Carbon awareness and return co-movement *

Lucas S. Li Xingjian Zheng †

November 10, 2023

Abstract

This paper documents a significant emission return co-movement as investors' carbon awareness increases. We show that firms that produce carbon emissions on a similar scale have a higher return correlation. This pattern became prominent after 2012 and has been steadily more significant ever since, whereas it was barely significant before 2012. We also show that this co-movement is driven by investor flows as investors purchase green stocks and divest brown stocks. We also adopt a state-level emission target shock to address the endogeneity issue. Overall, this paper examines whether investors care about carbon risk and the pricing of carbon risk from a different perspective.

JEL classification: G12, G14, G40.

Keywords: Carbon emission; Pairs trading; Investor flow

*Very preliminary draft and subject to change. Please do not circulate.

†Lucas S. Li (lucaslisj25@gmail.com) is from Shanghai American School, PR China. Xingjian Zheng (xjzheng.20@saif.sjtu.edu.cn) is from the Shanghai Advanced Institute of Finance, Shanghai Jiao Tong University, PR China. Click *here* for the latest manuscript.

1. Introduction

Climate change poses a significant and imminent risk to our socio-economic system (Giglio et al., 2021; Hong et al., 2020) (Hong et al., 2020). Previous literature has argued that institutional investors care about carbon risk and are determined to promote sustainable investments Bolton and Kacperczyk (2021c); Cohen et al. (2023); Kim and Yoon (2023). On the contrary, anecdotal evidence from the industry shows that many institutional investors, led by BlackRock and Vanguard, are still hesitating about the trade-offs between financial and ESG performance. As a consequence, it is unclear whether these green initiatives are just talks or do they lead to real actions and have real consequences for the resource allocations of society.

This paper attempts to examine this question with a novel approach. Instead of directly regressing stock returns on carbon emissions as in Bolton and Kacperczyk (2021a) or creating carbon portfolios like Cheema-Fox et al. (2021), we examine firms' return correlation following a pairs' trading approach. More specifically, we examine the relationship between firms' return correlation and their carbon emissions.

Our sample is on the firm-pair level. For each firm pair i and j, we calculate their return correlation coefficient over the past 52 weeks as the dependent variable and compute their emission similarity as the independent variable. The emission similarity is computed firstly by sorting firm pairs' emission differences, as measured by emission distance $GHG_i - GHG_j$, from high to low into deciles, then converting this continuous difference variable into a decile variable that denotes emission similarity. A higher emission similarity denotes the two firms produce carbon emissions in a more similar amount. We conjecture that, as investors' carbon awareness increases, the emission-return comovement would be more significant.

This approach is unique in the sense that it does not impose restrictions on the sign of carbon premium. The debate on whether carbon risk is positively or negatively priced in the stock return has been heated in the past several years. An influential research paper by Bolton and Kacperczyk (2021a) finds that stocks of firms with higher carbon emissions earn higher returns on average. However, others argue the emission-return relationship should be negative (Garvey et al., 2018; In et al., 2017; Matsumura et al., 2014; Zhang, 2023), and some even reach an inconclusive result (Aswani et al., 2023a; Monasterolo and De Angelis, 2020). As (Pástor et al., 2021) shows in a static model, green assets can outperform brown assets when positive shock changes investors' stochastic discount factor, but in equilibrium, green assets have lower expected returns than brown assets. Despite the insightful theory, it is difficult to empirically identify a market-wide ESG shock, since investors' awareness towards ESG has been gradually rising. Moreover, the road to ESG investment is not smooth, it has encountered many reversals, particularly highlighted by former President Trump's withdrawal from the Paris Agreement. Any positive shock towards ESG can be arguably counteracted with a negative shock, potentially leaving ambiguous results. As a result, our approach based on pairs trading is valuable in the sense that it does not impose any restriction on investors' preferences and only looks at their trading behaviour. If the emission-return co-movement has been more prominent in the past several years, then we could claim a convincing carbon awareness pattern.

Our empirical results show that investors' carbon awareness, as measured by the coefficients in front of the emission similarity, has been steadily increasing after 2012, and it was barely significant from 2004 to 2012. Besides, despite all the reversals and negative shocks to the ESG factor, e.g., Trump's unexpected decision to withdraw from the Paris Agreement, carbon awareness seems to be rather unaffected. This pattern is sustained with all scopes of data and with both emission levels and emission intensity. Furthermore, double sorting results based on other firm fundamentals like size, profitability, and financial constraint levels do not affect the result as well. Our analysis is also interesting because it produces the same results with whatever emission variable we choose. Prior literature Bolton and Kacperczyk (2021a), Aswani et al. (2023a) and Zhang (2023) find contrasting results with different emission variables, as emission level (logarithmic value of carbon emissions) yields totally different regression results as compared to emission intensity (scaled by firms' sales). Our research overcomes the inconsistency issue faced by previous research.

To empirically solve the identification problem, we use the roll-out of state-level emission reduction plans as shocks to examine whether increased emission similarity increases stock return co-movement. The underlying assumption is that when a state has announced emission reduction targets, firms would reduce their carbon emission mandatorily. As a consequence, we examine whether the return correlation increases after a high-emission firm within a firm pair has experienced an emission reduction shock whereas the low-emission firm does not. This exogenous variation enables us to perform a staggered DID regression. Additionally, this research design prohibits the potential concern of fundamental co-movement, which we will elaborate on in the empirical section.

We also examine the underlying mechanism that drives this return co-movement. With a 2SLS style approach, we show that investors simultaneously buy green stocks and sell brown stocks, and their investor flows jointly result in increased return correlations. Moreover, additional analyses show that their behaviour can also be explained by expecting firms' future growth and profitability to change. This just may imply that practising sustainable investment may help improve firms' financial performance.

Our paper contributes to the current research on carbon pricing in several ways. Firstly, we provide a novel method to examine how investors' carbon awareness affects asset prices. This method is unique since it does not impose any strict restrictions on investors' SDFs. Previous research that directly regresses stock returns on firms' carbon emissions implicitly assumes a constant SDF that is related to the ESG factor (Bolton and Kacperczyk, 2021a; Pástor et al., 2021). However, as investors have taken responsible investment more seriously, this assumption might not be very strong and cannot be easily tested. However, with this approach that simply regresses return correlation on emission similarity, we could easily test for any structural mutation point that indicates investors' attention towards sustainable investments. Moreover, this approach is quite robust and reliable for all emission variables, whereas previous research documents different results with different emission variables (Aswani et al., 2023a; Zhang, 2023). Secondly, we propose a clean identification strategy to examine the dynamics of carbon awareness. We show that states' emission reduction targets do have a real impact on firms' behaviour, and it is reflected in the stock market as well. This kind of regulation pressure from the government is rarely examined in the previous literature (Cohen et al., 2023). Finally, we complement the literature by showing that carbon awareness may only be important after 2012, which coincides with Acharya et al. (2022) and Skiadopoulos et al. (2023).

The remainder of this paper is as follows. In section 2 we review the related literature. In section 3 we describe our research design and data. In section 4 we report main empirical results, identification methodologies, and mechanism tests. In section 5, we provide further analyses. Section 6 concludes.

2. Related literature

Pástor et al. (2021, 2022) and Pedersen et al. (2021) are among the first to propose an ESG-CAPM framework that studies equilibrium asset pricing in the topic of climate finance. Their model shows that in equilibrium, green asses have low expected returns because investors prefer ESG assets to non-ESG assets as they either gain utility or hedge undesirable climate risk by holding these assets. In a similar vein, investors display a strong aversion to the so-called "sin" stocks (Blitz and Fabozzi, 2017; Hong and Kacperczyk, 2009; Luo and Balvers, 2017), investors dislike assets that have a bad ESG record. For example, if a company emits considerable carbon dioxide and sulfur dioxide, then the temperature and atmosphere around its peripheral are deteriorated, which induces obvious health risks to people's well-being. Moreover, these firms are more vulnerable to environmental penalties and litigations if they fail to comply with the environmental emission standard set by the authority (Hsu et al., 2022), which also makes the firm riskier to invest. The ESG preference and the large exposure to climate risk essentially determine the efficient risk-return frontier and are the reasons why brown assets have higher climate betas as compared to green assets (Bolton and Kacperczyk, 2021a; Choi et al., 2020; Engle et al., 2020). In Choi et al. (2020), they show that investors' preference for green stocks is affected by increased attention to abnormal temperatures. They provide a mechanism in which people pay more attention to climate change and ESG issues when experiencing abnormal weather, and this experience affects their behaviour in the financial markets through home bias. Investors' aversion to stocks of firms that have unfavourable ESG records is widely documented in the literature. The "sin" stocks, stocks of firms that produce products against social norms such as alcohol, tobacco, and gaming, are boycotted by stock market investors (Hong and Kacperczyk, 2009) and thus outperform non-sin stocks. This sin-premium is also observed in the VC market, where venture capital and private equity investors would sacrifice investment returns for social impact, and their preference is stronger for investors who join the United Nations Principles of Responsible Investment (UNPRI) signatories (Barber et al., 2021). These large institutional investors drive nonpecuniary utility from investing in funds that both consider financial returns and social impacts. In the fixed-income market, investors also display a strong preference for ESG bonds, where green bonds tend to be priced at a premium, offering lower yields than brown bonds (Baker et al., 2018; Duan et al., 2021; Zerbib, 2019). In other related studies, El Ghoul et al. (2011) and Chava (2014) documents that firms with better ESG records face lower financing costs. Furthermore, improvement in responsible employee relations, environmental policies, and product strategies contributes substantially to reducing firms' cost of equity.

However, there exist opposing findings that firms with better ESG scores outperform the opposites. For example, in a well-recognized paper by Edmans (2011), the author showed that firms that were awarded the prize of "100 Best Companies to Work for in America" earned an annual four-factor alpha of 3.5% from 1984 to 2009, and the results are robust after controlling for other factors. This result is consistent with the others (Cremers and Ferrell, 2014; Gompers et al., 2003), suggesting that firms have better governance by having higher employee satisfaction, stronger shareholder rights, or ESG ratings are more favoured by investors.

This contradiction of the ESG premium documented in previous research is heated within a more specific ESG topic, i.e., carbon emissions, led by a series of works by Bolton et al. (2022a); Bolton and Kacperczyk (2021a,b); Bolton et al. (2022b); Bolton and Kacperczyk (2020a,b, 2021c); Bolton et al. (2022c), who mainly documents a positive link between stock returns and emissions, and other firm outcomes, which is consistent with the risk compensation hypothesis. The idea behind this hypothesis is quite straightforward, as firms with higher greenhouse gas emissions are more vulnerable to state environmental penalties or other environmentally related risks. As a result, investors require a higher rate of return for extra risk compensation. Their influential research is well-sustained by Oestreich and Tsiakas (2015) and Ilhan et al. (2021).

However, some other researchers believe this phenomenon is entirely driven by vendor-estimated emissions, which makes the estimation results quite unreliable Aswani et al. (2023a,b), even though the carbon emission provided by different data vendors are quite persistent Busch et al. (2022). On the other hand, Monasterolo and De Angelis (2020) examined the carbon premia by constructing a carbon portfolio and found that investors have started to consider low-carbon assets as an appealing investment opportunity more recently. Matsumura et al. (2014), Garvey et al. (2018), and In et al. (2017) document similar results. Duan et al. (2021) examines the pricing of a firm's carbon risk in the corporate bond market and finds that bonds of more carbon-intensive firms earn significantly lower returns than their industry peers. Their empirical results are more robust because multiple existing bonds exist for a single firm, making time-series estimation available. In Cheema-Fox et al. (2021), researchers construct a decarbonization factor that goes long low-carbon intensity firms and shorts high-carbon intensity firms. This decarbonization factor yields significantly positive returns, especially in Europe. In a recent analysis with global evidence, Choi et al. (2022), researchers find that high-emission firms tend to have lower price valuation ratios than low-emission firms, and the devaluation of high-emission firms' phenomena is most prominent in recent years. Their empirical analyses mainly focus on the valuation ratios such as PE, PS, and PB, instead of stock returns.

3. Data and methodology

We examine investors' carbon awareness following the pairs trading strategy in Chen et al. (2019), which involves both return co-movement and emission distance within a firm pair. The main story is that the returns of brown stocks co-move as investors jointly divest from them, and the returns of green stocks co-move as investors tilt their portfolios towards them.

