Corporate Social Responsibility and Financial Reporting Quality: Evidence from CSR-Disasters

Juan Manuel Garcia Lara, Beatriz Garcia Osma, Irina Gazizova*
Universidad Carlos III de Madrid

ABSTRACT

We introduce the concept of CSR-disasters and argue that these events trigger attention for CSR performance. CSR-disasters are large technological disasters that can be attributed to particular firms. We propose that after CSR-disasters, stakeholders become more critical of CSR performance in firms operating in disaster-affected industries. We show that firms with low pre-disaster CSR performance significantly improve CSR in the post-disaster period, mainly through window dressing practices. For these firms, we find that higher window dressing CSR leads to higher earnings management (proxied by absolute value of discretionary accruals) and lower quality of narrative disclosure (proxied by Bog Index). This suggest that exogenously-driven CSR practices may lead to unfavorable outcomes, in particular impair transparency.

Keywords: Corporate Social Responsibility, Financial Reporting Quality, Earnings Management, CSR-Disasters, Technological Disasters.

*Corresponding author: Irina Gazizova Dept. Economia de la Empresa. Universidad Carlos III de Madrid, calle Madrid 126, 28903 Getafe, Madrid (Spain).E-mail: igazizov@emp.uc3m.es

1 Introduction

The debate over the desirability of corporate social responsibility (CSR) practices (Ferrell et al., 2016; Hart and Zingales, 2017) has resulted in the mandatory implementation of social obligations for the private sector in some countries.[1] In the recent spate of technological disasters, proponents of obligatory state-regulated CSR raise with new vigor. For instance, after the BP oil spill of 2010, President Obama argued that "*one of the lessons we've learned from this spill is that we need better regulations, better safety standards, and better enforcement when it comes to offshore drilling*."[2] Further, Mr. Obama "*acknowledged that federal agencies had failed to ensure that safety and environmental standards were being met and announced a thorough review of the oversight process*."[3] This implied that not only BP, but all *other* firms dedicated to offshore drilling were also under suspicion of lacking social responsibility.

Despite the salience of CSR issues, there is a shortage of empirical evidence on whether enforced and voluntary CSR practices have identical properies and lead to similar social surplus. In this study, we introduce the concept of CSR-disasters and argue that these disasters enable us to isolate plausibly exogenous increases in socially enforced CSR performance for firms that operate in affected industries. These settings allow us to study how socially enforced increases in CSR performance affects financial reporting quality (FRQ).

Prior studies have explored the link between CSR and FRQ, focusing in particular on earnings management practices. For instance, Kim et al. (2012) argue that CSR firms are more critical of ethical issues including incidents of earnings management. For CSR firms, these authors find limited evidence of aggressive earnings management through abnormal accruals and real activities manipulation. However, one could argue that in equilibrium, firms voluntarily and

[1]For instance, in 2013, the Government of India (Ministry of Corporate Affairs) enacted Section 135 of the Indian Companies Act that obligates all companies that meet specified financial thresholds to spend 2% of average net profits on CSR. https://www.mca.gov.in/Ministry/pdf/FAQ_CSR.pdf

[2]The Telegraph (June 16th, 2010) https://telegraph.co.uk/finance/newsbysector/energy/oilandgas/7831643/BP-oil-spill-Barack-Obamas-speech-in-full.html

[3]The New York Times (May 14th, 2010) https://nytimes.com/2010/05/15/us/politics/15obama.html

simultaneously choose to engage in CSR and avoid earnings management. If that is the case, it might be problematic to infer whether high CSR performance in fact affects FRQ.

To address this issue, we focus on a simple framework. We argue that CSR-disasters plausibly provide a randomly assigned increase in CSR performance (see following discussion) and thereby help to eliminate the problem of simultaneity. In other words, this setting allows us to attribute any change in FRQ to the influence of CSR. Further, observing *ex-post* the sign of the change (or no change) in FRQ helps to distinguish between "true"[4] and window dressing CSR, where, following prior literature, we define CSR as voluntary actions that improve social conditions and that are not required by the law and extend beyond firm's profit maximization (McWilliams and Siegel, 2000; Godfrey et al., 2009).

We use a set of technological disasters as our proxy for CSR-disasters. By definition, a technological disaster refers to a failure of a technological structure or/and human error in controlling or using the technology.[5] Examples of technological disasters can be explosions, chemical spills, or gas leaks.[6] We focus on major technological disasters that were not caused by the formal violation of the law or malice, but rather, by a set of failures in meeting technological, environmental and ethical standards. As an illustrative example, consider the garment factory collapse in Rana Plaza, Dhaka on April 24, 2013. If the clothing companies that operated in the Bangladesh factory (Primark or Canada's Loblaw) had gone further than just meeting their formal obligations, it is likely that the collapse could have been prevented. For instance, instead of blindly accepting the building certificate for a Bangladesh factory, clothing firms could have sent "people to check

[4]CSR that is associated with real changes in corporate ethical standards.

[5]Following the definition proposed by The Institute of Food and Agricultural Sciences (IFAS) a technological disaster "...Is an event caused by a malfunction of a technological structure and/or some human error in controlling or handling the technology. The effects of a disaster on families and individuals may be long lasting and can endure for years. However, symptoms may appear gradually, and impacts may not be seen immediately". For more details follow `http://edis.ifas.ufl.edu/`. In this study terms "disaster", "disaster event", and "catastrophe" are used interchangeably.

[6]For example, Pek et al. (2018) classify the following events as technological disasters: chemical spills, collapses, explosions, fires, gas leaks, poisonings, radiation leaks, and large-scale transportation accidents. Perhaps the most famous technological disasters are The BP oil spill on April 20, 2010 and The Fukushima Daiichi nuclear incident on March 11, 2011.

every pillar."[7] Key to our identification strategy is that the collapse provoked a reaction not only from the firms that directly operated in Rana Plaza, but also from other firms in the industry. More than 150 companies signed a legally binding agreement (The Accord on Fire and Safety in Bangladesh) and 27 US brands signed their own non-legally binding industry-led version (The Alliance for Bangladesh Worker Safety).[8] These agreements facilitate worker-management committees in factories and obligate companies to independently inspect factories, provide transparent reports, and, if necessary, financially contribute to fix detected problems.[9]

In this study, we define CSR-disasters as technological disasters that (1) are sufficiently large to affect the whole disaster-affected industry; and (2) that, plausibly, could have been prevented or mitigated if a firm had gone further than just meeting the minimum formal obligations imposed by law. As a result of these CSR-disasters, same-industry firms (treated firms) (1) are exposed to a negative stakeholder reaction (negative shock); and (2) may undertake efforts to mitigate it (Blacconiere and Patten, 1994; Desai, 2011; Diestre and Rajagopalan, 2014; Pek et al., 2018).

In our tests, we exploit a set of major industrial catastrophes, as reported by The International Disaster Database[10] (EM-DAT) that occurred between 2004 and 2012 in the US. We measure CSR performance using KLD data. Using a research design similar to Flammer (2015), we apply a differences-in-difference approach to estimate the effect of the disasters on CSR. More specifically, if a firm operates in an industry that is exposed to technological disaster, we compute the difference in CSR before and after the catastrophe. Then we compare this difference with the

[7]The Economist (May 4th, 2013). https://www.economist.com/leaders/2013/05/04/disaster-at-rana-plaza.

[8]The Accord on Fire and Safety is a "legally binding agreement between companies and unions where companies commit to independent inspections and transparent reporting, including developing strong worker-management committees in factories" The Guardian (April 2nd, 2014). https://theguardian.com/sustainable-business/bangladesh-factory-collapse-10-things-changed

[9]In addition to the agreements, a dozen global brands (as the Gap, H&M, Mango, and Walmart) have contributed $21.5m to Rana plaza Donors Trust Fund, which was set up to award compensations to victims and their families. (The Guardian, April 24th, 2015). https://theguardian.com/sustainable-business/2015/apr/24/bangladesh-factories-building-collapse-garment-dhaka-rana-plaza-brands-hm-gap-workers-construction.

[10]http://www.emdat.be/ EM-DAT, 2018 EM-DAT is widely used in the literature (e.g. Evan et al., 2011; Lutz et al., 2014; Lesk et al., 2016) and well-known as one of the most comprehensive databases on disaster events in the world (Voigt et al., 2015).

corresponding difference in industries that are not affected by the catastrophe.

We first document that treated firms experience decreases in CAR, ROA, ROE, and sales growth after CSR-disasters. Our results are consistent with previous event studies (Blacconiere and Patten, 1994; Heflin and Wallace, 2017; Pek et al., 2018), and serve to validate our expectations over the effect of CSR-disasters on the *other* firms in the industry.

