

Does Greenwashing Pay Off?

An Empirical Analysis of Cheap-Talk in
Firm Behaviour.

Jakub Przewoski

Administrative Number: u431929

Student Number: 2092491

Bachelor of Science in Economics

Tilburg University

Department of Economics

Supervisor: David Schindler

Date of Submission: June 6, 2025

Number of words:

Abstract

Your abstract goes here.

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 4 |
| 2 | Background | 5 |
| 3 | Data | 9 |
| 3.1 | Cheap-Talk Indicator | 9 |
| 3.2 | Refinitiv | 11 |
| 3.3 | Merging | 12 |
| 4 | Methodology | 13 |
| 4.1 | Default Difference-in-Differences | 13 |
| 4.2 | Multivalued Difference-in-Differences | 14 |
| 4.3 | Validity of Assumptions — Stock Prices | 17 |
| 4.4 | Validity of Assumptions — Sales | 17 |
| 5 | Results | 19 |
| 5.1 | Stock Prices Response to Grenewashing | 19 |
| 5.2 | Sales Response to Grenewashing | 19 |
| 6 | Discussion | 20 |
| A | Data | 27 |
| A.1 | Tables | 27 |
| A.2 | Cheap-Talk Index | 31 |
| A.3 | Merging | 31 |
| B | Models | 31 |
| B.1 | Formal assumptions | 31 |
| B.1.1 | Parallel Trends (PT) | 32 |
| B.1.2 | No anticipation | 32 |

List of Figures

| | | |
|---|--|----|
| 1 | Growth in New York Times Coverage by Topic | 6 |
| 2 | Google Search Trends of “Fridays For Future” | 7 |
| 3 | Mean Cumulative Abnormal Return (Daily) | 18 |
| 4 | Year-over-Year Change in Quarterly Sales | 19 |
| 5 | Event Study Results — Stock Prices | 22 |
| 6 | Event Study Results — Sales | 23 |

List of Tables

| | | |
|---|---|----|
| 1 | Results of Estimation | 20 |
| 2 | Robustness Checks Results | 21 |
| 3 | Average Causal Responses | 21 |
| 4 | Number of reports per year | 27 |
| 6 | Allocation of countries to regional markets | 28 |
| 5 | Sample Composition | 30 |
| 7 | Estimation periods for each frequency | 31 |

1 Introduction

In recent years, climate change has become an issue of a global importance. Policymakers, international organizations as well as consumers have been growing more aware of the problems facing the world in the upcoming years. In turn, heightened awareness has led to more and more people becoming educated on the topic of sustainability and climate change prevention. This awareness has lead to a rapid development of sustainable finance and a creation of a large body of ESG investment funds. Companies, trying to answer the needs of consumers have reoriented their supply chains and started transitioning to more sustainable methods of production. At the same time, this structural shift in demand created an avenue for greenwashing. In this setting, some companies can choose to appear as sustainable through making statements or campaigns aimed at sustainable consumers/investors, while not carrying through with their promises.¹ Recent theoretical literature suggests that there exist equilibria where it is optimal for companies to greenwash (Cartellier et al., 2023; Wu et al., 2020), as the investments in sustainability are largely unobserved. In sectors with large costs of the transition, this could potentially allow them to benefit from this shift in demand, while not bearing the costs.

This paper tries to answer the question: Is greenwashing a beneficial strategy for company performance in the presence of a positive sustainability demand shock? My approach consists of two steps. Firstly, I collect and process the sustainability reports for a sample of public companies, for which such reports are available. After processing, the reports are used to create a Cheap-Talk-Index (CTI) that gauges the credibility of climate commitments made by companies (Bingler et al., 2024). This index allows for measuring the likelihood of greenwashing following the contributions of Coen et al., 2022. Secondly, I try to estimate a causal response of cumulative abnormal returns (CAR) of companies stock prices' to greenwashing in presence of a demand shock to sustainability. Using the world-wide September 2019 strikes organized by Fridays For Future (20 & 27 September 2019)² and the occurrence of the “Global Week for Future”, as the shock and

¹Write the definition more explicitly and reference papers.

²A series of 4500 strikes across over 150 countries on the last two Fridays of September 2019

the CTI as the treatment, I test whether the markets have responded differently to firms more likely to greenwash. I find that ...This result can be explained by

Finally, I use the same procedure to investigate the effects on sales of the above-mentioned companies. I argue that companies trying to be perceived among investors as environmentally friendly, are likely to try to be perceived in the same way by their customers. Using the CTI, I test whether the sales of greenwashing companies have been substantially impacted following the shock. Results indicate that ...This result can be explained by ...

This study is related to the **X** streams of literature. First, it contributes to the literature on market reactions to greenwashing, by showing how the market has re-priced the value of the greenwashing firms in the presence of a shock.

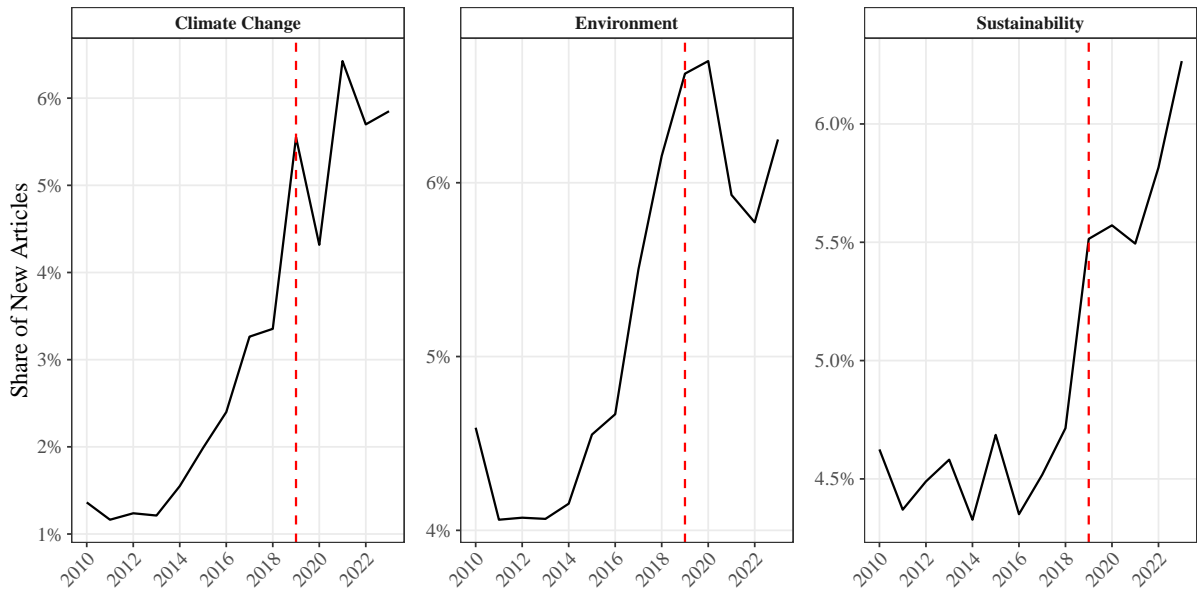
Secondly it draws on the literature concerning the relation between principals and managers, by checking whether the market is able to see through the greenwashing of firms. Studies concerning themselves with the impact of company environmental performance and its effects on the financial performance are many. Ilhan et al., 2023, suggests that

Finally, it contributes to literature on the link between advertising and consumer behaviour, by investigating the effects of greenwashing on the reactions of consumers.

