Do Corporate Carbon Emissions Data Enable Investors to Mitigate Climate Change?

Vitali Kalesnik, Marco Wilkens, and Jonas Zink

Vitali Kalesnik
is a partner and the director of research at Research Affiliates Global Advisors (Europe), Ltd. in London, UK.
kalesnik@rallc.com

Marco Wilkens
is professor of finance and banking at the University of Augsburg in Augsburg, Germany.
marco.wilkens@wiwi.uni-augsburg.de

Jonas Zink
is a research associate, doctoral candidate and the chair of finance and banking at the University of Augsburg in Augsburg, Germany.
jonas.zink@wiwi.uni-augsburg.de

KEY FINDINGS

- Without mandatory carbon emissions reporting, many companies do not self-report. We find that data providers' estimated carbon emissions are not a satisfactory substitute for self-reporting. Estimated emissions data are 2.4 times less effective than self-reported data.
- We find that forward-looking carbon scores from data providers lack the power to predict future company emissions.
- We find that third-party estimated emissions are not, as many investors believe, a satisfactory substitute for company-reported emissions. Mandatory and audited reporting is needed to provide investors quality emissions data.

ABSTRACT

Investors play an important role in combating climate change. The authors examine several types of currently available carbon emissions data in their capacity to enable investors to incentivize companies to reduce emissions. The authors evaluate the information content of estimated current and forward-looking carbon emissions data from four popular data providers. Absent mandatory reporting and although many companies report their carbon emissions, much emissions data are estimated by data providers. Despite the providers' claims of accuracy, the authors find the data on estimated current emissions (often comprising >50% of observations) mostly capture only basic company information (e.g., company size and industry). The authors find that forward-looking carbon scores from different data providers do not have any power in predicting future changes in emissions. The authors' analyses suggest that estimated emissions are at least 2.4 times less effective than self-reported emissions. Their results debunk the belief that third-party estimated emissions are a satisfactory substitute for company-reported emissions and call for mandatory and audited carbon emissions disclosure.

In this article, we analyze the quality of corporate carbon emissions data. Accuracy in emissions data is extremely important to investors who desire to impact climate change (Chowdhry, Davies, and Waters 2019; Barber, Morse, and Yasuda 2020; Oehmke and Opp 2020).[1] Investors need accurate data to identify green and brown

[1] Investors with more than US$100 trillion assets under management (AUM) are signatories of the Principles for Responsible Investment (PRI), which aim, among other goals, to tackle climate-change issues (PRI 2020). Furthermore, investors who represent about US$52 trillion in AUM have joined Climate Action 100+, an organization more directly focused on coping with climate change (Climate Action 100+ 2020).

companies in order to make appropriate selections for their portfolios, including strategies to incentivize greener corporate behavior, such as switching investments from brown to green companies and engaging in activist measures (Broccardo, Hart, and Zingales 2020).[2]

Without mandatory reporting, as is currently the case, some companies do not report their emissions. Data providers attempt to close this availability gap by estimating GHG emissions. GHG emissions comprise all greenhouse gases defined in the Kyoto Protocol that cause anthropogenic climate change including carbon dioxide, methane, and nitrous oxide (United Nations 1998). Each greenhouse gas (GHG) contributes to the greenhouse effect differently, thus they are standardized into carbon emissions equivalents. This article analyzes the accuracy of the estimated data currently available to investors.[3] Our findings suggest that the status quo, in which GHG reporting is voluntary and data providers estimate the missing data, is inadequate. Overall, the estimated data capture most of the variation of the reported emissions, but conservative estimates suggest that investor actions are at least 2.4 times more diluted when investors use estimated emissions compared to reported emissions. Our results have important implications for investors, companies, regulators, standard setters, data providers, and researchers.

EMISSIONS DATA QUALITY AND THE IMPLICATIONS FOR INVESTORS

Our framework evaluates the quality of carbon data, stipulating five criteria that enhance the effectiveness of investor actions: 1) high data coverage, 2) comparability between companies, 3) consistency across data providers, 4) predictive power of forward-looking information, and 5) accuracy in reflecting true emissions. We use these criteria to compare the carbon data available to investors from four major carbon data providers, which we label DP_A, DP_B, DP_C, and DP_D.[4]

We show these data sharply differ in terms of market cap and emissions coverage. In the absence of mandatory reporting, some companies choose not to report, leaving data providers to estimate carbon emissions for the nonreporters. Estimated emissions often compose a large fraction of the data providers' datasets.

Many investors view estimated emissions as a satisfactory substitute for company-reported emissions (henceforth, reported emissions), revealing an implicit assumption that data providers are successfully closing the data availability gap. We examine this assumption by analyzing the ability of both reported and estimated corporate carbon emissions data to help investors mitigate climate change. We test

[2] If carbon emissions data are noisy and of low quality, the real carbon performance of companies is less likely to be reflected in stock prices. In order to urge companies to become greener, collective and concentrated actions by investors are required (Fama and French 2007; Gollier and Pouget 2014; Pástor, Stambaugh, and Taylor 2021). Only then can investors exert sufficient pressure on stock prices to incentivize companies to reduce emissions (Heinkel, Kraus, and Zechner 2001; De Angelis, Tankov, and Zerbib 2020). Emissions data accuracy is also needed for shareholder engagement activities, which can be very time consuming and costly to be successful (Naaraayanan, Sachdeva, and Sharma 2021; Akey and Appel 2020).

[3] Several earlier studies analyze which factors contribute to a company's decision to disclose corporate GHG emissions (e.g., Prado-Lorenzo, Gallego-Alvarex, and Garcia-Sanchez 2009; Liesen et al. 2015) and how GHG emissions are priced in capital markets (e.g., Matsumura, Prakash, and Vera-Muñoz 2014; Lee, Park, and Klassen 2015). In the latter context, some studies distinguish between mandatorily or voluntarily disclosed GHG emissions (e.g., Busch and Lewandowski 2018) and some analyze the effect of using different data providers (e.g., Busch et al. 2018; Berg, Koelbel, and Rigobon 2020; Li and Polychronopoulos 2020).

[4] We anonymize the data providers. Our goal is to evaluate the overall data landscape available to investors and not to compare the quality of the data providers.

whether estimated emissions by data providers accurately approximate true emissions when compared to reported emissions.

For a subset of companies that started reporting in our sample period, using simple Monte Carlo simulations, we find that currently available estimated data identify the worst 5% of emitters (currently responsible for 80% of overall emissions[5]) at least 2.4 times less efficiently than reported data. In other words, investor actions are at least 2.4 times less effective when based on the currently available estimated data. We further show that estimated emissions almost exclusively reflect the companies' operating industries and size-related information, which complicates investors' ability to identify green companies in brown industries.

Our findings have several important implications. First, investors who use noisy data lower the effectiveness of their efforts to mitigate climate change. Second, to improve data accuracy, investors should consider exerting more pressure on nonreporting companies to disclose and/or to avoid investing in nonreporting companies. Third, an investor's choice of data provider can have meaningful portfolio selection implications that deviate from their climate mitigation goal. Fourth, a mandatory and standardized regulatory framework for the measurement and reporting of emissions, including audits of reported data, can vastly improve data quality. We call on regulators and standard setters to step forward in this regard. Fifth, in the absence of mandatory reporting, investors play a critical role in incentivizing companies to voluntarily disclose emissions and increase the transparency of the environmental impact of their business activities. Last, data providers should be cautious in making claims about the quality of their carbon emissions estimates for nonreporting companies.[6]

DESCRIPTION OF CARBON DATASETS

We analyze annual carbon data from four major carbon data providers (DP_A, DP_B, DP_C, and DP_D) for the seven-year period from 2010 to 2016 for which we have overlapping data from all data providers. The data providers supply investors with very different data. Although DP_C solely provides climate-related data, investors can obtain additional ESG (environmental, social, and governance) data from the other data providers. The additional ESG data are, similar to the carbon data, partly based on often voluntary self-reported data by the companies but also contain subjective assessments and estimates by the data providers. In addition to ESG data, DP_B and DP_C also offer investors a variety of financial metrics. We provide more details on the offered data and their coverage in Appendices A, B, and C.

The available carbon data broadly fall into two categories: 1) historical (reported and estimated) GHG emissions and 2) carbon scores (and ratings)[7] that contain forward-looking information (e.g., emission reduction targets).[8] Historical emissions data help identify more- and less-GHG-efficient companies. Forward-looking data provide investors information about which companies are engaged in reducing emissions and which are increasing emissions.

We access three databases that provide historical carbon emissions by emission category in accordance with the GHG Protocol (WBCSD and WRI 2015b). The fourth

[5] In our sample, the worst 300 emitters (5% of sample size) together account for 80% of total carbon emissions.

[6] Importantly, we do not interpret our results as suggesting that data providers are sloppy estimators of various datasets. Instead, likely because of information asymmetry, these estimates are the best that data providers can make.

[7] For example, an emission reduction score.

[8] More information about the forward-looking scores and ratings is in Appendix A.

data provider, DP_D, solely provides carbon emission scores. The scores capture not just the raw emissions estimates but also several other types of (forward-looking) information that DP_D deems essential.[9] The score-based data of DP_D are not easily compared to the other datasets, so we exclude these data from our baseline analyses but use these data when we examine forward-looking information.

The GHG Protocol classifies reporting into three categories of carbon emissions: 1) direct GHG emissions (Scope 1); 2) indirect GHG emissions (Scope 2) basically due to energy (e.g., electricity) used by the company; and 3) other indirect emissions (Scope 3) for which reporting is optional. Scope 3 emissions arise from the activities of the company, but occur from sources not owned or controlled by the company (WBCSD and WRI 2011, 2015a, 2015b). The Scope 3 category often represents the largest source of GHG emissions, in some cases accounting for up to 90% of the total GHG impact.[10]

Very few companies report Scope 3 emissions (only about 1,500 companies). For those that do report, Scope 3 data are available on average for just 5 of the 15 Scope-3 categories. Great variability exists in the categories companies report on, and data comparability of Scope 3 emissions is quite challenging; hence, although Scope 3 emissions are extremely important, they are not yet appropriate for an empirical analysis.[11]

Most data providers rely on self-disclosed information (e.g., from company websites or company sustainability reports) for gauging company-level GHG emissions. DP_C follows a different approach and annually sends an electronic questionnaire for companies to voluntarily complete and does not gather information from any other sources. Because the information collected through the DP_C questionnaire is very comprehensive and detailed, DP_A and DP_B also use DP_C data as a source for their own datasets. DP_A and DP_B complement the DP_C data with publicly available information.[12]

FRAMEWORK FOR CARBON DATA EVALUATION

We argue that GHG data should satisfy the following five criteria: 1) wide availability, 2) comparability between companies, 3) consistency across data providers, 4) forward-looking information should have predictive power, and 5) all data should accurately reflect true emissions. If GHG data do not fulfill these criteria, basing investment decisions on them may not lead to the desired outcome of lower carbon emissions.

---

[9] DP_D also publishes emissions data but only in the form of scores, which are not comparable to the data provided by the other data providers.

[10] https://ghgprotocol.org/sites/default/files/standards_supporting/FAQ.pdf.

[11] Moreover, Scope 3 emissions are heavily subject to double counting. In certain cases, two or more companies may account for the same emissions within Scope 3 (WBCSD and WRI 2011). Consequently, investors also face the challenge of inconsistent Scope 3 data and often do not consider them. Due to their heavy impact, we encourage companies and regulators to improve the disclosure on Scope 3 emissions, alongside the Scope 1 and Scope 2 emissions. Given the limitations related to Scope 3 emissions, our analysis focuses on the sum of Scope 1 and Scope 2 emissions, although many of the concerns we raise for these two emission types also apply to Scope 3 emissions. Further, we only focus on carbon data from publicly listed companies. Importantly, because GHG data offerings are rapidly evolving, the Scope 3 data quality available to investors today could be of somewhat higher quality than the data we use in our study.

[12] In Appendix B, we provide summary statistics for the data available from the various providers. DP_A only estimates carbon emissions jointly for the two emission categories (Scope 1 and 2), whereas DP_B provides estimates separately for the categories. Companies reporting to DP_C are on average larger as indicated by higher net sales and higher number of employees.

