NBER WORKING PAPER SERIES

NOTHING BUT THE TRUTH? PRIVATE INFORMATION AND REPORTING ON CORPORATE SOCIAL RESPONSIBILITY

Jean-Etienne de Bettignies
Hua Fang Liu
David T. Robinson

Working Paper 28159
http://www.nber.org/papers/w28159

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2020, Revised March 2024

An earlier and significantly different version of this paper was circulated under the title "Corporate Social Responsibility and Imperfect Regulatory Oversight: Theory and Evidence from Greenhouse Gas Emissions Disclosures." We are particularly thankful to Heski Bar-Isaac, David P. Baron, Kevin Bryan, Alberto Galasso, Ricard Gil, Mitch Hoffman, Olena Ivus, Tom Lyon, John Maxwell, and Brian Silverman, for their suggestions; and to seminar participants at Baruch College, Harvard Business School, Simon Fraser University, University of Toronto, Western University, the Strategy and the Business Environment Conference (Stanford University, 2019), the Markets, Contracts and Organizations Conference (Canberra, 2023), the Asian Meeting of the Econometric Society (Singapore, 2023), and the Society for Institutional and Organizational Economics (SIOE) Conference (Frankfurt, 2023) for their comments. All remaining errors are ours. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Nothing but the Truth? Private Information and Reporting on Corporate Social Responsibility
Jean-Etienne de Bettignies, Hua Fang Liu, and David T. Robinson
NBER Working Paper No. 28159
December 2020, Revised March 2024
JEL No. D62,J31,M14,M52,Q58

ABSTRACT

We develop and test a model in which firms can make non-verifiable statements about their CSR engagement, and hence may have incentives to mislead markets with exaggerated CSR claims. Firms may privately receive a signal correlated with their CSR engagement - e.g. a measure of greenhouse gas emissions - and have discretion over whether to publicly disclose it. Based on whether a signal is disclosed to them, and on what the signal is, markets form beliefs about the firm's CSR activities and the truthfulness of its claims. The model illustrates the disciplining effect of ex post private signal availability on firms' ex ante reporting on their CSR engagement. We test the model using a difference-in-differences approach that exploits special features of the introduction of the UK Companies Act of 2013, and find evidence supporting the model's predictions.

Jean-Etienne de Bettignies
University of Toronto
Department of Management
Toronto, Ontario, M1C 1A4
CANADA
jean.debettignies@rotman.utoronto.ca

Hua Fang Liu
Baruch College
Zicklin School of Business
New York City, NY 10010
huafang.liu@baruch.cuny.edu

David T. Robinson
Fuqua School of Business
Duke University
100 Fuqua Drive
Durham, NC 27708
and NBER
davidr@duke.edu

1 Introduction

In the sixty years since Milton Friedman (1962) famously advocated for a model of corporate social responsibility (CSR) succinctly captured by Levitt's (1958) quip, "the business of business is profits," there has been a wholesale shift away from shareholder capitalism towards a broader, more inclusive, multi-stakeholder view of the firm.[1] Today, firms purport to care not only about profits; they claim to behave with a broad social and environmental purpose. Indeed, among the 150 largest non-financial U.S. firms by revenue, reporting on wider societal objectives in letters to shareholders has increased from 20% in the 1980s to 90% in 2020 (Rajan *et al.*, 2023).

How accurate are the claims that firms make about their behavior in this multi-stakeholder environment? Measuring profits may be straightforward, but measuring *engagement* in environmental, social and governance (ESG) activities is not. And while some measures of ESG *performance* - e.g. greenhouse gas (GHG) emissions - may be available, this information is usually private to organizations, who have the option - if it is unflattering - of claiming that it was not in fact obtainable. Without verifiable standards of ESG engagement, and ESG performance measures that are private to organizations, complex tradeoffs between alternative stakeholder considerations leave significant room for firms to exaggerate their CSR engagement activities.[2] Indeed, they may describe themselves as multi-stakeholder organizations focused on "growing the pie" (Edmans, 2020) even if their actions do not live up to the standards they espouse. When can firms' CSR claims be trusted if they can exaggerate their engagement and choose the information that they reveal to the market?

To study this question, we develop and test a model in which a firm can be one of two types, either high-CSR engagement or low-CSR engagement. At the beginning of the game, the firm announces its type to the market, but this announcement is unverifiable. For example, the firm could publish a shareholder letter stating that it pursues a broad range of social objectives, and as such maintains a fundamental commitment to social responsibility. Or it could instead simply report that it is a straightforward profit-maximizer. Because the market places a premium on social responsibility, the firm has an incentive to claim that it engages in CSR even when it does not.

After this announcement is made, the firm may - with some probability - receive private infor-

[1] To wit, in 2019 the Business Roundtable went so far as to redefine corporate purpose away from shareholder primacy and toward commitment to all stakeholders: see `https://opportunity.businessroundtable.org`.

[2] And firms do exaggerate. For example, Kim and Lyon (2011) find differences between the greenhouse gas emissions reported by US electric utilities to the US Department of Energy through the Voluntary GHG Registry, and actual emissions calculated using fuel consumption data. Of the 5,296 products sold in large retailer chains in the US and Canada and investigated by TerraChoice, an environmental marketing firm, 95% of the "greener" products were found to make claims that were either false or misleading (TerraChoice, 2010).

mation about its CSR performance, which is correlated with its CSR engagement. For example, the firm may or may not learn the amount of carbon released into the atmosphere through its corporate decisions. This CSR performance measure is really a "signal" that is (perfectly) correlated with the firm's type, and if it chooses to disclose it to the market, the signal is completely verifiable. The market then uses this signal or its non-disclosure to update its beliefs about the firm's type and truth-telling.

The key insight of the model is that the equilibrium of the *ex post* disclosure game disciplines behavior in the *ex ante* "soft information" game. In the disclosure game, a high-CSR firm that receives a (good) signal always discloses it, because it perfectly confirms its high type to stakeholders. In contrast, a low-CSR firm never discloses its (bad) signal when it receives it - lest it reveal its low type to the market;[3] and instead prefers to pool with a high-CSR firm that received no signal. Moving backward, the availability of the *ex post* signal affects the low-CSR firm's *ex ante* incentives to truthfully announce its CSR-engagement type in the soft-information game (in contrast to the high-CSR firm which always tells the truth *ex ante* and is unaffected by the *ex post* game). When signal availability is low, the market imposes low penalties on silence, because it infers that this non-disclosure may well be the result of a high-type firm having not received the signal. Anticipating this, even CSR-disengaged firms have an incentive to make *ex ante* claims that they are CSR-engaged, and to later argue that they have no CSR-performance information to report. But when signal availability is high, the lack of a report is interpreted by the Bayesian market as likely evidence that the firm is CSR-disengaged and has chosen not to disclose its bad signal. The fear of being punished for making exaggerated *ex ante* claims that cannot be substantiated *ex post* creates an incentive for low-CSR firms to report their low CSR-engagement type truthfully *ex ante*.

Of course, the idea that *ex post* punishment disciplines *ex ante* behavior is common to many dynamic games. The novelty here - and the contribution of the model - is the disciplining mechanism itself. Our model shows that in certain equilibrium settings, but not others, market beliefs can discipline *ex ante* behavior even when there is no intrinsic mechanism to force the disclosure of non-verifiable information. The insights this offers about the disciplining effects of *ex post* type-correlated signals have broad application in other settings, both for applied theory and for policy makers, which we discuss in the conclusion of this paper.

We test the model in the context of environmental CSR, where GHG emissions capture the

[3]The firm's concern that disclosing its bad signal - e.g. high GHG emissions - may induce "punishment" by the market is consistent with the recent empirical work of Bolton and Kacperczyk (2021), which shows that worst carbon performers suffer an increase in cost of capital, upon disclosing their emissions.

potentially (un)available private signal correlated with firms' CSR engagement activities. We use the passage of *The UK Companies Act* of 2013, a law in the UK that mandated the reporting of GHG emissions as part of standard annual financial reporting. Although a number of recent papers (see literature review) have used the passage of the law as a shock to mandatory GHG emissions disclosure, our analysis turns on a largely overlooked feature of *The UK Companies Act*. Specifically, the law states that the disclosure requirements apply "only to the extent that it is practical for the company to obtain the information in question [...]" (sub-paragraph (4), Part 7). Thus the policy actually requires firms to *at least try* to obtain information about their emissions, but leaves them with the option not to disclose their GHG emissions by claiming that it is not "practical" for them to obtain this information. As a result of the law, more firms will attempt to measure their GHG emissions, leading to a greater availability of engagement-correlated signals across firms on average. Comparing UK firms with similar firms from 15 other European countries before and after the policy change in a difference-in-differences (DiD) framework, we find that reporting on environmental CSR engagement in the UK increased by a smaller amount after the policy change relative to CSR engagement reported in the other European countries which were not affected by the policy. The negative and significant treatment effect of private signal availability on reported CSR engagement is consistent with the main prediction of our model, and survives a large number of robustness checks.

In order to test whether private signal availability has a differentially larger negative impact on CSR-engagement reporting among low-type firms than among high-type firms, we exploit the fact that in our model high-type firms are much more likely than low-type firms to disclose their private signals - GHG emissions - that may correlated with their true CSR engagement. To that end, we utilize a proprietary database, Carbon Disclosure Project (CDP), to construct a control group composed of (high-type) UK firms which frequently reported their GHG emissions (before the policy change in the UK) and a treatment group composed of (low-type) firms which did not. We use these control and treatment groups in a DiD analysis with *The UK Companies Act* as an exogenous shock to private signal availability; and find that, consistent with our model prediction, private signal availability affects low-type firms in the treatment group more negatively than high-type firms in the control group.

Our paper is related to recent theoretical and empirical work examining *greenwashing*; that is, the corporate practice of hyping its environmental stewardship or social performance in order to boost the firm's image or valuation in the eyes of various market participants. The precise manner in which the concept of greenwashing is operationalized, however, varies in important ways.

The theoretical work of Wu *et al.* (2020) considers greenwashing to be selective investment in publicly observable CSR activities by profit-maximizing firms in an attempt to masquerade as socially responsible firms. In contrast, Lyon and Maxwell (2011) treat greenwashing as the act of revealing environmental successes to the market, but hiding environmental failures. They consider a setting in which a firm may fully disclose, may greenwash, or may disclose no news at all, and examine the disciplinary role that activists can play to curtail greenwashing. They find that while the threat of punishment by activists dissuades firms from engaging in greenwashing, it can come at the expense of fewer firms engaging in full disclosure. This is related to the disclosing of *ex post* signals in our model; but we instead combine elements of *ex post* disclosure games (Grossman and Hart, 1980; Grossman, 1981, Milgrom, 1981) with *ex ante* messaging of soft information about firm type, and show that optimal behavior in the former can impose discipline in the latter.[4]

Our empirical analysis is related to Kim and Lyon (2011), who as mentioned above found differences between actual GHG emissions by US electric utilities and their emissions reported to the US Department of Energy; and to Ramus and Montiel (2005), who investigate four different industry sectors, and find a discrepancy between commitment to, and implementation of, environmental policies. Marquis *et al.* (2016) use a sample of public firms from 45 countries obtained from the Trucost database, to propose a global study of greenwashing measured as "selective disclosure magnitude" - which captures firms' strategic disclosure of "beneficial or relatively benign performance indicators to obscure their less impressive overall record." Their key finding is that firms causing more environmental damages are more exposed to scrutiny and global norms, and hence are less likely to engage in greenwashing. Kim and Yoon (2023) focus on greenwashing in the asset-management industry and find that while signatories of the United Nations Principles for Responsible Investment (PRI) benefit from an inflow of funds, their investments actually do not yield higher fund-level ESG scores.

Our empirical analysis is also related to the recent work of Grewal *et al.* (2022), which uses *The UK Companies Act* as an exogenous shock to mandatory carbon reporting and examines its impact on greenwashing. In Grewal *et al.* (2022), greenwashing is operationalized as selective disclosure à la Marquis *et al.* (2016); they examine the impact of mandatory carbon disclosure on the amount of extraneous environmental information contained in accounting reports. Thus, they

[4]In Lyon and Maxwell's (2011) framework, a favorable disclosure triggers punishment through scrutiny, which incentivizes silence; whereas in our framework, silence triggers belief updating, which can discipline *ex ante* behavior.

consider tradeoffs associated with disclosure of *ex post* signals with different informational content, whereas we focus on the tradeoff between *ex post* management of private information and *ex ante* reporting on CSR engagement.

More generally, *The UK Companies Act* has been the subject of broad empirical interest in recent years. Krüger (2015) interprets the mandatory disclosure policy as an indirect reduction in GHG emissions and investigates its impact on firm value. Grewal (2017), Bolton and Kacperczyk (2021), Downar *et al.* (2021), and Jouvenot and Krüger (2021) examine the impact of mandatory carbon disclosure on GHG emissions and firm performance. These articles, as well as Grewal *et al.* (2022), focus on sub-paragraphs (2) and (3) of Part 7 of the law, which stipulate mandatory reporting of GHG emissions for all publicly traded firms in the UK. We interpret the impact of *The UK Companies Act* differently: we are the first (to our knowledge) to shine a light on the fact that these carbon reporting requirements are only mandatory when it is "practical" for the firm to collect this information. Thus, in our empirical framework, this policy induces firms to try to collect information about their GHG emissions - an increase in the availability of a signal correlated with CSR engagement in the language of our model - but gives firms the option not to report their GHG emissions by simply arguing that they were not able to collect the relevant information.

The remainder of the paper is organized as follows. Section 2 presents the basic elements of the model. Section 3 derives the key theoretical results regarding private signal availability and CSR-engagement reporting in equilibrium. Section 4 discusses *The UK Companies Act* and its implications for our analysis. Section 5 presents the main empirical analysis of the relationship between private signal availability and reporting on CSR engagement, and discusses robustness checks. Section 6 explores additional implications of the model and provides further corroborating evidence based on CDP data and consistent with our model predictions. Section 7 concludes.

2 Model Setup

The main elements of the model are as follows.

Firm Type and CSR Reporting. We consider a firm that can be of one of two CSR-engagement types $T \in \{T_h, T_l\}$: a firm of "high" type T_h engages in pro-social, CSR-related activities; while a firm of "low" type T_l does not. For simplicity we do not model shareholders, managers and workers inside the firm, and abstract from issues related to hidden information or hidden action within the organization, in order to focus instead on the information asymmetry issues between the firm and the market. The firm is of type T_h with exogenous probability q and of type T_l with probability $1 - q$; and maximizes expected profits, conditional on its type.

After privately observing its type T, the firm makes a public announcement $t \in \{t_h, t_l\}$ about it: the firm may make announcement t_h that it is of type T_h, for example by publishing a report about its CSR activities or a shareholder letter outlining its social goals (Rajan *et al.*, 2023); or it may make announcement t_l that it is of type T_l. Statements about the firm's "type" t_j are inherently unverifiable, and hence the firm need not truthfully report its type. Here we have in mind pro-social statements and other non-verifiable information often included in CSR reports and on company websites.

Signals, Observability, and Disclosure. After making its public announcement t, the firm may receive a signal $S \in \{S_h, S_l\}$ that is positively correlated with its type: $P(S_i \mid T_i) > 1/2 > P(S_j \mid T_i)$, with $i \in \{h, l\}$ and $j \neq i$. A natural way to interpret this signal is as a measure of the firm's CSR impact. For example, the firm might learn how much carbon it has released, how much fossil fuel it has consumed, or how many compliance violations its factories has received. Thus, while the firm's type T captures its level of *CSR engagement* (i.e. an input measure), the signal S received can be interpreted as representing the firm's *CSR performance* (i.e. an output measure).

Importantly, this signal is privately observable to the firm, and is received by the firm with probability $\gamma \in (0, 1)$. With probability $1 - \gamma$ the firm receives no signal. Thus, parameter γ captures *private signal availability* or *private performance measure availability* in our model. Because the signal is privately observed by the firm, the market (e.g. consumers, activists, regulators, lawmakers) does not know if the firm even received a signal unless the firm chooses to disclose it, in which case the signal announcement is assumed to be hard, verifiable information. Hence, if the firm receives a signal, it need not reveal that it did, but conditional on disclosing its signal, it cannot lie about which signal it received. It then follows that after observing a signal S_i, $i \in \{h, l\}$, the firm makes one of two possible "signal" announcements: either it makes announcement s_i that it received signal S_i, or it makes announcement s_0 that it did not receive any signal. After observing no signal, the firm's only option is to make announcement s_0.

Without loss of generality, we assume that S_i, $i \in \{h, l\}$ is a perfect signal of the firm's type T_i: $P(S_i \mid T_i) = 1$. This allows us to keep the analysis as simple and clear as possible, and to focus - in the model and in the empirical analysis that follows - on the impact of signal availability γ (rather than signal precision) on the firm's CSR status disclosure.

A firm's GHG emissions are a good example of the type of signal or CSR performance measure we have in mind here, for several reasons. First, GHG emissions are likely to be positively correlated with firms' environmentally-focused CSR activities, in that firms which engage in environmentally-friendly activities are likely to have lower GHG emissions (i.e. higher CSR performance), *ceteris paribus*. Second, the (in)feasibility and cost of obtaining an accurate and publicly verifiable measure of GHG emissions may vary significantly across firms, and indeed for some firms (a fraction $1-\gamma$ in our model), it may be technically infeasible or prohibitively costly to do so. Finally, GHG emissions publicly disclosed by a firm can arguably be verified in a court of law; but if a firm does not make any GHG emissions disclosure, the public cannot verify whether it is because the firm is unable or unwilling to obtain this information, or because it chooses not to disclose it.

Preferences and Profits. The firm's profits depend on how the market "views" it: consumers with a positive view of the firm have a higher willingness to pay for its products, for example; and activists and regulators with a positive view of the firm are less likely to launch a campaign against it, or audit it, respectively. For simplicity, and without loss of generality, we focus here on consumers' valuation of the firm's products, V_{ij}, which depends on the firm's type T_i, $i \in \{h, l\}$, and on its public announcement t_j, $j \in \{h, l\}$. We posit that $V_{hh} > V_{hl} > V_{ll} > V_{lh}$: this rank ordering of preferences captures the idea that consumers, and markets more generally, value pro-social high-type firms more than low-type firms; and that conditional on firm type, they prefer truth-telling over lying. Of course, consumers' willingness to pay (WTP) for the firm's products, V, will also depend on their beliefs about the firm's type, which they do not observe directly.[5]

We contend that firms' profits depend positively on consumers' willingness to pay for their products, and here for simplicity we assume that the firm has zero costs of production, and can extract as profits consumers' entire willingness to pay for its products: $\Pi = V$.

Timing of the Game. The timing of the game can be described as follows:

- Date 0: Nature determines the firm to be of type T_h with probability q and of type T_l with probability $1-q$. The firm privately observes its type. Consumers have prior beliefs $p_0 = q$ and $1 - p_0 = 1 - q$ that the firm is of types h and l, respectively.

[5] For simplicity, we assume the consumers care about firm CSR-engagement types, but not about signals. In doing so, we are attempting to capture the idea that consumers value firms *trying* to be socially responsible, recognizing that impact may be affected by a number of exogenous factors and may be quite noisy. In any case, our conjecture is that allowing consumer preferences to depend on signals as well as types would reduce the analytical tractability of the model but would not alter the main results.

- Date 1: The firm makes type announcement t_j, $j \in \{h, l\}$, that it is of type T_j. Based on the firm's type announcement t_j and on prior belief p_0, consumers form interim beliefs $p_1(t_j, p_0)$ and $1 - p_1(t_j, p_0)$ that the firm is of types T_h and T_l, respectively.

  - If $p_1(t_j, p_0) = 1$ or $p_1(t_j, p_0) = 0$, consumers are certain about the firm's type, and their willingness to pay for the firm's products is $V = V_{hj}$ or $V = V_{lj}$, respectively. The firm extracts as profits $\Pi = V$ the consumers' WTP, and the game ends.
  - If $p_1(t_j, p_0) \in (0, 1)$, the game moves on to date 2.

- Date 2: With probability γ, the T_i-type firm receives signal S_i; and either makes signal announcement s_i that it received signal S_i, or makes announcement s_0 that it received no signal. With probability $1 - \gamma$, the firm receives no signal and makes announcement s_0 that no signal was obtained.

  Based on the firm's signal announcement s_k, $k \in \{i, 0\}$, and on interim belief p_1, consumers form posterior beliefs $p_2(s_k, p_1)$ and $1 - p_2(s_k, p_1)$ that the firm is of types T_h and T_l, respectively.

  Consumers then determine their WTP V for the firm's products, as a function of their posterior beliefs $p_2(s_k, p_1)$ and $1 - p_2(s_k, p_1)$ about the firm's type T_i, the firm's type announcement t_j, and valuations V_{ij}, with $i, j \in \{h, l\}$. Finally, the firm extracts as profits $\Pi = V$ the consumers' WTP, and the game ends.

Total Surplus and the First Best. As a benchmark, we briefly consider the case in which there is no information asymmetry and consumers can observe the firm's type. In this trivial scenario, signal availability γ is irrelevant, and the firm always announces its true type since consumers unambiguously value truth-telling: $V_{hh} > V_{hl}$ and $V_{ll} > V_{lh}$. Thus, the (expected) total surplus generated, from a date 0 point of view, is:

$$TS_{FB} = qV_{hh} + (1 - q) V_{ll}. \tag{1}$$

We will refer back to this benchmark in our main analysis, when we examine the impact of information asymmetry and signal availability on value creation in equilibrium.

3 Theoretical Results

In this section, we derive the perfect Bayesian equilibria of the game. We proceed by backward induction and start by determining the equilibrium in signal announcements at date 2, taking interim beliefs as given. We then move backward and examine the firm's equilibrium type announcements at date 1.

3.1 Equilibrium in Signal Announcements

At date 2, with probability γ the firm obtains a private signal S_i, $i \in \{h, l\}$, that is perfectly correlated with its type T_i; in which case it can make a verifiable and truthful signal announcement $s_i = S_i$ about its type, or it can make announcement s_0 that it did not receive any signal. With probability $1 - \gamma$ the firm receives no signal and cannot but make announcement s_0.

Based on these signal announcements s_k, $k \in \{i, 0\}$, and taking their interim beliefs $p_1(t_j, p_0), 1 - p_1(t_j, p_0) \in (0, 1)$ as their new prior beliefs, consumers determine their posterior beliefs $p_2(s_k, p_1) = P(T_h \mid s_k, p_1)$ and $1 - p_2(s_k, p_1) = P(T_l \mid s_k, p_1)$ about the firm's type. These posterior beliefs, together with the firm's type announcement t_j, $j \in \{h, l\}$, made at date 1, enable consumers to derive their willingness to pay $V(t_j, p_2(s_k, p_1))$ for the firm's products, which coincides with profits $\Pi_i(t_j, p_2(s_k, p_1))$ for the type-T_i firm:

$$\Pi_i(t_j, p_2(s_k, p_1)) = V(t_j, p_2(s_k, p_1)) = p_2(s_k, p_1) V_{hj} + (1 - p_2(s_k, p_1)) V_{lj}. \tag{2}$$

Anticipating profits as expressed in (2), it is easy to show that the firm chooses the following signal announcement strategy in equilibrium:

Lemma 1 *For all interim beliefs $p_1(t_j, p_0) \in (0, 1)$, there exists a unique pure-strategy perfect Bayesian equilibrium in signal announcement, in which:*

- *The type-T_h firm always tells the truth: it makes announcement s_0 when it does not receive a signal and makes announcement s_h when it receives signal S_h;*
- *The type-T_l firm always pretends not to have received a signal: it makes announcement s_0 regardless of whether or not it receives signal S_l.*

Proof. See appendix. ■

The intuition behind this result is simple: for a type-T_h firm, truthfully disclosing its good signal S_h - e.g. low GHG emissions - is useful because it convinces consumers that the firm is indeed a high-type firm, which has a positive effect on their valuation of the firm's products and in turn on the firm's profits.

