Greenwashing Premium

Eric Tham*[1,2]

[1]School of Science and Technology, James Cook University

[2] EDHEC Business School

Mar 2023

ABSTRACT

Greenwashing is framed as a signalling game between firms and investors and occurs due to cheap news signals. A worst social outcome occurs with a pooling equilibrium. Firms greenwashed if ex-ante they are in the top decile of a portfolio sorted by ESG news scores from NLP, and ex-post fell out from the top decile for ESG performance scores. The greenwashing premium for ESG in USA is historically not significant but episodic. The premium in 2020 was largely due to consumer and green firms at 2.9% and 4.2% respectively. It was larger for the 'E' and 'G' than the 'S' pillar, except amongst brown firms. A mechanism design is proposed to conceptually incentivise firms to keep to their ESG goals.

Keywords: Greenwashing, Game Theory, Signalling Game, ESG finance, Cho-Kreps criterion, Mechanism Design

JEL classification: C70, C72, D02, D04, Q50, Q51, Q54, Q58

*erictham@hotmail, Special thanks to Richard Petersen and Liu CJ from MarketPsych, Eric Fischkin and Radha Pendyala from Refinitiv for the provision of data. He also thanks Marat Molyboga, Kang Young-Dae and Jonathan Harris for their helpful comments. All errors are solely mine.

1 Introduction

Recently, greenwashing has taken centre stage in combating global warming. Greenwashing is defined in Wikipedia as green marketing being used to persuade the public that an organization's products, aims and policies are environmentally and socially friendly. With the recent increase in public ESG concerns, greenwashing has become an urgent matter to understand and mitigate. In Nov 2022, Europe's top regulators launch a study to better understand greenwashing.[1] Similarly, the United Nations released a report condemning the practice at the COP 27 in Nov 2022.[2] It cited that net-zero pledges by non-state players cannot be a cover-up and the promises must be accountable and transparent. At the firm level, a study by the non-profit New Climate Institute in Feb 2022 in Nature[3] showed that 25 of the largest companies in the world accounting for 5% of global emissions are actually making much less of their commitments. Only 3 firms actually lived up to their commitments.

Evidence of greenwashing in fund intermediaries for example in hedge funds and asset managers are cited in Liang, Sun, and Teo (2022). The Financial Times[4] remarked in Jun 2022 that funds mislead investors by overstating the greenness of their funds. S&P Global[5] remarked that sustainable bond issuance including green, social, sustainability, and sustainability-linked bonds have collectively exceed $1 trillion in 2021. However, a lack of consistency in instrument labelling and post-issuance disclosure has raised investor fears that sustainability claims made by issuers might be overstated or unreliable. An Economist[6] article wrote that sustainable funds are stuffed full of polluter and sin stocks.

Liang et al. (2022) wrote that greenwashing is an agency problem for hedge funds, which act as intermediaries for investment. This paper further frames greenwashing as an agent's *signalling game*. In this signalling game, 'good' firms are those that commit and deliver on their green promises. 'Bad' firms are those that do not. They masquerade as far as possible to be 'good' to enjoy the benefits of the good firms – increased public investment with lower costs of capital, greater liquidity of public funds and better equity performance.

Another channel of greenwashing is firms influencing consumers' tastes through advertising, projects or public announcements of subsequent 'blank promises'. This is a much-studied

topic in consumer marketing. In both cases, greenwashing has detrimental financial and social implications. It is to the advantage of firms to commit to responsible behavior and appear as 'green'. These firms could potentially outperform after investors' ESG concerns are factored in. Bolton and Kacperczyk (2021a) wrote that green firms experienced higher equity returns in the prior years before 2020 due to increased episodic public interest in environmental concerns in the US and globally in Bolton and Kacperczyk (2021b). Pastor, Stambaugh, and Taylor (2022) cited that green assets delivered high returns in recent years due to unexpectedly strong increases in environmental concerns, and not high expected returns. These episodes occur as the episodic greenwashing premium found in this paper that firms expect to get from making 'positive news' on their ESG claims to investors and customers.

In this signalling framework firms are incentivised to greenwash due to 'cheap news signals' cited in Bolton and Kacperczyk (2022). Bolton and Kacperczyk (2022) had used NLP (Natural Language Processing) to analyse firm's corporate climate disclosures in annual reports, defining 'cheap' talk as the share of precise versus imprecise climate commitments. Due to the *verbal nature of greenwashing*, it is difficult to quantify a firms' propensity of greenwashing. Jong, Huluba, and Beidad (2019) wrote that often firms engage in half-truths and half-lies making it hard to decipher true green claims. This paper thus uses NLP to compute the firms' ESG news sentiment scores.

To determine firms' incentive to greenwash or the greenwashing premium, I use a double portfolio sorting method. The first portfolio in this method sorts firms at the beginning of a period on their ESG *news scores*. The top decile in this portfolio are firms that send through the media the most positive ESG messages. The second portfolio sorts firms based on ex-post ESG performance scores. If the firms live up to their initial promises, they should be reflected in the top decile of the ESG *performance scores*. Otherwise, they would be deemed to have greenwashed. The one year equally weighted excess returns of the firms in the portfolio excluding firms in the second portfolio are then obtained excluding the usual control variables. If the two decile portfolios are the same, there would be no greenwashing. The ex-post nature of this second portfolio is necessary due to the information asymmetry of the game. As such,

the premium computed cannot be hedged in the conventional sense.

The paper postulates that the present regulatory regime without transparent and robust standards inclines the signalling game towards a pooling equilibrium. Under this equilibrium, the investors/ consumers cannot distinguish between 'good' and 'bad' players by their messages. This type of equilibrium has potentially several responses from the firms under the Cho-Kreps criterion. The equilibrium depends on the a priori belief of the investors/ consumers on the percentage of the 'good' firms. A worst possible social outcome occurs when they categorise all firms as 'bad'. Even originally 'good' firms could then degenerate towards behaving like the 'bad' players. This is detrimental towards the efforts of climate change. It is no surprise that the chair of the UN expert group[7] reiterated that the 'planet cannot afford delays, excuses, or more greenwashing'.

The signalling methodology has the secondary advantage of revealing the percentage of firms that eventually 'greenwash' after initial 'optimistic' news messages. In the empirical section, this percentage of firms has shown a steady decline across green firms - decreasing from more than 90% in the 2005 to about 70% in 2020s. However, for the brown and consumer firms this proportion had been relatively unchanged at $\approx 80\%$ for the same period. Further, the game equilibrium outcomes also predict 'greenhushing' - a phenomenon when firms abstain from making public commitments on their climate efforts. This occurs when the penalties for greenwashing are overly harsh. This has been reported in a Oct 2022 Bloomberg article.[8]

The Refinitiv ESG news and performance data are used for the study. The data are granular and allow the segmentation of the ESG scores according to individual Environmental, Social and Governance pillars. This methodology is applied on US firms, in particular to brown, green and consumer-orientated firms to determine the greenwashing premium in these industries on a yearly basis since 2005 to 2021. The paper finds that the yearly greenwashing premium is historically not statistically significant by the t-test hypothesis. Instead it is episodic with highs in 2012 (warmest weather on record in US), 2016 (post Paris COP) and 2019-2020.