We argue that a CSR-disaster is a negative shock to the relationship between a firm and its stakeholders. Given that stakeholders' positive attitude in the form of firm's social capital has a valuable effect on firm's financial performance (Lev et al., 2010; Cheng et al., 2014; Shiu and Yang, 2017), managers have to undertake actions to restore this relationship. Thus, we expect that firms improve CSR performance in the post-disaster period. We propose two possible mechanisms through which strengthening CSR performance may lead to strengthening the relation between a firm and its stakeholders. First, because high CSR performance can help to differentiate treated firms from the guilty firm by signaling the low operational risk and high preparation for the possible regulatory changes associated with the disaster (Heflin and Wallace, 2017). Second, an increase in CSR performance can signal firms' social awareness and high environmental and social standards that do not necessarily mitigate firms' risk. Prior literature suggests that CSR helps to build social capital and to form stakeholders' positive attitudes, which mitigates negative market reaction at the time of the disaster (Godfrey, 2005; Godfrey et al., 2009). We find strong evidence that treated firms improve CSR performance in the post-disasters periods.

Next, we address whether firms' response to a CSR-disaster is sensitive to the pre-disaster level of CSR. Prior literature suggests that firms with previously accumulated social capital in the form of high CSR performance can mitigate negative market reactions because (1) market participants expect that these firms have lower costs associated with the disaster (Godfrey et al., 2009); and (2) because these firms have social trust and stakeholders' loyalty (Godfrey, 2005; Shiu and Yang, 2017). Further, incremental increases in CSR performance may not be equally useful for firms with different pre-disaster CSR performance. For instance, Clarkson et al. (2004)

show that in the pulp and paper industry only low-polluting firms [11] extract incremental economic benefit from environmental expenditures. We show that only firms with low pre-disaster CSR improve CSR in the post-treatment periods. Specifically, firms in the low pre-treatment CSR quintiles add (eliminate) two strength (concerns) dimensions to (from) the total CSR score.

Finally, we study whether enforced CSR performance leads to better FRQ. Following Kim et al. (2012) we argue that improvement in "true" CSR performance leads to enhancement in all corporate ethical standards. Firms that exert efforts and spend resources to achieve high social, ethical, and environmental standards may apply these standards to all their business decisions, including financial reporting. For instance, Atkins (2006) argue that being socially responsible means being transparent in firms' financial reporting. Conversely, if CSR-disasters generate entirely window-dressing improvements in CSR performance, then there will be a negative or no relationship between CSR and FRQ in the post-disaster period.

We use earnings management and narrative disclosure quality as two alternative proxies for financial reporting quality. We measure earnings management using discretionary accruals (Jones, 1991; Subramanyam, 1996; DeFond and Subramanyam, 1998; Kothari et al., 2005) and real activities manipulation proxies (Roychowdhury, 2006). We employ Bog Index as a proxy for the quality of narrative disclosure (Bonsall et al., 2017). Our results show that for firms with low pre-disaster level of CSR, there is a negative relationship between CSR and FRQ in the post-disaster period. These firms increase earnings management through abnormal discretionary accruals and have lower readability of their disclosure.

Our findings indicate that treated firms respond to CSR-disasters by improving subsequent CSR performance and that this exogenous increase in CSR leads to deterioration in FRQ. We contribute to the literature in several ways. First, we add to the literature that studies how technological disasters shape firms behavior (Desai, 2011; Diestre and Rajagopalan, 2014; Heflin and Wallace, 2017). Specifically, we show that CSR is a channel through which firms can respond to disasters. Second, our paper extends prior research that studies the link between CSR and

[11] e.g. firms with high environmental CSR performance

FRQ (Kim et al., 2012). Our contribution over Kim et al. (2012) is that we eliminate the problem of simultaneity. In other words, Kim et al. (2012) show that, on average, firms with high CSR performance are associated with better FRQ. This study focuses on the question whether *changes* in CSR performance lead to improvements in firms' FRQ. We show that increases in CSR performance driven by subset of technological disasters (CSR-disasters) lead to lower levels of FRQ. This result is important in light of the current discussion about whether firms should be required to spend a share of their profits on CSR policies. Finally, our results contribute to the literature that studies why during crisis periods some firms perform better than their industry-peers (Godfrey et al., 2009; Lins et al., 2017).

The paper proceeds as follows. In Section II we discuss the related literature and develop our hypotheses. In section III we describe the data and methodology, while we present results in Section IV. We conclude in Section V.

2 Literature Review and Hypotheses Development

To derive theoretical predictions on the firms' reaction to the CSR disasters, we draw from different strands of literature. We begin this section by analyzing prior literature to argue that affected firms can use CSR as a response to CSR-disasters. Next, we hypothesize how the pre-disaster level of CSR influences the response to these disasters. Finally we investigate how changes in CSR performance affect firms' FRQ.

2.1 Firms' Response to CSR-Disasters through Corporate Social Responsibility

Prior literature documents that natural disasters generate waves of corporate donations and subsequent increases in CSR performance (Muller and Kräussl, 2011; Madsen and Rodgers, 2015). In this section, we ask whether CSR-disasters provoke subsequent improvement in CSR performance in firms that are exposed to the negative spillover effects.

CSR-disasters also generate negative spillover effects that can be substantial for firms in the same-industry.[12] For instance, as a result of CSR-disasters, same-industry firms (treated firms) are exposed to negative consequences such as negative abnormal stock returns (Diestre and Rajagopalan, 2014; Heflin and Wallace, 2017) and higher scrutiny from regulators (Desai, 2011). In our framework, if it is revealed through a highly visible CSR-disaster that one firm in the industry neglects CSR standards in environmental, safety or technological related matters, *other* firms in the same industry become suspect of also violating these norms. Thus, we study how firms that experience negative spillover effect respond to CSR disasters.

A CSR-disaster is plausibly a negative shock for the relation between the firm and its stakeholders. We propose two possible mechanisms that explain why an improvement in CSR performance in the post-disaster period may lead to an improvement in the firm-stakeholders relation.

First, firms may improve CSR performance to signal that they are less risky and have high quality operational processes and, overall, to differentiate themselves from the guilty firm in terms of safety. Heflin and Wallace (2017) propose that large technological disasters update investors' expectations on the likelihood of the recurrence of the disaster and the following regulatory changes. Further, the authors study the case of the BP oil spill in 2010 and show that firms in the oil and gas industry improve their environmental performance in the post-disaster period. Heflin and Wallace (2017) argue that firms improve their CSR performance in the post-disaster period to signal their readiness for possible regulatory changes. Overall, this leads to the empirical prediction that "true" CSR-performance increases after the CSR-disasters.

Further, we draw from the research that argues that CSR can contribute to firms' positive image and by doing so can add to the relationship with stakeholders. In other words, we argue that to

[12]The phenomenon that one firm's deviant behavior can result in the punishment of other (not responsible) firms in the same industry is discussed in prior studies (Desai, 2011; Diestre and Rajagopalan, 2014; Liang and Renneboog, 2017). Diestre and Rajagopalan (2014) suggest that in the short run, market participants tend to form their beliefs, based on the highly visible and available information, such as belonging to one industry. The authors provide two reasons for spillover. First, firms in one industry may have the same third party relationship (e.g. relationship with suppliers) which may cause the accident. Second explanation goes through sociocognitive literature. In short, external audiences predict organizations' future behavior based on the other firms' behavior in the same industry (for more details, please follow Diestre and Rajagopalan (2014), p. 1130-1131.).

be useful, improvement in CSR performance does not have to directly reduce firms' operational risk. Firms can simply improve CSR to signal their ethical and environmental standards (window dressing CSR). Uzzi (1997) and Godfrey et al. (2009) argue that CSR performance creates the moral capital that helps to improve the relationship with stakeholders. Further, using a risk management model, Godfrey (2005) shows that this moral capital can contribute to shareholders' wealth in the time of disasters. We hypothesize that in firms with higher CSR performance stakeholders are more loyal in the post-disaster period and, thus, impose lower penalties.

Based on prior literature, our prediction is that after CSR-disasters, affected firms (treated firms) increase CSR to signal their operational quality and social awareness and, thus, to improve their relationship with stakeholders and subsequent financial performance. This hypothesis, stated in its alternative form, is as follows:

H1: *Firms in the industry that experience a CSR-disaster (treated firms) improve their CSR performance in the post-disaster period.*

2.2 Pre-disaster Level of CSR Performance

Firms with different pre-disaster levels of CSR may differently react to CSR-related negative externalities, and improve their CSR to different extents in a post-disaster period. This would hold, first, if firms with high pre-disaster CSR performance do not suffer from negative market reactions. Then, these firms would not be incentivized to change their CSR performance. Godfrey (2005) and Godfrey et al. (2009) introduce the view that CSR performance has an "insurance-like effect," whereby CSR performance limits the potential negative impact on stock price of negative events related to corporate operations. In other words, CSR expenditures act as an insurance premium that the firm pays to avoid market losses in the case of a negative event. Firms with strong reputation for CSR suffer less because (1) they are expected to have lower costs associated with the disaster in the future (Godfrey et al., 2009); and (2) because they have accumulated "moral reputational capital" (Godfrey, 2005; Shiu and Yang, 2017).