As with financial data, nonfinancial information needs to be available for all investment opportunities for investors to make sound investment decisions. Using data from multiple carbon datasets, we test whether GHG data are widely available for listed companies. We find that, due to voluntary reporting, GHG reporting is still in its infancy.

For the reported GHG emissions data that are available, these data need to be comparable between companies to be of value to investors. We find that due to the lack of a uniform reporting standard and generous leeway within the existing standards, reported GHG emissions are not perfectly comparable between companies.

Usually, investors obtain ESG-related nonfinancial data from a specialized ESG data provider. Preferably, reported GHG emissions would not differ among the data providers. We find some discrepancies, however, between reported emissions from various data providers. Because every data provider applies its own estimation model, estimated emissions differ more strongly.

Many investors are interested not only in current levels of emissions but also in how companies will change their carbon emissions in the future. To provide this information, data providers have started to collect information about a corporation's anticipated future carbon performance (e.g., does a company have emission reduction targets?). To be helpful for investors, this forward-looking information should have the power to predict future changes in carbon emissions.

Finally, we argue that GHG data should accurately reflect the true emissions. If investors act on noisy data, they identify the polluting companies with less accuracy, which leads to ineffective investor actions.

We illustrate this by examining the ways investors can impact the real economy. Investors follow two main approaches to realize their goal of having a positive impact on the environment: 1) shifting capital from brown (exit) to green companies (e.g., selling brown stocks and buying green ones) and 2) using a shareholder activist approach (voice, stewardship, engagement).[13]

Dordi and Weber (2019) provided evidence that divestment announcements by investors targeting brown firms result in lower share prices for those companies.[14] Heinkel, Kraus, and Zechner (2001) argued that higher prices for brown companies result in a higher cost of capital that stifles brown activities. Further, Rohleder, Wilkens, and Zink (2022) showed that concentrated divestment actions by mutual funds toward brown firms result in a change in the carbon-emitting behavior of affected firms. Conversely, providing cheaper capital to green companies has the effect of lowering the cost of capital and helps to close funding gaps. Further, because management compensation is linked to share prices (Edmans, Gabaix, and Jenter 2017), the downward price pressure incentivizes managers to align their company's operations with investors' interests.

The second approach of investors who wish to drive positive change in the environment involves shareholder activism (voice). Investors have begun to engage with a company's management to push them toward a lower negative impact on the environment.[15] In contrast to the first approach, these investors aim to invest in brown companies and then try to exert pressure on the management from inside the company. Akey and Appel (2020) showed that hedge fund activism targeting companies' environmental behavior is associated with a 17% drop in chemical emissions at the

[13] For a detailed literature review, see Koelbel et al. (2020).

[14] Bolton and Kacperczyk (2021) observed significant divestment transactions of institutional investors. Similarly, Boermans and Galema (2019) showed that a sample of Dutch pension funds have been actively decarbonizing their portfolios.

[15] Often, investors must push companies to even disclose their environmental practices (Cotter and Najah 2012; Flammer, Toffel, and Viswanathan 2019).

plants of targeted firms. The findings of Naaraayanan, Sachdeva, and Sharma (2021) supported these results.

Investors' impact channels are severely limited. For instance, when investors sell shares of brown companies to purchase the shares of green ones, they are not denying the brown company access to capital. Given that for every seller there is a buyer, these actions simply reallocate ownership. Only if the selling pressure is extremely high, putting downward pressure on the share price, will the brown company face a higher cost in raising new capital. In contrast, green firms benefit from access to cheaper capital. Because substantial reallocation is required to move the cost of capital, more-concentrated and higher-volume investor actions are necessary to achieve a strong impact. Further, GHG emissions are highly concentrated. According to CDP (2017), 71% of all global GHG emissions since 1988 can be traced to just 100 fossil fuel producers, whereas only 5% of the entire universe of emitters is responsible for 80% of emissions. Therefore, because the worst emitters are responsible for the lion's share of emissions and because investor actions quickly lose efficacy if applied in an unfocused manner, investors must have accurate emissions data.

Another important consideration is the type of information GHG emissions data are based on. Although emissions are concentrated in certain industries, investors need to distinguish between companies within a brown industry to identify the green(er) companies. When investors act on industry information alone, they indiscriminately penalize the whole industry and thus punish green(er) companies within the industry. This stifles the green projects in the brown sectors—the opposite of what investors desire. Consequently, companies may be discouraged from pursuing green investment if the investment is not rewarded—and may even be punished—by investors.

GHG DATA COVERAGE

Until now, GHG reporting for listed companies has been voluntary in most jurisdictions, and only a very few instances call for mandatory reporting. One of the most progressive countries is the United Kingdom, where over 1,300 of the largest companies and financial institutions will have to disclose climate-related financial information on a mandatory basis from April 2022 on.[16] In Europe, the EU Directive on Corporate Sustainability Reporting, published in April 2021, seeks to replace and expand existing ESG reporting obligations (EU Commission 2021). However, until the regulation becomes effective, GHG reporting remains voluntary. In contrast, little progress is evident in the United States, where the Securities and Exchange Commission (SEC) only declared its intention to revise the guidance on climate-related disclosures and is currently just at the point of collecting feedback from the public on how to regulate (SEC 2021a, 2021b; US Congress 2021).[17] In other jurisdictions like China, Japan, or emerging market countries, which account for a high proportion of the world's carbon emissions, GHG reporting is also still voluntary to date.

However, even if companies do report to regulators, data are often hard to access and not investor friendly.[18] The voluntary basis for reporting significantly lowers data coverage and introduces a potential self-selection bias (Matsumura, Prakash, and Vera-Muñoz 2014) causing companies to underreport their emissions and investors

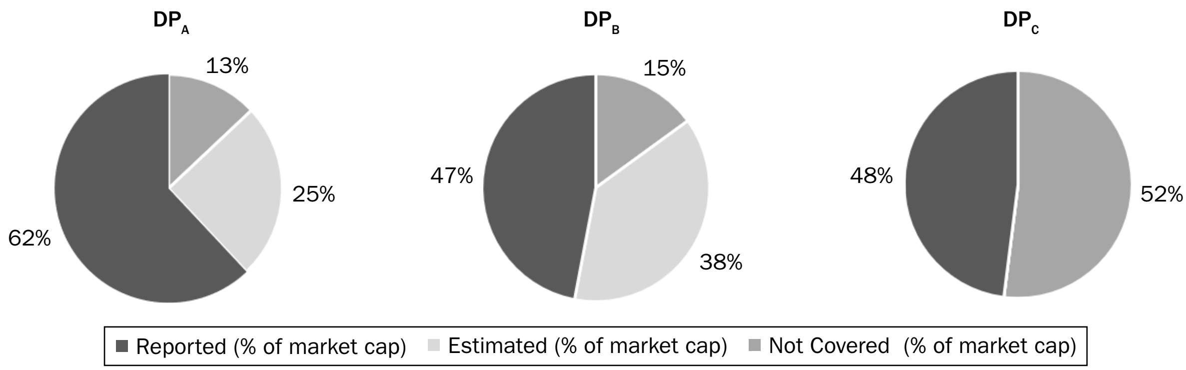
[16] https://www.gov.uk/government/news/uk-to-enshrine-mandatory-climate-disclosures-for-largest-companies-in-law.

[17] Comments on SEC's proposed climate-change disclosure are publicly available at https://www.sec.gov/comments/climate-disclosure/cll12.htm.

[18] For instance, companies that fall under the EU Emissions Trading System report their carbon emissions at plant level, not at firm level.

EXHIBIT 1

GHG Emissions: Comparison of Coverage by Market-Capitalization, 2010–2016



NOTES: Reported (% of market cap) represents the fraction of market capitalization that is on average covered with reported GHG data in the respective dataset. Estimated (% of market cap) represents the fraction of market capitalization that is on average covered with estimated GHG data in the respective dataset. Not covered (% of market cap) represents the fraction of market capitalization that is on average not covered with any GHG data in the respective dataset. The numbers reflect the time-series average of cross-sectional means for the period 2010–2016.

SOURCE: Research Affiliates, LLC and University of Augsburg, based on anonymized data from GHG emissions data providers.

to have undesired exposure to climate risk (Karagozoglu 2021). We argue that carbon emissions are the most useful for investors if they are widely available in the investment universe. To test this, we compare the GHG data coverage of different carbon data providers by: 1) market capitalization and 2) amount of carbon emissions.

Market-Capitalization Coverage

In Exhibit 1, we display the breakdown of the market-capitalization coverage of GHG data for the three data providers that provide carbon emissions for the seven-year period 2010–2016 (Appendix C provides more detail by year).

Overall, DP_A captures reported carbon emissions for around 62% of all listed companies by market capitalization. DP_B's and DP_C's coverage is significantly smaller at 47% and 48%, respectively. To complement the datasets, DP_A and DP_B use models to estimate emissions for 25% and 38% of the companies, respectively.[19] We do not have estimated data for DP_C.[20] Consequently, the remaining fractions of market capitalization not covered with GHG data range from 13% to 52%. The insufficient coverage of GHG data can be attributed to the voluntary reporting of climate-related information.

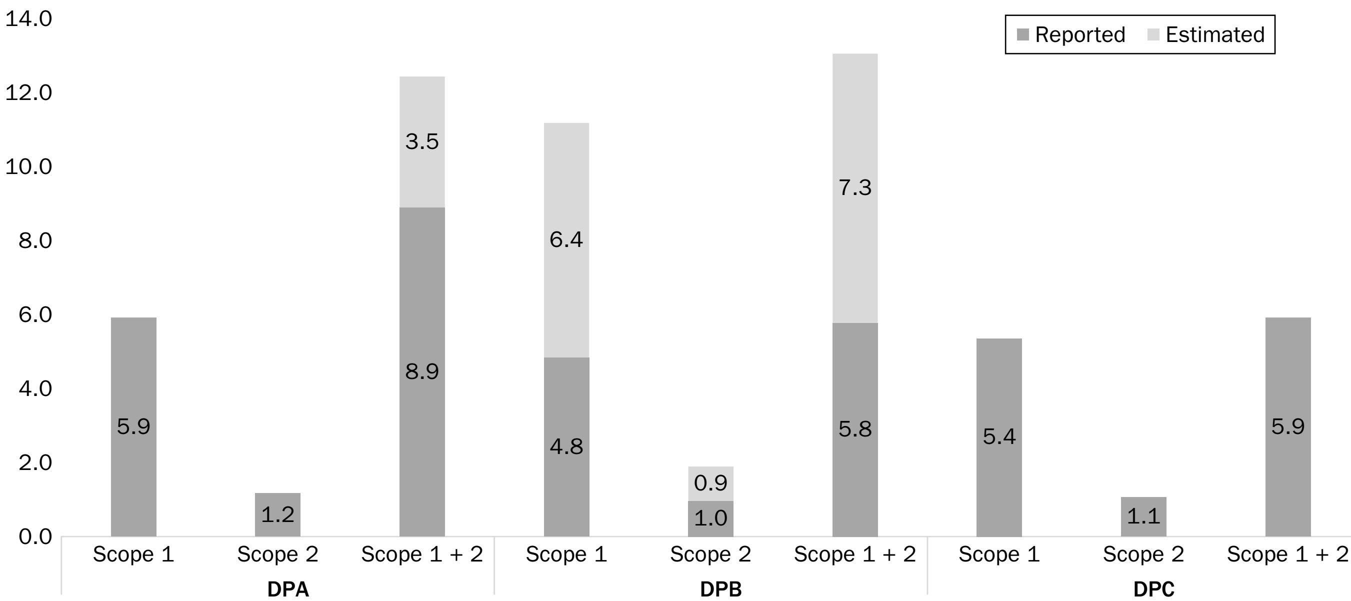
Carbon Emissions Coverage

For investors to achieve the greatest impact possible, GHG data must be available for all investments and especially for the largest emitters. Heavily emitting companies have the greatest potential for emission reductions. We compare the datasets by their respective carbon emissions coverage (in metric gigatons[21]). In Exhibit 2, we display the average annual amount of covered reported and estimated emissions by emission category (scope) for each data provider. We exclude from the exhibit those companies not covered with GHG data.

