20.1	25.0	33.9	22.0	29.0	35.0	10.9	14.1	21.4
Product resp.	36.1	40.6	46.1	44.7	45.7	46.6	35.0	39.8	45.0
Workforce	50.0	50.0	50.5	50.0	50.0	51.0	50.2	50.2	50.2
Governance	43.9	45.4	47.9	50.9	49.2	49.6	49.3	47.8	47.2
CSR strategy	26.4	28.5	31.5	39.1	39.3	43.5	27.3	26.3	27.9
Management	47.4	48.3	50.9	53.4	50.6	50.2	52.8	51.7	50.7
Shareholders	43.8	47.2	48.5	50.4	51.2	51.6	52.2	48.9	48.2
	Health Care			**Technology**			**Utilities**		
ESG	39.6	38.5	36.3	42.3	42.2	42.3	44.1	44.3	47.5
Environment	26.2	22.2	17.6	34.0	30.4	28.2	40.4	41.5	44.8
Emissions	29.6	24.5	20.3	35.8	33.0	32.5	47.3	46.2	47.2
Innovation	9.9	6.1	4.8	30.2	25.3	20.7	29.2	34.8	41.4
Resource use	30.8	27.2	20.5	37.0	34.7	34.1	43.3	42.5	45.4
Social	39.9	42.2	43.3	40.6	44.0	45.0	41.8	42.8	46.3
Community	50.4	50.0	49.8	49.9	50.0	49.9	49.9	50.0	50.0
Human rights	15.4	18.7	17.7	18.9	27.3	30.9	18.9	21.6	35.9
Product resp.	41.1	43.6	45.6	42.8	44.9	45.3	41.5	42.9	44.0
Workforce	50.4	50.1	49.9	50.4	50.1	50.1	50.0	50.0	50.4
Governance	47.3	44.4	40.6	50.3	47.3	46.8	54.1	50.8	53.5
CSR strategy	25.7	21.0	16.5	30.4	28.5	27.7	50.8	46.7	55.0
Management	51.4	47.2	43.4	53.2	50.5	49.5	56.4	52.5	54.7
Shareholders	47.7	50.6	47.2	53.7	49.3	50.1	48.6	47.9	48.6

Note: This table reports the average scores for each sector in the Refinitiv database. The sample covers the period from 2010 to 2019.

Table 8. Summary Statistics on Exclusion Portfolio – Global Exclusion / Reinvestment

Category	Prop. firms zeros	Prop. mkt cap zeros	Prop. firms excluded	Prop. mkt cap excluded	Bench-mark score	Port-folio score	Score gain	Ann. return	Ann. vola-tility	Sharpe ratio	Ann. track. error
MSCI ACWI	–	–	0.0	0.0	–	–	–	9.71	14.07	0.69	–
Benchmark	–	–	5.2	2.4	–	–	–	9.91	14.34	0.69	0.94
Panel A: 10% **threshold**											
ESG	0.0	0.0	26.2	9.9	64.0	68.4	4.4	9.85	14.41	0.68	1.16
Environment	8.9	4.1	22.6	9.8	61.2	67.3	6.1	9.64	14.37	0.67	1.08
Emissions	13.6	6.2	21.8	9.9	66.9	73.9	7.0	9.57	14.37	0.67	1.11
Innovation	40.3	26.1	40.3	26.1	44.1	59.8	15.7	10.05	14.30	0.70	1.53
Resource Use	13.9	6.0	23.7	9.9	69.2	76.4	7.2	9.74	14.36	0.68	1.06
Social	0.0	0.0	27.1	9.9	65.4	70.2	4.8	10.02	14.38	0.70	1.13
Community	0.6	0.2	27.7	9.9	73.4	79.6	6.2	10.19	14.48	0.70	1.20
Human Rights	39.8	21.9	40.2	22.1	50.3	64.5	14.2	9.67	14.46	0.67	1.29
Product Resp.	13.0	4.5	20.4	10.0	61.0	67.0	5.9	10.03	14.25	0.70	1.05
Workforce	0.0	0.0	23.3	9.8	72.8	78.4	5.7	9.84	14.41	0.68	1.07
Governance	0.0	0.0	20.0	9.9	63.7	68.3	4.7	9.67	14.35	0.67	1.16
CSR Strategy	18.7	9.5	20.2	10.3	61.9	69.0	7.1	9.60	14.37	0.67	1.09
Management	0.0	0.0	18.2	9.9	65.9	71.7	5.8	9.84	14.34	0.69	1.13
Shareholders	0.0	0.0	13.0	9.8	57.4	62.9	5.5	9.79	14.38	0.68	1.14
Panel B: 25% **threshold**											
ESG	0.0	0.0	50.4	24.7	64.0	73.0	9.1	9.81	14.33	0.68	1.40
Environment	8.9	4.1	47.2	24.7	61.2	74.3	13.1	9.58	14.31	0.67	1.37
Emissions	13.6	6.2	46.4	24.7	66.9	81.3	14.4	9.60	14.37	0.67	1.33
Innovation	40.3	26.1	40.5	26.3	44.1	60.0	15.9	10.06	14.32	0.70	1.55
Resource Use	13.9	6.0	49.2	24.8	69.2	83.9	14.7	9.79	14.40	0.68	1.33
Social	0.0	0.0	49.6	24.7	65.4	75.5	10.1	10.12	14.37	0.70	1.37
Community	0.6	0.2	49.7	24.8	73.4	85.8	12.4	10.60	14.48	0.73	1.45
Human Rights	39.8	21.9	47.6	26.5	50.3	67.3	17.0	9.69	14.52	0.67	1.44
Product Resp.	13.0	4.5	37.7	25.3	61.0	75.2	14.2	9.64	14.15	0.68	1.34
Workforce	0.0	0.0	43.7	24.5	72.8	84.2	11.5	9.54	14.48	0.66	1.36
Governance	0.0	0.0	40.6	24.7	63.7	73.8	10.1	9.95	14.38	0.69	1.33
CSR Strategy	18.7	9.5	44.4	26.1	61.9	78.3	16.5	9.19	14.42	0.64	1.46
Management	0.0	0.0	38.1	24.7	65.9	78.7	12.7	10.10	14.39	0.70	1.22
Shareholders	0.0	0.0	30.8	24.7	57.4	70.4	13.0	9.78	14.35	0.68	1.23
Panel C: 33% **threshold**											
ESG	0.0	0.0	60.2	33.0	64.0	75.2	11.2	9.70	14.29	0.68	1.60
Environment	8.9	4.1	55.7	32.6	61.2	77.1	15.9	8.99	14.16	0.64	1.62
Emissions	13.6	6.2	55.2	32.9	66.9	84.4	17.5	9.33	14.51	0.64	1.59
Innovation	40.3	26.1	49.5	33.5	44.1	64.8	20.7	10.12	14.36	0.70	1.56
Resource Use	13.9	6.0	59.1	32.9	69.2	86.9	17.7	9.64	14.42	0.67	1.50
Social	0.0	0.0	58.4	32.9	65.4	78.1	12.7	9.62	14.42	0.67	1.64
Community	0.6	0.2	57.4	32.8	73.4	88.4	15.0	10.53	14.45	0.73	1.62
Human Rights	39.8	21.9	55.0	33.1	50.3	71.7	21.4	9.76	14.44	0.68	1.51
Product Resp.	13.0	4.5	48.2	33.1	61.0	79.2	18.2	9.56	14.12	0.68	1.39
Workforce	0.0	0.0	53.1	32.9	72.8	86.8	14.1	9.45	14.51	0.65	1.49
Governance	0.0	0.0	49.6	32.9	63.7	76.4	12.7	9.97	14.40	0.69	1.41
CSR Strategy	18.7	9.5	52.8	33.4	61.9	81.6	19.7	8.99	14.36	0.63	1.68
Management	0.0	0.0	46.7	32.8	65.9	82.0	16.1	10.35	14.42	0.72	1.37
Shareholders	0.0	0.0	40.0	33.0	57.4	74.4	17.0	9.95	14.32	0.69	1.34