3.1. Return comovement

We first calculate the return co-movement between each firm pair. It relies on pairwise stock return correlations for all stocks listed in the US equity market (that are traded on NYSE, AMEX, and NASDAQ) in the CRSP database. For stock j of any firm, we calculate its 52-week return correlation with any other stock $CORR_{<i,j>,t}$. We require stocks to have full weekly observations in each period and derive a large correlation matrix in each cross-section in our sample period from 2004 to 2019 in each month.

Since we are computing correlation matrices on each cross-section, which requires heavy computation power given the vast number of stocks, we use two methods to decrease the computation complexity. First, we perform our main analysis with the observations at the end of each year (on an annual basis) instead of on a monthly basis. In other words, we are using the cross-sectional correlation matrices only in December of each year instead of the whole year from January to December. This allows us to perform a pooled OLS regression by reducing computation to 1/12 of its original size. But we also use monthly observations as well in the empirical analyses to perform the structural mutation test. Second, we use the stocks that exist in the year 2004 as our sample group that also can be matched with the Trucost database (the 2004 stock sample), which has 797 firms. This also reduces the computation by roughly 1/7.

3.2. Emission distance and other fundamentals

As for the independent variable of interest, we define a variable $GHGSIM_{<i,j>,t}$ that measures the distance of carbon emissions produced by firm i and j in each year. To do so, we first compute the difference of carbon emission between firm i and firm j, and sort all the firm pairs based on the emission difference. Next, we transform this continuous variable into a decile variable from 1 to 10, where 10 denotes the difference in carbon emissions two firms produce on the same scale, meaning that they could be both brown firms, or they could both be green firms. On the contrary, a value of 1 denotes the two firms have distinct carbon profiles, where one firm could be very green and the other could be very brown. The transformation from 1 to 10 is mainly designed for easier interpretation of regression results, we also tried with logarithmic difference and the results are largely the same.

We mainly use Scope 1 and Scope 2 carbon emissions, as well as their intensity variable which is computed by scaling raw emissions by firm sales in the concurrent year. We also include firms' scope 3 emissions (both upstream and downstream, respectively) as well as their emission intensities in the robustness as well. However, the scope 3 emission data in our sample only begins after 2017. As a result, we do not rely heavily on this scope.

As for other control variables, we are using the same methodologies to compute the fundamental distance for each firm pair including firm size $SIZE_{<i,j>,t}$, the book-to-market ratio $B2M_{<i,j>,t}$, the leverage ratio $LEVERAGE_{<i,j>,t}$, the investment-to-asset ratio $INVEST2A_{<i,j>,t}$, $ROE_{<i,j>,t}$, the Herfindahl-Hirschman Index $HHI_{<i,j>,t}$, net Property, Plant, and Equipment $PPE_{<i,j>,t}$, sales growth rate $SALESGR_{<i,j>,t}$, and earnings per share growth rate $EPSGR_{<i,j>,t}$. Each fundamental variable is a decile variable where 10 denotes the highest similarity and 1 denotes the lowest similarity. For example, if firm i and j have $SIZE_{<i,j>,t}$ value of 10, then it means they are in the same size group and they could be both large firms or both small firms. Additionally, we control for industry fixed effects in our empirical designs by identifying a dummy variable that indicates whether two firms are in the same industry $INDUS_{<i,j>,t}$.

The emission data is obtained from the Trucost database, and the firm fundamentals data is obtained from COMPUSTAT. We match firms' carbon emissions and their firm fundamentals with the link table provided by WRDS and match the merged dataset with stock return data from CRSP. The sample period for emission and firm fundamentals is also from 2004 to 2019. We deliberately chose this period by excluding observations after 2019 to avoid the impact of Covid-19.

We present summary statistics in table 1. In panel A, we report firms' pair-wise return correlations on an annual basis, which are computed by calculating the 52-week return coefficient of each firm pair at the end of each year. We have 1526995 firm pair-year level correlation observations, with a mean correlation of 0.33 and a standard deviation of 0.22. In panel B, we report the scope 1, 2, and 3 carbon (both up-and-down) emissions for each firm. We scale the raw emission by its concurrent firm sales to derive emission intensity. For scope 1 and 2 emission data, we have a total of 7208 firm-year observations, which contrasts with scope 3 data with only 1891 firm-year observations. In panel C, we report firm fundamentals including flow intensity, which is computed by scaling investor flow by firm sales, its leverage ratio, book-to-market ratio, investment-to-asset ratio, ROE, Herfindahl-Hirschman Index, logarithmic value of net PPE, sales growth rate, and earnings growth rate. In the main regressions, we convert these continuous variables for each firm pair into decile firm-pair level variables that indicate the similarity of each fundamental.

[**Insert Table 1 near here**]

Similar to the carbon emission itself, the emission similarity measure is also highly persistent and has a high auto-correlation coefficient. To show this, we regress the future 1/3/5/7 years' emission similarity on the similarity in the current year, plus a dummy variable that indicates whether the two firms are in the same industry, other pair-level control variables and a year-fixed effect. We use scope 1 and 2 emissions, and scope 1 emission intensity and scope 2 emission intensity. The regression results are shown in table 2.

[**Insert Table 2 near here**]

As can be seen from table 2, for the emission similarity pair computed with scope 1 emission, the 1 year ahead auto-correlation coefficient is 0.9624 (with a t-statistic of 369.55), and the magnitude remains large and significant from year 3 to year 7. The autocorrelation of emission similarity for firm pairs is higher than the autocorrelation of the emission variable of itself (Bolton and Kacperczyk, 2021a; Li and Zheng, 2023) and the variation is insignificant across time. Similarly, for emission intensity computed with other carbon emissions, the coefficients are also close to zero, especially for the intensity-based measures. This implies that emission largely co-moves with the scaling variable as well, and firms that emit higher carbon emissions tend to have higher sales.

We also explore the determinants of emissions similarity. We follow Bolton and Kacperczyk (2021a) by regressing carbon emission intensity on an industry dummy and a set of fundamental difference measures, and we also control for year-fixed effect. We use emission distance computed with scope 1 and 2 emissions, and scope 1 emission intensity and scope 2 emission intensity data. The standard errors are clustered at the industry-pair level. The sample group is the 2004 stock sample, and the time period is from 2004 to 2019. The results are shown in table 3.

[Insert Table 3 near here]

As can be seen from table 3, the dummy variable $INDUS$ which denotes whether two firms are in the same emission decile has insignificant explanatory power of scope 1 and scope 2 emission similarity in columns 1 and 2, whereas it has significant regression coefficients where the dependent variable are scope 1 and 2 emission intensity. For the $SIZE$ variable, regression results suggest that size similarity is highly correlated with emission similarity in columns 1 and 2, similar to the coefficients in front of the $LOGPPE$ variable. This is quite intuitive as firms' emission level is highly related to firm size. However, once we scale the emission by firm sales in columns 3 and 4, the magnitude of the regression results becomes much smaller and even turns negative four scope 2 emission intensity.

3.3. Empirical design

We examine the relationship and its dynamics between return co-movement and emission similarity, which we define as carbon awareness. If stocks of firms that produce carbon emissions on a similar scale have higher co-movement patterns in recent years, whereas this co-movement pattern is barely significant in the early years, then it suggests that responsible investors are really walking their talks by jointly selling high-carbon stocks and buying low-carbon stocks.

To examine our conjecture, we perform the following regression, where on the left-hand side is the return correlation between firm pair i and j, and the variable of interest on the right-hand side is firms' emission similarity. We include a set of control variables including a dummy that indicates whether the two firms are in the same industry, along with other decile variables that indicate the distance between firm i and j fundamentals, and a higher decile variable indicates higher similarity. The fundamental decile variables include firm size, book-to-market ratios, leverage ratios, investment-to-asset ratio, ROE, Herfindahl-Hirschman Index, logarithmic value of net PPE, sales growth rate, and earnings growth rate. The standard errors are clustered at the industry-pair level. The emission similarity is measured by scope 1 and 2 emissions, and scope 1 emission intensity and scope 2 emission intensity. We also include time-fixed effects and cluster robust standard errors at the industry pair levels. The sample group is the 2004 stock sample, and the time period is from 2004 to 2019. The sample group is the 2004 stock sample, and the time period is from 2004 to 2019.

$$\rho_{<i,j>,t} = \alpha + \beta \cdot GHGSIM_{<i,j>,t} + \Gamma' X_{i,t} + \mu_t + \varepsilon_{<i,j>,t} \tag{1}$$

We perform this main regression in two ways. Firstly, we perform rolling regressions on a monthly basis, where the dependent variable on the left-hand side is the monthly return correlation, and the year-fixed effect is dropped out. We examine the coefficients in front of $GHGSIM$ on each cross-section level. Subsequently, we look for a structural mutation point from 2004 to 2019. We use a boosting-based machine learning algorithm to fit the implied trend, and test for potential mutation points. We also use more traditional methods including a Z-test and a Kolmogorov–Smirnov test to validate the difference in investors' carbon awareness before and after the mutation point.

Secondly, we perform the regression with a pooled OLS, where the observations are all on the annual level. We partition the sample period before and after the mutation point and perform sub-sample regressions. We test for differences in regression coefficients for the two sub-samples, so as to examine the difference in investors' raising carbon awareness.

4. Empirical results

After computing the variable that measures firms' emission similarity, we first perform the rolling regressions on a monthly basis to examine the structural mutation point, and then we use a pooled OLS further to validate the disparity in investors' carbon awareness.

4.1. Main results

Firstly, we derive the regression coefficients in front of emission similarity by equation 1 on a monthly frequency. This rolling regression produces a series of coefficients from each cross-sectional regression, and we define the coefficients as "carbon awareness", as it measures how stock returns are associated with firms' carbon emissions. We use four variables to measure carbon emissions, including scope 1 emission GHG1, scope 2 emission GHG2, scope 1 emission intensity, and scope 2 emission intensity. We plot the dynamics in figure 1. The black solid line denotes the coefficients obtained from each regression, and the yellow line denotes the fitted trend from a boosting algorithm, which implies the implied dynamics of carbon awareness.

[Insert Figure 1 near here]

In subfigure A, we plot the awareness dynamics with emission similarity computed with scope 1 carbon emissions. The regression coefficients of the black line suggest that, prior to 2012, the coefficients were insignificant and close to 0. After the year 2012, the regression coefficients gradually increased, implying a rising carbon awareness. The regression coefficient in the year 2020 is roughly 0.07, suggesting that an increase in emission similarity decile increases the return correlation coefficient by 0.007, this can be translated into a 2.12% (0.007/0.33) increase in investors' awareness, and the highest emission similarity pairs versus the lowest similarity pair has a difference in return correlation of 21.21%. Additionally, the boosting algorithm (the yellow line) suggests a structural mutation point around 2012. In subfigure B where emission similarity is computed with scope 2 carbon emission, the dynamics are largely the same, and the coefficients for the black line are quite volatile for the post-2012 period. The dynamics in subfigure 3 and 4 show similar results, where the implied trends remain relatively close to 0 and rose after 2012, and it has been increasing ever since. Moreover, the awareness measured by Scope 1 carbon emissions is significantly higher than that measured by Scope 2 carbon emissions, suggesting that investors care more about the direct emissions produced by the firm itself instead of indirect emissions resulting from the generation of purchased electricity, heat, or steam consumed by the organization.

We further split the time series of coefficients before and after 2012 and examine their differences for the carbon awareness estimated with the four different carbon emissions. We first examine the difference in coefficient means. For carbon awareness estimated with scope 1 carbon emissions, the sample mean before 2012 is -0.0016 (with a standard deviation of 0.0035) and 0.0053 after 2012 (with a standard deviation of 0.0037). The difference is -0.0069 and is statistically significant. We also perform a Two-sample Kolmogorov–Smirnov test, which is a statistical test used to compare two independent samples or datasets. It is a non-parametric test that assesses whether two samples share the same continuous distribution. The KS statistic is 0.6875 and is also significantly positive. As for carbon awareness estimated from scope 2 carbon emissions, the difference is -0.0026, and the KS statistic is 0.5927, all being statistically significant. Similarly, when we use carbon awareness computed with scope 1 emission intensity and scope 2 emission intensity, the differences are also statistically significant, where the regression coefficient is more prominent after 2012.

[**Insert Table 4 near here**]

After deriving the mutation point from previous analyses, we next perform a pooled OLS regression where the observations are all on the annual level. We pool the cross-sectional observations in December of each year into a large panel data and perform the regression formula as equation 1. The results are shown in table 5, where we report the carbon awareness estimated from four different emission measures, including scope 1 and 2 emissions, scope 1 intensity, and scope 2 intensity.