In line with this view, Lins et al. (2017) argue that high CSR performance accumulates firm-specific social capital in the form of trust between a firm and both its stakeholders and investors. This social capital pays off during periods when the overall level of trust in corporations is low. Further, these authors show that firms with high CSR performance outperform their peers during the 2008-2009 financial crisis. In related research, Muller and Kräussl (2011) show that the more a firm is known for socially irresponsibility the greater was the negative impact of Hurricane Katrina on the stock prices and greater the probability that this firm improved its CSR performance in the post-Katrina period (through corporate philanthropic disaster response). Heflin and Wallace (2017) argue that firms in the oil and gas industry with high environmental disclosure before the BP oil spill of 2010 experience less negative equity share price changes because market participants expect that the costs of the disaster will be lower for these firms. Further, they show that firms with poor environmental disclose in the pre-spill period improve their disclosure in the post-disaster period. The authors also show that this improvement in disclosure is not entirely window dressing and that it is associated with an improvement in environmental performance.

A second reason why firms may differently react to CSR-disasters is that marginal increases in CSR performance may be not equally useful for firms with different pre-disaster CSR performance. For instance, Clarkson et al. (2004) show that the market does not equally value environmental expenditure for different firms in the pulp and paper industry. Specifically, only low-polluting firms extract incremental economic benefit from environmental expenditures. The market does not value environmental expenditures for high-polluting firms and further assess them by the existence of unbooked environmental liabilities.

Based on this prior literature we assume that firms with lower-CSR performance before the disaster have more incentives to improve firm-stakeholders relationships and, thus, they will improve their CSR more in the post-treatment period. Given these, our second hypothesis is:

H2: *Firms in the industry that experience a CSR-disaster (treated firms) with lower pre-disaster CSR performance improve their CSR more in the post-disaster period.*

2.3 Corporate Social Responsibility and Financial Reporting Quality

Following Kim et al. (2012) we argue that improvements in CSR performance lead to enhancement in all corporate ethical standards. Firms that exert efforts and spend resources to achieve high social, ethical, and environmental standards may use these standards in all their business decisions, including financial reporting. For instance, Atkins (2006) argue that being socially responsible means being transparent in financial reporting. Thus, if CSR-disasters cause increases in "true" CSR performance, we should observe an improvement of FRQ.

Alternatively, if CSR-disasters induce purely "window dressing" CSR, then there will be a negative or no effect on FRQ. The main reason for this could be the fact that "window dressing" CSR does not lead to improvement in corporate ethical standards or may even be associated with managerial misbehavior. For instance, Petrovits (2006) suggests that managers manipulate payments to their corporate foundations to manage financial reporting targets. In particular, firms use a cookie jar strategy to maintain strings of increasing earnings. Prior et al. (2008) show that managers that engage in earnings management have incentives to use CSR to please stakeholders and thereby enlist their support. Given this mixed evidence on the relationship between CSR and FRQ we formulate the following hypothesis:

H3: *Increases in CSR performance that are induced by CSR-disasters affect FRQ.*

3 Data and Sample Selection

3.1 International Disasters Data

To construct a set of CSR-disasters we use major technological disasters in US we use The International Disaster Database (EM-DAT). This database provides information about natural (geophysical, meteorological, hydrological, etc.) and technological (industrial, transport, and miscellaneous) disasters. Each event is accompanied by information on date and type of the event, country name, location, total deaths, total number of people affected, and total damage

(USD).[13] Each event in EM-DAT meets at least one of the following criteria: over 10 deaths, over 100 people affected (and/or injured, homeless), and a request for international assistance and/or declaration by the government of a state of emergency.

For the period from 2004 to 2012 EM-DAT provides three major (classified by total damage) technological disasters.[14] The most harmful one is the BP oil spill of April 20, 2010 (EVENT I). The rig explosion was owned by Transocean and drilling for BP. This explosion killed eleven people and caused a damage of over $20 billion. After the oil well explosion, 4.9 million barrels of oil and gas leaked into the Gulf of Mexico. Following prior studies (Heflin and Wallace, 2017), we consider industries with SIC codes 13 (Oil and Gas Extraction) and 29 (Petroleum Refining and Related Industries) as "treated" by this disaster.

According to EM-DAT, the second largest technological disaster is the Georgia Sugar Refinery Explosion. On February 7, 2008, dust exploded on the Imperial Sugar refinery in Port Wentworth, Georgia. The accident caused 13 deaths, injuries to 40 people, and damage over $323 million. According to the investigation by the Occupational Safety and Health Administration (OSHA) and the Bureau of Alcohol, Tobacco, Firearms and Explosives sugar dust was the fuel for the fire. The disaster occurred in the Cane Sugar Refining industry (SIC Code 2062) and was directly caused by the process of sugar refinery. Other industries within SIC code 20 are not linked to refinery processes, therefore only SIC 2062 firms are considered to be affected by the disaster. Given that observations with SIC code 2062 are missing in our sample we do not include this disaster in our event study.[15]

The third largest event is the San Bruno Gas Pipeline Explosion. On September 9, 2010, nat-

[13] Total deaths (definition considered in EM-DAT): " [. . .] it is the sum of deaths and missing". Total affected (definition considered in EM-DAT): " [. . .] it is the sum of the injured, affected and left homeless after a disaster" (`http://www.emdat.be`).

[14] The search result is presented in Appendix B

[15] Industries within SIC code 20 (Food and Kindred Products): 201 – Meat products; 202 – Dairy Products; 203 – Canned, Frozen, and Preserved Fruits, Vegetables, and Food Specialties; 204 – Grain Mill Products; 205 – Bakery Products; 206 – Sugar and Confectionery Products; 207 – Fats and Oils; 208 – Beverages; 209 – Miscellaneous Food Preparations and Kindred Products. Within SIC code 206 only industry 206 only industry 2062 relates to sugar refinery. Rest industries are the following: 2061 – Cane Sugar, except Refinery; 2063 – Beet Sugar; 2065 – Candy and other Confectionery Products; 2066 – Chocolate and Cocoa Products; 2067 – Chewing Gum; 2068 – Salted and Roasted Nuts and Seeds.

ural gas pipeline, owned by Pacific Gas and Electric (PG&E, primary SIC Code 4931 – Electric and other Services Combined) company, exploded in San Bruno, California (EVENT II). The explosion killed eight people. In 2012, an independent audit from the State of California issued a report claiming that PG&E illegally used over $100 million from a fund used for safety operations. The company spent this money for executive compensation and bonuses. PG&E is not able to approach the source of the explosion and find out the cause of the accident. In 2014 federal grand jury in U.S. District Court, San Francisco, indicted PG&E for violations of the Natural Gas pipeline Safety Act of 1968. We consider all firms with 49 SIC code as "treated" because the source of the explosion was not found.

Finally, we include one event that is not presented in the EM-DAT as a technological disaster, but to our knowledge, perfectly suits to our research settings. Our third event is the Fukushima Daiichi Nuclear Disaster that occurred on March 11, 2011 in Japan (EVENT III). EM-DAT does not classify this event as "technological" for the following reasons. First, this event occurred in Japan. However, the consequences of the disaster affected the energy sector worldwide. Second, initially, it was caused by the tsunami which, in turn, was caused by The Great East Japan earthquake. However, according to the Fukushima Nuclear Accident Independent Investigation Commission, the nuclear accident was foreseeable. The plant operator, Tokyo Electric Power Company (TEPCO) failed to take basic security measures. Further, Naomi Hirose, TEPCO president, admitted that "even if a tsunami caused the accident, we are the operator of the Fukushima nuclear power plant and we do take responsibility."[16] Thus, this event meets our two necessary conditions that are a) to be caused by business entities, and b) to be sufficiently large. We consider all firms with 49 SIC code as "treated."

[16]How Can Companies Take Responsibility for Major Accidents? (2015, August 4). Yale Insights. https://insights.som.yale.edu/insights/how-can-companies-take-responsibility-for-major-accidents.

3.2 Firm-level Data

CSR data is obtained from MSCI (formerly KLD). KLD covers the largest 3000 U.S. publicly traded companies by market capitalization.[17] KLD ratings are well known and widely used in the CSR literature (e.g., Godfrey et al. (2009); Barnett and Salomon (2012); Flammer (2015); Lins et al. (2017); Fernando et al. (2017)). KLD provides information on how firms address the needs of their stakeholders along different social dimensions, such as environment, community, human rights, employee relations, diversity, products, corporate governance, and controversial business issues, including alcohol, gambling, firearms, military, nuclear power, and tobacco. Each social dimension is twofold and has both strength and concern components.

We obtain accounting data from Compustat. For data on financial variables, we use CRSP. Bog index is coming from Bonsall et al. (2017). Consistent with the previous research (Kim et al., 2012), we exclude financial firms (SIC codes 6000-6999). All continuous variables are winsorized at the top and bottom 1 percent of their distributions. Although the exact number of observations depends on the specific regression, the baseline sample for which we estimate the equations contains 16,281 firm-year observations for the period 2003 - 2013.

