The anatomy of decarbonizing firms

Guillaume Coqueret* Thomas Giroux† Borui Qiu‡§

January 5, 2026

Abstract

We study the firm-level drivers of corporate decarbonization using a global panel combining financial characteristics, balance-sheet indicators, ESG metrics, macroeconomic variables, and firms' greenhouse gas emissions. Emission trajectories exhibit strong persistence, with past emissions the dominant predictor of future performance. However, firms with larger carbon footprints also display greater abatement potential, creating opportunities for targeted stewardship and capital reallocation. We identify adoption of Science-Based Targets (SBT) as a significant catalyst for emission reductions and provide causal evidence that credible, externally validated climate commitments accelerate decarbonization. We develop machine-learning models that outperform heuristic benchmarks in forecasting one-year-ahead abatement, especially in hard-to-abate sectors, and we show that such forecasts can help further curtail portfolio-level emissions. Nevertheless, our results indicate that, given current emission trends in the cross-section of firms, building portfolios with net zero objectives for 2050 remains an open challenge.

1 Introduction & Motivation

1.1 Context

The world is rapidly approaching the critical threshold of 1.5°C global warming,[1] with a 70% likelihood this limit will be breached in the near future. As a result, the financial sector is under increasing pressure to help drive the transition to a low-carbon economy (Bolton and Kacperczyk, 2021; Krueger et al., 2020). By 2024, over 300 institutions, managing $60T in assets, had joined the Net Zero Asset Managers initiative, reflecting investors' growing commitment to climate action.

Translating these commitments into real-world impact requires robust measurement of current emissions and credible forecasts of future trajectories. Yet, accurate and consistent

*EMLYON Business School, 144 avenue Jean Jaures, 69007 Lyon, FRANCE. E-mail: coqueret@em-lyon.com
†ETH Zurich, SFI, Zurichbergstrasse 18, 8032 Zurich, Switzerland. E-mail: thomas.giroux@mtec.ethz.ch
‡EMLYON Business School, 144 avenue Jean Jaures, 69007 Lyon, FRANCE. E-mail: qiu@em-lyon.com
§COACTIS, Université Lyon 2, 14/16 av Berthelot, 69363 Lyon Cedex 07 FRANCE.

[1]See WMO Global Annual to Decadal Climate Update.

emissions data remain difficult to obtain, if only because disclosure standards and regulatory requirements vary widely across jurisdictions. For example, the EU's Corporate Sustainability Reporting Directive (CSRD) mandates GHG disclosure from 2025 for over 50,000 firms, while recent U.S. regulations focus on large companies and material emissions, with staggered implementation timelines. Moreover, only a handful of countries have truly mandatory GHG reporting regimes (World Resources Institute).

This fragmented disclosure landscape complicates investors' ability to assess emissions at scale and integrate climate risk into investment processes. Data providers attempt to fill these gaps with proprietary estimation models, but their transparency and accuracy are limited. Consistent with the literature on ESG score dispersion (Berg et al., 2022; Dimson et al., 2020), recent research documents significant discrepancies in GHG estimates—especially for modeled Scope 2 and 3 emissions—raising concerns about the reliability of available data (Busch et al., 2022; Chen et al., 2025; Kalesnik et al., 2022; Papadopoulos, 2022).

This fragmented disclosure landscape complicates investors' ability to assess emissions at scale and integrate climate risk into investment processes. Data providers attempt to fill these gaps with proprietary estimation models, but their transparency and accuracy are limited. Consistent with the literature on ESG score dispersion (Berg et al., 2022; Dimson et al., 2020), recent research documents significant discrepancies in GHG estimates—especially for modeled Scope 2 and 3 emissions—raising concerns about the reliability of available data (Busch et al., 2022; Kalesnik et al., 2022; Papadopoulos, 2022).

Developing accurate, transparent, and systematic methods for forecasting firm-level emissions is therefore essential to enable effective capital allocation and to support investors in meeting their net-zero objectives.[2] In the next subsection, we introduce a simple theoretical framework to illustrate why reliable future emissions forecasts are indispensable for asset managers seeking to align portfolios with net-zero goals.

1.2 Theoretical motivation

Consider an asset manager whose sole objective is to minimize the future footprint f of her portfolio. The investment universe consists of a continuum of assets (e.g., stocks), each characterized at time t by a footprint f_t, modeled as a random variable representing the cross-sectional distribution of footprints. Let w_t denote the portfolio weights chosen at time t, and let $f_t^{(w)}$ be the corresponding aggregate portfolio footprint. The relative change in the portfolio footprint is then

$$\Delta f_{t+1}^{(w)} = \frac{f_{t+1}^{(w)} - f_t^{(w)}}{f_t^{(w)}}.$$

For simplicity, we assume a constant negative rate of change and derive the decarbonization trajectories shown in Figure 1. Starting from the 2025 footprint, we plot a continuous decline under three scenarios corresponding to different decay rates. Due to compounding, these rates lead to substantially different long-term outcomes. In particular, achieving a 90% reduction in the footprint over a 25-year horizon requires a constant annual decay rate of

[2] See, for example, Barahhou et al. (2022); Bolton et al. (2022); Cenedese et al. (2023); Fraser and Fiedler (2023); Le Guenedal et al. (2022); Le Guenedal and Roncalli (2022); Roncalli (2024).

approximately 9%. As discussed below, sustaining such a rate is especially difficult because of the interaction between two key factors.

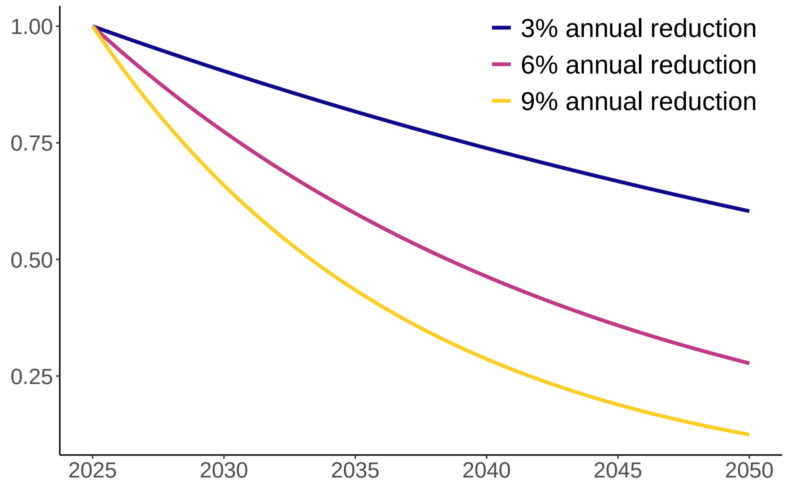


Figure 1: **Decarbonization paths**.

The first factor is largely exogenous: the decarbonization of an aggregate portfolio is determined by the carbon trajectories of the individual firms it contains. If all firms exhibit increasing carbon footprints, then any convex combination of these assets will necessarily result in an increasing aggregate footprint. At the same time, however, the asset manager can shape the investment universe by selecting firms that are, on average, more likely to reduce their emissions. The second challenge is therefore primarily technical: forecasting future carbon footprints is inherently difficult, even when firms' business models display substantial inertia.

We lay out a simple theory that shows the interplay of these two effects. To this purpose, we write $\hat{f}_{t+1}$ for the manager's estimates of the footprints at time $t+1$. By construction, these estimates are imperfect, and what will matter will be the degree to which they are aligned with the future realizations of the footprints of the firms.

Since the manager's objective is to reduce the aggregate portfolio footprint, she adjusts her portfolio positions based on her predictions in the following manner. Let

$$\tilde{\Delta}\hat{f}_{t+1} = \Delta\hat{f}_{t+1} - c_w, \quad \left(\text{with} \quad \Delta\hat{f}_{t+1} = \frac{\hat{f}_{t+1} - f_t}{f_t}\right) \tag{1}$$

denote the (shifted) predicted relative change in footprint, where c_w is a deviation constant which we define below. The manager want to increase (*resp.* decrease) her positions in the assets for which the footprint is forecasted to decrease (*resp.* increase). If we assume a proportional adjustment, this gives:

$$\Delta w_t = \frac{w_t}{w_{t-1}} - 1 = -\alpha\tilde{\Delta}\hat{f}_{t+1} \quad \Longleftrightarrow \quad w_t = w_{t-1}\left(1 - \alpha\tilde{\Delta}\hat{f}_{t+1}\right). \tag{2}$$

In Equation (1), c_w is the constant that ensures that the budget constraint $\mathbb{E}[w_t] = 1$ is satisfied, i.e., it is equal to $c_w = \mathbb{E}[w_{t-1}\Delta\hat{f}_{t+1}]$. The parameter $\alpha > 0$ allows the manager to tune the portfolio adjustments, based on her confidence in her predictions and depending on her willingness to decarbonize her allocation smoothly (low α) or aggressively (large α). It can then be shown under a mild assumption (see Appendix A) that the average decarbonization rate is

$$\delta = \mathbb{E}\left[\Delta f^{(w)}_{t+1}\right] = \underbrace{\mu}_{\substack{\text{decarbonization potential} \\ \text{in investment universe}}} - \underbrace{\kappa}_{\substack{\text{adjustment} \\ \text{speed}}} \times \underbrace{\mathbb{C}\text{or}(\Delta f_{t+1}, \Delta \hat{f}_{t+1})}_{\rho\text{ = predictive skill}}, \tag{3}$$

where $\mu = \mathbb{E}[\Delta f_{t+1}]$ is the average decarbonization potential in the investment universe that is considered by the manager. The constant $\kappa = \alpha\sigma_{\Delta f}\sigma_{\Delta \hat{f}}$ is the adjustment rate, scaled by the standard deviations of Δf and $\Delta \hat{f}$. Indeed, we need to adapt α to the magnitude of footprint variations (and their forecasts). If they are too limited, the portfolio update will be insignificant and decarbonization very slow.

The above expression implies that expected rate of footprint variation depends on the relative importance of two terms:

1. the average rate of decay of the investable universe, μ: this is the raw potential for decarbonization;
2. a term that captures the accuracy of forecasts: it is driven by the manager's ability to correctly predict the evolution of footprints. We will refer to this as the *skill* factor.

Simply put, the manager wants δ to be as negative as possible. When $\rho = 0$ in Equation (3), predictions are unrelated to realizations (the manager has zero skill), and the expected decarbonization rate reduces to μ. By contrast, if forecasts are positively correlated with realized changes (i.e., $\rho > 0$ and the manager is skilled), the manager will achieve a larger decarbonization rate than μ.

Figure 2 illustrates δ as a function of the two parameters: the correlation (forecasting skill) ρ in the x-axis and the adjustment speed κ (solid versus dotted lines). An original speed of $\alpha = 1$ in Equation (2) means that if a firm has a forecasted relative decarbonization change of -10%, then its weight in the portfolio will increase by 10%. This speed seems reasonable, but for the sake of completeness we also consider a more aggressive adjustment factor of $\alpha = 3$ as well.

Two sets of curves are displayed. The upper set (in **brown**) corresponds to a pessimistic scenario in which the manager has an investment universe in which the average relative variation is $\mu = +6\%$, i.e., when firms pollute *more* on average every year. The lower curves, in **green**, pertain to a more favorable situation in which the manager focuses on a universe with average decarbonization rate of -4%. The full lines pertain to a porfolio adjustment speed of $\alpha = 10$ in Equation (3), while the dotted lines correspond to the smoother case $\alpha = 5$. In the plot, the two vertical dashed lines mark two skill levels observed in our empirical study, one which we characterize as "*medium*" and one which is the best we were able to reach ("*high*" skill).

Let us now turn to the quantitative implications. The +6% scenario (upper curves in Figure 2) corresponds to the average decarbonization rate observed in the full sample of our study. Taking this as the baseline and assuming a (vigorous) adjustment speed of $\alpha = 3$, Figure 2 shows that the average footprint change that can be achieved with high forecasting skills is +2% (brown circle), meaning a reduction of -4% in absolute value compared the

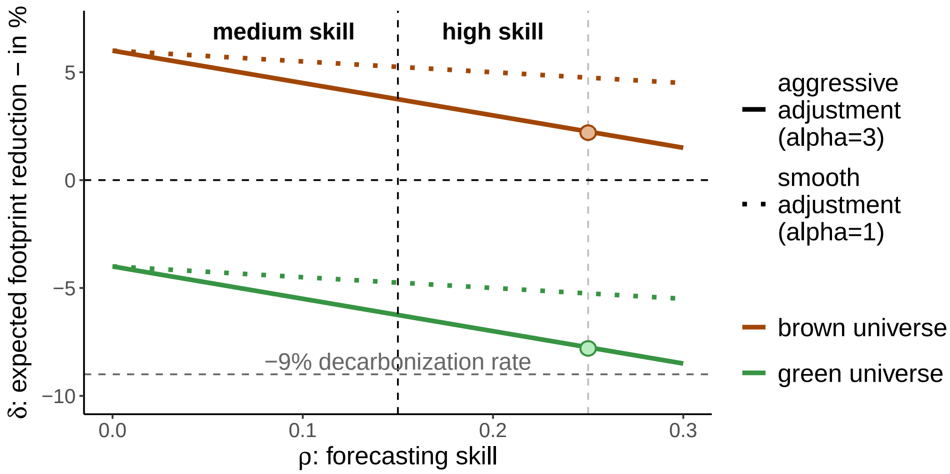


Figure 2: **Decarbonization rates**. We plot the expected decarbonization rate (δ in Equation (3)) as a function of the covariance (skill factor) ρ. The upper lines show the case when $\mu = +6\%$, whereas the lower lines pertain to $\mu = -4\%$. The solid (*resp.* dotted) lines correspond to an adjustment speed of $\alpha = 3$ (*resp.* $\alpha = 1$). The bottom dashed horizontal line marks the -9% decarbonization objective. Based on our empirical results, the product of standard errors in κ is set at $\sigma_{\Delta f}\sigma_{\Delta \hat{f}} = 0.05$.

original starting point. This is an increase in footprint nonetheless and and a value that is very distant from the initial goal of -9%.

By contrast, the second parametrization, with $\mu = -4\%$, nearly attains this target (green circle). Interestingly, this case coincides with the average trajectory of a specific subgroup within our sample: firms that have received approval from the Science-Based Target initiative (SBTi). Restricting attention to this smaller set of corporations, we obtain an expected decarbonization rate of nearly -8%, bringing us much closer to the desired objective. In sum, to reach an ambitious target requires both a thoroughly chosen investment universe to start with and an accurate model for footprint predictions.

In Equation (3), there are three quantities: κ is decided by the investor, ρ, the forecasting skill, hinges on the investor but is not chosen by her strictly speaking, and, finally, μ may or may not depend on the investor (think of a fund with a specific mandate). The two main empirical sections of the paper seek to shed light on two of these key parameters. First, Section 3 identifies drivers of decarbonization in the cross-section of firms and eventually proposes a causal mechanism that helps determine a universe with negative μ. Second, Section 4, through a machine learning prediction exercise, showcases the levels of ρ that can reasonably be assumed for the skill of investors. The combination of these calibrations for μ and ρ is analyzed in Section 4.3.

1.3 Related literature

Our study sits at the intersection of several strands of research on corporate decarbonization, sustainable finance, and emissions dynamics.

Firm-level drivers of decarbonization. A growing body of literature investigates the firm-level determinants of sustainability policies and GHG emissions reductions. Within the ESG context, Martiny et al. (2024) find that greater ownership by socially responsible investors predicts higher future CSR scores, while Kahn et al. (2023) document a similar association for corporate emissions. Financial instruments also appear relevant: firms issuing green bonds are found to subsequently reduce emissions (Flammer, 2021), and green bonds lower the cost of capital, potentially facilitating firms' transition (Zerbib, 2019). These mechanisms are reinforced by innovation capacity, as investments in green R&D and clean technologies are linked to sustained emissions reductions over time (Habiba et al., 2022; Lee and Min, 2015). Beyond ownership and financing, voluntary corporate commitments have received considerable attention. Firms adopting Science-Based Targets tend to set more stringent emissions goals and increase investment in abatement projects (Freiberg et al., 2021). More broadly, firms that publicly commit to reducing their emissions subsequently exhibit declines in their carbon footprint and intensify these efforts when expectations of future climate regulation rise (Bolton and Kacperczyk, 2023; Ramadorai and Zeni, 2024).

Questioning the real-economy impact. However, a growing body of work challenges the economic significance of these effects. Atta-Darkua et al. (2025); Heath et al. (2023); Rink et al. (2024) argue that sustainable investors exert only a limited influence on real-economy decarbonization. They show that portfolio-level decarbonization is largely achieved through reallocation away from high emitters toward firms that are already relatively green, implying that observed reductions primarily reflect portfolio tilting rather than changes in firms' production decisions. While such ownership can encourage public climate commitments, it does not necessarily translate into meaningful emissions abatement. Consistent with this view, Berk and Van Binsbergen (2024) find that divestment has little effect on firms' cost of capital, challenging the mechanism proposed by Pástor et al. (2021). From a financing perspective, Feldhütter and Pedersen (2023) show that firms' incentives to invest in green projects are largely independent of financing structure, and Lam and Wurgler (2024) estimate that only about 2% of corporate green bond proceeds finance genuinely new green investments. Related evidence also tempers the effectiveness of voluntary commitments. Bolton and Kacperczyk (2023) document that the aggregate impact of Science-Based Targets on global emissions is limited, as participating firms tend to be lower emitters and more climate-conscious ex-ante, whereas the largest emitters either abstain or adopt weaker, intensity-based targets. Moreover, Berg et al. (2024) show that observed emissions reductions among firms with approved SBT are driven primarily by those obtaining third-party assurance of their carbon accounting, rather than by target-setting alone.

The central role of public policy. In contrast to the mixed evidence on investor- and firm-driven mechanisms, several studies emphasize the central role of public policy. Leffel et al. (2024) argue that state-level climate measures, such as subsidies and incentives for energy efficiency, have a stronger impact on decarbonization than voluntary corporate initiatives. Complementing this view, Ramadorai and Zeni (2024) show that firms' beliefs about future climate regulation, together with reputational concerns and peer effects, are key drivers of abatement activity. At a more general level, Pedersen (2024) identify carbon pricing as the most efficient policy instrument to maximize social welfare, and Adamolekun (2024) provide empirical evidence that carbon pricing in the European Union significantly reduces corporate emissions. In sum, current evidence points to several credible channels through which firms can reduce emissions, while also underscoring meaningful heterogeneity in corporate responses to decarbonization incentives.

Forecasting ESG performance and GHG emissions. Alongside this predominantly causal literature, a separate strand focuses on the empirical challenge of forecasting ESG performance and GHG emissions. Goldhammer et al. (2017) introduce early firm-level forecasting models using Ordinary Least Squares and Gamma Generalized Linear Regression, improving upon simple extrapolation but for a limited sample of European firms. Building on this work, Nguyen et al. (2021) apply machine learning techniques and develop a meta-learner combining linear models, penalized regressions, neural networks, tree-based methods, and clustering algorithms, substantially improving in- and out-of-sample accuracy. Nevertheless, they highlight persistent challenges, including imperfect separation between direct and indirect emissions and pronounced industry-level heterogeneity in predictive performance. For instance, forecast accuracy for Scope 1 emissions in highly polluting industries remains modest, with average R^2 values below 50%. More recently, Michalski and Low (2024) extend similar methods to ESG score prediction, while Pastor et al. (2024) estimate the present value of future sector-level emissions and show that emissions dynamics are predictable from firms' past emissions, investment behavior, climate scores, and valuation characteristics.