[**Insert Table 5 near here**]

In each sub-panel, we report regression results of emission similarity with and without control variables. In the first sub-panel where we use scope 1 carbon emissions, the regression coefficients are 0.0011 and 0.0104 (with t-stats of 1.56 and 8.02) for the before and after sample periods, respectively. We report the difference in regression coefficients in the bottom line, where their difference is -0.0093 (with a t-stat -6.31), suggesting an increase in investors' carbon awareness. When we control for other pair-level variables like industry pairs and size pairs, the regression coefficients are 0.0003 and 0.0092 (with t-stats 0.61 and 11.15). The difference in regression coefficients is -0.0089 (with a t-stat of -9.55). The regression coefficient for the after-2012 sample suggests that a one-unit increase of emission similarity increases firms' return correlation by 0.0092 units, a 2.78% (0.0092/0.33) increase. Moreover, most emission-similar pairs have a 27.87% higher return correlation than the lower emission-similar pairs. Interestingly, controlling for industry fixed effect and firm size pair decreases the magnitude for the coefficients before the 2012 period, while it has little effect on the coefficients after the 2012 period, and it also increases the statistical significance of carbon awareness we observed.

In the other sub-panels where we use scope 2 emissions, scope 1 emission intensity, and scope 2 emission intensity measures, the results are also largely the same, where regression coefficients in front of emission similarity are significantly higher in the sample period after 2012 than that of before 2012. The differences in the coefficients, as shown in the bottom line, are also significant, implying an increase in investors' carbon awareness.

To validate the robustness of the increased carbon awareness after the year 2012, we follow Lou et al. (2019) by performing double-sorting analyses on the correlation coefficients.

We first sort firm pairs into deciles from low to high based on firm fundamental similarities such as firm size, book-to-market ratios, leverage ratios, investment-to-asset ratio, ROE, Herfindahl-Hirschman Index, logarithmic value of net PPE, sales growth rate, and earnings growth rate. Then, we sort firm pairs by their emission similarity $GHGSIM$ and report the difference in return correlation coefficient between the highest $GHGSIM$ and the lowest $GHGSIM$ groups. We perform the double sorting methodology in each year and report mean values of correlation differences of each year. We use four different measures of carbon emissions including scope 1 and 2 emissions, scope 1 emission intensity, and scope 2 emission intensity to compute emission similarity. The sample group is the 2004 stock sample, and the time period is from 2012 to 2019. The results are shown in table 6.

[Insert Table 6 near here]

In panel A of Table 6, we report the double-sorting results for correlation coefficients. Results show that the return co-movement is consistently higher for the Hi-minus-low emission similarity portfolio, where the difference in coefficients in the last column are all positive around 0.10, suggesting that the high emission similarity group on average has a 0.10 higher return correlation than the lowest emission similarity group. In the first row where we first sort firm pairs by their size similarity, the Hi-minus-low coefficients are all significantly positive, and are relatively stable around 0.0981, suggesting that the pattern we have observed is not affected by firm size. In the remaining rows, the differences in coefficients are also significantly positive in each decile group, which adds to the robustness of our result. We also performed double-sorting analyses for the 2004 stock group using the before-2012 sample period, and the results are shown in the appendices. The correlation coefficients on average a much lower than that in the table 6.

One potential concern for rising carbon awareness may be that the increase in return co-movement is purely driven by fundamental co-movements. Since firms' carbon emissions are highly associated with their size, sales, and other unobserved industry heterogeneity, investors would include similar firms in their portfolios at the same time. In other words, stock returns co-move not because of investors' attention to sustainable investment, but because of their pure desire for financial returns.

We use several methods to address and alleviate this concern. Firstly, we control for the industry fixed effects, which may result in industry momentum and within-industry correlations. We include a dummy variable that indicates whether two firms are in the same industry using the 4-digit SIC codes in equation 1. Regression results in table 5 show that the coefficients are all significantly positive. In columns 3 and 4, regression coefficients are 0.1759 and 0.2512 (t stats are 8.88 and 9.92, respectively), and the magnitudes are comparable before and after 2012. Secondly, we control for other fundamental similarity decile variables in the regressions. Most notably, we control for the firm size where the variable $SIZE$ denotes whether two firms share are on a similar size scale. The regression coefficients in columns 3 and 4 of table 5 are also significant, with values of 0.0017 and 0.0064 (t stats are 8.93 and 20.56, respectively). This suggests that the size also matters for return correlation.

However, simply controlling for fundamentals does not solve the endogeneity issue. As a result, we resort to a state-level regulation shock for identification. We use exogenous shocks that affect firms' carbon emissions to see if they have a marginal impact on their stock return correlations as well.

States in the US have been actively engaged in promoting low-carbon styles. A wide range of policies has been adopted at the state and regional levels to reduce greenhouse gas emissions, develop clean energy resources, promote alternative fuel vehicles, and promote more energy-efficient buildings and appliances, among other things. According to the C2ES database, twenty-four states plus the District of Columbia have adopted specific greenhouse gas emissions targets by the end of 2022. While each state has adopted a target and baseline year that suits its circumstances, the prevalence of these targets shows the widespread support for climate action.

We took advantage of these state-level regulations as exogenous variations to firms' carbon emissions. We conjecture that, after a state has announced an emission reduction target, firms that operate within the state's border produce fewer carbon emissions, probably through greener technologies or less production. On the contrary, firms that operate within the states that have yet to announce an emission reduction target are unaffected. As a consequence, if a high-emission firm within a firm pair is subject to this emission regulation shock, whereas the low-emission firm within the firm pair does not, then we expect to see they should have a smaller difference in their carbon emissions and higher emission similarity $GHGSIM$. Following this regulation shock, we examine whether the regulation-induced change in emission similarity affects stock return correlations.

To perform the identification analyses, we define a dummy variable $GHGREGU$ that indicates whether the high-emission firm within a firm pair has experienced a regulation shock whereas the low-emission firm does not. We regress the correlation coefficients on this dummy variable and examine whether this coefficient gives positive results as shown in regression 2. We include a set of control variables and time effects as same as equation 1, and the standard errors are clustered at the industry pair level.

$$\rho_{<i,j>,t} = \alpha + \beta \cdot GHGREGU_{<i,j>,t} + \Gamma' X_{i,t} + \mu_t + \varepsilon_{<i,j>,t} \tag{2}$$

Since the $GHGREGU$ variable is computed based on whether the higher emission firm has experienced a regulation shock, we use four different measures of carbon emissions including scope 1 and 2 emissions, scope 1 emission intensity and scope 2 emission intensity to identify the high emission firms within a firm pair. The sample period is from 2012 to 2019, and we use the 2004 stock sample. The regression results are shown in table 7.

[Insert Table 7 near here]

In table 2, we show whether an increase in emission similarity that is caused by the states' regulations has a positive effect on firms' carbon emissions. In the first column where we use scope 1 emission to determine the high-emission stocks, the regression coefficient is 0.0111 (with a t-stat of 4.34). This positive regression coefficient suggests that after the emission similarity improves, the return co-movement of the two firms also increases. Besides, the regression coefficient is comparable in terms of magnitude to the regression results in table 5 (0.0092 versus 0.0111). We report regression results using $GHGREGU$ identified with scope 2 emissions, scope 1 emission intensity, and scope 2 emission intensity in columns 2, 3, and 4, respectively. The results are largely the same. When the high-emission firm within a firm pair reduces carbon emissions whereas the low-emission firm does not, then their return co-moves more significantly.

Our identification strategy is exempt from other concerns driven by firm fundamentals. The reason is that when the high-emission firm is subject to more stringent emission regulations, it may switch to an alternative cleaner technology or just simply produce fewer products. This may potentially detriment the cash flows for sure. However, the cash flow for the low-emission firm, which has yet to experience such a shock, is unaffected. As a consequence, if the alternative explanation based on firm financial performance is correct, we should observe a negative correlation coefficient result instead of a positive one. Moreover, since we have controlled for the industry-pair fixed effects and most of the firms in our sample are not concentrated within the same industry, the competition effect on the product market also does not account for the positive regression results we have documented.

5. Further analyses

After documenting the main results, we move on to examine the underlying mechanism that drives the results and perform additional robustness tests to further validate our findings.

5.1. Mechanism and heterogeneity results

We rely on a 2SLS-style system of equations to pin down the mechanism to show that investor flow leads to return co-movement. We use institutional investors' money flow as the mediation variable that links emission similarity and return co-movement. The investor flow is obtained by estimating equation 3 following van der Beck (2022), where we subtract the total investor holding amount at year t minus the expected cumulative investor holding amount. The $A_{i,t}$ is obtained from institutional investors' 13F files, and this measure is a simple adaption from the flow measure in the mutual fund literature. We scale the investor flow by the firm sales. We then transform this flow measure into a similarity measure $FLOW_{<i,j>,t}$ for any firm pair as well. A higher value of flow similarity indicates a similar flow pattern (either inflow or outflow), and a lower flow distance indicates a different flow pattern.

$$FLOW_{i,t} = A_{i,t} - A_{i,t} \times (1 + R_{i,t}) \tag{3}$$

Then, we perform the following system of equations, where we first regress flow similarity on emission similarity, plus a set of control variables and fixed effects as shown in equation 4. Then, we use the predicted value of flow similarity from the first stage regression as independent variables in the second stage, and regress return correlation on the fitted flow similarity as shown in equation 5. The regression coefficients of interest are the β_{1st} and β_{2nd} in the first and second stage regressions, respectively. This system of equations does treat the emission similarity as an exogenous instrumental variable in the first stage regression as normally the 2SLS would do, as carbon emission itself is clearly dependent on the firm's fundamental and industry heterogeneity. Instead, this set of regression mainly shows that the return co-movement originated from institutional investors' money flow co-movement, which is induced by investors' increasing awareness of ESG and sustainable investment. We use the the 2004 stock sample and the sample period is from 2012 to 2019. We cluster robust standard errors at the industry pair level for both equations.

$$FLOWSIM_{<i,j>,t} = \alpha + \beta_{1st} \cdot GHGSIM_{<i,j>,t} + \Gamma' X_{i,t} + \mu_t + \varepsilon_{<i,j>,t} \tag{4}$$

$$CORR_{<i,j>,t} = \alpha + \beta_{2nd} \cdot FLOW\hat{SIM}_{<i,j>,t} + \Gamma' X_{i,t} + \mu_t + \varepsilon_{<i,j>,t} \tag{5}$$

[Insert Table 8 near here]

The regression results are shown in table 8, where we report emission similarity computed with scope 1 and 2 carbon emissions, scope 1 emission intensity, and scope 2 emission intensity data. In columns 1 and 2, where we rely on scope 1 greenhouse gas for emission awareness, the regression coefficient of interest in the first stage is 0.0698 (with a t-stat of 16.11). This suggests that similar emission firms attract similar investor flows. If two firms have high emission similarity and are both brown firms, then investors tend to divest from these two firms at the same time. On the contrary, if the two firms produce less emissions, then investors are more inclined to simultaneously purchase stocks of these two firms. In the second stage, the regression result is 0.1321 (with a t-stat of 8.09), suggesting that emission-induced flow co-movement is positively related to return co-movement. This is consistent with the findings by Lou (2012) and Chen et al. (2019), as investor flow is highly associated with stock returns.

In columns 2, 3, and 4 where we compute emission similarity based on scope 2 emission, scope 1 emission intensity, and scope 2 emission intensity, the regression results are largely the same, implying that the return correlation induced by institutional investors' fund flow is indeed associated with investors' carbon awareness. Moreover, this effect is prevalent in both the emission level cohort (scope 1 and 2 emissions) and the intensity cohort (scope 1 and 2 emission intensity), which is different from previous findings as in Bolton and Kacperczyk (2021a) and Zhang (2023).

We further examine the heterogeneity behind the flow-induced co-movements. More specifically, we examine whether carbon awareness is more prominent among the low-carbon stock groups. We define a dummy variable that indicates whether the two firms within any firm pair both belong to the low-emission group. The low-emission group is defined as firms in the lowest emission quintile every year. We use scope 1 emissions, scope 2 emissions, scope 1 emission intensity, and scope 2 emission intensity to identify the low-emission firms. Then, we regress firms' return correlation coefficients on this dummy variable following equation 1 and examine whether it returns positive regression results. We use the 2004 stock sample and the time period is from 2012 to 2019. The regression results are shown in table 9.