Historically, the average premium for consumer firms is higher and about 2.3% for the G

pillar and 3.1% for the E pillar during the period 2005 to 2021. The green firms have a 2% premium for the S pillar and a small 1.1% for the E pillar. Brown firms have a premium of -0.002% (almost zero) for the E pillar but instead could be rewarded for their G pillar efforts.

The literature on greenwashing have largely been on consumer marketing and perspectives, and on firms' Corporate Social Responsibility (CSR), and less so in the financial literature. Jong, Harkink, and Barth (2018) established that consumers recognised the green claims made by green marketing, but it has limited benefits and no true competitive advantage (purchase interest). This paper makes similar findings. Historically the greenwashing premium (from equity returns) had been inconsistent and episodic. Delmas and Burbano (2011) wrote that lax and uncertain regulation as a key driver of greenwashing, and poor environmental performance and positive communication as key drivers to greenwashing. The paper remarked that it is difficult for firms to adjust their environmental performance in the short term, but communication through the media is much 'easier'. It is cheap in other words to send signals via the news media channel. Yu, Luu, and Chen (2020) uses the relative position of firms in Bloomberg ESG Transparency scores and Refinitiv ASSET4 scores and consider firms to greenwash if they are 'overly' transparent and reveal large quantities of ESG data but perform poorly in ESG aspects. This paper similarly compares between 2 industry ESG scores but also uses news as key signals to the public.

The paper is organised as follows. In the following section 2, the data from Refinitiv is first described. In section 3, the signalling game is discussed. Section 4 discusses empirical results on the greenwashing premium and also the economic implications. It looks at a mechanism design approach to the signalling game such that firms are incentivised to reveal their true intentions. Finally the last section concludes.

2 Data

Two sets of data from Refinitiv are used - firm news data from MarketPsych ESG Analytics (thenceforth TRMA) and ASSET4 data. MarketPsych uses natural language processing (NLP) engine to scour the news and social media for firm-related ESG news. The TRMA

analyses both news and social media news and uses them to obtain sentiment scores for ESG-related news for firms ranging from 1 (most negative) to 100 (most positive). TRMA sources are extensive coming from 2000 news and 800 social media platforms around the world. The sources include all notable news outlets from Wall Street Journal, CNBC, Bloomberg and Reuters. The inclusion of the social media is advantageous since they would also include advertisements made by the firms in promotional efforts, and not necessarily formal ESG news announcements. The news sentiment scores are divided into topics like the environmental, social and governance pillars, environmental innovations, corporate controversies and others.

The firms' ESG performance scores are from Refinitiv ASSET4. Unlike the TRMA data which are from news data, these scores are verifiable and the methodology is transparent on the website[9]. The ASSET4 scores are similarly divided into ESG overall scores and also the individual pillars - Environmental, Social and Governance. Only companies in both data sources in the years they are common are analysed. Data from Jan 2005 to Dec 2021 are used. Different green, brown and consumer-related industry portfolios are created under the Thomson Reuters Business Classifications (TRBC). The Table 2 shows the number of firms in each industry to be increasing over the years, similar to Bolton and Kacperczyk (2021a).

[Insert Table 2: No of US Firms common in TRMA and ASSET4 by Industry Types]

Figure 3 shows the time series of the average ESG, Environmental, Social and Governance pillars scores for the two series. Generally, the TRMA scores are higher than the ASSET4 standard scores. This means that companies' public news announcement could be more optimistic than its actual performance. However, this should not construed as greenwashing as it could be an artefact of the individual data construction. A general improvement in the ASSET4 ESG scores could be observed up till 2021. This is in line with the increased ESG awareness amongst the general public and corporate sector.

[Insert Figure 3: Time Series of average TRMA and ASSET4 scores from 2005 to 2021 Dec]

It is interesting to delve into example individual headline news for each of the topics for

some of the firms with some news headlines dating back to 2018 in Table 1. They are categorised into individual ESG pillars. These news are exemplary of firms' announcements and commitments to be ESG responsible. Logically, they should translate to improved ESG performance further down the horizon.

Table 1. Examples of Firm Positive News Announcement and Signals

No	Firm	Headline	Pillar	Year	Industry
1	Glatfelter	Leadership Council to Guide Commercialization and Open Access to Safe Lithium Battery Architecture	Environmental	2018	Basic Materials
2	Corteva Inc	Corteva Agriscience announces award to galvanise food resilience and promote agricultural sustainability	Environmental	2021	Consumer
3	Omnicron Media	OMG UK certified with the Trustworthy Accountability Group (TAG) Brand Safety Certified Program	Governance	2020	Consumer
4	Tesla	Tesla Tapped by Australia anew in Virtual Power Plant Plan	Environmental	2018	Technology
5	Ingevity	Ingevity elects two new members to its board of directors	Social	2019	Basic Materials

[1] News articles or public announcements relate to firm ESG-related topics - across the different E, S and G pillars.
[2] The news are categorised into the green, brown and consumer-orientated industries based on their relative emission intensities. The basic materials is under the brown classification, the consumer under the consumer-orientated and the technology under the green category.

2.1 News as Signals

A firm signals to the markets of its ESG goals through public news. Studies like Boykoff and M.Boykoff (2007) and Hale (2010) have shown the news media in US as increasing social awareness on environmental issues. It can influence a population's perceptions in two ways: (i) by informational content that could affect the consumer's channel (ii) by diverting investors' attention to the topic. Suppose at time $t = 0$, ESG news score sentiment $>> 0$ or there is positivity on firms enacting on its ESG goals. The firm is sending to the market that is a good firm. On the contrary, suppose the ESG news score sentiment < 0, or there is negativity in the ESG-related news. The firm could be interpreted as a non ESG-compliant firm.

3 Model

I use the classic signalling game model to model greenwashing. There are two players in this game – firms (the sender) and investors/ consumers (the receiver). Investors/ consumers have two choices – I (invest/ buy) or E (Exit).

Figure 1. Greenwashing as a signalling game in extensive form

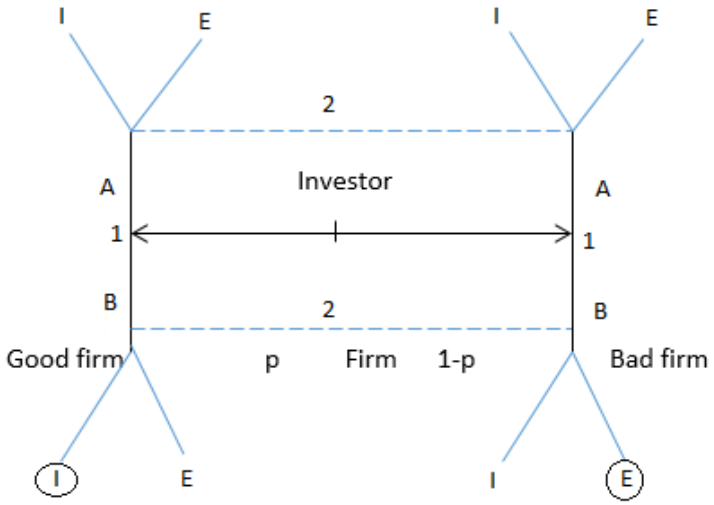


Figure 1. This is a 3-stage signalling game. At the beginning (stage 0), the firm knows their type – either good or bad. Good firms want to comply with ESG goals, unlike bad firms which key motivation is more pecuniary. Of the firm population, proportion p are good firms, and $1-p$ are bad firms. In stage 1, the good and bad firms decide their course of action and send individual signals s_i as news regards their ESG goals and actions. In stage 2, investors observe these signals but not knowing ex-ante if the firm is good or bad, and choose either to I (Invest) or E (Exit).