Note: This table reports summary statistics for exclusion portfolios based on the 10%, 25%, and 33% thresholds. The first two columns report the proportion of firms with zero scores and the fraction of the market value with zero scores. The next two columns report the proportion of excluded firms and of the excluded market value relative to the MSCI ACWI. The sample includes firms that belong to the MSCI ACWI with an ESG score over the period from 2010 to 2019. Financial performance measures are computed from 2011 to 2020.

Table 9. Summary Statistics on Exclusion Portfolio – 25% exclusion

Category	Prop. firms excluded	Prop. mkt cap excluded	Bench-mark score	Port-folio score	Score gain	Ann. return	Ann. vola-tility	Sharpe ratio	Ann. track. error
MSCI ACWI	0.0	0.0	–	–	–	9.71	14.07	0.69	–
Benchmark	5.2	2.4	–	–	–	9.91	14.34	0.69	0.94
Panel A: Regional exclusion and reinvestment									
ESG	45.0	24.4	64.0	72.2	8.2	9.72	14.25	0.68	1.47
Environment	41.6	24.4	61.2	73.0	11.8	9.45	14.21	0.67	1.53
Emissions	41.1	24.3	66.9	79.9	13.0	9.64	14.30	0.67	1.32
Innovation	41.7	28.0	44.1	60.7	16.7	10.04	14.30	0.70	1.91
Resource Use	43.2	24.4	69.2	82.4	13.2	9.75	14.20	0.69	1.24
Social	44.0	24.4	65.4	74.5	9.1	10.01	14.20	0.71	1.39
Community	39.1	24.5	73.4	83.7	10.3	9.78	14.45	0.68	1.50
Human Rights	47.5	28.5	50.3	67.2	16.9	9.79	14.30	0.68	1.48
Product Resp.	37.6	25.4	61.0	74.7	13.7	9.78	14.05	0.70	1.34
Workforce	40.9	24.3	72.8	83.2	10.4	9.60	14.38	0.67	1.40
Governance	37.0	24.3	63.7	73.3	9.7	9.79	14.32	0.68	1.46
CSR Strategy	41.5	25.7	61.9	77.2	15.3	9.12	14.28	0.64	1.61
Management	35.3	24.3	65.9	78.2	12.2	10.12	14.30	0.71	1.33
Shareholders	28.4	24.4	57.4	70.0	12.7	9.72	14.40	0.67	1.41
Panel B: Sectoral exclusion and reinvestment									
ESG	49.2	24.3	64.0	72.7	8.7	9.88	14.33	0.69	1.71
Environment	46.7	24.0	61.2	73.5	12.3	9.73	14.29	0.68	1.69
Emissions	45.8	24.1	66.9	80.6	13.7	9.70	14.35	0.68	1.54
Innovation	45.9	30.7	44.1	62.6	18.5	10.06	14.40	0.70	1.70
Resource Use	47.9	24.4	69.2	83.3	14.1	9.83	14.40	0.68	1.56
Social	48.3	24.1	65.4	75.1	9.7	9.97	14.40	0.69	1.63
Community	47.5	25.3	73.4	85.4	11.9	10.53	14.44	0.73	1.54
Human Rights	50.0	29.1	50.3	68.7	18.4	9.54	14.51	0.66	1.96
Product Resp.	41.8	29.6	61.0	77.0	16.0	9.55	14.25	0.67	1.54
Workforce	42.7	24.0	72.8	83.7	11.0	9.77	14.36	0.68	1.63
Governance	39.1	24.0	63.7	73.2	9.5	10.00	14.34	0.70	1.38
CSR Strategy	44.4	25.5	61.9	77.3	15.4	9.51	14.42	0.66	1.50
Management	36.6	24.2	65.9	78.1	12.2	10.09	14.37	0.70	1.30
Shareholders	30.2	24.0	57.4	69.9	12.5	9.85	14.28	0.69	1.30
Panel C: Sector-region exclusion and reinvestment									
ESG	42.0	23.3	64.0	74.2	10.3	9.56	14.34	0.67	1.50
Environment	40.4	22.8	61.2	75.2	14.0	9.42	14.26	0.66	1.43
Emissions	39.7	22.5	66.9	81.9	15.0	9.67	14.29	0.68	1.25
Innovation	45.4	31.2	44.1	66.7	22.6	10.03	14.47	0.69	1.49
Resource Use	41.5	23.1	69.2	84.2	15.0	9.61	14.19	0.68	1.23
Social	41.1	22.7	65.4	76.7	11.3	9.54	14.33	0.67	1.47
Community	37.4	23.3	73.4	85.4	11.9	9.90	14.43	0.69	1.34
Human Rights	48.8	31.0	50.3	73.3	23.0	9.87	14.43	0.68	1.56
Product Resp.	40.3	30.0	61.0	80.6	19.5	9.86	14.29	0.69	1.48
Workforce	38.5	22.4	72.8	84.6	11.9	9.60	14.30	0.67	1.35
Governance	34.5	22.4	63.7	75.3	11.7	9.82	14.40	0.68	1.28
CSR Strategy	40.9	24.9	61.9	79.8	18.0	9.34	14.28	0.65	1.42
Management	32.9	22.5	65.9	80.5	14.6	10.02	14.37	0.70	1.23
Shareholders	28.1	22.1	57.4	72.9	15.6	9.99	14.38	0.69	1.25