[Insert Table 9 near here]

In table 9, regression results suggest that low-emission firm pairs do have higher return correlations. In column 1 where we identify the green firm pairs with scope 1 carbon emissions, the regression coefficient is 0.0620 (with a t-stat of 7.19). This coefficient is way larger than the regression coefficient in the main result (0.0092 versus 0.0620), implying that the co-movement within the green assets is most notably strong. Investors are simultaneously buying green assets, which increases their return co-movements. The result is also consistent in column 3, where the regression coefficient is 0.0835 (t-stat 7.98). However, the regression coefficients are not significant for the scope 2 carbon emission pairs, despite the signs that their regression coefficients are positive. This may imply that investors are mainly relying on direct carbon emissions to make investment decisions.

We next examine the correlation between firms' carbon similarity and future financial fundamental performance. This directly speaks to the concern about whether investors' decision is rational or not. To do so, we first regress the firm pair's future sales similarity and EPS similarity on current emission similarity, we include a set of control variables and fixed effects following 1. We use four different emission variables to compute emission similarity. We use sample the 2004 stock sample and the sample period is from 2012 to 2019. The regression results are shown in table 10.

[Insert Table 10 near here]

In column 1 of table 10, the regression coefficient in front of $GHGSIM$ is 0.0361 (with a t-stat of 8.56). This suggests that firms' emission co-movements predict future sales co-movement. The regression coefficient in front of the current $SALESGR$ is also significant, with a coefficient value of 0.2035 (with a t-stat of 68.77). Moreover, in column 5 where the dependent variable is future EPS growth rate, the regression result is also significantly positive, with a coefficient of 0.0687 and a t-stat of 15.58. This result provides evidence supporting the notion that firms' carbon emissions, as well as environmental performance, have predictive power for business operations and profitability. Based on the regression results from table 9, the positive regression results imply that firms with lower carbon emissions tend to have better future performance. The results in the remaining columns from columns 2 to 4 and columns 6 to 8 show the same results.

Overall, our empirical results show that, as investors' carbon awareness increases, they simultaneously purchase green stocks and sell brown stocks, which drives return co-movement. Additionally, we show that carbon awareness coincides with fundamental co-movements, as low-emission firms tend to have better future performance.

5.2. Carbon awareness and predictability

We also examine the predictability of carbon awareness. Since the return correlation and emission similarity are highly persistent variables, we expect to see strong predictability for carbon awareness.

We examine this problem by regressing firms' future return correlations on the emission similarity in the current year. We also add other control variables and fixed effects, with the same sample group and sample period as in equation 1. The regression results are shown in table 11.

[Insert Table 10 near here]

In table 11, the regression results in column 1, where we use scope 1 carbon emissions to compute emission similarity, are significantly positive, with a regression coefficient of 0.0093 (and a t-stat of 11.02). This suggests that as the emission similarity increases by 1 unit, the return correlation in the next year should increase by 0.0093 units. In the remaining columns, regression results produce similar results, the emission similarity shows high predictive power for future return correlations.

5.3. Carbon awareness with scope 3 emission data

The scope 3 emission data in our sample originates after 2017, which makes it difficult to examine the dynamics of carbon awareness with this scope data over a long sample period. Still, we examine the correlation-return relationship for the scope 3 data with our limited data sample and examine whether this coefficient is significantly distinct from zero.

We use scope 3 upstream emissions, downstream emissions, scope 3 upstream emission intensity, and downstream emission intensity data in this section, where the intensity is measured by scaling raw emission over firm sales. We include control variables and fixed effects in equation 1, and we use the 2004 stock sample. Finally, the sample period is from 2017 to 2019. The regression results are shown in table 12.

[Insert Table 12 near here]

In table 12, we show that post-2017, carbon awareness is also prominent for all types of scope 3 emission data. In columns 1 and 2 where we use scope 3 upstream carbon emission as a variable of interest to compute emission similarity, the regression coefficients are 0.0042 and 0.0043 (with t-stats 4.31 and 7.64, respectively) without and with control variables. The regression coefficients are smaller in terms of magnitude than the regression results in the main regressions, and yet, still positively significant. In the remaining columns from column 3 to 8, the regression coefficients are largely the same, suggesting that investors buying green stocks and selling brown stocks also happens within the scope 3 emission group.

6. Conclusion

This paper presents a unique approach to examining investor carbon awareness and its impact on asset prices, providing some key findings. Our empirical results underscore a steady increase in investors' carbon awareness post-2012, remaining largely unaffected by notable reversals such as the U.S. withdrawal from the Paris Agreement. This connection holds when controlling for various firm fundamentals and persists across different emission variables, an improvement over previous research which indicates conflicting results amid different emission data.

Our research explores the dynamics of investor behaviour, providing evidence that investors both buy green stocks and sell brown stocks, contributing to an increase in return correlations. Additionally, this behaviour seems driven by expectations of changes in firms' future growth and profitability, implying sustainable investment practices may bolster financial performance.

We use state-level emission reduction plans as an identification strategy, illuminating the real impact of regulation pressure on firms' behaviour and their subsequent reflection in the stock market. A significant observation highlights that carbon awareness became particularly pertinent after 2012, aligning with previous works.

Our study thus provides a fresh perspective to the carbon pricing literature, alleviating methodological constraints from previous studies, offering a consistent approach across different emission variables, and revealing the impact of regulatory pressure on firm behaviour. These contributions shed further light on how increasing environmental consciousness is shaping investor behaviour and the broader financial landscape.

As the focus on sustainable investing continues to grow, research like this will undoubtedly become more vital, helping to inform both policy and practice towards achieving more sustainable economic systems.

References

Acharya, V. V., Johnson, T., Sundaresan, S., Tomunen, T., 2022. Is physical climate risk priced? evidence from regional variation in exposure to heat stress. Tech. rep., National Bureau of Economic Research.

Aswani, J., Raghunandan, A., Rajgopal, S., 2023a. Are carbon emissions associated with stock returns? Review of Finance, forthcoming .

Aswani, J., Raghunandan, A., Rajgopal, S., 2023b. Are carbon emissions associated with stock returns? reply. Review of Finance, Forthcoming .

Baker, M., Bergstresser, D., Serafeim, G., Wurgler, J., 2018. Financing the response to climate change: The pricing and ownership of us green bonds. Tech. rep., National Bureau of Economic Research.

Barber, B. M., Morse, A., Yasuda, A., 2021. Impact investing. Journal of Financial Economics 139, 162–185.

Blitz, D., Fabozzi, F. J., 2017. Sin stocks revisited: Resolving the sin stock anomaly .

Bolton, P., Halem, Z., Kacperczyk, M., 2022a. The financial cost of carbon. Journal of Applied Corporate Finance 34, 17–29.

Bolton, P., Kacperczyk, M., 2021a. Do investors care about carbon risk? Journal of financial economics 142, 517–549.

Bolton, P., Kacperczyk, M., 2021b. Global pricing of carbon-transition risk. Tech. rep., National Bureau of Economic Research.

Bolton, P., Kacperczyk, M., Samama, F., 2022b. Net-zero carbon portfolio alignment. Financial Analysts Journal 78, 19–33.

Bolton, P., Kacperczyk, M. T., 2020a. Carbon premium around the world .

Bolton, P., Kacperczyk, M. T., 2020b. Signaling through carbon disclosure. Available at SSRN 3755613.

Bolton, P., Kacperczyk, M. T., 2021c. Carbon disclosure and the cost of capital. Available at SSRN 3755613 .

Bolton, P., Kacperczyk, M. T., Wiedemann, M., 2022c. The co2 question: Technical progress and the climate crisis. Available at SSRN .

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

Chava, S., 2014. Environmental externalities and cost of capital. Management science 60, 2223–2247.

Cheema-Fox, A., LaPerla, B. R., Serafeim, G., Turkington, D., Wang, H. S., 2021. Decarbonization factors. The Journal of Impact and ESG Investing .

Chen, H., Chen, S., Chen, Z., Li, F., 2019. Empirical investigation of an equity pairs trading strategy. Management Science 65, 370–389.

Choi, D., Gao, Z., Jiang, W., 2020. Attention to global warming. The Review of Financial Studies 33, 1112–1145.

Choi, D., Gao, Z., Jiang, W., Zhang, H., 2022. Carbon stock devaluation. Available at SSRN 3589952 .

Cohen, S., Kadach, I., Ormazabal, G., 2023. Institutional investors, climate disclosure, and carbon emissions. Journal of Accounting and Economics p. 101640.

Cremers, M., Ferrell, A., 2014. Thirty years of shareholder rights and firm value. The Journal of Finance 69, 1167–1196.

Duan, T., Li, F. W., Wen, Q., 2021. Is carbon risk priced in the cross-section of corporate bond returns? Available at SSRN 3709572 .

Edmans, A., 2011. Does the stock market fully value intangibles? employee satisfaction and equity prices. Journal of Financial economics 101, 621–640.

El Ghoul, S., Guedhami, O., Kwok, C. C., Mishra, D. R., 2011. Does corporate social responsibility affect the cost of capital? Journal of banking & finance 35, 2388–2406.

Engle, R. F., Giglio, S., Kelly, B., Lee, H., Stroebel, J., 2020. Hedging climate change news. The Review of Financial Studies 33, 1184–1216.

Garvey, G. T., Iyer, M., Nash, J., 2018. Carbon footprint and productivity: does the "e" in esg capture efficiency as well as environment. J Invest Manag 16, 59–69.

Giglio, S., Kelly, B., Stroebel, J., 2021. Climate finance. Annual Review of Financial Economics 13, 15–36.

Gompers, P., Ishii, J., Metrick, A., 2003. Corporate governance and equity prices. The quarterly journal of economics 118, 107–156.

Hong, H., Kacperczyk, M., 2009. The price of sin: The effects of social norms on markets. Journal of financial economics 93, 15–36.

Hong, H., Karolyi, G. A., Scheinkman, J. A., 2020. Climate finance. The Review of Financial Studies 33, 1011–1023.

Hsu, P.-H., Li, K., Tsou, C.-Y., 2022. The pollution premium. Journal of Finance, Forthcoming .

Ilhan, E., Sautner, Z., Vilkov, G., 2021. Carbon tail risk. The Review of Financial Studies 34, 1540–1571.

In, S. Y., Park, K. Y., Monk, A., 2017. Is "being green" rewarded in the market? an empirical investigation of decarbonization risk and stock returns. International Association for Energy Economics (Singapore Issue) 46.

Kim, S., Yoon, A., 2023. Analyzing active fund managers' commitment to esg: Evidence from the united nations principles for responsible investment. Management Science 69, 741–758.

Li, F., Zheng, X., 2023. Carbon emission and asset prices: new evidence from machine learning. Available at SSRN 4400681 .

Lou, D., 2012. A flow-based explanation for return predictability. The Review of Financial Studies 25, 3457–3489.

Lou, D., Polk, C., Skouras, S., 2019. A tug of war: Overnight versus intraday expected returns. Journal of Financial Economics 134, 192–213.

Luo, H. A., Balvers, R. J., 2017. Social screens and systematic investor boycott risk. Journal of Financial and Quantitative Analysis 52, 365–399.

Matsumura, E. M., Prakash, R., Vera-Munoz, S. C., 2014. Firm-value effects of carbon emissions and carbon disclosures. The accounting review 89, 695–724.

Monasterolo, I., De Angelis, L., 2020. Blind to carbon risk? an analysis of stock market reaction to the paris agreement. Ecological Economics 170, 106571.

Oestreich, A. M., Tsiakas, I., 2015. Carbon emissions and stock returns: Evidence from the eu emissions trading scheme. Journal of Banking & Finance 58, 294–308.

Pástor, L., Stambaugh, R. F., Taylor, L. A., 2021. Sustainable investing in equilibrium. Journal of Financial Economics 142, 550–571.

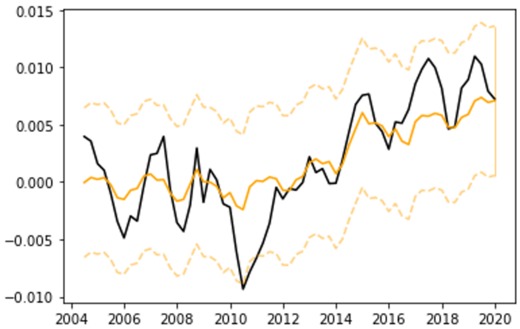
Pástor, L., Stambaugh, R. F., Taylor, L. A., 2022. Dissecting green returns. Journal of Financial Economics 146, 403–424.