2 Background

The year 2019 was an important year for climate change awareness as well as for climate change policy. The Fridays For Future (FFF) movement started in August 2018 by Greta Thunberg has been picking up in popularity following the release of a previous IPCC report on the commitments needed to stop the rising world temperatures. At the same time, 2019 witnessed many climate-related disasters. It was the second-hottest year since the start of the record keeping in 1880 (National Centers for Environmental Information (NCEI), 2019). According to The Guardian, extreme events such as floods, droughts, storms and wildfires have accounted for more than \$100bn worth of damages (Harvey,

Figure 1: Growth in New York Times Coverage by Topic



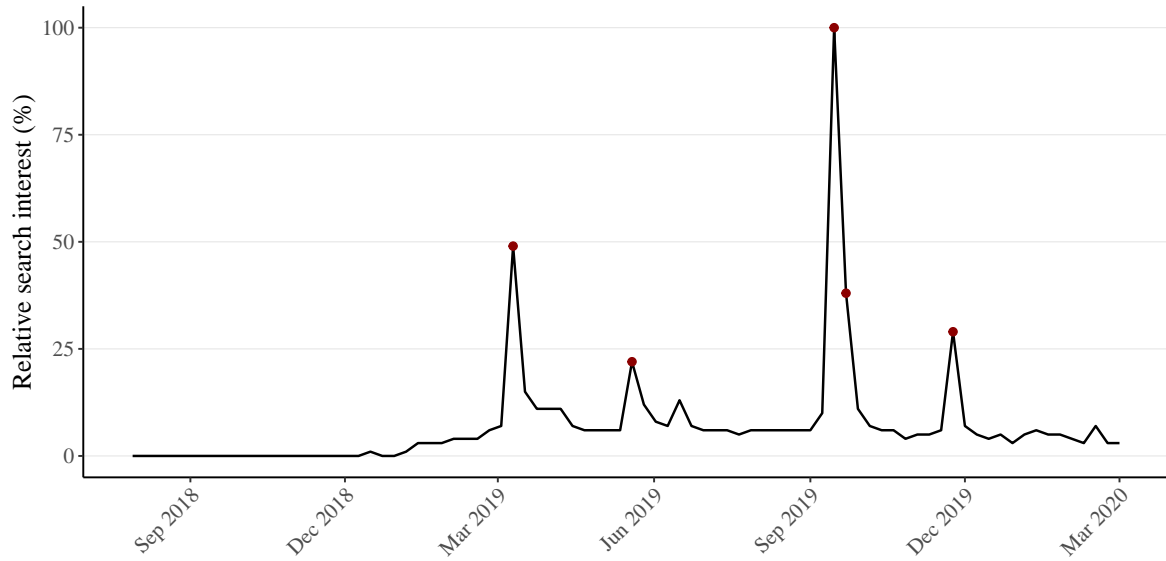
Number of new articles published per year (2010–2023). Red line indicates the year of the shock. Source: New York Times

2019). As evidenced by the Figure 1, 2019 marked a significant increase in new articles related to climate change compared to the previous years. While the overall number of new articles created by the New York Times has been in decline from 2012 until 2019, authors have been writing more articles revolving around climate change. Sustainability and Environment have also been more pronounced, however their temporary rise may be associated with the fact that more articles were written in general in 2020.

The pressure exerted on policymakers culminated in September during a series of strikes titled “Global Week for Future” (23–27 September 2019), during which the United Nations Climate Action Summit (UNCAS) took place. In the prior and during that week, FFF organized two global strikes across +270 countries that were attended by 4 million (for both the 20–09–2019 and 27–09–2019 ones) according to the organizers (Fridays for Future, n.d.). The number of participants as well as the interest were an order of magnitude larger than the previously organized events. Moreover, it was the first time when the youth movement was joined by adults with Watts and Laville of The Guardian describing it as:

For the first time since the school strikes for climate began last year, young people

Figure 2: Google Search Trends of “Fridays For Future”



Proportion of searches relative to the most searched, for the whole world. Red dots show major strikes organized by FFF. These strikes happened on the following dates (from left to right): 15-03-2019, 20-09-2019, 27-09-2019, 29-11-2019. Source: Google Trends

called on adults to join them – and they were heard. Trade unions representing hundreds of millions of people around the world mobilised in support, employees left their workplaces, doctors and nurses marched and workers at firms like Amazon, Google and Facebook walked out to join the climate strikes. (Watts & Laville, 2019)

As shown on the Figure 2, strikes related to that week attracted over twice as much attention online as any other strikes organized by FFF. This effect has been amplified by speeches given by Greta Thunberg at the climate strike in New York City and at the UNCAS in which she motivated the strikers as well as attacked the governmental officials for not taking appropriate action on climate change (Thunberg, 2019).

These events had an impact on financial markets. In a 2023 study, Schuster et al. found that the FFF global climate strikes significantly affected the short-term stock performance of firms from the S&P 500 and STOXX Europe 600 indices. They estimated a financial event study and find that the two global climate strikes in September 2019 generally hurt the performance of the companies in these indices.³

Overall, the action taken by the FFF grasped the imagination of the public with many

³Describe more of the study here

people choosing to walk out onto the streets and support their causes. The events gathered news and online attention and can be said to have had an impact on the performance of companies in that period.⁴

⁴Describe here also the events that happened afterwards: the COP of that year (December) and the IPCC report that came out earlier

3 Data

3.1 Cheap-Talk Indicator

The ESG reports were scraped from the ResponsibilityReports.com. Additionally, the website allowed me to retrieve information on company headquarters location and the approximate number of employees. Table 4 displays the number of reports available per year. As can be seen, the number of reports is non-homogenous across the years, as with time, communicating sustainability has become more important. As the analysis concerns the events of 2019, the scraped data was subset for companies reporting in 2018, to guarantee that the reports did not originate in response to the upcoming shock.

Sustainability reports contain information meant for investors and stakeholders. They inform the reader on initiatives undertaken by the company to manage its impact on the environment, its production standards, as well as the climate risk a company is exposed to. To illustrate, 3M Company’s 2018 Sustainability Report titled “Improving Lives” contains (among others) the following sections: “Our Values, Vision, and Strategies”, “Enterprise Risk”, “Environmental Management”.

As these reports do not fall under the oversight of any institutions (unlike annual earnings reports), the companies issuing them, may overstate their environmental achievements and as such use them to appeal to investors/consumers rather than present unbiased information. In essence, these reports allow for the companies to manipulate the perception of their environmental efforts (or their lack). Therefore, they introduce the possibility of greenwashing and deception.

To identify this phenomenon in relation to the environmental claims, the text was extracted from the reports. This has been done by using blocks contained in the PDF files to preserve the structure of the paragraphs.⁵ To further refine the data, each block has been classified as a paragraph, if it spanned at least 3 lines, contained at least 30 words and had at least one full stop. This has been done following the rule-based approach of Bingler et al. (2024).

⁵By *blocks*, I mean an object within the structure of a PDF file that usually determines the location of a paragraph on a page.

Using this data, I constructed the Climate Cheap-Talk Index (*CTI*) following the methods of Bingler et al. (2024). In this study, the authors utilize text from annual financial reports of a sample of companies to construct and train machine learning models that may be used to examine the level of cheap-talk in a given report. Their method consist of two steps. Firstly, they fine-tune an instance of the ClimateBert model developed by Webersinke et al. (2022) for classifying paragraphs as *climate related*. Then, they train another instance of the same model to classify paragraphs as specific (containing specific commitments) and non-specific (containing only vague commitments).⁶ The procedure that the authors used for training both of the models is described in Bingler et al. (2022, 2024).