Macroeconomic determinants of decarbonization. Finally, our analysis relates to a more mature literature on the macroeconomic determinants of decarbonization. While much of this work focuses on individual countries, several studies adopt panel approaches across groups of countries. This literature consistently highlights the importance of trade openness (Coskuner et al., 2020; Dogan and Seker, 2016b; Nguyen et al., 2021; Sharma, 2011), income levels and economic growth (Coskuner et al., 2020; Nguyen et al., 2021; Sharma, 2011), and energy consumption patterns—particularly the composition of fossil versus renewable energy (Dogan and Seker, 2016a; Jiang and Guan, 2016). Environmental regulation (Puertas and Marti, 2021) and financial development (Dogan and Seker, 2016b) also emerge as robust predictors of emissions dynamics.

1.4 Summary of contributions

In this paper, we contribute to the emerging literature on decarbonization in several ways. First, we examine the characteristics of the firms most susceptible to curtail their GHG output. A key insight is that no single narrative explains which variables drive emissions reductions. Both sustainability-linked factors and traditional financial metrics matter at the firm-level, while external context also plays a meaningful role. However, we do not find evidence that institutional ownership or active "green" funds are associated with lower future emissions, challenging the view that sustainable investors consistently drive decarbonization through capital allocation. Overall, our findings indicate that real-world GHG reductions arise from the interaction of firm-level strategy, financial strength, and broader structural forces. Effective corporate decarbonization therefore requires not only internal efforts, such as higher innovation or the adoption of Science-Based Targets, but also supportive energy systems, market structures, and favorable policy environments.

Second, we document a substantial discrepancy between firms' reported emissions and those estimated by data providers, who rely heavily on backward-looking accounting models and place limited weight on forward-looking commitments such as Science-Based Targets. This is critical, as the adoption of such targets is one of the strongest predictors of future emissions reductions. Our causal analysis suggests that Science-Based Target adoption leads to an additional 1.26% annual reduction in emissions intensity.

Third, we find that complex nonlinear models using panel data do not consistently outperform simpler approaches in forecasting emissions levels. Emissions adjust slowly, making recent historical values highly informative for short term forecasts. However, when forecasting abatement rates, nonlinear approaches deliver significant gains. These forecasts can help portfolio managers tilt capital toward firms with declining emissions trajectories, hereby strengthening the financial system's capacity to support the net-zero transition.

Finally, we show that prediction accuracy varies meaningfully across industries. Hard-to-abate sectors such as Energy, Materials, Utilities, and Industrials are more predictable, increasing the value of our data-driven approach in these sectors.

2 Data

2.1 Sources

We collect carbon emission footprints under scopes 1, 2 and 3 for listed equities worldwide, covering a broad universe of developed and emerging markets for a period spanning from 2014 to 2023. We rely on ISS data because it is one of the main provider and it covers a very large database of both reported and estimated emissions. We also consider MSCI as additional source only if a firm is not present in the ISS database. The rationale for this is that we do not want dual sources for a given stock and we prioritize ISS as our main source for emissions.

GHG intensities are computed by scaling raw emissions by revenues, which are sourced from Bloomberg. This approach has become standard in the industry, as well as academic research on the topic. Table 7 of the Appendix displays the name, source, and definition of each variable.

Our panel sample covers over 22,000 unique companies worldwide (North America, Europe, Japan, Asia ex Japan and Emerging markets) over the period 2014 to 2023. We underline that emissions for 2023 are only available in 2024. As shown in Figure 3, North America dominates in terms of number of companies and total market capitalization since 2015, followed by Europe and Asia.

As is shown in Table 1, our sample has a bias toward small companies, reflecting the true nature of financial markets and the economy more generally.

An increasing number of companies have begun reporting their direct GHG emissions over time, rising from 13% of the overall sample in 2014 to 55% in the most recent year, with a notable jump in 2022. Interestingly, the share of reported direct emissions as a proportion of the total universe footprint has followed a more modest upward trend, albeit from a much higher starting point. In 2014, the 13% of reporting companies accounted for 55% of total

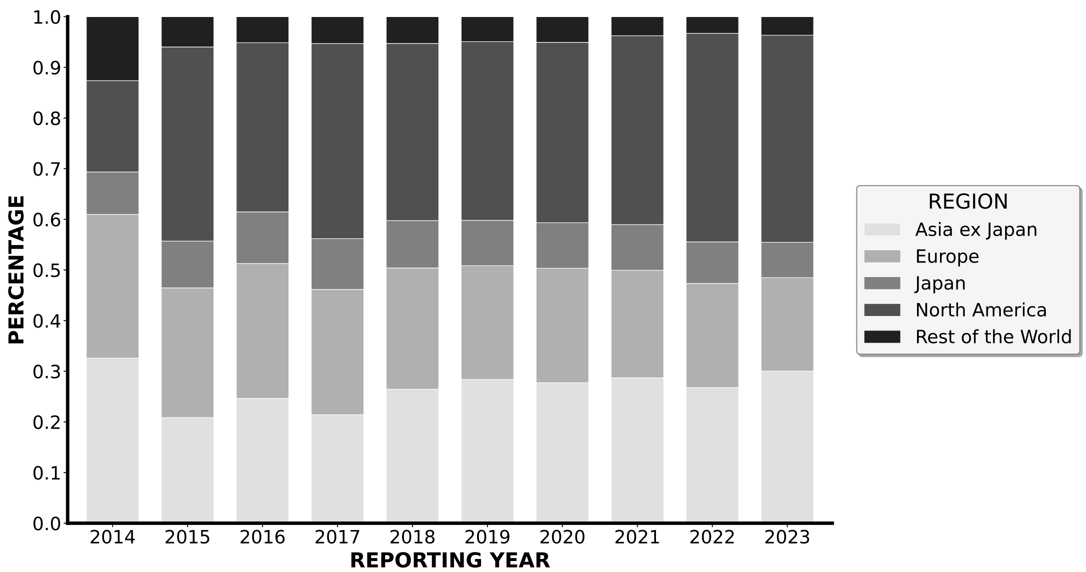


Figure 3: **Share of total market capitalisation by region and year**

Region	min	10%	50%	mean	90%	max
Asia ex Japan	0.50	58.00	736.00	4.33×10^3	7.37×10^3	1.74×10^6
Europe	0.50	48.00	990.00	1.03×10^4	1.83×10^4	7.86×10^5
Japan	3.70	55.00	376.00	3.94×10^3	6.04×10^3	9.80×10^5
North America	0.50	158.00	2.24×10^3	1.57×10^4	2.98×10^4	2.74×10^6
Rest of the World	0.60	65.00	1.06×10^3	5.97×10^3	1.08×10^4	2.92×10^5

Table 1: **Companies market capitalisation** (Million USD)

direct GHG emissions. By 2023, although 55% of companies reported, they represented only 83% of the total footprint. This indicates that more recent reporters tend to have relatively smaller carbon footprints than earlier ones. Nevertheless, focusing on reported Scope 1 and 2 emissions still captures over 80% of the total universe footprint, making such analysis broadly representative.

Figure 4 summarizes this finding over time, and Table 9 provides descriptive comparison statistics for the two groups.

In addition, we identify a broad list of candidate drivers of firm decarbonization based on the literature. For clarity, we grouped them into the following categories.

Corporate Sustainability Indicators. We identify the ESG profile of companies as indicative of their awareness of climate-related issues and their ability to carry out change. and collect E, S, and G scores from ISS. Furthermore, because climate change and the biodiversity crisis are deeply intertwined, we take advantage of Iceberg-Datalab to include corporate biodiversity footprint (CBF) as a proxy of firm impact on biodiversity. Finally, we include firms' GHG reduction targets collected by MSCI. This includes whether targets are validated

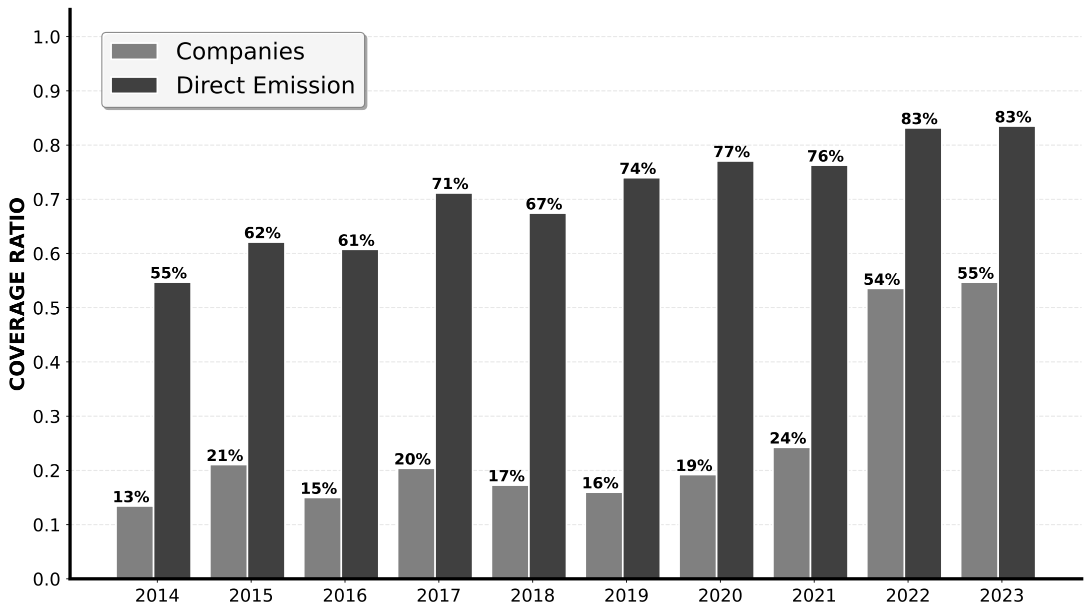


Figure 4: **Distribution of cumulative reported emissions over the period.**

by the Science Based Target Initiative (SBTi), among others.

Corporate Fundamental & Market Indicators. We consider the financial ability of companies to engage in decarbonization and gather market and fundamental variables from Bloomberg, including geography and industry segmentation (GICS), market capitalization, price to book, debt to equity, free-cash-flows, long-term investments, R&D spending and others.

Corporate Ownership. We include the percentage of shares held by institutional investors and the percentage of insider shares' outflow as provided by Bloomberg to account for the ownership structure. In addition, we compute our own estimate of active sustainable shareholders by aggregating annual holdings from a comprehensive list of 780 active sustainable equity funds computed by Exane Research (2023). Details about the funds' sample can be found in the online Appendix[3].

Regional Indicators. Finally, we also consider the carbon intensity of energy mix from the IEA, and an environmental policy score from the OECD. It is worth noting the US are not covered by the OECD data and are therefore considered as laggard in advocating ambitious environmental policies during the 2014 – 2024 period.

A summary of all variables names, sources and interpretations can be found in Table 8 in the Appendix. Because corporate carbon emissions are updated on a yearly frequency, we keep this reporting frequency in the construction of our dataset. For variables with higher

[3]ESG Fund Lists - Exane 2023

frequency such as market variables, we compute averages of past quarterly observations when relevant. Descriptive statistics of the universe are provided in Section 2.3.

In the present study, we use GHG emissions and intensities as the dependent variables and lag all independent variables by one year in order to conduct a predictive analysis.

2.2 Data construction

In the following section, we provide further details on the intermediary steps taken to construct a consistent and clean dataset from the various sources previously mentioned.

Treatment of outliers. We begin by excluding observations with non-positive GHG emissions. Next, we identify and remove outliers based on year-over-year changes in either absolute emissions or emissions intensity (normalized by sales). An observation is flagged as an outlier if it exhibits a reduction greater than 50% or an increase exceeding 200% compared to the previous year's reported value. Finally, to mitigate the influence of extreme values, all independent variables are winsorized at the 0.1^{th} and 99.9^{th} percentiles within each sector.

Panel construction. To be included in our panel dataset, firms must have at least four consecutive years of valid GHG data reported. If there are breaks in the data due to outlier treatment, we keep the most recent uninterrupted run only. We then divide all firms into two subsets: *i) Reported* sample, which includes firms that have reported their own GHG emissions for at least one year for both Scope 1 and Scope 2; and *ii) Modeled or Estimated* sample, which includes firms for which all GHG data are estimated by the data provider.

Data imputation. We perform data imputation for independent variables with a specific approach for each group. For sustainability indicators, we apply linear interpolation, while for the financial and market variables, following Chen and McCoy (2024), we resort to cross-sectional mean imputation at the sector and size level. In practice, for all ratios and percentage variables, we group the observations by GICS sector and fill the missing values with the cross-sectional mean of the sector each year. For absolute variables, we use the same approach, but we also take into account the market capitalization of the company compared to the sector average while grouping.

Omitted variables due to colinearity. Figure 13 in the Appendix presents the correlation matrix of all independent variables as defined above. We observe several groups of highly correlated variables and therefore curate our dataset to avoid multicollinearity issues in the estimation of linear models. First, we find that market capitalization, revenue, enterprise value, and lagged GHG emissions are highly correlated. To address this, we retain only lagged emissions, which we consider the most relevant predictor of firms' future emissions and a way to mitigate multicollinearity. Second, all regional indicators from the OECD display strong mutual correlations. For linear models, we retain only the cross-sectional policy score, which serves as a summary indicator. We also include the GDP growth rate of the country where each firm's headquarters is located, as a macroeconomic control. Finally, energy mix variables from the IEA are also highly correlated. For linear models, we retain only coal, oil, and renewable shares; guided by the prior belief that these sources are more indicative of a firm's potential to reduce emissions. In contrast, for nonlinear models such as random forests, we keep the full set of energy mix variables and allow the model to determine the most relevant predictors.

2.3 Descriptive statistics

Following the grouping described in Section 2.2, we split the analysis between companies belonging to the *i) reported* sample and those with *ii) modeled or estimated* GHG data.

We consider four types of dependent variables for each Scope:

1. Raw emissions measured in tons CO2 equivalent (tCO2e) - which we refer to as **GHG**;
2. Emission intensities (**INT**), reflecting tCO2e per million of revenue;
3. Relative changes in absolute emissions (Δ**GHG**);
4. Relative changes in emissions intensities (Δ**INT**).

Formally, if $\mathrm{GHG}_{t,i}$ is the emission value of firm i and time t, then we define $\Delta\mathrm{GHG}_{t,i} = \mathrm{GHG}_{t,i}/\mathrm{GHG}_{t-1,i} - 1$, and similarly for intensities.

We focus our baseline analysis on reported Scope 1+2 emissions (Direct Emissions) because they are less prone to disagreement and divergence, and because current indirect emissions (Scope 3) highly rely on the modeling process done by the data vendor. As an illustration, 96% of Scope 3 GHG emission in our sample are modeled. We run robustness tests on estimated emissions for Scope 1+2 and both reported and estimated Scope 3 emissions. Figure 5 below and Table 9 in the Appendix provide an overview of the distribution of these variables for Scope 1+2.

We observe that reported Scope 1 and 2 GHG emissions tend to be higher than modeled estimates, although their year-over-year changes display similar distributions. Several factors may explain this discrepancy. First, mandatory disclosure requirements do not apply universally and may introduce size and sectoral biases, as larger firms and industries with higher climate-related financial materiality are typically prioritized by regulators. Moreover, voluntary disclosure requires financial resources and technical expertise that smaller firms may lack. Conversely, modeled emissions may underestimate actual emissions due to methodological limitations, such as sectoral aggregation or reliance on financial proxies. Nonetheless, concerns have also been raised about biases in carbon accounting and voluntary disclosure. For instance, Chen et al. (2025) document a phenomenon of "green silence", whereby firms strategically choose whether to disclose emissions: high emitters tend to remain silent to exploit potential model underestimation, while lower emitters disclose to correct overestimates. Similarly, Berg et al. (2024) show that firms often underestimate their emissions when disclosure lacks third-party assurance, suggesting that taking reported emissions at face value may inadvertently penalize firms that engage more rigorously in carbon accounting.

Turning to the exogenous variables used as potential predictors in our analysis, we present their main descriptive statistics in Table 10. We confirm the suggested bias toward larger companies, trading at higher valuations and generating higher profit margins on average for our reported sample compared to companies in the modeled sample. Reporting companies also tend to have better ESG and environmental scores, and more ambitious Science-Based

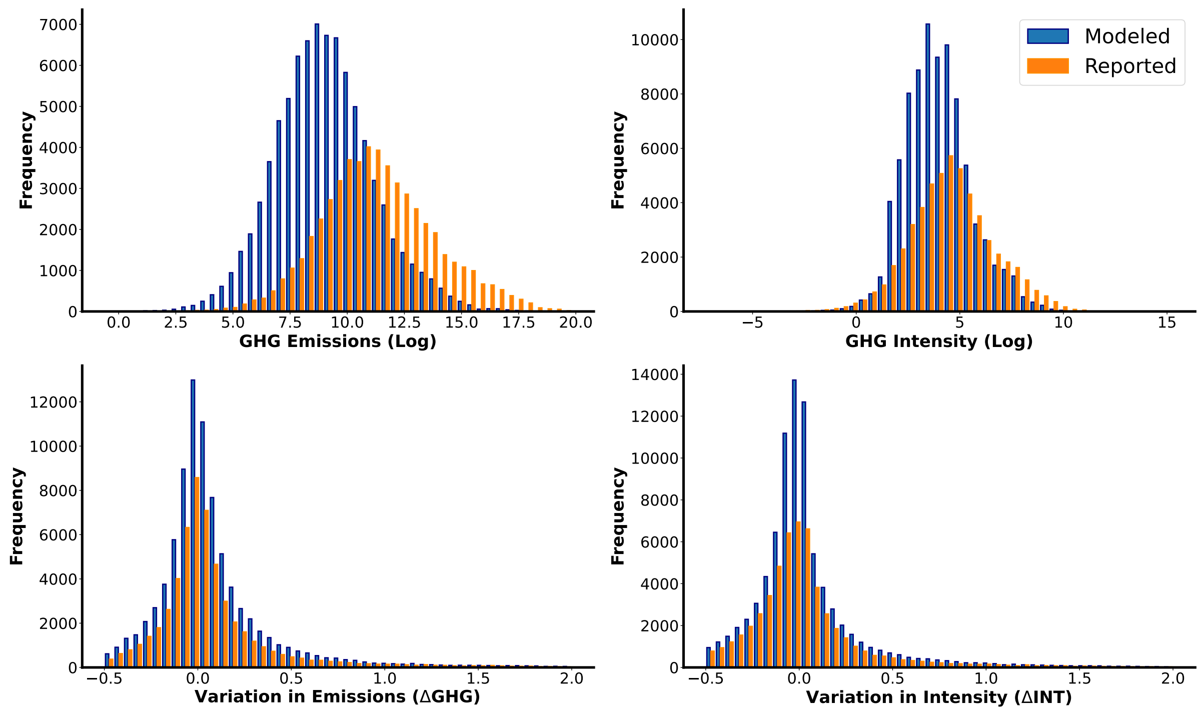


Figure 5: **Distribution of dependent variables - Scope 1+2**.