Pedersen, L. H., Fitzgibbons, S., Pomorski, L., 2021. Responsible investing: The esg-efficient frontier. Journal of Financial Economics 142, 572–597.

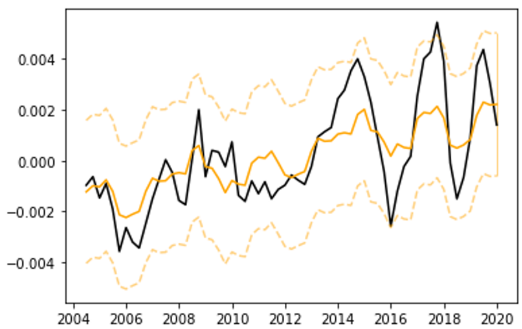
Skiadopoulos, G., Faccini, R., Matin, R., 2023. Dissecting climate risks: Are they reflected in stock prices? Journal of Banking and Finance .

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

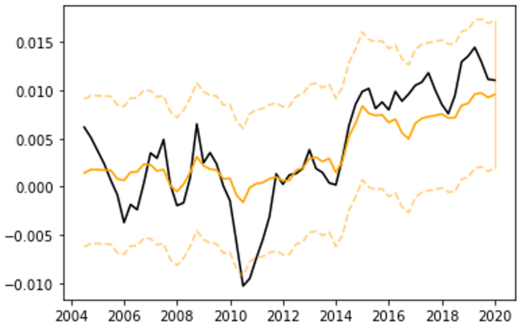
Zhang, S., 2023. Carbon returns across the globe. Available at SSRN 4378464 .



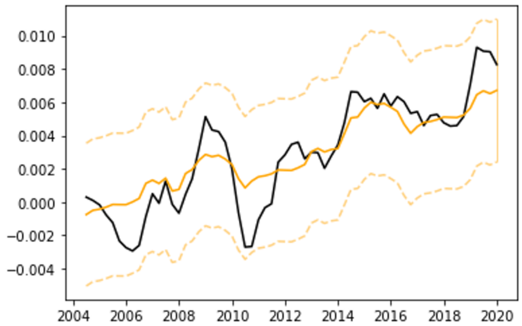
(a) GHG1



(b) GHG2



(c) GHG1 intensity



(d) GHG2 intensity

Fig. 1. Carbon awareness dynamics. This set of figures plots rolling regression coefficients which regress monthly stock return correlations on emission similarities, plus a set of control variables and fixed effects. The dependent variable $CORR$ is the pairwise Pearson correlation coefficient of the past 52 weeks' returns of stock i and j, and the independent variable of interest $GHGSIM$ is a decile variable that indicates the distance of carbon emissions of i and j, and a higher variable indicates the firm i and j produce carbon emission on a similar scale. If this variable has a value of 10 (the highest decile), then firms i and j are in the most "similar" decile. In contrast, a lower variable indicates the two firms have contrasting carbon emissions. If this variable has a value of 0 (the lowest decile), then firm i and j are in the most distant decile. We use four variables to measure carbon emissions, including scope 1 emission $GHG1$, scope 2 emission $GHG2$, scope 1 emission intensity $GHG1INTEN$, and scope 2 emission intensity $GHG2INTEN$. The control variables include a dummy variable that indicates whether firms are in the same industry, and other decile variables of firm fundamentals including Size, B2M ratio, leverage ratio, investment ratio, ROE, HHI, PPE, sales growth rate, and EPS growth rate. We control for industry-pair fixed effects in each cross-sectional regression. The standard errors are clustered at the industry-pair level. Finally, we perform the rolling regression by month and plot regression coefficients. The black solid line denotes the coefficients obtained from each regression, and the yellow line denotes the fitted trend from a boosting algorithm, which implies the implied dynamics of carbon awareness. The sample period is from 2004 to 2019.

Table 1: Summary statistics

	N	Mean	Std	Min	Q1	Median	Q3	Max
	Panel A: Return correlation							
Firm pair-year level								
CORR	1526995	0.33	0.22	-0.60	0.18	0.33	0.48	0.97
	Panel B: Carbon emission							
Firm-year level								
GHG1	7208	2803672.31	7664027.62	20.04	23480.49	127467.84	756107.65	36728360.74
GHG1_INTEN	7208	2.55	6.82	0.00	0.04	0.18	0.97	35.21
GHG2	7208	482270.73	816266.92	47.50	36436.99	130271.73	490432.51	3446871.86
GHG2_INTEN	7208	0.37	0.53	0.01	0.09	0.19	0.42	3.75
GHG3UP	1891	2463347.24	3801227.46	4660.95	279717.99	943506.71	2584748.45	17473029.44
GHG3DOWN	1891	6904126.67	15255811.99	0.04	59951.16	612910.78	4156686.29	60099686.31
GHG3UP_INTEN	1891	1.62	1.54	0.15	0.50	1.10	2.38	10.11
GHG3DOWN_INTEN	1891	7.77	22.25	0.00	0.10	0.90	4.36	139.91
	Panel C: Firm fundamentals							
Firm-year level								
FLOW_INTEN	7208	-22860.87	203941.79	-1124538.78	-80257.61	1321.11	74247.29	584872.37
LEVERAGE	7208	15.83	0.88	10.56	15.19	16.12	16.62	16.62
B2M	7208	0.53	0.46	0.07	0.25	0.42	0.70	6.68
LEVERAGE	7208	0.62	0.20	0.08	0.49	0.62	0.77	1.05
INVEST2A	7208	0.04	0.04	0.00	0.01	0.03	0.06	0.26
ROE	7208	0.15	0.29	-2.01	0.08	0.14	0.23	1.00
HHI	7208	0.27	0.23	0.01	0.09	0.20	0.36	1.00
LOGPPE	7208	7.31	1.48	1.11	6.27	7.45	8.86	9.01
SALESGR	7208	0.07	0.18	-0.50	-0.01	0.05	0.13	1.27
EPSGR	7208	0.14	1.49	-7.88	-0.19	0.09	0.34	5.16

This table reports summary statistics of the main variables in the empirical analysis. In panel A, we report firms' pair-wise return correlations on an annual basis, which are computed by calculating the 52-week return coefficient of each firm pair at the end of each year. In panel B, we report the scope 1, 2, and 3 carbon (both up-and-down) emissions for each firm. We scale the raw emission by its concurrent firm sales to derive emission intensity. In panel C, we report firm fundamentals including flow intensity, which is computed by scaling investor flow by firm sales, its leverage ratio, book-to-market ratio, investment-to-asset ratio, ROE, Herfindahl-Hirschman Index, logarithmic value of net PPE, sales growth rate, and earnings growth rate. In the main regressions, we convert these continuous variables for each firm into decile firm-pair level variables that indicate the distance of each fundamental. We use the 2004 stock sample as our sample group, which has 797 firms. The sample period is from 2004 to 2019.

Table 2: Auto-correlation of emission similarity

Measured by	GHGSIM															
	GHG1				GHG2				GHG1_INTEN				GHG2_INTEN			
	L1	L3	L5	L7	L1	L3	L5	L7	L1	L3	L5	L7	L1	L3	L5	L7
GHGSIM	0.9624	0.9233	0.8946	0.8745	0.9133	0.8224	0.7577	0.7100	0.9587	0.9171	0.8916	0.8714	0.8721	0.7523	0.6772	0.6277
	(369.55)	(224.29)	(181.49)	(143.09)	(170.09)	(99.31)	(98.24)	(107.22)	(423.15)	(222.54)	(162.28)	(134.77)	(125.48)	(58.04)	(41.16)	(32.33)
INDUS	0.0768	0.1989	0.2887	0.3786	0.1039	0.2501	0.3563	0.3684	0.1325	0.3064	0.4445	0.5913	0.2841	0.5954	0.8578	1.0182
	(1.70)	(2.14)	(2.21)	(2.46)	(1.49)	(1.79)	(1.81)	(1.63)	(5.22)	(7.79)	(8.37)	(9.02)	(3.90)	(4.19)	(4.23)	(4.29)
SIZE	-0.0049	-0.0111	-0.0150	-0.0197	-0.0111	-0.0259	-0.0413	-0.0538	-0.0043	-0.0095	-0.0132	-0.0151	0.0012	-0.0004	-0.0070	-0.0144
	(-5.12)	(-5.84)	(-7.04)	(-5.59)	(-5.65)	(-5.67)	(-7.31)	(-6.84)	(-3.97)	(-6.54)	(-5.98)	(-5.20)	(0.43)	(-0.07)	(-0.92)	(-1.52)
B2M	0.0069	0.0092	0.0140	0.0149	-0.0006	-0.0115	-0.0200	-0.0251	0.0069	0.0150	0.0262	0.0352	0.0072	0.0047	0.0051	0.0092
	(3.58)	(2.53)	(3.24)	(3.12)	(-0.16)	(-1.62)	(-1.56)	(-1.37)	(3.11)	(4.40)	(5.44)	(4.46)	(1.54)	(0.61)	(0.59)	(1.48)
LEVERAGE	-0.0049	-0.0090	-0.0107	-0.0114	-0.0067	-0.0126	-0.0134	-0.0118	-0.0023	-0.0023	-0.0017	-0.0009	0.0008	0.0026	0.0073	0.0118
	(-4.59)	(-4.23)	(-3.58)	(-2.85)	(-5.37)	(-3.85)	(-2.76)	(-1.96)	(-2.07)	(-0.97)	(-0.59)	(-0.27)	(0.29)	(0.55)	(1.25)	(1.57)
INVEST2A	0.0043	0.0025	0.0009	-0.0057	0.0019	-0.0036	-0.0067	-0.0089	0.0112	0.0205	0.0209	0.0221	0.0163	0.0276	0.0232	0.0253
	(1.57)	(0.60)	(0.22)	(-1.07)	(1.03)	(-1.31)	(-1.87)	(-1.53)	(5.06)	(5.70)	(4.28)	(4.42)	(4.11)	(5.30)	(3.12)	(3.38)
ROE	0.0014	0.0047	0.0060	0.0119	0.0113	0.0281	0.0402	0.0594	-0.0045	-0.0073	-0.0112	-0.0137	0.0006	0.0125	0.0170	0.0272
	(1.03)	(1.42)	(1.83)	(3.60)	(4.08)	(5.73)	(6.69)	(6.90)	(-4.00)	(-2.86)	(-3.71)	(-4.54)	(0.21)	(2.04)	(2.09)	(2.20)
HHI	0.0011	0.0015	0.0008	0.0002	0.0101	0.0220	0.0314	0.0410	-0.0023	-0.0054	-0.0076	-0.0100	0.0013	0.0052	0.0098	0.0142
	(0.87)	(0.64)	(0.24)	(0.06)	(4.49)	(5.85)	(5.62)	(6.22)	(-2.37)	(-2.69)	(-2.56)	(-2.67)	(0.52)	(0.98)	(1.35)	(1.65)
LOGPPE	0.0047	0.0096	0.0118	0.0135	0.0048	0.0031	-0.0001	-0.0046	0.0002	0.0010	-0.0004	-0.0023	-0.0053	-0.0108	-0.0132	-0.0146
	(4.97)	(4.28)	(4.29)	(4.65)	(1.72)	(0.42)	(-0.01)	(-0.33)	(0.29)	(0.54)	(-0.18)	(-0.96)	(-2.87)	(-2.54)	(-2.22)	(-2.14)
SALESGR	-0.0021	-0.0003	-0.0011	-0.0014	-0.0037	0.0000	-0.0032	-0.0026	-0.0002	-0.0004	-0.0026	-0.0037	-0.0009	0.0075	0.0031	-0.0051
	(-1.20)	(-0.11)	(-0.26)	(-0.24)	(-1.59)	(0.00)	(-0.75)	(-0.39)	(-0.05)	(-0.10)	(-0.69)	(-0.73)	(-0.24)	(1.33)	(0.43)	(-0.75)
EPSGR	0.0010	0.0018	0.0026	0.0067	-0.0022	0.0008	0.0058	0.0124	0.0029	0.0021	0.0053	0.0016	0.0042	0.0005	0.0163	0.0189
	(0.50)	(0.70)	(0.69)	(1.21)	(-1.11)	(0.18)	(0.93)	(1.77)	(1.62)	(0.68)	(1.45)	(0.31)	(1.24)	(0.08)	(2.49)	(3.33)
Const	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
Year FE	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
R2	0.93	0.86	0.81	0.77	0.84	0.69	0.59	0.53	0.93	0.86	0.81	0.78	0.77	0.57	0.47	0.40
N	1386789	1136046	917776	722229	1386789	1136046	917776	722229	1386789	1136046	917776	722229	1386789	1136046	917776	722229

This table examines the auto-correlation of emission similarity with scope 1 and 2 emissions, and scope 1 emission intensity and scope 2 emission intensity. The pair-level emission similarity variable is derived by converting the emission distance between firm i and j into a decile variable, where a higher value denotes carbon emissions on a similar scale. For emission similarity computed by each measure, we regress the future 1/3/5/7 years' emission similarity on the similarity in the current year, plus a dummy variable that indicates whether the two firms are in the same industry, other pair-level control variables and a year-fixed effect. The standard errors are clustered at the industry-pair level. The sample group is the 2004 stock sample, and the time period is from 2004 to 2019.