By applying their models to my data, I was able to classify the scraped reports according to their methodology and construct the *CTI* by following their approach. Equation 1 presents the formula for the index:

$$CTI = \frac{CLIM_i \cap NONSPEC_i}{CLIM_i} \quad (1)$$

Where for a sustainability report of company i , $CLIM_i$ represents the number of climate-related paragraphs, $NONSPEC_i$ represents the number of paragraphs containing specific commitments, and $CLIM_i \cap NONSPEC_i$ represents the intersection between the two. This approach is motivated by the work of Coen et al., 2022, which finds that companies are less likely to follow-up on their commitments if these are stated in a vague and non-specific manner. Therefore, the *CTI*, interpreted as a measure of non-specificity of companies climate commitments, is used as a proxy for greenwashing. However, it must be noted that this is only an approximation, as greenwashing itself usually cannot be observed, unless a given company is involved in a scandal related to it. Nonetheless, as investors and consumers act on the *perceived* greenwashing of a company (as they also cannot observe it), this method was deemed as the best approximation available.⁷

Finally, the *CTI* was aggregated into quartiles, to create the final treatment-dose

⁶For more information on the models, examples of the researchers' data, and general information regarding the construction of the *CTI* see Appendix A.2

⁷Include the arguments of the authors from the paper!!!

variable D_i . The aggregation was carried out for two reasons. Firstly, one could argue that investors/consumers would not see the difference in behaviour of two companies whose score is very close. Therefore, dividing the continuous variable into separate bins would resemble the reasoning of the agents more closely and reduce the noise in the estimation. Secondly, as the index itself is very sensitive to the number of climate related paragraphs, grouping the treatment variable again reduces noise.

3.2 Refinitiv

The Refinitiv dataset has been used to retrieve data on firms' characteristics, their price indices and their quarterly earnings. Moreover, it was also used to retrieve the prices for regional market indices (e.g. S&P500, MSCI Europe, etc.) that were needed to estimate the normal returns for each company. The characteristic variables include the stock exchange of the listing, country of incorporation, country of domicile, the company's sector and the company's industry. The price indices (defined as the stock price expressed as a fraction of the initial listing price), were log-differenced to transform them into percentage changes. The same has been done for the quarterly earnings and for the values of regional stock indices. The data has been retrieved at weekly frequency for stock prices and at quarterly frequency for the earnings. Table 7 shows the total number of periods that were retrieved.

In order to plausibly estimate the effect of the treatment on stock returns, I needed to separate the effects of the market movement from the movement of the stock of the company itself. These abnormal returns were estimated following the market model approach outlined in Mackinlay (1997). This approach relies on estimating the expected stock return by predicting it using the return of the whole market. As the sample contains companies from multiple continents which markets may differ, each company was allocated to a geographical market index by their country of domicile. The allocation can be viewed in Table 6 in the Appendix A. The model has been separately estimated for each company using observations that predate the estimation window of interest. Intervals that were used are reported in Table 7. Below I report the regression specification

used for the estimation:

$$RN_{it} = \alpha_i + \beta_i RI_{mt} + \varepsilon_{it} \quad (2)$$

Where, RN_{it} is the return of the security i at time t predicted by the return of its geographical market m , α_i is the regression constant of for the company i , RI_{mt} is the return of the said market, and $\varepsilon_{i,t}$ is the noise. Finally, as this regression estimates the return of the security predicted by the return of the market, the abnormal return for i in a given time period t is given by the residual ε_{it} .

This type of modelling is sufficient for my use case, as per the work of Mackinlay, employing more sophisticated models (e.g. factor models) yields limited improvement in studies that do not rely on within-group variation, such as within-sector variation.

Moreover, the abnormal returns for each of the securities were later cumulatively summed to arrive at the Cumulative Abnormal Return (CAR). This was done, as aggregation across time is needed in order to draw useful inference from the results of the actual study. This results from the fact that estimating the average abnormal returns directly would not take into account the path of the data following treatment, but only the instantaneous effect of the treatment dose.

3.3 Merging

In the last step of the data preparation, the two datasets were merged. As this process was difficult and required multiple steps, its description is included in the Appendix A.3. Overall, the final dataset contained 795 companies, with 57 being dropped due to inability to be merged.⁸

⁸I believe that the 57 is incorrect, I still need to resolve it. Note: I am yet to include descriptive statistics here

4 Methodology

This section describes the methodology used in the attempt to estimate the causal effect of greenwashing on company performance. My initial approach was to use a classic Difference-in-Differences (DiD) methodology with binary treatment, however upon consideration I decided to implement a variation of DiD with multivalued treatment following Callaway et al. (2024). This approach is motivated by the fact, that the market may react differently to firms that use more cheap talk in their communication. This method also allows for estimation of experiments in settings without an untreated group, where the group with the lowest treatment-intensity is used as a comparison.

Given the use of newly developed methods in this study, I chose to contrast their methodology with the well-understood traditional DiD framework to better illustrate how this approach can improve upon the default approach. Therefore, the methodology is divided into three parts. Firstly, I will cover the conventional difference-in-differences design. Then I will contrast it with the implemented design using multivalued treatment. Finally, I will describe the differences in the taken approach when estimating the effect on sales of companies instead of stock prices.

4.1 Default Difference-in-Differences

Conventional Difference-in-Differences studies aim to estimate the causal effect of a policy by comparing units over time and then comparing those changes across groups. In order to separate the causal effect, they require two main assumptions explained below.⁹

Parallel trends (PT) assumption, which requires that in the absence of treatment, the outcomes of the units assigned to the treatment group would have evolved in the same way (in terms of the trend) as the outcomes of the units that were not assigned to treatment. In my example, it implies that in the absence of cheap-talk of a company, their stock price/sales would have evolved in the same way as the stock price/sales of those that do not engage in cheap talk.

No-anticipation assumption, imposes that the individuals do not act in anticipa-

⁹A formal treatment of these assumptions can be found in Appendix B.1

tion of the treatment. In my example, it implies that before the shock, the investors do not expect the greenwashing firms to perform differently from non-greenwashing firms after the shock, and therefore choose not to adjust their portfolios in advance of it. In other words, belonging to the treatment or the control group should not matter for the outcomes of the companies before the climate strikes shock.

Under these two assumptions and an additional assumption on independent sampling, the treatment effect can be recovered using the following regression specification utilizing two-way-fixed-effects (TWFE):

$$Y_{i,t} = \alpha_i + \Phi_t + \beta^{twfe} D_i \cdot Post_t + \varepsilon_{i,t} \quad (3)$$

Where Y_{it} is the outcome variable, α_i is the entity fixed effect, Φ_t is the time fixed effect, D_i is the binary treatment dummy, $Post_t$ is the post-treatment period dummy. In this setting, β^{twfe} recovers the average treatment effect.¹⁰

Although the interpretation of this treatment effect has been well described, it is not useful when analysing cases with multiple treatment doses, as it can be shown that the TWFE parameter becomes a weighted average of dose-*ATT*'s, with weights that deprive this parameter of its causal interpretation (Callaway et al., 2024). Therefore, in the following section, I describe the methodology of an adjusted framework.

4.2 Multivalued Difference-in-Differences

In order to incorporate the multivalued treatment into my study, I implemented the framework developed by Callaway et al. (2024). This is not the only approach present in the literature (see De Chaisemartin et al., 2024), however I deemed it the most suitable based on the required setting. This method was developed in the context of estimating the average treatment effect in settings where there is no treatment in pre-event periods. De Chaisemartin et al. differs in this regard, as it relies on treatment being positive in the periods prior to the event, and it requires a change in intensity in the event-period, to retrieve the average treatment effect. As in my case the treatment is fixed for a unit, but

¹⁰Note: Cite studies that mention problems with TWFE estimators?

varies in intensity between them, this requires the Callaway et al. (2024) approach. It proposes the use of a multivalued treatment, where the single dummy D_i can be replaced with multiple dummies $D_{i,d} \in \mathcal{D}$ that show that the unit i has received the dose d , where d refers to the “bins” that signify different treatment intensity. Then, this approach compares the dose of a group $D_{i,d}$, to the untreated group $D_{i,0}$.