[19] We analyze the quality of the estimated emissions later in the article.
[20] DP_C began to estimate emissions after our period of analysis.
[21] A gigaton is 10^9 tons.

EXHIBIT 2

GHG Emissions: Comparison of Coverage by Dataset and Emission Category, 2010–2016



NOTES: This exhibit compares the total amount of covered carbon emissions by dataset and emission category. Reported represents the total amount of covered reported carbon emissions (in gigatons) and estimated represents the total amount of covered estimated carbon emissions (in gigatons). The numbers reflect the time-series average of annual total covered emissions for the period 2010–2016.

SOURCE: Research Affiliates, LLC and University of Augsburg based on anonymized data from GHG emissions data providers.

DP_A and DP_C cover a similar amount of reported Scope 1 emissions at 5.9 and 5.4 gigatons, respectively. Recall that DP_B separately estimates Scope 1 and Scope 2 emissions. Therefore, DP_B reports a much higher total covered direct emissions at 11.2 gigatons. The reported Scope 2 emission coverage is similar for all three datasets but slightly higher for DP_B after we include estimated emissions. Considering both emission categories together (Scope 1 and 2), DP_A displays the highest reported carbon emissions coverage (50% more than DP_C), whereas DP_C and DP_B show similar coverage of 5.9 and 5.8 gigatons, respectively. Finally, DP_B estimates the most Scope 1 and 2 emissions at 7.3 gigatons.

Next, we test whether each company identifies the same companies as the largest emitters based on the sum of all available Scope 1 and 2 emissions. We find that the 50 largest emitters of DP_A only overlap with 23 of the largest emitter of DP_C. This may be because DP_A estimates emissions for nonreporting companies, whereas DP_C does not. However, if we compare DP_A to DP_B, which both estimate emissions, we also only find an overlap of 34 out of 50. This suggests that an investor, who, for instance, follows a popular exclusion strategy based on the 50 largest overall emitters, will exclude 27 (16) different companies from his investable universe only by using different data providers. Obviously, the choice of database can have significant portfolio implications for green investors as well as for researchers.[22] For instance, some studies suggest that companies with higher emissions experience higher stock returns reflecting a carbon premium (e.g., Bolton and Kacperczyk 2022).

[22] Moreover, different coverages and the divergence of carbon emissions (especially for estimated emissions) across data providers may reduce the informative power of empirical research because the choice of the database may alter the findings and the conclusions. For instance, there are mixed signals from studies analyzing the performance of carbon-intensive stocks (e.g., Trinks et al. 2018; Cheema-Fox et al. 2019; In, Park, and Monk 2019; Bolton and Kacperczyk 2022; Goergen et al. 2020), which may at least partially be attributable to different GHG datasets.

However, this relationship is primarily attributable to emissions estimated by data providers, as opposed to emissions disclosed by the companies themselves (Aswani, Raghunandan, and Rajgopal 2021).

COMPARABILITY BETWEEN COMPANIES

No universally accepted reporting standard for GHG emissions exists. The TCFD was founded to provide guidance on which information is relevant to investors, lenders, insurers, and other stakeholders and how the information should be disclosed by reporting companies.[23] In terms of the measurement and disclosure of carbon emissions, the TCFD refers to the GHG Protocol, which for many years has been one of the most commonly followed standards, although other reporting standards exist (e.g., US Environmental Protection Agency [EPA] Center for Corporate Climate Leadership GHG Inventory Guidance, UK Environmental Reporting Guidelines, and China National Development and Reform Commission (NDRC) GHG accounting and reporting guidelines).[24] The ability to choose the reporting standard potentially introduces a self-serving bias in the reported data and hinders comparability between companies. In 2018, only 33% of companies reported GHG emissions in line with TCFD recommendations (TCFD 2019). Furthermore, the academic ClimateDisclosure100.info initiative recognizes only 21 firms worldwide to have reported 100% of their Scope 1 GHG emissions.[25]

Finally, the standards themselves often give companies significant leeway in measuring and disclosing their emissions. For instance, the GHG Protocol allows for reporting carbon emissions using either the equity share or the financial control approach.[26] As a result, reported carbon emissions can significantly differ based on the approach used.

CONSISTENCY ACROSS DATA PROVIDERS

Most investors obtain their carbon data from a specialized carbon data provider. In addition to different data coverage, each data provider has unique features, which can potentially lead to different investment decisions. For instance, data providers deal differently with corporate events, such as mergers and acquisitions. Some adjust carbon emissions for corporate actions (e.g., DP_A and DP_B claim to do so), whereas others do not (DP_C). Further, some data providers correct obvious reporting errors (e.g., use of the wrong unit), whereas others do not.

To illustrate why the database matters, we can consider the following example. As mentioned earlier, Scope 2 emissions quantify the emissions from the generation of acquired and consumed electricity, heat, or cooling. The GHG Protocol allows companies to use the market-based or location-based approach to calculate their Scope 2 emissions.[27] Often, companies report emissions compliant to both

[23] https://www.fsb-tcfd.org/.

[24] For the US EPA Center for Corporate Climate Leadership GHG Inventory Guidance, see https://www.epa.gov/climateleadership/center-corporate-climate-leadership-greenhouse-gas-inventory-guidance. For the UK Environmental Reporting Guidelines, see: https://www.gov.uk/government/publications/environmental-reporting-guidelines-including-mandatory-greenhouse-gas-emissions-reporting-guidance. For the China NDRC GHG accounting and reporting guidelines, see: https://www.wri.org/our-work/top-outcome/china-moves-toward-mandatory-corporate-ghg-reporting.

[25] https://web.archive.org/web/20210622092028/https://climatedisclosure100.info/.

[26] Under the equity share approach, a company accounts for GHG emissions from operations according to its share of equity in the operation. Under the control approach, a company accounts for 100% of the GHG emissions from operations over which it has control (WBCSD and WRI 2015b).

[27] The location-based method quantifies Scope 2 emissions based on average energy-generation emission factors for defined geographic locations. The market-based method quantifies Scope 2 emissions based on GHG emissions emitted by the generators from whom the reporting company contractually purchases electricity (WBCSD and WRI 2015a).

EXHIBIT 3

Correlation of Reported and Estimated GHG Emissions between Data Providers, 2010–2016

	DP_A	DP_B	DP_C		DP_A	DP_B	DP_C
Panel A: Reported Scope 1 Emissions				**Panel B: Reported Scope 2 Emissions**			
DP_A	1			DP_A	1		
DP_B	0.987	1		DP_B	0.976	1	
DP_C	0.986	0.995	1	DP_C	0.979	0.993	1
Panel C: Reported Scope 1 + 2 Emissions				**Panel D: Estimated Scope 1 + 2 Emissions**			
DP_A	1			DP_A	1		
DP_B	0.984	1		DP_B	0.847	1	
DP_C	0.984	0.996	1	DP_C	–	–	1

NOTE: This exhibit shows rank correlations between various emission categories and between carbon data sets for the period 2010–2016. The number of observations varies within and across each panel.

SOURCE: Research Affiliates, LLC, and University of Augsburg based on anonymized data from GHG emissions data providers.

approaches, but some data providers report only one approach to investors.[28] The choice of reporting appears arbitrary and leads to considerable differences in reported Scope 2 emissions between data providers. For some observations, we find differences in reported Scope 2 of more than 30%.

Our access to multiple carbon datasets allows us to test whether reported carbon emissions data are consistent across data providers. Although consistency itself does not guarantee good quality data, it serves as an indirect proxy for data accuracy. To evaluate the consistency of emissions across datasets, we compare rank correlations of emission levels. Ranks are very important because many investment companies and initiatives encourage the exclusion of, for instance, the 200 most polluting companies from their portfolios (negative screening).

In Exhibit 3, we compare emissions pairwise from two data providers if both cover the respective company in the respective year. In addition, we distinguish between the Scope of emissions and whether the emissions were self-reported by the companies or estimated by the data providers (Panel D).

Exhibit 3 suggests that the ranks are relatively consistent across GHG datasets for the reported data: The rank correlations are almost 1 between DP_B and DP_C across all emission categories. The rank correlations of Scope 2 emissions between DP_A and DP_C and between DP_A and DP_B are lower due to their different treatments of Scope 2 emissions (location-based versus market-based). Estimated carbon emissions are more inconsistent across data providers (0.847) than reported emissions and can be attributed to the differences in their estimation models. Busch et al. (2018) supported our findings as they observed similar rank correlations.[29]

Several studies have already shown that data providers can come to very different outcomes when producing ESG ratings (Dimson, Marsh, and Staunton 2020; Avramov et al. 2022; Brandon, Philipp, and Schmidt 2021). According to Berg, Koelbel, and Rigobon (2020) the differences in ESG ratings are based on three different sources. *Weight divergence* occurs when rating agencies hold different views on the relative importance of attributes when aggregating multiple indicators. *Scope divergence* refers to the situation in which different sets of attributes are used as a basis for

[28] Our analysis suggests that DP_A is more likely to report Scope 2 emissions in accordance with the market-based approach as opposed to DP_B, which uses the location-based approach.

[29] In Appendix D, we conduct another analysis in which we calculate percentage deviations between the data providers. In this analysis, we find some inconsistencies, but they do not lead to significant changes in rank.

forming ratings. *Measurement divergence* occurs when rating agencies measure the same attribute using different indicators.

Analogous to ESG ratings, the differences in carbon emission estimates may be attributable to the same sources. First, data providers use different approaches to obtain estimates. For instance, DP_B estimates carbon emissions in one step using a simple regression model and thus weights all attributes equally, whereas DP_A uses a multilevel estimation procedure with different weights on the attributes (weight divergence). Second, data providers may use different indicators to describe the same attribute. For instance, the size of a company can be proxied by either its net sales or its number of employees (scope divergence). Third, data providers often choose different attributes in estimating carbon emissions. For example, DP_A uses the attribute *energy consumption* in its estimation procedure, whereas DP_B does not use this attribute at all (measurement divergence). Consequently, subjective ESG ratings and estimated carbon emissions may have the same sources of heterogeneity.

In summary, reported data are largely consistent across data providers. Estimated data exhibit some inconsistencies between data providers. The information content of estimates deserves careful examination.

PREDICTIVE POWER OF FORWARD-LOOKING INFORMATION

Investors who desire to mitigate climate change are particularly interested in which companies plan to reduce their future emissions. Access to a company's forward-looking information is therefore precious. Data providers have also started to supply data with information on how companies are expected to change their carbon-emitting behavior in the future (e.g., based on corporate emissions reduction targets). This information is not supplied to investors in terms of future estimates of carbon emissions (in metric gigatons) but rather through a score or rating.[30] We examine whether these data are helpful for investors in predicting future levels and changes in emissions and whether carbon scores with forward-looking information have any predictive power, as often claimed by data providers (e.g., emission reduction score). Appendix A describes the carbon scores we use in our analysis.

Predictability of Future Reported Emission Levels

We expect carbon emissions levels to be persistent over time because changes in emissions are not readily observable. We test the expectation that future emissions levels are predictable by running regressions on future reported carbon emissions, which are explained by simple fixed effects in predictive Models (1) and (2) and by historical emissions (autoregressive model) in predictive Models (3) and (4). After calibrating the regression models, we study how well investors can forecast future emission levels by examining rank correlations between model-predicted future emissions and real reported future emissions. In Exhibit 4, we display the results.

The adjusted R^2 of the regression using both fixed effects and historical emissions is 97.2%; that is, we can almost perfectly predict future emissions by a simple regression model. In addition, the rank correlation between predicted future emissions and reported future emissions is very high at 98.7%. Overall, our results suggest that carbon emissions are stable from one year to another and that investors are well served by taking the latest reported emissions as the proxy for next year's emissions. In this case, forward-looking carbon scores from the data providers are not needed.

[30] Data providers condense corporate information into carbon scores or carbon ratings but do not estimate the amount of future carbon emissions.