3.3 Measurement of CSR and FRQ

3.3.1 CSR Measures

To construct our CSR proxy (CSR_SCORE), we follow Kim et al. (2012) and subtract concern-related measures from strength-related ones among five social dimensions: environment, community, employee relations, diversity, and product. In 2010 the industry-based key issue rating model was introduced to KLD.[18] We circumvent this potential problem by scaling each KLD

[17] Prior 2003 the composition of the covered firms was different. For more details follow Appendix C

[18] Prior 2010 all of the positive ESG performance indicators were searched for all of the companies. Starting from 2010, all companies are assessed for limited set of industry specific positive ESG indicators. For more details follow Appendix C.

dimension by the maximum value of this dimension within the year.[19]

The use of the aggregate proxy for CSR performance (CSR_SCORE) has been widely criticized (Entine, 2003; Godfrey et al., 2009). The primary criticism stresses that the information value in the individual social categories can be destroyed as a result of subtracting concerns from strengths, as well as summing different social dimensions. Further, Lins et al. (2017) emphasize that some components of CSR_SCORE could be more critical to enhancing the trust of all of a firm's stakeholders than others. Thus, following Lins et al. (2017), we desegregate CSR_SCORE into two parts: those that mainly beneficial for external stakeholders (EXTERNAL) and those that mainly beneficial for internal stakeholders (INTERNAL).

To measure financial reporting quality (FRQ), we use a number of earnings management and narrative disclosure proxies that we describe in detail next.

3.3.2 Earnings Management Measures

We measure earnings management using discretionary accruals proxies (Jones, 1991; Subramanyam, 1996; DeFond and Subramanyam, 1998; Kothari et al., 2005) as wall as real activities manipulation proxies based on the work of Roychowdhury (2006). According to Healy and Wahlen (1999), "Earnings management occurs when managers use judgment in financial reporting and in structuring transactions to alter financial reports to either mislead some stakeholders about the underlying economic performance of the company or to influence contractual outcomes that depend on reported accounting numbers" (p. 368). Following prior literature on CSR and FRQ (Kim et al., 2012) we measure earnings management through discretionary accruals and employ a cross-sectional version of the modified Jones model with the past-year return on asset (ROA). To address the issue that earnings management can be based on income-increasing or income-decreasing accruals (Warfield et al., 1995; Klein, 2002), we use absolute value of discre-

[19]For instance, in 2010 the maximim value of community strength dimension (com_str_num) is 4 (firm with cusip = 45814010). Then, in 2010 we scale all community strength values by 4 (com_str_num/4).Thus, the scaled KLD dimensions have values from 0 to 1. To have comparable with prior literature summary statistics in Table (2) we present unscaled CSR_SCORE.

tionary accruals for our analysis (ABS_DA).

Roychowdhury (2006) defines real activities manipulation (RAM) as "management actions that deviate from normal business practices, undertaken with the primary objective of meeting certain earnings thresholds" (p. 336). Following (Roychowdhury, 2006) we measure RAM through (1) abnormal cash flow from operations (AB_CFO), (2) abnormal production costs (AB_PROD), and (3) abnormal discretionary expenses (AB_EXP). Further, following Kim et al. (2012) we calculate fourth RAM proxy (COMBINED_RAM) as a linear combination of above three[20].

3.3.3 Narrative Disclosure Quality

Lexical properties, or readability, of disclosure is an important dimension of FRQ (Li, 2008). The term "readability" refers to how complex a text is and how difficult it is to read it and extract necessary information.[21] In 1998 SEC issued the handbook promoting companies to use "plain English" in writing all publicly disclosed documents. The SEC encourages the use of short sentences, everyday words and active voice. Authors should be confident that the final version of a document captures the original meaning, and it is written in the easiest possible way. With this document SEC emphasizes importance of lexical properties of financial disclosure and the effect it may have on investors and markets.

We use the Bog Index as our main proxy for readability (Bonsall et al., 2017). This Index captures the processing costs associated with the type of language used in financial disclosure. Bog Index captures linguistic attributes such as sentence length, passive voice, weak verbs, overused words, complex words, and jargon (Bonsall et al., 2017). In contrast to Fog Index (Li, 2008;

[20] $COMBINED\_RAM = AB\_CFO - AB\_PROD + AB\_EXP$ By the construction, $COMBINED\_RAM$ increases as firms constrain their RAM.

[21] The Cambridge dictionary suggests the following meaning: "Easy and enjoyable to read" (http://dictionary.cambridge.org/dictionary/english/readable). The Oxford dictionary offers a similar, but slightly different definition: "The quality of being legible or decipherable" (http://www.oxforddictionaries.com/definition/english/readability).

Loughran and McDonald, 2011), which counts all multi-syllabic words as complex, Bog Index measures words familiarity base on a proprietary list of over 200,000 words. A higher value of Bog Index refers to a lower level of financial disclosure readability.

3.4 Empirical Models

3.4.1 Differences-in-Difference Approach

First, we validate whether firms experience a deterioration in profitability and growth opportunity. Heflin and Wallace (2017) document that firms in oil and gas industries after the BP oil spill in 2010 experienced a negative stock price reaction (proxied by cumulative abnormal returns). Muller and Kräussl (2011) find that majority of US firms experience negative abnormal stock returns after Hurricane Katrina in 2005. The negative impact is stronger for firms with low CSR performance (irresponsible firms) before the hurricane. To capture the market reaction on the treated firms, we estimate the following model:

$$\begin{aligned} REACTION_t = {} & \beta_0 + \beta_1 aftertreat\_sic2\_\sharp + \beta_2 SIZE_{t-1} + \beta_3 MB_{t-1} + \beta_4 LEV_{t-1} \\ & + \beta_5 CH_{t-1} + \beta_6 COMBINED\_RAM_{t-1} + \beta_7 ABS\_DA_{t-1} \\ & + \beta_8 RD\_INT_{t-1} + \beta_9 AD\_IND\_INT_{t-1} + \beta_{10} BIG4_{t-1} \\ & + \beta_{11} FIRM\_AGE_{t-1} + \epsilon_t, \end{aligned} \tag{1}$$

where, $REACTION$ is ROA, ROE or sales growth rate (SALE_G); $aftertreat\_sic2\_\sharp$ ($aftertreat\_sic2\_0$ or $aftertreat\_sic2$) is a dummy variable that equals 1 for treated industries in the year of a disaster (years after the disasters (including the years of the disasters)). The rest of the variables are as described in Appendix A. We use firm and year fixed effects and cluster standard errors at the two-digit SIC level[22][23]. The coefficient of interest is β_1, which measures the

[22] Here and after, following Bertrand et al. (2004) we cluster standard errors at the dimension of the treatment (two-digit SIC level).

[23] Variable $aftertreat\_sic2\_\sharp$ equals interaction between time dummy ($time$) and industry dummy ($treated$). Dummy $time$ and time fixed effect are collinear. Dummy $treated$ and firm fixed effect are collinear. For that reason,

difference in market reaction between treated and control firms after the technological disasters. If our tests support the hypothesis that treated firms suffer from the negative consequences of the disasters, the coefficient β_1 is expected to be negative.

Next, we measure how firms adjust their CSR after the disasters take place. We estimate the following simple regression:

$$CSR = \beta_0 + \beta_1 aftertreat\_sic2 + \beta_2 SIZE_{t-1} + \beta_3 ROA_{t-1} + \beta_4 MB_{t-1} + \beta_5 LEV_{t-1} + \beta_6 CH_{t-1} + \epsilon_t, \quad (2)$$

where, $aftertreat\_sic2$ is a dummy variable that equals 1 for treated industries in the years after the disasters (including the years of the disasters); CSR is alternatively one of the following proxies: CSR_SCORE, EXTERNAL, or INTERNAL in t, $t+1$, $t+2$, and $t+3$[24]. These are as previously defined. The rest of the variables are as described in Appendix A. The set of control variables is consistent with the previous research (Flammer, 2015). We use year and firm fixed effects and cluster standard errors at the two-digit SIC level. The coefficient of interest is β_1, which measures the difference in CSR performance between treated and control firms after the technological disaster. If our tests support the hypothesis that firms boost CSR performance after the catastrophes, the coefficient β_1 is expected to be positive.

we do not include $time$ and $treated$ dummies in the regressions. Year and firm fixed effects help us to control for omitted time- and firm-specific variables. Thus, as a main specification we use regression with year and firms fixed effects and without $year$ and $treated$ dummies. We repeat all tests with $time$ and $treated$ dummies and without year and firm fixed effects (i.e. the following specification: $Dependent\_variable = \beta_0 + \beta_1 aftertreat\_sic2\_\sharp + \beta_2 time + \beta_3 treated + \Sigma CONTROLS + \epsilon_t$). All results are qualitatively and quantitatively the same. For brevity, we do not tabulate them.

[24] We awareness that a larger time window increases the risk that CSR reactions may be contaminated by information that is unrelated to the CSR-disasters.