In contrast, for the type-T_l firm, truthfully disclosing its bad signal S_l - e.g. high GHG emissions - is not optimal because it would eliminate all doubt in consumers' minds that the firm is in fact a low-type firm, which would have a negative effect on their valuation of the firm's products and in turn on the firm's profits. Firm T_l is better off *not* disclosing its signal, to create doubt in consumers' minds that it may in fact be a type-T_h firm that simply did not receive a signal.

Thus in equilibrium, upon reaching date 2, if consumers observe signal announcement s_h they deduce that the firm is of type T_h with probability $p_2(s_h, p_1) = 1$. Taking this posterior belief and the firm's type announcement t_j, $j \in \{h, l\}$ from date 1 into account, we can express their willingness to pay for the firm's products, and hence the firm's profits, as: $V(t_j, 1) = \Pi_h(t_j, 1) = V_{hj}$.

On the other hand, if consumers observe announcement s_0 that no signal was received - an event that could occur either because the firm is of type T_h but did not receive a signal, or because the firm is of type T_l - consumers form posterior belief $p_2(s_0, p_1)$ that the firm is of type T_h:

$$p_2(s_0, p_1) = P(T_h \mid s_0) = \frac{P(s_0 \mid T_h)\, p_1}{P(s_0 \mid T_h)\, p_1 + P(s_0 \mid T_l)(1 - p_1)} = \frac{(1-\gamma)\, p_1}{(1-\gamma)\, p_1 + (1 - p_1)}. \tag{3}$$

Using this posterior belief and the firm's type announcement t_j, $j \in \{h, l\}$ from date 1, we can express consumers willingness to pay for the firm's products, and hence profits for firm T_i, $i \in \{h, l\}$, as:

$$V(t_j, p_2(s_0, p_1)) = \Pi_i(t_j, p_2(s_0, p_1)) = \frac{(1-\gamma)\, p_1}{(1-\gamma)\, p_1 + (1-p_1)} V_{hj} + \frac{(1-p_1)}{(1-\gamma)\, p_1 + (1-p_1)} V_{lj}. \tag{4}$$

Importantly, note that both $p_2(s_0, p_1)$ and $V(t_j, p_2(s_0, p_1))$ are strictly decreasing in γ: the higher the degree of signal availability, the lower the probability that the type-T_h firm fails to receive a signal; and hence the lower the probability that announcement s_0 is coming from that firm, and the lower the willingness to pay for the firm's products.

3.2 Equilibrium in Type Announcements

Moving back to date 1, having observed its type T_i, $i \in \{h, l\}$, the firm makes type announcement t_j, $j \in \{h, l\}$ to maximize its expected profits from a date 1 point of view (henceforth referred to as date 1 profits for simplicity): $E_i(t_j, p_1(t_j, p_0))$. In making this announcement, the firm anticipates its impact on consumers' interim belief $p_1(t_j, p_0)$, and in turn - if the game continues beyond date 1 - on the equilibrium date 2 signal announcements and profits derived above. In order to analyze equilibria in type announcements, we consider separating, pooling, and semi-separating equilibria in turn.

Separating Equilibria. Suppose that firm T_h made announcement t_j with $j \in \{h, l\}$, and that firm T_l made announcement t_z with $z \neq j$. If this were an equilibrium, based on announcements j and z, consumers' interim beliefs would be $p_1(t_j, p_0) = 1$ and $p_1(t_z, p_0) = 0$, respectively, and the game would end at date 1. With these beliefs, firm T_h would obtain payoff $E_h(t_j, 1) = V_{hj}$ and firm T_l would get $E_l(t_z, 0) = V_{lz}$. But since $V_{hj} > V_{lz}$ for all $j, z \in \{h, l\}$, firm T_l would have an incentive to deviate from t_z and make announcement t_j, thus making a separating equilibrium impossible. Thus:

Proposition 1 *There are no separating equilibria in type announcement in this game.*

Proof. Follows directly from above. ∎

Pooling Equilibria. If a pooling equilibrium exists in which both types of firm make the same type announcement t_j, $j \in \{h, l\}$, then based on announcement t_j, consumers form interim belief $p_1(t_j, p_0) = p_0$; and based on announcement t_z, $z \neq j$, they form out-of-equilibrium interim belief η_z that the firm is of type T_h. Using these interim beliefs, we can express the type-T_l firm's date 1 profits from announcements t_j and t_z, respectively:

$$\begin{aligned} E_l(t_j, p_0) &= \Pi_l(t_j, p_2(s_0, p_0)) = p_2(s_0, p_0) V_{hj} + (1 - p_2(s_0, p_0)) V_{lj} \\ &= \frac{(1-\gamma) p_0}{(1-p_0) + (1-\gamma) p_0} V_{hj} + \frac{(1-p_0)}{(1-p_0) + (1-\gamma) p_0} V_{lj} \end{aligned} \tag{5}$$

$$\begin{aligned} E_l(t_z, \eta_z) &= \Pi_l(t_z, p_2(s_0, \eta_z)) = p_2(s_0, \eta_z) V_{hz} + (1 - p_2(s_0, \eta_z)) V_{lz} \\ &= \frac{(1-\gamma) \eta_z}{(1-\eta_z) + (1-\gamma) \eta_z} V_{hz} + \frac{(1-\eta_z)}{(1-\eta_z) + (1-\gamma) \eta_z} V_{lz} \end{aligned} \tag{6}$$

In words, upon making equilibrium type announcement t_j, firm T_l anticipate that it will make signal announcement s_0 at date 2, and that its payoff will depend on consumers' posterior belief $p_2(s_0, p_1)$ as defined in (3), which itself depends on interim belief $p_1 = p_0$. Firm T_l's date 1 profits take a similar form if the firm make out-of-equilibrium announcement t_z, except with out-of-equilibrium interim belief η_z.

Similarly, we can also use these interim beliefs $p_1(t_j, p_0) = p_0$ and η_z to express firm T_h's date 1 profits from announcements t_j and t_z, respectively, as follows:

$$\begin{aligned} E_h(t_j, p_0) &= \gamma\, \Pi_h(t_j, p_2(s_h, p_0)) + (1-\gamma)\,\Pi_h(t_j, p_2(s_0, p_0)) \\ &= \gamma\, V_{hj} + (1-\gamma)\left[p_2(s_0, p_0)\, V_{hj} + (1 - p_2(s_0, p_0))\, V_{lj}\right] \\ &= \gamma\, V_{hj} + (1-\gamma)\left[\frac{(1-\gamma)\,p_0}{(1-p_0)+(1-\gamma)\,p_0} V_{hj} + \frac{(1-p_0)}{(1-p_0)+(1-\gamma)\,p_0} V_{lj}\right] \end{aligned} \tag{7}$$

$$\begin{aligned} E_h(t_z, \eta_z) &= \gamma\, \Pi_h(t_z, p_2(s_h, \eta_z)) + (1-\gamma)\,\Pi_h(t_z, p_2(s_0, \eta_z)) \\ &= \gamma\, V_{hz} + (1-\gamma)\left[p_2(s_0, \eta_z)\, V_{hz} + (1 - p_2(s_0, \eta_z))\, V_{lz}\right] \\ &= \gamma\, V_{hz} + (1-\gamma)\left[\frac{(1-\gamma)\,\eta_z}{(1-\eta_z)+(1-\gamma)\,\eta_z} V_{hz} + \frac{(1-\eta_z)}{(1-\eta_z)+(1-\gamma)\,\eta_z} V_{lz}\right] \end{aligned} \tag{8}$$

Upon making equilibrium type announcement t_j (resp. t_z), firm T_h anticipate that with probability γ it will receive signal S_h at date 2. In this case, as discussed in Section 3.1, the firm will make signal announcement s_h, thus fully revealing its type and obtaining profits V_{hj} (resp. V_{hz}). With probability $1-\gamma$, firm T_h will receive no signal and will have to make announcement s_0, in which case its payoff will depend on consumers posterior belief $p_2(s_0, p_1)$, which itself depends on interim belief $p_1 = p_0$ (resp. η_z).

Pooling on type announcement t_j, $j \in \{h, l\}$ is an equilibrium if there exists an out-of-equilibrium belief $\eta_z \in [0,1]$ such that $E_l(t_j, p_0) \geq E_l(t_z, \eta_z)$ and $E_h(t_j, p_0) \geq E_h(t_z, \eta_z)$. Since, as illustrated (6) and (8), out-of-equilibrium payoffs $E_l(t_z, \eta_z)$ and $E_h(t_z, \eta_z)$ are strictly increasing in out-of-equilibrium belief η_z, necessary and sufficient conditions for the existence of a pooling equilibrium on t_j are simply $E_l(t_j, p_0) \geq E_l(t_z, 0)$ and $E_h(t_j, p_0) \geq E_h(t_z, 0)$ (since in that case an equilibrium would exist at least for $\eta_z = 0$). Using these conditions and the above payoffs, one can readily verify that:

Proposition 2 *A threshold private signal availability level $\gamma^* = 1 - \frac{1-p_0}{p_0}\frac{V_{ll}-V_{lh}}{V_{hh}-V_{ll}}$ exists such that:*

- *At low levels of signal availability $\gamma \leq \gamma^*$, a unique set of pooling equilibria exists and survives the D_1 criterion: equilibria in which both types of firm, T_h and T_l, make announcement t_h that they are of type T_h at date 1.*

- *At high levels of signal availability $\gamma > \gamma^*$, there are no pooling equilibria in this game.*

Proof. See appendix. ■

To understand the intuition behind this proposition, two questions must be addressed. First, why can we have pooling on t_h but not pooling on t_l? And second, why is pooling on t_h an equilibrium only at low degrees of signal availability? We address each question in turn.

Pooling on t_h versus pooling on t_l. As expressions (5), (6), (7), and (8) illustrate, a key difference between the two types of firm is that the type-T_h firm has a stronger incentive than the type-T_l firm to make announcement t_h, *ceteris paribus*. This is because it anticipates that with probability γ it will receive signal S_h and that by making signal announcement s_h it will be able to credibly reveal its type to consumers. At that point having made truthful type announcement t_h at date 1 will yield a higher valuation V_{hh} than the valuation V_{hl} associated with having made untrue type announcement t_l.

As a result of this, even if a pooling on t_l equilibrium did exist, firm T_h would have a *stronger* incentive than firm T_l to deviate to t_h, and indeed the set of out-of-equilibrium beliefs that would make deviation to t_h strictly preferable to equilibrium choice t_l would be greater for firm T_h than for firm T_l. Hence under the D_1 criterion (Banks and Sobel, 1987), upon observing out-of-equilibrium type announcement t_h, consumers would assign belief $\eta_h = 1$ to the possibility of facing a type-T_h firm, which would lead both to deviate, thus precluding the survival of pooling on t_l as an equilibrium.

Conversely, in a pooling on t_h equilibrium, firm T_h would have a *weaker* incentive than firm T_l to deviate to t_l, and under the D_1 criterion, upon observing out-of-equilibrium type announcement t_l, consumers would assign belief $\eta_h = 0$ to the possibility of facing a type-T_h firm. Neither firm would have an incentive to deviate, thus ensuring the survival of pooling on t_h as an equilibrium.

Impact of signal availability γ on existence of pooling equilibrium. Recall from above that necessary and sufficient conditions for the existence of a pooling equilibrium on t_h are simply that equilibrium payoffs for firms T_l and T_h be greater than their *minimum* out-of-equilibrium payoffs: $E_l(t_h, p_0) \geq E_l(t_l, 0)$ and $E_h(t_h, p_0) \geq E_h(t_l, 0)$. An increase in signal availability γ has no impact on minimum out-of-equilibrium payoffs: $E_l(t_l, 0)$ and $E_h(t_l, 0)$. But it does unambiguously decrease equilibrium payoffs $E_l(t_h, p_0)$ and $E_h(t_h, p_0)$, because it reduces consumers' posterior belief $p_2(s_0, p_0)$, upon observing a no-signal announcement s_0 at date 2, that the firm is of type T_h, and hence their willingness to pay for its products.

At low levels of signal availability $\gamma \leq \gamma^*$, consumers' posterior belief $p_2(s_0, p_0)$, upon observing a no-signal announcement s_0 at date 2, and in turn equilibrium payoffs, are sufficiently high for the pooling on t_h equilibrium to hold. At high levels of signal availability $\gamma > \gamma^*$, this is not the case and the equilibrium cannot be sustained.

Semi-Separating Equilibria. Suppose that firm T_l makes type announcement t_j, $j \in \{h, l\}$ while firm T_h randomizes between announcements, announcing t_j with probability $m \in (0,1)$ and announcing t_z, $z \neq j$, with probability $1 - m$. Then, upon observing type announcement t_j, consumers form interim belief

$$p_1(t_j, p_0) = \frac{P(t_j \mid T_h)\, p_0}{P(t_j \mid T_h)\, p_0 + P(t_j \mid T_l)(1 - p_0)} = \frac{m p_0}{m p_0 + (1 - p_0)}, \tag{9}$$

that the firm is of type T_h; while upon observing t_z, $z \neq j$, they form interim belief $p_1(t_z, p_0) = 1$. Based on these beliefs, firm T_l anticipates that its profits from a date 1 point of view will be

$$\begin{aligned}
E_l(t_j, p_1(t_j, p_0)) &= \Pi_l(t_j, p_2(s_0, p_1(t_j, p_0))) \\
&= p_2(s_0, p_1(t_j, p_0))\, V_{hj} + (1 - p_2(s_0, p_1(t_j, p_0)))\, V_{lj} \qquad (10) \\
&= \frac{(1-\gamma)\, p_1(t_j, p_0)}{(1 - p_1(t_j, p_0)) + (1-\gamma)\, p_1(t_j, p_0)} V_{hj} + \frac{(1 - p_1(t_j, p_0))}{(1 - p_1(t_j, p_0)) + (1-\gamma)\, p_1(t_j, p_0)} V_{lj} \\
&= \frac{(1-\gamma)\, m p_0}{(1 - p_0) + (1-\gamma)\, m p_0} V_{hj} + \frac{1 - p_0}{(1 - p_0) + (1-\gamma)\, m p_0} V_{lj}
\end{aligned}$$

if it plays equilibrium strategy t_j; and $E_l(t_z, 1) = V_{hz}$ if it plays strategy t_z. Note that $E_l(t_j, p_1(t_j, p_0)) \in (V_{lj}, V_{hj})$. Similarly, firm T_h anticipates date 1 payoff

$$\begin{aligned}
E_h(t_j, p_1(t_j, p_0)) &= \gamma\, \Pi_h(t_j, p_2(s_h, p_1(t_j, p_0))) + (1-\gamma)\, \Pi_h(t_j, p_2(s_0, p_1(t_j, p_0))) \\
&= \gamma\, V_{hj} + (1-\gamma)\left[p_2(s_0, p_1(t_j, p_0))\, V_{hj} + (1 - p_2(s_0, p_1(t_j, p_0)))\, V_{lj}\right] \qquad (11) \\
&= \gamma\, V_{hj} + (1-\gamma)\, E_l(t_j, p_1(t_j, p_0)) \\
&= \gamma\, V_{hj} + (1-\gamma)\left[\frac{(1-\gamma)\, m p_0}{(1 - p_0) + (1-\gamma)\, m p_0} V_{hj} + \frac{1 - p_0}{(1 - p_0) + (1-\gamma)\, m p_0} V_{lj}\right]
\end{aligned}$$

if it plays equilibrium strategy t_j; and $E_h(t_z, 1) = V_{hz}$ if it plays strategy t_z. Furthermore, since firm T_h randomizes between the two strategies, the equilibrium value m^h_{jz} of mixing probability m must be such that it is indifferent between the two strategies, i.e. such that $E_h(t_j, p_1(t_j, p_0)) = E_h(t_z, 1) = V_{hz}$.

Now suppose that firm T_h makes type announcement t_j, $j \in \{h, l\}$ while firm T_l randomizes between announcements, announcing t_j with probability $m \in (0,1)$ and announcing t_z, $z \neq j$, with probability $1-m$. Then, upon observing type announcement t_j, consumers form interim belief

$$p_1(t_j, p_0) = \frac{P(t_j \mid T_h)\, p_0}{P(t_j \mid T_h)\, p_0 + P(t_j \mid T_l)(1-p_0)} = \frac{p_0}{p_0 + m(1-p_0)}, \tag{12}$$

that the firm is of type T_h; while upon observing t_z, $z \neq j$, they form interim belief $p_1(t_z, p_0) = 0$. Based on these beliefs, firm T_l anticipates that its profits from a date 1 point of view will be

$$\begin{aligned}
E_l(t_j, p_1(t_j, p_0)) &= \Pi_l(t_j, p_2(s_0, p_1(t_j, p_0))) \\
&= p_2(s_0, p_1(t_j, p_0))\, V_{hj} + (1 - p_2(s_0, p_1(t_j, p_0)))\, V_{lj} \qquad (13)\\
&= \frac{(1-\gamma)\, p_1(t_j, p_0)}{(1 - p_1(t_j, p_0)) + (1-\gamma)\, p_1(t_j, p_0)} V_{hj} + \frac{(1 - p_1(t_j, p_0))}{(1 - p_1(t_j, p_0)) + (1-\gamma)\, p_1(t_j, p_0)} V_{lj} \\
&= \frac{(1-\gamma)\, p_0}{m(1-p_0) + (1-\gamma)\, p_0} V_{hj} + \frac{m(1-p_0)}{m(1-p_0) + (1-\gamma)\, p_0} V_{lj}
\end{aligned}$$

if it plays equilibrium strategy t_j; and $E_l(t_z, 0) = V_{lz}$ if it plays strategy t_z. Note that $E_l(t_j, p_1(t_j, p_0)) \in (V_{lj}, V_{hj})$. Furthermore, since firm T_l randomizes between the two strategies, the equilibrium value m^l_{jz} of mixing probability m must be such that it is indifferent between the two strategies, i.e. such that $E_l(t_j, p_1(t_j, p_0)) = E_l(t_z, 0) = V_{lz}$. Similarly, firm T_h anticipates date 1 payoff

$$\begin{aligned}
E_h(t_j, p_1(t_j, p_0)) &= \gamma\, \Pi_h(t_j, p_2(s_h, p_1(t_j, p_0))) + (1-\gamma)\, \Pi_h(t_j, p_2(s_0, p_1(t_j, p_0))) \\
&= \gamma\, V_{hj} + (1-\gamma)\left[p_2(s_0, p_1(t_j, p_0))\, V_{hj} + (1 - p_2(s_0, p_1(t_j, p_0)))\, V_{lj}\right] \qquad (14)\\
&= \gamma\, V_{hj} + (1-\gamma)\, E_l(t_j, p_1(t_j, p_0)) \\
&= \gamma\, V_{hj} + (1-\gamma)\left[\frac{(1-\gamma)\, p_0}{m(1-p_0) + (1-\gamma)\, p_0} V_{hj} + \frac{m(1-p_0)}{m(1-p_0) + (1-\gamma)\, p_0} V_{lj}\right]
\end{aligned}$$

if it plays equilibrium strategy t_j; and $E_h(t_z, 0) = V_{lz}$ if it plays strategy t_z.

Based on these payoffs, we derive the following results:

Proposition 3 *A threshold private signal observability level* $\gamma^* = 1 - \frac{1-p_0}{p_0} \frac{V_{ll} - V_{lh}}{V_{hh} - V_{ll}}$ *exists such that:*

- *At low levels of signal availability* $\gamma \leq \gamma^*$*, no semi-separating equilibria exist in this game.*
- *At high levels of signal availability* $\gamma > \gamma^*$*, there exists a unique semi-separating equilibrium in which firm* T_h *always truthfully reports its type, and firm* T_l *randomizes between truthfully reporting its type and pretending to be of type* T_h*. Moreover, the higher the degree of signal availability* γ*, the greater the probability of truthful reporting by firm* T_l*.*

Proof. See appendix. ■

To understand the intuition behind this proposition, again two questions must be addressed. First, why is there only one type of semi-separating equilibrium? Second, why does this equilibrium only exist at high degrees of signal availability? We address each question in turn.

Different types of semi-separating equilibria. In order to have a semi-separating equilibrium, two conditions must be met: 1) there must exist mixing probabilities $m, (1-m) \in (0,1)$ such that the randomizing firm is indifferent between the pooling strategy which obfuscates its type and the separating strategy which perfectly reveals it; and 2) given mixing probabilities m and $(1-m)$, the pooling firm must prefer the pooling strategy over the separating strategy. It is easy to see that 3 of the 4 possible semi-separating scenarios fail to satisfy conditions 1) and 2).

First, suppose that the type-T_l firm makes type announcement t_l while the type-T_h firm randomizes. Clearly, in making type announcement t_h the type-T_h firm would perfectly reveal its type and tell the truth in doing so, thus obtaining profits $E_h(t_h, 1) = V_{hh}$. These profits are strictly superior to the profits $E_h(t_l, p_1(t_l, p_0))$ from pooling with the type-T_l firm, regardless of $m, (1-m) \in (0,1)$, and hence the type-T_h cannot be indifferent between the two strategies: condition 1) is not met and this scenario cannot be an equilibrium.

Conversely, suppose that the type-T_h firm makes type announcement t_l while the type-T_l firm randomizes. Clearly, in making type announcement t_h, firm T_l would perfectly reveal its type even though it would be untruthful in its announcement, thus obtaining profits $E_l(t_h, 0) = V_{lh}$. These profits are strictly inferior to the profits $E_l(t_l, p_1(t_l, p_0))$ from pooling with the type-T_h firm, regardless of $m, (1-m) \in (0,1)$, and hence the type-T_l cannot be indifferent between the two strategies: this outcome cannot be an equilibrium either.

Now consider the case in which the type-T_l firm makes type announcement t_h while the type-T_h firm randomizes. A key difference between the two types of firm here is that the type-T_h firm has a stronger incentive than the type-T_l firm to make announcement t_h, *ceteris paribus.* The type-T_h firm anticipates that - in pooling on t_h - with probability γ it will receive signal S_h and will be able to credibly reveal its type to consumers (with signal announcement s_h). At that point having made truthful type announcement t_h at date 1 will yield a higher valuation V_{hh} than the valuation V_{hl} associated with having made untrue announcement t_l. Thus, if firm T_h is indifferent between announcements t_h and t_l (condition 2)), then firm T_l must strictly prefer announcement t_l, which is inconsistent with the semi-separating equilibrium under consideration.

Impact of private signal availability γ on existence of semi-separating equilibrium. Finally, consider the scenario in which the type-T_h firm makes truthful type announcement t_h; while the type-T_l firm randomizes between truthfully revealing its type and obtaining profits $E_l(t_l, 0) = V_{ll}$, and obfuscating the truth by pooling with the type-T_h firm. In this latter case, the type-T_l firm's payoff depends on consumers date 2 posterior beliefs that it is a high-type firm that simply did not receive a signal, or a low-type firm that lied about its type. When private signal availability is low ($\gamma \leq \gamma^*$), consumers are quite willing to believe that the firm is of high-type following no signal at date 2, and anticipating this the low-type firm may strictly prefer to pool with the high-type firm regardless of $m, (1-m) \in (0, 1)$. In this case this semi-separating equilibrium cannot hold.