EXHIBIT 4

Predictability of Future Reported Carbon Emission Levels, 2010–2016

	(1) Log(Emissions$_{t+1}$)	(2) Log(Emissions$_{t+1}$)	(3) Log(Emissions$_{t+1}$)	(4) Log(Emissions$_{t+1}$)
Log(Emissions$_t$)			0.98**	0.94**
Constant	14.37**	14.79**	0.27**	0.77**
	(27.55)	(17.45)	(8.15)	(6.27)
Industry fixed effects	Yes	Yes	No	Yes
Year fixed effects	No	Yes	No	Yes
Country fixed effects	No	Yes	No	Yes
Number of observations	11,984	11,984	11,984	11,984
Adjusted R^2	60.4%	65.4%	97.1%	97.2%
Rank correlation between predicted future emissions and real reported future emissions	78.0%	81.0%	98.7%	98.7%

NOTES: This exhibit reports regression results of past reported carbon emissions on real reported future carbon emissions (scope 1 and 2). All variables are logarithmized due to their skewness. Emissions reflect a company's Scope 1 and 2 emissions. Industry fixed effects are based on a company's Thomson Reuters Business Classification (TRBC) activity code. Country fixed effects are based on a country's International Organization for Standardization (ISO) 3166 code. ** and * indicate significance at the 1% and 5% level, respectively. Standard errors are clustered at the company level. The values in the parentheses are *t*-statistics.

SOURCE: Research Affiliates, LLC and University of Augsburg based on anonymized data from GHG emissions data providers.

Predictability of Future Reported Emission Changes Using Forward-Looking Information

Next, we test whether future changes in emissions are predictable and whether carbon scores from the data providers have explanatory power in predicting future changes in emissions.[31] Again, we use a regression approach, as shown in Exhibit 5, to explain future changes in emissions by past changes (autoregressive model) and to add provider-specific carbon scores[32] with forward-looking information to the model.

Independent of the model specification, predicting future changes in emissions is difficult as indicated by a low (negative) adjusted R^2, mainly due to the random character of changes in emissions. Even data provider–specific carbon emission scores with forward-looking content cannot overcome this problem. The rank correlation between model-predicted future changes in emissions and real reported future emissions changes for any model specification is very low (below 8%). We repeat the analysis focusing on future changes in carbon intensities (Appendix F) and get similar results.

We conclude that future changes in carbon emissions are difficult to forecast. We find no evidence that the forward-looking scores from the data providers carry any useful forecasting information. We believe the data providers are doing an honest job collecting data from companies. The lack of predictability is likely driven by the use of nonscientifically verified estimation methods and by cheap talk from companies engaged in greenwashing. We argue that forward-looking information should be externally verified before being published. Until scientific verification demonstrates that a specific type of forward-looking data is predictive of future emissions changes, investors would be prudent to operate as if these data contain no useful forward-looking information.

[31] In Appendix E, we examine whether changes in carbon emissions (Panel A) and carbon intensities (Panel B) are persistent over time. We find the likelihood of last year's trend (e.g., strong emissions reduction) continuing in the future is low.

[32] Appendix A gives an overview of carbon scores with forward-looking information.

EXHIBIT 5

Predictability of Future Reported Carbon Emission Changes, 2010–2016

	Δ% Emissions$_{t+1,t}$	Δ% Emissions$_{t+1,t}$	Δ% Emissions$_{t+1,t}$	Δ% Emissions$_{t+1,t}$	Δ% Emissions$_{t+1,t}$	Δ% Emissions$_{t+1,t}$
Δ%Emissions$_{t,t-1}$		–0.04*	–0.04*	–0.04*	–0.04*	–0.04*
DP_C A list$_t$			–0.09			–0.02
DP_A emission reduction score$_t$				–0.00		–0.00
DP_D carbon emission score$_t$					–0.08	–0.07
Constant	0.72	0.74	0.74	1.02*	0.94	1.04*
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	5,917	5,917	5,917	5,917	5,917	5,917
R^2	3.6%	3.7%	3.7%	3.7%	3.8%	3.8%
Adjusted R^2	–3.0%	–2.9%	–2.9%	–2.9%	–2.9%	–2.9%
Rank correlation between predicted changes in future emissions and real reported changes in future emissions	6.6%	7.0%	7.2%	6.6%	7.0%	6.9%

NOTES: This exhibit reports regression results of past reported carbon emission changes on real reported future carbon emission changes (scope 1 and 2). Δ% Emissions$_{t+1,t}$ reflects the percentage change of a company's Scope 1 and 2 emissions from one year to another. DP_C A list reflects a binary variable, which specifies whether a company was in the DP_C A list (highest rated category) in the respective year. Industry fixed effects are based on a company's TRBC activity code. Country fixed effects are based on a country's ISO 3166 code. ** and * indicate significance at the 1% and 5% level, respectively. Standard errors are clustered at the company level.

SOURCE: Research Affiliates, LLC and University of Augsburg based on anonymized data from GHG emissions data providers.

ACCURACY OF ESTIMATED EMISSIONS

The voluntary reporting of GHG emissions data results in only 48%–62% of companies being covered when based on market capitalization (see Exhibit 1). Data providers estimate the missing data but have limited visibility into companies' green efforts and activities. Nevertheless, data providers claim their estimated emissions are generated by highly sophisticated estimation models. These representations lead to a widespread misunderstanding that estimated emissions are of similar quality as reported emissions and can be used with equal confidence in decision making. In this section, we analyze the accuracy of estimated emissions.

We run several tests to assess the information content of the estimated data using reported data as the best available proxy for true emissions. Estimated emissions are by definition a noisy proxy of true emissions, which are unobserved. Data providers often base their estimates on broad business metrics and industry affiliations and, to our best knowledge, do not engage with companies in estimating emissions.

Calibrating Model Accuracy

We test how much power simple financial business metrics and fixed effects have in explaining reported carbon emissions. We use the model R^2—a statistical metric representing the percentage of total emissions explained by the model-independent variables—as a measure of model accuracy. A very high explanatory power (R^2 close to 100%) suggests that estimates properly reflect a company's true emissions. A lower R^2 indicates the estimates are not fully capturing all information contained in reported emissions, but even reported emissions are not a perfect proxy for a company's true emissions. Using estimated emissions, which can be at most as good as reported emissions (if R^2 is 100%), requires special caution.

We set up various specifications of regression models to test whether simple business metrics and fixed effects have any power in explaining reported emissions. We use DP_C's emissions in the model calibration as the dependent variable because DP_C collects the bulk of the reported data firsthand and the other data providers use these data as the main source for their own datasets. We use quite simple financial indicators, such as the logarithm of net sales (to proxy for the size of a company's output), as independent variables. Next, we use a company's industry affiliation, country of domicile, and the year as fixed effects (Appendix G shows that carbon emissions differ highly across countries and industries). In Appendix H, we provide a detailed definition and rationale for each variable we use. We take logs of carbon emissions and business metrics due to their high skewness. We use an ordinary least squares regression to calibrate the models. Equation 1 is our baseline regression model.

$$\text{Log}(\text{Carbon emissions}_{i,t}^{Reported}) = \alpha + \sum_{j}^{n} \beta_j \text{Log}(\text{Business metric}_{i,t}) + \sum_{k}^{m} \beta_k \text{Fixed and time effects}_{i,t} + \varepsilon_{i,t} \quad (1)$$

We display the model calibration results in Exhibit 6. Models (1) through (4) show that company size and industry explain most of the company variations in reported emissions. Specifically, Model (1), which uses the logarithm of net sales as the proxy for company size, explains around half of the emission variation; the model-adjusted R^2 is 45.7%. Similarly, Model (2), which uses only industry fixed effects, has an adjusted R^2 of 54.2%. Combining the measure of size, industry, and other fixed effects in Model (4), the adjusted R^2 rises to 83.8%.

In Model (5), we add a few more company-specific financial ratios as explanatory variables. The most important variables, as reflected by their *t*-stats, are 1) property, plant, and equipment (PP&E) value scaled by net sales (Log(PP&E/Net Sales (NS))), which proxies for the equipment intensiveness of production and 2) the number of employees scaled by net sales (Log(Employees/NS)), which measures the employee intensiveness of production. Both variables positively predict the variations in emissions, suggesting that the more-innovative and greener companies are also less well equipped and less employee intensive.

Model (5) captures some information beyond the broad correlates: The adjusted R^2 increases from 83.8% for Model (4) to 87.0% for Model (5). If we compare our inputs to the regression model with the data providers' inputs to their estimation models, we find we typically include more or at least equal information compared to the data providers. Consequently, we can assume the data providers are unlikely to explain the variation of reported emissions better than we do with Model (5).

We conclude that industry and size characteristics play a major role in explaining carbon emissions. Models incorporating this type of information as well as information on a company's efficiency have very good explanatory power with an R^2 around 87%. We do not observe perfect explanatory power (R^2 is 100%). Importantly, this means that reported emissions contain other information not captured by these high-level business metrics.

Green investors are likely more interested in the fraction of emissions variation (residual of regression) not explained by the broad correlates (87% of the variation) because the residual information is not observable from the outside. This hidden information includes, for instance, a company that uses green versus conventional electricity. Estimated emissions, solely based on high-level business metrics and

EXHIBIT 6

Calibration of Model Accuracy Using Reported Carbon Emissions, 2010–2016

	(1) Log(Emissions$_t$)	(2) Log(Emissions$_t$)	(3) Log(Emissions$_t$)	(4) Log(Emissions$_t$)	(5) Log(Emissions$_t$)
Log(Net sales$_t$)	1.06**			1.10**	1.11**
Log(Employees/NS$_t$)					0.29**
Log(Market Cap/NS$_t$)					–0.18**
Log(EBT/NS$_t$)					–0.03
Log(FFO/NS$_t$)					0.28**
Log(OI/NS$_t$)					–0.08
Log(PP&E/NS$_t$)					0.43**
Log(COGS/NS$_t$)					0.18**
Constant	–10.96**	13.99**	13.94**	–9.94**	–6.47**
Industry fixed effects	No	Yes	Yes	Yes	Yes
Year fixed effects	No	No	Yes	Yes	Yes
Country fixed effects	No	No	Yes	Yes	Yes
Number of observations	7,227	7,227	7,227	7,227	7,227
Adjusted R^2	45.7%	54.2%	61.2%	83.8%	87.0%

NOTES: This exhibit displays panel regression results of various company characteristics on carbon emissions (scope 1 and 2). All variables are logarithmized due to their skewness. Emissions reflect a company's Scope 1 and 2 emission reported to DP_C for the period 2010–2016. The independent variables are described in Appendix H. Industry fixed effects are based on a company's TRBC activity code. Country fixed effects are based on a country's ISO 3166 code. ** and * indicate significance at the 1% and 5% level, respectively. Standard errors are clustered at the company level.

SOURCE: Research Affiliates, LLC and University of Augsburg based on anonymized data from GHG emissions data providers.

widely used in the financial industry, are not likely to contain new, relevant information for green investors. In the following section, we illustrate how using estimated emissions based on an R^2 lower than 100% can affect the impact of investor actions.

Model Accuracy and Investor Ability to Identify the Worst Emitters

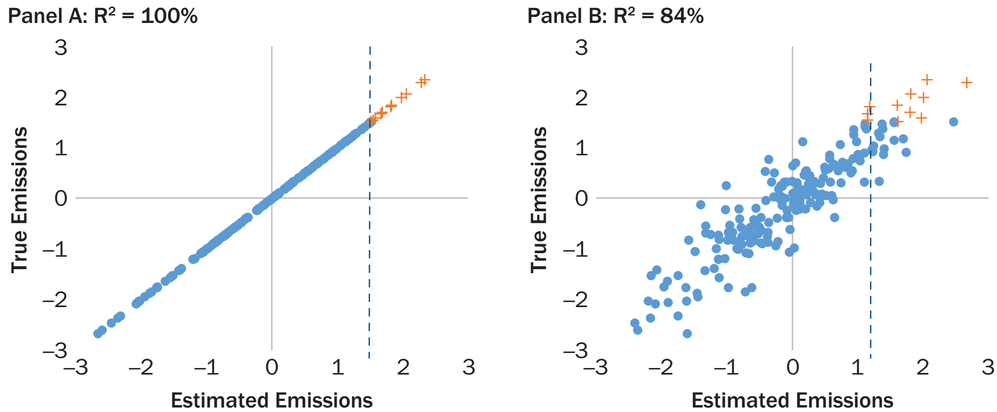
Carbon emissions are highly concentrated. Only 5% of companies are responsible for 80% of total emissions. How precisely investors can identify these 5% largest emitters using noisy estimated data?