3.4.2 Instrumental variable (IV) Regressions

To study the effect of enforced CSR - instrumented by the CSR-disasters - on FRQ, we use two-stage least squares (2SLS). In the first stage, we regress EXTERNAL[25] on the CSR-disasters. Specifically, we estimate the following model:

$$\begin{aligned} EXTERNAL_{t+1} = \beta_0 + \beta_1 aftertreat\_sic2 + \beta_2 SIZE_{t-1} + \beta_3 ROA_{t-1} + \beta_4 MB_{t-1} \\ + \beta_5 LEV_{t-1} + \beta_6 CH_{t-1} + \beta_7 COMBINED\_RAM_{t-1} \\ + \beta_8 RD\_INT_{t-1} + \beta_9 AD\_IND\_INT_{t-1} + \beta_{10} GOVERNANCE_{t-1} \\ + \beta_{11} BIG4_{t-1} + \beta_{12} FIRM\_AGE_{t-1} + \epsilon_t, \end{aligned} \tag{3}$$

where, $aftertreat\_sic2$ is as previously defined and all other variables are as described in Appendix A. We use firm and year fixed effects and cluster standard errors at the two-digit SIC level. The predicted values from Equation 3 provide the "instrumented" EXTERNAL (f_EXTERNAL_w_hat) - i.e., the exogenous component of the EXTERNAL. In the second stage, we estimate the following equation using f_EXTERNAL_w_hat in lieu of EXTERNAL :

$$\begin{aligned} EM = \beta_0 + \beta_1 f\_EXTERNAL\_w\_hat_t + \beta_2 SIZE_{t-1} + \beta_3 ROA_{t-1} + \beta_4 MB_{t-1} \\ + \beta_5 LEV_{t-1} + \beta_6 CH_{t-1} + \beta_7 COMBINED\_RAM_{t-1} + \beta_8 RD\_INT_{t-1} \\ + \beta_9 AD\_IND\_INT_{t-1} + \beta_{10} GOVERNANCE_{t-1} + \beta_{11} BIG4_{t-1} \\ + \beta_{12} FIRM\_AGE_{t-1} + \epsilon_t, \end{aligned} \tag{4}$$

where, EM is a proxy for absolute value of discretionary accruals (ABS_DA) in $t+2$, $t+3$, $t+4$, and $t+5$. The coefficient of interest is β_1, which provides an estimate of the effect of CSR

[25] As it is shown in Section 4, on average, firms in our sample improve their CSR performance through EXTERNAL component of CSR_SCORE. Thus, using EXTERNAL component (instead of CSR_SCORE) in the first stage of IV (equation 3) helps us to better satisfy the inclusion restriction (the treatment (i.e., CSR-disasters) need to trigger relevant changes in CSR perfomance).

on FRQ[26].

4 Results

4.1 Descriptive Statistics

In Table 1, we present the sample distribution by the two-digit SIC code industry. The most heavily represented industry is Business Services (SIC code 73, 13.6%), followed by Chemical and Allied Products (SIC code 28, 11.39%), and Electronic and Other Electronic Equipment (SIC code 36, 8.74%). Industry distribution in the sample is consistent with prior research (Kim et al., 2012).

Table 2 reports descriptive statistics for selected variables. All variables are defined in Appendix A. On average, firms in the sample are socially irresponsible and have CSR_SCORE less than 0 (CSR_SCORE mean is -0.17)[27]. In other words, the average firm in our sample has more concerns than strengths, consistent with Kim et al. (2012) and Lins et al. (2017). By construction, the means of earnings management proxies are 0.

The mean value of ADJ_ROA is 0.03, indicating that, on average, our sample firms are more profitable than their industry peers. 90% of the firms is audited by the Big 4 accounting firms. On average, firms' R&D (advertising) expenditures are 17% (1%) of their net sales. FIRM_AGE 2.63 means, that the average age of the firms in our sample is 13 years.

Table 3 presents Pearson correlations. CSR_SCORE has negative correlation with the absolute value of discretionary accruals and abnormal production costs. There is a positive correlation between CSR proxy and abnormal cash flow from operations (AB_CFO), abnormal discretionary expenses (AB_EXP), and COMBINED_RAM. Overall, our descriptive statistics and correlations

[26] Equations for the first and second stages estimations for the other dependent variables are presented in Appendix D.

[27] To have comparable with prior research summary statistics (e.g. Kim et al. (2012)) in Table (2) we tabulate results for unscaled CSR_SCORE. Scaled CSR_SCORE (that is used in all tabulated regressions) has the following descriptive statistics: obs = 16753; mean = -0.206; std.dev. = 0.542; p25 = -0.55 p50 = -0.25 p75 = 0.

are consistent with the prior research (Kim et al., 2012).

4.2 Market Reaction and the Relation between CSR and FRQ

Table 4 presents the mean $CARs_{it}$ and the associated p-values for the hypothesis that the mean is less than zero. The means for the three-day $CARs$ are negative and significant at the one percent level. The means for the five-day $CARs$ are negative and significant at the one percent level for EVENTs I and II. The fifteen-day mean is positive and insignificant for EVENTs I and III and suggests a reversal over time. Further, both in the short and long run, treated firms experience decrease in ROA, ROE, and sales growth (Table 5). The effect is strong for the long run. These results supports the idea that firms in the treated industries experience a negative spillover effect.

To provide perspective on the effect of CSR-disasters on CSR, Figure 1 plots the evolution of average EXTERNAL in treated (blue solid line) and control (red dashed line) groups before and after the treatment. This figure provides three insights. First, EXTERNAL is trending upward in both treatment and control groups. This trend is consistent with observations in the previous studies (Flammer, 2013, 2015) and emphasizes the importance of using a control group - not accounting for changes in KLD index at the control group would overstate the effect of CSR-disasters on CSR performance (proxied by KLD index), as it would capture some of the time trend[28]. Second, before the treatment the difference in the EXTERNAL in treatment and control groups is relatively constant. Third, following the CSR-disasters, the two lines diverge: firms in the treated industries increase their EXTERNAL substantially more compared to the control group. Further, Figure 1 provides evidence that two years after the CSR-disasters treated firms steadily increase their EXTERNAL. However, due to data availability we cannot follow the long-term dynamics of the treatment effect.

Table 6 presents the main results (equation 2). In Columns 1-3 (4-6, 7-9, and 10-12) the dependent variable is in the year of the CSR-disasters (one year ($f\_\sharp$), two years ($f2\_\sharp$), and

[28] Partially, we solve this potential source of error by scaling our CSR proxy (CSR_SCORE, EXTERNAL, or INTERNAL). Please find additional discussion of KLD trends and composition in Appendix C.

three years ($f3\_\sharp$) after the CSR-disasters). The dependent variable in columns 1, 4, 7, and 10 is CSR_SCORE, while in columns 2, 5, 8, and 10 (3, 6, 9, and 12), it is EXTERNAL (INTERNAL). In all models, we include industry fixed effect (two-digit SIC level) because some industries may be more likely to invest in CSR than others and may have been deferentially affected by the CSR-disasters. We also control for time-varying omitted variables by including year fixed effect. In Table 7 we re-estimate equation 2, but instead of including all observations, we divide treated firms into ones with high and low CSR performance before the CSR-disasters[29].

For each specification, Tables 6 and 7 report the coefficient on CSR-disasters dummy ($aftertreat\_sic2$) and its standard error in parentheses. As can be seen, the coefficient on CSR-disasters is positive and statistically significant for specifications with EXTERNAL as dependent variable. Results in Table 7 show that improvement in CSR performance is mainly driven by firms with low pre-disaster CSR performance.

Following Lins et al. (2017) we examine whether it is a firm's CSR performance in aggregate (CSR_SCORE) or a specific component of CSR that is important in the post-disasters period. In Tables 6 and 7, we separately estimate the change of EXTERNAL and INTERNAL components of CSR_SCORE after the CSR-disasters, and find that the improvement in CSR_SCORE is entirely driven by EXTERNAL component (Environment, Community, and Human Rights). Further, the results suggest that firms consistently improve their CSR performance in three years after the CSR-disasters. This result is consistent with the idea, that CSR variables are very sticky and it takes time to significantly improve CSR performance. Overall, the evidence in Tables 6 and 7 support the hypothesis that treated firms respond to the CSR-disasters by increasing CSR performance mainly through EXTERNAL component.

Next, we examine whether improvement in CSR performance leads to change in FRQ. Tables 8 and 9 present the results of IV regression analyses of discretionary accruals and the quality of

[29]We define treated firms as with high (low) CSR performance if there EXTERNAL performance is greater (lower or equal) than the medium withing industry-year one year before the CSR-disasters. Results are qualitatively and quantitatively similar if we define treat firms as with high/low CSR according to pre-disaster CSR_SCORE (instead of EXTERNAL).

narrative disclosure. We find a negative relation between CSR and FRQ. Specifically, in Table 8 the estimated coefficient on predicted value of EXTERNAL (f_EXTERNAL_w_hat) is positive and significant ($p < 0.01$), indicating that treated firms manage earnings more through accruals[30]. We observe similar results from the regressions of narrative disclosure (Table 8, columns 7 and 8). Together evidence suggest that enforced CSR leads to deterioration in FRQ. In turn, this result means that firms that are enforced to increase CSR performance do it mainly through "window dressing" CSR.

5 Robustness Checks

Table 10 and Figure 2 support robustness of the main results of the paper. Following prior studies (Atanasov and Black, 2016; Flammer and Kacperczyk, 2015) we construct leads and lags model. Table 10 presents treatment dynamics of the CSR-disasters on the change in CSR performance (proxied by EXTERNAL component). The results show that treatment is not anticipated by the firms. However, our results reinforces the presumption that treated and control firms have different CSR performance before the CSR-disasters (coefficients before the treatment are negative and significant)[31].