The signal action 'B' is truthful reporting for both good and bad firms for their ESG actions. The signal action 'A' is exaggerated or unfulfilled 'greenwash' claims. This is not trivial due to a lack of standardisation of reports, transparency nor disclosure. A firm can be a signatory to the United Nations Principles for Responsible Investment (PRI) as a signal. But Kim and Yoon (2022) wrote that US mutual funds after signing the United Nations Principles for Responsible Investment (PRI) did not have improvements in fund-level ESG scores, or

follow-through actions in responsible ESG implementation. Instead, the funds use the PRI affiliation widely on company websites, marketing materials, and fund documents as a form of advertisement. Note that the bad firms are not necessarily 'sin' firms – those in polluting industries for example.

Let the good firm's signal F_g costs be $s_g \in [0, s_g]$ and the bad firm F_b be $s_b \in [0, s_b]$. This cost of signal using news at present is cheap at least from a monetary perspective. It must be noted that the investor would know the industry type of the firm (trivially), but do not know the true intention or its commitment to the Net Zero initiative. The investor has 2 avenues of action – either E or I at the decision node after observing the signal from the good and bad firms. The investor updates this belief of the firm type through Bayesian expectations. The 'E' action would be an exit or abstain strategy. This is not an uncommon strategy as the Financial Times[10] reported that pension funds, churches and universities have actively broken from pollutive companies. The 'I' represents an investment into the firm or consumer buy-in to the firm's products or services.

The investor updates his Bayesian beliefs from observing the firm i signal s_i where $P(F_g|s_g)$ is his belief for a good firm and $P(F_b|s_b)$ for a bad firm. In a separating equilibrium, the investor is able to identify the firm type from its signals. This is an ideal scenario for regulators and investors alike, since they can fairly reward both types of firms for their ESG commitments.

3.1 Pooling Equilibrium is detrimental to social good

In an environment with 'cheap' signals and lack of transparency and regulation, the equilibrium type is likely a pooling equilibrium, where investors are not able to distinguish the good and bad firms ex-ante. There is a continuum of equilibrium for the pooling case. The Cho-Kreps criterion in Cho and Kreps (1987) or the Intuitive Criterion reduces the possible outcome scenarios by restricting the possible firm types to types who could obtain higher utility levels by deviating to off-the-equilibrium messages, and to types for which the off-the-equilibrium message is not equilibrium dominated. The Cho-Kreps criterion could imply a worst possible outcome in terms of public social good ex-ante. Good firms may deprecate to

put in less efforts for its ESG goals as that of 'bad' firms, as the bad firms take a 'slice of the pie'. There is evidence of this form of equilibrium occurring in the hedge fund industry in Liang et al. (2022) with *investors not able to discriminate between greenwashers and genuinely green funds.*

Regulatory action and disclosures should target a separating equilibrium for the good and bad firms. Most of the regulatory actions in its present state are for the environmental ('E') aspect and target funds instead of firms. These social and governance aspects are less well 'regulated' due to its implicity. For example, issues like gender diversity on boards or employee satisfaction are encouraged in firms but never tightly regulated or imposed.

3.2 Equilibrium payoff

The firms are risk-neutral and their expected reward utility is μ_i (with $\mu_g > \mu_b$ by definition) which are netted off their signalling cost s_i, other intangible risks including reputational, regulatory and legal risks, encapsulated by l_i. In an unregulated market without strict disclosure rules or penalties, $l_i \to 0$.

$$firm\ payoff\ = \mu_i - s_i - l_i - p_i \tag{1}$$

The $s_i \approx 0$ due to the cheap news signals. The project costs to fulfil claims are represented by p_i and measures the incentive/ disincentive for firms to carry out their promises. Brown or 'sin' firms face a much higher cost of p_i than green firms. For example, brown firms by being a signatory to the Net Zero Initiative have to make costlier commitments. They may need to employ environmental technologies like carbon capture technologies, dispose profitable stranded assets or exit from profitable markets due to tighter regulations.

3.3 How will good firms behave?

In a separating equilibrium, the good and bad firms have incentive compatibility to act its true status with better payoffs. For the good firms, this imply:

$$\mu_g - l_g > \mu_b \qquad \textit{Incentive compatibility} \tag{2}$$

On the LHS, the good firm makes ex-ante positive news messages consistent with its intended ESG goals, and rewarded with μ_g. On the RHS, the good firm is identified as a bad firm to being ESG non-compliant with $\mu_b \leq 0$ as the investors/ consumers unlikely reward bad behaviour. Note the p_g term cancels out on both sides as the good firm commits to its ESG goals. The $l_g = 0$ on the LHS since the good firm by not making false news claims is unlikely to be liable for potential legal and regulatory penalties. However suppose with onerous high penalties or fine where $l_g \uparrow$ the good firm may choose to 'greenhush' to be not liable for potential legal and regulatory penalties so that $l_g = 0$ on the RHS. Greenhushing means the good firms send no signals to the market whatsoever and keeps their climate goals secretive as they are afraid of backlash against them.

3.4 How will bad firms behave?

For the bad firm, its incentive compatibility is:

$$\mu_g - l_b > \mu_b = 0 \qquad \textit{Incentive compatibility} \tag{3}$$

On the LHS, it sends a misleading signal of committing to its ESG goals, and enjoys the benefits of the good firms with μ_g. It faces potential legal and regulatory penalties $l_b > l_g$ likely more than the good firms. For low regulatory penalties $l_b \to 0$, they would be incentivised to just make false claims. On the RHS it is truthful in sending it would not commit to any ESG goals. However since $\mu_b \leq 0$, they would rather not make such news so that $\mu_b = 0$. They would also not be liable for any potential damage/ costs ($l_b = 0$) by not making any claim.

The Inequalities 2 and 3 thence reduce to:

$$\mu_g - l_i > 0 \qquad for\ i \in [g, b] \tag{4}$$

The model implies the good firms either send true positive news of their ESG goals or greenhush. The bad firms either send misleading news signals or do not participate from sending signals at all. The normal form of the separating equilibrium of the game is in Table 3.