This framework shares the no-anticipation and the independent sampling assumptions with the conventional DiD literature, however the differences arise in the parallel-trends assumption. In this setting, the traditional PT assumption allows for the estimation of each of the individual intensity parameters (i.e. ATT for each treatment is still recovered), however, a stronger assumption is needed to compare the effects across these estimators. Therefore, this assumption is required in order to assess the effect of higher greenwashing intensity of a firm.

The **Strong Parallel Trends (SPD) assumption** requires that the average evolution of outcomes for the entire treated population if all experienced dose d is equal to the path of outcomes that dose group d actually experienced (Callaway et al., 2024). This assumption is called the strong parallel trends, as it implies that the average response to the treatment would be the same for all dose groups, if they were to receive (for example) a lower dose d . It is more restrictive than the regular parallel trends assumption, and the researchers themselves argue that it may be unreasonable in many cases due to treatment effect heterogeneity of different groups. However, under this assumption, one can estimate the average causal response parameter, which the researchers denote as $ACRT$. They show that the $ACRT$ is equal to the difference in average potential outcomes between dose level d and the next lowest dose, scaled by the difference between these two doses.

Furthermore, the authors extend this framework to allow for estimation of treatment effects, where there are no untreated units, implying a setting where all dose-groups are compared to the lowest-dose group. While in the case with an untreated group the ATT estimates the effect of switching from no treatment to a treatment of dose d , they show that in this case, the ATT compares the change from low-intensity treatment to

high intensity treatment. This introduces a selection bias which cannot be solved with a standard parallel trends assumption, and therefore calls for the use of the strong parallel trends. Due to the fact that in my setting there are no explicitly untreated companies, but only companies that are in the lowest quartile of the CTI index, this assumption is required for the whole estimation process.

In order to retrieve the causal effect, I estimate the following multivalued specification:

$$Y_{i,t} = \alpha_i + \phi_t + \sum_{d=2}^4 \beta_d \cdot 1\{D_{i,d} = d\} \cdot Post_t + \varepsilon_{i,t} \quad (4)$$

Where $Y_{i,t}$ is the outcome variable, α_i are the unit fixed effects, ϕ_i are the time fixed effects, $1\{D_i = d\}$ is a dummy variable equal to 1 if the unit i belongs to the treatment group d , $Post_t$ is the dummy variable equal to one in the post-treatment period, and $\varepsilon_{i,t}$ is the standard normal noise.¹¹ Note that $d = 1$ is omitted, as the regression is estimated relative to the lowest dose group. In this specification, each β_d estimates the difference between $ATT(d)$ and the $ATT(d_L)$, which is average difference in outcomes when treated with d instead of d_L . Formally, it can be expressed as:

$$\beta_d = E[Y|D = d] - E[Y|D = d_L] = ATT(d) - ATT(d_L) = E[Y_{t=2}(d) - Y_{t=2}(d_L)] \quad (5)$$

This is a common approach adopted by many applied researchers (Acemoglu & Finkelstein, 2008; Deschênes & Greenstone, 2011). While this approach does not summarize the effect of the policy with one parameter, in the same way as the β^{twfe} does, Callaway et al. argue for this as their preferred specification. According to their study, estimating these parameters separately omits the weighting problem faced by specifications such as:

$$Y_{i,t} = \alpha_i + \phi_t + \beta^{twfe} Post_t \cdot (d_j - d_{j-1}) + \varepsilon_{i,t} \quad (6)$$

Where the β^{twfe} tries to aggregate the dose-specific ATT 's into a single number. They propose their own aggregation scheme that estimates ...¹²

¹¹Note: forgot to talk about the event window and the standard errors!

¹²Explain that the researchers have their own preferred way, but it does not they do not show whether it can be used with no-untreated unit scenario. State that therefore, you limit yourself to showing the

4.3 Validity of Assumptions — Stock Prices

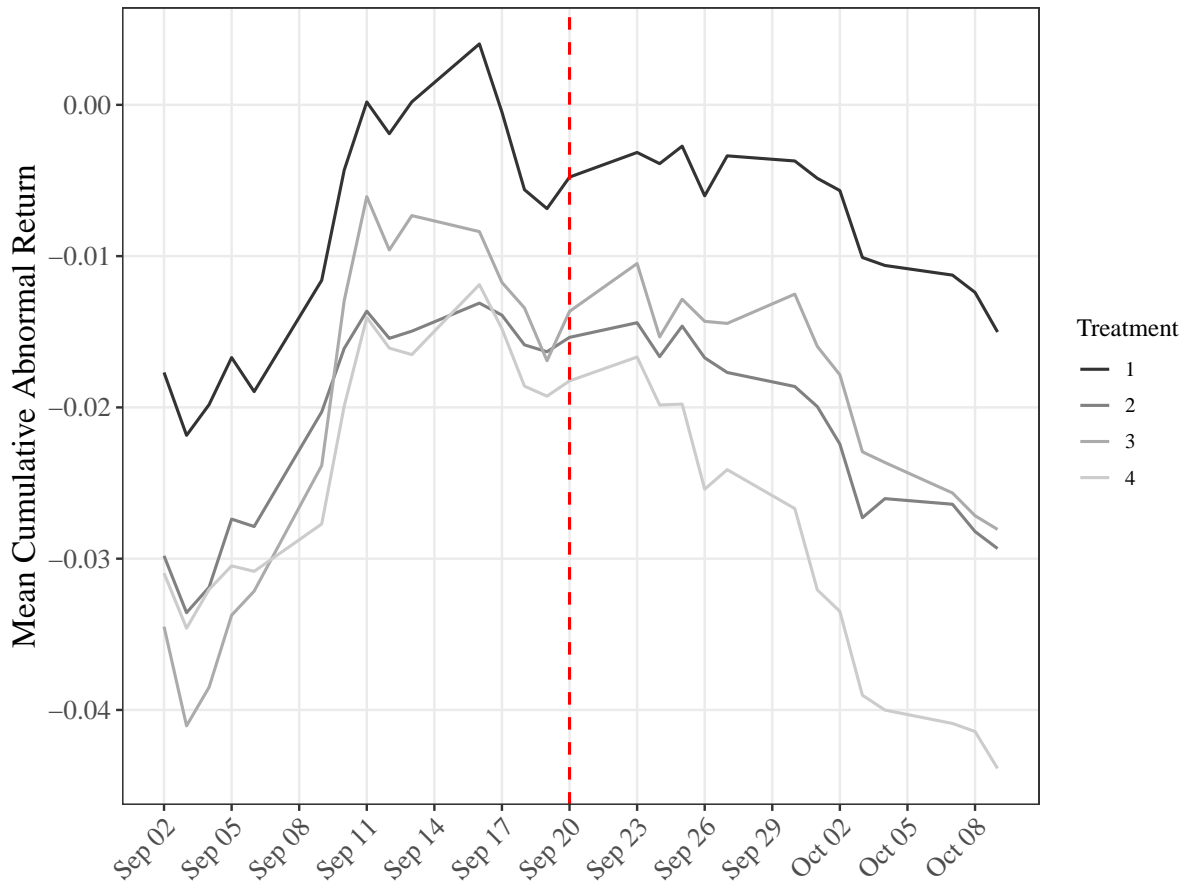
The chosen framework requires no-anticipation and the strong parallel trends assumptions. With respect to the SPT, this assumption may be concerning, as in the same way as in Koenig and Schindler (2023), the companies self-select into being treated. To alleviate this concern, Figure 3 shows that the firms stock prices followed similar paths no matter their treatment assignment. Furthermore, an event study is run in the next section to evaluate the possibility of pre-trends. However, it has to be noted that the authors of Callaway et al. (2024) state that this method of evaluating the existence of pre-trends has low power, and therefore the lack of significance within the pre-trends does not necessarily imply that they actually do not exist. Nonetheless, the event study is estimated as per the standard practice. With respect to the no-anticipation assumption, concerns are more pronounced. In their event study, Schuster et al. (2023) found evidence of anticipation of the stock market within a period of 20 days prior to the FFF strikes. In order to alleviate this concern, an analysis is performed with the shock date being shifted to multiple different dates preceding the shock.