Table 3: Determinants of emission similarity

	GHGSIM			
Measured by	GHG1	GHG2	GHG1_INTEN	GHG2_INTEN
INDUS	1.3331	0.5982	2.0039	1.5360
	(1.28)	(1.12)	(2.10)	(3.10)
SIZE	-0.1366	-0.0880	-0.0307	0.0107
	(-22.91)	(-18.81)	(-5.65)	(2.43)
B2M	0.1277	-0.0125	0.1590	0.0295
	(11.97)	(-2.08)	(13.91)	(4.14)
LEVERAGE	-0.1229	-0.0826	-0.0842	-0.0082
	(-10.59)	(-10.45)	(-7.63)	(-0.93)
INVEST2A	0.2094	0.0720	0.2766	0.1569
	(17.87)	(9.26)	(22.35)	(16.44)
ROE	-0.0403	0.0467	-0.0577	0.0069
	(-5.19)	(9.84)	(-7.62)	(1.42)
HHI	0.0248	0.0930	-0.0280	-0.0010
	(1.46)	(9.87)	(-1.63)	(-0.10)
LOGPPE	0.1363	0.1201	0.0035	-0.0435
	(13.42)	(13.71)	(0.35)	(-7.52)
SALESGR	0.0067	-0.0068	0.0456	0.0365
	(1.59)	(-2.11)	(10.51)	(11.05)
EPSGR	0.0273	0.0036	0.0463	0.0577
	(5.99)	(1.03)	(9.61)	(15.64)
Const	T	T	T	T
Year FE	T	T	T	T
R2	0.11	0.04	0.12	0.04
N	1526995	1526995	1526995	1526995

This table examines the determinants of emission similarity with scope 1 and 2 emissions, and scope 1 emission intensity and scope 2 emission intensity. We report regression coefficients by regressing emission distance variables on a dummy variable that indicates whether the two firms are in the same industry and other pair-level decile variables plus a year-fixed effect. The standard errors are clustered at the industry-pair level. The sample group is the 2004 stock sample, and the time period is from 2004 to 2019.

Table 4: Break point test

	Before	After	Before-After	P-value	KS statistic	P-value
GHG1	-0.0016	0.0053	-0.0069	0.0000	0.6875	0.00
	(0.0035)	(0.0037)				
GHG2	-0.0011	0.0015	-0.0026	0.0000	0.5927	0.00
	(0.0012)	(0.0021)				
GHG1_INTEN	-0.0002	0.0078	-0.0080	0.0000	0.6875	0.00
	(0.0043)	(0.0043)				
GHG2_INTEN	0.0003	0.0054	-0.0051	0.0000	0.7752	0.00
	(0.0023	(0.0019)				

This table examines a structural mutation of carbon awareness in our whole sample period from 2004 to 2019. To do so, we first compute a time series of regression coefficients that regresses return correlations on emission similarities on a monthly level, plus other control variables. This contrasts with the main regressions and other analyses where we only use the pair-level variables of the month at the end of each year. We report the mean and standard deviation (in parentheses) of estimated coefficients for the periods before and after the year 2012 in each row. We report: (1) a difference test for the Before-after coefficients in columns 3 and 4; and (2) a Kolmogorov–Smirnov test that examines the maximum difference between the cumulative distributions between two coefficient samples in columns 5 and 6. The sample group is the 2004 stock sample, and the time period is from 2004 to 2019.

Table 5: Main results: emission similarity and return co-movement

Measured by	CORR															
	GHG1				GHG2				GHG1_INTEN				GHG2_INTEN			
Sample Period	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After	Before	After
GHGSIM	0.0011	0.0104	0.0003	0.0092	0.0010	0.0037	0.0004	0.0030	0.0020	0.0120	0.0010	0.0097	0.0020	0.0072	0.0006	0.0049
	(1.56)	(8.02)	(0.61)	(11.15)	(2.29)	(5.27)	(1.16)	(6.83)	(2.62)	(8.86)	(2.39)	(11.95)	(3.17)	(7.07)	(1.69)	(8.41)
INDUS			0.1759	0.2512			0.1759	0.2594			0.1737	0.2451			0.1748	0.2553
			(8.88)	(9.92)			(8.94)	(8.88)			(8.60)	(9.76)			(8.85)	(8.95)
SIZE			0.0017	0.0064			0.0017	0.0053			0.0017	0.0051			0.0017	0.0047
			(8.93)	(20.56)			(9.13)	(16.42)			(9.24)	(16.82)			(9.14)	(15.02)
B2M			0.0011	0.0035			0.0012	0.0044			0.0010	0.0032			0.0012	0.0041
			(3.64)	(7.16)			(3.98)	(8.78)			(3.14)	(6.37)			(3.96)	(8.32)
LEVERAGE			0.0020	0.0012			0.0020	0.0004			0.0020	0.0009			0.0019	0.0002
			(5.21)	(2.08)			(5.09)	(0.65)			(5.28)	(1.54)			(5.00)	(0.37)
INVEST2A			0.0012	0.0029			0.0012	0.0051			0.0010	0.0022			0.0012	0.0045
			(3.21)	(4.42)			(3.19)	(6.57)			(2.82)	(3.47)			(3.15)	(5.92)
ROE			0.0028	0.0040			0.0028	0.0036			0.0029	0.0041			0.0028	0.0037
			(13.17)	(7.53)			(12.97)	(6.24)			(13.68)	(7.97)			(12.87)	(6.46)
HHI			0.0001	-0.0006			0.0000	-0.0007			0.0001	-0.0001			0.0001	-0.0004
			(0.13)	(-0.65)			(0.07)	(-0.68)			(0.21)	(-0.12)			(0.14)	(-0.39)
LOGPPE			0.0028	0.0001			0.0028	0.0006			0.0028	0.0012			0.0029	0.0013
			(7.62)	(0.12)			(8.32)	(1.10)			(8.36)	(2.30)			(8.63)	(2.39)
SALESGR			0.0023	0.0043			0.0023	0.0046			0.0023	0.0040			0.0023	0.0044
			(11.87)	(15.76)			(11.80)	(16.34)			(11.83)	(14.38)			(11.93)	(15.79)
EPSGR			0.0007	0.0020			0.0007	0.0025			0.0007	0.0019			0.0007	0.0022
			(3.98)	(5.45)			(3.98)	(6.46)			(3.92)	(5.06)			(3.94)	(5.93)
Const	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
Year FE	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
R2	0.00	0.02	0.31	0.32	0.00	0.00	0.31	0.31	0.00	0.02	0.31	0.33	0.00	0.01	0.31	0.32
N	587796	852879	587796	852879	587796	852879	587796	852879	587796	852879	587796	852879	587796	852879	587796	852879
Difference	0.0093		0.0089		0.0027		0.0027		0.0101		0.0087		0.0052		0.0042	
Z-test	(6.31)		(9.55)		(3.26)		(4.85)		(6.45)		(9.49)		(4.41)		(6.08)	

This table examines investors' carbon awareness before and after the year 2012. We report regression coefficients that regress stocks' return correlation pairs on emission similarities at the end of each year, plus a set of control variables and a year-fixed effect. The control variables include a dummy that indicates whether the two firms are in the same industry, along with other decile variables that indicate the distance between firm i and j fundamentals, and a higher decile variable indicates higher similarity. The fundamental variables include firm size, book-to-market ratios, leverage ratios, investment-to-asset ratio, ROE, Herfindahl-Hirschman Index, logarithmic value of net PPE, sales growth rate, and earnings growth rate. The standard errors are clustered at the industry-pair level. The emission similarity is measured by scope 1 and 2 emissions, and scope 1 emission intensity and scope 2 emission intensity. In the bottom line, we report differences in regression coefficients in front of the emission similarity variables and Z-test statistics. The sample group is the 2004 stock sample, and the time period is from 2004 to 2019.

Table 6: Double sorting results

	1	2	3	4	5	6	7	8	9	10	Mean
					Panel A: GHG1						
SIZE	0.0543	0.0823	0.1002	0.1093	0.1047	0.1052	0.1036	0.1150	0.1088	0.0974	0.0981
	(3.61)	(6.45)	(6.41)	(8.03)	(7.03)	(6.58)	(7.32)	(7.38)	(6.68)	(5.60)	
B2M	0.0627	0.0532	0.0920	0.1009	0.1032	0.1038	0.0998	0.0885	0.0904	0.0867	0.0881
	(79.62)	(1.15)	(8.43)	(6.49)	(5.43)	(5.87)	(6.44)	(5.40)	(5.84)	(5.79)	
LEVERAGE	0.0987	0.0848	0.0832	0.0822	0.0945	0.0908	0.0837	0.0856	0.1116	0.1671	0.0982
	(6.13)	(6.98)	(7.06)	(5.70)	(5.89)	(6.06)	(6.18)	(5.77)	(6.80)	(9.34)	
INVEST2A	0.0405	0.0288	0.0508	0.0616	0.0740	0.0761	0.0845	0.0932	0.0991	0.2038	0.0812
	(1.77)	(2.07)	(5.22)	(5.76)	(6.58)	(4.99)	(5.32)	(5.72)	(6.29)	(9.92)	
ROE	0.0621	0.0701	0.0774	0.0743	0.0920	0.0902	0.1006	0.1183	0.1410	0.1505	0.0976
	(4.24)	(5.24)	(6.00)	(5.18)	(5.76)	(6.26)	(7.74)	(6.87)	(6.76)	(6.92)	
HHI	0.0688	0.0597	0.0969	0.0837	0.0934	0.0910	0.1015	0.0829	0.1125	0.1905	0.0981
	(6.95)	(4.77)	(6.37)	(6.23)	(5.77)	(5.05)	(5.10)	(6.53)	(4.57)	(9.88)	
LOGPPE	0.0550	0.0784	0.0887	0.0993	0.1062	0.1078	0.1088	0.1055	0.1224	0.0949	0.0967
	(5.05)	(6.35)	(6.79)	(6.13)	(6.43)	(6.66)	(6.02)	(6.15)	(8.58)	(6.14)	
SALESGR	0.0557	0.0791	0.0867	0.0960	0.1024	0.1050	0.1075	0.1115	0.1146	0.1172	0.0976
	(3.90)	(6.44)	(8.26)	(6.98)	(7.29)	(6.56)	(6.22)	(5.95)	(5.60)	(5.68)	
EPSGR	0.0888	0.0821	0.0657	0.0773	0.0821	0.0868	0.0891	0.1157	0.1322	0.1464	0.0966
	(4.45)	(8.55)	(7.19)	(4.58)	(6.18)	(6.68)	(5.71)	(5.98)	(6.21)	(6.15)	
					Panel B: GHG2						
SIZE	0.0543	0.0823	0.1002	0.1093	0.1047	0.1052	0.1036	0.1150	0.1088	0.0974	0.0981
	(3.61)	(6.45)	(6.41)	(8.03)	(7.03)	(6.58)	(7.32)	(7.38)	(6.68)	(5.60)	
B2M	0.0627	0.0532	0.0920	0.1009	0.1032	0.1038	0.0998	0.0885	0.0904	0.0867	0.0881
	(79.62)	(1.15)	(8.43)	(6.49)	(5.43)	(5.87)	(6.44)	(5.40)	(5.84)	(5.79)	
LEVERAGE	0.0987	0.0848	0.0832	0.0822	0.0945	0.0908	0.0837	0.0856	0.1116	0.1671	0.0982
	(6.13)	(6.98)	(7.06)	(5.70)	(5.89)	(6.06)	(6.18)	(5.77)	(6.80)	(9.34)	
INVEST2A	0.0405	0.0288	0.0508	0.0616	0.0740	0.0761	0.0845	0.0932	0.0991	0.2038	0.0812
	(1.77)	(2.07)	(5.22)	(5.76)	(6.58)	(4.99)	(5.32)	(5.72)	(6.29)	(9.92)	
ROE	0.0621	0.0701	0.0774	0.0743	0.0920	0.0902	0.1006	0.1183	0.1410	0.1505	0.0976
	(4.24)	(5.24)	(6.00)	(5.18)	(5.76)	(6.26)	(7.74)	(6.87)	(6.76)	(6.92)	
HHI	0.0688	0.0597	0.0969	0.0837	0.0934	0.0910	0.1015	0.0829	0.1125	0.1905	0.0981
	(6.95)	(4.77)	(6.37)	(6.23)	(5.77)	(5.05)	(5.10)	(6.53)	(4.57)	(9.88)	
LOGPPE	0.0550	0.0784	0.0887	0.0993	0.1062	0.1078	0.1088	0.1055	0.1224	0.0949	0.0967
	(5.05)	(6.35)	(6.79)	(6.13)	(6.43)	(6.66)	(6.02)	(6.15)	(8.58)	(6.14)	
SALESGR	0.0557	0.0791	0.0867	0.0960	0.1024	0.1050	0.1075	0.1115	0.1146	0.1172	0.0976
	(3.90)	(6.44)	(8.26)	(6.98)	(7.29)	(6.56)	(6.22)	(5.95)	(5.60)	(5.68)	
EPSGR	0.0888	0.0821	0.0657	0.0773	0.0821	0.0868	0.0891	0.1157	0.1322	0.1464	0.0966
	(4.45)	(8.55)	(7.19)	(4.58)	(6.18)	(6.68)	(5.71)	(5.98)	(6.21)	(6.15)	