To further enhance the credibility of our results, we next conduct a placebo test. For each year of the events we randomly assign treated industry. Then we estimate Equation 2. We repeat this exercise 1000 times and plot the discretized probability density of the placebo coefficients in Figure 2. The graph shows that the placebo coefficient largely follows a normal distribution centered at zero (mean = -0.116).

[30] In untabulated tests we do not find evidence that improvement in CSR performance leads to change in real activities manipulation (AB_CFO, AB_PROD, AB_EXP, and COMBINED_RAM)

[31] Ancillary results show that using CSR_SCORE as a dependent variable (instead of EXTERNAL) does not change the conclusion that treatment is not anticipated by the firms. In this specification coefficients before the treatment are insignificant and close to zero. For brevity, we do not tabulate this test.

6 Additional Analyses

6.1 Investment in CSR

The arguments provided so far indicate that, when faced with CSR-disasters, treated firms increase their CSR to improve their reputation and differentiate themselves from the "guilty" firm. Further we show that this improvement in CSR has a negative impact on FRQ. Thus, as we discuss previously, we conclude that CSR-disasters induce window dressing CSR that is not associated with real changes in corporate ethical standards. However, Owens et al. (2016) show that idiosyncratic shocks can misspecify discretionary accruals models. For instance, because of investing in CSR projects, a firm could have a cash flow level that is lower than the average in the industry (i.e. this firm would have negative AB_CFO, which is a proxy for real earnings management). Also, this firm may have higher Bog Index because with initiated new projects to increase CSR, managers may have to use more technical terminology (i.e. difficult words that increase Bog Index). Overall, lower FRQ (that is proxied by discretionary accruals models and Bog Index) potentially could be driven by "true" CSR.

To alleviate this concern, we investigate whether increase in CSR after the CSR-disasters is associated with financial investments in CSR. Following Di Giuli and Kostovetsky (2014), we study whether improvement in CSR is associated with higher levels of Selling, General, and Administrative expenses (SG&A). Evidence in Table 11 suggests that ,on average, improvement in CSR is not associated with increase in SG&A expenses. Only firms that have high CSR before the CSR-disasters improve their CSR through investing in SG&A. The conclusion from this result is twofold. First, consistent with the main finding of this paper, CSR-disasters induce window-dressing CSR that does not entail positive outcomes, such as improvement in FRQ. Second, this result supports the idea that voluntary CSR (those firms that have high CSR before the CSR-disasters) have real CSR (that is associated with investments in CSR), and this CSR performance is associated with high ethical corporate standards (Kim et al., 2012).

6.2 Risk Management VS Greenness

The above findings provide evidence of a negative relation between CSR and FRQ in the post-disaster period, when firms are forced to improve their CSR performance. Further, we find that improvement in CSR performance is not associated with investment in CSR. Thus, we conclude that improvement in CSR comes mainly from window dressing practices. To provide more granularity in our results, we repeat our analyses separately for STRENGTHS and CONCERNS dimensions of CSR_SCORE.

Several recent studies have emphasized the importance of distinguishing between STRENGTHS and CONCERNS (e.g. Kacperczyk (2009); Kim et al. (2014); Ioannou and Serafeim (2015)). From this perspective, "doing good" (STRENGTHS) and "doing no harm" (CONCERNS) are fundamentally different and reflect different underlying mechanisms. Thus, subtracting CONCERNS from STRENGTHS would not be reasonable, as these variables are not perfectly substitutable. Further, Fernando et al. (2017) argue that it is only decreasing CONCERNS that is associated with firms' risk and financial cost reduction. On the other hand, increasing STRENGTHS cannot be explained by any risk management actions[32]. In our settings, it means that if treated firms aim to improve their CSR performance to decrease the likelihood of the disaster repetition (i.e. "real" CSR) they would do it through decreasing CONCERNS. In turn, if treated firms had reputational purpose (i.e. window dressing CSR), they would increase CSR via increasing STRENGTHS.

Table 12 reports the results of estimating equation 2 separately for STRENGTHS and CONCERNS as dependent variables. The results suggest that firms improve their CSR performance through increasing STRENGTHS. Coefficient for CONCERNS is negative but insignificant. These results provide additional evidence that treated firms improve their CSR performance

[32]More precisely, Fernando et al. (2017) argue that ENVIRONMENTAL strength and concerns have different impact on firms' environmental risk exposure and the likelihood of potential losses. We repeat our analysis with only ENVIRONMENTAL strength and concerns as dependent variables (instead of composite STRENGTH and CONCERNS. See variable definitions in Appendix A). Our results are qualitatively and quantitatively similar those that we tabulate for STRENGTHS and CONCERNS below.

through window dressing activities that are not related to risk-management practices and do not reduce possible future losses associated with the disasters.

6.3 Propensity Score Matching

Next, we conduct propensity score matching (PSM) aimed at better alleviating the concern that firms in treated and control groups are different.

To construct a sample of firms that are similar to the treated firms, we match each treated firm to a control firm on the basis of firm-level characteristics using the following procedure. The nearest neighbor is calculated based on six firm-level characteristics: CSR performance (CSR_SCORE), size ($SIZE$), market-to-book (MB), leverage ratio (LEV), return on asset (ROA), and cash holdings (CH). Matching variables are computed as average in the three years preceding the disasters. This matching procedure is used in prior literature to construct a set of comparable firms (e.g.Flammer (2015)).

We repeat all main test with the PSM sample[33]. All results remain qualitative and quantitatively similar to those shown in the main part of the paper.

7 Conclusion

This paper extends our understanding of how firms respond to CSR-disasters. Using a set of technological disasters as a proxy for CSR-disasters, we study the reaction of firms that operate in the involved industries. We argue that there are at least two reasons why affected firms would improve their CSR in the post-disaster periods. First, firms can decrease their operational risk (that is associated with the likelihood of CSR-disaster repetition) by investing in CSR ("true" CSR). Second, firms can build positive image and signal their social and environmental awareness through CSR (window dressing CSR). In both scenarios, CSR would help to improve the relationship with stakeholders. In turn, these firms would be less penalized by their stakeholders

[33] Please find results of PSM in Appendix E

in the post-disaster period.

Using differences-in-difference approach, we show that firms in the affected industries improve their CSR performance in the post-disasters periods. This result is mainly driven by the firms with low pre-disaster CSR performance. Then, using instrumental variables (IV) approach, we show that increase in CSR performance leads to lower FRQ proxied by earnings management (absolute value of discretionary accruals) and narrative disclosures (Bog Index). We do not find evidence that improvement in CSR performance has an impact on the level of real activities manipulation (RAM).

Taken together, our evidence suggests that in some particular cases, when managers are forced to improve firms' CSR performance, they do so by implementing window dressing practices that are not associated with real changes in firms' ethical or safety standards. Although firms with low CSR performance before the CSR-disasters significantly increase their CSR rating, we do not find evidence that these firms in fact do increase their CSR expenditures (proxied by SG&A expenses). The only exception are firms, with high CSR performance before the CSR-disasters. These firms improve their CSR rating through real investments in CSR. Our results reinforce the awareness that some CSR practices may lead to unexpected unfavorable outcomes, in particular deterioration in FRQ.

References

Atanasov, V. and B. Black (2016). Shock-based causal inference in corporate finance and accounting research. *Critical Finance Review 5*(2), 207–304.

Atkins, B. (2006). Is corporate social responsibility responsible. *Corp. Govern. Advert 14*, 28–29.

Barnett, M. L. and R. M. Salomon (2012). Does it pay to be really good? addressing the shape of the relationship between social and financial performance. *Strategic Management Journal 33*(11), 1304–1320.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *The Quarterly journal of economics 119*(1), 249–275.

Blacconiere, W. G. and D. M. Patten (1994). Environmental disclosures, regulatory costs, and changes in firm value. *Journal of accounting and economics 18*(3), 357–377.

Bonsall, S. B., A. J. Leone, B. P. Miller, and K. Rennekamp (2017). A plain english measure of financial reporting readability. *Journal of Accounting and Economics 63*(2-3), 329–357.

Cheng, B., I. Ioannou, and G. Serafeim (2014). Corporate social responsibility and access to finance. *Strategic Management Journal 35*(1), 1–23.

Clarkson, P. M., Y. Li, and G. D. Richardson (2004). The market valuation of environmental capital expenditures by pulp and paper companies. *The accounting review 79*(2), 329–353.

DeFond, M. L. and K. Subramanyam (1998). Auditor changes and discretionary accruals. *Journal of accounting and Economics 25*(1), 35–67.

Desai, V. M. (2011). Mass media and massive failures: Determining organizational efforts to defend field legitimacy following crises. *Academy of Management Journal 54*(2), 263–278.

Di Giuli, A. and L. Kostovetsky (2014). Are red or blue companies more likely to go green? politics and corporate social responsibility. *Journal of Financial Economics 111*(1), 158–180.

Diestre, L. and N. Rajagopalan (2014). Toward an input-based perspective on categorization: Investor reactions to chemical accidents. *Academy of Management Journal 57*(4), 1130–1153.

Entine, J. (2003). The myth of social investing: A critique of its practice and consequences for corporate social performance research. *Organization & Environment 16*(3), 352–368.