To understand the exercise, consider the following example. Suppose an investor wants to avoid investing in the 5% largest emitters in a universe of 10,000 companies and solely uses estimated data to identify these 5% of heavy-emitting companies. With perfect data (reported or estimated with an R^2 of 100%), an investor would be able to exactly identify the 500 worst companies (100% confidence). Panel A of Exhibit 7 illustrates a simulated distribution of true and estimated emissions with an R^2 of 100%. In this case, estimated emissions allow the perfect identification of all 5% of the worst emitters (data points marked with crosses). If the data are noisy and do not perfectly reflect a company's true emissions (R^2 below 100%), an investor would need to exclude more companies from their portfolio in order to ensure all of the 500 largest emitters are excluded.

To show the effects of using estimated data, we perform a Monte Carlo simulation. We assume an R^2 of 84% as in Model (4) and calculate that an investor would need to exclude 1,250 companies to remove at least 95% of the largest emitters. In Panel B of Exhibit 7, to the right of the dashed line, the dots represent the less-severe emitters together with the crosses that represent the 5% worst emitters. In this case, the number of noninvestable stocks increases by 2.5 relative to the 500 stocks that would be excluded with perfect insight, reducing the effectiveness of the investor

EXHIBIT 7

Simulated Distribution of True (Unobservable) and Estimated Emissions For Different Levels of R^2

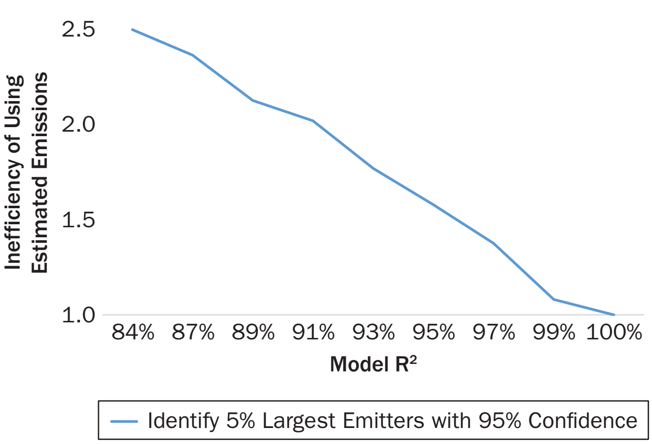


NOTES: This exhibit shows the simulated distribution of normally randomly distributed true (unobservable) and estimated (observable) emissions. We mark with crosses the 5% worst emitters based on the true emissions. We also mark with the dashed line the cutoff value of the estimated emissions necessary to select at least 95% of the 5% worst emitters based on the estimated emissions.

SOURCE: Research Affiliates, LLC and University of Augsburg.

EXHIBIT 8

Inefficiency of Using Estimated Emissions to Target Largest Carbon Emitters



NOTES: This exhibit shows the results of Monte Carlo simulations with a sample size of 10,000 using normally distributed random variables in all parts of the simulations. For a model with a given R^2, we estimate the size of the sample sufficiently large to identify the 5% largest emitters with 95% confidence.

SOURCE: Research Affiliates, LLC and University of Augsburg.

action by 2.5 times (i.e., the investor does not invest in 1,250 companies versus 500). We call 2.5 the inefficiency of using estimated emissions in comparison to reported emissions. We repeat our simulation for various levels of R^2 and display the results in Exhibit 8.

Compared to Model (4), using Model (5) with an R^2 of 87%, the estimates are not much better. An investor would need to pick about 2.4 times (compared to 2.5 times) as many stocks to exclude the 5% worst stocks, about 1,190 versus 500 using reported data. The higher R^2 of Model (5) lowers the inefficiency of the investor action but only minimally. A model R^2 of 98% or higher is likely the point at which the multipliers begin to converge to 100%, and investors would need a sample of no more than 1.2 times to target the 5% worst emitting companies.

Accuracy and Information Content of Estimated Emissions by Data Providers

To further test the accuracy of estimated data, we use a unique subset of companies to run the horse race between the data of the two data providers DP_A and DP_B and the estimates obtained from Models (4) and (5). This subset of companies began self-reporting after a period of nonreporting, during which the data providers had estimated the companies' emissions. Given the high persistence of emissions as described earlier, the subsequently reported data are an excellent proxy for historical emissions.[33]

[33] We expect some variation between estimated and reported emissions because of the one-year time lag between the estimated (t) and the reported emissions ($t + 1$). Therefore, we focus on correlations rather than absolute deviations.

EXHIBIT 9

Accuracy of Estimated and Model-Predicted Emissions, 2010–2016

	N	Rank Correlations	Log-Level Correlations
Panel A: Data Provider Switches from Estimated to Reporting Emissions at the Same Time			
DP_A	100	0.84	0.79
DP_B	100	0.78	0.75
Model (4)	100	0.83	0.79
Model (5)	100	0.87	0.84
Panel B: Data Provider–Specific Accuracy			
DP_A	860	0.69	0.67
DP_B	1,479	0.79	0.79
Model (4)	1,900	0.78	0.75
Model (5)	1,900	0.83	0.80

NOTES: This exhibit reports both rank correlations and log-level correlations between data providers' estimates on carbon emissions (or predicted values) (in *t*) and company-reported emissions (in *t* + 1) for the period 2010–2016. Panel A reports correlations for observations when DP_A and DP_B simultaneously switched from estimated emissions (in *t*) to reported emissions (in *t* + 1) in their datasets. Panel B ignores other databases and reports correlations for all observations in which a switch occurred from estimated to reported emissions in the respective dataset.

SOURCE: Research Affiliates, LLC and University of Augsburg based on anonymized data from GHG emissions data providers.

EXHIBIT 10

Cross-Rank Correlations of Estimated Data vs. Model Predictions, 2010–2016

Rank Correlations	DP_A	DP_B	Model (4)	Model (5)
DP_A	1.00			
DP_B	0.85	1.00		
Model (4)	0.78	0.82	1.00	
Model (5)	0.82	0.86	0.95	1.00

NOTES: This exhibit reports rank correlations between data providers' estimates on carbon emissions and the predicted values from the regression models. We also analyze log-level correlations and find similar results, which are available upon request.

SOURCE: Research Affiliates, LLC and University of Augsburg based on anonymized data from GHG emissions data providers.

We argue that if one data provider has superior information compared to the other, its estimates would be more highly correlated to the (later) reported emissions. To test this expectation, we compare correlations between the last known estimated emissions by the data providers and the first reported emissions by the companies. In Exhibit 9, we report the rank correlations and log-level correlations between the different types of estimates. Because estimating emissions may be harder for smaller companies due to less-available information, we show results for two different samples. Panel A shows correlations for observations in which both data providers switch from estimated to reported emissions at the same time (i.e., we compare the same companies). Panel B relaxes this condition and shows results for all observations within a dataset. As the emissions of DP_C are used to calibrate the models, we benchmark Model (4) and Model (5) predictions against reported emissions from DP_A and DP_B.

All models have roughly similar correlations. Relative to Model (4), the estimates from data providers on average display similar, and in rare cases, higher correlation levels with subsequently reported emissions. Relative to Model (5), they display correlations a notch lower in all cases, but the difference is small.

We cannot directly translate the correlations with the subsequently reported emissions into the model R^2 for the current emissions, which serves as our measure of model accuracy. We have, however, estimated the R^2 for Models (4) and (5) at 83.7% and 87.0%, respectively. Given that the correlations of data from DP_A and DP_B are on average similar to the correlations of Model (4) and lower than for Model (5), we can conclude that the equivalent R^2 should be on par with 83.7% and lower than 87.0%. Thus, the estimates from data providers are likely to capture similar information to the inputs of our regression models, strengthening our conclusion that estimates from the data providers are at least 2.4 times less effective in identifying the worst emitters compared to the reported data.

Estimate consistency. We also test whether our model predictions from Model (4) and Model (5) capture similar information to the estimates by the data providers by comparing rank correlations, displayed in Exhibit 10.

The rank correlation of 0.85 between the DP_A and DP_B estimates is a notch higher than the Model (4) correlations, 0.78 and 0.82, with data from DP_A and DP_B, respectively. The Model (5) correlation with DP_B estimated data is 0.86, a notch higher than 0.85; the correlation with DP_A estimated data is 0.82, a notch lower than 0.85. Overall, the correlations between the estimates from data providers and Model (5) are similar when compared to our model predictions. The fact that we observe similar correlations between the DP_A and DP_B estimates versus their correlations with Model (5), combined with the fact that Model (5) scores quite high in terms of accuracy, suggest that the Model

EXHIBIT 11

Information Content in Estimated Data, 2010–2016

	Model (2)		Model (4)		Model (5)	
	DP_A Log(Emissions$_t$)	DP_B Log(Emissions$_t$)	DP_A Log(Emissions$_t$)	DP_B Log(Emissions$_t$)	DP_A Log(Emissions$_t$)	DP_B Log(Emissions$_t$)
Log(Net sales$_t$)			0.95**	0.95**	0.98**	0.97**
Log(Employees/NS$_t$)					0.33**	0.24**
Log(Market Cap/NS$_t$)					–0.01	–0.03**
Log(EBT/NS$_t$)					0.01	–0.00
Log(FFO/NS$_t$)					0.03	0.05**
Log(OI/NS$_t$)					–0.09**	–0.03**
Log(PP&E/NS$_t$)					0.08**	0.21**
Log(COGS/NS$_t$)					–0.04*	0.06**
Constant	15.10**	13.73**	–5.85**	–5.86**	–2.78**	–3.54**
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	No	Yes	Yes	Yes	Yes
Country fixed effects	No	No	Yes	Yes	Yes	Yes
Observations	8,665	19,365	8,665	19,365	8,665	19,365
Adjusted R^2	64.9%	55.6%	87.7%	87.6%	89.2%	89.6%

NOTES: This exhibit displays panel regression results of various company characteristics on carbon emissions (scope 1 and 2). All variables are logarithmized due to their skewness. Emissions reflect a company's Scope 1 and 2 emissions. Column DP_A reports the estimates from the DP_A dataset, whereas column DP_B reports the estimates from the DP_B dataset. *Models* refer to Exhibit 9. The independent variables are described in Appendix H. Industry fixed effects are based on a company's TRBC activity code. Country fixed effects are based on a country's ISO 3166 code. ** and * indicate significance at the 1% and 5% level, respectively. Standard errors are clustered at the company level.

SOURCE: Research Affiliates, LLC and University of Augsburg based on anonymized data from GHG emissions data providers.

(5) estimates could serve as the simple and transparent substitute for estimates from data providers.

Information content of the estimates. In addition to the need for accuracy, emissions data should capture information beyond simple correlates, such as industry and size. Failure to capture this information stops investors from identifying the green companies in brown sectors and leads to counterproductive investor actions. By construction, estimates from Models (4) and (5) almost exclusively rely on information about company size and industry as well as a few other simple correlates; Model (4) uses only company size and industry, implying the estimates from these models will not help investors identify the green companies in brown sectors. Will the estimates from DP_A and DP_B be more successful in this regard?

To analyze the information content of the estimates from DP_A and DP_B, we use the same regression, Equation 1, as in Models (4) and (5), but in Equation 2, we use estimated emissions instead of reported emissions. In Exhibit 11, we provide the results of the regressions.

$$\text{Log(Carbon emissions}_{i,t}^{Estimated}) = \alpha + \sum_j^n \beta_j \text{Log(Business metric}_{i,t}) + \sum_k^m \beta_k \text{Fixed effects}_{i,t} + \varepsilon_i \quad (2)$$

Company size, as well as industry, year, and country fixed effects, capture most of the variations in estimates. The specification of the model that captures only industry fixed effects, Model (2), explains 64.9% (55.6%) of DP_A's (DP_B's) estimated data variation. As reported in Exhibit 6, industry fixed effects explain 54.2% of the reported data variation, the adjusted R^2 of Model (2). Because the adjusted R^2 is significantly higher for the DP_A estimates, DP_A seems to rely more heavily on simple industry affiliations in estimating carbon emissions. The specification of Model (4), which captures the company size and fixed effects (largely capturing the between-industry variation), draws a similar picture. The fraction of explained variation is higher using estimated emissions as the dependent variable as opposed to using reported emissions.