3.5 Bayesian Beliefs of the Investors and Consumers

A perfect Bayesian equilibrium depends on the beliefs of the investors and consumers. A pooling equilibrium results if they are not able to distinguish the good/ bad firms through their messages. There are many potential pooling equilibrium depending on the Bayesian beliefs on the distribution of the good/ bad firms in Equation 5. In a pooling equilibrium, each firm is rewarded uniformly as μ_{bg} as μ_g becomes μ_{bg} in Inequalities 2 and 3.

$$\mu_{bg} = x_b \mu_b + x_g \mu_g \tag{5}$$

The x_b and $x_g = 1 - x_b$ are the proportions of bad and good firms respectively in the competitive market. They are obtained in our empirical study with $x_b \approx 0.6$ to 0.9 in Figures 4 to 6. When $x_b \uparrow 1$ and with $\mu_b < 0$, the $\mu_{bg} \to 0$. This has a knock-on effect on Inequalities 2 and 3 such that the LHS $<$ RHS, and good firms behave as bad. The empirical method section quantifies μ_{bg} as the average 1-year equity returns at each industry level and obtains it as the greenwash premium.

4 Empirical Study

I use the aforesaid methodology to compute the greenwashing premium. Ex-ante the firm's ESG news scores from NLP are sorted and the top decile obtained. A τ period is used as the 'fulfilment' period for firms to oblige on their 'promises'. A $\tau = 1$ year is used for convenience as an annual return and further, the ASSET4 data granularity for most of the firms are

largely on a yearly basis. The portfolio sorting method was popularised by the seminal paper on momentum in Jegadeesh and Titman (1993). In climate finance, Engle, Giglio, Kelly, Lee, and Stroebel (2020) also used it to construct factor mimicking portfolios on climate change. Ex-post, the firms in the top decile sorted by ASSET4 ESG performance scores are obtained.

The average returns of these 'greenwashed' firms are obtained, after controlling for firm-level variables in Equation 6. These control variables ($CTRL$) include factors that are known to affect returns. These include the momentum factor (MOM) proxied by the average of the last 12 months returns, the value factor proxied by the book to market ratio (BM) and size factor proxied by the log of the market capitalisation (MKT_CAP). The BM values are obtained directly from the Wharton Research Data Services (WRDS) and the MKT_CAP from the number of shares and equity prices from CRSP. The momentum factor is derived from the average of the firm's past 12 months monthly returns from CRSP. The use of these control variables are similar to Bolton and Kacperczyk (2021a) and Bolton and Kacperczyk (2021b). The omission of the α constant from Equation 6 is deliberate, as there is evidence from Pastor et al. (2022) that part of the unaccounted returns inside α is ESG-related.

$$r_{i,t\to t+\tau} = \beta_{1,i}MOM + \beta_{2,i}BM + \beta_{3,i}MKT\_CAP + \epsilon_i \tag{6}$$

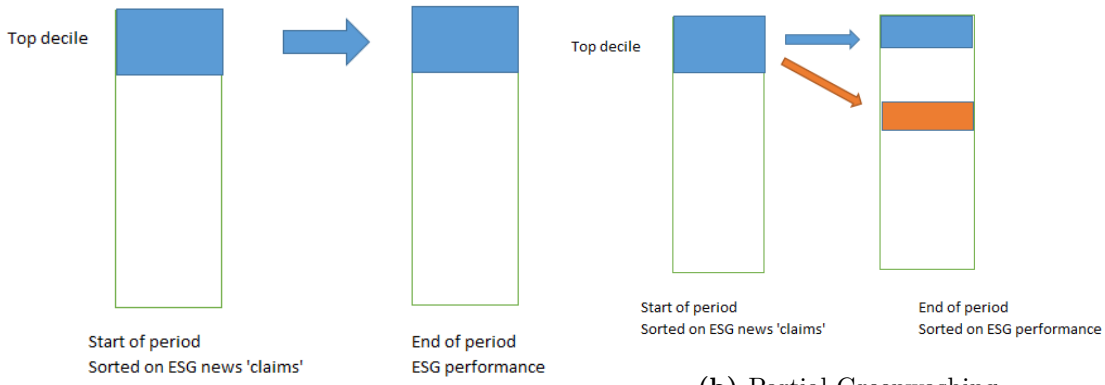
The excess premium over period of τ year attributable to greenwashing is thence:

$$Greenwashing\ premium_{\tau\ yr} = \frac{1}{n}\sum_{i=1}^{n}\Big[r_{i,t\to t+\tau} - (\beta_{1,i}MOM_i + \beta_{2,i}BM_i + \beta_{3,i}MKT\_CAP_i)\Big]_{i\in\mathbb{A}\notin\mathbb{B}} \tag{7}$$

- Set $\mathbb{A}$ - top decile of portfolio ranked by ESG news sentiment scores at time t
- Set $\mathbb{B}$ - top decile of portfolio ranked by ESG standard scores at time $t+\tau$

One scenario that this methodology is not able to capture the greenwashing premium is when the top decile firms remain the same and all firms in the portfolio collude to not oblige on their ESG obligations. In this case, the premium is zero, although this 'collusion' is a highly unlikely scenario in a competitive environment. Greenwashing is only known ex-post

when a firm *not* subsequently fulfils its earlier commitments. The premium also does not distinguish between bona fide and extraneous difficult circumstances for firms. Firms may make a genuine attempt to fulfil their ESG promises made at the beginning of the period, but may be prevented to do so due to the economic situation or other difficult circumstances. This is one reason why the greenwashing premium was high during the Lehman crisis period of the 2008-2010. The risk-free rate discounting has been ignored in Equation 7 for simplicity. The model implication is best illustrated in the Figure 2.



(a) No greenwashing

(b) Partial Greenwashing

Figure 2. Illustration of no greenwashing and partial greenwashing

Illustration of no greenwashing and (partial) greenwashing in an industry. Firms are deemed to greenwash (in brown) if their initial 'positive' (top decile) news sentiment are not reflected in the subsequent top decile of the ESG performance rankings.

4.1 Results

Figures 4 to 6 show the time series of this greenwashing premium for firms in the green, brown and consumer-orientated industries for the US. The green industries are quintessentially financial, technology and healthcare. The brown industries are utilities, energy and basic materials. The consumer industries include consumer cyclical goods and consumer non-cyclical goods. In Ardia, Bluteau, Boudt, and Inghelbrecht (2022), these brown firms are stated to have greenhouse emissions that are almost 100x more than the 'green' firms.

[Insert Figure 4: Greenwashing Premium & percent of greenwashing firms for Green Industries]

[Insert Figure 5: Greenwashing Premium & percent of greenwashing firms for Brown Industries]

[Insert Figure 6: Greenwashing Premium & percent of greenwashing firms for Consumer Industries]

Table 4 shows the t-test with the null hypothesis that the greenwashing premium = 0 from 2005 Jan to 2021 Jan. The null hypothesis is not rejected at the confidence interval of 95%, meaning the greenwashing premium has not been statistically significant during this period. However, the greenwashing premium is episodic. Green firms witnessed a greenwashing premium in 2019 of 6.7% which decreased to 4.2% in 2020. In 2019, this premium was due to the individual ESG pillars. On the contrary, consumer firms have a greenwash premium of 9.8% in 2019, largely due to the 'E' pillar at 9.7%. These two industries unlike the brown industries were the main culprits of greenwashing during in 2019/2020. On the contrary, the brown industry has a negative greenwashing premium during these two years of about -12% and -1.5% in 2019 and 2020 respectively for the E pillar. The brown industry had a 5.0% greenwashing premium in the 'G' pillar in 2019 which decreased to -9.8% in 2020.