This table reports double sorting results of return correlation. We first sort firm pairs into deciles from low to high based on firm fundamental similarities such as firm size, book-to-market ratios, leverage ratios, investment-to-asset ratio, ROE, Herfindahl-Hirschman Index, logarithmic value of net PPE, sales growth rate, and earnings growth rate. Then, we sort firm pairs by their emission similarity $GHGSIM$ and report the difference in return correlation coefficient between the highest $GHGSIM$ and the lowest $GHGSIM$ groups. We perform the double sorting methodology in each year and report mean values of correlation differences of each year. We use four different measures of carbon emissions including scope 1 and 2 emissions, scope 1 emission intensity and scope 2 emission intensity to compute emission similarity. The sample group is the 2004 stock sample, and the time period is from 2012 to 2019.

Table 6: Cont'd

	Panel C: GHG1_INTEN										
SIZE	0.0543	0.0823	0.1002	0.1093	0.1047	0.1052	0.1036	0.1150	0.1088	0.0974	0.0981
	(3.61)	(6.45)	(6.41)	(8.03)	(7.03)	(6.58)	(7.32)	(7.38)	(6.68)	(5.60)	
B2M	0.0627	0.0532	0.0920	0.1009	0.1032	0.1038	0.0998	0.0885	0.0904	0.0867	0.0881
	(79.62)	(1.15)	(8.43)	(6.49)	(5.43)	(5.87)	(6.44)	(5.40)	(5.84)	(5.79)	
LEVERAGE	0.0987	0.0848	0.0832	0.0822	0.0945	0.0908	0.0837	0.0856	0.1116	0.1671	0.0982
	(6.13)	(6.98)	(7.06)	(5.70)	(5.89)	(6.06)	(6.18)	(5.77)	(6.80)	(9.34)	
INVEST2A	0.0405	0.0288	0.0508	0.0616	0.0740	0.0761	0.0845	0.0932	0.0991	0.2038	0.0812
	(1.77)	(2.07)	(5.22)	(5.76)	(6.58)	(4.99)	(5.32)	(5.72)	(6.29)	(9.92)	
ROE	0.0621	0.0701	0.0774	0.0743	0.0920	0.0902	0.1006	0.1183	0.1410	0.1505	0.0976
	(4.24)	(5.24)	(6.00)	(5.18)	(5.76)	(6.26)	(7.74)	(6.87)	(6.76)	(6.92)	
HHI	0.0688	0.0597	0.0969	0.0837	0.0934	0.0910	0.1015	0.0829	0.1125	0.1905	0.0981
	(6.95)	(4.77)	(6.37)	(6.23)	(5.77)	(5.05)	(5.10)	(6.53)	(4.57)	(9.88)	
LOGPPE	0.0550	0.0784	0.0887	0.0993	0.1062	0.1078	0.1088	0.1055	0.1224	0.0949	0.0967
	(5.05)	(6.35)	(6.79)	(6.13)	(6.43)	(6.66)	(6.02)	(6.15)	(8.58)	(6.14)	
SALESGR	0.0557	0.0791	0.0867	0.0960	0.1024	0.1050	0.1075	0.1115	0.1146	0.1172	0.0976
	(3.90)	(6.44)	(8.26)	(6.98)	(7.29)	(6.56)	(6.22)	(5.95)	(5.60)	(5.68)	
EPSGR	0.0888	0.0821	0.0657	0.0773	0.0821	0.0868	0.0891	0.1157	0.1322	0.1464	0.0966
	(4.45)	(8.55)	(7.19)	(4.58)	(6.18)	(6.68)	(5.71)	(5.98)	(6.21)	(6.15)	
	Panel D: GHG2_INTEN										
SIZE	0.0543	0.0823	0.1002	0.1093	0.1047	0.1052	0.1036	0.1150	0.1088	0.0974	0.0981
	(3.61)	(6.45)	(6.41)	(8.03)	(7.03)	(6.58)	(7.32)	(7.38)	(6.68)	(5.60)	
B2M	0.0627	0.0532	0.0920	0.1009	0.1032	0.1038	0.0998	0.0885	0.0904	0.0867	0.0881
	(79.62)	(1.15)	(8.43)	(6.49)	(5.43)	(5.87)	(6.44)	(5.40)	(5.84)	(5.79)	
LEVERAGE	0.0987	0.0848	0.0832	0.0822	0.0945	0.0908	0.0837	0.0856	0.1116	0.1671	0.0982
	(6.13)	(6.98)	(7.06)	(5.70)	(5.89)	(6.06)	(6.18)	(5.77)	(6.80)	(9.34)	
INVEST2A	0.0405	0.0288	0.0508	0.0616	0.0740	0.0761	0.0845	0.0932	0.0991	0.2038	0.0812
	(1.77)	(2.07)	(5.22)	(5.76)	(6.58)	(4.99)	(5.32)	(5.72)	(6.29)	(9.92)	
ROE	0.0621	0.0701	0.0774	0.0743	0.0920	0.0902	0.1006	0.1183	0.1410	0.1505	0.0976
	(4.24)	(5.24)	(6.00)	(5.18)	(5.76)	(6.26)	(7.74)	(6.87)	(6.76)	(6.92)	
HHI	0.0688	0.0597	0.0969	0.0837	0.0934	0.0910	0.1015	0.0829	0.1125	0.1905	0.0981
	(6.95)	(4.77)	(6.37)	(6.23)	(5.77)	(5.05)	(5.10)	(6.53)	(4.57)	(9.88)	
LOGPPE	0.0550	0.0784	0.0887	0.0993	0.1062	0.1078	0.1088	0.1055	0.1224	0.0949	0.0967
	(5.05)	(6.35)	(6.79)	(6.13)	(6.43)	(6.66)	(6.02)	(6.15)	(8.58)	(6.14)	
SALESGR	0.0557	0.0791	0.0867	0.0960	0.1024	0.1050	0.1075	0.1115	0.1146	0.1172	0.0976
	(3.90)	(6.44)	(8.26)	(6.98)	(7.29)	(6.56)	(6.22)	(5.95)	(5.60)	(5.68)	
EPSGR	0.0888	0.0821	0.0657	0.0773	0.0821	0.0868	0.0891	0.1157	0.1322	0.1464	0.0966
	(4.45)	(8.55)	(7.19)	(4.58)	(6.18)	(6.68)	(5.71)	(5.98)	(6.21)	(6.15)	

Table 7: Identification: emission regulations and return comovement

	CORR			
Measured by	GHG1	GHG2	GHG1_INTEN	GHG2_INTEN
GHGREGU	0.0111	0.0132	0.0116	0.0135
	(4.34)	(6.25)	(4.56)	(5.90)
INDUS	0.2222	0.2222	0.2222	0.2221
	(12.04)	(12.03)	(12.05)	(12.05)
SIZE	0.0023	0.0023	0.0023	0.0023
	(10.01)	(10.05)	(9.89)	(9.84)
B2M	0.0022	0.0022	0.0023	0.0023
	(5.53)	(5.52)	(5.55)	(5.55)
LEVERAGE	0.0007	0.0007	0.0008	0.0008
	(1.60)	(1.57)	(1.66)	(1.66)
INVEST2A	0.0035	0.0035	0.0035	0.0035
	(6.85)	(6.88)	(6.83)	(6.83)
ROE	0.0034	0.0034	0.0034	0.0034
	(9.00)	(9.01)	(9.00)	(9.00)
HHI	0.0002	0.0002	0.0002	0.0002
	(0.27)	(0.27)	(0.27)	(0.28)
LOGPPE	0.0019	0.0019	0.0020	0.0020
	(4.82)	(4.76)	(4.89)	(4.93)
SALESGR	0.0029	0.0029	0.0029	0.0029
	(13.67)	(13.76)	(13.63)	(13.68)
EPSGR	0.0018	0.0018	0.0018	0.0018
	(6.57)	(6.53)	(6.58)	(6.55)
Const	T	T	T	T
R2	0.03	0.03	0.03	0.03
N	1526995	1526995	1526995	1526995

This table reports the identification results. We use whether a state has released an emission reduction initiative/statutory target as a regulation shock and perform staggered DID regressions to justify the casualty. To do so, we define a dummy variable that indicates whether the high-emission firm within a firm pair experiences a regulation shock that promotes cleaner production and reduces carbon emission subsequently, whereas the low-emission firm does not experience such a shock. We report regression coefficients by regressing return correlations on the dummy, along with other control variables. We use four different measures of carbon emissions including scope 1 and 2 emissions, scope 1 emission intensity and scope 2 emission intensity to compute emission similarity. The standard errors are clustered at the industry-pair level. The sample group is the 2004 stock sample, and the time period is from 2004 to 2019.

Table 8: Emission, investor flow, and return comovement

	CORR							
Measured by	GHG1		GHG2		GHG1_INTEN		GHG2_INTEN	
FLOW_INTEN		0.1321		0.0373		0.9691		0.2556
		(8.09)		(6.22)		(2.14)		(5.95)
GHGSIM	0.0698		0.0812		0.0100		0.0191	
	(16.11)		(28.55)		(2.23)		(6.49)	
INDUS	-0.6847	0.3416	-0.6369	0.2832	-0.6319	0.8575	0.5973	0.1026
	(-3.74)	(7.10)	(-3.83)	(8.13)	(-4.50)	(2.63)	(4.67)	(2.84)
SIZE	-0.1855	0.0309	-0.1851	0.0122	-0.1969	0.1959	0.1968	-0.0456
	(-56.80)	(9.79)	(-58.51)	(10.05)	(-61.21)	(2.19)	(62.10)	(-5.48)
B2M	-0.0267	0.0071	-0.0213	0.0052	-0.0211	0.0237	0.0186	-0.0006
	(-7.30)	(9.85)	(-6.08)	(9.83)	(-5.79)	(2.62)	(5.29)	(-0.44)
LEVERAGE	-0.0092	0.0025	-0.0114	0.0008	-0.0165	0.0169	0.0174	-0.0042
	(-2.36)	(2.43)	(-3.04)	(1.17)	(-4.26)	(1.90)	(4.48)	(-3.58)
INVEST2A	-0.0167	0.0051	-0.0047	0.0053	-0.0009	0.0030	-0.0057	0.0060
	(-4.66)	(4.93)	(-1.22)	(6.34)	(-0.24)	(0.79)	(-1.53)	(5.49)
ROE	0.0014	0.0038	-0.0037	0.0037	-0.0002	0.0043	0.0007	0.0035
	(0.48)	(5.22)	(-1.26)	(6.25)	(-0.07)	(1.43)	(0.23)	(3.78)
HHI	-0.0068	0.0003	-0.0131	-0.0002	-0.0048	0.0046	0.0052	-0.0017
	(-1.59)	(0.23)	(-3.04)	(-0.20)	(-1.12)	(0.95)	(1.18)	(-1.23)
LOGPPE	-0.0363	0.0049	-0.0395	0.0021	-0.0289	0.0292	0.0301	-0.0064
	(-11.61)	(5.86)	(-13.30)	(3.79)	(-9.52)	(2.24)	(9.92)	(-4.08)
SALESGR	-0.0455	0.0104	-0.0440	0.0063	-0.0438	0.0464	0.0421	-0.0064
	(-20.02)	(11.61)	(-19.10)	(13.47)	(-18.94)	(2.33)	(18.10)	(-3.37)
EPSGR	-0.0598	0.0099	-0.0576	0.0047	-0.0563	0.0564	0.0541	-0.0116
	(-22.96)	(8.39)	(-21.76)	(7.60)	(-21.75)	(2.18)	(20.84)	(-4.92)
Const	T	T	T	T	T	T	T	T
Year FE	T	T	T	T	T	T	T	T
F-Stat	8345.66	10021.18	8946.86	34398.48	8425.91	288.16	8526.51	4071.51
R2 (First stage)	0.05		0.06		0.05		0.05	
N	852879	852879	852879	852879	852879	852879	852879	852879