Targets commitments than non-disclosing companies. However, we do not observe any significant differences in terms of institutional ownership nor sustainable active ownership.

3 The drivers of decarbonization

This section seeks to identify the key determinants of decarbonization across the full sample. We begin in Section 3.1 by estimating linear models, using simple panel regressions. As noted in Section 2.1, all explanatory variables are lagged by one year to preserve a predictive structure. In Section 3.2, we then employ random forest models to capture potential nonlinear relationships between firm characteristics and changes in absolute emissions or emission intensities. Finally, Section 3.3 moves beyond correlation analysis and implements a Difference-in-Differences design to assess the causal impact of some of the key drivers found above.

3.1 Linear panel models

Our analysis starts with linear models, considering the four dependent variables outlined in Section 2.3. Henceforth, we write $y_{t+1,i}$ for the time $t+1$ value of one of these variables for firm i. The model thus reads:

$$y_{t+1,i} = \alpha_g + a_t + \sum_{k=1}^{K} x_{t,i}^{(k)} b^{(k)} + e_{t+1,i}, \tag{4}$$

where α_g are industry and region fixed effect, α_t is the time fixed-effect, and $x_{t,i}^{(k)}$ are the K predictors retained for the analysis, as listed in Table 10 in the Appendix. Note that these predictors include the lagged values of the dependent variables ($y_{t,i}$ may be useful to forecast $y_{t+1,i}$).

Table 2 shows the corresponding estimates. To ease interpretation, we only report t-statistics, as they provide the key message we are interested in: the sign of the relationship, and its statistical strength. We tested two estimators for standard errors, namely Newey and West (1987) and Beck and Katz (1995), but as the latter yielded more conservative values (smaller t-statistics in absolute value), we only provide this version. Significance decisions are emphasized with a green background in the results Table.

	GHG (log)		**INT (log)**		**ΔGHG**		**ΔINT**	
	Reported	Modeled	Reported	Modeled	Reported	Modeled	Reported	Modeled
Direct_GHG_Emission	463.722	710.684			−21.273	−21.211		
Direct_GHG_Intensity			314.076	424.625			−16.812	−23.272
ISS_Environment_Score	1.571	−4.477	0.314	−2.723	−3.576	0.322	−2.872	2.978
ISS_Governance_Score	−2.690	5.906	−1.525	5.220	−2.696	2.623	−0.409	−0.205
ISS_Social_Score	1.653	−2.055	1.063	−1.224	−1.519	−2.825	−2.205	−1.425
Dummy_SBTi_Approved	3.718	−0.888	2.206	−1.012	−9.156	−1.102	−9.089	−2.225
Years_to_Closest_Target	1.928	1.469	1.113	0.332	1.907	2.485	0.812	−0.615
Biodiversity_Score	0.622	2.825	−0.025	−3.130	−1.062	1.775	−1.161	−6.040
Debt_to_Equity	0.957	−0.962	0.749	1.286	0.342	−0.741	−1.848	1.276
Free_Cash_Flow	1.610	0.686	−1.042	0.752	−0.867	−0.700	−0.399	−0.420
Interest_Coverage	−2.093	0.887	1.905	−1.422	−5.536	0.794	−2.624	−1.694
Longterm_Debt	5.844	−1.573	2.525	−1.749	−0.508	0.793	−2.739	−1.172
Longterm_Investment	2.074	1.753	0.777	0.399	−1.168	0.527	0.747	−0.376
Price_to_Book	−1.477	0.045	−2.155	−0.179	−0.296	0.116	−4.824	−0.761
Price_to_Earning	−1.641	−0.465	−0.680	2.349	−0.596	−0.844	−1.931	0.800
Profit_Margin	−2.038	−6.159	0.502	3.744	−0.628	0.240	−0.861	−0.851
R&D_Expenditure	4.709	1.944	−2.318	−0.597	0.310	1.565	−4.761	−3.084
12M_Stock_Return	1.613	1.395	−7.369	−0.136	11.076	0.887	−6.214	−1.451
Institutional_Share	−0.626	−0.185	8.865	0.175	−2.321	−0.721	1.953	−0.431
Internal_Share_Outflow	−0.689	−0.572	−1.769	−2.856	3.227	4.922	−0.020	−1.088
Share_Repurchase	−1.579	0.475	−1.475	−1.952	−1.264	−1.124	0.558	−1.343
Green_Fund_Holding	−0.974	0.112	1.084	0.932	−2.402	1.542	−0.384	0.156
Coal_Energy_pct	−5.681	−6.988	−9.357	−15.797	−1.128	4.900	−4.978	−3.350
Oil_Energy_pct	−5.845	−8.230	−7.956	−12.752	−1.432	−0.843	−4.966	−5.072
Renewable_Energy_pct	−2.043	−0.502	−5.119	−6.434	−1.392	2.145	−7.138	−6.768
Policy_Stringency	2.844	2.236	4.244	6.537	−3.994	0.870	−3.293	5.212

Table 2: **Panel Regression Results - Time + Region + Industry Fixed Effects**. We report the t-statistics for models with time, region, and industry fixed effects - all independent variables are lagged. The regressions control for time effects (year), region effects (country), and industry effects (GICS sector). Standard errors are robust to heteroskedasticity. Colors indicate significance levels: light green for 1% level, dark green for 0.1% level.

Our findings reveal several consistent and robust patterns across both reported and modeled data. A first and most salient result is that past GHG emissions are consistently a strong predictor of future emissions, highlighting the inherent inertia in emission trajectories. Lagged values of absolute emissions (GHG) and intensities (INT) exhibit the largest statistics, confirming that firms with high current emissions tend to maintain high emission levels in subsequent years. At the same time, higher GHG emissions and intensities are associated with greater potential for reduction in the following periods, as reflected by negative t-statistics when explaining future variations. This result aligns with the main findings of Pastor et al. (2024), who document similar persistence and reversion patterns in GHG emission forecasts.

A second key takeaway is that a broad set of variables contributes to explaining both the levels and the changes in absolute GHG emissions and intensities. Nearly every category described in Section 2.1 includes at least one significant predictor. However, notable discrepancies emerge between reported and modeled data, raising concerns about the robustness of current carbon accounting in reported data and estimation models provided by data vendors.

Corporate Sustainability Indicators. ESG-related variables display mixed significance across reported and modeled datasets. On the one hand, environmental scores are significantly associated with lower modeled emissions, while governance scores are positively correlated with higher modeled emissions. On the other hand, these relationships do not hold for the reported sample—where the associations turn negative—indicating that stronger governance and environmental performance are linked to greater reductions in GHG emissions and intensities among disclosing firms. This contrast underscores important discrepancies between firm-disclosed and provider-estimated emissions and aligns with the ambiguity in the academic literature regarding the predictive value of ESG scores, with recent studies such as Kalesnik et al. (2022) and Pastor et al. (2024) reporting diverging results.

In addition, ambitious reduction targets reported to the SBTi have a significant effect on lowering future emissions, both in absolute terms and relative intensities, in line with Bolton and Kacperczyk (2023) and Berg et al. (2024). Companies with ambitious Science-Based Targets reduce their emissions faster than peers without such commitments and by a greater absolute amount, even after controlling for firm fixed effects. It is noteworthy to underline that this effect does not appear in modeled estimations, revealing a potential limitation of current modeling approaches. In addition, in the final year of our dataset, only 16% of firms had approved SBTi targets. This illustrates the potential for a much broader climate impact if such initiatives were to be scaled further. Interestingly, the time horizon of the nearest target does not have a significant effect, raising questions about whether the proximity of target deadlines truly influences decarbonization efforts, or whether other factors such as implementation capacity, regulatory context, or investor pressure play a more decisive role. These potential drivers are explored in greater detail in the following sections.

We do not find corporates' biodiversity footprint to be a significant driver in the reported data sample while it is only associated with lower future emission intensities in the modeled sample. This is calling into question the practical links between climate and biodiversity impacts and could explain previous evidence from Coqueret et al. (2025) suggesting that investors do not yet fully internalize the interconnection between climate and biodiversity risks.

Corporate Fundamental & Market Indicators. Free cash flows show no significant effect, despite expectations that it might facilitate investment in transformative projects. In contrast, long-term financing and financial resilience appear to play a meaningful role. Long-term debt and interest coverage are positively associated with both absolute emission reductions and intensity-based abatement, suggesting that firms with stronger financial buffers and stable financing structures are better positioned for the green transition. Such firms are likely more capable of investing in capital-intensive low-carbon technologies and benefit from easier access to green financing.

Recent financial performance, proxied by 12-month stock returns, reinforces this finding for emission intensities but is positively associated with increased absolute emissions. This asymmetry suggests that while better-performing firms may become more carbon-efficient, this does not necessarily translate into a reduction in absolute total emissions. Such divergence is particularly relevant to the carbon premium debate (Aswani et al., 2024; Bolton and Kacperczyk, 2021, 2024), as it highlights the distinct mechanisms driving improvements in emission intensity versus reductions in absolute levels. This further underscores the importance of distinguishing between these two metrics when assessing real-world climate impact.

Financial indicators such as profit margins, price-to-book, and price-to-earnings ratios are either insignificant or only significant in explaining estimated but not reported emissions. This aligns with findings from Pastor et al. (2024), which suggest that data providers may incorporate such financial ratios into their emissions estimation models, even though these variables appear to have limited practical effect on reported emissions.

Corporate Ownership. The influence of institutional ownership and sustainable fund holdings on GHG emission levels and reduction dynamics appear largely insignificant in both samples, consistent with previous findings from Atta-Darkua et al. (2025); Heath et al. (2023); Ramadorai and Zeni (2024). By contrast, internal share outflows (reflecting reductions in insider or internal ownership) are positively associated with higher corporate GHG emissions. This relationship underscores the importance of governance structures and insider alignment, suggesting that firms with stronger internal ownership may be better equipped to commit to, and deliver on, emission reduction efforts.

Regional Indicators. In line with Coskuner et al. (2020) and Puertas and Marti (2021), our results confirm that the macroeconomic environment plays a central role in shaping emission reduction dynamics. The relative use of coal, oil, and renewables consistently emerges as significant drivers of both absolute and relative emissions. The strong statistical significance observed suggests that the energy transition operates through multiple and complementary channels that collectively support long-term net-zero objectives. This pattern is consistent with the observed effect of policy stringency. The OECD Environmental Policy Stringency Index is positively associated with reductions in both absolute GHG emissions and emission intensities. However, this relationship is not confirmed in the modeled data, highlighting again the discrepancies between the two information sources.

Overall, the results underscore that real-world GHG reductions arise from the interplay between firm-level strategies, financial strength, and broader structural forces. Effective corporate decarbonization requires not only internal efforts, such as innovation and Science-

Based Targets commitment, but also supportive energy systems, market conditions, and policy environments.

As a robustness check, we report in Table 11 (Section C.1 in the Appendix) the regression results using individual (firm) fixed effects. The findings are broadly consistent with those obtained under the industry and region fixed-effects specification, confirming the robustness of the main results. In both models, past GHG emissions remain the strongest predictor of future levels and changes, reflecting the persistence and partial mean reversion of corporate emission trajectories. Most significant predictors remain so, with the exception of R&D expenditure, while institutional ownership and sustainable fund holdings become significant in the individual fixed-effects model, consistent with part of the literature (Kahn et al., 2023; Martiny et al., 2024). The comparison suggests that although firm-specific characteristics explain part of the variation in decarbonization outcomes, structural, sectoral, and regional factors play an equally important role.

In Sections C.2 and C.3 in the Appendix, we deepen our analysis on the variables that explain the trajectories of corporate GHG emissions. The former analysis uses LASSO regressions to select only the most relevant variables in linear models. The latter method reveals the attributes of pools of firms that are grouped according to the variations of their emissions (i.e., decarbonizers versus polluters). We classify firms based on their realized emission trajectories, and identify which characteristics most clearly differentiate firms that successfully reduce their footprint from those whose emissions continue to rise. Both studies mostly confirm our baseline findings which is why they are postponed to the Appendix.

3.2 Inference from nonlinear models - Random forests

To further explain emissions and intensity changes over time, we rely on random forests. The rationale for this choice is two-fold. First, tree-based supervised learning performs very well on tabular data (Grinsztajn et al. (2022), Januschowski et al. (2022), Shwartz-Ziv and Armon (2022)). Second, these methods are readily interpretable via the feature importance metrics that are computed once the models are trained (see, e.g., Molnar (2020)). In addition, statistical tests have also been developed to evaluate the significance of features in these models (see, e.g., Mentch and Hooker (2016)), as well as confidence intervals based on resampling and deleted-d jacknife estimators (Ishwaran and Lu (2019)). We follow this approach in our analysis. In Figure 6, we report the confidence regions for the feature importance, computed over 100 bootstraps. Only the top 10 features are shown.

This section mostly focuses on changes in GHG emissions and intensities, since absolute levels are predominantly determined by their historical values. Accordingly, nonlinear models are expected to offer limited additional insight for absolute emissions but may more effectively capture the underlying dynamics of their variations. Our results confirm the predictive power of many variables identified in Section 3.1. For both GHG abatement and intensity reduction, past GHG emissions remain the most important predictor. Other key variables include variants of the SBTi approved target, 12 month stock returns, fundamental ratios such as price-to-earnings, price-to-book and free cash flows, as well as components of the IEA energy mix.

It is worth noting that firm fundamentals exhibit much stronger predictive power in the

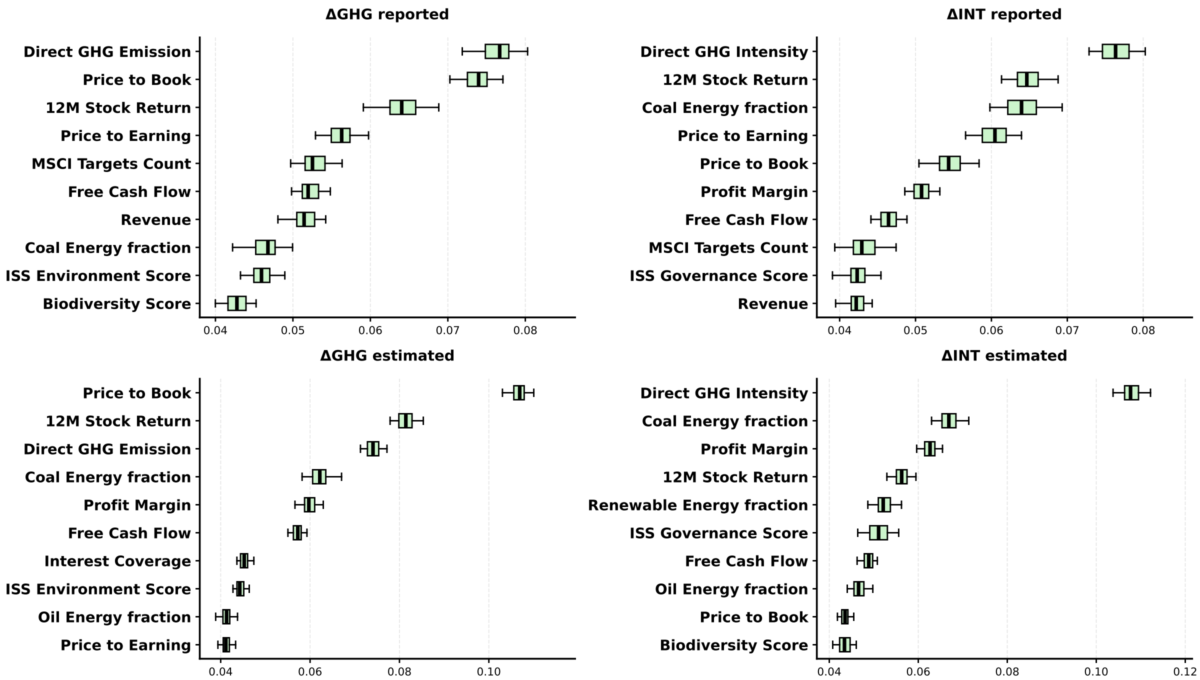


Figure 6: **Confidence regions for feature importance**. We plot the median (center), inter-quartile range (boxes) and 95% interval (whiskers) for the variable importance, following Ishwaran and Lu (2019). The number of bootstrap samples is 100. The forests were fitted separately on the estimated (left panels) and reported (right) variations, as well as on raw emissions (top panels) and intensities (bottom).

nonlinear setting. This suggests that nonlinear relationships exist between financial characteristics and carbon abatement potential, which linear approaches fail to capture. By contrast, environmental policy stringency no longer appears as a key variable, possibly because its effect is already reflected in the energy mix data.

Finally, although the set of key variables remains broadly consistent, strong discrepancies persist between reported and estimated samples, with the relative importance of predictors differing in magnitude. For example, while past direct GHG emissions are the top predictor of absolute emission changes in the reported sample, they fall to the third rank in the modeled sample, where financial variables such as price to book ratios and 12 month stock returns take precedence. This pattern resonates with Pastor et al. (2024), who suggest that data providers may incorporate such financial ratios into their emissions estimation models, even though these variables appear to have limited practical influence on reported emissions.

The above analysis, while crucial to understand associations between heuristic variables and decarbonization paths, only reveals *correlations*. The next natural stage is to try to uncover *causal* relationships. This task is inherently challenging, as no perfect random experiment is available for us to exploit. Nevertheless, among the top predictors identified, the adoption of ambitious Science-Based Targets offers a particularly promising setting for causal analysis. We develop this investigation in the following Section.

3.3 The causal impact of Science-Based Targets on future corporate emissions

Being able to determine the drivers of voluntary corporate climate commitments is a central question in sustainable finance. Among these, the Science-Based Targets initiative (SBTi) has emerged as a leading framework for aligning corporate greenhouse gas (GHG) emissions with the goals of the Paris Agreement. The number of companies adopting Science-Based Targets (SBT) has grown exponentially since 2015, now covering a substantial share of global market capitalization (Bendig et al., 2023). Yet, despite this rapid diffusion, empirical evidence on whether adopting SBTs leads to measurable emissions reductions remains limited.

We identify the adoption of Science-Based Targets as a key factor in explaining the cross-section of firms' GHG emissions levels and abatement. Our findings suggest that firms with SBTi-validated targets are, on average, higher absolute emitters but subsequently reduce both their absolute emissions and their emission intensity faster than peers without ambitious climate targets. These findings are consistent with the SBTi Progress Report (2021), which estimates that companies with approved SBT targets curtailed their emissions by 12% in 2020, compared to an average 10% reduction across their respective economies. Yet, while these patterns are indicative of a relationship between SBT adoption and emissions trajectories, they do not yet reveal the causal impact of adopting ambitious voluntary targets.