But when the degree of private signal availability is high ($\gamma > \gamma^*$), consumers are likely to believe, following no signal at date 2, that the firm is of low-type and lied. Anticipating this the low-type firm will be willing to randomize between truth telling and pooling with the high-type firm. As γ increases, the payoff from pooling diminishes, *ceteris paribus*, and the low type firm responds by telling the truth more frequently.

We illustrate these results in Figure 1, which depicts the average CSR claim as a function of signal availability γ.

[Insert Figure 1 here.]

There are three curves of interest in Figure 1. First, the long-dotted curve captures $\hat{t}_h$, the average CSR claim for the type-T_h firm. This is a flat line at t_h because the type-T_h firm always truthfully reports its type, regardless of signal availability γ. Second, the short-dotted curve captures $\hat{t}_l$, the average CSR claim for the type-T_l firm. At low levels of signal availability $\gamma < \gamma^*$, the type-T_l firm announces t_h in an attempt to masquerade as a type-T_h firm, subsequently claiming not to have received a signal. As signal availability γ increases beyond γ^*, the expected cost of doing so becomes too large, and the type-T_l firm begins to randomize between announcing t_h and pooling with the type-T_h firm, and telling the truth by announcing t_l. As γ increases further, the type-T_l firm tells the truth more and more often, and hence the average CSR claims of the type-T_l firm decrease with signal availability γ; ending with the firm telling the truth with certainty when $\gamma = 1$. Finally, the solid curve represents the average CSR claims from a date-0 point of view, before the firm's type is realized. Since the probabilities of the firm being of types T_h and T_l are q and $(1-q)$, respectively, the date-0 average CSR claims can be expressed as $\hat{t} = q\hat{t}_h + (1-q)\,\hat{t}_l$. These date-0 average CSR claims gradually decrease with signal availability γ, at the type-T_l firm evolves from always lying about its type, to always telling the truth.

3.3 Testable Implications

The model delivers three main testable implications. The main insight of the model, which follows from Propositions 1-3, is that while low-type firms - which do not engage in CSR - may be tempted to make false, unverifiable claims about their CSR engagement, the mere availability of an *ex post* private but disclosable signal correlated with their CSR engagement - e.g. a CSR performance measure - may be sufficient to discipline them into *ex ante* truth-telling about their CSR engagement. High-type firms, on the other hand, do not need the discipline. They are unaffected by private signal availability and always tell the truth about their CSR engagement.

The prediction is then that at low levels of private (engagement-correlated) signal availability, high-type firms truthfully report their level of CSR engagement but low-type firms exaggerate their CSR engagement to masquerade as high-type firms. Then, as signal availability increases, high-type firms continue to truthfully report their CSR engagement while low-type firms gradually begin to report their true lower levels of CSR engagement more and more. Thus, the availability of private signals correlated with CSR engagement - such as CSR performance measures - is predicted to have a negative impact on reported CSR engagement. This result is illustrated by curve $\hat{t}$ in Figure 1.

In the empirical analysis presented below, we test this prediction in the context of *environmental CSR*, where *GHG emissions* represent a possibly unavailable private *measure* of CSR *performance* and private *signal* of CSR *engagement*: they are likely to be correlated with a firm's engagement in environmental activities; may or may not be available to firms or readily estimated by them; and when available are privately observed. We make use of the *UK Companies Act* of 2013; a policy change which we describe in detail in the next section and interpret as an exogenous increase in the availability of GHG emissions measures on average across firms, i.e. as an increase in signal availability γ. The fact that the policy change was implemented in the UK, but not in 15 other European countries with companies listed on local stock exchanges, facilitates a difference-in-differences (DiD) analysis of UK firms vs. non-UK firms, as we describe in the following pages.

Two additions empirical predictions emerge from our model. First, the reported CSR engagement by high-type firms (which are genuinely engaged in CSR) should be higher on average than reported CSR engagement by low-type firms (which are not engaged in CSR). Second, the availability of private CSR-engagement signals should not affect reporting on CSR engagement by high-type firms, as they always truthfully report their CSR engagement, but should decrease CSR reporting by low-type firms, as they gradually reduce exaggerated claims about their CSR engagement.

To test these additional predictions, we use data from the UK Carbon Disclosure Project (CDP) and proxy for firm type by examining the degree of voluntary disclosure before *The UK Companies Act*. As discussed in Section 6, CDP was a voluntary carbon disclosure effort that already operated in the UK before the policy change. High-type firms are more likely to disclose their CSR performance (if they are able to measure it), because this disclosure will likely reveal to consumers that they are high-type firms; while in contrast, low-type firms would prefer not to disclose their CSR performance (even if they could measure it), because doing so would likely reveal to consumers that they are in fact a low-type firm. Indeed, our model shows that in equilibrium, consumers - and us as data analysts here - infer that the firm is more likely to be high-type following disclosure of private engagement-correlated signals, and more likely to be low type following non-disclosure.

4 The UK Companies Act of 2013

In July 2013, the UK government passed *The Companies Act 2006 (Strategic Report and Director's Report) Regulation 2013* (*The UK Companies Act*), which required all listed UK firms to report their GHG emissions in their annual reports. Two years before that, the Department of Environment, Food and Rural Affairs (DEFRA) conducted a public consultation of about two thousand stakeholders to seek views about the introduction of mandatory GHG reporting regulation; and an impact assessment that included the results of this public consultation, cost/benefit analyses, and DEFRA's recommendation of mandatory reporting for all publicly traded companies, was made available to members of parliament and policy makers from August 31, 2011. Furthermore, some of the stakeholders who had participated in the public consultation - and who were predominantly in favor of mandatory reporting - made their responses public during the summer of 2011 (Krüger, 2015). Thus, although the bill itself was made publicly available on July 25, 2012, was approved by the UK House of Commons on July 16, 2013, and came into effect on October 1, 2013, the public, organizations and other stakeholders could have anticipated its effects as early as summer 2011.

Sub-paragraphs (2) and (3) of Part 7 of *The UK Companies Act* require that all quoted companies state, in their yearly director's report, "the annual quantity of emissions in tonnes of carbon dioxide equivalent from activities for which the company is responsible including (a) the combustion of fuel; and (b) the operation of any facility;" as well as "the annual quantity of emissions in tonnes of carbon dioxide equivalent resulting from the purchase of electricity, heat, steam or cooling by the company for its own use." The law further requires that the report state "the methodologies used to calculate the information disclosed;" and "at least one ratio which expresses the quoted company's annual emissions in relation to a quantifiable factor associated with the company's activities." Thus, the passage of the bill has been interpreted as an exogenous shock to carbon disclosure requirements (Grewal 2017; Downar *et al.* 2021; Jouvenot and Krüger 2021; Grewal *et al.* 2022).

Important for our purpose, however, is sub-paragraph (4) of Part 7 of the Act, which stipulates that the disclosure requirements apply "only to the extent that it is practical for the company to obtain the information in question; but where it is not practical for the company to obtain some or all of that information, the report must state what information is not included and why."

Thus, the policy requires public firms to *attempt* to obtain information on their GHG emissions, which will lead to this information becoming *available to more firms on average*; however, the policy allows for selective non-disclosure. Whether or not firms do in fact have information about their GHG emissions, they retain the option not to disclose it by claiming that it is not "practical" for them to obtain this information. In fact, as of 2020, i.e. seven years after the passage of the bill, only 75% of the 150 largest firms (by market capitalization) of the London Stock Exchange (LSE) were clearly disclosing their GHG emissions in their financial reports (ClientEarth, 2020). This percentage is likely much lower among the 1100 or so smaller, less visible firms quoted on the LSE.

Thus, building on the fact that, as discussed in Section 3.3, GHG emissions can be interpreted as a private measure of CSR performance and a signal of CSR engagement as discussed in our model (specifically, of CSR engagement in environmental activities); *The UK Companies Act* can then be thought of as an exogenous shock to the *availability* γ of this privately observed but disclosable signal of CSR engagement among publicly traded firms in the UK.

We conclude this section with a final comment about *The UK Companies Act*: what if we interpret it as in prior work on the subject, as a "forced" increase in disclosure of GHG emissions by firms? It is important to underline that such an interpretation is still perfectly consistent with our model. Indeed, eliminating firms' discretion over their GHG emissions disclosure and forcing them to make these emissions public is equivalent to a special case in our model where γ is exogenously increased to 1. In this extreme scenario, all firms disclose their type (assuming a nominal penalty $\epsilon \to 0$ for non-disclosure), markets can perfectly infer firm type at the signal disclosure stage, and anticipating perfect *ex post* type revelation prompts firms to tell the truth about their CSR engagement *ex ante*. Importantly, our analysis in this paper offers a markedly richer environment in which to study *puffery* in CSR engagement, as it does allow - both theoretically and empirically - for discretion over signal disclosure.

5 Evidence from UK Firms vs. EU Firms

5.1 Methodology and Data

Our DiD approach draws from Krüger (2015), and as mentioned above turns on the fact that *The UK Companies Act* was implemented in the UK but not in continental Europe. Our control group is a set of firms from other European stock exchanges, in countries that include Ireland, Belgium, Denmark, France, Germany, Italy, Luxembourg, Netherlands, Norway, Sweden, Switzerland, Austria, Portugal, Spain, and Finland.[6] We use 2011 as our UK policy intervention year, to account for the fact that although *The UK Companies Act* was officially passed in 2013 it was made publicly available and circulated as of August 2011.[7] We specify the following DiD regressions:

$$\begin{aligned} Env\ Reporting_{i,t} &= \beta_0 + \beta_1 \cdot Treat_i + \beta_2 \cdot After_t + \beta_3 \cdot Treat_i \times After_t \\ &\quad + \gamma' \cdot Control_{i,t} + \delta_t + \varphi_i + \varepsilon_{i,t}, \end{aligned} \tag{15}$$

where $Env\ Reporting_{i,t}$ denotes firm i's "report-based" environmental score in year t: this score is based on self-reported information by each firm, and hence captures firms' reported CSR engagement in environmental activities - which corresponds to the firm's type announcement $t \in \{t_h, t_l\}$ in our model. The right-hand side of expression (15) includes $Treat_i$, our treatment dummy variable equal to one for UK (treated) firms, and to zero for non-UK (untreated) firms; $After_t$, our intervention dummy variable equal to one for years in the post-intervention period (including and after 2011), and zero for years in the pre-intervention period (before 2011); $Control_{i,t}$, a set of control variables that may affect firms' reporting on environmental initiatives; as well as year and firm fixed effects and error term $\varepsilon_{i,t}$. Our coefficient of interest is β_3, which captures the average treatment effect of private CSR-engagement signal availability on report-based environmental CSR-engagement scores of treated (UK) firms.

Reported CSR engagement in environmental activities. To construct our measure of reporting on environmental CSR engagement for firms from the UK and other European countries, we use data from the Refinitiv ESG database over the 2005-2015 time period. Refinitiv ESG provides ratings on environmental, social and governance (ESG) performance for over 7,000 public firms around the globe, with panel data going back to 2002. Refinitiv ESG first selects and groups

[6]In constructing our control group, we restrict our attention to other European firms and do not consider firms from other regions around the world. Our argument here is that the UK and other European countries are more likely to have similar unobservable features such as culture and institutions, which may affect firms' CSR reporting.

[7]We also use 2012 as the intervention year in the robustness checks, and we obtain similar results.

more than 400 ESG measures into 10 categories, and then aggregates them into 1) a total ESG score, and 2) a pillar score for each of the three ESG dimensions - environmental, social, and corporate governance - for each firm in each year in the sample. These scores can range from 0 to 100 and can be interpreted as percentage scores. Unlike other ESG databases which tend to use equal-weighted ratings, Refinitiv ESG adopts data-driven category weights and applies percentile rank scoring methodology to ensure that the ESG scores are comparable across companies and industries.

Refinitiv ESG's environmental score is aggregated from 137 indicators associated with 3 of the 10 categories alluded to above: an environmental-innovation category, a resource-use category, and an emissions-reduction category. This environmental score exhibits two noteworthy characteristics for our purpose. First, and most importantly, this environmental score is based on *self-reported* information from firms - collected from CSR reports, CSR sections in annual reports, company websites, and stock exchange filings, for example - rather than on actual, objective performance measures. Indeed, Refinitiv ESG is very clear about the fact that "company disclosure is at the core of [their] methodology"(Refinitiv, 2022).

Second, the information conveyed by this score is largely *unverifiable* by third-party users of the data. Indeed, the construction of the environmental score from the 137 indications mentioned above would make it very difficult for consumers or other stakeholders to "reverse-engineer" a firm's environmental claims from its environmental score, let alone verify the validity of these claims. Furthermore, even if reverse-engineering were possible, the indicators seem largely based on claims - e.g. activities, initiatives, investments aimed at reducing environmental footprint - that may be difficult to quantify and in turn verify.[8]

The self-reported and non-verifiable nature of the information associated with this environmental score make it a compelling measure of firms' reported CSR engagement in environmental activities - itself representing firms' *ex ante* non-verifiable type announcements examined in our

[8]The emissions-reduction category appears to include more quantifiable measures than the environmental-innovation category and the resource-use category. Indeed the emissions-reduction category comprises indicators on GHG emissions, which in our model is private but *ex post* disclosable information that can be verified if disclosed, and is different from the *ex ante* non-verifiable CSR engagement reporting that we are attempting to capture with this environmental score. We do not expect the presence of these GHG emissions indicators to affect our empirical results, because 1) as mentioned above it would be difficult to reverse-engineer and reliably verify GHG emissions information from the environmental score, and 2) only a very small fraction of firms in the sample report their GHG emissions anyway. Nevertheless, as a robustness check we did re-run our main regression separately for each environmental score category. We find very similar results for the environmental-innovation and resource-use categories, and slightly weaker results for the emissions-reduction category. This is in line with our model: the more quantifiable and verifiable the reported information is, the less firms may be tempted to exaggerate their claims, regardless of the availability of *ex post* CSR performance signals.

model. Thus, we use the Refinitiv ESG environmental score as the main dependent variable in our empirical analysis.

Control Variables. Using Compustat Global Annual Files for both UK firms and firms from the other 15 European countries listed above, we construct control variables to capture firm- and industry-level characteristics which may affect firms' decision to report their environmental initiatives. These include *Firm Size* measured by the natural logarithm of one plus a firm's total assets; *Profitability* measured by return on assets (*ROA*); *Leverage* measured by total debt to total assets ratio; *Sales* measured by the natural logarithm of one plus the firm's total sales; *R&D Intensity* measured by firms' R&D expenditure scaled by total assets; and *Market Competition* based on a Herfindahl index computed from annual sales. We provide detailed definition for all variables in Table 1. All variables are winsorized at the 1st and 99th levels to minimize the effect of outliers.

[Insert Table 1 here.]

Summary Statistics. We merge the Refinitiv ESG data with the Compustat Global data to form our final sample. In Table 2, we present summary statistics for our variables in Panel A; and the distribution of observations by industry in Panel B, within the treatment group (UK firms) and the control group (other European firms). We note from Panel A, that statistics for covariates are similar across treatment and control groups; and Panel B shows that the top four industries in which firms operate are the same in the two groups: manufacturing, finance, transportation, and service. Correlations are reported in the Online Appendix.

[Insert Table 2 here.]

5.2 Baseline Difference-in-Differences Estimation

In Figure 2 we plot average report-based environmental CSR scores[9] for firms in the treatment group and the control group, for each year in our sample period (2005 - 2015).

[Insert Figure 2 here.]

This figure illustrates two main points. First, it shows that report-based environmental CSR scores in the treatment group evolve in parallel to the corresponding scores in the control before our UK

[9] For convenience, throughout the paper we use "report-based environmental CSR scores," "CSR reporting on environmental initiatives," and "environmental reporting" interchangeably.

policy intervention year of 2011, suggesting - consistent with DiD's "parallel trends" identifying assumption - that they would have continued to evolve similarly in the post-treatment period *if no treatment had occurred* (Heckman *et al.* 1997; Abadie 2005).

Second, Figure 2 also shows that report-based CSR environmental scores in the treatment and control groups begin to diverge from 2011 onward, with the difference between average scores in the treatment and control groups becoming more negative post-intervention. This is consistent with the UK policy intervention and its associated increase in private CSR-engagement signal availability having a negative treatment effect on report-based environmental CSR scores in UK firms, a result to which we turn more formally now.

The results of our baseline DiD regression specified in expression (15) are reported below in Table 3, where in all specifications standard errors are adjusted for within-firm clustering.

[Insert Table 3 here.]

Column (1) reports the results of the simplest DiD specification, with no additional control variables and no fixed effects. In column (2), we introduce firm-level and industry-level controls: *Firm Size*, *ROA*, *Leverage*, *Sales*, *R&D Intensity*, and *Market Competition*. In columns (3), (4), and (5) we sequentially add year fixed effects, industry fixed effects, and country fixed effects. Finally, in column (6) we add firm fixed effects.

Our main coefficient estimate of interest in these regressions is the coefficient on $Treat \times After$, as it captures the average treatment effect of private signal availability on report-based environmental CSR scores of treated firms. This coefficient estimate is negative and statistically significant at the 1% significance level in all six specifications. Moreover, the value of the coefficient estimate remains fairly constant around -5 in all but one specifications (in column (6), where firm fixed effects are introduced, its value is around -3.2).

Taken together these results indicate that the *The UK Companies Act* had a strong causal impact on CSR reporting. The increase in private CSR-engagement signal availability on UK firms' reporting on their environmental CSR engagement leads to a 5 percentage point decrease in their report-based environmental CSR scores, compared to what their scores would have been absent the UK policy change. This evidence provides strong support for the main prediction of our model, namely that greater availability of private CSR-engagement signals ought to reduce firms' reported CSR engagement in environmental initiatives.

We end our baseline analysis by providing more formal empirical support for the parallel trends assumption discussed above. To that end, we consider the dynamic treatment effects of the UK policy intervention on report-based environmental CSR scores: we regress firms' environmental scores on a treatment group dummy, year dummies with 2010 as the benchmark year, and interactions between the treatment group dummy and each of the year dummies (as well as control variables and firm fixed effects). For clarity and brevity purposes, we omit the regression table here, and instead represent the treatment-year coefficient estimates graphically, along with 95% confidence intervals, in Figure 3 below.

[Insert Figure 3 here.]

Figure 3 illustrates statistically insignificant treatment-year coefficient estimates for the pre-intervention years (2005 - 2010), lending further support to the parallel trends assumption; and negative and significant estimates in the post-intervention years (2011-2015), consistent with the negative treatment effect identified in Table 3 and discussed above.

5.3 Difference-in Differences Estimation with Propensity Score Reweighting

The baseline analysis presented in the previous section provides empirical support for the validity of DiD estimation's identifying parallel trends assumption. In order to further improve the plausibility of this assumption, pre-processing methods are often used prior to DiD estimation. These methods involve either reweighting or dropping units to make the treatment and control groups more comparable - or "balanced" - in terms of their covariates, and in turn to improve the treatment variable's independence from background characteristics (Hainmueller 2012).

There are a number of balancing methods available. In this section, we follow the recent applied work of Guadalupe *et al.* (2012) and Cunningham *et al.* (2021), and focus on *propensity score reweighting (PSR)* (Hirano *et al.* 2003, Busso *et al.* 2014), as our main pre-processing method;[10] and estimate the propensity score function using the following covariates:[11] *Firm Size*, *ROA*, *Leverage*, *Sales*, *R&D Intensity*, and *Market Competition*.[12]

[10] We also consider other pre-processing methods as robustness checks (see below), and obtain similar results.

[11] This choice of covariates is based on the correlations between these variables and *Env Reporting*, which are reported in the Online Appendix; and on the existing literature (e.g. Konar and Cohen 2001; Krüger, 2015; Ioannou, Li, and Serafeim, 2016; Flammer, 2015).

[12] PSR involves first using a logit regression of treatment on these covariates and using the predicted value from the regression as estimated probability of treatment - or propensity score - for each observation in the control and treatment groups. Having done that, each observation is weighted by the inverse of the probability of receiving the treatment it did actually receive - in order to make the treated and control groups more similar. Thus, an observation in the treatment group with a propensity score $\hat{p}$ receives a weight of $1/\hat{p}$; while an observation in the control group with the same propensity score $\hat{p}$ receives a weight of $1/(1-\hat{p})$ (Huntington-Klein 2022, p.280). To mitigate the impact of possible extreme values of weights, we winsorize the weights at the 1st and 99th levels.

The results of our PSR - DiD regressions, which are derived from the same specification in each column as in the baseline analysis, are reported below in Table 4, where in all specifications standard errors are adjusted for within-firm clustering.

[Insert Table 4 here.]

Once again, the coefficient estimate on $Treat \times After$ is negative and statistically significant at the 1% significance level, and remains fairly constant around -5, in all six specifications. Taken together these results offer further evidence of a statistically and economically strong causal impact of *The UK Companies Act* and its associated increase in private signal availability on UK firms' reporting on their environmental CSR engagement, consistent with the main prediction of our model.

As in the baseline scenario, we evaluate the parallel trends assumption by considering the dynamic treatment effects of the UK policy intervention on report-based environmental CSR scores. Figure 4 below represents the treatment-year coefficient estimates, along with 95% confidence intervals, graphically.

[Insert Figure 4 here.]

The statistically insignificant treatment-year coefficient estimates for all pre-intervention years (2005 - 2010) except 2007, and the negative and significant estimates for the post-intervention years (2011 - 2015), provide empirical support for the parallel trends assumption and the negative treatment effect identified in Table 4, respectively.

5.4 Robustness Considerations

In order to provide additional evidence in support of our main empirical results, our Online Appendix contains a battery of robustness checks. We consider placebo tests based on 1) a sample with pre-intervention period only and using years 2007 and 2008 as pseudo-intervention years; and 2) different dependent variables, specifically report-based CSR engagement on *non-environmental* activities. With placebo intervention years, we find no evidence of a decrease in UK firms' reporting on environmental initiatives. When we consider placebo dependent variables, we find that while (as shown above) a reduction in UK firms' CSR reporting due to the UK policy intervention holds for environmental reporting, results are weaker or insignificant for reporting on non-environmental activities. Taken together, these results further corroborate our main modeling prediction.

Further robustness checks include using alternative pre-processing matching methods - namely entropy balancing, propensity score matching, and propensity score stratification; addressing serial correlation concern by collapsing the sample into two periods or clustering standard errors at higher levels; using two modified samples that take out firms from countries with confounding policies around 2011 and firms associated with EU Emissions Trading System, respectively; using 2012 as an intervention year, or dropping the years (2011-2012) surrounding the intervention, and controlling for lagged dependent variable to mitigate time-variant unobserved heterogeneity concerns; using three alternative dependent variables to measure firms' reporting on environmental activities, namely report-based environmental-innovation scores, resources-use scores, and emissions-reduction scores; focusing on a subsample of firms that publish separate CSR reports; and finally, using Sustainalytics as an alternative data set to Refinitiv ESG to capture firms' self-reporting on CSR engagement.

We report the results of these robustness checks in our supplementary empirical analysis in the Online Appendix. In all cases, we find that the coefficient estimate on $Treat \times After$ continues to be negative; as well as broadly consistent in magnitude across regressions and economically and statistically significant; thus providing additional support for our main empirical results.