4.4 Validity of Assumptions — Sales

The argument with respect to the sales data follows similar reasoning, however seems to be less plagued by the necessary assumptions. The relative lack of interest in climate change and sustainability seems to validate the fact that there should be no differences prior to the shock in the quarterly sales of treated and untreated companies. Therefore, as long as the treatment is evenly distributed across different firms (Figure 4), the SPT assumption should hold. Nonetheless, it is analysed with an event study. With regard to the no-anticipation assumption, one could argue that consumers may need more time to re-orient their spending to more sustainable sources, following such an increase in interest. In such case they would be less likely to adjust their consumption prior to the FFF climate strikes, assuming that they would have heard about these protest before they happened. Therefore, in this case the assumptions seem more probable. On the

dynamic $ATT(D_i)$ for each treatment group.

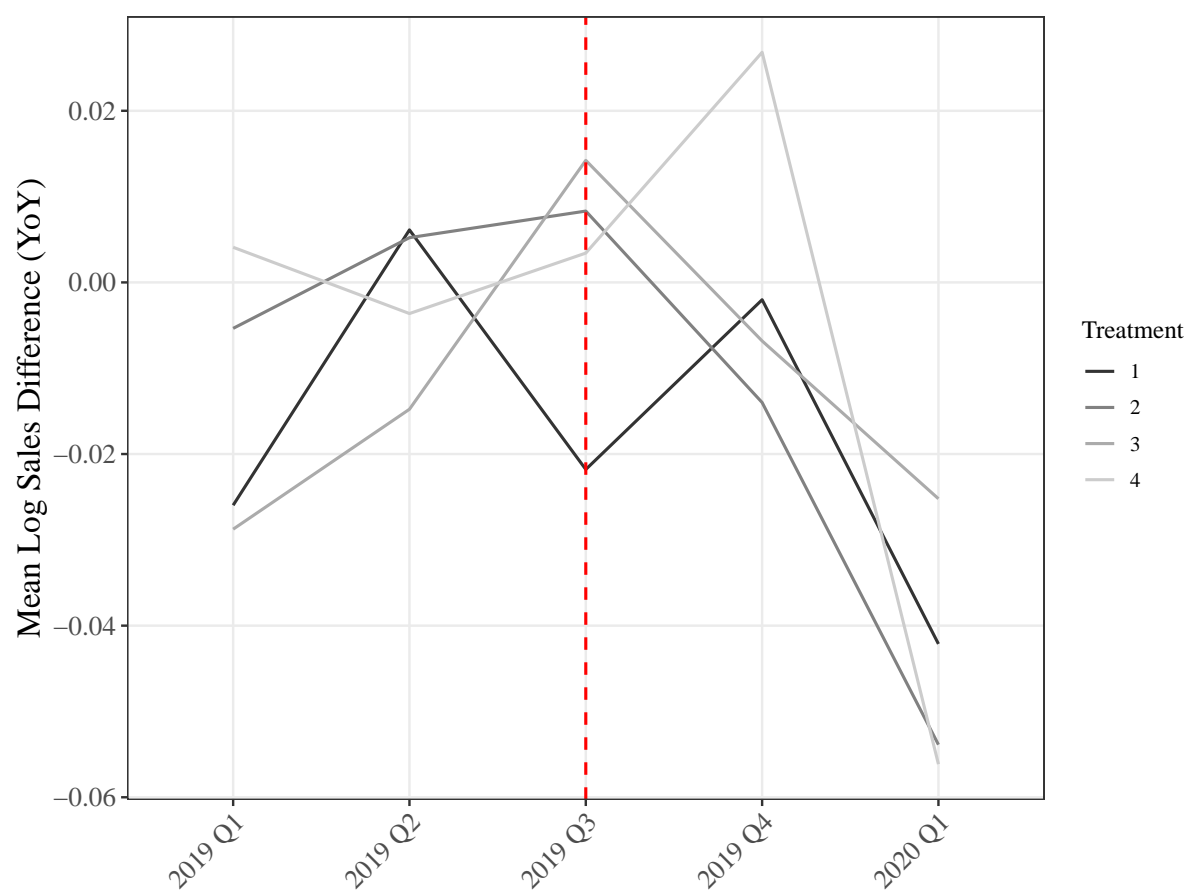
Figure 3: Mean Cumulative Abnormal Return (Daily)



Treatment Groups indicated with numbers from (lowest dose) 1–4 (highest dose). Red line indicates the first Fridays For Future climate strike during the Global Week For Future. Source: Refinitiv

other hand, one must acknowledge, that in this scheme, the CTI works only as a proxy to the belief of consumers about the greenwashing-intensity of the companies. This is not an obvious assumption, as firms may try to be more (or less) deceitful to consumers than to investors. However, given the lack of public information on marketing spending in relation to ESG, it is reasonable to assume that firms communicate similarly with both consumers and investors, and therefore that their ESG reports would reflect their sustainability-oriented marketing.

Figure 4: Year-over-Year Change in Quarterly Sales



Treatment Groups indicated with numbers from (lowest dose) 1–4 (highest dose). Red line indicates the quarter which witnessed the Global Week For Future. Source: Refinitiv

5 Results

5.1 Stock Prices Response to Grenewashing

5.2 Sales Response to Grenewashing

Result table:

test test

Testing the event study plot:

Table 1: Results of Estimation

| Variables | CAR (1) | Δ_4 Log of Sales (2) |
|-------------------|----------------------|--------------------------------|
| $Post \times D_2$ | -0.0020 (0.0054) | -0.0077 (0.0150) |
| $Post \times D_3$ | 0.0008 (0.0058) | 0.0279* (0.0158) |
| $Post \times D_4$ | -0.0091* (0.0054) | 0.0032 (0.0160) |
| Entity FE | ✓ | ✓ |
| Time FE | ✓ | ✓ |
| Observations | 17,752 | 2,515 |
| Within R^2 | 0.00266 | 0.00185 |
| R^2 | 0.90320 | 0.09844 |

Estimation Results of regressing the difference in differences on cumulative abnormal returns and the 4-quarter log difference of sales. Time FE is at the day level for (1) and Quarter for (2) **Add p-values and more description**

6 Discussion

List of things to add to the discussion:

- Could've weighted my regression somehow to make it more accurate?
- Better greenwashing measurements are needed

Table 2: Robustness Checks Results

| Variables: | CAR | |
|-------------------------|---------------------|---------------------|
| | (1) | (2) |
| $POST_{-5} \times D_2$ | -0.0018 (0.0056) | |
| $POST_{-5} \times D_3$ | 0.0017 (0.0059) | |
| $POST_{-5} \times D_4$ | -0.0070 (0.0056) | |
| $POST_{-10} \times D_2$ | | -0.0030 (0.0071) |
| $POST_{-10} \times D_3$ | | 0.0058 (0.0073) |
| $POST_{-10} \times D_4$ | | -0.0067 (0.0069) |
| Entity FE | ✓ | ✓ |
| Day FE | ✓ | ✓ |
| Observations | 17,752 | 17,752 |
| Within R^2 | 0.00169 | 0.00242 |
| R^2 | 0.90311 | 0.90318 |