These observations imply the brown industry is unlikely to be rewarded for 'making positive claims' with improved equity performance. The public and investors alike are less inclined or already wearied to believe greenwashing claims in the brown industry in the US especially for the 'E' pillar. This could partly be due to existential lawsuits for brown energy firms like Chevron, Marathon, Exxon Mobil and Hess[11]. On the contrary, there is incentive for green and consumer firms to make 'positive news claims'.

Further, the peaks in the greenwashing premium are generally observed in the years 2012/2013, 2016 and in 2019/2020. The temperatures recorded in 2012/2013 were one of the 10 warmest years on record globally[12] according to National Oceanic and Atmospheric Administration (NOAA). This was partly due to the la Nina effect. This could indirectly renew consumers' interest in ESG concerns, motivating firms to 'greenwash'. A study by Addoum, Ng, and Ortiz-Bobea (2020) showed that temperature had an effect on the sales and profitability of firms. Similarly, investors' interest in the firm could be affected due to behavioral changes from extreme temperatures in Cao and Wei (2007). In 2016, the landmark Paris COP 21 was also concluded, igniting investors' and consumers' interest in ESG concerns.

The proportion of firms that greenwashing (after making news claims and not fulfilling them) has witnessed a general decline of 86% in 2011 to 70% in 2021. This ratio is relatively stable for the brown industry from 78% to 84% from 2005 to 2021, whilst it is from 90% to 69% for the green industries and 90% to 68% for the consumer industries. Across the pillars, the environmental 'E' pillar is the lowest at 72% compared to 78% for the social 'S' pillar and 90% for the governance 'G' pillar. This could be attributable to the 'E' being relatively more verifiable with the carbon disclosures unlike the harder to quantify 'S' and 'G' pillars as cited in Po-Hsua, Li, and Tsou (2022). The environmental 'E' pillar has emissions disclosures through the Greenhouse Protocol and the CDP online, which are made mandatory in 2022 by the SEC.

4.2 Policy Implications

This paper shows there could be positive equity rewards for firms that greenwash, especially for green and consumer industries in the 'S' and 'G' pillars. Many existing regulations largely increase transparency and consistency of disclosures/ reporting. This helps to reduce the information asymmetry for investors/ consumers, and the premium due to adverse selection. These measures are implemented somewhat in Dec 2022 by the Swiss Federal Council[13]. The measures stressed 1) transparency and comparability 2) accountability 3) credibility that is verifiable. The points 1 and 3 help to reduce information asymmetry in the signalling game whilst point number 2 would adjust the l_i targeting the regulatory penalties.

More could possibly be done on the media front as this is the primary means of communication to consumers and investors. In fact, this has resulted in several law cases on tort[14]. Tort in a narrow sense is defined legally as an act or omission that gives rise to injury or harm to another and amounts to a civil wrong for which courts impose liability. An example in the footnote link is HSBC Bank. They are alleged to claim transition to net-zero while at the same time financing industries that emit notable levels of greenhouse gases. The Advertising Standards Authority (ASA) has since banned this forms of media ads, and at the same time enforce that future communications of such claims not omit material information. These

measures increase the l_i which are the legal and reputation risks in Equation 1 and Equation 4.

This study finds that the brown industry is less incentivised to 'greenwash' possibly due to more public sceptism than the green and consumer industries and also due to the large cost of p_i to implement 'green' measures. There could be differentiating measures between brown and green/ consumer industries in regulation against greenwashing.

A potential application of this study is mechanism design or reverse game theory in Harsanyi (1967). The mechanism design should adjust the incentive compatibility (IC) conditions such that good and bad firms are incentivised to only signal their true states in a separating equilibrium than prefer a pooling equilibrium. The problem is complicated by cheap talk or signals in Farrell and Rabin (1996) for $s_i \approx 0$ in Equation 1. The regulators could target l_i at l^* - the legal, reputation and regulatory costs to differentiate between the good and bad firms. For the good firm, its payoffs would then become in Equation 8. The l^* is applied in a pooling equilibrium on the RHS, whilst the LHS is the incentive condition for it to act as 'good' firm in a separating equilibrium.

$$\mu_g - (l_g + p_g) > \mu_{bg} - (l^* + p_g) \tag{8}$$

$$\implies l^* - l_g > \mu_{bg} - \mu_g \tag{9}$$

For the bad firm, its payoffs are such that its IC makes it such that it prefers a separating equilibrium on the LHS than a pooling equilibrium payoff on the RHS.

$$\mu_b - (l_b + p_b) > \mu_{bg} - (l^* + p_b) \tag{10}$$

$$\implies l^* - l_b > \mu_{bg} - \mu_b \tag{11}$$

The inequalities thence provide a range for the l^* for a separating equilibrium to occur between the good and bad firms. l^* has a lower bound of $max(l_b + \mu_{bg} - \mu_b, l_g + \mu_{bg} - \mu_g)$. The actual implementation would be complicated by the intangible legal and reputational costs of l^*.

5 Conclusion

This paper makes a few contributions. It first frames greenwashing in ESG as a signalling game framework with the players as investors/ consumers and the firms on opposite sides. Signals are sent through cheap news talk by firms on their ESG claims. The existing anecdotal evidence is that a pooling equilibrium has resulted such that the market cannot distinguish the good and bad firms ex-ante. Regulators should target a separating equilibrium such that the good and bad firms are incentivised to act their true selves. The paper suggests to manage the non-compliance penalties for the game design such that the good and bad firms are incentivised to act their true selves.

The paper also contributes empirically by quantifying the greenwash premium. Using a double portfolio sorting method, the study estimates the greenwashing premium for the different 'E', 'S' and 'G' pillars and for the green, brown and consumer-orientated industries. It finds this greenwashing as not statistically significant but is episodic in nature from 2005 to 2021. The green and the consumer industries are more culpable of greenwashing than the brown firms, especially in the 'E' and 'S' pillar in 2019/20. This paper hopefully helps to shed some light on the nature of greenwashing in ESG.

References

Addoum, Jawad M., David Ng, and Ariel Ortiz-Bobea, 2020, Temperature shocks and establishment sales, *Review of Financial Studies* 33, 488–500.

Ardia, David, Keven Bluteau, Kris Boudt, and Koen Inghelbrecht, 2022, Climate change concerns and the performance of green versus brown stocks, *SSRN working paper* .

Bolton, Patrick, and Marcin Kacperczyk, 2021a, Do investors care about carbon risk?, *Journal of Financial Economics* 142, 517–549.

Bolton, Patrick, and Marcin Kacperczyk, 2021b, Global pricing of carbon transition, *NBER working paper* .

Bolton, Patrick, and Marcin Kacperczyk, 2022, Cheap talk and cherry-picking: What climatebert has to say on corporate climate risk disclosures, *Finance Research Letters* 47.

Boykoff, Maxwell T., and Jules M.Boykoff, 2007, Climate change and journalistic norms: A case-study of us mass-media coverage, *Geoforum* 38, 1190–1204.

Cao, Melanie, and Jason Wei, 2007, Stock market returns: A note on temperature anomaly, *Journal of Banking and Finance* 29, 1559–1573.