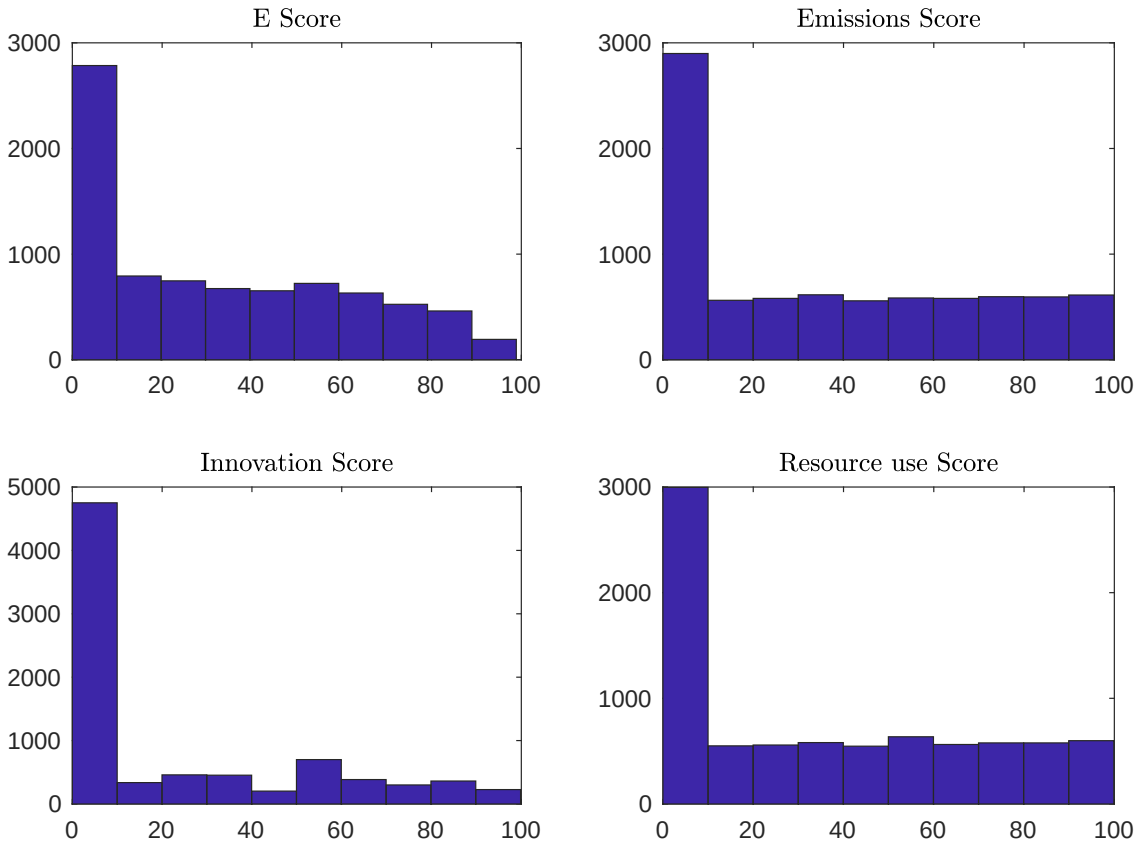
Note: This table reports summary statistics for portfolios based on sectoral, regional, and region-sector exclusion at the 25% threshold. The first two columns report the proportion of excluded firms and of the excluded market value relative to the MSCI ACWI. The sample includes firms that belong to the MSCI ACWI with an ESG score over the period from 2010 to 2019. Financial performance measures are computed from 2011 to 2020.

Table 10. Summary Statistics on Exclusion Portfolio – 33% exclusion

Category	Prop. firms excluded	Prop. mkt cap excluded	Bench-mark score	Port-folio score	Score gain	Ann. return	Ann. vola-tility	Sharpe ratio	Ann. track. error
MSCI ACWI	0.0	0.0	–	–	–	9.71	14.07	0.69	–
Benchmark	5.2	2.4	–	–	–	9.91	14.34	0.69	0.94
Panel A: Regional exclusion and reinvestment									
ESG	54.5	32.4	64.0	74.1	10.2	9.51	14.19	0.67	1.72
Environment	51.0	31.9	61.2	75.4	14.2	9.20	14.12	0.65	1.70
Emissions	50.2	32.4	66.9	82.7	15.7	9.46	14.24	0.66	1.45
Innovation	46.9	33.7	44.1	64.2	20.1	9.89	14.45	0.68	1.94
Resource Use	53.2	32.4	69.2	85.0	15.8	9.74	14.13	0.69	1.30
Social	53.4	32.4	65.4	76.7	11.3	9.59	14.18	0.68	1.59
Community	48.3	32.4	73.4	86.2	12.8	9.62	14.46	0.67	1.54
Human Rights	52.3	34.0	50.3	69.5	19.2	9.90	14.16	0.70	1.49
Product Resp.	47.4	33.1	61.0	78.2	17.2	9.99	13.99	0.71	1.35
Workforce	50.1	32.5	72.8	85.5	12.8	9.76	14.33	0.68	1.35
Governance	46.9	32.4	63.7	75.8	12.1	9.80	14.29	0.69	1.57
CSR Strategy	50.0	33.5	61.9	80.5	18.6	8.90	14.22	0.63	1.84
Management	44.8	32.5	65.9	81.4	15.5	10.18	14.33	0.71	1.49
Shareholders	37.2	32.5	57.4	73.7	16.4	9.71	14.41	0.67	1.36
Panel B: Sectoral exclusion and reinvestment									
ESG	58.7	32.3	64.0	74.7	10.8	9.72	14.37	0.68	1.83
Environment	55.7	31.9	61.2	76.4	15.2	9.64	14.34	0.67	1.77
Emissions	54.7	32.1	66.9	83.6	16.7	9.63	14.39	0.67	1.76
Innovation	53.5	37.4	44.1	66.0	21.9	9.80	14.40	0.68	1.80
Resource Use	56.9	32.4	69.2	86.2	16.9	9.58	14.46	0.66	1.72
Social	56.7	31.9	65.4	77.5	12.1	9.84	14.45	0.68	1.77
Community	55.9	32.7	73.4	87.8	14.4	10.62	14.52	0.73	1.70
Human Rights	55.8	34.6	50.3	71.7	21.4	9.62	14.60	0.66	2.00
Product Resp.	49.7	36.0	61.0	79.9	18.8	9.52	14.17	0.67	1.59
Workforce	51.3	32.3	72.8	86.3	13.5	9.58	14.35	0.67	1.68
Governance	47.9	32.3	63.7	75.7	12.0	10.09	14.36	0.70	1.50
CSR Strategy	53.2	33.6	61.9	80.5	18.7	9.36	14.46	0.65	1.67
Management	45.4	32.2	65.9	81.3	15.4	10.19	14.40	0.71	1.44
Shareholders	39.4	32.4	57.4	73.8	16.4	10.12	14.31	0.71	1.35
Panel C: Sector-region exclusion and reinvestment									
ESG	51.6	31.0	64.0	75.8	11.9	9.50	14.30	0.66	1.54
Environment	49.2	30.4	61.2	77.5	16.3	9.24	14.26	0.65	1.49
Emissions	48.2	30.6	66.9	84.4	17.5	9.77	14.13	0.69	1.34
Innovation	50.8	37.4	44.1	68.5	24.4	9.74	14.41	0.68	1.56
Resource Use	50.0	30.9	69.2	86.5	17.3	9.42	14.11	0.67	1.39
Social	49.8	30.7	65.4	78.7	13.3	9.54	14.32	0.67	1.49
Community	45.7	31.4	73.4	87.6	14.2	9.71	14.43	0.67	1.39
Human Rights	53.5	36.2	50.3	74.3	24.0	9.77	14.45	0.68	1.60
Product Resp.	48.2	36.2	61.0	82.3	21.3	9.54	14.18	0.67	1.47
Workforce	47.6	30.9	72.8	87.0	14.2	9.65	14.23	0.68	1.34
Governance	43.5	30.7	63.7	77.6	14.0	9.99	14.36	0.70	1.40
CSR Strategy	49.5	33.0	61.9	82.5	20.6	8.91	14.29	0.62	1.60
Management	41.5	30.5	65.9	83.4	17.5	10.14	14.43	0.70	1.34
Shareholders	36.3	30.7	57.4	76.6	19.2	10.25	14.50	0.71	1.41