Similarly to voluntary carbon disclosure (Berg et al., 2024; Chen et al., 2025), SBT adoption is likely endogenous. Freiberg et al. (2021) contend that "*firms with a track record of setting and achieving ambitious carbon targets are more likely to set science-based targets*", and Bjørn et al. (2022) emphasize that adoption has accelerated among large firms in high-income countries with established reputations for managing climate impacts. Consistently with these findings, Bolton and Kacperczyk (2023) document that firms announcing SBT commitments tend to be lower emitters and more climate-conscious ex-ante. While these studies suggest that SBT adoption is associated with increased climate action, such as greater investment in carbon-reduction projects and higher expected emission and cost savings, they, too, do not identify the causal effect of adoption on realized emissions, and at best imply that this effect may be limited. Disentangling this causal impact is therefore essential to assess whether voluntary corporate commitments can serve as effective instruments to accelerate the transition toward a net-zero economy.

The following section contributes to filling this empirical gap by examining the causal impact of adopting SBT on firms' GHG emissions performance using a difference-in-differences research design.

3.3.1 Identification Strategy

We address the endogeneity of voluntary climate commitments using an instrumental variable (IV) approach within a difference-in-differences (DiD) framework. Our strategy exploits the staggered rollout of the unified CDP–ICLEI regional reporting system, initiated in 2018.[4] This reform established an authoritative disclosure platform at the country level, enhancing the visibility of local climate action and strengthening normative and peer pressures on firms within participating jurisdictions—mechanisms expected to increase the likelihood of

[4] ICLEI and CDP launch unified reporting system. (04/24/2018)

voluntary climate commitments. Empirical evidence from Seo (2021) supports this assumption, showing that firms are more likely to disclose when their industry peers do, particularly when they rely more heavily on external financing, implying that peer disclosure amplifies reputational and visibility pressures in capital markets.

The staggered adoption across jurisdictions provides quasi-exogenous variation in the timing of this external pressure. We define treatment cohorts based on the year the CDP-ICLEI system was implemented in a firm's headquarters country: 2018 (covering 42 countries including major economies),[5] 2019 (5 additional countries), 2021 (2 countries), and 2023 (1 country).[6] This variation is plausibly uncorrelated with firms' pre-existing emissions trajectories after conditioning on observable characteristics.

A critical concern for our identification strategy is that country-level CDP implementation may coincide with other environmental policy changes that independently influence firm emissions, potentially violating the exclusion restriction. We address this concern, along with related threats, through three complementary validity tests summarized in Table 14 in the Appendix.

First, Panel A assesses instrument exogeneity by regressing 19 pre-treatment firm characteristics on the CDP exposure indicator. In particular, we test whether CDP rollout correlates with environmental policy stringency, as measured by the OECD Environmental Policy Stringency Index. The instrument passes 18 of 19 balance tests (a 94% pass rate), with only the internal share outflow variable showing significance—likely due to high imputation rates for missing data. Importantly, the CDP treatment indicator and its lags are statistically indistinguishable from zero when regressed on policy stringency (*Independence of CDP rollout and environmental policy*).

Second, Panel B implements a triple-difference specification that interacts CDP exposure with policy stringency levels. The insignificant triple interaction terms ($p = 0.586$ for direct emissions; $p = 0.062$ for intensity) suggest that the instrument's effect does not systematically differ between eventual SBTi adopters and non-adopters, nor across varying policy regimes. This finding further supports the exclusion restriction (*Policy stringency does not alter the influence of CDP encouragement across corporate groups*). Finally, Panel C examines differential pre-trends between future SBTi adopters and non-adopters within not-yet-treated comparison groups. A joint test of pre-period leads reveals no significant differences ($p > 0.79$ for both outcomes), supporting the parallel trends assumption. Figure 7 visually confirms this result, showing no significant pre-treatment effects in the first stage.

Taken together, these results suggest our identification strategy is robust. Accordingly, we exclude the OECD Environmental Policy Stringency Index from our control variables to avoid potential multicollinearity with country-level indicators.

The staggered nature of the policy rollout, while providing identification power, introduces methodological challenges. Traditional two-way fixed effects (TWFE) DiD estimators can produce biased estimates when treatment timing varies and effects are heterogeneous, as already-treated units may receive negative weights when used as controls for newly-treated units. We therefore implement the robust estimator from Callaway and Sant'Anna (2021), which constructs valid counterfactuals using only not-yet-treated or never-treated firms for

[5] The 2018 cohort represents 88% of treated firms in our sample. Detailed country-year rollout information is provided in Appendix Table 15.

[6] Historic review of Cities, States and Regions reporting through CDP-ICLEI

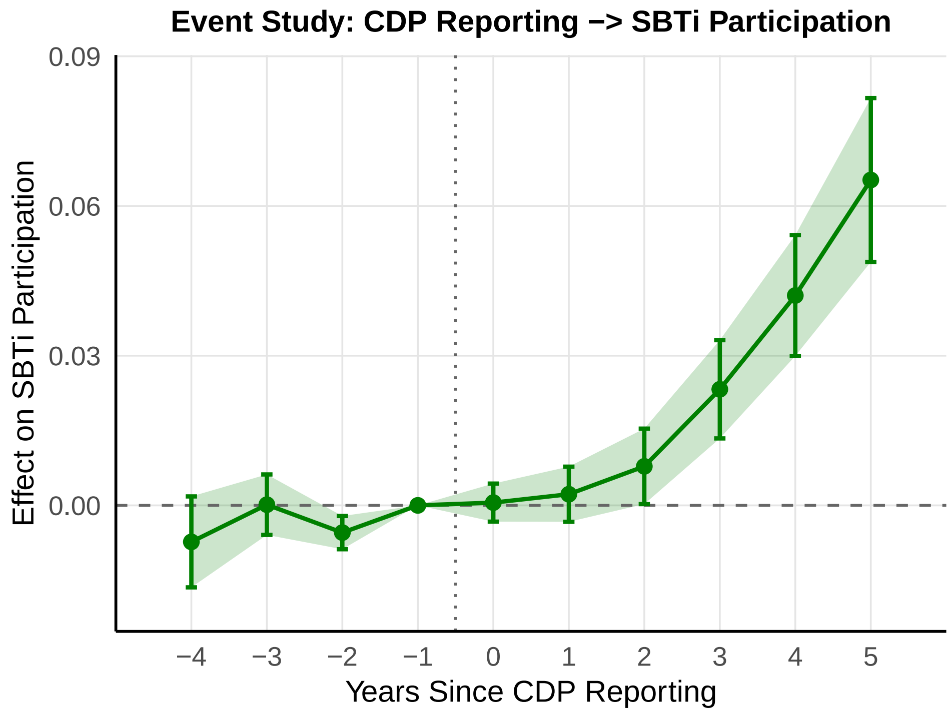


Figure 7: **Event study**. We plot of the effect of the CDP regional reporting program on SBTi adoption. The plot displays the dynamic coefficients with 95% confidence intervals. The absence of statistically significant effects in the pre-treatment periods (t < 0) supports the parallel trends assumption for the first stage of our IV design.

each cohort-time comparison. This approach avoids the "bad control" problem inherent in pooled TWFE specifications while allowing for transparent aggregation of treatment effects. Due to fuzzy adoption, we follow De Chaisemartin and D'HaultfŒuille (2018) and form a Wald-DiD at the cohort–time level:

First Stage (Cohort-specific):

$$\text{SBTi}_{i,g,t} = \pi_{g,t} \cdot \text{CDP}_{i,g,t} + X'_{i,g,t}\Gamma_{g,t} + \epsilon_{i,g,t} \tag{5}$$

Reduced Form (Cohort-specific):

$$\Delta\text{Emissions}_{i,g,t} = \beta_{g,t} \cdot \text{CDP}_{i,g,t} + X'_{i,g,t}\Theta_{g,t} + u_{i,g,t} \tag{6}$$

where $\text{CDP}_{i,g,t}$ equals 1 for treated firms in cohort g at time $t \geq g$, and 0 for control firms. We deploy the Callaway and Sant'Anna (2021) procedure to avoid any bad control. The cohort-specific Local Average Treatment Effect is $LATE_{g,t} = \beta_{g,t}/\pi_{g,t}$ using the R package coded by the authors.

To validate the CS-WALD aggregation and ensure robustness to heterogeneous effects, we also implement the stacked two-period 2SLS (STS) estimator proposed by Miyaji (2024). Following their approach, we construct cohort-specific 2×2 comparisons using not-yet-treated controls. We exclude event-time periods $\ell \in \{0, 1\}$ to avoid vague implication impact and require $\ell \geq 2$ with a 2-year anticipation gap. We then aggregate these estimates using precision-compliance weights:[7]

$$\text{LATE}_{STS} = \frac{\sum_{g,\ell} w_{g,\ell} \cdot \beta_{g,\ell}}{\sum_{g,\ell} w_{g,\ell} \cdot \pi_{g,\ell}} \tag{7}$$

where $w_{g,\ell} = \frac{\max(\pi_{g,\ell},0) \cdot n_{g,\ell}}{\text{SE}(\beta_{g,\ell})^2}$, and we restrict to cells with $F_{g,\ell} > 10$ and $\pi_{g,\ell} > 0$.

3.3.2 Main Results

Our primary IV–DiD results, presented in Table 3, indicate a significant causal effect of SBTi adoption on firms' emissions intensity. The first-stage estimates (Panel A) confirm the instrument's relevance: exposure to the CDP program increases the likelihood of SBT adoption by 3.7 percentage points. This finding supports the validity of our identification strategy and highlights the influence of government signaling, heightened scrutiny, and peer pressure in shaping corporate climate behavior.

The reduced-form estimate (Panel B) shows that the CDP program leads to a 0.047 percentage point annual reduction in GHG intensity. Combining these estimates, the Local Average Treatment Effect (LATE, Panel C) reveals that for firms induced to participate by the program ("compliers"), SBTi adoption causes an additional 1.26% annual reduction in emissions intensity.

To address concerns about heterogeneous treatment effects and weak identification, Panel D reports results from an alternative stacked two-stage (STS) estimator as a robustness check. The STS estimate for GHG intensity (-1.29%) closely aligns with our main estimate, while the point estimate for absolute emissions (-0.81%) suggests potential reductions, though statistical power remains limited.

While the effect on absolute emissions does not reach statistical significance, it indicates that firms initially prioritize efficiency improvements in their decarbonization strategies. The negative point estimate from the STS method implies that absolute emission reductions may eventually occur over time, though limited post-treatment observations and sample size constraints prevent definitive conclusions.

3.3.3 Robustness tests

Table 15 in the Appendix reveals important heterogeneity across adoption cohorts. The 2018 cohort, representing 90% of treated firms, provides strong identification (first-stage F-statistic on the excluded instruments $F = 87.3$) with positive compliance. In contrast, the 2019 cohort exhibits weak instruments ($F = 6.2$) and perverse first-stage effects, while later cohorts contain too few observations for reliable inference. This concentration of identifying variation in early adopters motivates our use of strong-instrument filtering in both estimation approaches.

Overall, our findings indicate that adopting ambitious Science-Based Targets, particularly under quasi-regulatory pressure, reduces uncertainty and reflects genuine alignment

[7]The STS aggregation uses approximately 25 strongly-identified (g, ℓ) cells from the 2018 cohort, ensuring robust identification while avoiding contamination from weak or reverse-compliance cohorts.

Panel & Variable	Estimate	Std. Error	95% Conf. Interval
Panel A: First Stage			
CDP $\rightarrow$ SBTi Adoption (π)	0.037***	0.004	[0.029, 0.045]
Panel B: Reduced Form			
CDP $\rightarrow$ Direct GHG Emissions (%)	0.001	0.008	[-0.015, 0.016]
CDP $\rightarrow$ GHG Intensity (%)	-0.047***	0.008	[-0.063, -0.031]
Panel C: Wald LATE (SBTi $\rightarrow$ Emissions)			
SBTi $\rightarrow$ Direct GHG Emissions (%)	0.013	0.218	[-0.414, 0.441]
SBTi $\rightarrow$ GHG Intensity (%)	-1.264***	0.254	[-1.761, -0.767]
Panel D: STS Outcome (SBTi $\rightarrow$ Emissions)			
SBTi $\rightarrow$ Direct GHG Emissions (%)	-0.813	0.5228	[-1.848, 0.222]
SBTi $\rightarrow$ GHG Intensity (%)	-1.291***	0.580	[-2.428, -0.154]

Table 3: **Main IV-DID Results using Aggregated Callaway-Sant'Anna Estimator.** This table reports aggregated treatment effects. The LATE is calculated as the ratio of the reduced-form estimate to the first-stage estimate. Standard errors are clustered at the firm-level via cluster bootstrap that re-estimates FS and RF jointly (Equations (5) and (6)). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Significance stars for panels A and B are based on the 95% confidence interval not covering zero.

with net-zero objectives. The notable drop in emissions intensity among complying firms suggests that these commitments lead to real actions rather than symbolic pledges. Yet, the effect appears to unfold gradually or coincide with structural shifts in SBTi governance, supporting the call of Tilsted et al. (2023) for a strengthened science-based approach. Nonetheless, such a framework remains helpful for investors to identify credible signals of future efficiency gains compared to voluntary, unverified targets.

4 Forecasting & Credibility Assessment

As highlighted in our introduction, sharp forecasting skills can considerably help investors seeking portfolios decarbonization in accordance with Net-Zero pledges (see, e.g., Bolton et al. (2022) and Le Guenedal et al. (2022)). This section analyzes the benefits that can be brought by machine learning in predicting corporate emission trajectories. In order to evaluate our results, we will consider two heuristic benchmarks for future emissions. The first one, the *"status quo"* benchmark considers that emissions or intensities remain unchanged from a reporting year to the following year. It is an agnostic view that refrains from making any directional guess. The second benchmark is a *"trend"* extrapolation. We fit a linear model on six years of data and predict a value that follows the straight line.

However, as can be seen in Figures 8 and 9, none of these benchmarks are fully satisfactory in practice. Indeed, while for large companies (see Figure 8) the linear extrapolation properly fits the data, it is not the case for many other companies, like small firms or companies operating in hard-to-abate sectors. We therefore investigate if alternative approaches can add value by generating more credible emissions forecasts and, in turn, help investors

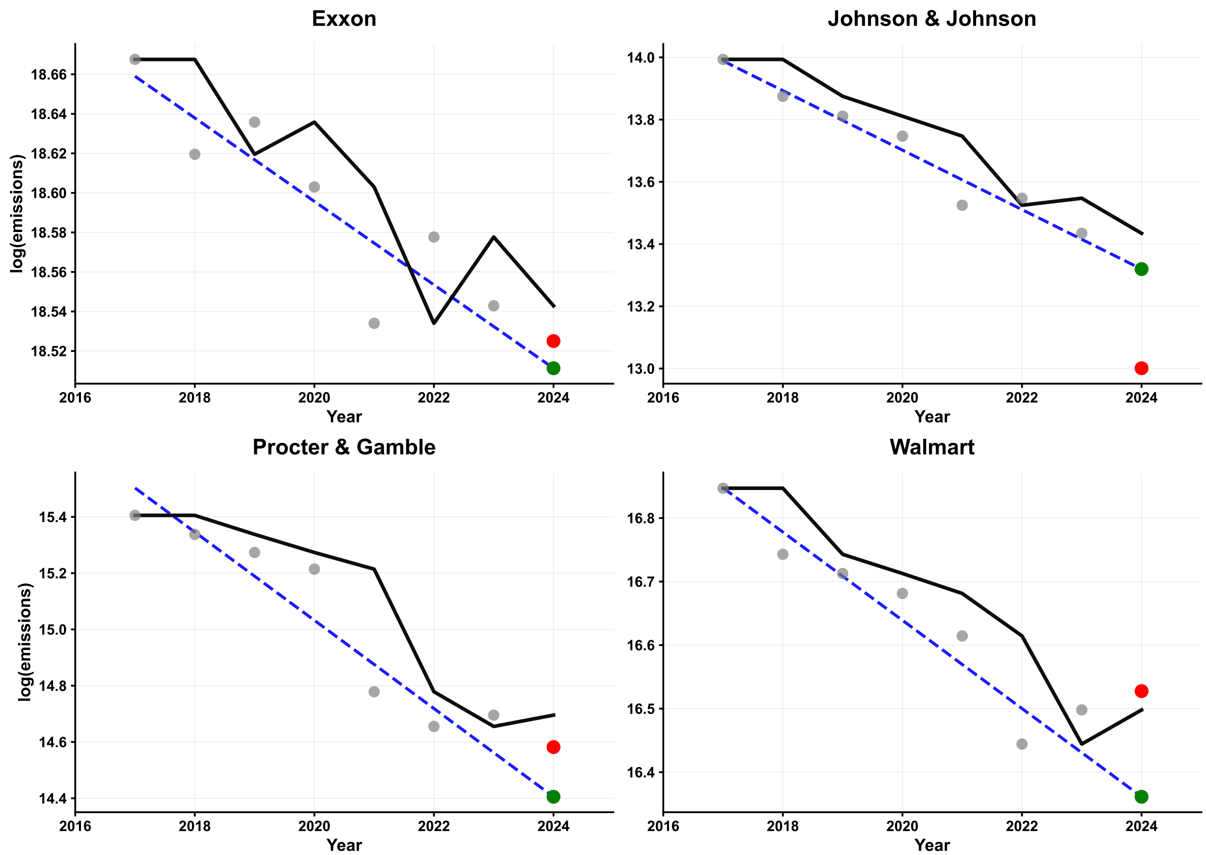


Figure 8: **Extrapolation for large firms**. We plot emissions for the four largest firms that are fully covered in the sample (i.e., with zero missing data). Grey dots are reported values through 2023. The blue dashed line is an OLS trend fitted on observations up to 2023 and extrapolated to 2024 (green dot). The realized 2024 value, when available, is shown in red.

construct portfolios with more efficient decarbonization outcomes.