6 Evidence from the UK Carbon Disclosure Project

To test the remaining predictions of our model we turn next to data from the Carbon Disclosure Project (CDP), a data set that allows us to distinguish high-type firms from low-type firms using private signal disclosure - i.e. disclosure of GHG emissions - as a proxy for firm type, since as shown in the model and discussed in Section 3.3, high-disclosing firms are likely to be predominantly high-type firms committed to environmental CSR while low-disclosing firms are likely to be predominantly low-type firms uncommitted to CSR.

CDP is a UK-registered charity which operates a global carbon disclosure system in which companies and cities voluntarily report data on environmental performance. These data are then analyzed and transformed into metrics that investors can use to make investment decisions. According to the CDP website (`http://www.cdp.net`) the project has enrolled over 7,000 firms globally. CDP uses a questionnaire to collect information on disclosure of GHG emissions from firms in both emerging and developed markets. Firms can choose to respond to CDP's request by marking their response status as either "Public" or "Private"; or they can choose not to respond, which yields a status of "NA". Public responses from firms indicate that these firms will disclose their GHG emissions to the general public; whereas private responses allow CDP to include that firm's data in broader regional and industry indices, but not to make the firm's identity known. Our sample period with the CDP data is 2010-2013.

Following Krüger (2015), we construct treatment and control groups within the UK using CDP, based on firms' response statuses regarding their GHG emissions before *The UK Companies Act* came into effect in late 2013.[13] In particular, we first calculate the frequency of firms' response status as "Public" during 2010 - 2013 for each UK firm in the sample. We then construct a dummy variable - high disclosers - coded one if the number of times that a firm publicly discloses their CSR information over the four-year period is 3 or more; and coded zero otherwise (corresponding to low disclosers). Accordingly, the control group consists of high-type, "high-disclosing" UK firms, and our treatment group includes low-type, "low-dislosing" firms in our UK sample. We use the same baseline DiD and PSR - DiD approaches as in Section 6, and report the results below.

Merging data from CDP with Refinitiv ESG and Global Compustat, we construct two samples of UK-only firms: 1) a baseline within-UK sample of 2,479 firm-year observations with 301 UK firms operating in 156 4-digit SIC industries over the period of 2005 - 2015, used for our baseline DiD estimation; and 2) a within-UK PSR sample used for our PSR-DiD estimation and obtained using the same method as in Section 6 and estimating propensity scores based on the following covariates: *Firm Size*, *ROA*, *Leverage*, *Sales*, *R&D Intensity*, and *Market Competition*.

The results are reported below in Table 5.

[Insert Table 5 here.]

Panel A presents summary statistics regarding CSR engagement scores based on self-reported information, for both control and treatment groups, in our baseline within-UK sample. Consistent with model predictions, mean and median CSR engagement scores are much lower in the treatment group of low-disclosers (low-type firms) than in the control group of high-disclosers (high-type firms). For example, in our within-UK sample the median CSR engagement score is 27/100 for low-type firms versus 54/100 for high-type firms.

Panel B reports the baseline DiD regression results in Columns (1)-(3), while Columns (4)-(6) report results from DiD regressions with propensity-score reweighting. In columns (1) and (4), the regressions include our firm-level and industry-level control variables - *Firm Size*, *ROA*, *Leverage*,

[13]Grewal (2017) also uses CDP to obtain emissions information for firms that disclosed voluntarily, and considers the impact of *The UK Companies Act* on these firms' subsequent GHG emissions.

Sales, *R&D Intensity*, and *Market Competition* - but no fixed effects. In columns (2) and (5), we add year fixed effects and in columns (3) and (6) we further include firm fixed effects. All standard errors are adjusted for within-firm clustering.

The coefficient estimate on the interaction term $Treat \times After$ is negative and statistically significant at the 5% or the 1% level in all but one specification; the exception being the baseline regression with year and firm fixed effects, in which case the coefficient estimate is significant at the 10% level. All in all, these results offer evidence - consistent with our model prediction - of a differential impact of *The UK Companies Act* and its associated increase in private CSR-engagement signal availability on reported CSR engagement in high-type firms versus low-type firms, with the latter being more negatively affected by the UK policy change than the former.

To explore the parallel trends assumption, we regress firms' report-based environmental scores on a treatment group dummy, year dummies with 2010 as the benchmark year, as well as interactions between the treatment group dummy and each of the year dummies. The regressions also includes the control variables discussed in Section 5.1, firm fixed effects, and standard errors adjusted for within-firm clustering. As discussed above, if the parallel trends assumption holds, we would expect the coefficient estimates on treatment-year interactions should be statistically insignificant for years in the pre-intervention period. The treatment-year coefficient estimates are reported graphically, along with 95% confidence intervals, for the baseline within-UK sample and PSR within-UK sample, in Figures 5a and 5b, respectively.

[Insert Figures 5a and 5b here.]

Figure 5a illustrates statistically insignificant treatment-year coefficient estimates for all pre-intervention years (2005 - 2010), suggesting an overall parallel evolution of environmental score outcomes between treatment and control groups before the 2011 intervention year; and lending further support to the validity of the parallel trends assumption in this context. We also observe negative and significant treatment-year coefficient estimates for the first 3 of the post-intervention years (2011 - 2015), consistent with the negative treatment effect identified in Table 3 and discussed above. Figure 5b illustrates mostly insignificant treatment-year coefficient estimates in the pre-intervention period, and negatively significant estimates in at least some post-intervention years.

7 Conclusion

Today, reporting on CSR engagement is commonplace. Of the 250 largest companies (by revenue) in the world, all but ten report their corporate social responsibility (CSR) and sustainability-related activities to external stakeholders (KPMG, 2020). But are these claims accurate? This paper develops and tests a model that studies the interplay between firms' incentives to mislead customers about their CSR engagement, and the availability of private engagement-correlated information that can be voluntarily revealed to markets. Based on whether information is disclosed to them, and on its content, markets form beliefs about the type of firm they are facing and the truthfulness of their claims. The market's *ex post* inferences have a disciplining effect on firms' *ex ante* reporting on their CSR engagement. The model predicts that the availability of private information that can be disclosed to markets reduces CSR-engagement claims, as low-engagement firms that once attempted to pool with high-engagement firms find it too costly to exaggerate their CSR claims.

We test this in the context of environmental CSR, where GHG emissions represent the potentially (un)available private information correlated with firms' environmental CSR engagement. We exploit *The UK Companies Act*, which encouraged publicly traded UK companies to at least try to obtain (and ideally disclose) measures of their GHG emissions - an event we interpret as an exogenous increase in the availability of private engagement-correlated information across firms. Using a difference-in-differences approach, we show that - consistent with our theory - firms in the UK make lower CSR claims after increased availability of private information, compared to firms from the other 15 European countries which did not experience the policy change. We also perform a number of robustness checks and explore other predictions of the model.

Overall, our work contributes to the literature on CSR by bringing together 1) a simple theory of communication of non-verifiable information, when information signals may be privately observed *ex post*; 2) a natural application of the theory to CSR-reporting context, when firms may obtain *ex post* GHG emissions signals correlated to CSR activities; and 3) an empirical analysis providing support for the main predictions of our model, using a variety of data sets and methods.

Our paper also has broader implications for applied theory: although our model is presented in the context of exaggerated CSR-engagement claims and availability of privately observed but disclosable CSR performance measures (GHG emissions), its insights about the disciplining effects of *ex post* privately observed but disclosable type-correlated signals on *ex ante* unverifiable messages could apply to a large number of situations in which *puffery* may be relevant.

Finally, our paper offers food for thought regarding ESG-related policy discussions. Misleading statements about CSR engagement and ESG initiatives may be very costly from a social point of view, because it may lead to an inefficient allocation of resources - by consumers making purchases, by workers selling their labor, by investors providing capital, etc. Our paper suggests that full-fledged CSR *performance* disclosure requirements (such as GHG emissions disclosure requirements) - which may be costly to implement and monitor - may not be necessary to induce truthful reporting of CSR *engagement* efforts and investments. A less onerous policy alternative may be to focus on increasing the private availability of CSR performance signals to firms, and let voluntary disclosure of private information promote better resource allocation through market discipline.

References

[1] Abadie, Alberto. 2005. "Semiparametric Difference-in-Differences Estimators." *Review of Economic Studies* **72 (1)**: 1-19.

[2] Bolton, Patrick and Marcin Kacperczyk. 2021. "Carbon Disclosure and the Cost of Capital" Working Paper, Imperial College.

[3] Busso, Matias, John DiNardo, and Justin McCrary. 2014. "New Evidence on the Finite Sample Properties of Propensity Score Reweighting and Matching Estimators." *The Review of Economics and Statistics* **96(5)**: 885–897.

[4] The Companies Act 2006 (Strategic Report and Directors' Report) Regulations 2013. https://www.legislation.gov.uk/ukdsi/2013/9780111540169/pdfs/ukdsi_9780111540169_en.pdf

[5] ClientEarth. 2020. "Accountability Emergency: A Review of UK-Listed Companies' Climate Change-Related Reporting (20129-2020)." https://www.clientearth.org/media/wbglw3r3/clientearth-accountability-emergency.pdf

[6] Cunningham, Colleen, Florian Ederer, and Song Ma. 2021. "Killer Acquisitions." *Journal of Political Economy* **129 (3)**: 649-702.

[7] Downar, Benedikt, Jürgen Ernstberger, Stefan Reichelstein, Sebastian Schwenen, and Aleksandar Zaklan. 2021. "The impact of carbon disclosure mandates on emissions and financial operating performance." *Review of Accounting Studies* **26**: 1137–1175.

[8] Edmans, Alex. 2020. *Grow the Pie*. Cambridge: Cambridge University Press.

[9] Flammer, Caroline. 2015. "Does Product Market Competition Foster Corporate Social Responsibility? Evidence from Trade Liberalization." *Strategic Management Journal* **36**: 1469-1485.

[10] Friedman, Milton. 1962. *Capitalism and Freedom*. University of Chicago Press.

[11] Grewal, Jody. 2017. "The Effects of Transparency Regulation on Voluntary Disclosers: Evidence from Mandatory Carbon Reporting." Working Paper, University of Toronto.

[12] Grewal, Jody, Gordon Richardson, and Jingjing Wang. 2022. "The Effects of Mandatory Carbon Reporting on Greenwashing." Working Paper, University of Toronto.

[13] Grossman, Sanford. 1981. "The Informational Role of Warranties and Private Disclosure about Product Quality." *Journal of Law and Economics* **24 (3)**: 461-483.

[14] Grossman, Sanford and Oliver Hart. 1980. "Disclosure Laws and Takeover Bids." *Journal of Finance* **35 (2)**: 323-334.

[15] Guadalupe, Maria, Olga Kuzmina, and Catherine Thomas. 2012. "Innovation and Foreign Ownership." *American Economic Review* **102 (7)**: 3594-3627.

[16] Hainmueller, J. 2012. "Entropy Balancing for Causal Effects: A multivariate reweighing method to produce balanced samples in observational studies." *Political Analysis* **20 (1)**: 25-46.

[17] Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies* **64 (4)**: 605-654.

[18] Hirano, Keisuke, Guido W. Imbens, and Geert Ridder. 2003. "Efficient Estimation of Average Treatment Effects using the Estimated Propensity Score." *Econometrica* **71 (4)**: 1161-1189.

[19] Huntington-Klein, Nick. 2022. *The Effect: An Introduction to Research Design and Causality.* CRC Press.

[20] Ioannou, Ioannis, Shelley Xin Li, and George Serafeim. 2016. "The Effect of Target Difficulty on Target Completion: The Case of Reducing Carbon Emissions." *The Accounting Review* **91(5)**: 1467–1492.

[21] Jouvenot, Valentin and Philipp Krüger. 2021. "Mandatory Corporate Carbon Disclosure: Evidence from a Natural Experiment" Working Paper, University of Geneva.

[22] Kim, Eun-Hee and Thomas P. Lyon. 2011. "Strategic Environmental Disclosure: Evidence from the DOE's Voluntary Greenhouse Gas Registry." *Journal of Environmental Economics and Management* **61(3)**: 311-326.

[23] Kim, Soohun and Aaron Yoon. 2023. "Analyzing Active Fund Managers' Commitment to ESG: Evidence from the United Nations Principles for Responsible Investment." *Management Science* **69(2)**: 741-758.

[24] KPMG IMPACT. 2020. "The Time Has Come: The KPMG Survey of Sustainability Reporting 2020." https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2020/11/the-time-has-come.pdf.

[25] Konar, Shameek and Mark A. Cohen. 2001. "Does the Market Value Environmental Performance?" *Journal of Economic Literature* **83(2)**: 281-289.

[26] Krüger, Philipp, 2015. "Climate Change and Firm Valuation: Evidence from a Quasi-Natural Experiment." Working Paper, Swiss Finance Institute.

[27] Levitt, Theodore. 1958. "The Dangers of Social Responsibility." *Harvard Business Review* **36(5)**: 41–50.

[28] Lyon, Thomas P. and John W. Maxwell. 2011. "Greenwash: Corporate Environmental Disclosure under Threat of Audit." *Journal of Economics and Management Strategy* **20(1)**: 3-41.

[29] Marquis, Christopher, Michael W. Toffel, and Yanhua Zhou. 2016. "Scrutiny, Norms, and Selective Disclosure: A Global Study of Greenwashing." *Organization Science* **27 (2)**: 483-504.

[30] Maxwell, John W., Thomas P. Lyon, and Steven C. Hackett. 2000. "Self-Regulation and Social Welfare: The Political Economy of Corporate Environmentalism." *Journal of Law and Economics* **43 (2)**: 583-617.

[31] Milgrom, Paul. 1981. "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics* **12 (2)**: 380-391.

[32] Rajan, Raghuram, Pietro Ramella, and Luigi Zingales. 2023. "What Purpose Do Corporations Purport? Evidence from Letters to Shareholders," Working Paper, ECGI.

[33] Ramus, Catherine A. and Ivan Montiel. 2005. “When Are Corporate Environmental Policies a Form of Greenwashing? ” *Business & Society* **44 (4)**: 377 - 414.

[34] Refinitiv. 2022. “Environmental, Social and Governance Scores from Refinitiv,” Report, LSEG, London, UK. https://www.lseg.com.

[35] TerraChoice. 2010. “The Sins of Greenwashing: Home and Family Edition,” Report, Terrachoice/UL, Ottawa, ON, Canada. http://sinsofgreenwashing.org/index35c6.pdf.

[36] Wu, Yue, Kaifu Zhang, and Jinhong Xie. 2020. “Bad Greenwashing, Good Greenwashing: Corporate Social Responsibility and Information Transparency. ” *Management Science* **66 (7)**: 3095-3112.

Table 1: Variable Definitions

This table presents the definitions of the variables used in our empirical analysis. All variables are measured annually at the firm level, except market competition which is measured at the industry level.

Variables	Definitions	Sources
Env Reporting	The weighted average relative rating of a firm based on the reported environmental information and the resulting three environmental category scores in environmental innovation, resource use, and emissions reduction.	Refinitiv ESG
Firm Size	Natural logarithm of one plus a firm's total assets, measured at the end of fiscal year.	Compustat Global
ROA	Operating income before depreciation divided by book value of total assets, measured at the end of fiscal year.	Compustat Global
Leverage	Book value of debt divided by book value of total assets measured at the end of fiscal year.	Compustat Global
Sales	Natural logarithm of one plus a firm's total sales, measured at the end of fiscal year.	Compustat Global
R&D Intensity	A firm's R&D expenditure scaled by its total assets, measured at the end of fiscal year.	Compustat Global
Market Competition	One minus the Herfindahl–Hirschman Index (HHI) based on four-digit SIC industries.	Compustat Global
Non-Env Reporting	The average of relative rating of a firm based on the reported social information and governance practices.	Refinitiv ESG
Environmental-Innovation Score	Environmental innovation category score reflects a company's capacity to reduce the environmental costs and burdens for its customers, and thereby creating new market opportunities through new environmental technologies and processes or eco-designed products.	Refinitiv ESG
Resource-Use Score	Resource use category score reflects a company's performance and capacity to reduce the use of materials, energy or water, and to find more eco-efficient solutions by improving supply chain management.	Refinitiv ESG
Emissions-Reduction Score	Emission category score measures a company's commitment and effectiveness towards reducing environmental emission in the production and operational processes.	Refinitiv ESG

Table 2: Summary Statistics and Sample Distribution

In Panel A of this table, we report summary statistics of all the variables in our sample by treatment and control group. Panel B presents the sample distribution by industry sector. The sample period of 2005 - 2015. Variable definitions are provided in Table 1.

Panel A: Summary Statistics for Treatment and Control Groups

	Treatment Group			Control Group		
variable	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.
Env Reporting	41.29	39.19	26.40	49.16	51.71	29.16
Firm Size	7.711	7.436	1.776	9.404	9.173	1.826
ROA	0.127	0.119	0.096	0.106	0.10	0.084
Leverage	0.238	0.213	0.187	0.259	0.247	0.164
Sales	5.626	6.638	3.303	6.799	8.025	3.795
R&D Intensity	0.011	0	0.033	0.016	0	0.035
Market Competition	0.585	0.634	0.334	0.621	0.679	0.319
Non-Env Reporting	46.73	46.44	18.02	49.49	49.54	19.93
Environmental-Innovation Score	18.17	0	27.91	31.74	21.63	33.85
Resource-Use Score	46.69	47.32	32.33	51.96	58.58	34.08
Emissions-Reduction Score	46.72	47.34	30.64	52.18	58.79	34.21

Panel B: Sample Distribution by Industry Sector for Treatment and Control Groups

	Total		Treatment Group		Control Group	
Industry	Freq.	Percent	Freq.	Percent	Freq.	Percent
Mining	426	4.86%	253	8.65%	173	2.97%
Manufacturing	3,413	38.97%	746	25.51%	2,667	45.71%
Construction	322	3.79%	147	5.03%	185	3.17%
Transportation	1,266	14.46%	333	11.39%	933	15.86%
Wholesale Trade	242	2.76%	118	4.04%	124	2.13%
Finance	2,022	23.09%	745	25.48%	1,277	21.89%
Service	965	11.02%	567	19.39%	398	6.82%
Public Administration	92	1.05%	15	0.51%	77	1.32%

Table 3: Private Signal Availability and Reporting on Environmental Initiatives: Baseline

This table presents coefficients estimates of our baseline regressions examining the effect of private signal availability on a firm's reported environmental CSR engagement. We use *The UK Companies Act* as an exogenous increase in private signal availability, with 2011 as our intervention year. We construct a UK-EU sample of UK firms and firms from 15 European countries over the 2005-2015 time period. The dependent variable is *Env Reporting*. $Treat$ is a dummy variable equal to 1 for UK-listed firms and to 0 for EU firms. $After$ is a dummy variable equal to 0 for all years in the pre-intervention period (2005-2010), and to 1 for all years in the post-intervention period (2011-2015). $Treat \times After$ is the interaction term that provides the DiD estimate. All variables are defined in Table 1. Standard errors are reported in parentheses below the coefficient estimates and adjusted for within-firm clustering. Significance at the 10%, 5%, and 1% level is indicated by *,**, ***, respectively.

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Treat	-5.359***	9.487***	10.013***	8.541***	20.457***	
	(1.662)	(1.417)	(1.417)	(1.515)	(2.838)	
After	13.222***	12.634***				
	(0.741)	(0.685)				
$Treat \times After$	-5.253***	-4.843***	-5.686***	-4.736***	-4.921***	-3.179***
	(1.228)	(1.138)	(1.137)	(1.111)	(1.097)	(1.083)
Firm Size		8.426***	8.392***	8.352***	8.920***	2.605**
		(0.330)	(0.329)	(0.710)	(0.680)	(1.163)
ROA		-4.329	2.691	11.841*	13.573**	-1.443
		(6.411)	(6.414)	(6.223)	(5.819)	(5.386)
Leverage		0.499	-0.096	1.931	-0.766	0.181
		(3.392)	(3.397)	(3.826)	(3.625)	(3.249)
Sales		1.343***	1.307***	1.641**	2.061***	1.176
		(0.214)	(0.215)	(0.744)	(0.702)	(0.728)
R&D Intensity		11.693	6.840	61.497**	68.250***	-21.783
		(17.196)	(17.193)	(25.199)	(21.658)	(21.974)
Market Competition		0.652	0.579	-0.184	0.400	1.451
		(2.285)	(2.279)	(3.503)	(3.542)	(3.412)
Year FE	No	No	Yes	Yes	Yes	Yes
Industry FE	No	No	No	Yes	Yes	No
Country FE	No	No	No	No	Yes	No
Firm FE	No	No	No	No	No	Yes
Observations	9,328	8,783	8,783	8,783	8,783	8,783
R-squared	0.059	0.358	0.407	0.582	0.618	0.851
Clusters	1,182	1,133	1,133	1,133	1,133	1,133

Table 4: Private Signal Availability and Reporting on Environmental Initiatives: PSR Sample

This table presents coefficients estimates of our regressions examining the effect of private signal availability on a firm's reported environmental CSR engagement, using a propensity-score-reweighted sample (PSR sample) over the 2005-2015 time period. We use *The UK Companies Act* as an exogenous increase in private signal availability, with 2011 as our intervention year. The dependent variable is *Env Reporting*. $Treat$ is a dummy variable equal to 1 for UK-listed firms and to 0 for EU firms. $After$ is a dummy variable equal to 0 for all years in the pre-intervention period (2005-2010), and to 1 for all years in the post-intervention period (2011-2015). $Treat \times After$ is the interaction term that provides the DiD estimate. All variables are defined in Table 1. Standard errors are reported in parentheses below the coefficient estimates and adjusted for within-firm clustering. Significance at the 10%, 5%, and 1% level is indicated by *,**, ***, respectively.

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Treat	11.392***	9.520***	9.636***	8.934***	20.551***	
	(2.327)	(1.351)	(1.336)	(1.398)	(2.883)	
After	14.027***	12.664***				
	(0.929)	(0.766)				
$Treat \times After$	-4.927***	-5.334***	-5.499***	-5.276***	-5.590***	-4.853***
	(1.396)	(1.208)	(1.189)	(1.173)	(1.166)	(1.235)
Firm Size		8.458***	8.405***	8.154***	8.326***	3.129***
		(0.309)	(0.304)	(0.707)	(0.697)	(1.061)
ROA		1.330	8.600	13.254**	12.574**	0.521
		(6.540)	(6.528)	(6.012)	(5.872)	(5.920)
Leverage		1.426	0.670	-0.341	-2.298	-2.002
		(3.616)	(3.610)	(3.867)	(3.729)	(3.435)
Sales		1.164***	1.101***	1.919***	2.513***	1.150
		(0.208)	(0.207)	(0.733)	(0.684)	(0.793)
R&D Intensity		30.489*	25.611	56.060**	65.852***	-10.016
		(17.274)	(17.257)	(25.107)	(24.335)	(25.929)
Market Competition		-1.079	-1.055	-0.468	-0.119	1.360
		(2.486)	(2.486)	(3.448)	(3.391)	(3.487)
Year FE	No	No	Yes	Yes	Yes	Yes
Industry FE	No	No	No	Yes	Yes	No
Country FE	No	No	No	No	Yes	No
Firm FE	No	No	No	No	No	Yes
Observations	8,783	8,783	8,783	8,783	8,783	8,783
R-squared	0.067	0.443	0.484	0.635	0.660	0.861
Clusters	1,182	1,133	1,133	1,133	1,133	1,133

Table 5: Private Signal Availability and Reporting on Environmental Initiatives: High-Type Firms vs. Low-Type Firms

Panel A of this table presents summary statistics based on reporting status for UK companies that provide environmental data to the CDP project. "High-disclosers" are firms that allowed CDP to publicly disclose their reported GHG emissions more frequently than the median number of years firms in the sample allowed CDP to publicly disclose their reported GHG emissions. Conversely, "low-disclosers" are firms that allowed CDP to publicly disclose their reported GHG emissions less frequently than the median number of years firms allowed public disclosure of their GHG emissions. Panel B of this table presents coefficients estimates of regressions which examine the differential effect of private signal availability on reporting on environmental initiatives by high-disclosing (high-type) in the control and by low-disclosing (low-type) firms in the treatment group, over the 2005-2015 time period. Columns (1) - (3) and (4) - (6) present the results of the within-UK sample (UK sample) and the within-UK propensity-score- reweighted sample (UK PSR sample), respectivley. We use *The UK Companies Act* as an exogenous increase in private signal availability, with 2011 as our intervention year. The dependent variable *Env Reporting*. $Treat$ is a dummy variable equal to 1 for low-disclosers and to 0 for high-disclosers. $After$ is a dummy variable equal to 0 for all years in the pre-intervention period (2005-2010), and to 1 for all years in the post-intervention period (2011-2015). $Treat \times After$ is the interaction term that provides the DiD estimate. All variables are defined in Table 1. Standard errors are reported in parentheses below the coefficient estimates and adjusted for within-firm clustering. Significance at the 10%, 5%, and 1% level is indicated by *,**, ***, respectively.