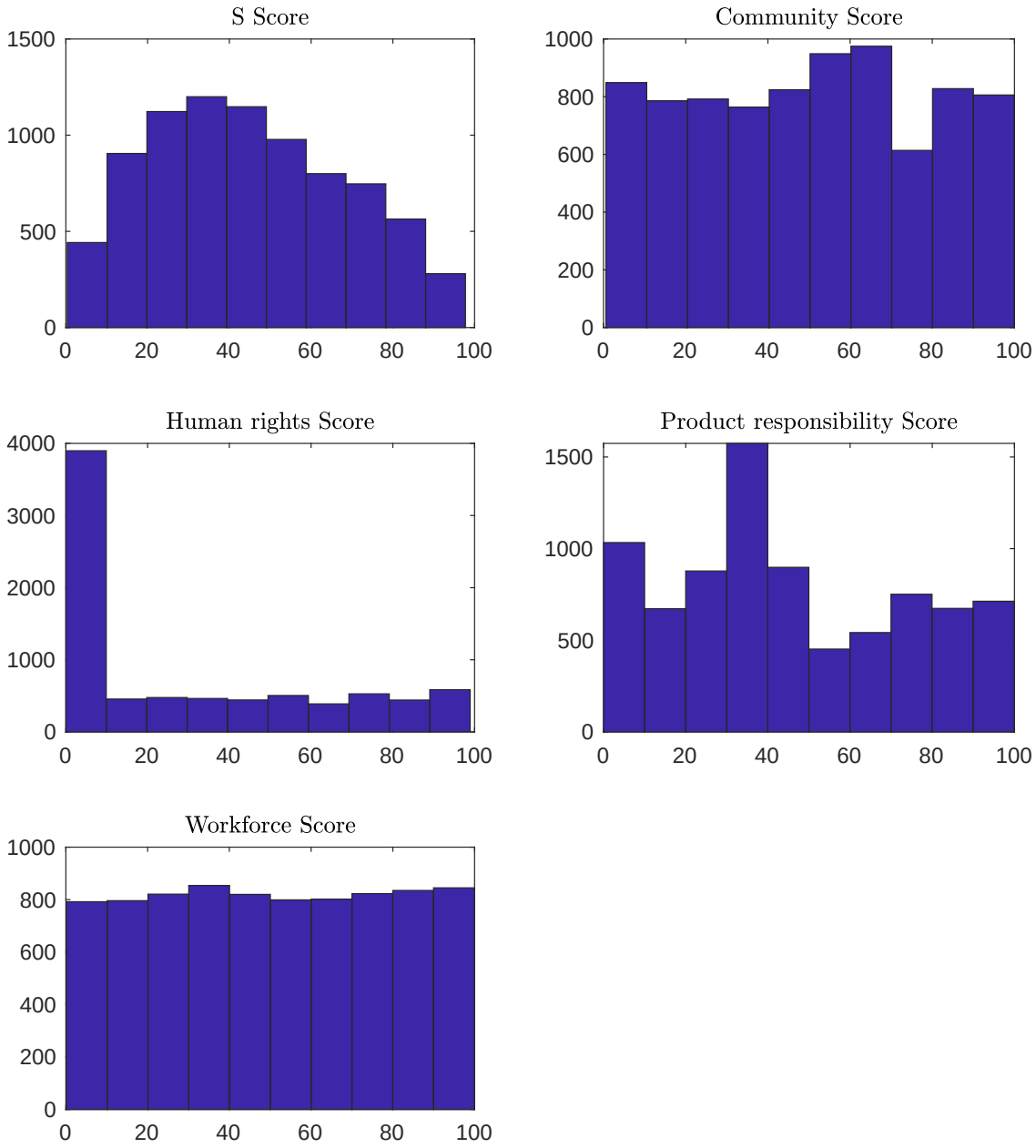
Note: This table reports summary statistics for portfolios based on sectoral, regional, and region-sector exclusion at the 33% threshold. The first two columns report the proportion of excluded firms and of the excluded market value relative to the MSCI ACWI. The sample includes firms that belong to the MSCI ACWI with an ESG score over the period from 2010 to 2019. Financial performance measures are computed from 2011 to 2020.

Figure 1. Distribution of the E Score and Categories – 2019



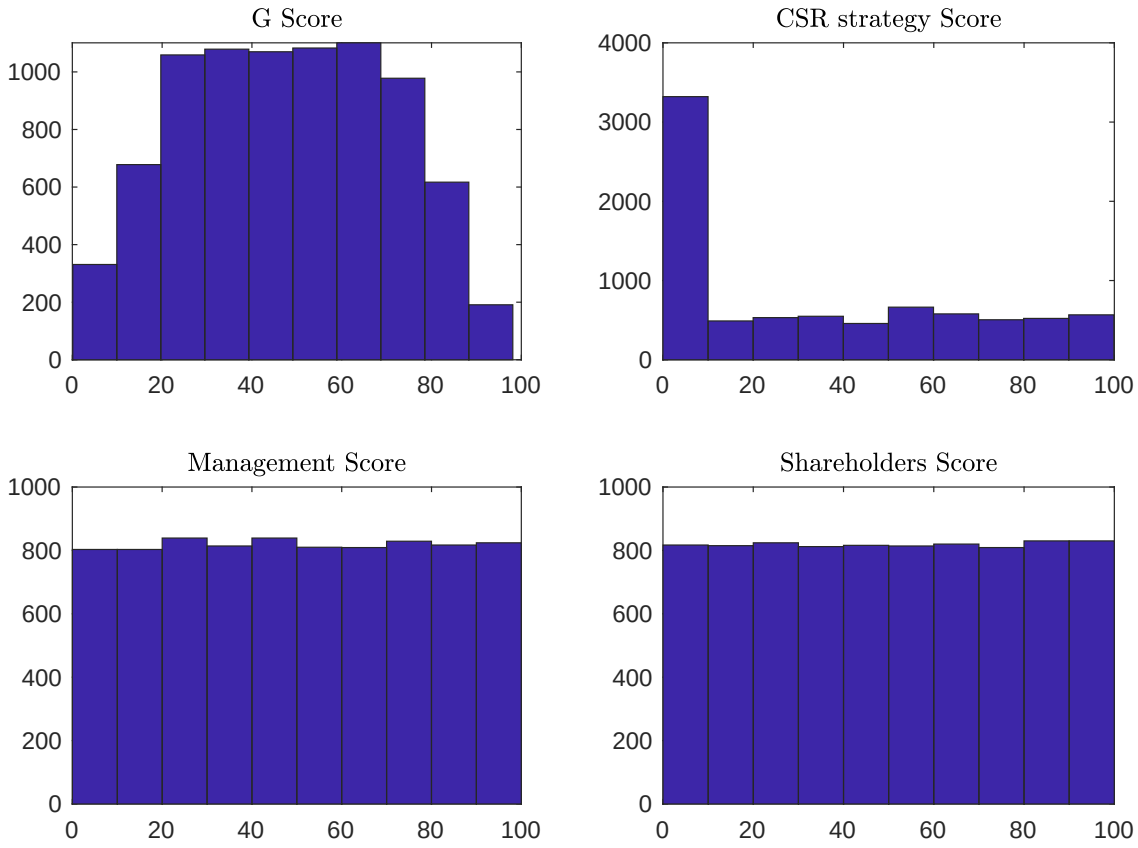
Note: This figure displays the cross-section distribution of scores for the E pillar and its categories for 2019.