Estimation Results of regressing the difference in differences on cumulative abnormal returns and the 4-quarter log difference of sales. **Add p-values and more description**

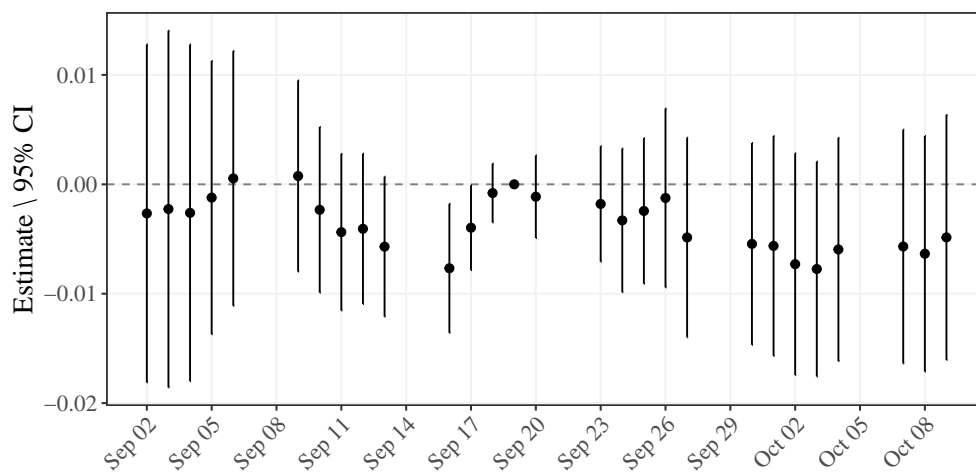
Table 3: Average Causal Responses

| Variables | CAR | Δ_4 Log of Sales |
|-------------|----------------------|-------------------------|
| | (1) | (2) |
| $ACRT(D_3)$ | 0.0028 (0.0053) | 0.0356** (0.0169) |
| $ACRT(D_4)$ | -0.0099* (0.0053) | -0.0247 (0.0178) |

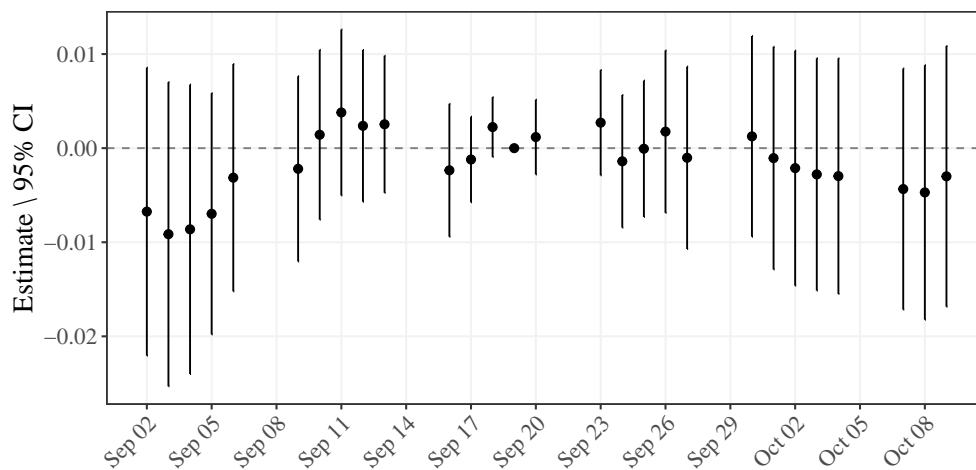
Estimation of ACRT resulting from the regression. $ACRT(D_2)$ is omitted as it can't be estimated (D_1 is the lowest dose)

Figure 5: Event Study Results — Stock Prices

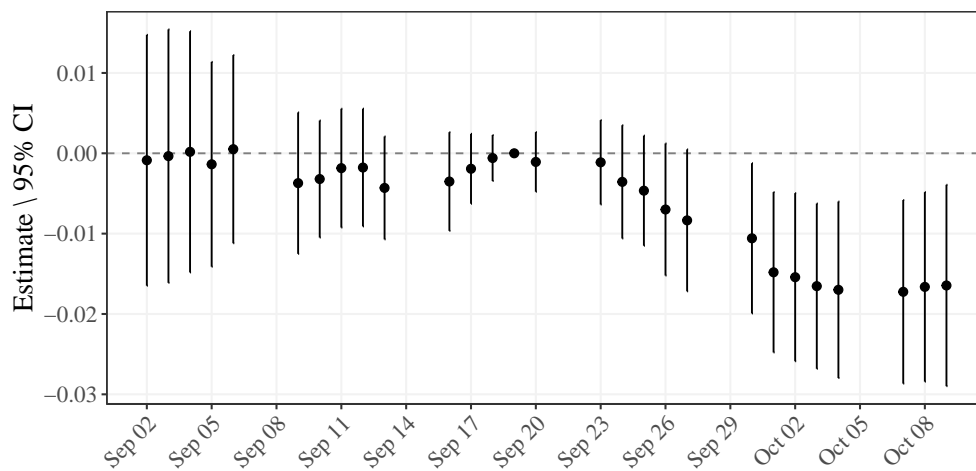
(a) Group 2



(b) Group 3



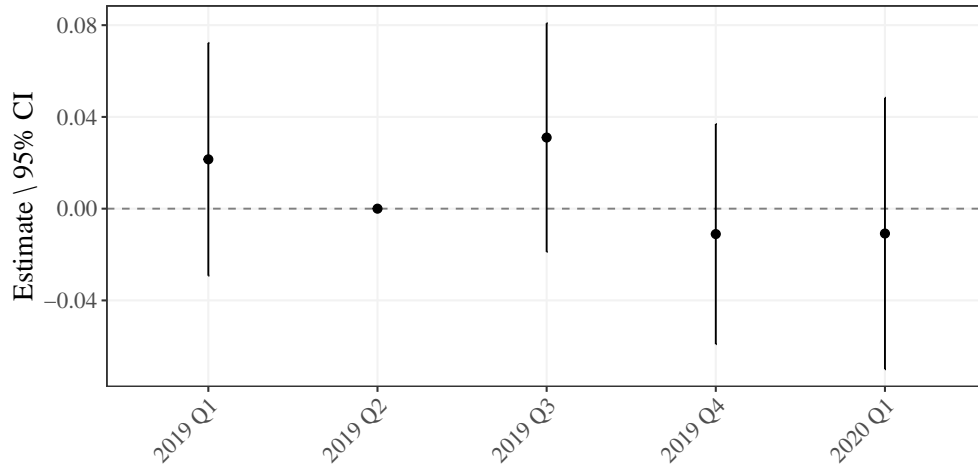
(c) Group 4



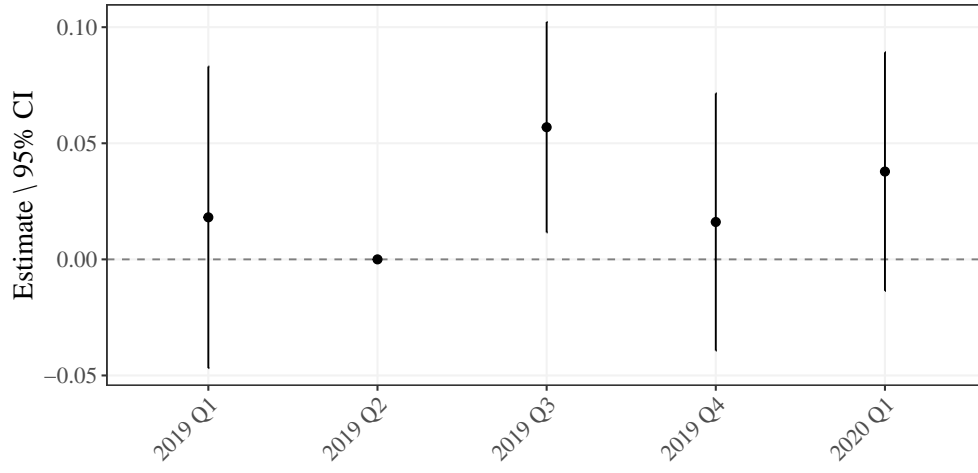
Stock price performance event studies for individual treatment groups. Groups 2,3,4 received their respective dose. **NOTE: Make sure that the final version is prettier**