Panel A: CDP Data

Disclosure Status	Mean	Median	SD	Min	Max	Firm-Yrs
High Disclosers	53.34	54.13	24.79	0	95.53	1,283
Low Disclosers	29.19	26.82	21.22	0	90.58	1,179

Panel B: Private Signal Availability and Reporting on Environmental Initiatives

	UK Sample			UK PSR Sample		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Treat	-12.812***	-13.498***		-11.699***	-12.526***	
	(2.331)	(2.354)		(2.323)	(2.351)	
After	11.414***			12.298***		
	(1.419)			(1.668)		
$Treat \times After$	-5.001**	-4.694**	-3.542*	-6.269***	-5.560**	-5.560***
	(2.024)	(2.022)	(1.916)	(2.245)	(2.234)	(2.030)
Firm Size	7.419***	7.335***	1.348	8.048***	7.990***	0.618
	(0.630)	(0.643)	(2.246)	(0.736)	(0.760)	(2.392)
ROA	13.943	22.274*	6.118	12.390	20.924	5.218
	(11.614)	(11.537)	(8.714)	(13.293)	(13.122)	(9.661)
Leverage	5.869	5.968	-1.427	2.876	2.694	-1.671
	(4.948)	(5.047)	(5.440)	(5.893)	(6.094)	(5.983)
Sales	0.074	-0.050	1.731	-0.149	-0.274	2.396
	(0.346)	(0.352)	(2.099)	(0.391)	(0.401)	(2.271)
R&D Intensity	21.926	16.436	-28.648	28.933	22.199	-30.511
	(18.412)	(18.339)	(38.345)	(21.179)	(21.274)	(36.209)
Market Competition	-1.946	-2.043	1.444	-2.914	-2.891	-0.154
	(3.780)	(3.797)	(6.351)	(4.217)	(4.196)	(7.158)
Year FE	No	Yes	Yes	No	Yes	Yes
Firm FE	No	No	Yes	No	No	Yes
Observations	2,313	2,313	2,313	2,313	2,313	2,313
R-squared	0.442	0.497	0.831	0.377	0.442	0.819
Clusters	281	281	281	281	281	281

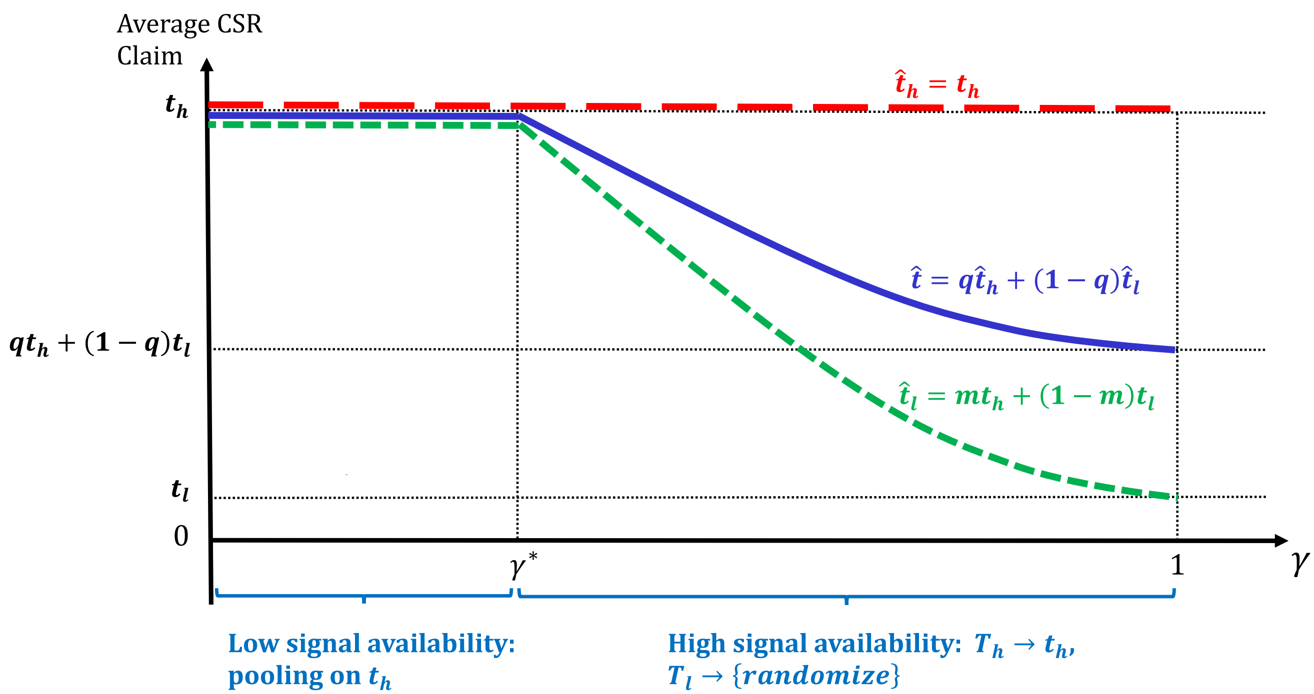
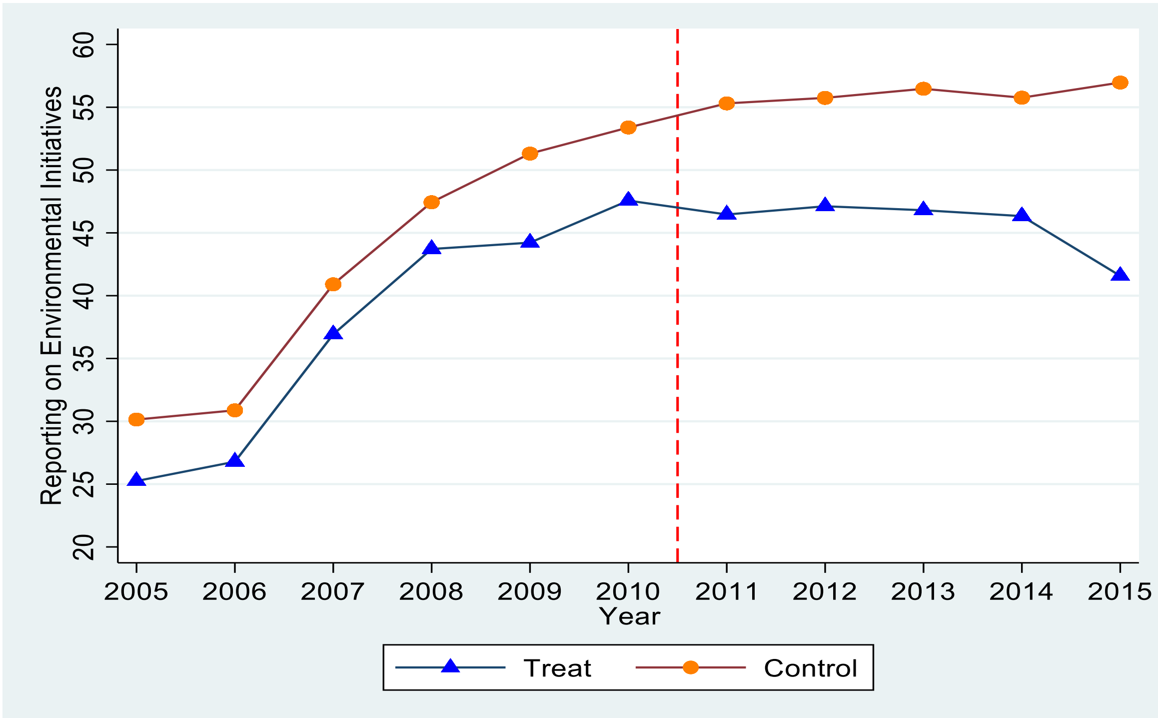
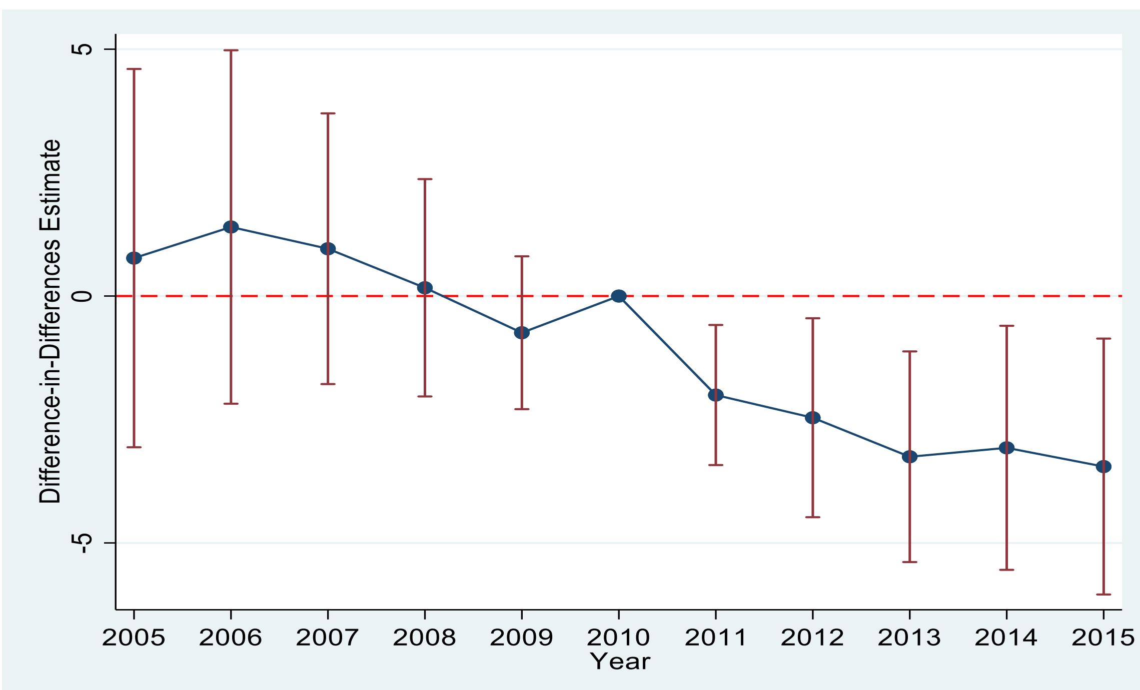


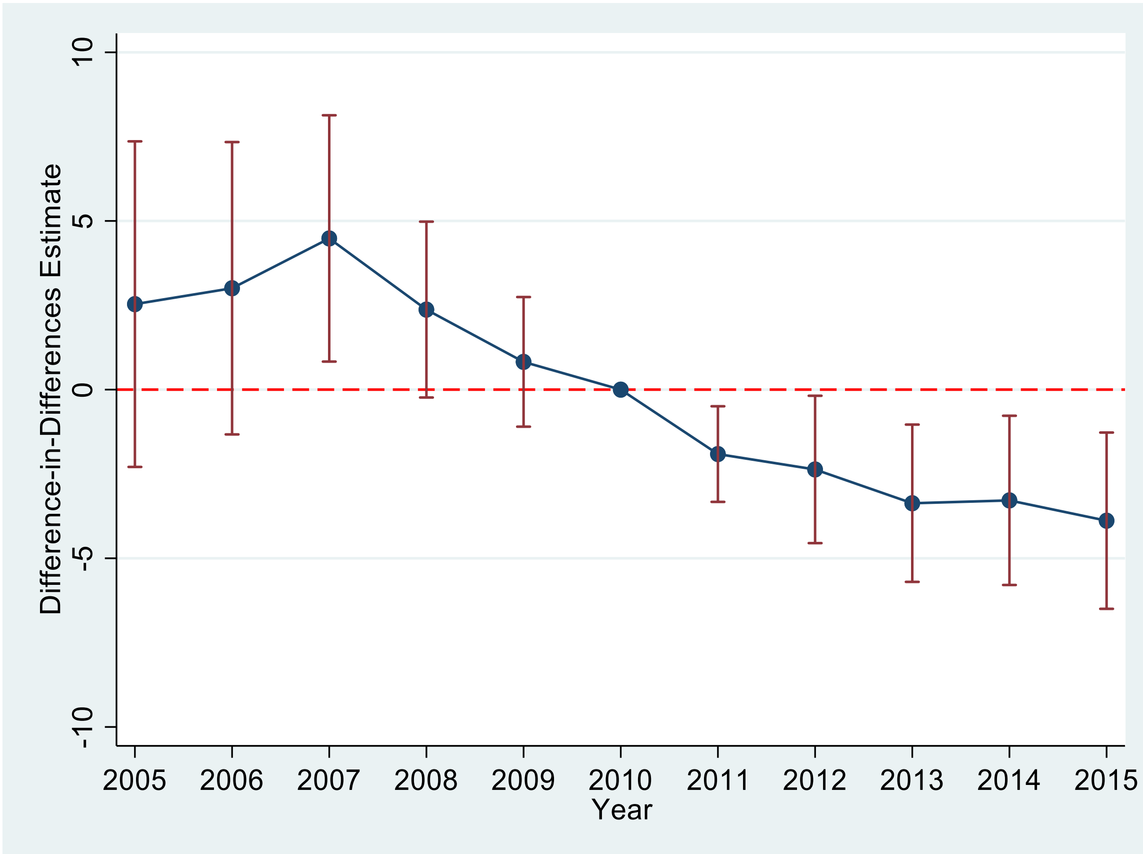
Figure 1: Private CSR-Engagement Signal Availability and CSR-Engagement Reporting



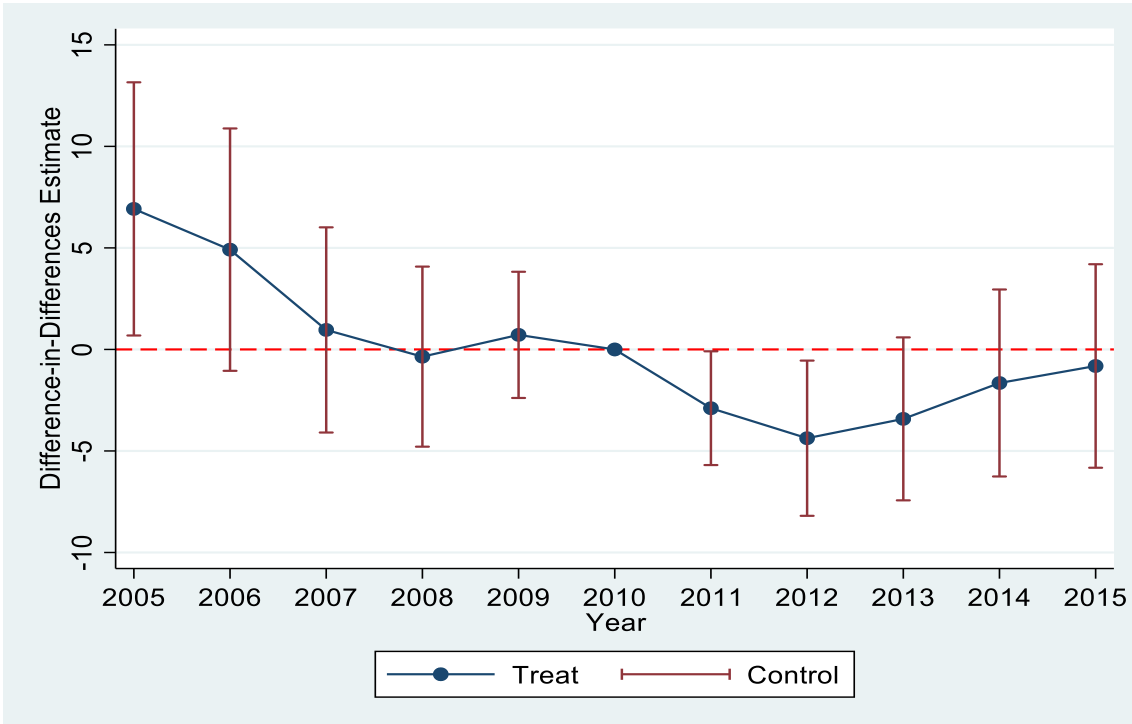
<u>Figure 2</u>: Reporting on Environmental Initiatives over 2005-2015 (UK-EU Sample)



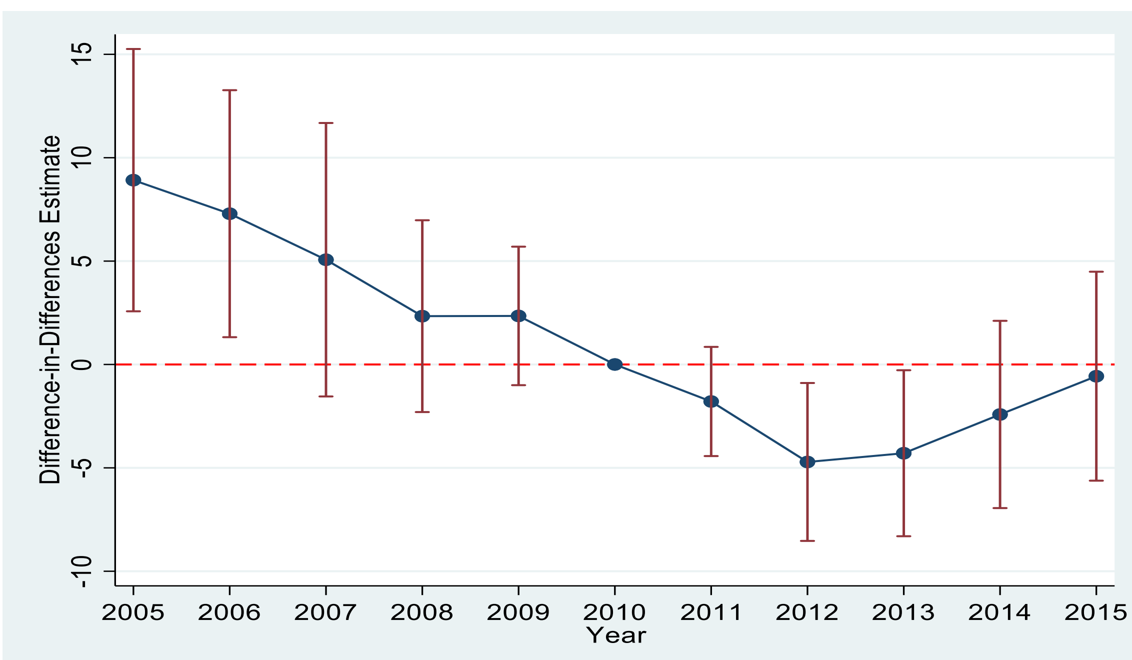
<u>Figure 3</u>: Dynamic Treatment Effects of Private CSR-Engagement Signal Availability on CSR-Engagement Reporting (UK-EU Sample)



<u>Figure 4:</u> Dynamic Treatment Effects of Private CSR-Engagement Signal Availability on CSR-Engagement Reporting (UK-EU PSR Sample)



<u>Figure 5a:</u> Dynamic Treatment Effects of Private CSR-Engagement Signal Availability on CSR-Engagement CSR Reporting (UK Sample)



<u>Figure 5b:</u> Dynamic Treatment Effects of Private CSR-Engagement Signal Availability on CSR-Engagement Reporting (UK PSR Sample)

Online Appendix for
Nothing but the Truth? Private Information and Reporting on Corporate Social Responsibility

Jean-Etienne de Bettignies
University of Toronto

Hua Fang Liu
Baruch College

David T. Robinson
Duke University and NBER

This Draft: February 26, 2024

Abstract

This online appendix contains proofs from the theoretical part of the paper, and additional robustness checks for the empirical analysis.

Key words: Corporate Social Responsibility Reporting; ESG; Signal Availability; Greenwashing; Greenhouse Gas Emissions; UK Companies Act of 2013

Email: jean.debettignies@rotman.utoronto.ca (Bettignies); huafang.liu@baruch.cuny.edu (Liu); davidr@duke.edu (Robinson).

1 Introduction

This appendix accompanies "Nothing But the Truth: Private Information and Reporting On Corporate Social Responsibility." In the next section, we provide proofs for the main Lemmas and Propositions contained in Section 3 of the paper. In Section 3 of this appendix, we propose a number of extensions and robustness checks to the empirical analysis presented in Sections 5 and 6 of the paper.

2 Theoretical Appendix - Proofs

2.1 Proof of Lemma 1

Recall that in any equilibrium firm T_i, $i \in \{h, l\}$, makes announcement s_0 whenever it does not receive a signal, since that is its only option in that case; and let us underline three additional points.

First, we cannot have an equilibrium in which firm T_h always makes announcement s_0 that it did not receive a signal, regardless of whether it did in fact receive one. To see this, suppose an equilibrium with these characteristics did exist. Then, upon observing announcement s_0, consumers' posterior belief that they are facing a firm of type T_h would be

$$p_2(s_0, p_1) = P(T_h \mid s_0) = \frac{P(s_0 \mid T_h)\, p_1}{P(s_0 \mid T_h)\, p_1 + P(s_0 \mid T_l)(1 - p_1)}, \tag{1}$$

with $P(s_0 \mid T_h) = 1$, and $P(s_0 \mid T_l) \in \{1 - \gamma, 1\}$ depending on firm T_l's equilibrium strategy. Evidently, regardless of firm T_l's equilibrium strategy, we have posterior belief $p_2(s_0, p_1) < 1$ for all prior beliefs $p_1 \in (0, 1)$. Hence, given signal announcement s_0 at date 2, consumers' willingness to pay for the firm's products, and in turn firm T_h's profits, can be expressed as: $\Pi_h(t_j, p_2(s_0, p_1)) = p_2(s_0, p_1)\, V_{hj} + (1 - p_2(s_0, p_1))\, V_{lj} < V_{hj}$.

In contrast, conditional on observing out-of-equilibrium announcement s_h, consumers assign out-of-equilibrium belief μ_h that the firm is of type h. Since announcement s_h is a verifiable disclosure of signals S_h, and since by assumption $P(S_h \mid T_h) = 1$, we can use Bayes' rule to derive belief $\mu_h = 1$. Thus, with these beliefs, upon obtaining signal S_h, firm T_h is better off deviating and reporting its signal by making announcement s_h, which would yield profits $\Pi_i(t_j, 1) = V_{hj} > \Pi_h(t_j, p_2(s_0, p_1))$. Hence this cannot be an equilibrium.