Figure 2. Distribution of the S Score and Categories – 2019



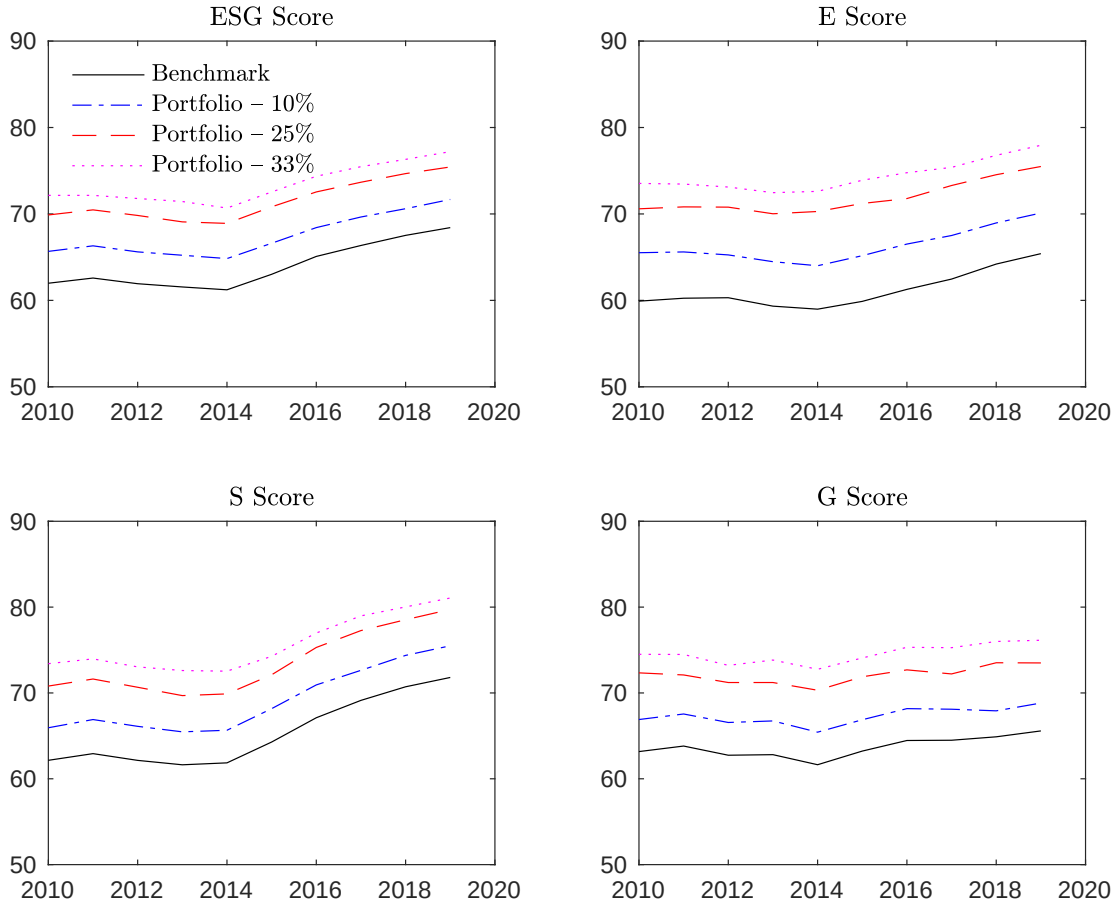
Note: This figure displays the cross-section distribution of scores for the S pillar and its categories for 2019.

Figure 3. Distribution of the G Score and Categories – 2019



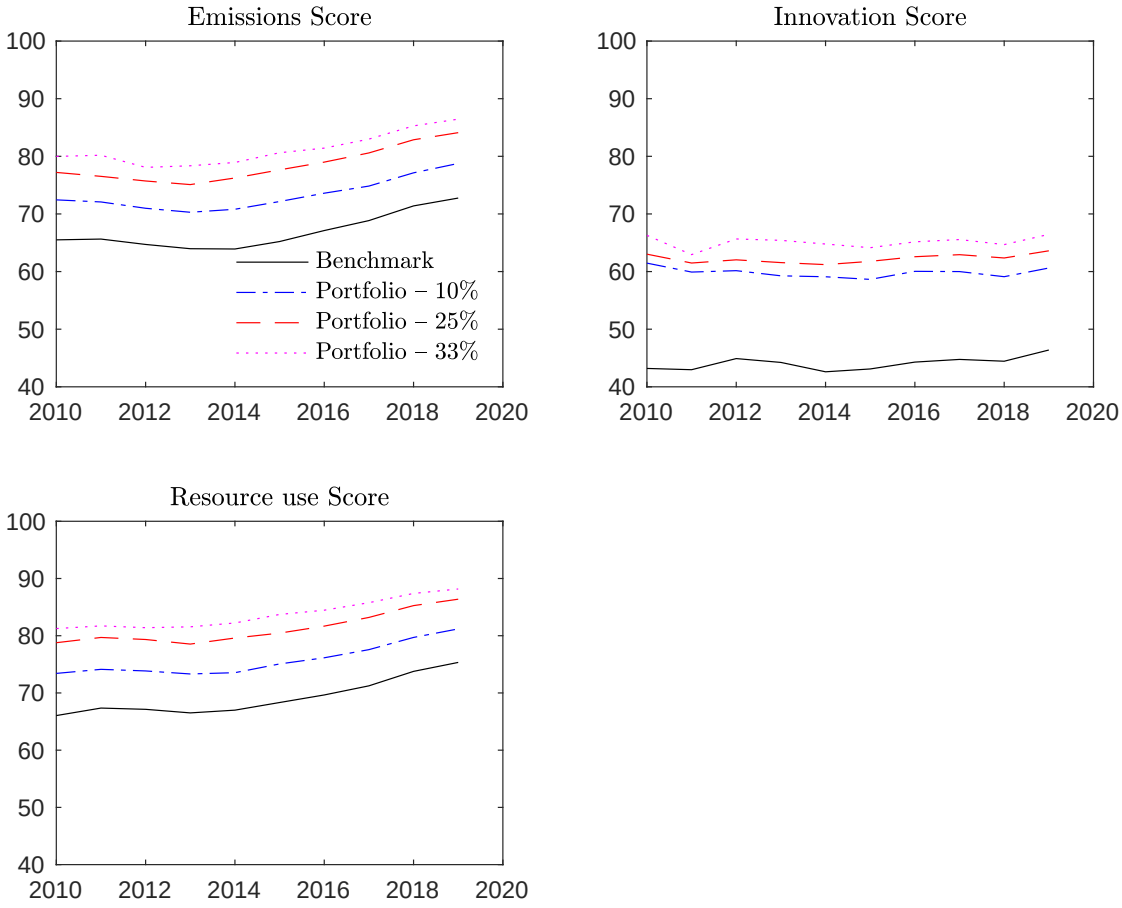
Note: This figure displays the cross-section distribution of scores for the G pillar and its categories for 2019.