Fernando, C. S., M. P. Sharfman, and V. B. Uysal (2017). Corporate environmental policy and shareholder value: Following the smart money. *Journal of Financial and Quantitative Analysis 52*(5), 2023–2051.

Ferrell, A., H. Liang, and L. Renneboog (2016). Socially responsible firms. *Journal of Financial Economics 122*(3), 585–606.

Flammer, C. (2013). Corporate social responsibility and shareholder reaction: The environmental awareness of investors. *Academy of Management Journal 56*(3), 758–781.

Flammer, C. (2015). Does product market competition foster corporate social responsibility? evidence from trade liberalization. *Strategic Management Journal 36*(10), 1469–1485.

Flammer, C. and A. Kacperczyk (2015). The impact of stakeholder orientation on innovation: Evidence from a natural experiment. *Management Science 62*(7), 1982–2001.

Flammer, C. and J. Luo (2017). Corporate social responsibility as an employee governance tool: Evidence from a quasi-experiment. *Strategic Management Journal 38*(2), 163–183.

Godfrey, P. C. (2005). The relationship between corporate philanthropy and shareholder wealth: A risk management perspective. *Academy of management review 30*(4), 777–798.

Godfrey, P. C., C. B. Merrill, and J. M. Hansen (2009). The relationship between corporate social responsibility and shareholder value: An empirical test of the risk management hypothesis. *Strategic management journal 30*(4), 425–445.

Hart, O. and L. Zingales (2017). Companies should maximize shareholder welfare not market value. *Journal of Law, Finance, and Accounting 2*(2), 247–275.

Healy, P. M. and J. M. Wahlen (1999). A review of the earnings management literature and its implications for standard setting. *Accounting horizons 13*(4), 365–383.

Heflin, F. and D. Wallace (2017). The bp oil spill: shareholder wealth effects and environmental disclosures. *Journal of Business Finance & Accounting 44*(3-4), 337–374.

Huang, X., S. H. Teoh, and Y. Zhang (2014). Tone management. *The Accounting Review 89*(3), 1083–1113.

Ioannou, I. and G. Serafeim (2015). The impact of corporate social responsibility on investment recommendations: Analysts' perceptions and shifting institutional logics. *Strategic Management Journal 36*(7), 1053–1081.

Jones, J. J. (1991). Earnings management during import relief investigations. *Journal of accounting research*, 193–228.

Kacperczyk, A. (2009). With greater power comes greater responsibility? takeover protection and corporate attention to stakeholders. *Strategic Management Journal 30*(3), 261–285.

Kim, Y., H. Li, and S. Li (2014). Corporate social responsibility and stock price crash risk. *Journal of Banking & Finance 43*, 1–13.

Kim, Y., M. S. Park, and B. Wier (2012). Is earnings quality associated with corporate social responsibility? *The Accounting Review 87*(3), 761–796.

Klein, A. (2002). Audit committee, board of director characteristics, and earnings management. *Journal of accounting and economics 33*(3), 375–400.

Kothari, S. P., A. J. Leone, and C. E. Wasley (2005). Performance matched discretionary accrual measures. *Journal of accounting and economics 39*(1), 163–197.

Lev, B., C. Petrovits, and S. Radhakrishnan (2010). Is doing good good for you? how corporate charitable contributions enhance revenue growth. *Strategic Management Journal 31*(2), 182–200.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and economics 45*(2-3), 221–247.

Liang, H. and L. Renneboog (2017). On the foundations of corporate social responsibility. *The Journal of Finance 72*(2), 853–910.

Lins, K. V., H. Servaes, and A. Tamayo (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance 72*(4), 1785–1824.

Loughran, T. and B. McDonald (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance 66*(1), 35–65.

Madsen, P. M. and Z. J. Rodgers (2015). Looking good by doing good: The antecedents and consequences of stakeholder attention to corporate disaster relief. *Strategic Management Journal 36*(5), 776–794.

Marano, V. and T. Kostova (2016). Unpacking the institutional complexity in adoption of csr practices in multinational enterprises. *Journal of Management Studies 53*(1), 28–54.

McWilliams, A. and D. Siegel (2000). Corporate social responsibility: A theory of the firm perspective. *Academy of management review 26*(1), 117–127.

Muller, A. and R. Kräussl (2011). Doing good deeds in times of need: A strategic perspective on corporate disaster donations. *Strategic Management Journal 32*(9), 911–929.

Owens, E. L., J. S. Wu, and J. Zimmerman (2016). Idiosyncratic shocks to firm underlying economics and abnormal accruals. *The Accounting Review 92*(2), 183–219.

Pek, S., C. H. Oh, and J. Rivera (2018). Mnc foreign investment and industrial disasters: The moderating role of technological, safety management, and philanthropic capabilities. *Strategic Management Journal 39*(2), 502–526.

Petrenko, O. V., F. Aime, J. Ridge, and A. Hill (2016). Corporate social responsibility or ceo narcissism? csr motivations and organizational performance. *Strategic Management Journal 37*(2), 262–279.

Petrovits, C. M. (2006). Corporate-sponsored foundations and earnings management. *Journal of Accounting and Economics 41*(3), 335–362.

Prior, D., J. Surroca, and J. A. Tribó (2008). Are socially responsible managers really ethical? exploring the relationship between earnings management and corporate social responsibility. *Corporate Governance: An International Review 16*(3), 160–177.

Roychowdhury, S. (2006). Earnings management through real activities manipulation. *Journal of accounting and economics 42*(3), 335–370.

Shiu, Y.-M. and S.-L. Yang (2017). Does engagement in corporate social responsibility provide strategic insurance-like effects? *Strategic Management Journal 38*(2), 455–470.

Subramanyam, K. (1996). The pricing of discretionary accruals. *Journal of accounting and economics 22*(1-3), 249–281.

Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative science quarterly*, 35–67.

Warfield, T. D., J. J. Wild, and K. L. Wild (1995). Managerial ownership, accounting choices, and informativeness of earnings. *Journal of accounting and economics 20*(1), 61–91.

Table 1: Sample Description: Distribution of Firm-Year Observations by Industry

Two-Digit SIC	# of Obs.	% of Sample	Cumulative Percent
10	72	0.43	0.43
13	717	4.28	4.71
15	145	0.87	5.58
16	69	0.41	5.99
20	459	2.74	8.73
23	186	1.11	9.84
24	83	0.5	10.33
25	116	0.69	11.02
26	206	1.23	12.25
27	202	1.21	13.46
28	1,909	11.39	24.86
29	91	0.54	25.4
30	165	0.98	26.38
31	10	0.06	26.44
32	95	0.57	27.01
33	275	1.64	28.65
34	246	1.47	30.12
35	1,194	7.13	37.25
36	1,464	8.74	45.99
37	548	3.27	49.26
38	1,209	7.22	56.47
39	163	0.97	57.45
42	167	1	58.44
44	11	0.07	58.51
45	137	0.82	59.33
48	595	3.55	62.88
49	882	5.26	68.14
50	388	2.32	70.46
51	213	1.27	71.73
53	165	0.98	72.72

Continued on next page

Table 1 *Continued from previous page*

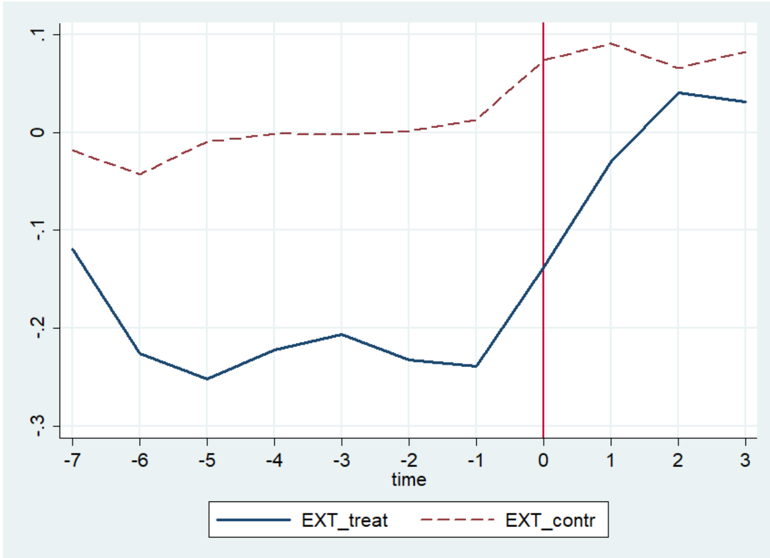
Two-Digit SIC	# of Obs.	% of Sample	Cumulative Percent
54	112	0.67	73.38
55	166	0.99	74.37
56	296	1.77	76.14
57	22	0.13	76.27
58	299	1.78	78.06
59	345	2.06	80.12
72	10	0.06	80.18
73	2,278	13.6	93.77
78	21	0.13	93.9
79	209	1.25	95.15
80	361	2.15	97.3
82	55	0.33	97.63
87	387	2.31	99.94
99	10	0.06	100
Total	16 753	100	