Second, we cannot have an equilibrium in which firm T_l makes announcement s_l when it receives signal S_l (and makes announcement s_0 when it does not receive any signal). To see this, suppose an equilibrium with these characteristics did exist. Since announcement s_l is a verifiable disclosure of signals S_l, and since by assumption $P(S_l \mid T_h) = 0$, upon observing announcement s_l, consumers would derive - using Bayes' rule - their posterior belief $p_2(s_l, p_1) = P(T_h \mid s_l, p_1) = 0$. Hence, given signal announcement s_l at date 2, consumers' willingness to pay for the firm's products, and in turn firm T_l's profits, can be expressed as: $\Pi_l(t_j, p_2(s_l, p_1)) = V_{lj}$.

In contrast, conditional on observing announcement s_0, consumers' posterior belief $p_2(s_0, p_1)$ would be as defined in (3) in the main text, with $P(s_0 \mid T_l) = 1 - \gamma$, and $P(s_0 \mid T_h) \in \{1 - \gamma, 1\}$ depending on firm T_h's equilibrium strategy. It then follows that regardless of firm T_l's equilibrium strategy, we have posterior belief $p_2(s_0, p_1) > 0$ for all prior beliefs $p_1 \in (0, 1)$. Thus, anticipating these beliefs and upon obtaining signal S_l, firm T_l is better off deviating and not reporting its signal by making announcement s_0, which would yield profits $\Pi_l(t_j, p_2) = p_2(s_0, p_1) V_{hj} + (1 - p_2(s_0, p_1)) V_{lj} > V_{lj}$. Hence this cannot be an equilibrium.

Finally, the foregoing discussion leaves only one candidate (pure strategy) equilibrium in signal announcement, in which a) upon receiving signal S_h, the T_h-type firm truthfully reports it by making announcement s_h; and b) upon receiving signal S_l, the T_l-type firm lies and makes announcement s_0 that it did not receive any signal. If this is an equilibrium, then the posterior beliefs follow directly from the above: $p_2(s_h, p_1) = 1$; $p_2(s_l, p_1) = 0$; and $p_2(s_0, p_1) \in (0, 1)$ for all $p_1 \in (0, 1)$ is as defined in (3) in the main text, with $P(s_0 \mid T_h) = 1 - \gamma$ and $P(s_0 \mid T_l) = 1$. Substituting these posterior beliefs in profit function (2) in the main text, one can readily verify that $\Pi_h(t_j, p_2(s_h, p_1)) = V_{hj} > p_2(s_0, p_1) V_{hj} + (1 - p_2(s_0, p_1)) V_{lj} = \Pi_h(t_j, p_2(s_0, p_1))$ and that $\Pi_l(t_j, p_2(s_0, p_1)) = p_2(s_0, p_1) V_{hj} + (1 - p_2(s_0, p_1)) V_{lj} > V_{lj} = \Pi_h(t_j, p_2(s_l, p_1))$.

In other words, conditional on receiving a signal, it is always optimal for the type-T_h firm to disclose it by making announcement s_h, and it is always optimal for the type-T_l firm not to disclose it and to pretend it did not receive it by making announcement s_0. This is consistent with the equilibrium under study, which therefore exists for all for all $p_1 \in (0, 1)$. ■

2.2 Proof of Proposition 2

If pooling on t_j, $j \in \{h, l\}$ is an equilibrium, then based on announcement t_j, consumers form interim belief $p_1(t_j, p_0) = p_0$; and based on announcement t_z, $z \neq j$, they form out-of-equilibrium interim belief η_z that the firm is of type T_h.

Using these interim beliefs, we can express firm T_h's date 1 profits from announcements t_j and

t_z, respectively, as follows:

$$\begin{aligned} E_h(t_j, p_0) &= \gamma\, \Pi_h(t_j, p_2(s_h, p_0)) + (1-\gamma)\, \Pi_h(t_j, p_2(s_0, p_0)) \\ &= \gamma\, V_{hj} + (1-\gamma)\left[p_2(s_0, p_0)\, V_{hj} + (1 - p_2(s_0, p_0))\, V_{lj}\right] \\ &= \gamma\, V_{hj} + (1-\gamma)\left[\frac{(1-\gamma)\, p_0}{(1-p_0) + (1-\gamma)\, p_0} V_{hj} + \frac{(1-p_0)}{(1-p_0) + (1-\gamma)\, p_0} V_{lj}\right] \end{aligned} \tag{2}$$

$$\begin{aligned} E_h(t_z, \eta_z) &= \gamma\, \Pi_h(t_z, p_2(s_h, \eta_z)) + (1-\gamma)\, \Pi_h(t_z, p_2(s_0, \eta_z)) \\ &= \gamma\, V_{hz} + (1-\gamma)\left[p_2(s_0, \eta_z)\, V_{hz} + (1 - p_2(s_0, \eta_z))\, V_{lz}\right] \\ &= \gamma\, V_{hz} + (1-\gamma)\left[\frac{(1-\gamma)\, \eta_z}{(1-\eta_z) + (1-\gamma)\, \eta_z} V_{hz} + \frac{(1-\eta_z)}{(1-\eta_z) + (1-\gamma)\, \eta_z} V_{lz}\right] \end{aligned} \tag{3}$$

In words, upon making equilibrium type announcement t_j, firm T_h anticipate that with probability γ it will receive signal S_h at date 2. In this case, as discussed in the main body of the paper, the firm will make signal announcement s_h, thus fully revealing its type and obtaining profits V_{hj}. With probability $1-\gamma$, firm T_h will receive no signal and will have to make announcement s_0, in which case its payoff will depend on consumers posterior belief $p_2(s_0, p_0)$, which itself depends on interim belief $p_1 = p_0$.

Firm T_h's date 1 profits take a similar form if the firm make out-of-equilibrium announcement t_z. If it receives signal S_h at date 2, Firm T_h can fully reveal its type and obtain profits V_{hz}. If it does not obtain a signal, its payoff will depend on consumers posterior belief $p_2(s_0, \eta_z)$, which itself depends on interim belief $p_1 = \eta_z$.

Similarly, we can also use these interim beliefs $p_1(t_j, p_0) = p_0$ and η_z to express firm T_l's date 1 profits from announcements t_j and t_z, respectively:

$$\begin{aligned} E_l(t_j, p_0) &= \Pi_l(t_j, p_2(s_0, p_0)) \\ &= p_2(s_0, p_0)\, V_{hj} + (1 - p_2(s_0, p_0))\, V_{lj} \\ &= \frac{(1-\gamma)\, p_0}{(1-p_0) + (1-\gamma)\, p_0} V_{hj} + \frac{(1-p_0)}{(1-p_0) + (1-\gamma)\, p_0} V_{lj} \end{aligned} \tag{4}$$

$$\begin{aligned} E_l(t_z, \eta_z) &= \Pi_l(t_z, p_2(s_0, \eta_z)) \\ &= p_2(s_0, \eta_z)\, V_{hz} + (1 - p_2(s_0, \eta_z))\, V_{lz} \\ &= \frac{(1-\gamma)\, \eta_z}{(1-\eta_z) + (1-\gamma)\, \eta_z} V_{hz} + \frac{(1-\eta_z)}{(1-\eta_z) + (1-\gamma)\, \eta_z} V_{lz} \end{aligned} \tag{5}$$

Date 1 profits for firms T_h and T_l differ only in the fact that the former does disclose its signal at date 2 if it receives it, and thus benefits from fully revealing its type in that case (an event which occurs with probability γ); while the latter always announces no having received a signal, even when in fact it did.

Importantly, note that out-of-equilibrium profits $E_h(t_z, \eta_z)$ and $E_l(t_z, \eta_z)$ are strictly increasing in out-of-equilibrium interim belief η_z. Intuitively, a higher interim belief leads to a higher posterior belief $p_2(s_0, \eta_z)$ that the firm is of type T_h, and in turn to a higher payoff for the firm.

Therefore, if there exist 1) a threshold $\eta_z^h \in [0,1]$ such that $E_h(t_j, p_0) \geq E_h(t_z, \eta_z)$ for all $\eta_z \in \left[0, \eta_z^h\right]$, and 2) a threshold $\eta_z^l \in [0,1]$ such that $E_l(t_j, p_0) \geq E_l(t_z, \eta_z)$ for all $\eta_z \in \left[0, \eta_z^l\right]$, then pooling on type announcement t_j, $j \in \{h, l\}$ is an equilibrium for all out-of-equilibrium beliefs $\eta_z \in \left[0, \min\left\{\eta_z^h, \eta_z^l\right\}\right]$.

There are two main points to highlight regarding pooling equilibria. First, *even if pooling on t_l does exist as an equilibrium for some out-of-equilibrium beliefs $\eta_h \in [0,1]$, this outcome would not survive the D_1 criterion* (Banks and Sobel, 1987). To see this, note that the incentive to deviate from t_l to t_h is always stronger for firm T_h than for firm T_l. Indeed, using the expressions in the paper, we can write:

$$E_h(t_h, \eta_h) - E_h(t_l, p_0) = \gamma \left[V_{hh} - V_{hl}\right] + (1-\gamma)\left[E_l(t_h, \eta_h) - E_l(t_l, p_0)\right]. \tag{6}$$

Since $V_{hh} - V_{hl} > 0$, then clearly at $\eta_h = \eta_h^l$ we have $E_l(t_h, \eta_h) - E_l(t_l, p_0) = 0$ and $E_h(t_h, \eta_h) - E_h(t_l, p_0) > 0$, and hence it must be that $\eta_h^h < \eta_h^l$. In turn, this implies that the set of out-of-equilibrium beliefs that would make deviation to t_h strictly preferable to equilibrium choice t_l is greater for firm T_h than for firm T_l: $\left(\eta_h^l, 1\right] \subset \left(\eta_h^h, 1\right]$. Hence under the D_1 criterion, upon observing out-of-equilibrium type announcement t_h consumers would assign belief $\eta_h = 1$ to the possibility of facing a type-T_h firm. But with $\eta_h = 1$ both types of firm would want to deviate to t_h to obtain profits $E_h(t_h, 1) = E_l(t_h, 1) = V_{hh}$, and hence this cannot be an equilibrium.

The second key point to highlight regarding pooling equilibria is that *pooling on t_h does exist for some parameter values, and does survive the D_1 criterion.* For expositional convenience, let us start by showing that the pooling on t_h equilibrium - if it exists - survives the D_1 criterion. Using the same logic as above, we note that the incentive to deviate from t_h to t_l is always stronger for firm T_l than for firm T_h. From the expressions in the main text we can derive

$$E_h(t_l,\eta_l) - E_h(t_h,p_0) = \gamma\left[V_{hl} - V_{hh}\right] + (1-\gamma)\left[E_l(t_l,\eta_l) - E_l(t_h,p_0)\right]. \tag{7}$$

Since $V_{hl} - V_{hh} < 0$, then clearly at $\eta_l = \eta_l^l$ we have $E_l(t_l,\eta_l) - E_l(t_h,p_0) = 0$ and $E_h(t_l,\eta_l) - E_h(t_h,p_0) < 0$, and hence it must be that $\eta_l^h > \eta_l^l$. In turn, this implies that the set of out-of-equilibrium beliefs that would make deviation to t_l strictly preferable to equilibrium choice t_h is greater for firm T_l than for firm T_h: $\left(\eta_l^h, 1\right] \subset \left(\eta_l^l, 1\right]$. Hence under the D_1 criterion, upon observing out-of-equilibrium type announcement t_l consumers would assign belief $\eta_h = 0$ to the possibility of facing a type-T_h firm. With $\eta_h = 0$ neither type of firm would want to deviate to t_l to obtain profits $E_h(t_l,0) = E_l(t_l,0) = V_{ll}$, and hence the pooling on t_h equilibrium, if it exists, does survive the D_1 criterion.

To determine existence conditions for this pooling on t_h equilibrium, consider condition

$$E_l(t_h,p_0) \geq E_l(t_l,0)\,, \text{which simplifies to } \gamma \leq 1 - \frac{1-p_0}{p_0}\frac{V_{ll}-V_{lh}}{V_{hh}-V_{ll}}. \tag{8}$$

Since out-of-equilibrium payoff $E_l(t_l,\eta_l)$ is strictly increasing in out-of-equilibrium belief η_l, condition (8), which stipulates that firm T_l prefers to play equilibrium strategy t_h when $\eta_l = 0$, is sufficient to ensure that $\eta_l^l \geq 0$, and that $E_l(t_h,p_0) \geq E_l(t_l,\eta_l)$ for all $\eta_l \in \left[0,\eta_l^l\right]$. Moreover, since as discussed above $\eta_l^h > \eta_l^l$, condition (8) is also sufficient to ensure that $E_h(t_h,p_0) > E_h(t_l,\eta_l)$ for all $\eta_l \in \left[0,\eta_l^h\right]$. Thus, condition (8) is sufficient to ensure that pooling on type announcement t_h is an equilibrium for all out-of-equilibrium beliefs $\eta_l \in \left[0,\eta_l^l\right]$.

Conversely, if condition (8) does not hold and $E_l(t_h,p_0) < E_l(t_l,0)$, there are no out-of-equilibrium beliefs $\eta_l \in [0,1]$ such that firm T_l would choose to play the equilibrium strategy t_h, and pooling on t_h cannot be an equilibrium. ■

2.3 Proof of Proposition 3

Semi-separating equilibria. Suppose that firm T_l makes type announcement t_j, $j \in \{h,l\}$ while firm T_h randomizes between announcements, announcing t_j with probability $m \in (0,1)$ and announcing t_z, $z \neq j$, with probability $1-m$. Then, upon observing type announcement t_j, consumers form interim belief

$$p_1(t_j, p_0) = \frac{P(t_j \mid T_h) p_0}{P(t_j \mid T_h) p_0 + P(t_j \mid T_l)(1-p_0)} = \frac{mp_0}{mp_0 + (1-p_0)}, \tag{9}$$

that the firm is of type T_h; while upon observing t_z, $z \neq j$, they form interim belief $p_1(t_z, p_0) = 1$. Based on these beliefs, firm T_l anticipates that its profits from a date 1 point of view will be

$$\begin{aligned} E_l(t_j, p_1(t_j, p_0)) &= \Pi_l(t_j, p_2(s_0, p_1(t_j, p_0))) \\ &= p_2(s_0, p_1(t_j, p_0)) V_{hj} + (1 - p_2(s_0, p_1(t_j, p_0))) V_{lj} \qquad (10) \\ &= \frac{(1-\gamma) p_1(t_j, p_0)}{(1-p_1(t_j, p_0)) + (1-\gamma) p_1(t_j, p_0)} V_{hj} + \frac{(1-p_1(t_j, p_0))}{(1-p_1(t_j, p_0)) + (1-\gamma) p_1(t_j, p_0)} V_{lj} \\ &= \frac{(1-\gamma) mp_0}{(1-p_0) + (1-\gamma) mp_0} V_{hj} + \frac{1-p_0}{(1-p_0) + (1-\gamma) mp_0} V_{lj} \end{aligned}$$

if it plays equilibrium strategy t_j; and $E_l(t_z, 1) = V_{hz}$ if it plays strategy t_z. Note that $E_l(t_j, p_1(t_j, p_0)) \in (V_{lj}, V_{hj})$.

Similarly, firm T_h anticipates date 1 payoff

$$\begin{aligned} E_h(t_j, p_1(t_j, p_0)) &= \gamma\, \Pi_h(t_j, p_2(s_h, p_1(t_j, p_0))) + (1-\gamma) \Pi_h(t_j, p_2(s_0, p_1(t_j, p_0))) \\ &= \gamma\, V_{hj} + (1-\gamma)\left[p_2(s_0, p_1(t_j, p_0)) V_{hj} + (1 - p_2(s_0, p_1(t_j, p_0))) V_{lj}\right] \qquad (11) \\ &= \gamma\, V_{hj} + (1-\gamma) E_l(t_j, p_1(t_j, p_0)) \\ &= \gamma\, V_{hj} + (1-\gamma)\left[\frac{(1-\gamma) mp_0}{(1-p_0) + (1-\gamma) mp_0} V_{hj} + \frac{1-p_0}{(1-p_0) + (1-\gamma) mp_0} V_{lj}\right] \end{aligned}$$

if it plays equilibrium strategy t_j; and $E_h(t_z, 1) = V_{hz}$ if it plays strategy t_z. Furthermore, since firm T_h randomizes between the two strategies, the equilibrium value m_{jz}^h of mixing probability m must be such that it is indifferent between the two strategies, i.e. such that $E_h(t_j, p_1(t_j, p_0)) = E_h(t_z, 1) = V_{hz}$.

There are two potential equilibria of this type to consider. First, let $j = l$ and $z = h$. In this case, firm T_l's equilibrium strategy t_l yields date 1 payoff $E_l(t_l, p_1(t_l, p_0)) \in (V_{ll}, V_{hl})$, while deviating to strategy t_h yields $E_l(t_h, 1) = V_{hh}$. Clearly, deviating is optimal for firm T_l, and hence this cannot be an equilibrium. Second, let $j = h$ and $z = l$. In that case, firm T_l's equilibrium strategy t_h yields date 1 payoff $E_l(t_h, p_1(t_h, p_0))$, while deviating to strategy t_l yields $E_l(t_l, 1) = V_{hl}$. However, equilibrium mixing strategy m_{hl}^h must be such that firm T_h's payoff from strategy t_h, $E_h(t_h, p_1(t_h, p_0)) = \gamma\, V_{hh} + (1-\gamma) E_l(t_h, p_1(t_h, p_0))$ is equal to its payoff $E_h(t_l, 1) = V_{hl}$ from strategy t_l. This in turn implies $E_l(t_h, p_1(t_h, p_0)) < V_{hl} = E_l(t_l, 1)$. Thus, deviating is optimal for firm T_l, and hence this cannot be an equilibrium.

Now suppose that firm T_h makes type announcement t_j, $j \in \{h, l\}$ while firm T_l randomizes between announcements, announcing t_j with probability $m \in (0,1)$ and announcing t_z, $z \neq j$, with probability $1 - m$. Then, upon observing type announcement t_j, consumers form interim belief

$$p_1(t_j, p_0) = \frac{P(t_j \mid T_h) p_0}{P(t_j \mid T_h) p_0 + P(t_j \mid T_l)(1 - p_0)} = \frac{p_0}{p_0 + m(1 - p_0)}, \tag{12}$$

that the firm is of type T_h; while upon observing t_z, $z \neq j$, they form interim belief $p_1(t_z, p_0) = 0$. Based on these beliefs, firm T_l anticipates that its profits from a date 1 point of view will be

$$\begin{aligned} E_l(t_j, p_1(t_j, p_0)) &= \Pi_l(t_j, p_2(s_0, p_1(t_j, p_0))) \\ &= p_2(s_0, p_1(t_j, p_0)) V_{hj} + (1 - p_2(s_0, p_1(t_j, p_0))) V_{lj} \qquad (13) \\ &= \frac{(1-\gamma) p_1(t_j, p_0)}{(1 - p_1(t_j, p_0)) + (1-\gamma) p_1(t_j, p_0)} V_{hj} + \frac{(1 - p_1(t_j, p_0))}{(1 - p_1(t_j, p_0)) + (1-\gamma) p_1(t_j, p_0)} V_{lj} \\ &= \frac{(1-\gamma) p_0}{m(1 - p_0) + (1-\gamma) p_0} V_{hj} + \frac{m(1 - p_0)}{m(1 - p_0) + (1-\gamma) p_0} V_{lj} \end{aligned}$$

if it plays equilibrium strategy t_j; and $E_l(t_z, 0) = V_{lz}$ if it plays strategy t_z. Note that $E_l(t_j, p_1(t_j, p_0)) \in (V_{lj}, V_{hj})$. Furthermore, since firm T_l randomizes between the two strategies, the equilibrium value m_{jz}^l of mixing probability m must be such that it is indifferent between the two strategies, i.e. such that $E_l(t_j, p_1(t_j, p_0)) = E_l(t_z, 0) = V_{lz}$.

Similarly, firm T_h anticipates date 1 payoff

$$\begin{aligned} E_h(t_j, p_1(t_j, p_0)) &= \gamma\, \Pi_h(t_j, p_2(s_h, p_1(t_j, p_0))) + (1-\gamma) \Pi_h(t_j, p_2(s_0, p_1(t_j, p_0))) \\ &= \gamma\, V_{hj} + (1-\gamma)\left[p_2(s_0, p_1(t_j, p_0)) V_{hj} + (1 - p_2(s_0, p_1(t_j, p_0))) V_{lj}\right] \qquad (14) \\ &= \gamma\, V_{hj} + (1-\gamma) E_l(t_j, p_1(t_j, p_0)) \\ &= \gamma\, V_{hj} + (1-\gamma)\left[\frac{(1-\gamma) p_0}{m(1 - p_0) + (1-\gamma) p_0} V_{hj} + \frac{m(1 - p_0)}{m(1 - p_0) + (1-\gamma) p_0} V_{lj}\right] \end{aligned}$$

if it plays equilibrium strategy t_j; and $E_h(t_z, 0) = V_{lz}$ if it plays strategy t_z.

There are two potential equilibria of this type to consider. First, let $j = l$ and $z = h$. In this case, firm T_l's strategy t_l yields date 1 payoff $E_h(t_l, p_1(t_l, p_0)) \in (V_{ll}, V_{hl})$, while strategy t_h yields $E_l(t_h, 1) = V_{lh}$. Clearly, since $V_{lh} < V_{ll}$, $E_l(t_h, 1) < E_h(t_l, p_1(t_l, p_0))$ for all $m \in (0,1)$: there is no value of m such that firm T_l is indifferent between strategies t_l and t_h, and hence this cannot be an equilibrium.

Second, let $j = h$ and $z = l$. In that case, firm T_l's strategy t_h yields date 1 payoff $E_l(t_h, p_1(t_h, p_0))$, while strategy t_l yields $E_l(t_l, 0) = V_{ll}$. Moreover, equilibrium mixing strategy m^l_{hl} must be such that firm T_l's payoffs from strategies t_h and t_l are equal, which implies $E_l(t_h, p_1(t_h, p_0)) = E_h(t_l, 0) = V_{ll} = E_h(t_z, 0)$. This in turn must means that $E_h(t_h, p_1(t_h, p_0)) = \gamma V_{hj} + (1-\gamma) E_l(t_j, p_1(t_j, p_0)) = \gamma V_{hj} + (1-\gamma) V_{ll} > V_{ll} = E_h(t_z, 0)$. In words, the equilibrium strategy t_h for firm T_h always dominates its deviation strategy t_l, and hence this semi-separating equilibrium is feasible, as long as an equilibrium mixing strategy $m^l_{hl} \in (0, 1)$ for firm T_h exists such that $E_l(t_h, p_1(t_h, p_0)) = E_h(t_l, 0)$. One can readily verify that the solution to this equation is

$$m^l_{hl} = (1-\gamma)\frac{p_0}{1-p_0}\frac{V_{hh} - V_{ll}}{V_{ll} - V_{lh}} > 0, \tag{15}$$

and that $m^l_{hl} < 1$ if and only if

$$\gamma > 1 - \frac{1-p_0}{p_0}\frac{V_{ll} - V_{lh}}{V_{hh} - V_{ll}}. \tag{16}$$

Thus, if condition (16) holds, there exists a semi-separating equilibrium in which firm T_h always makes type announcement t_h, and firm T_l randomizes between announcing t_h and t_l. Last but not least, the probability $1 - m^l_{hl}$ of telling the truth for firm T_l is strictly increasing in signal observability γ. ■

3 Empirical Appendix - Additional Statistics and Robustness Checks

3.1 Additional Statistics: Pearson Correlations

As noted in the main text, Propensity Score Reweighting (PSR) involves first using a logit regression of treatment on a selection of covariates and using the predicted value from the regression as estimated probability of treatment - or propensity score - for each observation in the control and treatment groups.