Figure 6: Event Study Results — Sales

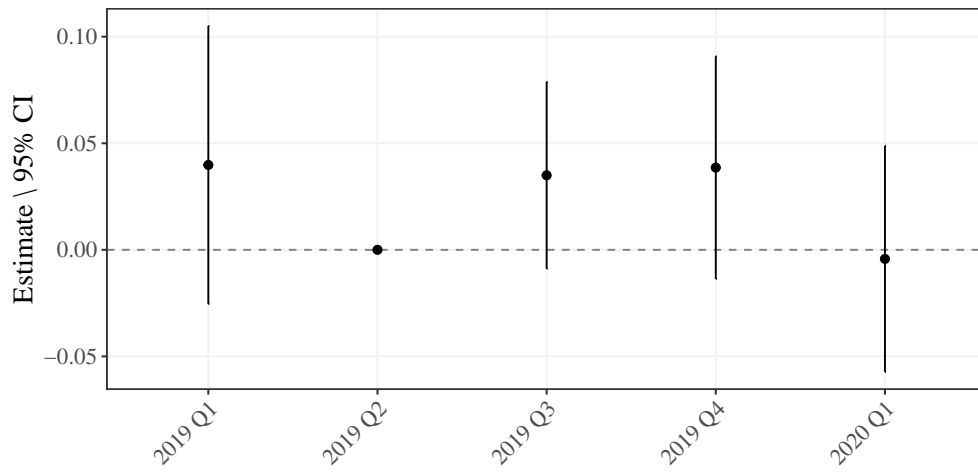
(a) Group 2



(b) Group 3



(c) Group 4



Sales performance event studies for individual treatment groups. Groups 2,3,4 received their respective dose. **NOTE: MAKE SURE THAT THE FINAL VERSION HAS QUARTERS AND NOT NUMS**

References

- 3M Company. (2018). *2018 sustainability report* (tech. rep.). 3M Company. St. Paul, MN.
<https://multimedia.3m.com/mws/media/1234567O/3m-2018-sustainability-report.pdf>
- Acemoglu, D., & Finkelstein, A. (2008). Input and technology choices in regulated industries: Evidence from the health care sector. *Journal of Political Economy*, *116*(5), 837–880. Retrieved June 1, 2025, from <http://www.jstor.org/stable/10.1086/595014>
- Bingler, J. A., Kraus, M., Leippold, M., & Webersinke, N. (2022). Cheap talk and cherry-picking: What ClimateBert has to say on corporate climate risk disclosures. *Finance Research Letters*, *47*, 102776. <https://doi.org/10.1016/j.frl.2022.102776>
- Bingler, J. A., Kraus, M., Leippold, M., & Webersinke, N. (2024). How cheap talk in climate disclosures relates to climate initiatives, corporate emissions, and reputation risk. *Journal of Banking & Finance*, *164*, 107191. <https://doi.org/10.1016/j.jbankfin.2024.107191>
- Callaway, B., Goodman-Bacon, A., & Sant’Anna, P. H. C. (2024). Difference-in-differences with a Continuous Treatment. *NBER WORKING PAPER SERIES*, (32117).
- Cartellier, F., Tankov, P., & Zerbib, O. D. (2023). Can Investors Curb Greenwashing? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4644741>
- Coen, D., Herman, K., & Pegram, T. (2022). Are corporate climate efforts genuine? an empirical analysis of the climate ‘talk–walk’ hypothesis. *Business Strategy and the Environment*, *31*(7), 3040–3059. <https://doi.org/10.1002/bse.3063>
- De Chaisemartin, C., D’Haultfœuille, X., & Vazquez-Bare, G. (2024). Difference-in-Difference Estimators with Continuous Treatments and No Stayers. *AEA Papers and Proceedings*, *114*, 610–613. <https://doi.org/10.1257/pandp.20241049>
- Deschênes, O., & Greenstone, M. (2011). Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the us. *American Economic Journal: Applied Economics*, *3*(4), 152–85. <https://doi.org/10.1257/app.3.4.152>

- Fridays for Future. (n.d.). List of countries – strike statistics [Accessed: 2025-05-17].
<https://fridaysforfuture.org/what-we-do/strike-statistics/list-of-countries>
- Harvey, F. (2019). Climate crisis linked to at least 15 \$1bn disasters in 2019 [Accessed: 2025-05-17]. *The Guardian*. <https://www.theguardian.com/world/2019/dec/27/climate-crisis-linked-to-at-least-15-1bn-plus-disasters-in-2019>
- Ilhan, E., Krueger, P., Sautner, Z., & Starks, L. T. (2023). Climate risk disclosure and institutional investors (I. Goldstein, Ed.). *The Review of Financial Studies*, 36(7), 2617–2650. <https://doi.org/10.1093/rfs/hhad002>
- Koenig, C., & Schindler, D. (2023). Impulse Purchases, Gun Ownership, and Homicides: Evidence from a Firearm Demand Shock. *Review of Economics and Statistics*, 105(5), 1271–1286. https://doi.org/10.1162/rest_a_01106
- Mackinlay, A. C. (1997). Event Studies in Economics and Finance. *Journal of Economic Literature*, 35(1), 13–39.
- National Centers for Environmental Information (NCEI). (2019). Global climate report - annual 2019 [Accessed: 2025-05-17].
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology*, 66(5), 688–701. <https://doi.org/10.1037/h0037350>
- Schuster, M., Bornhöft, S. C., Lueg, R., & Bouzzine, Y. D. (2023). Stock price reactions to the climate activism by Fridays for Future: The roles of public attention and environmental performance. *Journal of Environmental Management*, 344, 118608. <https://doi.org/10.1016/j.jenvman.2023.118608>
- Thunberg, G. (2019). Transcript: Greta thunberg’s speech at the u.n. climate action summit [Accessed: 2025-05-17].
- Watts, J., & Laville, S. (2019). Across the globe, millions join biggest climate protest ever [Accessed: 2025-05-17]. *The Guardian*. <https://www.theguardian.com/environment/2019/sep/21/across-the-globe-millions-join-biggest-climate-protest-ever>

- Webersinke, N., Kraus, M., Bingler, J., & Leippold, M. (2022). ClimateBERT: A Pre-trained Language Model for Climate-Related Text. *Proceedings of AAAI 2022 Fall Symposium: The Role of AI in Responding to Climate Challenges*. <https://doi.org/https://doi.org/10.48550/arXiv.2212.13631>
- Wu, Y., Zhang, K., & Xie, J. (2020). Bad greenwashing, good greenwashing: Corporate social responsibility and information transparency. *Management Science*, 66(7), 3095–3112. <https://doi.org/10.1287/mnsc.2019.3340>