When we add the Model (5) extra variables, which mostly capture the employee and equipment intensity of production, the adjusted R^2 rises to near 90% for both data providers. This estimate is higher than the Model (4) estimate of approximately 88%. The increase in the R^2 is good news, indicating these datasets do capture some, albeit very little, information beyond the general correlates.

More generally, the simple correlates, such as industry and size effects, explain the bulk of the variation in the estimated data. These models' accuracy is similar to Model (4) and is lower compared to Model (5). Given that both Models (4) and (5) almost exclusively rely on and therefore reflect industry and size information, the implication is that the DP_A and DP_B estimates are equally unlikely to help investors differentiate the green companies in brown sectors, further reducing the efficacy of investor climate-change mitigation activity. Overall, our results suggest that estimated emissions provide valuable information to green investors because these data contain information helpful in differentiating high from low emitters, even if the estimations are quite crude and largely capture basic company characteristics.[34] Our findings also suggest significant superiority of the reported data compared to the estimated data.

CONCLUSION AND POLICY IMPLICATIONS

Investors play a major role in combating climate change. To conduct effective investor actions and to have an impact on the real economy, high-quality GHG data are needed. We investigate the quality of the currently available GHG data from an investor's perspective.

We argue five criteria are necessary to help investors successfully mitigate climate change: 1) GHG data need to be widely available, 2) GHG data should be comparable between companies, 3) GHG data should be consistent across data providers, 4) forward-looking GHG information should have predictive power, and 5) GHG data should accurately reflect true company emissions.

We show that about half of the current emissions data are reported directly by companies. The concerns about the quality of these data include: 1) reporting is voluntary, which lowers data availability and introduces a potential self-reporting bias; 2) no single reporting standard has been adopted, which leads to incomparability of GHG emissions between companies; and 3) reported data are not perfectly consistent across data providers. Despite these drawbacks, the reported data are the best quality information currently available.

In addition to historical and current emissions, we also analyze data provider–specific carbon ratings and scores, which claim to capture forward-looking information. To be valuable to investors, these carbon ratings and scores should be able to explain future changes in emissions, but we find they have no predictive power.

We find overall that GHG estimated data capture most of the variation of reported emissions. Conservative estimates suggest, however, that investor actions are at

[34] The models performing on par with the estimated emissions from data providers in explaining the subsequent emissions do explain the majority of the emissions as suggested by their relatively high R^2 of 84–87%.

least 2.4 times more diluted when investors use estimated emissions compared to reported emissions. Further, we show that estimated emissions data are based mainly on industry and size information, providing limited valuable and relevant information for investors and possibly hindering their desired outcome. We uncover the misconception that estimated emissions can be equally as useful to investors as reported emissions.

Our findings suggest that the status quo, in which GHG reporting is voluntary and data providers estimate the missing data, is inadequate. Our results have important implications for investors, companies, regulators, standard setters, data providers, and researchers.

Investors' efficiency in their efforts to mitigate climate change might be significantly reduced by basing investment decisions on noisy data. Investors can reduce inefficiencies by only using reported emissions and insist that companies disclose.[35] From a portfolio perspective, the choice of the GHG dataset can have major implications; for instance, each dataset implies a different investable universe in the case of an exclusion strategy.

Companies can contribute to a better relationship with investors and other stakeholders by voluntarily disclosing carbon emissions in an investor-friendly way. Otherwise, they may face fewer financing opportunities due to a smaller investor base.

Our findings suggest that regulators may improve data quality by introducing an international regulatory initiative of mandatory reporting of audited GHG data to investors and other stakeholders, preferably based on a single measurement and accounting standard (e.g., the GHG Protocol).

Standard setters should ensure that data are comparable between companies and sectors and that the standard provides no leeway for companies to greenwash their environmental behavior. In the interim, data providers should increase transparency of their estimation models and highlight the limitations of the data.

Finally, researchers should be aware that their empirical findings and conclusions may be altered by the choice of GHG data provider due to different coverages, inconsistencies in reported emissions across data providers, and differences in the estimation of emissions for nonreporting companies.

APPENDIX A

DESCRIPTION OF CARBON SCORES BY DATA PROVIDER

Data Provider	Name of Score	Description	Range	Scored Companies in 2016
DP_A	Emission reduction score	Measures a company's commitment and effectiveness in reducing environmental emissions in the production and operational processes.	0–100	5,968
DP_C	Climate change score	Measures the comprehensiveness of disclosure, awareness, and management of environmental risks and best practices associated with environmental leadership, such as setting ambitions and meaningful targets.	From A to D (eight distinct categories)	2,112
DP_D	Carbon emission score	Evaluates the extent to which companies may face increased costs linked to carbon pricing or regulatory caps. Scores are based on exposure to GHG-intensive businesses and emerging regulations; carbon reduction targets and mitigation programs; and carbon intensity over time and versus peers.	0–10	14,798

[35] One way investors can incentivize companies to report is to assume the highest possible emissions given the company's size and industry, as proposed in the United Nations (1992) precautionary principle.

APPENDIX B

EXHIBIT B1

Summary Statistics of Carbon Data And Company Characteristics by Dataset, 2010–2016

	DP_A			DP_B			DP_C			DP_D		
	No. Obs.	Mean	p50	No. Obs.	Mean	p50	No. Obs.	Mean	p50	No. Obs.	Mean	p50
Panel A: Carbon Data												
Reported Scope 1 emissions	12,629	3,274,450	71,410	11,135	3,034,741	59,606	12,194	3,074,414	58,576	–	–	–
Reported Scope 2 emissions	12,477	651,793	111,240	11,119	602,332	107,140	12,099	614,282	105,616	–	–	–
Reported Scope 1 + 2 emissions	15,708	3,962,170	254,989	11,119	3,633,120	240,055	11,986	3,449,618	231,695	–	–	–
Estimated Scope 1 emissions	–	–	–	49,991	889,359	12,748	–	–	–			
Estimated Scope 2 emissions	–	–	–	49,991	130,546	18,146	–	–	–			
Estimated Scope 1 + 2 emissions	17,101	1,452,074	48,181	49,991	1,019,905	38,852	–	–	–	–	–	–
Carbon intensity	32,142	12.35	0.41	10,256	182.13	0.43	10,864	229.89	0.43	–	–	–
DP_A emission reduction score	31,588	50.6	50.8	–	–	–	–	–	–	–	–	–
DP_D carbon emission score	–	–	–	–	–	–	–	–	–	52,074	7.3	7.6
Panel B: Company Characteristics												
Net sales (in US$ millions)	32,316	9,100	2,590	54,469	5,490	1,010	10,877	15,500	5,340	19,844	11,100	3,170
BTM	32,168	0.71	0.55	51,243	0.56	0.56	10,761	0.70	0.56	19,702	0.56	0.52
Employees	28,035	26,735	8,100	44,800	17,263	3,872	10,063	39,931	14,529	17,944	29,658	8,700
Total assets/Net sales	32,134	125.79	1.74	53,471	14.63	1.62	10,821	16.31	1.44	19,672	16.53	1.63
R&D expenses/Net sales	12,750	230.9%	1.8%	21,593	328.4%	2.1%	5,391	4.1%	2.0%	8,317	167.4%	2.0%
PP&E/Net sales	31,756	4.02	0.28	52,761	3.33	0.26	10,754	0.92	0.26	19,474	2.36	0.29
EBT/Net sales	32,124	–2.14	0.10	53,609	–2.10	0.09	10,820	0.04	0.09	19,671	–1.28	0.10
FFO/Net sales	32,128	–1.86	0.15	53,588	–1.39	0.13	10,820	0.43	0.13	19,664	–0.48	0.15
OI/Net sales	32,102	–2.47	0.12	53,549	–1.90	0.10	10,810	0.05	0.10	19,653	–1.02	0.12
COGS/Net sales	28,217	0.22	0.60	47,262	0.77	0.61	9,572	0.60	0.62	17,584	0.64	0.60

NOTES: This exhibit provides summary statistics of variables used in this study for the period 2010–2016. Panel A shows yearly carbon-related data, and Panel B lists yearly company characteristics. Carbon data are explained in the text in detail. Carbon intensity is displayed in (tons/$) scaled by 10,000. Company characteristics are obtained from Refinitiv Datastream. No. obs. reflect the number of company-year observations. BTM reflects the book-to-market ratio. R&D expense represents all direct and indirect costs related to the creation and development of new processes, techniques, applications, and products with commercial possibilities. PP&E represents gross property, plant, and equipment less accumulated reserves for depreciation, depletion, and amortization. Earnings before taxes (EBT) reflects the pretax margin. Funds from Operations (FFO) represents the sum of net income and all noncash charges or credits. Operating Income (OI) represents the difference between sales and total operating expenses. Cost of Goods Sold (COGS) reflects the cost of goods sold excluding depreciation.

SOURCE: Research Affiliates, LLC and University of Augsburg based on anonymized data from GHG emissions data providers.

APPENDIX C

EXHIBIT C1

Market-Capitalization Coverage by Year, 2010–2016

	DP_A				DP_B				DP_C			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Year	Number of Covered Companies	Reported (%-age of Market Cap)	Estimated (%-age of Market Cap)	Not Covered (%-age of Market Cap)	Number of Covered Companies	Reported (%-age of Market Cap)	Estimated (%-age of Market Cap)	Not Covered (%-age of Market Cap)	Number of Covered Companies	Reported (%-age of Market Cap)	Estimated (%-age of Market Cap)	Not Covered (%-age of Market Cap)
2010	4,012	59%	26%	15%	7,959	38%	43%	19%	1,479	45%	0%	55%
2011	4,155	65%	27%	8%	8,246	43%	45%	13%	1,584	50%	0%	50%
2012	4,249	63%	25%	12%	8,460	41%	44%	15%	1,754	50%	0%	50%
2013	4,356	61%	25%	14%	8,576	51%	34%	15%	1,758	48%	0%	52%
2014	4,462	59%	27%	14%	8,614	50%	35%	15%	1,795	48%	0%	52%
2015	5,322	61%	26%	13%	8,564	51%	34%	15%	1,812	47%	0%	53%
2016	6,216	63%	24%	13%	10,691	51%	36%	13%	1,804	48%	0%	52%
Mean	4,682	62%	25%	13%	8,730	47%	38%	15%	1,712	48%	0%	52%

NOTES: This exhibit compares the coverage of companies with GHG data across data providers and over time. Column (1) reflects the number of companies covered with GHG data. Column (2) reflects the fraction of total world market capitalization covered with reported GHG data. Column (3) reflects the fraction of total world market capitalization covered with data providers' estimates on carbon emissions. Column (4) reflects the fraction of world total market capitalization not covered with any GHG data. Data on world market capitalization of listed companies is obtained from the World Bank (data.worldbank.org/indicator/CM.MKT.LCAP.CD).

SOURCE: Research Affiliates, LLC and University of Augsburg based on anonymized data from GHG emissions data providers.

APPENDIX D

RELATIVE EMISSIONS DEVIATION

The rank correlations are important but may mask big deviations in individual observations. We dig deeper into the percentage deviation of emissions and compare pairwise relative emission deviations. For each company, we compute relative emission deviations in emission levels of observations for which we have GHG data from multiple data providers and in Exhibit D1 display the percentage of companies that fall into various deviation ranges. Equation D1 displays the calculation formula of relative emission deviation (*K* and *L* reflect the respective carbon data provider),

$$\text{Relative emission deviation}_{i,t,K,L} = \left| \frac{\text{Carbon emissions}_{i,t,K} - \text{Carbon emissions}_{i,t,L}}{\text{Carbon emissions}_{i,t,L}} \right| \quad \text{(D1)}$$

The relative deviation between DP_C's and DP_B's reported emissions is relatively low. In the reported Scope 1 and 2 levels, 86% of all observations deviate less than 1%, and 94% deviate less than 10%; however, there are a few outliers. About 3% of all observations deviate by more than 30%, and 1% deviate by more than 70%. The reported emissions between DP_A and DP_C and between DP_A and DP_B share more inconsistencies. For instance, in more than 24% of cases, Scope 1 and 2 carbon emissions differ by more than 10% between DP_A and DP_C.