The choice of covariates for the logit regression depends in part on the existing literature (e.g. Konar and Cohen 2001; Krüger, 2015; Ioannou, Li, and Serafeim, 2016; Flammer, 2015a), and in part on the correlations between these covariates and *Env Reporting*. The correlations are reported in Table 1 below.

[Insert Table 1 here.]

3.2 Placebo Tests

We run placebo tests based on 1) a sample with pre-intervention period only and using years 2007 and 2008 as pseudo-intervention years; and 2) different dependent variables, specifically, and report-based CSR engagement on *non-environmental* activities. We report the results in Table 2.

[Insert Table 2 here.]

Columns (1) and (2), and (3) and (4), of Table 2 present the DiD regression results using the pre-intervention period sample (2005-2010), and with years 2007 and 2008 as intervention years, respectively. Across these four columns, the coefficient estimate on $Treat \times After$ is negative and insignificant. These results suggest that there is no evidence supporting a decrease in UK firms' reporting on environmental initiatives subsequent to the pseudo UK policy intervention years 2007 and 2008; and in turn that our observed decrease in UK firms' environmental engagement claims occurs after the UK policy intervention in 2011.

In columns (5) and (6), we turn our focus on using a pseudo outcome variable - firms' reporting on *non-environmental* engagement - constructed as the average of firms' reporting on social and governance related acclivities. The coefficient estimate on $Treat \times After$ is negative, and initially significant at the 10% level, but then becomes insignificant when firm fixed effects are controlled for. These results suggest that while (as shown above) a reduction in UK firms' CSR reporting due to the UK policy intervention holds for environmental reporting, results are weaker or insignificant for reporting on non-environmental activities. Taken together, the results in Table 2 lend further support for a negative effect of the UK policy intervention and its associated increase in private signal availability on firms' environmental CSR reporting.

3.3 Alternative Balancing Methods

In the main text we use *Propensity Score Reweighting* as our key pre-processing covariate balancing approach. One may be concerned that our results could be driven by the specific pre-processing technique we use. To mitigate this concern, in this section, we rerun our baseline regressions using three alternative balancing techniques: *entropy balancing*, *propensity score matching*, and *propensity score stratification.*

The entropy balancing approach relies on reweighing observations to minimize the difference between the treatment and control group. Specifically, entropy balancing technique selects observational weights to ensure that the first, second and third moments (mean, variance, skewness) of the distributions of our control variables (i.e., firm size, ROA, leverage, sales, R&D intensity, and market competition) are the same in the control group and the treatment group (Hainmueller, 2012), thus making the subsequent estimation of the treatment effects more reliable. We use this approach to form the entropy-balanced sample (EB sample).

We also calculate the propensity score (as discussed in the main text) - the scaled score of the conditional probability of receiving the treatment given the observed covariates - to balance the covariate between the treatment and control groups. In particular, we use two methods associated with propensity score - propensity score matching and propensity score stratification - to construct our propensity-score-matched sample (PSM sample) and propensity-score-stratification sample (PSS sample), respectively. Propensity scores are calculated for each observation in the sample based on a logit regression of the probability of being in the treatment group on a set of key covariates (i.e., firm size, ROA, leverage, sales, R&D intensity, and market competition).

Propensity score matching matches each UK firm in the treatment with non-UK firms in the control groups using single nearest-neighbor propensity score without replacement within a specified caliper width (Abadie and Imbens, 2011).[1] There are 234 matched treatment-control pairs after propensity score matching; and good country representation across the PSM sample, with no more than 17% of the sample from any one of the other 15 European countries.

Propensity score stratification divides the observations into strata that have similar propensity scores, with the objective of balancing the observed variables between treated and control units within each stratum. The treatment effect can then be estimated by combining stratum-specific estimates of treatment effect. Specifically, we divide our observations into 10 strata and we observe that there is a good balance between the treatment group and control group in all our covariates, i.e. firm size, ROA, leverage, sales, R&D intensity, and market competition.

Having constructed our EB sample, PSM sample, and PSS sample, we run the DiD regressions presented in the main text, and report the results in Table 3, where in all specifications standard errors are adjusted for within-firm clustering.

[Insert Table 3 here.]

Columns (1) and (2) report the results of the simplest DiD specification using the EB sample, with firm- and industry-level controls; and year, industry, country, and firm fixed effects gradually included. The coefficient estimate on $Treat \times After$ across the first two columns is negative and

[1] The specific caliper we use is 0.1 times the pooled standard deviation of the logit of the propensity score, which is argued to be the optimal caliper for matching (Rosenbaum and Robin, 1985).

significant at either the 1% or 10% significance level. Columns (3) and (4) present the results of DiD regression using the PSM sample. The coefficient estimate on $Treat \times After$ is negative and statistically significant at the 1% significance. Columns (5) − (6) report the results of the DiD regression based on the PSS sample. Again, the coefficient estimate on $Treat \times After$ remains negative and statistically significant at the 1% significance. Taken together these results offer further evidence of a statistically and economically strong causal impact of *The UK Companies Act* and its associated increase in signal availability on UK firms' reporting on their environmental CSR engagement.

We examine the parallel trends assumption by analyzing dynamic treatment effects of the UK policy intervention on report-based environmental CSR scores in the EB-DiD scenario, the PSM-DiD scenario, and the PSS-DiD scenario, respectively. We regress firms' environmental scores on a treatment group dummy, year dummies with 2010 as the benchmark year, as well as interactions between the treatment group dummy and each of the year dummies. The regressions also include our control variables, firm fixed effects, and standard errors adjusted for within-firm clustering. For clarity and space-saving purposes, we omit the regression table here, and instead report the treatment-year coefficient estimates graphically, along with 95% confidence intervals, in Figure A1 of this appendix.

[Insert Figure A1 here.]

Figure A1 illustrates statistically insignificant treatment-year coefficient estimates for all pre-intervention years (2005 - 2010), suggesting an overall parallel evolution of environmental score outcomes between treatment and control groups before the 2011 intervention year; and lending further support to the validity of the parallel trends assumption in this context. We also observe negative and significant treatment-year coefficient estimates for the post-intervention years (2011 - 2015), consistent with the negative treatment effect identified in Table 3 and discussed above.

3.4 Addressing Serial Correlation Concerns

Our DiD approach is based on multiple years of data and hence may be subject to serial correlation, underestimation of standard errors, and overestimation of t-statistics (Bertrand *et al.*, 2004). To address these concerns, in all our preceding DiD analyses, we clustered the standard errors at the firm level to account for the serial correlation within firms across multiple years. In this section, we take two further steps in addressing these concerns, and check whether our main results continue to hold a) when we collapse the data into two periods only and b) when we cluster standard errors at a higher - two-digit standard industrial classification (SIC) industry - level (Bertrand *et al.*, 2004; Athey and Imbens, 2017).

In our first step of addressing the concern related to serial correlation, we aggregate our data in into pre- and post- intervention periods to remove the time serial dimension of our data. In particular, for each variable in the treatment group and the control group, we calculate the average values across years in the pre-intervention period and in the post-intervention period. We then perform a DiD analysis using the constructed two-period sample, and we report the results in columns (1)-(3) of Table 4. Across these three columns, the coefficient estimate on $Treat \times After$ is negative and significant at either the 1% or 10% significance level, suggesting that the negative impact of signal availability on firms' CSR reporting on environmental initiatives continues to hold when we collapse the data into two periods to constrain the concern on serial correlation.

[Insert Table 4 here.]

Another solution to the serial correlation problem is go beyond clustering of standard errors at the firm level, and to cluster at levels higher than than the explanatory variable in the DiD approach (Angrist and Pischke, 2009). Following this approach, we cluster standard errors at the two-digit SIC industry level; allowing us to correct of any residual correlation within clusters - including the time series correlation. We report the results in columns (4) - (6) of Table 4. The coefficient estimate on $Treat \times After$ remains negative and significant at the 1% level, suggesting that our main result continue to hold when we make further corrections for potential serial correlation problems. Once again, the results provide support for a negative impact of signal availability on firms' reporting on environmental engagement.

3.5 Confounding Effects

One possible concern with the baseline analysis may be that the existence of confounding regulations in the other 15 European countries in our control group could preclude a proper identification of the impact of *The UK Companies Act* in the UK on CSR reporting on environmental initiatives. To address this issue, we first searched environmental regulations and policies in the other 15 European countries included in our control group,[2] and found four potentially relevant cases: (1) France: The

[2]We mainly obtained information on environmental regulations from the CSR disclosure efforts by national governments and stock exchanges from Harvard Kennedy School. http://iri.hks.harvard.edu/

Grenelle II Act was passed in 2012, which requires large firms to disclose their CSR information in their annual report. (2) Ireland: The carbon tax in Ireland started to cover almost all the polluting firms instead of only large emitters since 2012. (3) Norway: Legislation was passed on requiring large firms to disclose general CSR as well as GHG emissions around 2012. (4) Switzerland. The Swiss government implemented a regulation scheme in 2013 which provided explicit incentives for firms to reduce GHG emissions.

Having identified these potentially confounding regulations, we took out from our full sample the firms from France, Ireland, Norway, and Switzerland. We then perform the DiD analysis on the effect of signal availability on firms' environmental claims using this modified sample, and we report the results in columns (1) - (3) of Table 5.

[Insert Table 5 here.]

In column (1), the coefficient estimate on $Treat \times After$ is -4.977 and significant at the 1% significance level, with year and industry fixed effects included. In columns (2) and (3), we gradually include country and firm fixed effects. The coefficient estimate on $Treat \times After$ remains negative and significant at the 1% level. Thus, these results suggest that the negative impact of the signal availability on firms' CSR reporting on environmental engagement is robust to removing European firms with confounding policies from our sample.

Another possible confounding concern comes from the fact that firms in our sample may be covered by European Union Emissions Trading Scheme (EU ETS) and hence may be required to report their emissions to the EU ETS registry as well as financially incentivized to reduce their emissions through the cap-and-trade system under EU ETS. To exclude the potential impact of these firms' membership with EU ETS on their engagement in environmental initiatives, we re-ran the baseline model regressions, focusing on non-EU ETS firms only. To construct the non-EU ETS sample, we used firms' registration information provided by EU ETS registry and manually matched the names of the account holders under EU ETS registry with those of the firms in our sample. We report the results in columns (4) - (6) of Table 5. As shown in this table, the coefficient estimate on $Treat \times After$ is negative and significant at the 1% significance level in columns (4) and (5); the exception is for column (6) in which the coefficient estimate becomes insignificant when firm fixed effects are included.

Overall, the results in Table 5 suggest that a negative effect of signal availability on CSR reporting on environmental initiatives continues to hold when we take out from our sample the firms from countries which had confounding policies around 2011, and the firms registered with the EU ETS; and alleviate concerns that our baseline results might be driven by other policy changes.

3.6 Alternative Intervention Years and Lagged Dependent Variable

One could also be concerned that our main results may be driven by the particular intervention year we use. To address this concern, in this section, we first check whether or not our main results presented in Table 3 of the main body of the paper are robust to using year 2012 as the intervention year. We re-run the baseline regressions using 2012 as the cutoff year, and we report the results in columns (1) and (2) of Table 6. To further mitigate the impact of the choice of intervention years, we drop years 2011 and 2012 and rerun the baseline regressions, and report the results in columns (3) and (4) of Table 6.

[Insert Tables 6 here.]

Across the first four columns in Tables 6, the coefficient estimate on $Treat \times After$ is negative and significant at the 1% significance level, suggesting that our main results continue to hold if we use 2012 as the alternative intervention year, or drop observations during the intervention years 2011 and 2012 from our sample.

In all preceding DiD analyses, we controlled for firm fixed effects in an attempt to mitigate concerns related to omitted variables. The underlying assumption for this consideration is that the omitted variables are time-invariant and hence controlling for firm fixed effects can help address the concerns related to omitted variables. However, if the omitted variables are time-varying, controlling for firm fixed effects can only partially mitigate the issues related to omitted variables (Angrist and Pischke, 2009). To account for the impact of both time-invariant and time-varying omitted variables, we include both firm fixed effects and a lagged dependent variable, i.e., a lagged report-based environmental CSR score, into our regression.

The results of the baseline regressions with added lagged environmental reporting are presented in columns (5) and (6) of Table 6. The coefficient estimate on $Treat \times After$ in columns (5) and (6) is negative and significant at either the 1% or 5% significance level. These results suggest that the negative effect of signal availability on firms' CSR reporting on environmental initiatives continue to hold when controlling for a lagged dependent variable, which again provides support for our model predictions.

3.7 Alternative Dependent Variable and Sample

In the main text, we use the environmental score from Refinitiv ESG's database - which is "based on company-reported data"(Refinitiv, 2022) and includes information collected from CSR reports, CSR sections in annual reports, company websites, and stock exchange filings, for example - as our main measure of firms' reported CSR engagement in environmental activities and our main dependent variable. In this section, we consider an alternative type of dependent variable and an alternative sample.

Alternative measure of CSR reporting on environmental engagement. As a robustness check, we examine whether our main results continue to hold when we use alternative ways to measure firms' reporting on environmental initiatives. In particular, our main dependent variable, *Env Reporting*, is constructed by aggregating firms' engagement in three key sub-dimensional environmental activities, namely environmental-innovation activities, resource-use activities, and emissions-reduction activities. To gain insights on whether and how signal availability affects these three sub-dimensions of environmental reporting separately, we re-run the baseline regressions using each of these measures of reported environmental engagement separately as dependent variables. We report the results in Table 7.

[Insert Table 7 here.]

Columns (1) and (2), (3) and (4), and (5) and (6) present the results with environmental-innovation score, resource-use score, and emissions-reduction score, respectively, as dependent variable. Across all but one columns, we observe a negative and significant coefficient estimate on $Treat \times After$.[3] These results suggest that the negative effect of signal availability on firms' CSR reporting on environmental initiatives is robust to alternative dependent variables.

Alternative sample on firms with publishing CSR/sustainability reports. One may also worry that our main measure of reported CSR engagement may not perfectly capture the self-reported nature of the CSR engagement. To address this concern, we re-run the baseline regressions using a subsample that includes only firm-year observations associated with firms publishing a separate report, or a separate section in its annual report, on CSR, health and safety (H&S), or sustainability, in a given year. This subsample includes 79% of the observations present in the

[3] The only exception is in the last column which considers the impact of signal availability on emissions reduction activities when firm fixed effects are controlled. This is not surprising result because compared to the other two dimensions (environmental innovation and resource use), the category of emissions reduction involves concrete measures that are relatively easier for the public to verify.

original sample.[4] We report the results in Table 8.

[Insert Table 8 here.]

Column (1) reports the results of the basic DiD specification, while in columns (2) - (6), we sequentially introduce firm-level and industry-level controls, year fixed effects, industry fixed effects, country fixed effects, and firm fixed effects. Across all columns, our main coefficient estimate of interest - the coefficient on $Treat \times After$ - remains negative and significant either at the 1% or 5% significance level. We further examine the parallel trends assumption by analyzing dynamic treatment effects of the UK policy intervention on report-based CSR scores using this alternative dependent variable. These results, omitted here to save space, suggest an overall parallel evolution of environmental score outcomes between treatment and control groups before the 2011 (2005 - 2010), and negative and significant treatment-year coefficient estimates for the post-intervention years (2011 - 2015). These results are consistent with the negative treatment effect identified in Table 8.

3.8 Different Dataset on CSR Reporting: Sustainalytics

One may also worry that our main empirical result in the baseline analysis - the negative effect of signal availability on CSR reporting on environmental initiatives - may be driven by the specifics of the Refinitiv ESG database used in our analysis so far. We addressed this point by re-running the regressions from Table 3 of the main body of the paper with a different ESG data set - Sustainalytics - and checking whether our main results continue to hold.

Sustainalytics is a company that rates the sustainability of listed companies based on their ESG performance, and provides ESG and Corporate Governance research and ratings globally, from 76 countries over the 2009 - 2016 time period. Sustainalytics measures ESG along the usual environmental, social, and governance categories, and we aggregated the scores from these three dimensions to get overall ESG scores - our measure of ESG - for firms in the UK and in the 15 other European countries we consider, during the period 2009 - 2015. Importantly, the Sustainalytics scores are also based on firm-reported information.

[4] To be specific, in the original sample environmental CSR scores are based on various types of information disclosed by firms, such as information obtained from company websites, from stock exchange filings, from news sources, or from non-governmental organizations (NGOs); *and possibly* on information included in separately published reports or sections of their annual report, on CSR/H&S/sustainability. In contrast, in this subsample environmental CSR scores are also based on various types of information disclosed by firms, such as information obtained from company websites, stock filings, etc.; but are also based *with certainty* on information included in separately published reports or sections of their annual report, on CSR/H&S/sustainability.

We re-ran the baseline regressions with the Sustainalytics data, and report the results in Table 9, where in all specifications standard errors are adjusted for within-firm clustering.

[Insert Table 9 here.]

The results are very similar to our baseline results obtained using the Refinitiv ESG database. Across the first three columns in Table 9, the coefficient estimates on $Treat \times After$ are negative and significant at the 1% significance level, controlling for year, industry, and country fixed effects. In columns (4) - (6), we further include firm- and industry-level control variables, country fixed effects, and firm fixed effects, and the coefficient estimates on $Treat \times After$ remain negative and significant. Indeed, these results are consistent with a negative effect of private signal availability on environmental reporting in organizations; and mitigate the concern that our baseline results might be specific to the Refinitiv ESG database.

References

[1] Abadie, Alberto and Guido Imbens. 2011.“Bias-Corrected Matching Estimators for Average Treatment Effects.” *Journal of Business & Economic Statistics* **29 (1)**:1-11.

[2] Angrist, Joshhua D., and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press, 1st edition.

[3] Athey, Susan and Guido W. Imbens. 2017. “The State of Applied Econometrics: Causality and Policy Evaluation.” *Journal of Economic Perspective* **31 (2)**: 3-32.

[4] Bothello, J., Ioannou, I., Porumb, V.- A., and Zengin-Karaibrahimoglu, Y. 2023. “CSR decoupling within business groups and the risk of perceived greenwashing. ” *Strategic Management Journal* **44:** 3217–3251.

[5] Bertrand, Marianne, Esther Duflo, and Sendhil Muliainathan. 2014. “How much should we trust differences-in-differences estimates? ” *Quarterly Journal Economics* **February:** 249-275.

[6] Flammer, Caroline. 2015. “Does Product Market Competition Foster Corporate Social Responsibility? Evidence from Trade Liberalization.” *Strategic Management Journal* **36**: 1469-1485.

[7] Hainmueller, J. 2012. “Entropy Balancing for Causal Effects: A multivariate reweighing method to produce balanced samples in observational studies.” *Political Analysis* **20 (1)**: 25-46.

[8] Heckman James J., Lochner Lance C. and Petra E. Todd. 2006. “Earnings Functions, Rates of Return and Treatment Effects: the Mincer Equation and Beyond.” In: Hanushek E, Welch F (eds) Handbook of the economics of education, vol 1. Ch 7. North-Holland, Amsterdam, 307-458.

[9] Ioannou, Ioannis, Shelley Xin Li, and George Serafeim. 2016. “The Effect of Target Difficulty on Target Completion: The Case of Reducing Carbon Emissions.” *The Accounting Review* **91(5)**: 1467–1492.

[10] Konar, Shameek and Mark A. Cohen. 2001. “Does the Market Value Environmental Performance?” *Journal of Economic Literature* **83(2)**: 281-289.

[11] Krüger, Philipp, 2015. “Climate Change and Firm Valuation: Evidence from a Quasi-Natural Experiment.” Working Paper, Swiss Finance Institute.

[12] Refinitiv. 2022. “Environmental, Social and Governance Scores from Refinitiv,” Report, LSEG, London, UK. https://www.lseg.com.

[13] Rosenbaum, Paul R. and Donald B. Rubin. 1985. “Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. ” *The American Statistician* **39 (1)**: 33-38.

Table 1: Pearson Correlations

This table presents the Pearson correlations between our variables, namely *Env Reporting*, *Non-Env Reporting*, *Emissions-Reduction Score*, *Resource-Use Score*, *Environmental-Innovation Score*, *Firm Size*, *ROA*, *Leverage*, *Sales*, *R&D Intensity*, and *Market Competition*.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) Env Reporting	1										
(2) Non-Env Reporting	0.66***	1									
(3) Emissions-Reduction Score	0.87***	0.64***	1								
(4) Resource-Use Score	0.88***	0.65***	0.79***	1							
(5) Env-Innovation Score	0.70***	0.42***	0.46***	0.48***	1						
(6) Firm Size	0.53***	0.44***	0.43***	0.42***	0.38***	1					
(7) ROA	-0.14***	-0.04***	-0.06***	-0.06***	-0.13***	-0.32***	1				
(8) Leverage	0.05***	0.02**	0.05***	0.04***	-0.02*	0.11***	-0.09***	1			
(9) Sales	0.11***	0.21***	0.21***	0.21***	0.11***	-0.05***	0.34***	-0.02**	1		
(10) R&D Intensity	-0.06***	-0.01	-0.06***	-0.03***	0.06***	-0.15***	0.08***	-0.20***	0.16***	1	
(11) Market Competition	0.11***	0.03***	0.03***	-0.02	0.03***	0.33***	-0.27***	0.14***	-0.54***	-0.11***	1

Table 2: Private Signal Availability and Reporting on Environmental Initiatives: Placebo Tests

This table presents our placebo tests results. In columns (1) - (4), the dependent variable is *Env Reporting*, and the sample period is the pre-intervention period (2005-2010). Columns (1) and (2), and (3) and (4), present the results when 2007 and 2008 are used, as pseudo-intervention years, respectively, Columns (5) and (6) show the results using $Non - Env\ Reporting$ as pseudo-dependent variable, with the baseline sample period of 2005-2015 and 2011 as our intervention year. $Treat \times After$ is the interaction term that provides the DiD estimate. Standard errors are reported in parentheses below the coefficient estimates and adjusted for within-firm clustering. Significance at the 10%, 5%, and 1% level is indicated by *,**, ***, respectively.