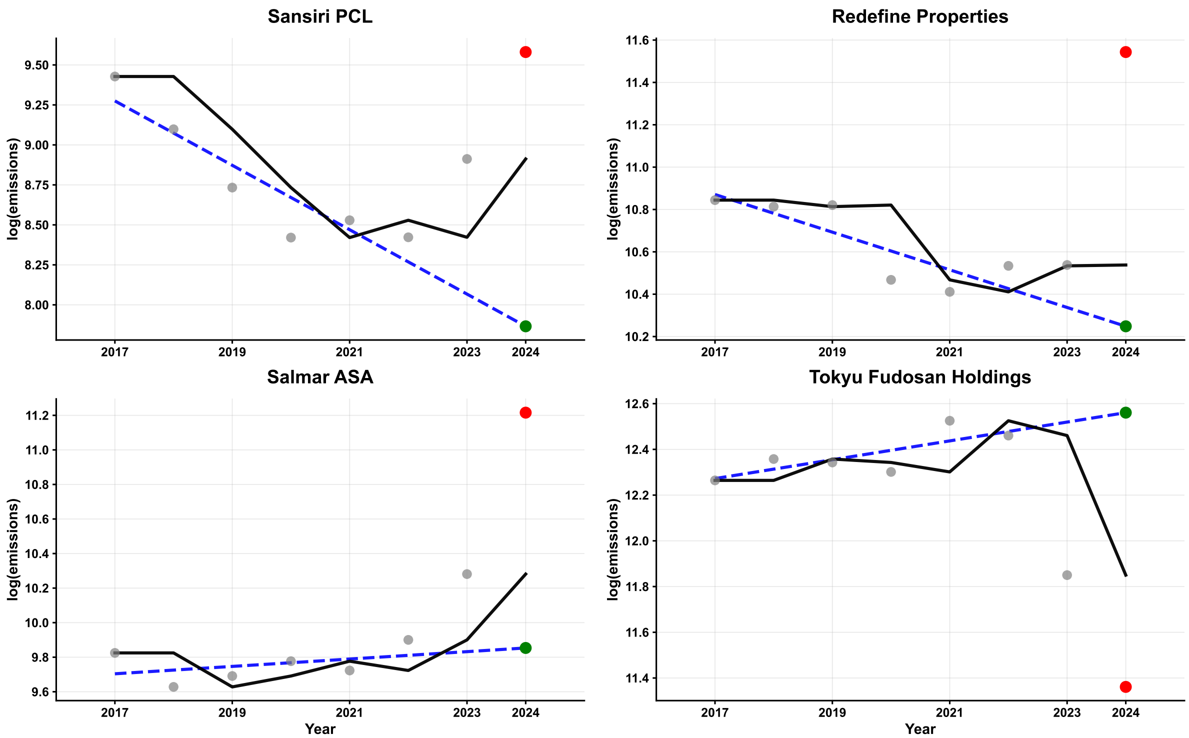
4.1 Out-of-sample analysis

Forecasting raw emissions and intensities. In this section, we analyze the out-of-sample accuracy of our emission forecasts. Importantly, given the discrepancies between reported and modeled emissions highlighted in Section 3, we restrain our analysis on companies with reported emissions. Due to data restriction, the training set consists of all observations, except from the last year (2023), while the test set encompasses all points from this last year.

Formally, the model is

$$y_{t+1,n} = f(\boldsymbol{x}_{t,n}) + e_{t+1,n}, \tag{8}$$

which is the panel specification commonly used in such settings (see, e.g., Gu et al. (2020) in the case of return prediction). In the model, f is a potentially nonlinear model and $e_{t+1,n}$ are the residuals (errors). The predictors $\boldsymbol{x}_t$ are the variables considered previously in the



Panel A: Small-size firms **Panel B: Hard-to-forecast sector**

Figure 9: **Large errors from extrapolation.** The left panel displays trajectories for small-sized firms that exhibit a sharp break in emissions in 2023. The right panel pertains to firms from the Real estate sector (GICS 60) whose emissions dynamics are particularly hard to forecast. The first six observations are in grey, the extrapolated value is shown in green, and the realized value for 2023 is shown in red.

paper. Note that we use the notation y as is customary in the literature; in our study, it will stand for (log) emissions, (log) intensities, and variations thereof. We work with logarithmic values to avoid the squared errors to be heavily biased towards the large emitters only. We also highlight the lag between the dependent variable (at time $t+1$) and the predictors (in t). In terms of algorithms, we stick to random forests, as our findings suggest that nonlinear relations are meaningful in explaining firms' emissions. As is customary in machine learning practice, we run the analysis on several values of the important parameters of the model, which are summarized in Table 16 in the Appendix.

Finally, in order to align with our theoretical motivation, we report the correlation ρ between our forecast and the true distribution of GHG emissions. While it is customary to focus on metrics like the mean squared errors (MSE), such indicators put a lot of focus on the variance of both the predictions and the realized values which are not of particular interest when it comes to picking decarbonizing firms. It is truly the correlation component of the MSE that matters (as is shown in Equation (3)).

To assess the quality of our forecast, we rely on the benchmarks presented above: the *"status quo"* forecast (the value of the preceding year) and the *"trend"* estimate, computed as the linear extrapolation of past values. For the latter benchmark, we must restrict the set to the firms for which all points between 2016 and 2022 are available so that the extrapolation makes sense. Following our theoretical model, we measure the performance of our forecasts on the basis of their correlation with the true distribution.

Our main results are summarized in Table 4. The first rows (Panel A) reports the correlation (ρ) between the predicted variables (in columns) and their realized outcome when predicting the four dependent variables (in columns). We report the correlation both on the full sample of 2022 and also on the subset of firms for which all points are available (the *"sub"* columns in the table). The next batch of rows (Panel B) provides the minimum, median and maximum values of ρ across the 108 hyperparameter combinations of Table 16.

dependent variable	**GHG**		**ΔGHG**		**INT**		**ΔINT**	
sample	full	sub	full	sub	full	sub	full	sub
PANEL A: benchmarks								
status quo (constant)	**0.945**	**0.995**	0.001	-0.002	**0.916**	**0.992**	0.014	-0.017
trend (extrapolation)	-	0.993	-	0.037	-	0.988	-	-0.022
PANEL B: random forests (statistics across 108 HP combinations)								
min	0.942	0.992	0.194	0.214	0.912	0.989	0.144	0.162
median	0.945	0.995	0.202	0.220	0.918	0.992	0.152	0.176
max	0.946	0.996	0.206	0.228	0.918	0.992	0.160	0.183
PANEL C: learning from errors (statistics across 108 HP combinations)								
min	0.944	0.996	-	-	0.916	0.991	-	-
median	0.944	0.996	-	-	0.917	0.992	-	-
max	0.945	0.996	-	-	0.917	0.992	-	-

Table 4: **Out-of-sample forecasting accuracy**. We gather the *correlations* between predicted and realized values for the year 2023, which serves as test set. In Panel A, we report two benchmarks: the constant value (from 2022) and the extrapolated one, from 2016 to 2022. For the extrapolation, the estimations are run on the set of firms for which all eight data points of reported emissions are available (we refer to this set as "*sub*"). In Panel B, we report the statistics (minimum, median and maximum) of the correlation when spanning the hyperparameter space with 108 combinations. Panel C shows results from the 'learning from errors' approach outlined below. Training is performed on 2016-2022 data with out-of-sample accuracy evaluated on 2023 data.

On the one hand, our results show that sophisticated algorithms do not systematically outperform heuristic benchmarks when forecasting annual levels of absolute GHG emissions or emissions intensities. Interestingly, the constant benchmark performs better, on average, than linear extrapolation, confirming our previous finding that corporate emissions adjust only slowly over time. It also reinforces the view that historical trend estimates are reliable only for a subset of companies (mostly large firms), rather than for the full investment universe.

On the other hand, we find that our forecasts substantially outperform heuristic benchmarks when predicting percentage changes in both absolute emissions and emissions intensities. Although the levels of correlations remain modest, they are close to the ones required in our theoretical simulations to achieve the -9% target rate of reduction in annual greenhouse gas emissions. Consistent with our framework, and as shown in Table 6 in the Appendix, this suggests that leveraging our nonlinear, data-driven approach to forecast one-year-ahead emissions abatement could enable investors to nearly double their portfolio decarbonization rate compared to using naïve linear extrapolation,[8] thereby meaningfully accelerating the transition toward a more resilient Net-Zero economy.

Forecasts built on relative changes. Building on the strong performance of our abatement forecasts, we examine whether level forecasts can be improved by reconstructing them from lagged emissions and predicted abatement rates. Formally, this yields

$$\log(\mathrm{GHG}_{i,t-1}) + \hat{\Delta}^{Level}_{i,t}, \quad \log(\mathrm{INT}_{i,t-1}) + \hat{\Delta}^{Int}_{i,t},$$

where $\hat{\Delta}_{i,t}$ is the abatement predicted from the model (either for emissions or for intensities).

Our results are reported in Panel C of Table 4. The correlation improves relative to Panel B, suggesting that this approach yields better forecasting performance compared with the direct random-forest models. However, the improvement is limited, and the results remain broadly comparable to those of the constant benchmark.

Forecasting binary changes. While predicting future levels of GHG emissions and intensities remains difficult—even just one year ahead—and our results already offer value for portfolio decarbonization, we believe it may also be useful to frame decarbonization as a binary outcome rather than as a continuous prediction.

A firm is labeled *"decarbonizing"* at year t if its emissions in t are lower than in year $t-1$. We will use the same designation for intensity reduction. The corresponding value for the dependent variable is one, and the only other possible value is zero if the firm is projected to increase its emissions in 2022. We then consider three benchmark options:

1. *"constant trend"*: a firm is forecasted to decarbonize in 2022 if it had reduced its footprint in 2021 compared to 2020, i.e., decarbonization is expected to continue.
2. *"regression forecast"*: we use the linear trend from 2016 to 2021 to forecast a value in 2022. If it is lower than that in 2021, then we predict that the firm will lower its emissions.
3. *"regression slope"*: given the above trend, we predict lower emissions if the slope is negative.

The forecasting accuracy of these benchmarks is reported in Panel A in Table 5. The accuracy is simply the proportion of correct predictions. When the samples are perfectly balanced, a natural yardstick is 50%, which corresponds to a coin toss.

In Panel B, we report the statistics (minimum, median and maximum) of the accuracy when spanning the hyperparameter space with 108 hyperparameter combinations. In this case, the Random Forest models are used to generate predictions for 2023. If the figure is

[8] Table 6 reports expected yearly decarbonization gains in percentage points for different portfolio adjustment intensities α from our model. In the full sample, naive linear extrapolation implies an increase in average emission and intensity, whereas using random forest forecasts results in the expected abatement of 6.4% p.p. titling gain. Our results are even better when considering emissions intensities (Panel B).

below that of 2022, then again, the firm is flagged (predicted) as a decarbonizer. In Panel C, it is the binary variable of decarbonization that is directly predicted.

Our classification models achieve accuracy rates above 50% and consistently outperform heuristic benchmarks, especially in Panel C. For absolute emissions, accuracy reaches, or exceeds, 60%, with slightly lower performance for emissions intensities, likely due to greater volatility in expected revenues, as shown in the plot next to Table 5. These results are practically useful, as they allow investors to flag firms that are unlikely to meet their near-term decarbonization trajectories and engage with them to strengthen short-term emissions reduction actions.

dependent variable	$1_{\Delta GHG<0}$		$1_{\Delta INT<0}$	
sample	all	sub	all	sub
PANEL A: Benchmarks				
constant trend	0.544	0.553	0.534	0.545
linear prediction	-	0.556	-	0.546
slope sign	-	0.543	-	0.529
PANEL B: Random Forest predictions (regression)				
min	0.525	0.519	0.494	0.482
median	0.535	0.537	0.511	0.509
max	0.540	0.550	0.527	0.534
PANEL C: Random Forest predictions (classification)				
min	0.582	0.633	0.533	0.556
median	**0.588**	**0.639**	**0.551**	**0.573**
max	0.592	0.645	0.562	0.585

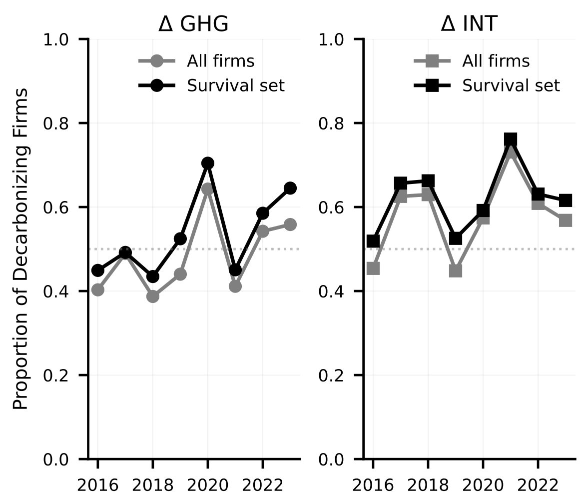


Table 5: **Classification of decarbonization**. In the left table, we report the accuracy of the predictive models for the year 2023. In Panel A, we report three benchmarks: i) the decarbonization status of 2022 is assumed to remain, ii) the linear prediction is used to determine if emissions or intensities will decrease and iii) the slope of the linear regression is used (a negative slope meaning decarbonization). For the latter two, the estimations are run on the set of firms for which all eight data points of reported emissions are available (we refer to this set as "*sub*"). In Panel B, we report the statistics (minimum, median and maximum) of the accuracy when spanning the hyperparameter space with 108 hyperparameter combinations. In this case, the RF models are used to generate predictions for 2023 and decarbonization is inferred from the realized 2022 emissions and intensities. In Panel C, it is the binary variable of decarbonization that is directly forecasted. In the right plot, we depict the evolution of the proportion of decarbonizing firms, both with respect to GHG and INT, and also discriminating between reported (black) and estimated (grey) figures.

4.2 A focus on hard-to-abate sectors

While our previous results, presented for the full universe of firms, already demonstrate value for investors seeking to reduce the GHG footprint of their portfolios, it is well established that decarbonization pathways vary substantially across sectors. A few industries are considerably harder to abate, yet they play an essential role in achieving the global net-zero transition.

These industries, commonly referred to as *"hard-to-abate"* sectors, are generally understood to include the Energy, Materials, Utilities, and Industrials sectors. Together, they account for a disproportionate share of global emissions (above 40% according to Net Zero Industry Tracker), and therefore require especially ambitious and innovative decarbonization strategies. This need is further underscored by our data (see Figure 10), which show that, except for Utilities, all hard-to-abate sectors were among the worst performers in terms of absolute emission reductions over 2016–2023. Recent analysis by ISS[9] also indicates that 56% of companies within hard-to-abate sectors have not set any GHG reduction targets by 2050 and only 22% of them have set GHG emission reduction targets encompassing all Scope 1, 2, and 3 emissions. Recognized as "priority sectors" by the Net-Zero Asset Owner Alliance (NZAOA), their successful transition is fundamental to the credibility and effectiveness of the global net-zero pathway.

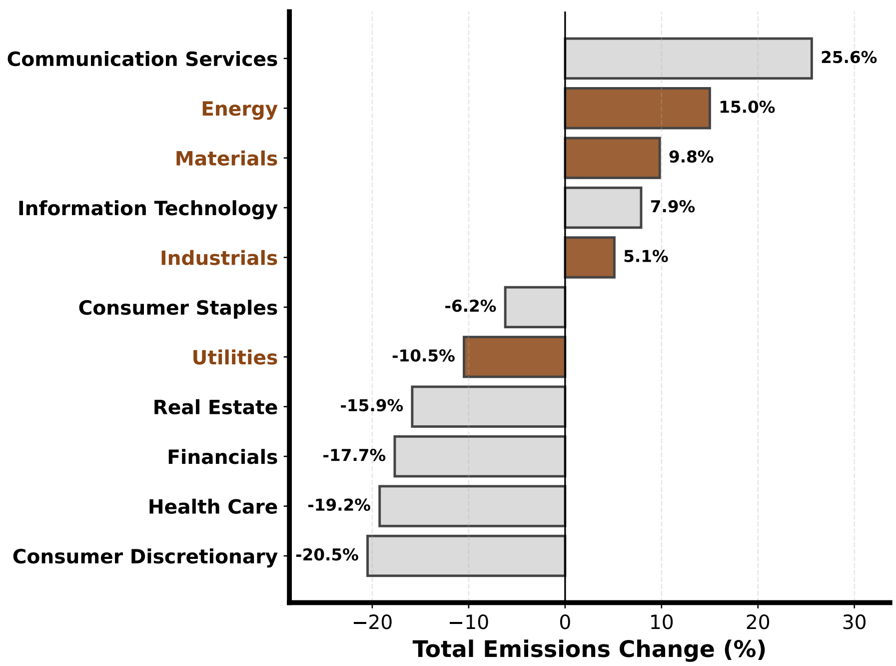


Figure 10: **Cumulative Emission Reductions by Sector**. This figure presents the cumulative emission reduction rates across sectors over the period 2016–2023. The sample is firms with complete emission reporting throughout the entire period. Hard-to-abate sectors (highlighted in brown) include energy, materials, utilities, and industrial sectors.

Therefore, we next conduct a sector-level analysis to evaluate whether applying our approach to hard-to-abate sectors can further enhance predictive accuracy and practical relevance.

Figure 11 shows the average predicted out-of-sample correlation aggregated at the sector level. The results reveal significant heterogeneity across sectors. Importantly, most hard-to-abate sectors rank within the top half in terms of forecasting accuracy (left panel) for absolute GHG emissions changes, which is particularly relevant for connecting our empirical findings to the theoretical framework.

[9] Beyond Net Zero Pledges: Navigating Hard-to-Abate Sectors, 2025

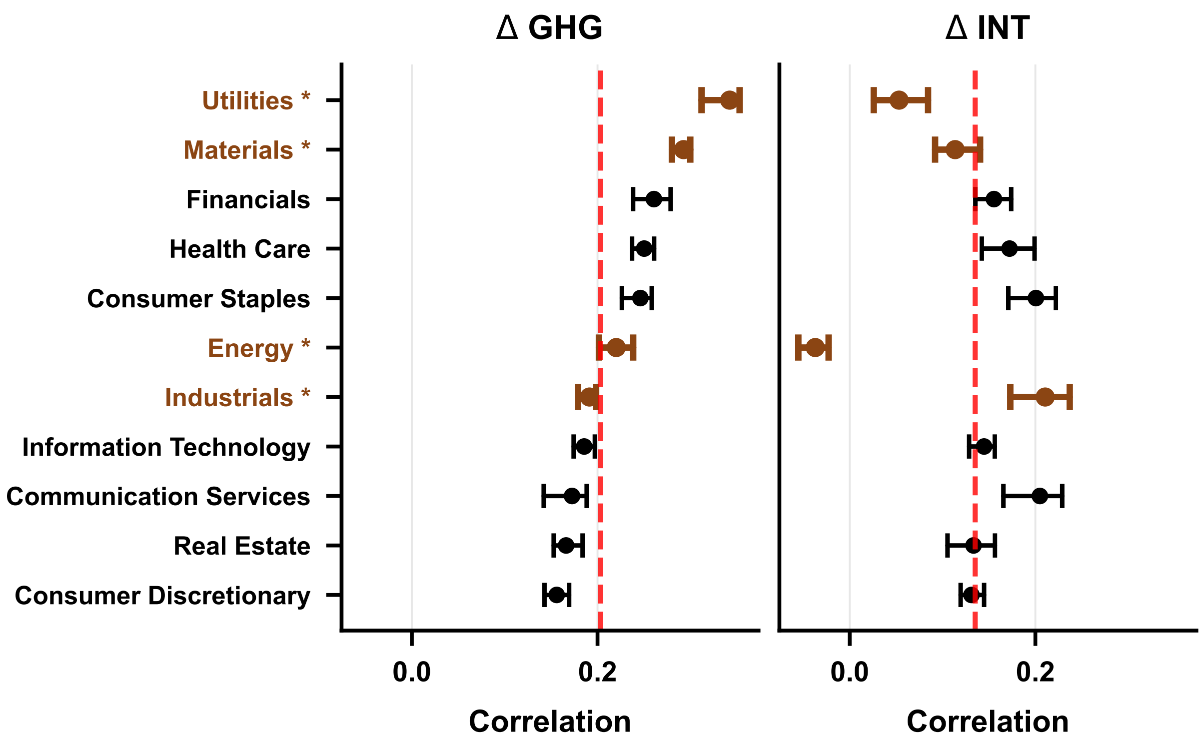


Figure 11: **Sector variations**. This plot shows the correlation of the out-of-sample predictive models from Table 4 (Panel B) aggregated at the sector level. The error bars mark the variations across the hyperparameters. The dashed red lines show the mean correlation over all sectors.