Figure 4. Scores of Exclusion Portfolio – ESG Pillars



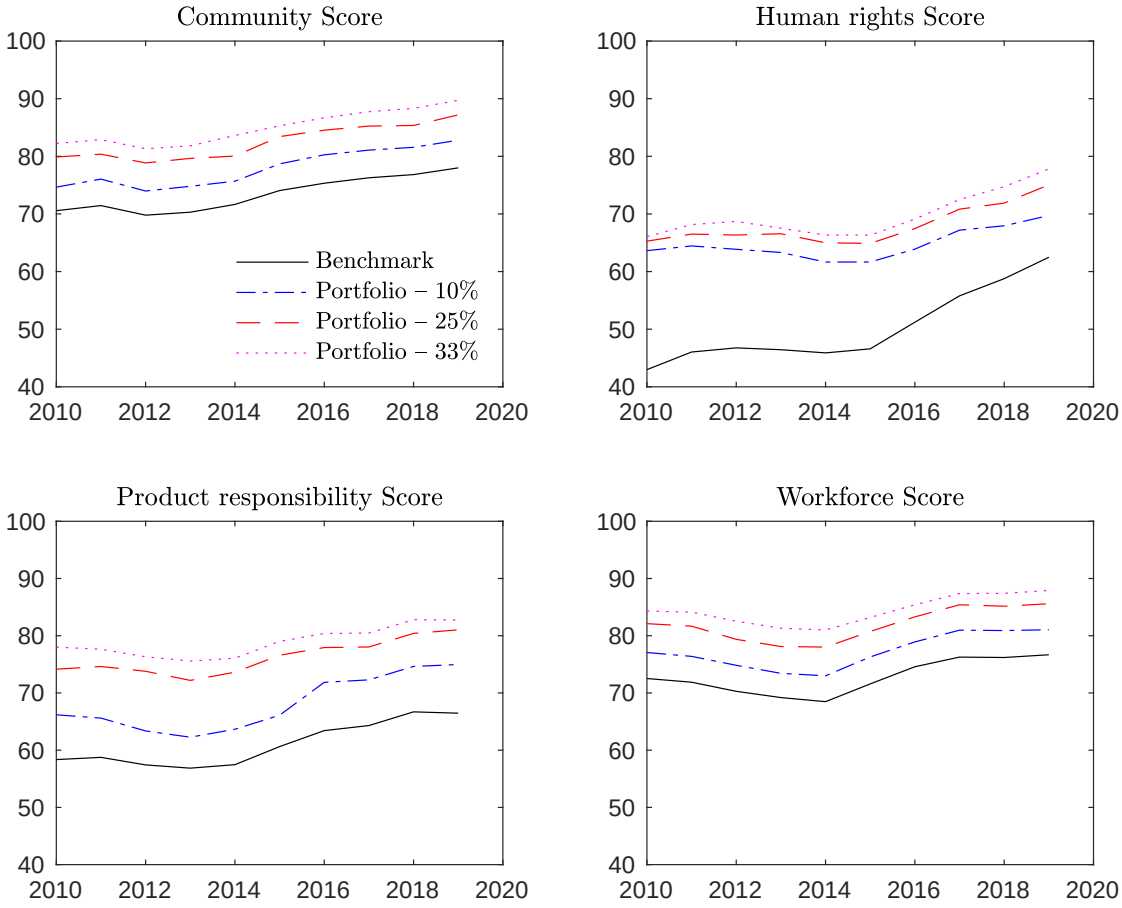
Note: This figure displays the temporal evolution of the score of the benchmark and the exclusion portfolios based on the 10%, 25%, and 33% thresholds, for the ESG, E, S, and G pillars.

Figure 5. Scores of Exclusion Portfolio – E Categories



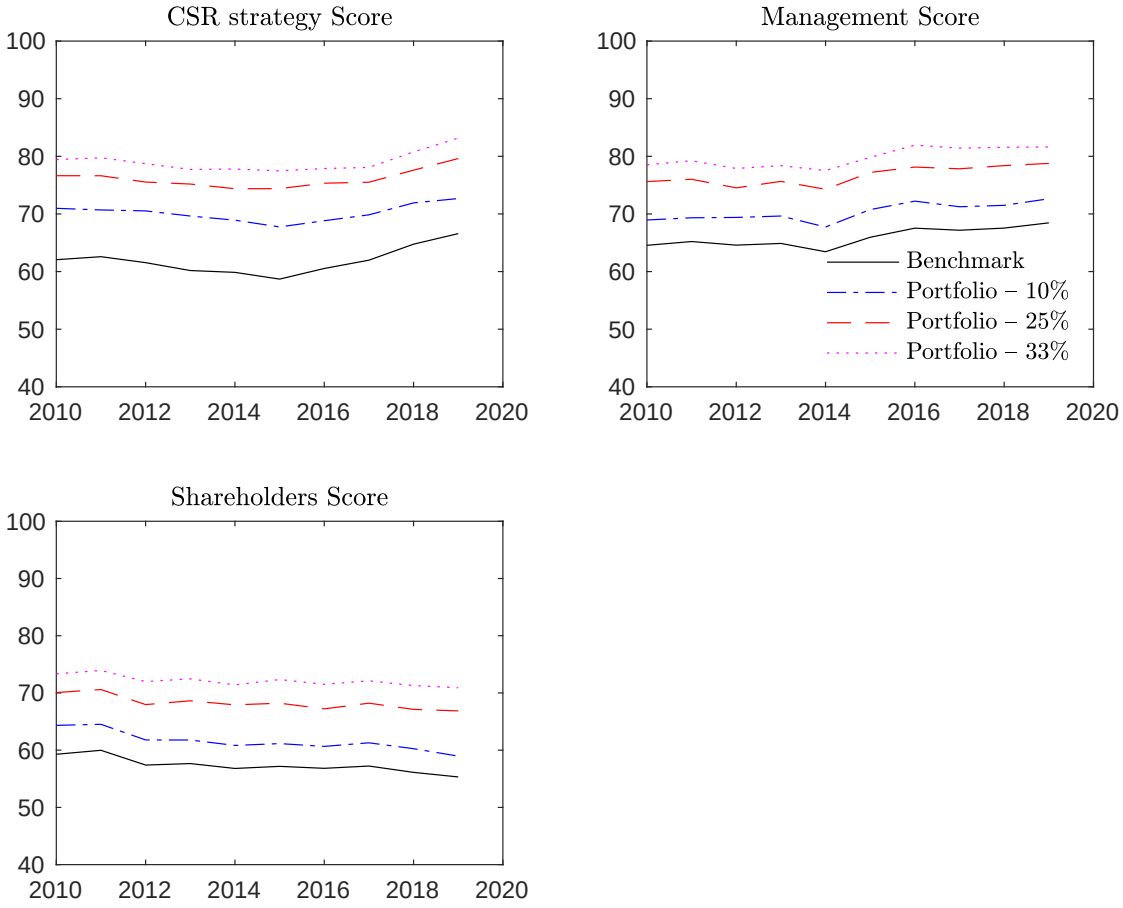
Note: This figure displays the temporal evolution of the score of the benchmark and the exclusion portfolios based on the 10%, 25%, and 33% thresholds, for the E pillar and its categories.