	Pseudo Cutoff (2007)		Pseudo Cutoff (2008)		Pseudo DV(Non-Env Rep.)	
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Treat	24.586***		24.259***		17.906***	
	(3.227)		(3.172)		(2.352)	
$Treat \times After$	-1.993	-1.049	-1.964	-1.099	-1.465*	-0.716
	(1.467)	(1.561)	(1.304)	(1.361)	(0.816)	(0.833)
Firm Size	9.653***	1.447	9.658***	1.464	7.594***	2.233**
	(0.754)	(1.719)	(0.754)	(1.723)	(0.587)	(0.913)
ROA	18.036***	-4.812	18.128***	-4.707	13.120***	2.048
	(6.758)	(6.533)	(6.766)	(6.530)	(4.248)	(3.775)
Leverage	-0.864	-1.403	-0.896	-1.452	-2.188	-1.502
	(4.126)	(4.749)	(4.126)	(4.755)	(2.650)	(2.493)
Sales	1.801**	0.110	1.800**	0.125	0.591	1.215*
	(0.789)	(1.201)	(0.790)	(1.202)	(0.670)	(0.634)
R&D Intensity	62.462***	-34.200	62.392***	-34.530	35.998*	9.831
	(22.547)	(27.291)	(22.557)	(27.354)	(19.951)	(16.693)
Market Competition	0.508	2.105	0.519	2.086	-2.412	0.617
	(4.705)	(5.008)	(4.701)	(5.003)	(2.628)	(2.472)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	No	Yes	No	Yes	No
Country FE	Yes	No	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes	No	Yes
Observations	4,583	4,583	4,583	4,583	8,783	8,783
R-squared	0.635	0.859	0.635	0.859	0.505	0.795
Clusters	943	943	943	943	1133	1133

Table 3: Private Signal Availability and Reporting on Environmental Initiatives: Alternative Balancing Methods

This table presents coefficients estimates of regressions which examine the effect of signal availability on a firm's reporting on its environmental initiatives, using entropy-balanced sample (EB sample), propensity-score-matched sample (PSM sample), and propensity score stratification sample (PSS sample). We use *The UK Companies Act* as an exogenous increase in private signal availability, with 2011 as our intervention year. The sample period is 2005-2015. The dependent variable is *Env Reporting*. $Treat$ is a dummy variable equal to 1 for UK-listed firms and to 0 for EU firms. $After$ is a dummy variable equal to 0 for all years in the pre-intervention period (2005-2010), and to 1 for all years in the post-intervention period (2011-2015). $Treat \times After$ is the interaction term that provides the DiD estimate. Columns (1) and (2), (3) and (4), and (5) and (6) present the results using EB sample, PSM sample, and PSS sample, respectively. Standard errors are reported in parentheses below the coefficient estimates and adjusted for within-firm clustering. Significance at the 10%, 5%, and 1% level is indicated by *,**, ***, respectively.

	EB Sample		PSM Sample		PSS Sample	
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Treat	15.775***		22.911***		20.428***	
	(3.335)		(5.855)		(2.853)	
$Treat \times After$	-4.468***	-2.619*	-4.600***	-4.355***	-4.681***	-3.179***
	(1.470)	(1.508)	(1.423)	(1.439)	(1.101)	(1.089)
Firm Size	7.916***	2.334*	8.388***	1.556	6.291***	1.529
	(0.799)	(1.388)	(1.046)	(1.349)	(1.215)	(1.432)
ROA	7.006	-2.852	19.820**	-1.592	15.510***	1.255
	(6.705)	(6.069)	(8.822)	(8.281)	(5.860)	(5.430)
Leverage	-0.435	0.054	-1.293	-2.174	-1.849	-0.853
	(3.783)	(3.457)	(5.628)	(4.280)	(3.793)	(3.250)
Sales	2.113***	1.412	2.652**	2.101	1.644**	0.897
	(0.773)	(1.052)	(1.080)	(1.335)	(0.736)	(0.742)
R&D Intensity	52.212**	-13.187	88.137***	-3.072	27.515	-40.803
	(21.371)	(18.766)	(33.719)	(33.078)	(26.788)	(25.214)
Market Competition	-4.899	-2.105	0.731	2.198	0.495	0.853
	(3.735)	(4.354)	(5.066)	(4.154)	(3.535)	(3.402)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	No	Yes	No	Yes	No
Country FE	Yes	No	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes	No	Yes
Observations	8,783	8,783	4,040	4,040	8,783	8,783
R-squared	0.625	0.846	0.659	0.849	0.621	0.851
Clusters	1,133	1,133	465	465	1,133	1,133

Table 4: Signal Availability and Reporting on Environmental Initiatives: Addressing Serial Correlation

This table presents coefficients estimates of regressions which examine the effect of private signal availability on a firm's reporting on its environmental initiatives, when collapsing data into two periods (pre- and post-intervention), and when clustering standard errors at industry level and country level. The dependent variable is *Env Reporting*. $Treat \times After$ is the interaction term that provides the DiD estimate. Columns (1) - (3) present the results showing the impact of signal availability on reporting on environmental initiatives for a sample with pre- and post- intervention periods only. Columns (4) - (6) report the results with standard errors clustered at industry level. The sample period is from 2005-2015. Standard errors are reported in parentheses below the coefficient estimates and adjusted for within-firm clustering in columns (1) - (3). Significance at the 10%, 5%, and 1% level is indicated by *,**, ***, respectively.

	Two Periods			Industry Clustering		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Treat	8.502***	19.807***		8.541***	20.457***	
	(1.547)	(2.944)		(2.207)	(3.958)	
$Treat \times After$	-6.279***	-6.512***	-2.676*	-4.736***	-4.921***	-3.179***
	(1.223)	(1.212)	(1.520)	(1.076)	(1.022)	(1.027)
Firm Size	8.481***	9.065***	2.319	8.352***	8.920***	2.605*
	(0.710)	(0.688)	(2.065)	(1.208)	(1.149)	(1.343)
ROA	11.671	15.332*	-3.402	11.841	13.573*	-1.443
	(8.440)	(8.272)	(14.054)	(7.525)	(6.999)	(5.533)
Leverage	-2.503	-4.591	1.460	1.931	-0.766	0.181
	(4.594)	(4.411)	(8.080)	(4.761)	(3.952)	(3.258)
Sales	1.544*	1.900**	2.085	1.641	2.061*	1.176
	(0.798)	(0.774)	(1.340)	(1.091)	(1.047)	(0.837)
R&D Intensity	54.104*	60.429**	-81.098	61.497***	68.250***	-21.783
	(29.374)	(26.677)	(67.467)	(15.146)	(21.872)	(28.009)
Market Competition	-7.721	-6.217	1.575	-0.184	0.400	1.451
	(5.833)	(5.576)	(8.511)	(3.681)	(3.578)	(3.417)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	No	Yes	Yes	No
Country FE	No	Yes	No	No	Yes	No
Firm FE	No	No	Yes	No	No	Yes
Observations	2,023	2,023	2,023	8,783	8,783	8,783
R-squared	0.602	0.641	0.947	0.582	0.618	0.851
Clusters	1,185	1,185	1,185	62	62	62

Table 5: Private Signal Availability and Reporting on Environmental Initiatives: Confounding Effects

This table presents coefficients estimates of regressions which examine the effect of signal availability on firms' reporting on environmental initiatives. We repeat the baseline regression but drop observations from the following four countries (France, Norway, Italy, and Switzerland) in columns (1) - (3), and take out firms associated with EU Emission Trading System (ETS) in columns (4) - (6). The dependent variable is *Env Reporting*. Standard errors are reported in parentheses below the coefficient estimates and adjusted for within-firm clustering. Significance at the 10%, 5%, and 1% level is indicated by *,**, ***, respectively.

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Treat	8.937***	19.645***		10.305***	16.848***	
	(1.777)	(2.932)		(2.010)	(4.512)	
$Treat \times After$	-4.977***	-5.117***	-3.706***	-4.292***	-4.382***	-2.375
	(1.198)	(1.192)	(1.172)	(1.649)	(1.649)	(1.620)
Firm Size	7.274***	8.158***	1.441	9.288***	9.289***	3.009
	(0.799)	(0.760)	(1.169)	(0.797)	(0.806)	(1.880)
ROA	13.642*	13.788**	1.596	7.175	9.989	2.971
	(6.956)	(6.497)	(5.406)	(9.339)	(8.951)	(7.668)
Leverage	3.737	0.163	0.180	-1.129	-1.773	0.532
	(4.406)	(4.248)	(3.553)	(5.323)	(5.043)	(4.995)
Sales	2.062**	2.532***	1.531**	1.374	1.483	0.530
	(0.821)	(0.780)	(0.669)	(0.963)	(0.911)	(1.296)
R&D Intensity	92.854***	92.909***	-13.002	89.623	64.906	-70.923
	(31.583)	(26.367)	(22.770)	(65.433)	(50.377)	(45.865)
Market Competition	0.290	0.743	3.406	1.482	1.713	0.603
	(4.263)	(4.304)	(4.025)	(5.474)	(5.480)	(5.369)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	No	Yes	Yes	No
Country FE	No	Yes	No	No	Yes	No
Firm FE	No	No	Yes	No	No	Yes
Observations	6,799	6,799	6,799	4,109	4,109	4,109
R-squared	0.581	0.614	0.849	0.578	0.614	0.848
Clusters	906	906	906	577	577	577

Table 6: Private Signal Availability and Reporting on Environmental Initiatives: Alternative Years and Controlling for Lagged Dependent Variable

This table presents coefficients estimates of regressions which examine the effect of private signal availability on firms' reporting on its environmental initiatives, using 2012 as the intervention year associated with *The UK Companies Act* in columns (1) and (2), dropping years 2011 and 2012 from the sample in columns (3) and (4), and controlling for lagged dependent variable in columns (5) and (6). The dependent variable is *Env Reporting*. $Treat \times After$ is the interaction term that provides the DiD estimate. The sample period is from 2005 - 2015. Standard errors are reported in parentheses below the coefficient estimates and adjusted for within-firm clustering. Significance at the 10%, 5%, and 1% level is indicated by *,**, ***, respectively.

	Cutoff Year 2012		Dropping 2011 and 2012		Controlling Lagged DV	
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Treat	19.929***		20.859***		3.153***	
	(2.828)		(2.791)		(0.794)	
$Treat \times After$	-4.711***	-2.982***	-5.723***	-3.537***	-1.179**	-1.655***
	(1.081)	(1.071)	(1.234)	(1.263)	(0.460)	(0.618)
Firm Size	8.911***	2.607**	8.798***	2.418**	1.655***	1.879***
	(0.680)	(1.166)	(0.682)	(1.232)	(0.223)	(0.664)
ROA	13.463**	-1.588	14.420**	-2.335	5.101**	3.979
	(5.816)	(5.385)	(5.911)	(5.975)	(2.057)	(3.636)
Leverage	-0.672	0.294	-1.138	0.321	-0.005	0.413
	(3.627)	(3.254)	(3.519)	(3.530)	(1.161)	(2.172)
Sales	2.070***	1.171	2.162***	1.556**	0.598***	0.829*
	(0.700)	(0.724)	(0.704)	(0.752)	(0.206)	(0.468)
R&D	68.072***	-22.592	66.143***	-25.185	20.225***	11.330
	(21.673)	(22.027)	(21.751)	(24.991)	(6.666)	(15.640)
Market Competition	0.558	1.540	2.232	1.735	-1.616	-1.761
	(3.540)	(3.410)	(3.630)	(3.570)	(1.568)	(1.969)
Lagged Env Reporting					0.795***	0.495***
					(0.009)	(0.014)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	No	Yes	No	Yes	No
Country FE	Yes	No	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes	No	Yes
Observations	8,783	8,783	7,139	7,139	7,709	7,709
R-squared	0.618	0.851	0.621	0.844	0.870	0.903
Clusters	1,133	1,133	1,129	1,129	1,014	1,014

Table 7: Private Signal Availability and Reporting on Environmental Initiatives: Alternative Dependent Variable

This table presents coefficients estimates of regressions which examine the effect of private signal availability on a firm's reporting on three sub-dimensions of environmental initiatives - environmental innovation, resource use, and emissions reduction. We use *The UK Companies Act* as an exogenous increase in private signal availability, with 2011 as our intervention year. The dependent variables for columns (1) and (2), (3) and (4), and (5) and (6) are firms' environmental-innovation score, resource-use score, and emissions-reduction score, respectively. $Treat$ is a dummy variable equal to 1 for UK-listed firms and to 0 for EU firms. $After$ is a dummy variable equal to 0 for all years in the pre-intervention period (2005-2010), and to 1 for all years in the post-intervention period (2011-2015). $Treat \times After$ is the interaction term that provides the DiD estimate. The sample period is from 2005-2015. Standard errors are reported in parentheses below the coefficient estimates and adjusted for within-firm clustering. Significance at the 10%, 5%, and 1% level is indicated by *,**, ***, respectively.

	Env-Innovation Score		Resource-Use Score		Emissions-Reduction Score	
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Treat	10.080***		25.787***		25.127***	
	(2.890)		(3.618)		(3.224)	
$Treat \times After$	-5.343***	-4.559***	-6.825***	-4.824***	-2.888**	-0.944
	(1.562)	(1.680)	(1.402)	(1.413)	(1.351)	(1.380)
Firm Size	7.783***	1.669	10.368***	2.015	10.682***	4.470***
	(0.915)	(1.507)	(0.887)	(1.390)	(0.767)	(1.263)
Leverage	-4.735	4.124	-1.366	-1.628	-2.713	-0.466
	(4.181)	(4.360)	(4.481)	(4.328)	(4.336)	(4.032)
ROA	0.278	4.540	13.894*	-4.803	19.399***	4.212
	(7.096)	(7.610)	(7.290)	(6.712)	(7.022)	(6.754)
Sales	-0.083	-0.197	1.910**	0.916	1.933**	1.590
	(0.873)	(1.118)	(0.907)	(0.783)	(0.845)	(1.014)
R&D Intensity	76.276***	-23.597	72.090***	-40.737	54.243**	-6.920
	(29.280)	(29.536)	(23.777)	(26.996)	(25.041)	(27.523)
Market Competition	-3.986	-3.472	-4.073	0.235	2.235	3.867
	(5.110)	(5.548)	(4.497)	(4.222)	(4.005)	(3.935)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	No	Yes	No	Yes	No
Country FE	Yes	No	Yes	No	Yes	No
Firm FE	No	Yes	No	Yes	No	Yes
Observations	8,777	8,777	8,773	8,773	8,773	8,773
R-squared	0.501	0.726	0.538	0.808	0.551	0.808
Clusters	1,132	1,132	1,132	1,132	1,132	1,132

Table 8: Private Signal Availability and Reporting on Environmental Initiatives: Focus on Subsample of Firms that Publish Separate CSR Reports

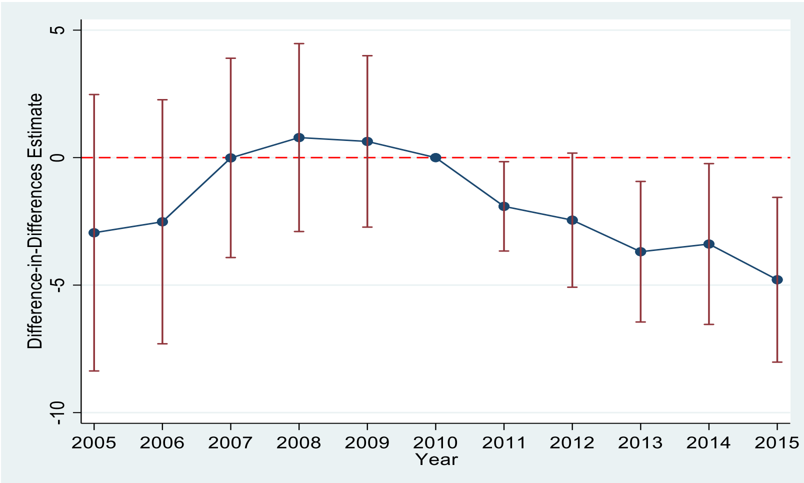
This table presents coefficients estimates of the baseline regressions examining the effect of private signal availability on a firm's reported environmental CSR engagement using a subsample that includes only firm-year observations associated with the firm publishing a separate report, or a separate section in its annual report, on CSR, health and safety, or sustainability, in a given year. This subsample includes 79% or the observations present in the original sample. We use *The UK Companies Act* as an exogenous increase in private signal availability, with 2011 as our intervention year. We construct a UK-EU sample of UK firms and firms from 15 European countries over the 2005-2015 time period. The dependent variable is a dummy variable equal to 1 if the firm publishes a CSR/sustainability report, either within its annual report or separately from them; and to 0 otherwise.. $Treat$ is a dummy variable equal to 1 for UK-listed firms and to 0 for EU firms. $After$ is a dummy variable equal to 0 for all years in the pre-intervention period (2005-2010), and to 1 for all years in the post-intervention period (2011-2015). $Treat \times After$ is the interaction term that provides the DiD estimate. All variables are defined in Table 1. Standard errors are reported in parentheses below the coefficient estimates and adjusted for within-firm clustering. Significance at the 10%, 5%, and 1% level is indicated by *,**, ***, respectively.

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Treat	-7.338***	3.093**	3.086**	2.219	14.911***	
	(1.624)	(1.427)	(1.446)	(1.464)	(2.643)	
After	1.710**	3.389***				
	(0.854)	(0.763)				
$Treat \times After$	-4.821***	-4.132***	-4.029***	-3.343***	-3.901***	-2.724**
	(1.423)	(1.283)	(1.282)	(1.197)	(1.178)	(1.106)
Firm Size		6.575***	6.687***	7.400***	8.132***	3.763***
		(0.308)	(0.311)	(0.815)	(0.808)	(1.311)
ROA		-11.611	-8.321	-0.780	2.950	9.357
		(7.416)	(7.461)	(7.470)	(7.070)	(5.961)
Leverage		5.288	5.289	2.628	-0.140	5.931
		(3.320)	(3.343)	(3.868)	(3.784)	(3.773)
Sales		0.621***	0.634***	0.430	0.963	-0.511
		(0.182)	(0.183)	(0.890)	(0.905)	(1.026)
R&D Intensity		19.582	19.177	79.713***	71.186***	-6.670
		(16.772)	(17.051)	(25.108)	(22.472)	(25.036)
Market Competition		2.388	2.279	-7.896*	-7.531*	-4.338
		(2.184)	(2.199)	(4.052)	(4.030)	(3.454)
Year FE	No	No	Yes	Yes	Yes	Yes
Industry FE	No	No	No	Yes	Yes	No
Country FE	No	No	No	No	Yes	No
Firm FE	No	No	No	No	No	Yes
Observations	6,017	5,758	5,758	5,758	5,758	5,758
R-squared	0.047	0.328	0.344	0.558	0.587	0.857
Clusters	933	911	911	911	911	911

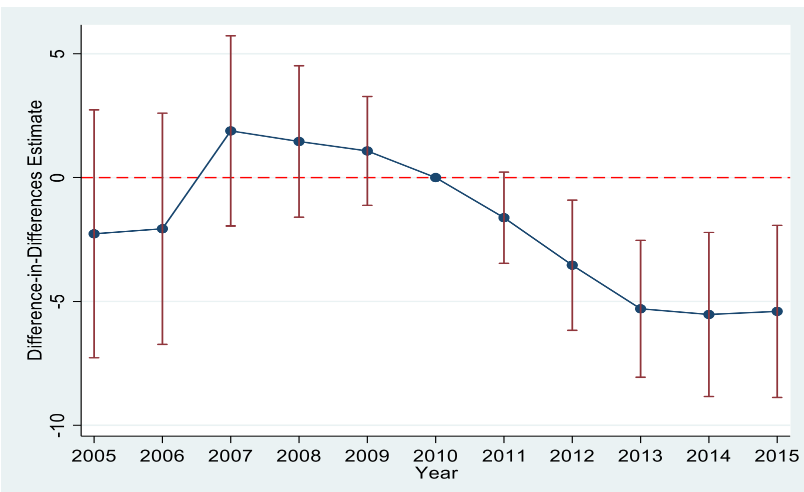
Table 9: Private Signal Availability and Reporting on Environmental Initiatives: Different CSR Data - Sustainalytics

This table presents coefficients estimates of regressions which examine the effect of private signal availability on firms' reporting on environmental initiatives, with different CSR data based on the Sustainalytics dataset. We use *The UK Companies Act* as an exogenous increase in private signal availability, with 2011 as our intervention year. The dependent variable is *EnvReporting*. *Treat* is a dummy variable equal to 1 for UK-listed firms and to 0 for EU firms. *After* is a dummy variable equal to 0 for all years in the pre-intervention period (2005-2010), and to 1 for all years in the post-intervention period (2011-2015). $Treat \times After$ is the interaction term that provides the DiD estimate. The sample period is from 2009 - 2015. Standard errors are reported in parentheses below the coefficient estimates and adjusted for within-firm clustering. Significance at the 10%, 5%, and 1% level is indicated by *,**, ***, respectively.

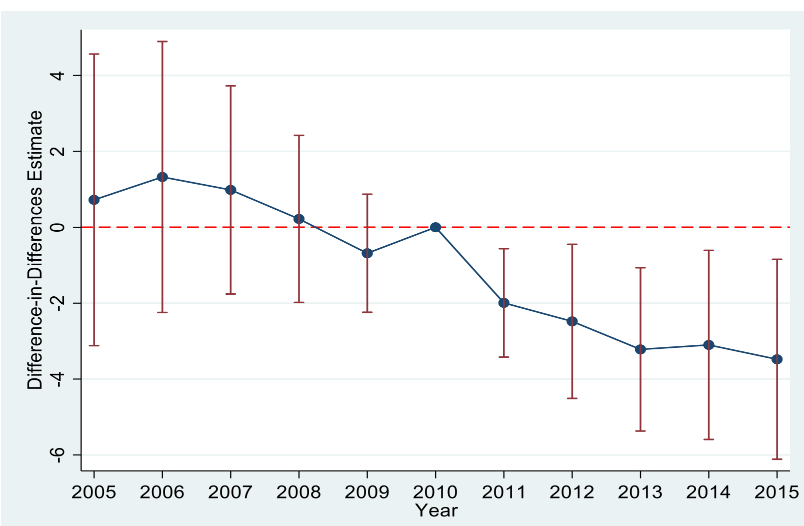
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Trea	4.616***	4.631***	2.736***	2.089***	2.431**	
	(0.876)	(0.877)	(0.828)	(0.744)	(0.978)	
After	1.923***	3.209***				
	(0.399)	(0.523)				
$Treat \times After$	-2.479***	-2.437***	-2.156***	-1.416**	-1.703***	-1.026*
	(0.777)	(0.777)	(0.652)	(0.609)	(0.605)	(0.551)
Firm Size				2.547***	2.577***	-0.161
				(0.389)	(0.376)	(0.343)
ROA				0.842	0.806	2.113
				(3.609)	(3.417)	(3.615)
Leverage				-2.056	-2.852	-1.274
				(1.952)	(1.968)	(1.785)
Sales				0.254	0.339	0.547
				(0.393)	(0.383)	(0.361)
R&D Intensity				29.125*	29.830*	4.660
				(16.044)	(16.255)	(12.368)
Industry Competition				-0.848	-0.445	0.463
				(1.527)	(1.569)	(1.594)
Year FE	No	Yes	Yes	Yes	Yes	Yes
Industry FE	No	No	Yes	Yes	Yes	No
Country FE	No	No	No	Yes	Yes	No
Firm FE	No	No	No	No	No	Yes
Observations	4,943	4,943	4,943	4,943	4,943	4,943
R-squared	0.029	0.038	0.411	0.509	0.531	0.917
Clusters	1,027	1,027	1,027	1,027	1,027	1,027



<u>Figure A1a:</u> Dynamic Treatment Effects of Private CSR-Engagement Signal Availability on CSR-Engagement Reporting (UK-EU EB sample)



<u>Figure A1b:</u> Dynamic Treatment Effects of Private CSR-Engagement Signal Availability on CSR-Engagement Reporting (UK-EU PSM sample)



<u>Figure A1c:</u> Dynamic Treatment Effects of Private CSR-Engagement Signal Availability on CSR-Engagement Reporting (UK-EU PSS sample)