These results indicate that data-driven methods can meaningfully enhance the identification of firms within hard-to-abate sectors that are likely to reduce future emissions. Such an approach enables investors to maintain exposure to these sectors while selecting firms with credible transition prospects, thereby strengthening portfolio alignment with Net Zero objectives.

4.3 Bottomline: achievable decarbonization rates

We are now endowed with reliable parameter values for μ and ρ, and we are thus able to compute the expected decarbonization rate δ from Equation (3). For μ, we consider two scenarios. First, the "*business as usual*" case with all firms kept in the universe and emission trajectories that keep the same distribution as in our sample; this implies $\mu = +6.4\%$, i.e., a negative decarbonization. Second, we only focus on firms that have adhered to the SBTi.

In Table 6, we report the related quantities for three values of α, which captures the investor's willingness to decarbonize smoothly ($\alpha = 1$), moderately fast ($\alpha = 2$), or rapidly ($\alpha = 3$). While we report results for intensities for the sake of completeness, the true yardstick for decarbonization is raw emissions.[10] The figures in the Panel B suggest that the best decarbonization rate we can achieve is -4.5% on the subset of firms that have adhered to the SBTi program. This is based on 2023 emission data compiled in 2024 and on random forest predictions that reach a skill level of 16.2%. Other forecasting methods yielded lower skill and are not reported in the table. It is worth noting that this best decarbonization rate is only half of the one required to substantially decarbonize the portfolio by 2050 (see Figure 1). A rate of -4.5% compounded over 25 years means -68.5%, which is both sizable, but also shy of the initial objective.

Panel A: $\triangle$GHG with random forest predictions

	Full sample: $\mu = +6.4\%$, $\rho = 0.203$		**SBTi**: $\mu = -3.9\%$, $\rho = 0.162$	
α	Gains (%)	Expected δ (%)	Gains (%)	Expected δ (%)
1	-2.1	$+4.3$	-0.2	-4.1
2	-4.3	$+2.1$	-0.4	-4.3
3	-6.4	$+0.0$	-0.6	-4.5

Panel B: $\triangle$INT with random forest predictions

	Full sample: $\mu = +1.8\%$, $\rho = 0.135$		SBTi ($\mu = -6.3\%$, $\rho = 0.134$	
α	Gains (%)	Expected δ (%)	Gains (%)	Expected δ (%)
1	-1.5	$+0.3$	-0.1	-6.4
2	-3.1	-1.3	-0.3	-6.6
3	-4.6	-2.8	-0.4	-6.7

Table 6: **Carbon-reduction gains with portfolio adjustment intensities**. We report δ defined in Equation (3). The gains from forecasting skills are defined as Gains $= -\kappa\rho$ (in percentage points). They are *negative* when the portfolio manager further *reduces* the footprint thanks to her forecasting skill. All quantities in Equation (3), i.e., μ, $\sigma_{\Delta f}$, $\sigma_{\Delta \hat{f}}$, are estimated from the last year of the dataset (test sample).

[10]This is because it is possible to reduce intensity while increasing emissions when corporate revenues grow faster than emissions.

5 Conclusion

This paper investigates the determinants of corporate decarbonization using a comprehensive global dataset that combines financial indicators, balance-sheet metrics, ESG characteristics, macroeconomic variables, and firms' greenhouse gas emissions. We show that emission trajectories exhibit strong persistence, with past emissions remaining the dominant predictor of future carbon performance. Yet, firms with higher current footprints also display greater abatement potential, highlighting meaningful opportunities for targeted climate stewardship and capital allocation.

We identify corporate commitments to Science-Based Targets (SBTi) as a key factor associated with accelerated emission reductions. Our causal analysis confirms that SBT adoption results in significantly stronger abatement outcomes, supporting the view that voluntary climate commitments, when credible and externally validated, can meaningfully contribute to real-economy decarbonization. We further demonstrate that data-driven forecasting methods outperform simple heuristics in predicting short-term emission trajectories, generating signals that can materially enhance portfolio-level decarbonization and alignment with net-zero objectives, especially in the case of hard-to-abate sectors.

All in all, we are able to derive realistic decarbonization rates for portfolios consisting of global equities. We find that, given recent emission trajectories and reasonable forecasting skills, low carbon portfolio can achieve annual reduction rates of close to -5%. While this may seem like an underwhelming figure, when consistently compounded over 25 years, it would curtail the portfolio's footprint by 70%. Yet, this would fall short of full decarbonization.

Our findings highlight the complexity of corporate transition pathways. No single variable fully explains decarbonization dynamics; rather, progress reflects the interaction between firm-level financial strength, climate strategy, policy context, and sector-specific constraints. We also document discrepancies between modeled and reported emissions, underscoring persistent measurement challenges and the risks of relying solely on modeled data.

While the latter part of our analysis focuses on short-term forecasting to support capital allocation and stewardship activities, credible long-horizon assessment remains essential for strategic investment and policy decisions. However, such assessments are currently largely qualitative and subject to significant uncertainty. An important avenue for future research is therefore the development of systematic, data-driven approaches to evaluate the long-term credibility of corporate transition plans, particularly as investors increasingly rely on forward-looking metrics, such as implied temperature rise and net-zero alignment scores, to guide strategic decisions. Further research could also explore the integration of physical-risk and policy-risk scenarios, as well as the role of supply-chain dependencies in shaping firm-level emission trajectories and the feasibility of corporate transition pathways.

A Theoretical motivation - Proof of Equation (3)

The identity (3) relies on the following assumption:

(A) – Portfolio weights w_{t-1} and predicted footprint changes $\Delta\hat{f}_{t+1}$ are independent.

Heuristically, there is no reason why the prediction model should suggest footprint variations that are correlated with the asset manager's positions. The assumption therefore seems reasonable. Given the budget constraint $\mathbb{E}[w_t] = \mathbb{E}[w_{t-1}] = 1$, it implies

$$\mathbb{E}[\Delta w_t] = \alpha\mathbb{E}[c_w - \Delta\hat{f}_{t+1}] = \alpha\left(\mathbb{E}[w_{t-1}\Delta\hat{f}_{t+1}] - \mathbb{E}[\Delta\hat{f}_{t+1}]\right) \tag{9}$$

$$= \alpha\left(\mathbb{E}[w_{t-1}]\mathbb{E}[\Delta\hat{f}_{t+1}] - \mathbb{E}[\Delta\hat{f}_{t+1}]\right) = 0 \tag{10}$$

First note that

$$\Delta f^{(w)}_{t+1} = \frac{w_t f_{t+1} - w_{t-1}f_t}{w_{t-1}f_t} = \overbrace{\left(\frac{w_t - w_{t-1}}{w_{t-1}}\right)}^{\Delta w_t}\overbrace{\left(\frac{f_{t+1} - f_t}{f_t}\right)}^{\Delta f_{t+1}} + \overbrace{\left(\frac{w_t - w_{t-1}}{w_{t-1}}\right)}^{\Delta w_t} + \overbrace{\left(\frac{f_{t+1} - f_t}{f_t}\right)}^{\Delta f_{t+1}},$$

and hence,

$$\begin{aligned}
\mathbb{E}\left[\Delta f^{(w)}_{t+1}\right] &= \mathbb{E}[\Delta w_t \Delta f_{t+1}] + \mathbb{E}[\Delta w_t] + \overbrace{\mathbb{E}[\Delta f_{t+1}]}^{=\mu} \\
&= \mathbb{C}\mathrm{ov}(\Delta w_t, \Delta f_{t+1}) + \mu + \overbrace{\mathbb{E}[\Delta w_t]}^{=0 \text{ by Eq. (10)}} \times (1 + \mathbb{E}[\Delta f_{t+1}]) \\
&= \mu - \alpha\mathbb{C}\mathrm{ov}(\Delta\hat{f}_{t+1}, \Delta f_{t+1}), \\
&= \mu - \underbrace{\alpha\sigma_{\Delta f}\sigma_{\Delta\hat{f}}}_{\kappa}\,\mathbb{C}\mathrm{or}(\Delta\hat{f}_{t+1}, \Delta f_{t+1})
\end{aligned}$$

QED.

B Data

B.1 Variable definitions

Parameter	Source	Description
I. GHG Emissions		
ISS_GHG_S1	**ISS ESG**	Scope 1 GHG emissions, expressed in tons equivalent CO2 (tCO2e)
ISS_GHG_S2	**ISS ESG**	Scope 2 GHG emissions, expressed in tons equivalent CO2 (tCO2e)
ISS_GHG_S3	**ISS ESG**	Scope 3 (upstream & downstream) GHG emissions, expressed in tons equivalent CO2 (tCO2e)
MSCI_GHG_S1	**MSCI**	Scope 1 GHG emissions, expressed in tons equivalent CO2 (tCO2e)
MSCI_GHG_S2	**MSCI**	Scope 2 GHG emissions, expressed in tons equivalent CO2 (tCO2e)
MSCI_GHG_S3	**MSCI**	Scope 3 (upstream & downstream) GHG emissions, expressed in tons equivalent CO2 (tCO2e)

Table 7: **Description of dependent variables.**

Parameter	Source	Description
I. Corporate Sustainability Indicators		
ISS_Environment_Score	**ISS ESG**	ISS Environmental Score
ISS_Social_Score	**ISS ESG**	ISS Social Score
ISS_Governance_Score	**ISS ESG**	ISS Governance Score
Biodiversity_GHG_Score	**Iceberg DataLab**	Corporate Biodiversity Footprint (CBF), expressed in $km2.MSA$
Dummy_SBTi_Approved	**MSCI**	Corporate voluntary emissions targets indicator
Years_to_Closest_Target	**MSCI**	Years to the closest emissions target
II. Corporate Fundamental & Market Indicators		
12M_Stock_Return	**Bloomberg**	Average annualised day-to-day return net of dividend
Revenue	**Bloomberg**	Sales revenue turnover, yearly fundamental
Price_to_Earning	**Bloomberg**	Price to earning ratio, calendar year average
Price_to_Book	**Bloomberg**	Price to book ratio, calendar year average
Debt_to_Equity	**Bloomberg**	Total debt to total equity ratio, yearly fundamental

Continued on next page

(Continued)

Parameter	Source	Description
Longterm_Debt	**Bloomberg**	Long term debt on the balance sheet, yearly fundamental
Profit_Margin	**Bloomberg**	Profit margin ratio, yearly fundamental
Interest_Coverage	**Bloomberg**	Interest coverage ratio, yearly fundamental
R&D_Expenditure	**Bloomberg**	Research and development expenditure, income statement yearly fundamental
Free_Cash_Flow	**Bloomberg**	Free cash flow, yearly fundamental
Longterm_Investment	**Bloomberg**	Long term investment, yearly fundamental
III. Corporate Ownership		
Internal_Share_Outflow	**Bloomberg**	P85ercentage of insider shares outflow at the end the reporting period
Institutional_Share	**Bloomberg**	Percentage of float shares held by institutions, calendar year average
Green_Fund_Holding	**Exane, Factset**	Percentage of shares held by ”*Active*” sustainable ESG funds as identified by Exane (2023).
Share_Repurchase	**Bloomberg**	Total value of shares repurchased, yearly fundamental
IV. Regional Indicators		
Policy_Stringency	**OECD**	OECD Cross sectional Environmental policy Stringency, No US information, considered 0
Coal_Energy_pct	**IEA**	Coal Energy Country Level Energy Consumption Ratio from IEA yearly
Oil_Energy_pct	**IEA**	Oil Energy Country Level Energy Consumption Ratio from IEA yearly
Renewable_Energy_pct	**IEA**	Renewables Energy Country Level Energy Consumption Ratio from IEA yearly

Table 8: **Description of explanatory variables.**

B.1.1 Descriptive statistics

	type	N	min	25%	median	mean	75%	max	std. dev.
GHG	estimated	83833	-0.63	7.63	8.97	9.05	10.36	19.14	2.16
GHG	reported	53708	-0.28	9.64	11.14	11.31	12.84	19.82	2.50
INT	estimated	83833	-5.07	2.89	3.87	4.00	4.89	15.29	1.59
INT	reported	53708	-7.26	3.28	4.51	4.58	5.79	12.03	2.02
ΔGHG	estimated	83833	-0.50	-0.08	0.02	0.08	0.14	2.00	0.31
ΔGHG	reported	53708	-0.50	-0.08	0.01	0.07	0.13	2.00	0.32
ΔINT	estimated	83833	-0.50	-0.10	-0.01	0.04	0.09	2.00	0.30
ΔINT	reported	53708	-0.50	-0.13	-0.02	0.03	0.10	2.00	0.32

Table 9: **Descriptive statistics of dependent variables for Scope 1+2** We report the baseline indicators for our four dependent variables. Emissions are initially in tons equivalent CO_2, and we then take the log. Variations in emissions and intensities ($\Delta \text{GHG}_{t,i}$ and $\Delta \text{INT}_{t,i}$) are trimmed below 50% and above 200%. Emissions are reported after the log. Intensities are computed as raw emissions divided by Revenue in USD.

.

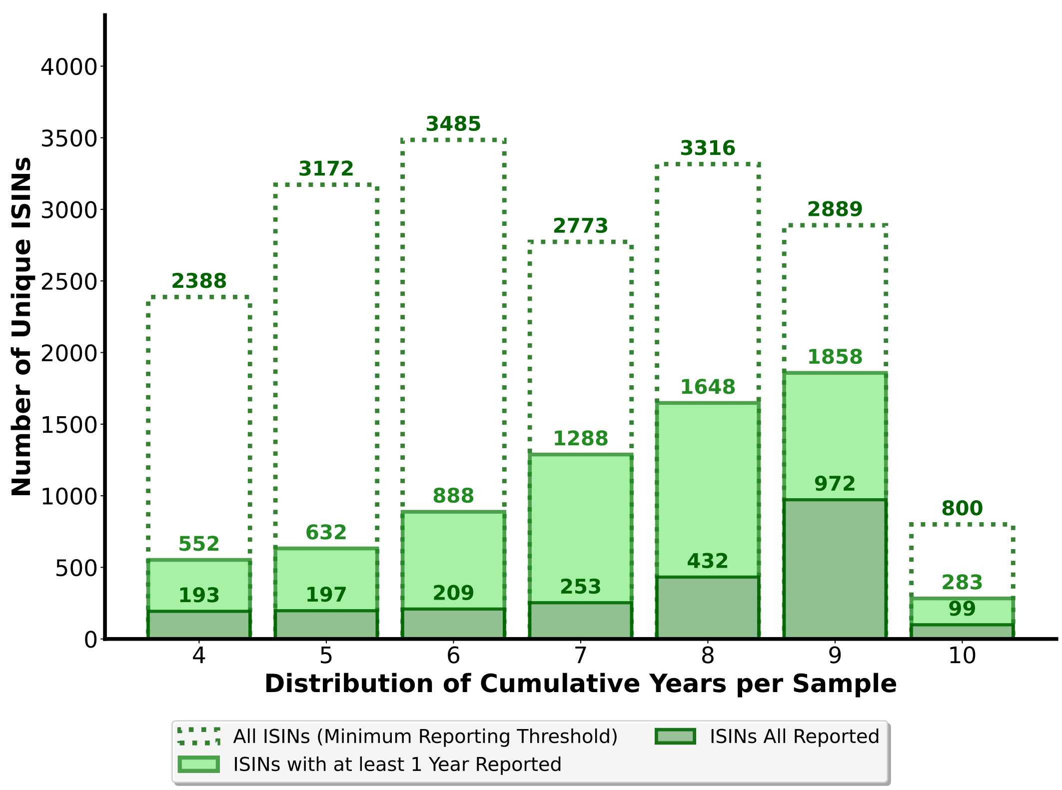


Figure 12: **Distribution of cumulative reported emissions over the period**

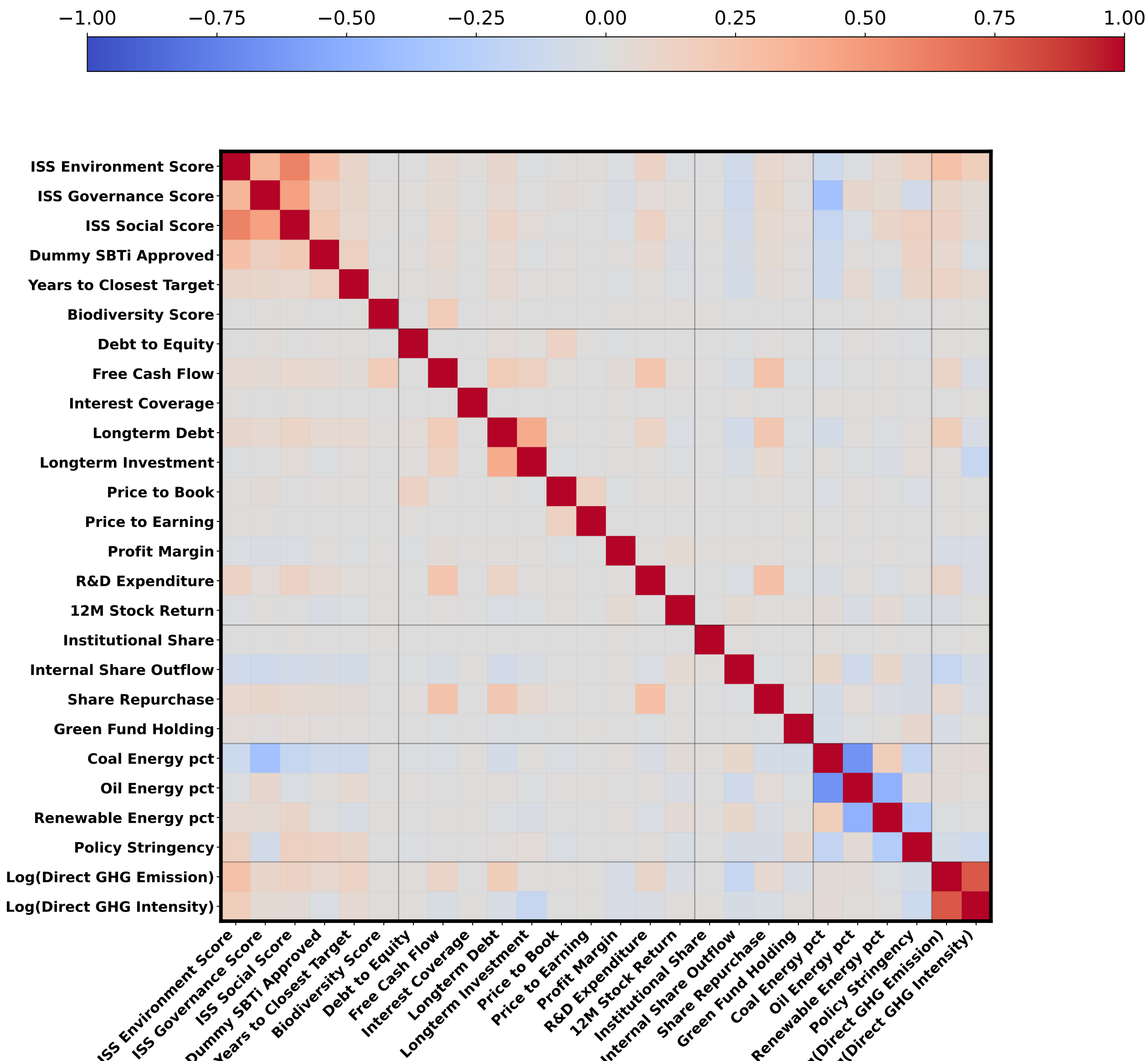


Figure 13: **Correlation matrix of all initial variables**.