Cho, In-Koo, and David M. Kreps, 1987, Signaling games and stable equilibria, *Quarterly Journal of Economics* 179-222.

Delmas, Magali A., and Vanessa Cuerel Burbano, 2011, The drivers of greenwashing, *California Management Review* 54.

Engle, Robert, Stefano Giglio, Bryan Kelly, Heebum Lee, and Johannes Stroebel, 2020, Hedging climate change news, *Review of Financial Studies* 33, 1184–1216.

Farrell, J., and M. Rabin, 1996, Cheap talk, *Journal of Economic Perspectives* 10, 103–118.

Hale, B.W., 2010, Using newspaper coverage analysis to evaluate public perception of management in river-floodplain systems, *Environmental Management* 45, 1155–1163.

Harsanyi, John C., 1967, Games with incomplete information played by "bayesian" players, *Management Science* 14, 159–182.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Jong, Menno De, Karen Harkink, and Susanne Barth, 2018, Making green stuff? effects of corporate greenwashing on consumers, *Journal of Business and Technical Communication* 32, 77–112.

Jong, Menno De, Gabriel Huluba, and Ardion Beidad, 2019, Different shades of greenwashing: Consumers' reactions to environmental lies, half-lies, and organizations taking credit for following legal obligations, *Journal of Business and Technical Communication* 34.

Kim, Soohun, and Aaron Yoon, 2022, Analyzing active fund managers' commitment to esg: Evidence from the united nations principles for responsible investment, *Management Science* .

Liang, Hao, Lin Sun, and Melvyn Teo, 2022, Greenwashing: Evidence from hedge funds, *SSRN working paper* .

Pastor, Lubos, Robert F. Stambaugh, and Lucian A. Taylor, 2022, Dissecting green returns, *SSRN working paper* .

Po-Hsua, Hsu, Kai Li, and Chi-Yang Tsou, 2022, The pollution premium, *Journal of Finance* .

Yu, Eleen, Bac Van Luu, and Catherine Chen, 2020, Greenwashing in environmental, social and governance disclosures, *Research in International Business and Finance* 52, 843–857.

Table 2. No of Firms common in both TRMA and ASSET4 scores by industry types

Year	Overall Industries	Green	Brown	Consumer-orientated
2005	438	169	92	69
2006	433	165	92	68
2007	444	162	104	67
2008	586	206	136	79
2009	640	230	146	83
2010	662	235	148	89
2011	689	240	144	104
2012	658	228	142	93
2013	667	230	144	96
2014	671	232	144	94
2015	1159	451	193	147
2016	1705	757	231	222
2017	1984	949	274	245
2018	2040	1005	282	229
2019	2136	1081	281	235
2020	2184	1133	279	233
2021	1750	996	181	185

[1] The number of firms has increased in coverage in both TRMA and ASSET4 since 2005 to 2021. The TRMA reflects the news sentiment scores for the firms' ESG claims and the ASSET4 the ESG performance scores one year later. Only firms with data at the beginning of the year - Jan are used and common in both datasets. Accordingly, the top decile of portfolio formed from these firms has also increased through the years. Note the numbers are for the overall ESG pillar, and are approximate for the other environmental, social and governance pillars.

Table 3. Game Payoffs in Normal Form for Separating Equilibrium

		Invest/ Increased Sales	Exit
Good Firm (g)	Greenhush	$\mu_g - p_g$	$\mu_b - p_g$
	True signal	$\mu_g - p_g$	$\mu_b - p_g$
Bad Firm (b)	False signal	$\mu_g - l_b - p_b$	$\mu_b - l_b - p_b$
	True signal	$\mu_g - p_b$	$\mu_b - p_b$

[1] This shows the normal form of the game between the good/ bad firms and the investor/ consumer. As the signals are costless, $s_i \approx 0$.

[2] When a bad firm sends a false signal, they become liable for l_b, which are intangible costs related to legal/ reputational/ regulatory costs for sounding out their commitments to their ESG goals. If it sends a true signal, they are not liable $l_b = 0$. A good firm would not send a false signal but instead greenhush by not communicating their ESG commitment plans

[3] Regardless, a good/ bad firm incurs p_g or p_b respectively on their project costs as they commit/ not commit to their ESG goals.

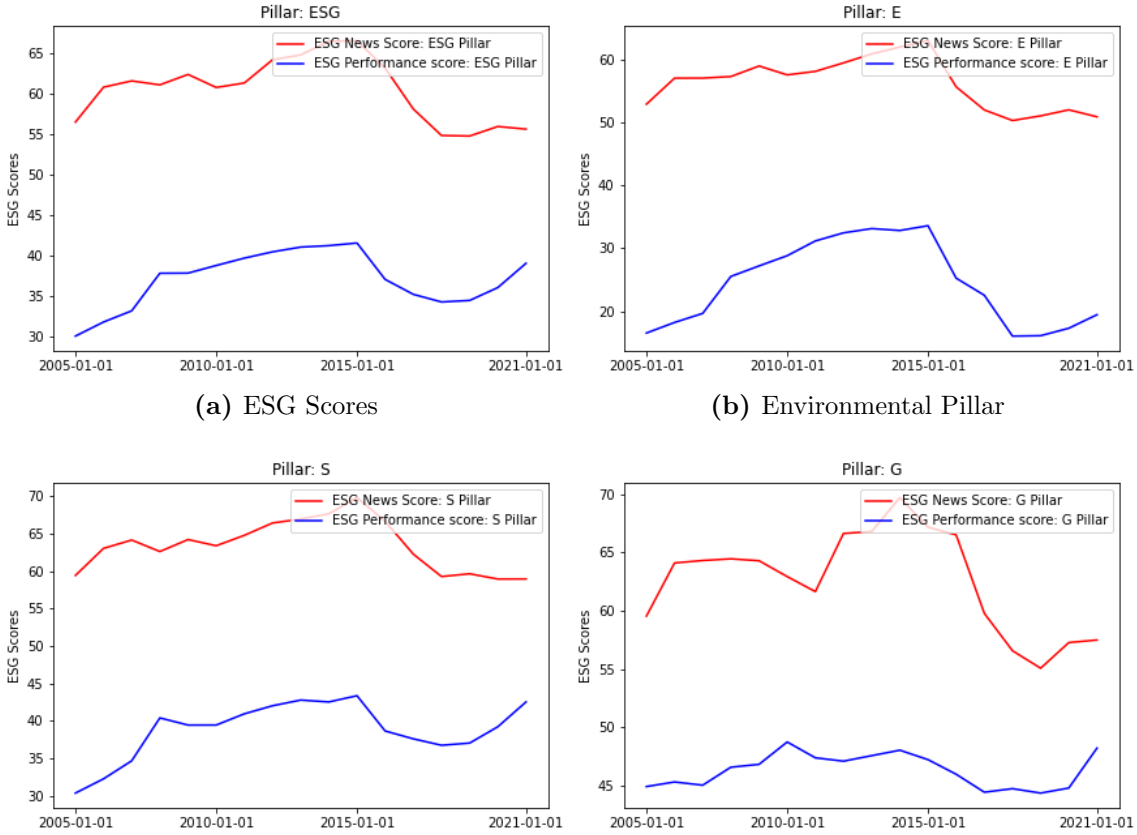
Table 4. t-stat test for positiveness of the greenwashing premium