Figure 6. Scores of Exclusion Portfolio – S Categories



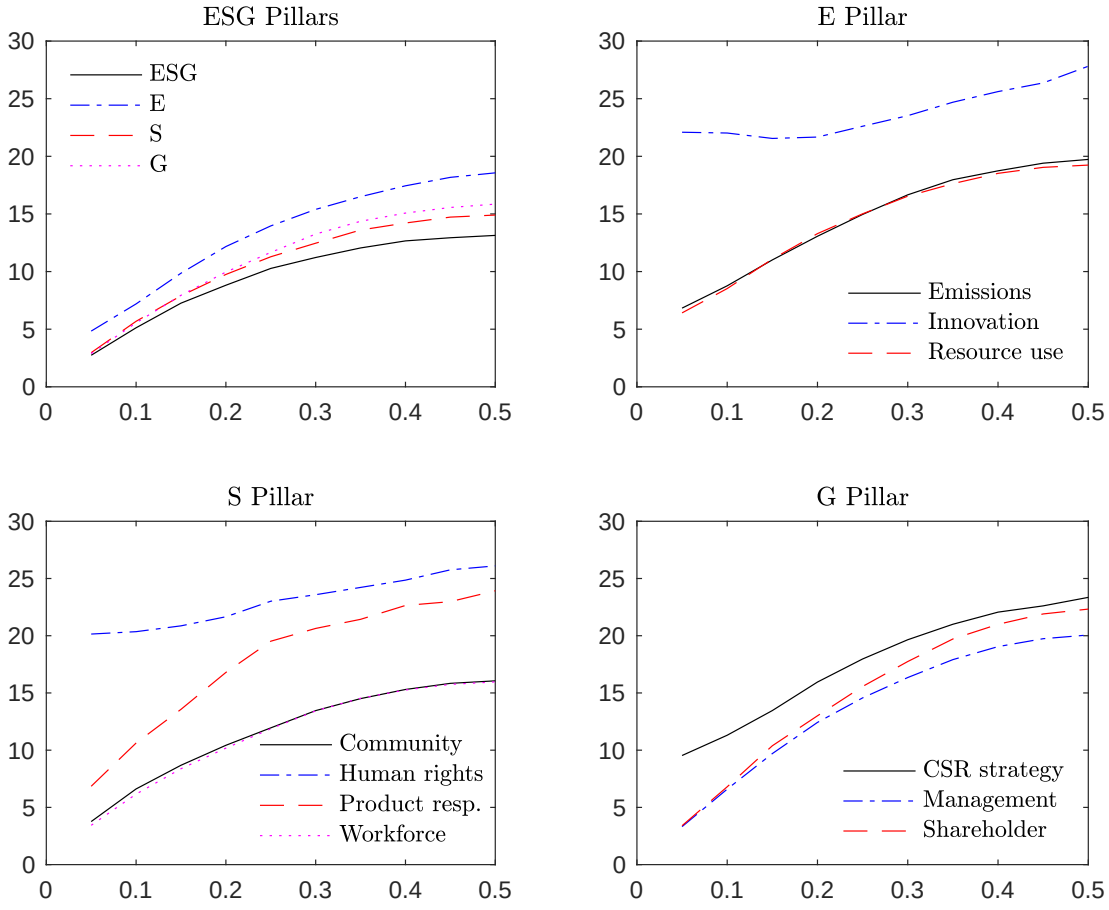
Note: This figure displays the temporal evolution of the score of the benchmark and the exclusion portfolios based on the 10%, 25%, and 33% thresholds, for the S pillar and its categories.

Figure 7. Scores of Exclusion Portfolio – G Categories



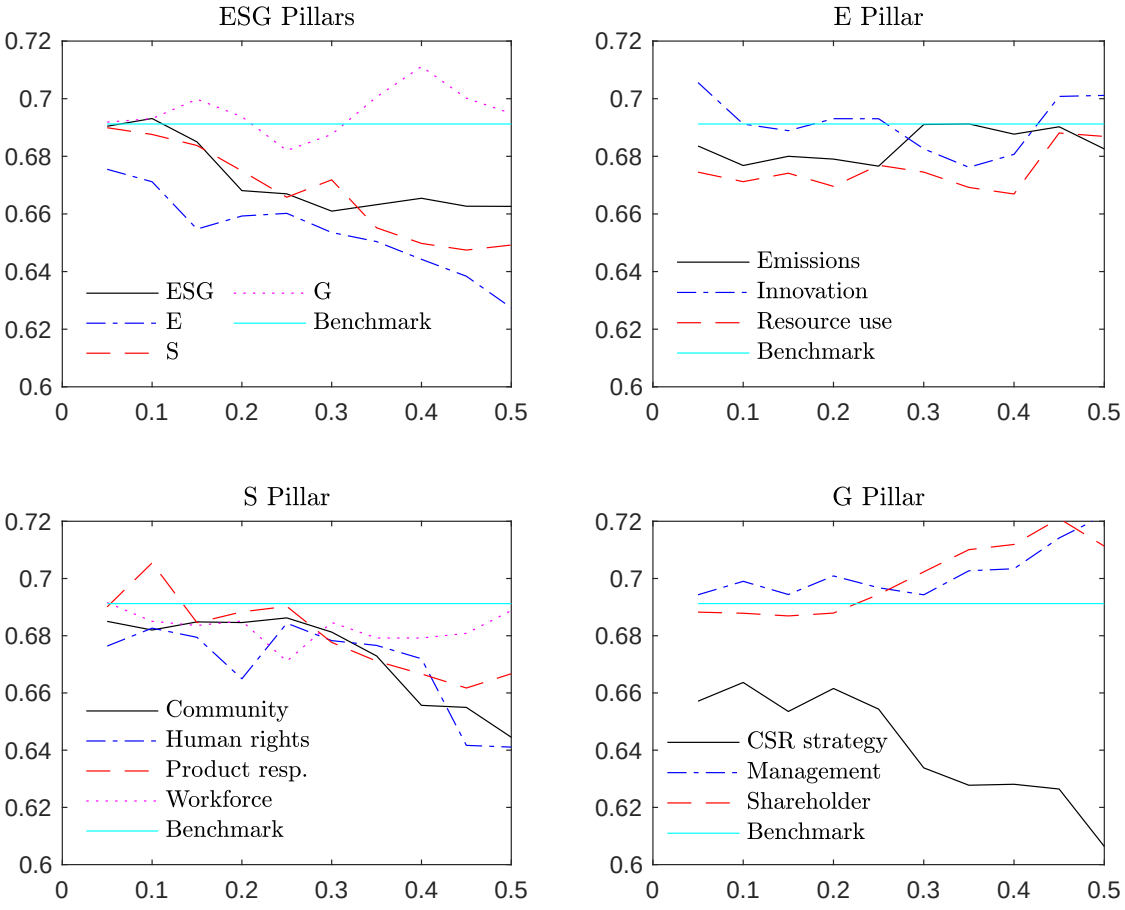
Note: This figure displays the temporal evolution of the score of the benchmark and the exclusion portfolios based on the 10%, 25%, and 33% thresholds, for the G pillar and its categories.