In the case of estimated emissions, we identify even larger differences that may be masked by relatively large correlation levels. For DP_A and DP_B, only 10% of the sample deviates by less than 10%. For the vast majority of the sample (72%), the difference is

EXHIBIT D1

Percentage of Companies in Emissions Deviation Ranges by Emission Category

		Joint Observations	Relative Deviation (%)					
			|0–1%|	|1–5%|	|5–10%|	|10–30%|	|30–70%|	|>70%|
Panel A: Reported Scope 1 Emissions								
DP_A	DP_C	7,804	60%	12%	7%	11%	5%	5%
DP_A	DP_B	7,566	61%	12%	7%	10%	5%	5%
DP_C	DP_B	9,052	88%	3%	3%	3%	1%	1%
Panel B: Reported Scope 2 Emissions								
DP_A	DP_C	7,796	57%	12%	9%	12%	7%	4%
DP_A	DP_B	7,543	60%	11%	8%	10%	6%	5%
DP_C	DP_B	9,111	90%	2%	2%	2%	2%	1%
Panel C: Reported Scope 1 + 2 Emissions								
DP_A	DP_C	8,637	50%	16%	10%	13%	7%	4%
DP_A	DP_B	8,353	54%	15%	9%	12%	6%	4%
DP_C	DP_B	9,114	86%	5%	3%	3%	2%	1%
Panel D: Estimated Scope 1 + 2 Emissions								
DP_A	DP_B	12,452	1%	4%	5%	18%	25%	47%

NOTES: This exhibit reports relative deviations in percentage terms for each emission category. Relative deviations reflect the percentage difference between the emissions from two different datasets for the same company in the same year. The percentage values within the exhibit reflect the fraction of joint observations in the respective relative emission deviation class. Numbers may not add up to 100% due to rounding.

SOURCE: Research Affiliates, LLC and University of Augsburg based on anonymized data from GHG emissions data providers.

larger than 30%, and for almost half of the sample (47%), the estimates from the two data providers disagree by more than 70%. These differences are large and caused by the different carbon emission estimation models.

APPENDIX E

PERSISTENCE OF CARBON EMISSIONS/INTENSITIES TRENDS

EXHIBIT E1

Stability of Carbon Emission Trends

$t/t+1$	Strong Reducer_{t+1}	Reducer_{t+1}	No Changer_{t+1}	Increaser_{t+1}	Strong Increaser_{t+1}
Strong Reducer_t	27.2%	23.9%	6.4%	17.6%	25.0%
Reducer_t	17.0%	36.0%	9.3%	22.9%	14.8%
No Changer_t	14.5%	33.6%	12.3%	24.4%	15.2%
Increaser_t	13.8%	29.9%	11.0%	28.1%	17.2%
Strong Increaser_t	20.6%	21.6%	7.2%	23.3%	27.4%

NOTES: The migration matrix shows from which category companies migrated from one year to the next year. We distinguish between strong reducers (emissions reduction of more than 10%), reducers (emissions reduction between 1% and 10%), no changer (overall emissions change smaller than 1%), increasers (emissions rise between 1% and 10%), and strong increasers (emissions rise more than 10%). Changes in emissions reflect the percentage deviation in companies' Scope 1 and 2 emissions for the period 2010–2016.

SOURCE: Research Affiliates and University of Augsburg, based on anonymized data from GHG emissions data providers.

EXHIBIT E2

Stability of Carbon Intensity Trends

$t/t+1$	Strong Reducer_{t+1}	Reducer_{t+1}	No Changer_{t+1}	Increaser_{t+1}	Strong Increaser_{t+1}
Strong Reducer_t	24.3%	20.8%	4.9%	16.0%	34.0%
Reducer_t	21.3%	28.1%	5.5%	19.0%	26.0%
No Changer_t	17.9%	26.6%	5.4%	24.1%	26.1%
Increaser_t	19.8%	23.7%	5.5%	21.8%	29.2%
Strong Increaser_t	24.6%	16.2%	5.0%	18.5%	35.7%

NOTES: The migration matrix shows from which category companies migrated from one year to the next year. We distinguish between strong reducers (emissions reduction of more than 10%), reducers (emissions reduction between 1% and 10%), no changer (overall emissions change smaller than 1%), increasers (emissions rise between 1% and 10%), and strong increasers (emissions rise more than 10%). Changes in emissions reflect the percentage deviation in companies' carbon intensities. Carbon intensities reflect the ratio of carbon emissions to net sales.

SOURCE: Research Affiliates, LLC and University of Augsburg based on anonymized data from GHG emissions data providers.

APPENDIX F

EXHIBIT F1

Predictability of Future Reported Carbon Intensity Changes

	Δ% Carbon $\text{Intensity}_{t+1,t}$	Δ% Carbon $\text{Intensity}_{t+1,t}$	Δ% Carbon $\text{Intensity}_{t+1,t}$	Δ% Carbon $\text{Intensity}_{t+1,t}$	Δ% Carbon $\text{Intensity}_{t+1,t}$	Δ% Carbon $\text{Intensity}_{t+1,t}$
Δ% carbon $\text{intensity}_{t1,t-1}$		–0.04*	–0.04*	–0.04*	–0.04*	–0.04*
DP_C A list_t			–0.08			–0.00
DP_A emission reduction score_t				–0.00		–0.00
DP_D carbon emission score_t					–0.08	–0.07
Constant	0.66	0.68	0.68	0.94	0.87	0.96
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	5,895	5,895	5,895	5,895	5,895	5,895
R^2	3.5%	3.6%	3.6%	3.6%	3.7%	3.7%
Adjusted R^2	–3.2%	–3.1%	–3.1%	–3.1%	–3.0%	–3.0%
Rank correlation between predicted changes in future emissions and reported changes in future emissions	10.9%	11.4%	11.3%	10.9%	10.2%	10.1%

NOTES: This exhibit reports regression results of past reported carbon emissions changes on future reported carbon emissions changes (scope 1 and 2). Δ% *Carbon* $\textit{intensity}_{t+1,t}$ reflects the percentage change of a company's carbon intensity (Scope 1 and 2 emissions divided by net sales) from one year to another. Industry fixed effects are based on a company's TRBC activity code. Country fixed effects are based on a country's ISO 3166 code. ** and * indicate significance at the 1% and 5% level, respectively. Standard errors are clustered at the company level.

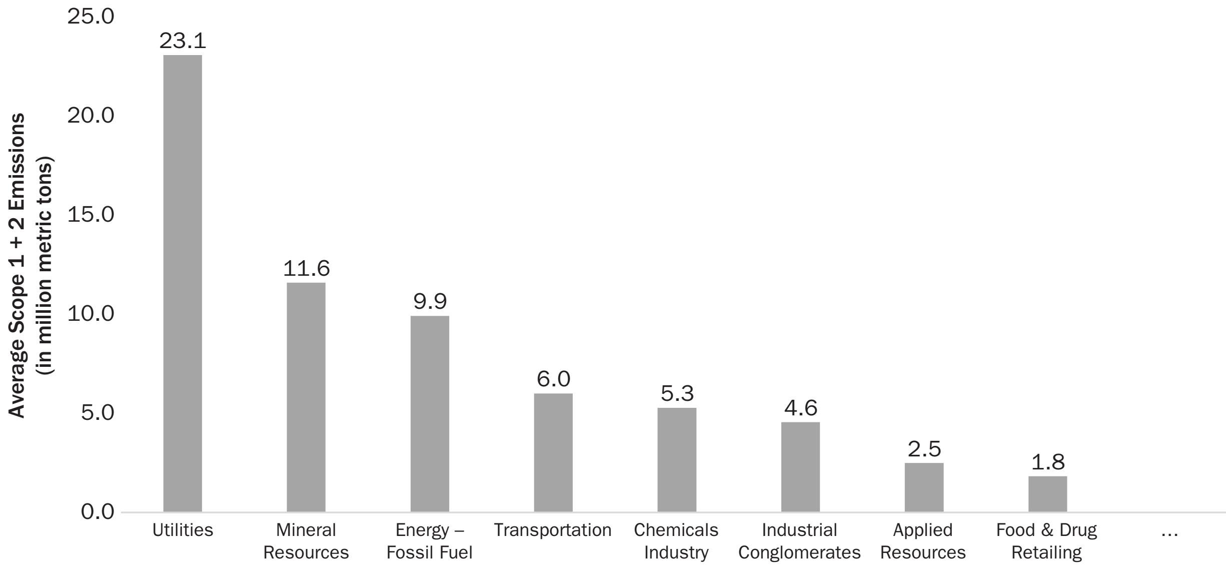
SOURCE: Research Affiliates, LLC and University of Augsburg based on anonymized data from GHG emissions data providers.

APPENDIX G

BREAKDOWN OF CARBON EMISSIONS BY INDUSTRY AND COUNTRY

EXHIBIT G1

Carbon Emissions by Industry: Averages of Scope 1 and 2 Emissions

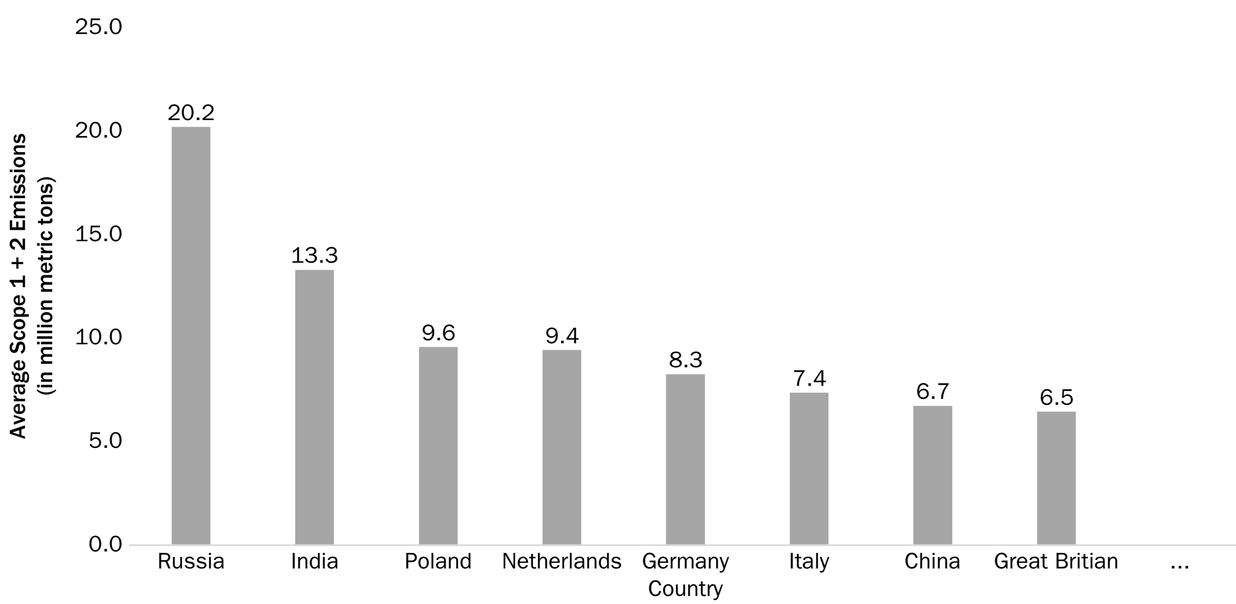


NOTES: This exhibit shows the cross-sectional average of Scope 1 and 2 emissions by industry. Industries are classified using Thomson Reuters Sector Code. Countries are classified according to their ISO 3166 standard. A minimum of 10 observations is required to calculate the average.

SOURCE: Research Affiliates, LLC and University of Augsburg based on anonymized data from GHG emissions data providers.