Metric	Brown	Consumer	Green	All Industries
Mean of Greenwashing Premium				
ESG	-0.007	-0.025	-0.003	-0.001
E	-0.002	0.031	0.011	0.024
S	0.011	-0.031	0.02	0.005
G	0.017	0.023	-0.001	0.012
Standard Deviation of Premium				
ESG	0.197	0.216	0.184	0.166
E	0.163	0.153	0.154	0.143
S	0.144	0.157	0.169	0.148
G	0.171	0.172	0.132	0.13
t-stat probability				
E	0.957	0.429	0.788	0.511
ESG	0.896	0.655	0.945	0.982
G	0.693	0.608	0.983	0.724
S	0.758	0.437	0.636	0.905

[1] The t-test null hypothesis is H_0: The greenwashing premium is $= 0$ from 2005 to 2021. The alternative hypothesis is greenwashing premium is $\neq 0$. The null hypothesis is not rejected with t-stats probability > 0.05 at the 95% confidence level.

[2] Whilst the average greenwashing premium is historically not significant across all industries from 2005 to 2021, there have been episodes when firms find it profitable to make overstatements in their ESG claims.

[3] The consumer industries have the environmental and the governance pillars with the largest premium. For the green industries, the environmental and social pillars have the largest premium. For brown industries, the premium are relatively smaller for the S and G pillars and negligible for the E pillar.

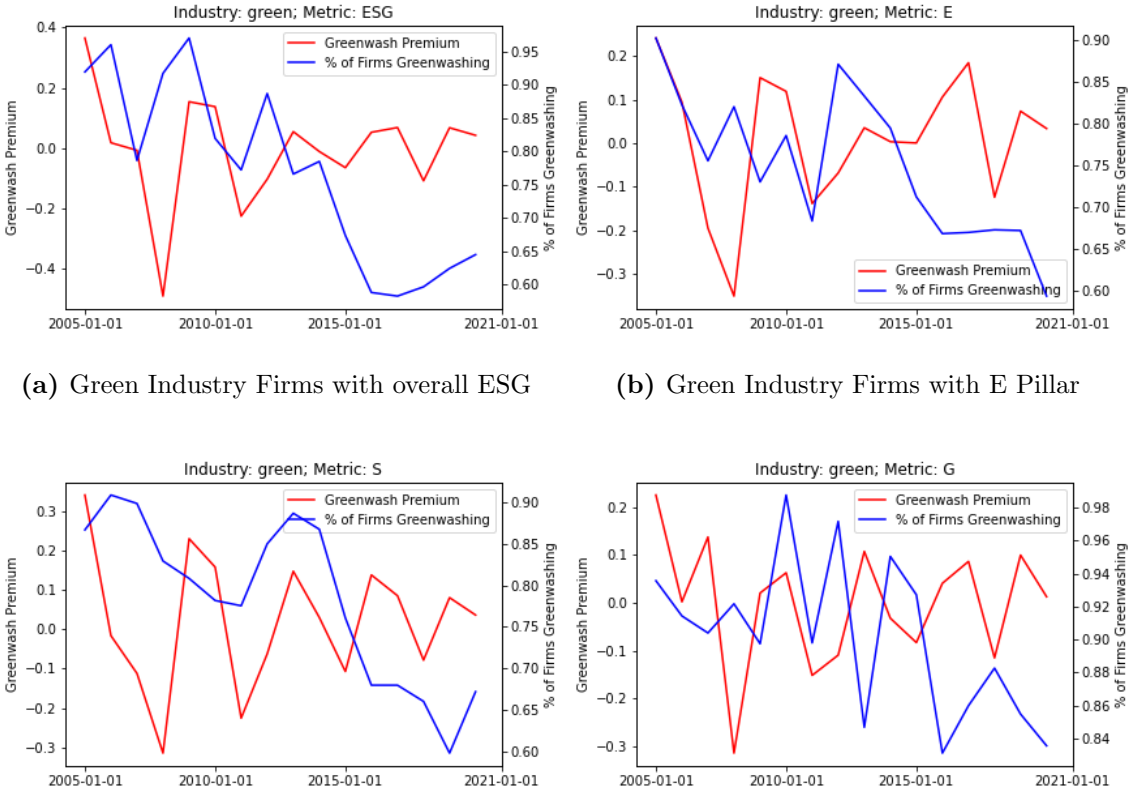


(a) ESG Scores **(b)** Environmental Pillar

(c) Social Pillar **(d)** Governance Pillar

Figure 3. Time Series of ESG Scores - TRMA and standard ASSET4 scores

The plots show the time series graphs of the average TRMA and standard ASSET4 scores from 2005 to 2021 Jan for the overall ESG scores and the individual pillars. The TRMA scores are averaged from firm-specific news sentiment scores on ESG topics obtained by NLP. The standard ASSET4 scores represent publicly verifiable scores on individual firms' ESG performance. The correlations between the 2 series for the ESG, E, S and G pillars are relatively high and are 0.63, 0.87, 0.50 and 0.53 respectively. The news sentiment scores are invariably higher than the performance scores. The standard deviation of these scores are similar in value and $\approx$ 17 to 23.