Figure 8. Impact of Exclusion Threshold on the Score Gain



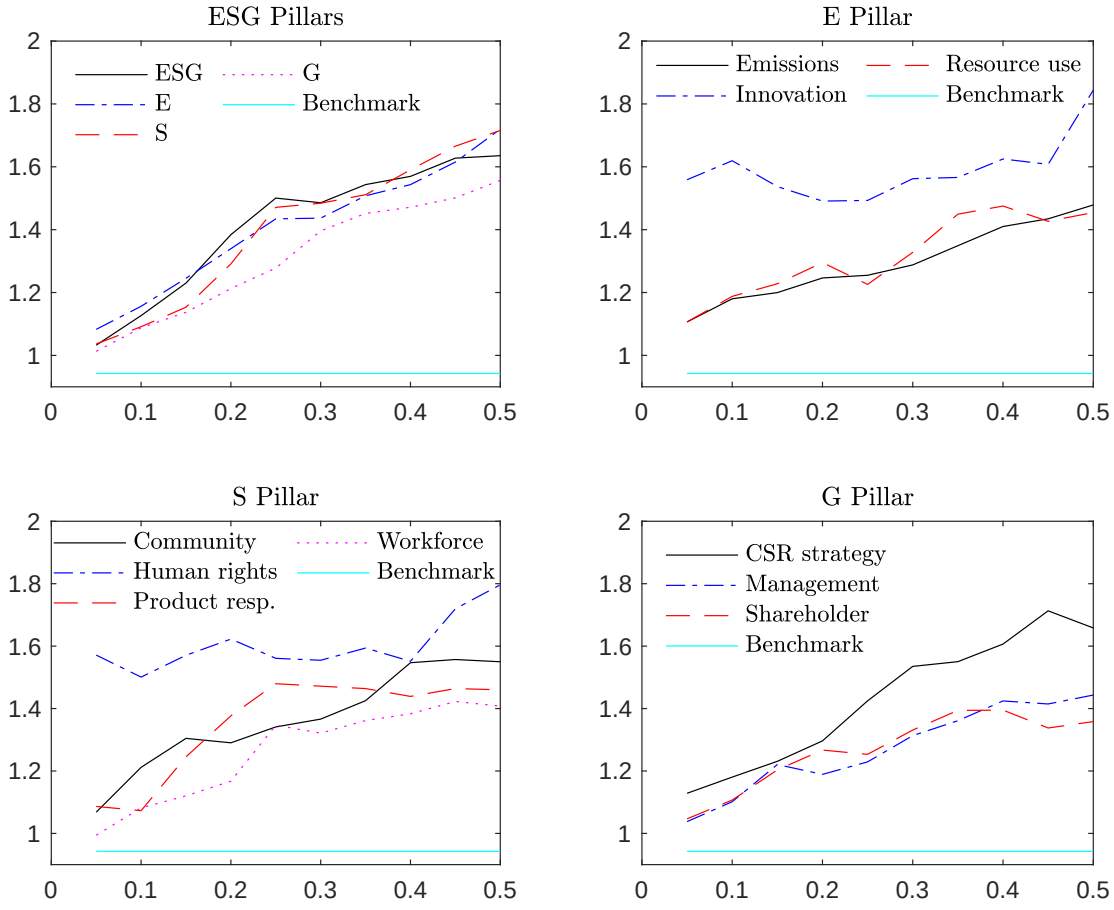
Note: This figure displays the gain in the score of the exclusion portfolios when the threshold is increased from 5% to 50%, for the various pillars and categories.

Figure 9. Impact of Exclusion Threshold on the Sharpe Ratio



Note: This figure displays the Sharpe ratio of the exclusion portfolios when the threshold is increased from 5% to 50%, for the various pillars and categories.

Figure 10. Impact of Exclusion Threshold on the Tracking Error



Note: This figure displays the annual tracking error of the exclusion portfolios when the threshold is increased from 5% to 50%, for the various pillars and categories.

Previous volumes in this series

1007 March 2022	Cross-border regulatory spillovers and macroprudential policy coordination	Pierre-Richard Agénor, Timothy P Jackson and Luiz A Pereira da Silva
1006 February 2022	Estimating conditional treatment effects of EIB lending to SMEs in Europe	Alessandro Barbera, Aron Gereben and Marcin Wolski
1005 February 2022	The NAIRU and informality in the Mexican labor market	Ana María Aguilar-Argaez, Carlo Alcaraz, Claudia Ramírez and Cid Alonso Rodríguez-Pérez
1004 February 2022	Original sin redux: a model-based evaluation	Boris Hofmann, Nikhil Patel and Steve Pak Yeung Wu
1003 February 2022	Global production linkages and stock market co-movement	Raphael Auer, Bruce Muneaki Iwadate, Andreas Schrimpf and Alexander Wagner
1002 February 2022	Exorbitant privilege? Quantitative easing and the bond market subsidy of prospective fallen angels	Viral V Acharya, Ryan Banerjee, Matteo Crosignani, Tim Eisert and Renée Spigt
1001 February 2022	Unequal expenditure switching: Evidence from Switzerland	Raphael Auer, Ariel Burstein, Sarah Lein and Jonathan Vogel
1000 February 2022	Dollar beta and stock returns	Valentina Bruno, Ilhyock Shim and Hyun Song Shin
999 February 2022	Shadow loans and regulatory arbitrage: evidence from China	Amanda Liu, Jing Liu and Ilhyock Shim
998 February 2022	Does IT help? Information technology in banking and entrepreneurship	Toni Ahnert, Sebastian Doerr, Nicola Pierri and Yannick Timmer
997 February 2022	Money markets, collateral and monetary policy	Fiorella De Fiore, Marie Hoerova and Harald Uhlig
996 January 2022	Monetary policy expectation errors	Maik Schmeling, Andreas Schrimpf and Sigurd A M Steffensen
995 January 2022	When uncertainty decouples expected and unexpected losses	Mikael Juselius and Nikola Tarashev
994 January 2022	Capital flows and institutions	Deniz Igan, Alexandre R Lauwers and Damien Puy

All volumes are available on our website www.bis.org.