EXHIBIT G2

Carbon Emissions by Country: Averages of Scope 1 and 2 Emissions



NOTES: This exhibit shows the cross-sectional averages of Scope 1 and 2 emissions by country over the period 2010–2016. Industries are classified using Thomson Reuters Sector Code. Countries are classified according to their ISO 3166 standard. A minimum of 10 observations is required to calculate the average.

SOURCE: Research Affiliates, LLC and University of Augsburg based on anonymized data from GHG emissions data providers.

APPENDIX H

OVERVIEW OF BUSINESS METRICS USED IN REGRESSION ANALYSES

No.	Variable	Variable Code	Description (of nominator)	Captures	Rationale
1	Log (Net sales)	WC01001	Net sales represent gross sales and other operating revenue less discounts, returns, and allowances.	Company size and production levels	Larger companies are more likely to be carbon intensive.
2	Log (Employees/ Net sales)	WC07011; WC01001	Employees represent the number of both full and part-time employees of the company.	Employee intensity of production	Companies with higher employee intensity are likely to be carbon intensive (e.g., need more office space).
3	Log (Market capitalization/ Net sales)	WC08001; WC01001	Market price at year end * Common shares outstanding	Market valuation of production	Companies with lower market valuations are more likely to be carbon intensive because emissions are negatively associated with market values on average (Matsumura, Prakash, and Vera-Muñoz 2017).
4	Log (Earnings before taxes/ Net sales)	WC08321; WC01001	Pretax income/Net sales or revenues * 100	Profitability	Companies with higher profitability have more capacity to invest in carbon-efficient technologies.
5	Log (Funds from operations/ Net sales)	WC04201; WC01001	Funds from operations represent the sum of net income and all noncash charges or credits. It is the cash flow of the company.	Liquidity	Companies with higher liquidity have more cash available to invest in carbon-efficient technologies.
6	Log (Operating income/Net sales)	WC01250; WC01001	Operating income represents the difference between sales and total operating expenses.	Profitability	Companies with higher profitability have more capacity to invest in carbon-efficient technologies.
7	Log (Property, plant, and equipment/ Net sales)	WC02501; WC01001	Net property, plant, and equipment represents gross property, plant, and equipment less accumulated reserves for depreciation, depletion, and amortization.	Tangible assets	Companies with a higher fraction of tangible assets are more likely to be carbon-intensive (e.g., energy to heat buildings, energy for production machines). In contrast, intangibles are less likely to be carbon-intensive (e.g., patents or brand values).
8	Log (Cost of goods sold/ Net sales)	WC01051; WC01001	For manufacturing companies, the cost of goods sold represents specific or direct manufacturing cost of material and labor used in the production of finished goods. For merchandise companies, the cost of goods sold represents the purchase price of items sold, as well as indirect overhead such as freight, inspection, and warehouse costs.	Production levels	Companies with higher production levels are more likely to be carbon intensive.

ACKNOWLEDGMENT

This work was supported by Stiftung Mercator: [grant number 19026202] for the project ‘Wissenschaftplattform Sustainable Finance’ (Rahmenprogramm Sustainable Finance).

REFERENCES

Akey, P., and I. Appel. 2020. “Environmental Externalities of Activism.” SSRN 3508808.

Aswani, J., A. Raghunandan, and S. Rajgopal. “Are Carbon Emissions Associated with Stock Returns?” Working paper, Columbia Business School, 2021.

Avramov, D., S. Cheng, A. Lioui, and A. Tarelli. 2022. “Sustainable Investing with ESG Rating Uncertainty.” *Journal of Financial Economics* 145 (2): 642–664.

Barber, B., A. Morse, and A. Yasuda. 2020. “Impact Investing.” *Journal of Financial Economics* 139 (1): 162–185.

Berg, F., J. Koelbel, and R. Rigobon. “Aggregate Confusion: The Divergence of ESG Ratings.” Working Paper 5822-19, MIT Sloan School of Management, 2020.

Boermans, M. A., and R. Galema. 2019. “Are Pension Funds Actively Decarbonizing Their Portfolios?” *Ecological Economics* 161: 50–60.

Bolton, P., and M. T. Kacperczyk. 2021. “Do Investors Care about Carbon Risk?” *Journal of Financial Economics* 142 (2): 517–549.

——. 2022. “Global Pricing of Carbon-Transition Risk.” SSRN 3550233.

Brandon, R. G., K. Philipp, and P. S. Schmidt. 2021. “ESG Rating Disagreement and Stock Returns.” *Financial Analysts Journal* 77 (4): 104–127.

Broccardo, E., O. Hart, and L. Zingales. “Exit vs. Voice.” Working paper no. 694/2020, European Corporate Governance Institute Finance, 2020.

Busch, T., M. Johnson, T. Pioch, and M. Kopp. *Consistency of Corporate Carbon Emission Data*. Hamburg: University of Hamburg/WWF Deutschland, 2018.

Busch, T., and S. Lewandowski. 2018. “Corporate Carbon and Financial Performance: A Meta-Analysis.” *Journal of Industrial Ecology* 22 (4): 745–759.

CDP. “The Carbon Majors Database: CDP Carbon Majors Report 2017.” 2017.

Cheema-Fox, A., B. LaPerla, G. Serafeim, D. Turkington, and H. Wang. 2019. “Decarbonization Factors.” SSRN 3448637.

Chowdhry, B., S. Davies, and B. Waters. 2019. “Investing for Impact.” *The Review of Financial Studies* 32 (3): 864–904.

Climate Action 100+. *2020 Progress Report*. 2020, http://www.climateaction100.org/.

Cotter, J., and M. M. Najah. 2012. “Institutional Investor Influence on Global Climate Change Disclosure Practices.” *Australian Journal of Management* 37 (2): 169–187.

De Angelis, T., P. Tankov, and O. Zerbib. 2020. “Environmental Impact Investing.” SSRN 3562534.

Dimson, E., P. Marsh, and M. Staunton. 2020. “Divergent ESG Ratings.” *The Journal of Portfolio Management* 47 (9): 75–87.

Dordi, T., and O. Weber. 2019. “The Impact of Divestment Announcements on the Share Price of Fossil Fuel Stocks.” *Sustainability* 11 (11): 1–20.

Edmans, A., X. Gabaix, and D. Jenter. 2017. "Chapter 7—Executive Compensation: A Survey of Theory and Evidence." In *The Handbook of the Economics of Corporate Governance*, vol. 1, North-Holland: Elsevier.

EU Commission. 2021. "EU Taxonomy, Corporate Sustainability Reporting, Sustainability Preferences and Fiduciary Duties: Directing Finance Towards the European Green Deal." COM(2021) 188 Final, April 21, 2021, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52021DC0188&from=EN.

Fama, E., and K. French. 2007. "Disagreement, Tastes, and Asset Prices." *Journal of Financial Economics* 83 (3): 667–689.

Flammer, C., M. W. Toffel, and K. Viswanathan. 2019. "Shareholder Activism and Firms' Voluntary Disclosure of Climate Change Risks." SSRN 3468896.

Gollier, C., and S. Pouget. "The 'Washing Machine': Investment Strategies and Corporate Behavior with Socially Responsible Investors." Working Paper 14-457, Toulouse School of Economics, 2014.

Goergen, M., A. Jacob, M. Nerlinger, R. Ryan, M. Rohleder, and M. Wilkens. 2020. "Carbon Risk." SSRN 2930897.

Heinkel, R., A. Kraus, and J. Zechner. 2001. "The Effect of Green Investment on Corporate Behavior." *Journal of Financial and Quantitative Analysis* 36 (4): 431–449.

In, Y., K. Park, and A. Monk. "Is 'Being Green' Rewarded in the Market?: An Empirical Investigation of Decarbonization and Stock Returns." Working paper, Stanford Global Project Center, 2019.

Karagozoglu, A. K. 2021. "Novel Risks: A Research and Policy Overview." *The Journal of Portfolio Management* 47 (9): 11–34.

Koelbel, J., F. Heeb, P. Falko, and T. Busch. 2020. "Can Sustainable Investing Save the World? Reviewing the Mechanism of Investor Impact." *Organization & Environment* 33 (4): 554–574.

Lee, S.-Y., Y.-S. Park, and R. D. Klassen. 2015. "Market Responses to Firms' Voluntary Climate Change Information Disclosure and Carbon Communication." *Corporate Social Responsibility and Environmental Management* 22 (1): 1–12.

Li, F., and A. Polychronopoulos. "What a Difference an ESG Ratings Provider Makes!" Research Affiliates Publications, January 2020, researchaffiliates.com.

Liesen, A., F. Figge, A. Hoepner, and D. M. Patten. 2015. "Climate Change and Asset Prices: Are Corporate Carbon Disclosure and Performance Priced Appropriately?" *Journal of Business Finance & Accounting* 44 (1–2): 35–62.

Matsumura, E. M., R. Prakash, and S. C. Vera-Muñoz. 2014. "Firm-Value Effects of Carbon Emissions and Carbon Disclosures." *Accounting Review* 89 (2): 695–724.

——. 2017. "To Disclose or Not to Disclose Climate-Change Risk in Form 10-K: Does Materiality Lie in the Eyes of the Beholder?" SSRN 2986290.

Naaraayanan, S. L., K. Sachdeva, and V. Sharma. "The Real Effects of Environmental Activist Investing." Finance Working Paper no. 743/2021, European Corporate Governance Institute, 2021.

Oehmke, M., and M. Opp. "A Theory of Socially Responsible Investment." Research Paper no. 20-2, Swedish House of Finance, 2020.

Pástor, L., R. F. Stambaugh, and L. A. Taylor. 2021. "Sustainable Investing in Equilibrium." *Journal of Financial Economics* 142 (2): 550–571.

Prado-Lorenzo, J.-M., I. Gallego-Alvarex, and I. M. Garcia-Sanchez. 2009. "Stakeholder Engagement and Corporate Social Responsibility Reporting: The Ownership Structure Effect." *Corporate Social Responsibility and Environmental Management* 16 (2): 94–107.

Principles for Responsible Investment (PRI). "Principles for Responsible Investment: An Investor Initiative in Partnership with UNEP Finance Initiative and the UN Global Compact." 2020, https://www.unpri.org/download?ac=10948.

Rohleder, M., M. Wilkens, and J. Zink. 2022. "The Effects of Mutual Fund Decarbonization on Stock Prices and Carbon Emissions." *Journal of Banking & Finance* 134: 1–13.

Securities and Exchange Commission (SEC). "Statement on the Review of Climate-Related Disclosure." 2021a, https://www.sec.gov/news/public-statement/lee-statement-review-climate-related-disclosure.

——. "Public Input Welcomed on Climate Change Disclosures." 2021b, https://www.sec.gov/news/public-statement/lee-climate-change-disclosures.

TCFD. "Task Force on Climate-Related Financial Disclosures: Status Report." 2019, https://www.fsb.org/wp-content/uploads/P050619.pdf.

Trinks, A., B. Scholtens, M. Mulder, and L. Dam. 2018. "Fossil Fuel Divestment and Portfolio Performance." *Ecological Economics* 146: 740–748.

United Nations. "Report of the United Nations Conference on Environment and Development." A/CONF.151/26, vol. 1, 1992, https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_CONF.151_26_Vol.I_Declaration.pdf.

——. "Kyoto Protocol to the UN Framework Convention on Climate Change." 1998, https://unfccc.int/sites/default/files/kpeng.pdf.

US Congress. "H.R. 2570." Report no. 117-39, 2021, https://www.congress.gov/bill/117th-congress/house-bill/2570/text.

WBCSD and WRI. "Corporate Value Chain (Scope 3) Accounting and Reporting Standard." 2011, https://ghgprotocol.org/sites/default/files/standards/Corporate-Value-Chain-Accounting-Reporing-Standard_041613_2.pdf.

——. "GHG Protocol Scope 2 Guidance." 2015a, https://ghgprotocol.org/sites/default/files/standards/Scope%202%20Guidance_Final_Sept26.pdf.

——. "The Greenhouse Gas Protocol—A Corporate Accounting and Reporting Standard." 2015b, https://ghgprotocol.org/sites/default/files/standards/ghg-protocol-revised.pdf.

Copyright of Journal of Portfolio Management is the property of With Intelligence Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.