(a) Green Industry Firms with overall ESG

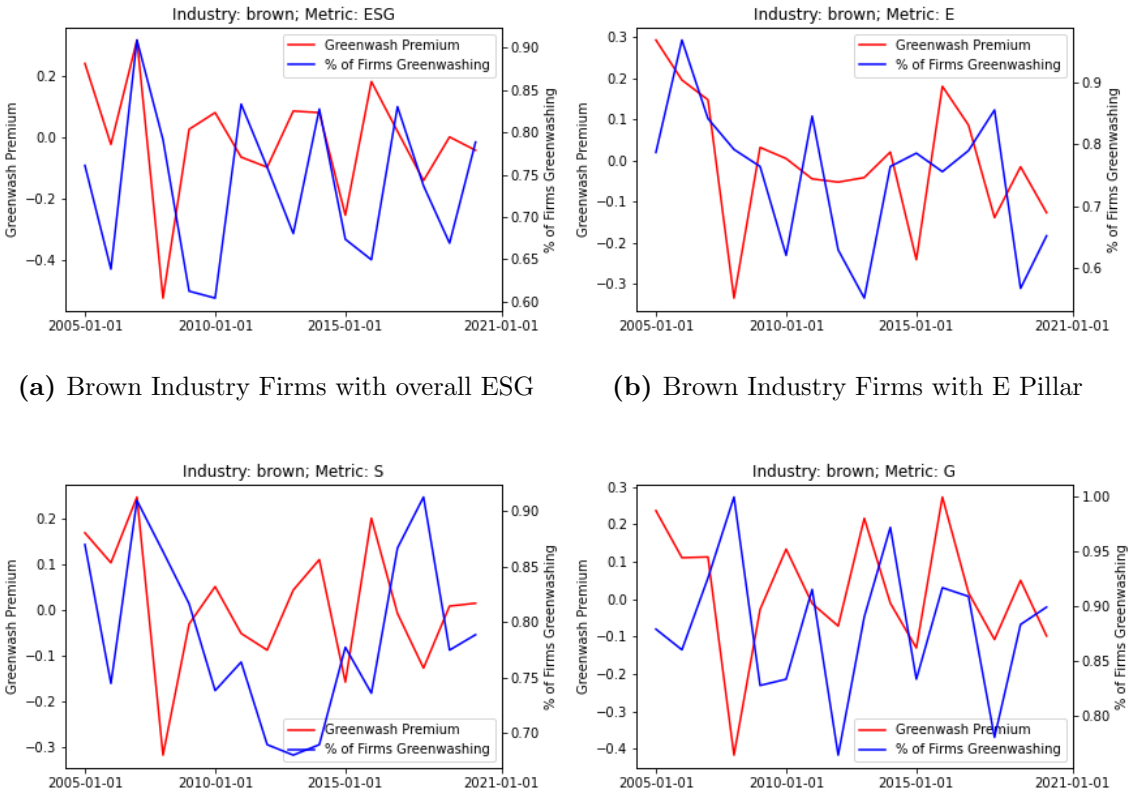
(b) Green Industry Firms with E Pillar

(c) Green Industry Firms with S Pillar

(d) Green Industry Firms with G Pillar

Figure 4. Time Series of Greenwashing Premium for Green Industry & % of Firms that greenwash

The plots show the time series graphs of the greenwashing premium for firms in green industry and % of firms that had greenwashed conditional on positive news sentiment announcement at beginning of period. The peaks are observed around the years 2012/2013 and 2016 (post Paris COP 21) notwithstanding the Lehman crisis years in 2008/2009..



(a) Brown Industry Firms with overall ESG

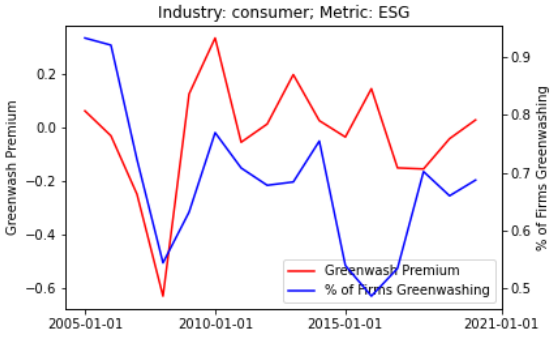
(b) Brown Industry Firms with E Pillar

(c) Brown Industry Firms with S Pillar

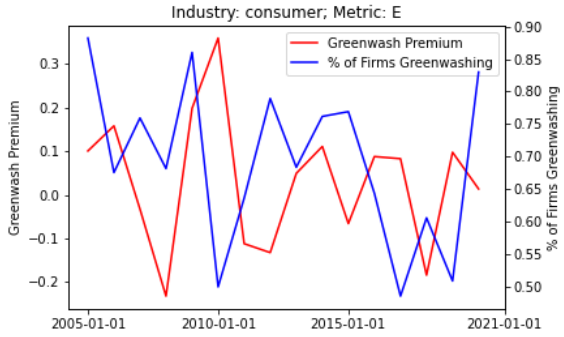
(d) Brown Industry Firms with G Pillar

Figure 5. Time Series of Greenwashing Premium for Brown Industry & % of Firms that greenwash

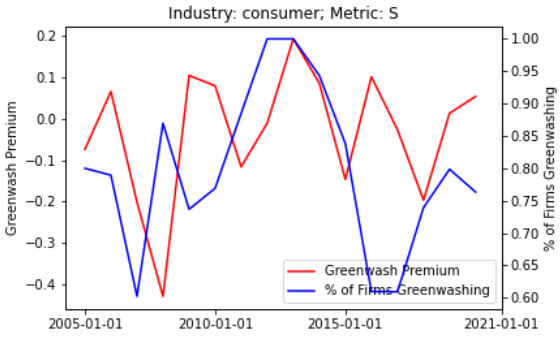
The plots show the time series graphs of the greenwashing premium for firms in brown industry and % of firms that had greenwashed conditional on positive news sentiment announcement at beginning of period. Similarly, peaks are observed around the years 2012/2013 and 2016 (post Paris COP 21) notwithstanding the Lehman crisis years in 2008/2009..



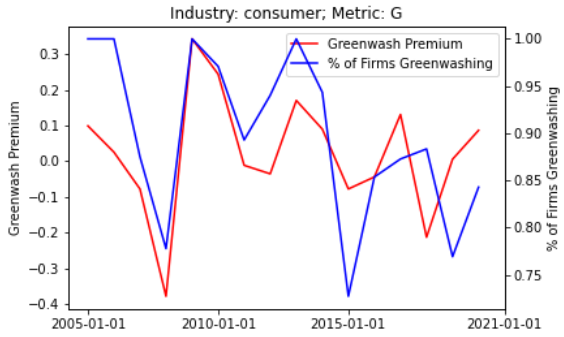
(a) Consumer-Orientated Firms with overall ESG



(b) Consumer-Orientated Industry Firms with E Pillar



(c) Consumer-Orientated Industry Firms with S Pillar



(d) Consumer-Orientated Industry Firms with G Pillar

Figure 6. Time Series of Greenwashing Premium for Consumer Industry & % of Firms that greenwash

The plots show the time series graphs of the greenwashing premium for firms in consumer industry and % of firms that had greenwashed conditional on positive news sentiment announcement at beginning of period. Similarly, peaks are observed around the years 2012/2013 and 2016 (post Paris COP 21) notwithstanding the Lehman crisis years in 2008/2009..

Notes

[1]https://www.esgtoday.com/european-regulators-launch-greenwashing-study/

[2]https://news.un.org/en/story/2022/11/1130317

[3]Climate pledges from top companies crumble under scrutiny https://www.nature.com/articles/d41586-022-00366-2

[4]https://www.ft.com/content/f1367ab4-ac6f-486d-8bd2-e7659448055d

[5]https://www.spglobal.com/ratings/en/research/articles/210823-the-fear-of-greenwashing-may-be-greater-than-the-reality-across-the-global-financial-markets-12074863

[6]https://www.economist.com/leaders/2021/05/22/sustainable-finance-is-rife-with-greenwash-time-for-more-disclosure

[7]https://news.un.org/en/story/2022/11/1130317

[8]https://www.bloomberg.com/news/articles/2022-10-17/companies-keep-climate-goals-secret-as-green-hushing-takes-off

[9]https://www.refinitiv.com/en/sustainable-finance/esg-scores

[10]https://www.ft.com/content/79851eee-d9e6-4ceb-be16-e9cf8b8c4ddf

[11]https://www.forbes.com/sites/ikebrannon/2022/06/14/litigation-on-us-energy-companies-threatens-more-than-just-them/

[12]https://www.climate.gov/news-features/understanding-climate/state-climate-2012-highlights

[13]https://www.newsd.admin.ch/newsd/message/attachments/74580.pdf

[14]https://truthinadvertising.org/articles/six-companies-accused-greenwashing/