Table 10: **Descriptive Statistics of Key Indicators.** We report summary descriptive statistics of independent variables. We also provide sample mean t-test significance results ($^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$) for the reported group versus the modeled group, and highlight in dark green if the value is larger for the reported sample, and in brown if the value is smaller.

Panel A: Reported

Parameter	count	mean	std	min	1%	50%	99%	max
I. Corporate Sustainability Indicators								
ISS Environment Score	53708	1.74^{***}	0.38	1.00	1.10	1.70	2.87	3.56
ISS Governance Score	53708	2.27^{***}	0.48	1.00	1.15	2.25	3.33	3.82
ISS Social Score	53708	1.81^{***}	0.33	1.02	1.16	1.76	2.82	3.32
Dummy SBTi Approved	53708	0.04^{***}	0.21	0.00	0.00	0.00	1.00	1.00
Years to Closest Target	53708	1.57^{***}	3.80	0.00	0.00	0.00	18.00	48.00
Biodiversity Score	53708	-441.43	2.44×10^{5}	-5.32×10^{6}	-2.62×10^{4}	-755.92	-0.04	3.95×10^{7}
II. Corporate Fundamental & Market Indicators								
Debt to Equity	53708	145.31^{***}	1.47×10^{3}	0.00	0.00	55.39	1.24×10^{3}	2.69×10^{5}
Free Cash Flow	53708	170.38^{***}	1.18×10^{3}	-1.90×10^{4}	-940.24	37.50	3.14×10^{3}	1.27×10^{5}
Interest Coverage	53708	1.19×10^{3}	1.54×10^{5}	-4.70×10^{5}	-99.00	11.18	3.87×10^{3}	3.55×10^{7}
Longterm Debt	53708	$3.60 \times 10^{3***}$	1.47×10^{4}	-0.07	0.00	403.05	5.79×10^{4}	3.62×10^{5}
Longterm Investment	53708	$4.83 \times 10^{3***}$	3.94×10^{4}	0.00	0.00	21.17	1.13×10^{5}	1.24×10^{6}
Price to Book	53708	5.48	64.19	0.04	0.26	1.68	39.04	5.84×10^{3}
Price to Earning	53708	260.76^{***}	2.37×10^{4}	0.10	2.70	17.62	433.86	3.42×10^{6}
Profit Margin	53708	9.29^{***}	61.42	-5.39×10^{3}	-59.51	6.80	101.24	1.99×10^{3}
R&D Expenditure	53708	44.71^{***}	364.38	0.00	0.00	0.00	841.80	1.99×10^{4}
12M Stock Return	53708	11.70^{**}	47.58	-91.93	-61.36	4.73	182.19	1.84×10^{3}
III. Corporate Ownership								
Institutional Share	53708	7.88×10^{5}	1.72×10^{8}	0.00	0.00	51.58	141.57	3.97×10^{10}
Internal Share Outflow	53708	4.85^{***}	11.15	0.00	0.00	0.47	57.97	110.59
Share Repurchase	53708	39.14^{***}	422.34	-859.45	0.00	0.00	831.50	3.14×10^{4}
Green Fund Holding	53708	0.14^{***}	0.86	0.00	0.00	0.00	2.08	62.91
IV. Regional Indicators								
Coal Energy pct	53708	0.08^{***}	0.08	0.00	5.01×10^{-3}	0.06	0.39	0.42
Oil Energy pct	53708	0.44^{***}	0.08	0.16	0.23	0.45	0.62	0.75
Renewable Energy pct	53708	0.08^{***}	0.07	0.00	0.02	0.06	0.29	0.69
Policy Stringency	53708	2.26^{***}	2.21	0.00	0.00	2.13	7.30	7.42

Continued on next page

(Continued)

Parameter	count	mean	std	min	1%	50%	99%	max
Panel B: Modeled								
I. Corporate Sustainability Indicators								
ISS Environment Score	83833	1.60***	0.26	1.00	1.07	1.63	2.38	3.48
ISS Governance Score	83833	2.15***	0.28	1.00	1.25	2.16	2.94	3.57
ISS Social Score	83833	1.71***	0.23	1.00	1.10	1.69	2.33	3.22
Dummy SBTi Approved	83833	1.30×10^{-4}***	0.01	0.00	0.00	0.00	0.00	1.00
Years to Closest Target	83833	0.10***	1.30	0.00	0.00	0.00	1.00	39.00
Biodiversity Score	83833	-1.72×10^{3}	1.05×10^{4}	-1.44×10^{5}	-1.17×10^{4}	-850.70	-3.24	4.90×10^{5}
II. Corporate Fundamental & Market Indicators								
Debt to Equity	83833	117.55 ***	1.02×10^{3}	0.00	0.00	36.48	1.02×10^{3}	9.18×10^{4}
Free Cash Flow	83833	16.75***	266.34	-1.92×10^{4}	-354.19	3.88	438.99	3.58×10^{4}
Interest Coverage	83833	788.32	3.67×10^{4}	-2.26×10^{6}	-518.60	7.60	9.64×10^{3}	6.49×10^{6}
Longterm Debt	83833	436.40 ***	1.67×10^{4}	0.00	0.00	12.26	3.90×10^{3}	2.34×10^{6}
Longterm Investment	83833	713.64 ***	2.99×10^{4}	0.00	0.00	3.12	7.29×10^{3}	2.96×10^{6}
Price to Book	83833	5.33	85.65	0.02	0.20	1.47	31.46	5.28×10^{3}
Price to Earning	83833	2.16×10^{3}***	7.69×10^{4}	0.02	1.91	19.65	1.13×10^{3}	1.02×10^{7}
Profit Margin	83833	-19.43 ***	2.00×10^{3}	-4.02×10^{5}	-196.72	4.52	87.65	1.16×10^{4}
R&D Expenditure	83833	6.92***	52.85	-2.67	0.00	0.00	98.10	5.92×10^{3}
12M Stock Return	83833	21.50**	961.59	-98.33	-68.21	1.94	237.22	1.05×10^{5}
III. Corporate Ownership								
Institutional Share	83833	5.67×10^{4}	1.10×10^{7}	0.00	0.00	16.22	132.33	2.60×10^{9}
Internal Share Outflow	83833	10.78***	16.13	0.00	0.00	2.95	68.66	127.02
Share Repurchase	83833	2.05***	47.31	-32.64	0.00	0.00	34.23	7.23×10^{3}
Green Fund Holding	83833	0.03***	0.53	0.00	0.00	0.00	0.62	63.14
IV. Regional Indicators								
Coal Energy pct	83833	0.11***	0.11	0.00	1.53×10^{-3}	0.07	0.41	0.42
Oil Energy pct	83833	0.42***	0.10	0.16	0.23	0.43	0.65	0.75
Renewable Energy pct	83833	0.07***	0.06	0.00	0.02	0.06	0.29	0.73
Policy Stringency	83833	1.93***	1.81	0.00	0.00	2.04	6.73	7.42

C Follow-up on the drivers of decarbonization

This section provides additional evidence on the determinants of emissions' dynamics.

C.1 Linear panel with time and individual fixed-effects

First, we gather the coefficients of the linear panel built with firm fixed-effects instead of industry fixed-effects in Table 11. Results are mostly consistent.

	GHG (log)		INT (log)		ΔGHG		ΔINT	
	Reported	Modeled	Reported	Modeled	Reported	Modeled	Reported	Modeled
Direct GHG Emission	51.895	71.883			−26.161	−35.699		
Direct GHG Intensity			47.419	58.816			−25.265	−33.826
ISS Environment Score	−0.238	0.487	−0.619	−0.414	−1.473	5.968	−2.034	4.699
ISS Governance Score	0.020	1.558	1.732	4.833	1.315	0.402	0.643	0.632
ISS Social Score	0.806	0.588	0.678	−2.637	−1.593	−0.085	−2.055	−1.559
Dummy SBTi Approved	−3.275	−0.672	−1.814	−0.650	−6.303	−4.751	−6.635	−2.819
Years to Closest Target	0.022	−0.120	−0.422	−1.368	1.190	−0.029	1.718	−0.817
Biodiversity Score	0.856	0.841	−0.095	−3.590	−0.316	0.896	−3.099	−5.513
Debt to Equity	0.016	−0.661	−1.279	0.203	1.049	−0.559	−1.076	−0.036
Free Cash Flow	−0.321	−0.446	−0.732	−0.842	−0.971	−0.260	1.363	−1.026
Interest Coverage	−4.916	−2.173	−2.608	−2.369	−4.499	0.373	−2.868	−0.759
Longterm Debt	0.686	0.358	0.121	0.006	−0.320	3.793	0.647	3.485
Longterm Investment	0.081	−0.299	0.258	−0.882	−1.560	1.861	−1.107	0.840
Price to Book	−0.346	0.283	−0.690	−0.598	0.854	0.091	−1.182	−1.579
Price to Earning	0.346	0.009	0.351	0.306	−1.115	0.265	−1.769	0.302
Profit Margin	1.118	−1.501	−0.380	3.878	0.881	2.404	−0.174	−1.757
R&D Expenditure	2.119	2.313	0.437	−0.564	−0.543	−0.517	0.846	0.068
12M Stock Return	−0.056	0.972	−6.275	−1.006	7.065	0.027	−5.105	−0.788
Institutional Share	7.722	0.176	4.699	−2.674	1.060	0.156	1.137	−0.546
Internal Share Outflow	1.312	−0.981	2.856	0.155	1.425	2.248	1.367	1.597
Share Repurchase	−1.499	−0.052	−0.750	−0.135	−0.234	−0.832	2.340	−0.887
Green Fund Holding	0.261	−1.080	0.705	−1.571	−1.134	−1.093	−0.999	−1.087
Coal Energy fraction	−10.171	−8.993	−9.091	−11.302	−3.264	0.875	−5.334	−2.572
Oil Energy fraction	−6.437	−6.327	−8.027	−9.889	−2.264	−1.318	−4.888	−4.409
Renewable Energy fraction	−7.402	−1.948	−8.556	−5.714	−4.039	0.511	−7.558	−5.107
Policy Stringency	−1.870	0.312	0.211	4.122	−4.830	−0.195	−3.402	4.867

Table 11: **Panel model - individual fixed effects**. We report the t-statistics for the panel models defined in Equation (4) - all independent variables are lagged. The regressions employ two-way fixed effects (TWFE) to account for unobserved heterogeneity. The overarching column names pertain to the dependent variables. The sub-column panels pertain to the type of emissions considered as dependent variable. Standard errors are computed following Beck and Katz (1995). Colors code when statistics are larger than 2.58 (**light green**, 1% confidence level) or 3.3 (**darker green**, 0.1% confidence level) in absolute value.

C.2 LASSO surviving rates without fixed effects

In order to determine which variables matter the most in explaining future footprint, we can leverage LASSO penalized regressions (Tibshirani (1996)), a common feature selection tool in statistics. We rewrite Equation (4) as

$$y_{t+1,i} = X_{t,i}b + e_{t+1,i},$$

where the matrix $X_{t,i}$ includes the fixed effects. The LASSO solves

$$b^* = \underset{b}{\text{argmin}}\, ||y - Xb||_2^2 + \lambda||b||_1,$$

where the shrinkage parameter, λ determines the stringency of the penalization. The larger it is, the more sparse the model becomes (more estimated coefficients are set to zero). By default, the {glmnet} package in the R language spans a large number of values of λ and generates a full matrix of estimates $\hat{B}_{k,j}$, where j is the index of λ_j. We are then interested in the proportion of times that a variable survives the penalization across all penalty values λ_j:

$$p_k = J^{-1} \sum_{j=1}^{J} 1_{\{\hat{B}_{k,j} \neq 0\}}. \tag{11}$$

We run the models separately on reported and estimated emissions. The corresponding values (in percents) are gathered in Table 12. We only list the top 15 variables for each category: raw footprint include GHG and INT, while changes are ΔGHG and ΔINT.

To be consistent with the initial specification in Equation (4), we include fixed effects in the model. However, they are rarely among the surviving variables and none even make it to the top 15. For the sake of completeness, we also ran models without fixed effects (i.e., omitting the dummy variables), and the results are qualitatively the same; the ranks are simply slightly altered (see Table 13 in the Appendix).

Although the results differ slightly from those obtained through standard panel regressions, they largely confirm the main findings presented in Table 2. Past emissions remain the strongest predictor of future values, reaffirming their central role in forecasting. ESG scores and SBT commitments also display strong predictive power, especially for future emission changes, consistent with the negative t-statistics observed in the fixed effects models. For voluntary commitments, the explanatory strength is driven not only by SBTi approval but also by the number of targets set, although target horizons continue to show no meaningful effect.

Financial variables, including 12 month stock returns, long-term investment, and price-to-earnings ratios, consistently rank among the top predictors for relative changes and for intensities in particular. In contrast, corporate ownership remains largely insignificant, with the exception of internal share outflows. Macroeconomic indicators, especially policy stringency and the national energy mix, also demonstrate robust predictive power. Coal and renewable energy shares are selected by the LASSO procedure in roughly 60 to 70 percent of reported data specifications, while policy stringency remains a stable and significant predictor for both absolute and intensity based reductions. Again, the persistent divergence between reported and modeled data emphasizes the need for continued refinement of provider-based emissions estimates.

		Raw footprint				Relative change			
		GHG (log)		INT (log)		ΔGHG		ΔINT	
rank		report.	estim.	report.	estim.	report.	estim.	report.	estim.
1	Direct GHG Emission	**99**	**99**			**99**	**99**		
2	Direct GHG Intensity			**99**	**99**			**99**	**99**
3	MSCI Target Count	36	6	0	1	74	50	80	47
4	12M Stock Return	0	6	62	24	78	0	71	36
5	Policy Stringency	50	37	52	40	73	61	70	38
6	Renewable Energy pct	69	77	67	39	66	76	75	63
7	Dummy SBTi Approved	46	25	45	37	70	50	70	45
8	BLG SALES	51	56	40	31	59	31	78	60
9	Coal Energy pct	71	41	66	44	63	85	63	59
10	ISS Environment Score	0	26	29	9	56	67	66	63
11	ISS Social Score	28	27	27	45	56	33	62	44
12	Internal Share Outflow	35	45	52	20	52	44	55	48
13	Price to Earning	14	0	18	12	48	6	51	23
14	Oil Energy pct	44	84	47	63	45	85	52	55
15	Longterm Investment	6	26	25	32	47	39	48	23

Table 12: **Lasso survival rate (%)**. We report the percentage of times that a given variable survives LASSO selection (see Equation (11)). The rank of the variable is determined by the average of the four columns. Fixed effects are included in the model.

		Raw footprint				Relative change			
		GHG (log)		INT (log)		ΔGHG		ΔINT	
rank		report.	estim.	report.	estim.	report.	estim.	report.	estim.
1	Direct GHG Emission	**99**	**99**			**97**	**99**		
2	Direct GHG Intensity			**99**	**98**			87	**99**
3	MSCI Target Count	0	0	21	0	**98**	62	**99**	69
4	Coal Energy pct	43	38	43	29	**96**	87	**91**	**93**
5	12M Stock Return	14	0	36	0	**94**	32	87	49
6	Policy Stringency	39	16	43	20	**91**	84	**90**	78
7	Dummy SBTi Approved	14	1	24	4	85	48	89	58
8	Longterm Debt	0	17	24	24	78	29	81	44
9	ISS Governance Score	14	10	28	0	60	44	89	80
10	Internal Share Outflow	10	0	0	0	85	87	60	61
11	ISS Environment Score	17	4	26	32	88	81	50	86
12	Longterm Investment	15	0	26	0	76	22	57	54
13	Oil Energy pct	12	32	25	27	76	**96**	53	86
14	Free Cash Flow	6	0	10	0	58	61	59	69
15	Profit Margin	17	22	0	26	54	49	54	61

Table 13: **Lasso survival rate without TWFE**. We report the percentage of times that a given variable survives LASSO selection. The rank of the variable is determined by lag variables first, then by the average of ΔGHG reported and ΔINT reported columns. No fixed effects are included.

C.3 The characteristics of decarbonizing firms

To further challenge and validate our previous findings on the main drivers of GHG emissions, we examine the characteristics of firms sorted by their realized GHG emission trajectories.

We follow Seyfi (2025) and first split the sample between firms with low versus high dependent variable—for instance large GHG emissions reduction versus low GHG emissions reduction—on a year-by-year and sector-by-sector basis. Then, we pool groups by year, and for each potential characteristic, we test if the average value is statistically different between the two groups.

Figure 14 displays our results for levels of absolute GHG emissions while Figure 15 extends the analysis to emissions intensities. We plot the absolute t-statistics of the tests for which the null hypothesis is that there is no difference between the two groups.