A Data

ADD SOURCES TO ALL OF THE TABLES + ADD GENERAL NOTES
SO THAT THEY ARE SELF-EXPLANATORY

A.1 Tables

Table 4: Number of reports per year

| Year | Number of Reports |
|------|-------------------|
| 2023 | 1859 |
| 2022 | 2020 |
| 2021 | 1733 |
| 2020 | 1355 |
| 2019 | 1081 |
| 2018 | 852 |
| 2017 | 686 |
| 2016 | 558 |
| 2015 | 475 |
| 2014 | 415 |
| 2013 | 357 |
| 2012 | 300 |
| 2011 | 252 |
| 2010 | 216 |

Table 6: Allocation of countries to regional markets

| Country | Market Index |
|----------------|-----------------------|
| United States | S&P 500 Composite |
| Canada | S&P 500 Composite |
| Bermuda | S&P 500 Composite |
| Cayman Islands | S&P 500 Composite |
| Mexico | MSCI EM Latin America |
| Puerto Rico | MSCI EM Latin America |
| Costa Rica | MSCI EM Latin America |
| Barbados | MSCI EM Latin America |
| Panama | MSCI EM Latin America |
| Colombia | MSCI EM Latin America |
| Brazil | MSCI EM Latin America |
| Chile | MSCI EM Latin America |
| Peru | MSCI EM Latin America |
| Uruguay | MSCI EM Latin America |
| Argentina | MSCI EM Latin America |
| United Kingdom | MSCI Europe |
| Ireland | MSCI Europe |
| Switzerland | MSCI Europe |
| Netherlands | MSCI Europe |
| Greece | MSCI Europe |
| Germany | MSCI Europe |
| Belgium | MSCI Europe |
| Denmark | MSCI Europe |
| Monaco | MSCI Europe |
| Luxembourg | MSCI Europe |
| France | MSCI Europe |
| Sweden | MSCI Europe |
| Isle of Man | MSCI Europe |
| Spain | MSCI Europe |
| Finland | MSCI Europe |
| Romania | MSCI Europe |

Note: All indices denominated in dollars (but actually as % of their listing price I think)

| Country | Market Index |
|----------------------|--------------|
| Italy | MSCI Europe |
| Austria | MSCI Europe |
| Jersey | MSCI Europe |
| Guernsey | MSCI Europe |
| Turkey | MSCI Europe |
| Hong Kong | MSCI Pacific |
| Singapore | MSCI Pacific |
| Japan | MSCI Pacific |
| Australia | MSCI Pacific |
| New Zealand | MSCI Pacific |
| Papua New Guinea | MSCI Pacific |
| China | MSCI AC Asia |
| India | MSCI AC Asia |
| South Korea | MSCI AC Asia |
| Taiwan | MSCI AC Asia |
| Mongolia | MSCI AC Asia |
| Indonesia | MSCI AC Asia |
| Philippines | MSCI AC Asia |
| Israel | MSCI World |
| Kazakhstan | MSCI World |
| United Arab Emirates | MSCI World |
| South Africa | MSCI World |

Note: All indices denominated in dollars (but actually as % of their listing price I think)

Table 5: Sample Composition

| ICB Industry | | | | Number of Employees | | | | Country of Domicile | | | |
|------------------------|-----------|-----------|-----------|---------------------|-----------|-----------|-----------|---------------------|-----------|-----------|-----------|
| Industry | Total | Treated | Control | Employees | Total | Treated | Control | Country | Total | Treated | Control |
| Industrials | 120 (19%) | 59 (18%) | 61 (19%) | 10,000+ | 487 (62%) | 250 (63%) | 237 (61%) | United States | 376 (59%) | 164 (51%) | 212 (66%) |
| Consumer Discretionary | 99 (15%) | 30 (9.4%) | 69 (21%) | 1001–5000 | 134 (17%) | 70 (18%) | 64 (16%) | United Kingdom | 64 (10%) | 38 (12%) | 26 (8.1%) |
| Financials | 87 (14%) | 65 (20%) | 22 (6.9%) | 5001–10,000 | 120 (15%) | 54 (14%) | 66 (17%) | Canada | 55 (8.6%) | 30 (9.4%) | 25 (7.8%) |
| Basic Materials | 83 (13%) | 347 (15%) | 36 (11%) | 201–500 | 20 (2.5%) | 10 (2.5%) | 10 (2.6%) | Australia | 48 (7.5%) | 31 (9.7%) | 17 (5.3%) |
| Consumer Staples | 57 (8.9%) | 20 (6.3%) | 37 (12%) | 501–1000 | 13 (1.6%) | 5 (1.3%) | 8 (2.1%) | Other | 97 (15%) | 56 (17%) | 41 (12%) |
| Technology | 52 (8.1%) | 36 (11%) | 16 (5.0%) | 11–50 | 8 (1.0%) | 5 (1.3%) | 3 (0.8%) | Unknown | 155 | 84 | 71 |
| Utilities | 47 (7.3%) | 10 (3.1%) | 37 (12%) | 51–200 | 6 (0.8%) | 4 (1.0%) | 2 (0.5%) | | | | |
| Health Care | 46 (7.2%) | 28 (8.8%) | 18 (5.6%) | 1–10 | 1 (0.1%) | 1 (0.3%) | 0 (0%) | | | | |
| Telecommunications | 20 (3.1%) | 14 (4.4%) | 6 (1.9%) | Unknown | 6 | 4 | 2 | | | | |
| Energy | 17 (2.7%) | 7 (2.2%) | 10 (3.1%) | | | | | | | | |
| Real Estate | 12 (1.9%) | 3 (0.9%) | 9 (2.8%) | | | | | | | | |
| Unknown | 155 | 84 | 71 | | | | | | | | |
| Total | N = 795 | N = 403 | N = 392 | Total | N = 795 | N = 403 | N = 392 | Total | N = 795 | N = 403 | N = 392 |

Table 7: Estimation periods for each frequency

| Frequency | Normal Returns Estimation | Event Window |
|-------------------|---------------------------|-------------------------|
| Weekly | 2016-01-01 — 2019-06-30 | 2019-07-01 — yyyy-mm-dd |
| Quarterly (Sales) | | |

NOTE: THINK WHETHER TO INCLUDE NUMBER OF PERIODS OF ESTIMATION, THINK WHETHER THAT SHOULD BE SYMMETRICAL FOR EACH OF THE FREQUENCIES

A.2 Cheap-Talk Index

The models for estimation were retrieved from ...

The table below presents multiple examples of data classified by the researchers in training the model: ...

Reports that contained less than 3 climate related paragraphs were dropped due to having to little data, which was prone to producing outliers.

A.3 Merging

The merging of the data was a difficult process, as company names are often reported differently between different services. Multiple strategies were used to merge the data. Firstly, the data was merged on the names retrieved from the

ResponsibilityReports.com and the longest version of names available in the Refinitiv database. **This resulted in ... observations.** After that, the datasets were merged on the stock tickers conditional on being listed on the same exchange. **This resulted in ... observations.** Finally, the remaining companies were merged using fuzzy matching. **Unfortunately this has left ... observations missing.**

B Models

B.1 Formal assumptions

In order to formally describe the assumptions, I adopt the potential outcomes framework of Rubin, 1974

B.1.1 Parallel Trends (PT)

Formally it is given by:

$$\mathbb{E}[Y_{i,2}(0) - Y_{i,1}(0)|D_i = 1] = \mathbb{E}[Y_{i,2}(0) - Y_{i,1}(0)|D_i = 0] \quad (7)$$

B.1.2 No anticipation

Formally, the assumption is defined as:

$$Y_{i,1}(0) = Y_{i,1}(1) \text{ for all } i \text{ with } D_i = 1 \quad (8)$$