This table examines the mechanism that drives the observed return co-movement after 2012. We perform the following systems of equations, where we first regress a decile variable that indicates the flow similarity between firm i and firm j on emission similarity plus a set of control variables as well as a year-fixed effect. Next, we use the predicted value of flow similarity, which was estimated from the first stage regression, as independent variables in the second stage. We regress the return correlation between stocks of firm i and j on this fitted flow similarity to examine the mechanism. We use four different measures of carbon emissions including scope 1 and 2 emissions, scope 1 emission intensity and scope 2 emission intensity to compute emission similarity. The standard errors are clustered at the industry-pair level. The sample group is the 2004 stock sample, and the time period is from 2012 to 2019.

Table 9: Green firm pairs and return comovement

	CORR			
Measured by	GHG1	GHG2	GHG1_INTEN	GHG2_INTEN
BOTHGREEN	0.0620	0.0015	0.0835	0.0114
	(7.19)	(0.20)	(7.98)	(1.06)
INDUS	0.2529	0.2601	0.2399	0.2596
	(10.01)	(8.68)	(11.04)	(8.86)
SIZE	0.0052	0.0050	0.0047	0.0047
	(15.96)	(15.65)	(14.92)	(15.13)
B2M	0.0045	0.0044	0.0048	0.0044
	(9.19)	(8.66)	(10.02)	(8.71)
LEVERAGE	0.0006	0.0002	0.0004	0.0002
	(0.96)	(0.35)	(0.71)	(0.35)
INVEST2A	0.0047	0.0053	0.0044	0.0053
	(6.28)	(6.75)	(6.02)	(6.59)
ROE	0.0035	0.0037	0.0034	0.0037
	(6.22)	(6.51)	(6.08)	(6.55)
HHI	-0.0006	-0.0005	-0.0005	-0.0004
	(-0.56)	(-0.47)	(-0.48)	(-0.42)
LOGPPE	0.0012	0.0014	0.0015	0.0013
	(3.07)	(2.46)	(4.02)	(2.38)
SALESGR	0.0047	0.0046	0.0046	0.0046
	(16.55)	(16.25)	(16.82)	(16.37)
EPSGR	0.0025	0.0026	0.0023	0.0025
	(6.47)	(6.47)	(6.16)	(6.44)
Const	T	T	T	T
Year FE	T	T	T	T
R2	0.32	0.31	0.32	0.31
N	852879	852879	852879	852879

This table reports the heterogeneity results. We first sort firms by their carbon emissions from low to high, and then define a dummy variable that indicates whether both firms in a firm pair are in the low-emission group. We regress return correlations on this dummy variable, along with other control variables and a year-fixed effect. We use four different measures of carbon emissions to determine whether two firms are both considered to be ”green”, including scope 1 and 2 emissions, scope 1 emission intensity and scope 2 emission intensity. The standard errors are clustered at the industry-pair level. The sample group is the 2004 stock sample, and the time period is from 2012 to 2019.

Table 10: Emission similarity and future fundamental comovement

	SALESGR				EPSGR			
Measured by	GHG1	GHG2	GHG1_INTEN	GHG2_INTEN	GHG1	GHG2	GHG1_INTEN	GHG2_INTEN
GHG_SIM	0.0361	0.0390	0.0109	0.0167	0.0687	0.0552	0.0541	0.0537
	(8.56)	(8.83)	(3.97)	(5.93)	(15.58)	(11.77)	(18.48)	(15.54)
INDUS	0.3998	0.3545	0.4332	0.3890	0.3275	0.3063	0.3802	0.3382
	(2.66)	(1.51)	(2.55)	(1.50)	(2.42)	(1.37)	(2.51)	(1.40)
SIZE	0.0373	0.0868	0.0327	0.0824	0.0330	0.0818	0.0299	0.0790
	(14.51)	(36.97)	(12.96)	(35.77)	(13.26)	(34.79)	(12.04)	(33.99)
B2M	0.0357	0.0345	0.0393	0.0384	0.0302	0.0312	0.0359	0.0351
	(10.71)	(9.25)	(11.87)	(10.57)	(8.98)	(8.20)	(10.84)	(9.58)
LEVERAGE	0.0382	-0.0006	0.0348	-0.0039	0.0393	-0.0009	0.0346	-0.0045
	(11.04)	(-0.20)	(9.92)	(-1.16)	(11.61)	(-0.26)	(10.06)	(-1.37)
INVEST2A	0.0359	0.0067	0.0447	0.0159	0.0232	-0.0007	0.0364	0.0082
	(7.32)	(1.79)	(8.61)	(3.97)	(5.08)	(-0.19)	(7.43)	(2.19)
ROE	0.0009	0.0450	-0.0007	0.0432	0.0029	0.0463	-0.0001	0.0439
	(0.31)	(13.30)	(-0.23)	(12.96)	(1.00)	(13.64)	(-0.05)	(13.18)
HHI	-0.0323	-0.0086	-0.0325	-0.0093	-0.0294	-0.0060	-0.0313	-0.0076
	(-5.33)	(-1.92)	(-5.34)	(-2.07)	(-4.99)	(-1.36)	(-5.37)	(-1.74)
LOGPPE	0.0041	-0.0122	0.0065	-0.0102	0.0092	-0.0071	0.0109	-0.0052
	(1.03)	(-4.68)	(1.64)	(-3.81)	(2.43)	(-2.81)	(2.88)	(-2.05)
SALESGR	0.2035	0.0595	0.2046	0.0606	0.2003	0.0572	0.2019	0.0580
	(68.77)	(24.26)	(67.80)	(24.42)	(69.89)	(23.63)	(68.17)	(23.93)
EPSGR	0.0690	0.3035	0.0711	0.3056	0.0661	0.3017	0.0675	0.3021
	(28.04)	(73.94)	(28.34)	(74.59)	(26.71)	(72.51)	(26.88)	(73.34)
Const	T	T	T	T	T	T	T	T
Year FE	T	T	T	T	T	T	T	T
R2	0.07	0.14	0.07	0.14	0.07	0.14	0.07	0.14
N	852879	852879	852879	852879	852879	852879	852879	852879

This table examines the fundamental co-movement and emission similarity. We report regression coefficients where the dependent variables are future decile variables of firms' sales growth rate similarity and earning per share growth rate similarity in year $t+1$, and the independent variable of interest is the emission similarity in year t. We control for other firm fundamentals in the current year t along with a year-fixed effect. We use four different measures of carbon emissions including scope 1 and 2 emissions, scope 1 emission intensity and scope 2 emission intensity to compute emission similarity. The standard errors are clustered at the industry-pair level. The sample group is the 2004 stock sample, and the time period is from 2012 to 2019.

Table 11: Emission similarity and correlation predictability

	CORR			
Measured by	GHG1	GHG2	GHG1_INTEN	GHG2_INTEN
GHGSIM	0.0093	0.0026	0.0101	0.0052
	(11.02)	(5.92)	(12.38)	(9.00)
INDUS	0.2566	0.2652	0.2499	0.2605
	(10.34)	(9.12)	(10.22)	(9.26)
SIZE	0.0055	0.0043	0.0042	0.0038
	(17.24)	(12.79)	(13.39)	(11.64)
B2M	0.0029	0.0038	0.0025	0.0035
	(5.47)	(7.08)	(4.70)	(6.63)
LEVERAGE	0.0009	-0.0000	0.0006	-0.0001
	(1.41)	(-0.00)	(0.96)	(-0.21)
INVEST2A	0.0028	0.0050	0.0020	0.0044
	(4.06)	(6.11)	(3.00)	(5.39)
ROE	0.0027	0.0023	0.0029	0.0024
	(4.96)	(3.88)	(5.40)	(4.08)
HHI	-0.0009	-0.0010	-0.0004	-0.0007
	(-0.96)	(-0.94)	(-0.43)	(-0.68)
LOGPPE	0.0004	0.0010	0.0016	0.0017
	(0.70)	(1.79)	(2.99)	(3.10)
SALESGR	0.0027	0.0029	0.0023	0.0027
	(9.25)	(10.01)	(7.86)	(9.31)
EPSGR	0.0028	0.0033	0.0026	0.0030
	(7.53)	(8.56)	(7.11)	(8.05)
Const	T	T	T	T
R2	0.28	0.27	0.29	0.27
N	817175	817175	817175	817175

This table examines the future return correlation and emission similarity. We report regression coefficients where the dependent variables are future return correlation coefficients in year $t+1$, and the independent variable of interest is the emission similarity in year t. We control for other firm fundamentals in the current year t along with a year-fixed effect. We use four different measures of carbon emissions including scope 1 and 2 emissions, scope 1 emission intensity and scope 2 emission intensity to compute emission similarity. The standard errors are clustered at the industry-pair level. The sample group is the 2004 stock sample, and the time period is from 2012 to 2019.

Table 12: Return comovement with scope 3 emission

	CORR							
Measured by	GHG3UP		GHG3DOWN		GHG3UP_INTEN		GHG3DOWN_INTEN	
GHGSIM	0.0042	0.0043	0.0041	0.0032	0.0041	0.0029	0.0045	0.0028
	(4.31)	(7.64)	(4.52)	(5.89)	(3.86)	(5.52)	(4.24)	(5.06)
INDUS		0.2744		0.2769		0.2747		0.2752
		(7.35)		(7.18)		(7.17)		(7.13)
SIZE		0.0068		0.0066		0.0064		0.0063
		(17.52)		(16.66)		(16.21)		(15.90)
B2M		0.0023		0.0020		0.0021		0.0021
		(3.78)		(3.29)		(3.42)		(3.38)
LEVERAGE		-0.0004		-0.0004		-0.0005		-0.0005
		(-0.56)		(-0.59)		(-0.72)		(-0.73)
INVEST2A		0.0057		0.0055		0.0055		0.0055
		(7.66)		(7.36)		(7.54)		(7.47)
ROE		0.0021		0.0021		0.0022		0.0022
		(3.24)		(3.35)		(3.46)		(3.45)
HHI		-0.0003		-0.0000		0.0001		0.0001
		(-0.29)		(-0.05)		(0.07)		(0.13)
LOGPPE		0.0007		0.0010		0.0010		0.0011
		(1.22)		(1.62)		(1.74)		(1.81)
SALESGR		0.0044		0.0043		0.0043		0.0043
		(11.12)		(10.89)		(10.94)		(10.89)
EPSGR		0.0026		0.0026		0.0026		0.0026
		(4.97)		(4.84)		(4.92)		(4.87)
Const	T	T	T	T	T	T	T	T
Year FE	T	T	T	T	T	T	T	T
R2	0.00	0.04	0.00	0.04	0.00	0.04	0.00	0.03
N	416885	416885	416885	416885	416885	416885	416885	416885

This table examines the return correlation and emission similarity with scope 3 emission. We report regression coefficients where the dependent variables are return correlations, and the independent variable of interest is the emission similarity in year t. We control for other firm fundamentals in the current year t along with a year-fixed effect. We use four different measures of carbon emissions including scope 3 upstream and downstream, and their scaled intensity data to compute emission similarity, respectively. The standard errors are clustered at the industry-pair level. The sample group is the 2004 stock sample, and the time period is from 2017 (when the scope 3 data in our sample started) to 2019.