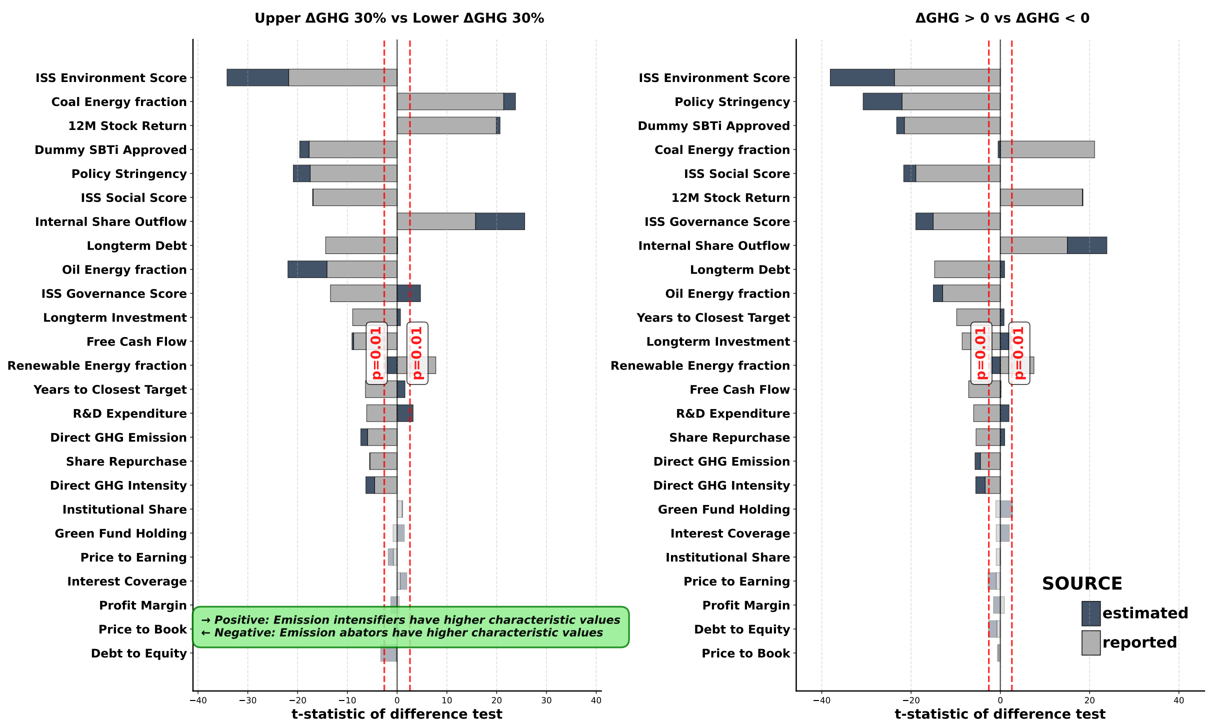


Figure 14: **Characteristics of decarbonizing firms**, pooled sample mean. This figure displays the absolute values of the t-statistics of the tests in differences in means between characteristics, **pooled by annual performance**, with high versus low ΔGHG.

Results in Figure 14 show that most determinants previously identified remain influential, although their relative importance changes. Variables such as coal use, stock returns, SBT commitments, policy stringency, internal share outflow, long-term debt, free cash flow, and R&D expenditure continue to exert effects consistent with earlier findings. However, past emissions are no longer the dominant driver, even if they remain statistically significant. In addition, stronger environmental and governance scores now display a clearer and more systematic association with firms that reduce their GHG emissions. Years to tar-

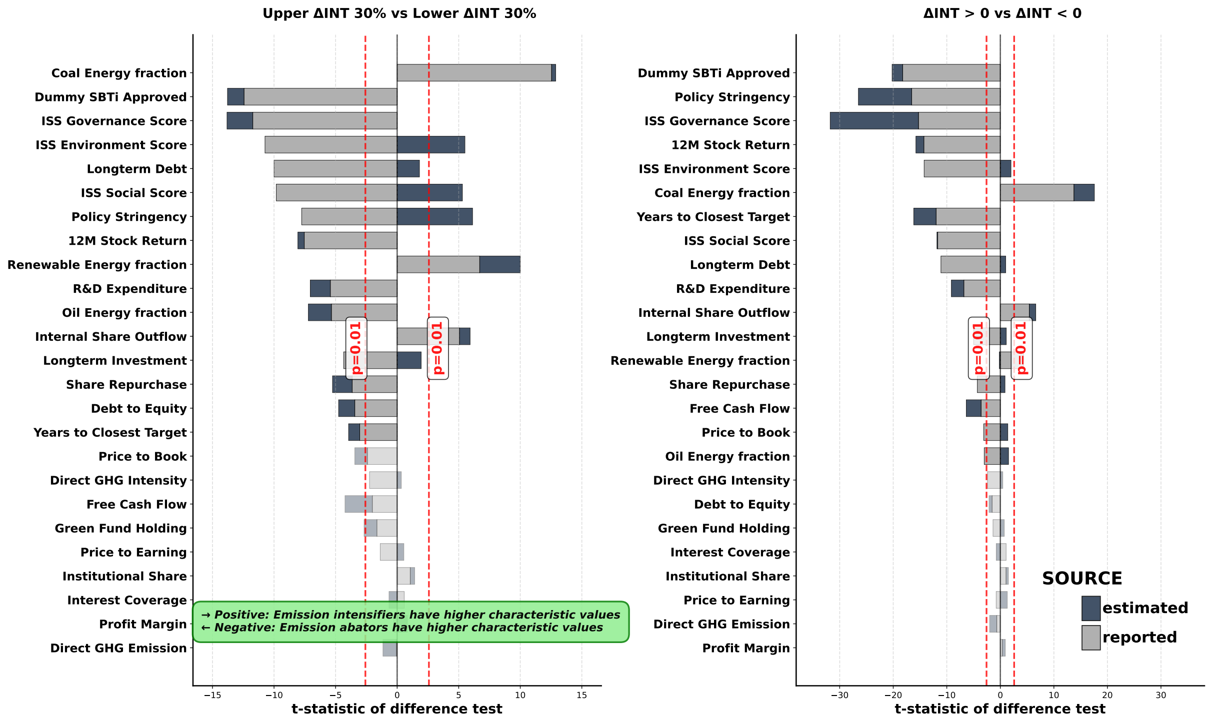


Figure 15: **Characteristics of (intensity) decarbonizing firms**, pooled sample mean. We plot the absolute values of the t-statistics of the tests in differences in means between characteristics, **pooled by annual performance**, with high versus low ΔINT.

get, which was previously insignificant, now emerges as a differentiating factor, implying that the temporal structure of climate commitments becomes more relevant once firms are grouped by decarbonization outcomes. Conversely, oil use in the energy mix shows an unexpected negative association with emissions, although its magnitude remains much smaller than the positive effect of coal. Finally, while discrepancies between reported and modeled data persist, they are less pronounced, as coefficient signs are generally consistent.

When focusing on emission intensities rather than absolute emissions (Figure 15), results remain coherent, but the gap between reported and estimated data widens substantially for the top 30 percent versus the bottom 30 percent of firms. The divergence is especially pronounced for ESG scores, financial ratios, and policy stringency, once again raising questions about the assumptions embedded in emissions estimation models.

D IV-DiD analysis

Panel A: Pre-treatment Outcomes Test

Variable Category	**Variables Tested**	**Passed Tests**	**Pass Rate**
Lagged Emissions	2	2	100%
Policy Stringency	1	1	100%
Financial Performance	6	6	100%
Capital Structure	4	4	100%
Market Valuation	2	2	100%
Ownership & Payouts	4	3	75%
Overall	**19**	**18**	**94%**

Panel B: Triple Difference Test

Outcome Variable	**Triple Interaction**	**P-value**	**Result**
Direct GHG Emissions (%)	0.0017	0.586	Pass
GHG Intensity (%)	0.0058	0.062	Pass

Panel C: STS Pre-treatment Placebo Tests ($\ell < 0$)

Outcome Variable	**Estimate**	**P-value**	**Result**
Direct GHG Emissions (%)	-0.002	0.855	Pass
GHG Intensity (%)	0.002	0.797	Pass

Table 14: **Validity Tests for IV-DID Identification.** Panel A reports results from regressing 18 pre-treatment firm characteristics on the instrumental variable; a test "passes" if the coefficient is statistically insignificant ($p > 0.05$). Panel B reports triple-difference results examining whether the instrument affects eventual SBTi adopters and non-adopters differentially. Panel C presents placebo tests using pre-treatment periods ($\ell < 0$) from the stacked two-stage estimator to verify parallel trends.

	First Stage		LATE Estimates	
Cohort	**π (CDP $\rightarrow$ SBTi)**	**F-stat**	**Direct GHG (%)**	**Intensity (%)**
2018 (N_{firm}=3,921)	0.042***	87.3	-0.039	-1.309***
	(0.004)		(0.191)	(0.222)
2019 (N_{firm}=420)	-0.015**	**6.2**	-1.524	-2.737
	(0.006)		(1.860)	(2.057)
2021 (N_{firm}=2)	-	-	-	-
2023 (N_{firm}=12)	-	-	-	-

Table 15: **Heterogeneity Analysis by Adoption Cohort using C&S Estimator.** This table reports cohort-specific IV-DID estimates using the Callaway & Sant'Anna (2021) method aggregated by group. Standard errors, clustered at the firm level, are in parentheses. The LATE is calculated as the ratio of the group-specific reduced-form estimate to the group-specific first-stage estimate. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Only the 2018 cohort exhibits both strong instrument properties (F>10) and positive first-stage effects, indicating that CDP reporting increased SBTi participation. The 2019 cohort shows weak instruments and perverse first-stage effects. The 2021, 2023 cohorts are not estimated due to low observations.

E Forecasting - Model hyper parameters

short name	**description**	**tested values**
ntree	number of trees	150, 500, 1500
mtry	% of columns used to train each tree	0.6, 0.9
sampsize	% of original sample used for each tree	0.6, 0.9
nodesize	max. size of leaves as % of training sample	0.0005, 0.001, 0.005
maxnodes	max. number of leaves per tree	1000, 2000, 5000

Table 16: **Random forest hyper-parameters**. We list the hyper-parameters used in the baseline cases, along with the tested values.

References

Adamolekun, G. (2024). Carbon price and firm greenhouse gas emissions. *Journal of Environmental Management 349*, 119496.

Aswani, J., A. Raghunandan, and S. Rajgopal (2024). Are carbon emissions associated with stock returns? *Review of Finance 28*(1), 75–106.

Atta-Darkua, V., S. Glossner, P. Krueger, and P. Matos (2025). Decarbonizing institutional investor portfolios: Helping to green the planet or just greening your portfolio? *SSRN Working Paper 4212568*.

Barahhou, I., M. Ben Slimane, T. Roncalli, and N. Oulid Azouz (2022). Net zero investment portfolios-part 1. The comprehensive integrated approach. *SSRN Working Paper 4283998*.

Beck, N. and J. N. Katz (1995). What to do (and not to do) with time-series cross-section data. *American Political Science Review 89*(3), 634–647.

Bendig, D., A. Wagner, and K. Lau (2023). Does it pay to be science-based green? the impact of science-based emission-reduction targets on corporate financial performance. *Journal of Industrial Ecology 27*(1), 125–140.

Berg, F., J. F. Koelbel, and R. Rigobon (2022). Aggregate confusion: The divergence of esg ratings. *Review of Finance 26*(6), 1315–1344.

Berg, F., J. Oliver Huidobro, and R. Rigobon (2024). On the importance of assurance in carbon accounting. *SSRN Working Paper 4734240*.

Berk, J. and J. H. Van Binsbergen (2024). The impact of impact investing. *SSRN Working Paper 3909166*.

Bjørn, A., J. P. Tilsted, A. Addas, and S. M. Lloyd (2022). Can science-based targets make the private sector paris-aligned? a review of the emerging evidence. *Current climate change reports 8*(2), 53–69.

Bolton, P. and M. Kacperczyk (2021). Do investors care about carbon risk? *Journal of Financial Economics 142*(2), 517–549.

Bolton, P. and M. Kacperczyk (2023). Firm commitments. Technical report, National Bureau of Economic Research.

Bolton, P. and M. Kacperczyk (2024). Are carbon emissions associated with stock returns? comment. *Review of Finance 28*(1), 107–109.

Bolton, P., M. Kacperczyk, and F. Samama (2022). Net-zero carbon portfolio alignment. *Financial Analysts Journal 78*(2), 19–33.

Busch, T., M. Johnson, and T. Pioch (2022). Corporate carbon performance data: Quo vadis? *Journal of Industrial Ecology 26*(1), 350–363.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. *Journal of econometrics 225*(2), 200–230.

Cenedese, G., S. Han, and M. T. Kacperczyk (2023). Carbon-transition risk and net-zero portfolios. *SSRN Working Paper 4565220*.

Chen, A. Y. and J. McCoy (2024). Missing values and the dimensionality of expected returns. *Journal of Financial Economics Forthcoming*.

Chen, C. Y.-H., A. Lioui, and O. Scaillet (2025). Green silence: Double machine learning carbon emissions under sample selection bias. *SSRN Working Paper 5512980*.

Coqueret, G., T. Giroux, and O. D. Zerbib (2025). The biodiversity premium. *Ecological Economics 228*, 108435.

Coskuner, C., M. K. Paskeh, G. Olasehinde-Williams, and S. S. Akadiri (2020). Economic and social determinants of carbon emissions: evidence from organization of petroleum exporting countries. *Journal of Public Affairs 20*(3), e2092.

De Chaisemartin, C. and X. D'HaultfŒuille (2018, April). Fuzzy Differences-in-Differences. *Review of Economic Studies 85*(2), 999–1028.

Dimson, E., P. Marsh, and M. Staunton (2020). Divergent esg ratings.

Dogan, E. and F. Seker (2016a). Determinants of co2 emissions in the European Union: the role of renewable and non-renewable energy. *Renewable energy 94*, 429–439.

Dogan, E. and F. Seker (2016b). An investigation on the determinants of carbon emissions for OECD countries: empirical evidence from panel models robust to heterogeneity and cross-sectional dependence. *Environmental Science and Pollution Research 23*, 14646–14655.

Feldhütter, P. and L. H. Pedersen (2023). Is capital structure irrelevant with ESG investors? *SSRN Working Paper*.

Flammer, C. (2021). Corporate green bonds. *Journal of financial economics 142*(2), 499–516.

Fraser, A. and T. Fiedler (2023). Net-zero targets for investment portfolios: An analysis of financed emissions metrics. *Energy Economics 126*, 106917.

Freiberg, D., J. Grewal, and G. Serafeim (2021). Science-based carbon emissions targets. *SSRN Working Paper 3804530*.

Goldhammer, B., C. Busse, and T. Busch (2017). Estimating corporate carbon footprints with externally available data. *Journal of Industrial Ecology 21*(5), 1165–1179.

Grinsztajn, L., E. Oyallon, and G. Varoquaux (2022). Why do tree-based models still outperform deep learning on typical tabular data? *Advances in Neural Information Processing Systems 35*, 507–520.

Gu, S., B. Kelly, and D. Xiu (2020). Empirical asset pricing via machine learning. *Review of Financial Studies 33*(5), 2223–2273.

Habiba, U., C. Xinbang, and A. Anwar (2022). Do green technology innovations, financial development, and renewable energy use help to curb carbon emissions? *Renewable Energy 193*, 1082–1093.

Heath, D., D. Macciocchi, R. Michaely, and M. C. Ringgenberg (2023). Does socially responsible investing change firm behavior? *Review of Finance 27*(6), 2057–2083.

Ishwaran, H. and M. Lu (2019). Standard errors and confidence intervals for variable importance in random forest regression, classification, and survival. *Statistics in mMdicine 38*(4), 558–582.

Januschowski, T., Y. Wang, K. Torkkola, T. Erkkilä, H. Hasson, and J. Gasthaus (2022). Forecasting with trees. *International Journal of Forecasting 38*(4), 1473–1481.

Jiang, X. and D. Guan (2016). Determinants of global CO2 emissions growth. *Applied energy 184*, 1132–1141.

Kahn, M. E., J. G. Matsusaka, and C. Shu (2023). Divestment and engagement: The effect of green investors on corporate carbon emissions. *SSRN Working Paper 4609861*.

Kalesnik, V., M. Wilkens, and J. Zink (2022). Do corporate carbon emissions data enable investors to mitigate climate change? *Journal of Portfolio Management 48*(10), 119–147.

Krueger, P., Z. Sautner, and L. T. Starks (2020). The importance of climate risks for institutional investors. *Review of Financial Studies 33*(3), 1067–1111.

Lam, P. and J. Wurgler (2024). Green bonds: New label, same projects. Technical report, National Bureau of Economic Research.

Le Guenedal, T., F. Lombard, T. Roncalli, and T. Sekine (2022). Net zero carbon metrics. *SSRN Working Paper 4033686*.

Le Guenedal, T. and T. Roncalli (2022). Portfolio construction with climate risk measures. *SSRN Working Paper 3999971*.

Lee, K.-H. and B. Min (2015). Green R&D for eco-innovation and its impact on carbon emissions and firm performance. *Journal of Cleaner Production 108*, 534–542.

Leffel, B., T. P. Lyon, and J. P. Newell (2024). Filling the climate governance gap: Do corporate decarbonization initiatives matter as much as state and local government policy? *Energy Research & Social Science 109*, 103376.

Martiny, A., F. Testa, J. Taglialatela, and F. Iraldo (2024). Determinants of Environmental, Social and Governance (ESG) performance: A systematic literature review. *Journal of Cleaner Production Forthcoming*, 142213.

Mentch, L. and G. Hooker (2016). Quantifying uncertainty in random forests via confidence intervals and hypothesis tests. *Journal of Machine Learning Research 17*(26), 1–41.

Michalski, L. and R. K. Y. Low (2024). Determinants of corporate credit ratings: Does esg matter? *International Review of Financial Analysis Forthcoming*, 103228.

Miyaji, S. (2024). Instrumented Difference-in-Differences with Heterogeneous Treatment Effects. Version Number: 5.

Molnar, C. (2020). *Interpretable machine learning*. Lulu.

Newey, W. and K. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*(3), 703–708.

Nguyen, D. K., T. L. D. Huynh, and M. A. Nasir (2021). Carbon emissions determinants and forecasting: Evidence from g6 countries. *Journal of Environmental Management 285*, 111988.

Nguyen, Q., I. Diaz-Rainey, and D. Kuruppuarachchi (2021). Predicting corporate carbon footprints for climate finance risk analyses: a machine learning approach. *Energy Economics 95*, 105129.

Papadopoulos, G. (2022). Discrepancies in corporate ghg emissions data and their impact on firm performance assessment.

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

Pastor, L., R. F. Stambaugh, and L. A. Taylor (2024). Carbon burden. *SSRN Working Paper 4998860*.

Pedersen, L. H. (2024). Carbon pricing versus green finance. *SSRN Working Paper 4382360*.

Puertas, R. and L. Marti (2021). Eco-innovation and determinants of GHG emissions in OECD countries. *Journal of Cleaner Production 319*, 128739.

Ramadorai, T. and F. Zeni (2024). Climate regulation and emissions abatement: Theory and evidence from firms' disclosures. *Management Science*.

Rink, S., M. Dumrose, and Y. Matheis (2024). Consistency or transformation? finance in climate agreements. *SSRN Working Paper 4835984*.

Roncalli, T. (2024). Portfolio alignment and net zero investing. *SSRN Working Paper 4891271*.

Seo, H. (2021). Peer effects in corporate disclosure decisions. *Journal of Accounting and Economics 71*(1), 101364.

Seyfi, S. (2025). Essence of the cross section. *SSRN Working Paper 4466972*.

Sharma, S. S. (2011). Determinants of carbon dioxide emissions: empirical evidence from 69 countries. *Applied Energy 88*(1), 376–382.

Shwartz-Ziv, R. and A. Armon (2022). Tabular data: Deep learning is not all you need. *Information Fusion 81*, 84–90.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology 58*(1), 267–288.

Tilsted, J. P., E. Palm, A. Bjørn, and J. F. Lund (2023). Corporate climate futures in the making: Why we need research on the politics of science-based targets. *Energy Research & Social Science 103*, 103229.

Zerbib, O. D. (2019). The effect of pro-environmental preferences on bond prices: Evidence from green bonds. *Journal of banking & finance 98*, 39–60.