Table 2: Descriptive Statistics of Selected Variables

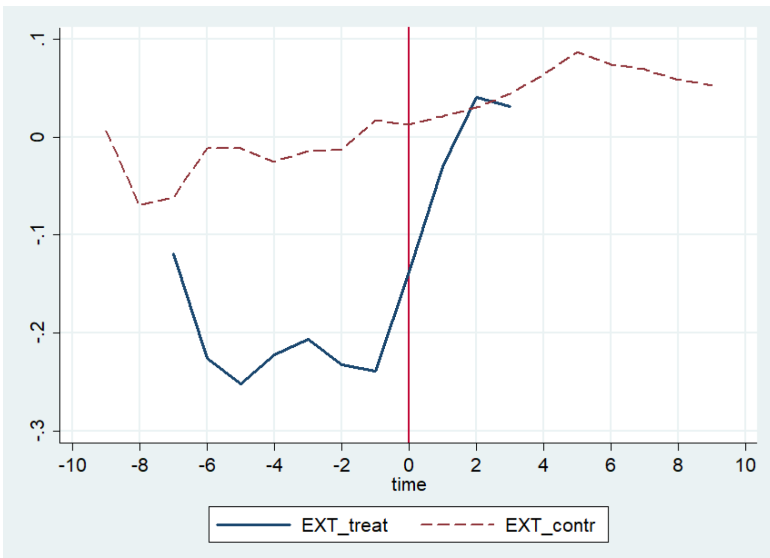
Variable	Obs	Mean	Std. Dev.	P25	P50	P75
CSR_SCORE	16753	-.166	2.211	-2	0	1
ABS_DA	16572	.054	.063	.015	.035	.068
AB_CFO	16753	0	.099	-.046	-.001	.048
AB_PROD	16753	-.001	.176	-.089	.001	.086
AB_EXP	16753	-.002	.189	-.096	0	.062
COMBINED_RAM	16753	-.001	.361	-.194	-.021	.172
SIZE	16735	7.121	1.525	5.998	6.913	8.039
MB	16733	3.142	4.285	1.471	2.276	3.725
ADJ_ROA	16753	.029	.155	.009	.05	.096
LEV	16683	.207	.202	.009	.175	.325
RD_INT	16665	.167	.733	0	.005	.082
AD_IND_INT	16665	.011	.026	0	0	.009
CH	16753	.206	.216	.041	.125	.3
ROA	16753	.021	.142	.009	.045	.085
ROE	16751	.071	.438	.018	.109	.195
SALE_G	16641	.144	.324	.004	.09	.21
BIG4	16753	.897	.303	1	1	1
FIRM_AGE	16753	2.631	.975	2.079	2.708	3.332
SGA	14996	5.299	1.457	4.255	5.14	6.189
SGA_AD	14990	5.248	1.445	4.216	5.095	6.147
SGA_R	14921	.297	.251	.128	.236	.387

Table 3: Correlations among CSR, Earnings Management Proxies, and Other Selected Variables

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1)CSR_SCORE	1.000											
(2)ABS_DA	-0.070	1.000										
	(0.000)											
(3)AB_CFO	0.093	-0.059	1.000									
	(0.000)	(0.000)										
(4)AB_PROD	-0.123	0.024	-0.464	1.000								
	(0.000)	(0.002)	(0.000)									
(5)AB_EXP	0.098	0.098	-0.045	-0.580	1.000							
	(0.000)	(0.000)	(0.000)	(0.000)								
(6)COMBINED_RAM	0.135	0.022	0.481	-0.915	0.786	1.000						
	(0.000)	(0.004)	(0.000)	(0.000)	(0.000)							
(7)SIZE	0.379	-0.200	0.174	-0.125	0.075	0.147	1.000					
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)						
(8)MB	0.070	0.085	0.077	-0.115	0.142	0.149	0.133	1.000				
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
(9)LEV	-0.027	-0.070	-0.110	0.062	-0.053	-0.089	0.108	-0.067	1.000			
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)				
(10)CH	-0.000	0.275	0.012	-0.074	0.182	0.134	-0.210	0.191	-0.358	1.000		
	(0.961)	(0.000)	(0.101)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)			
(11)ROA	0.093	-0.333	0.457	-0.233	-0.092	0.196	0.309	-0.006	-0.105	-0.318	1.000	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.442)	(0.000)	(0.000)		
(12)FIRM_AGE	0.164	-0.193	0.034	-0.001	-0.076	-0.028	0.315	-0.090	0.048	-0.266	0.152	1.000
	(0.000)	(0.000)	(0.000)	(0.939)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	



(a) 2010 is assigned as treatment year for the control group



(b) For each control industry treatment year is randompl assigned between 2004 and 2012

Figure 1: Evolution of average EXTERNAL in treatment and control group. Treatment group (EXT_treat) - industries with sic2 = 13, 29, and 49. Control group (EXT_contr)- all industries except sic2 = 13, 29, and 49. *time* - year relative to treatment.

Table 4: Market Reaction (CAR)

Event	Event Window	Estimation Window	CAR	p-stat (Ha: mean < 0)
EVENT I	[-1;1]	[-10;-100]	-0.00509	Pr(T< t) = 0.0000
EVENT I	[-2;2]	[-10;-100]	0.001656	Pr(T < t) = 0.8944
EVENT I	[-1;1]	(-1;-126]	-0.0033582	Pr(T < t) = 0.0111
EVENT I	[-7;7]	(-1;-126]	0.0121059	Pr(T < t) = 1.0000
EVENT II	[-1;1]	[-10;-100]	-0.00499	Pr(T < t) = 0.0001
EVENT II	[-2;2]	[-10;-100]	-0.00719	Pr(T < t) = 0.0000
EVENT II	[-1;1]	(-1;-126]	-0.0035608	Pr(T < t) = 0.0015
EVENT II	[-7;7]	(-1;-126]	-0.0378907	Pr(T < t) = 0.0000
EVENT III	[-1;1]	[-10;-100]	-0.00808	Pr(T < t) = 0.0000
EVENT III	[-2;2]	[-10;-100]	-0.01037	Pr(T < t) = 0.0000
EVENT III	[-1;1]	(-1;-126]	-0.0084889	Pr(T < t) = 0.0000
EVENT III	[-7;7]	(-1;-126]	0.0059551	Pr(T < t) = 1.0000
EVENTS I-III	[-1;1]	[-10;-100]	-0.00648	Pr(T < t) = 0.0000
EVENTS I-III	[-2;2]	[-10;-100]	-0.0045209	Pr(T < t) = 0.0000
EVENTS I-III	[-7;7]	[-10;-100]	-0.0103254	Pr(T < t) = 0.0000
EVENTS I-III	[-1;1]	(-1;-126]	-.0052385	Pr(T < t) = 0.0000

This table shows abnormal returns cumulated across either 3, 5, or 15 days around the Events. Abnormal returns are the prediction errors from the market model. Market model is the following: $R_{it} = \alpha_i + \beta_i R_{mt} + \epsilon_{it}$, where R_{it}-return for firm i on day t, α_i - intercept for firm i, R_{mt} - return on the CRSP equal-weighted market portfolio on day t, and ϵ_{it} - error term with mean zero. We calculate daily abnormal returns (AR_{it}) as follows: $AR_{it} = R_{it} - (\hat{\alpha}_i + \hat{\beta}_i R_{mt})$. We calculate the cumulative abnormal returns ($CAR_{i,T}$) for each firm (i) and event window (T) by summing across the T days in an event window. EVENT I - BP oil spill on April 20, 2010. EVENT II - San Bruno Gas Pipeline Explosion on September 9, 2010. EVENT III - Fukushima Daiichi Nuclear Disaster on March 11, 2011.

Table 5: Long-run performance after the accidents

VARIABLES	(1) ROA	(2) ROE	(3) SALE_G	(4) ROA	(5) ROE	(6) SALE_G
aftertreat_sic2_0	-0.00830	-0.0491**	-0.0349			
	(0.00693)	(0.0209)	(0.0339)			
aftertreat_sic2				-0.0146***	-0.0668***	-0.0629***
				(0.00481)	(0.0171)	(0.0154)
l_SIZE_w	0.0250***	-0.00545	0.0396**	0.0253***	-0.00412	0.0409**
	(0.00324)	(0.0207)	(0.0153)	(0.00319)	(0.0206)	(0.0153)
l_MB_w	0.00191***	0.0238*	0.00378**	0.00190***	0.0238*	0.00375**
	(0.000428)	(0.0141)	(0.00152)	(0.000427)	(0.0141)	(0.00152)
l_LEV_w	0.0205	0.0304	0.0677	0.0206	0.0304	0.0678
	(0.0191)	(0.0949)	(0.0641)	(0.0193)	(0.0957)	(0.0642)
l_CH_w	0.0540**	-0.00891	0.0661	0.0537**	-0.00965	0.0651
	(0.0262)	(0.0791)	(0.0440)	(0.0263)	(0.0789)	(0.0450)
l_COMBINED_RAM	0.0397***	0.0544	-0.0189	0.0396***	0.0541	-0.0192
	(0.00575)	(0.0345)	(0.0165)	(0.00571)	(0.0343)	(0.0166)
l_ABS_DA	0.0395**	0.193	-0.0156	0.0399**	0.194	-0.0147
	(0.0170)	(0.175)	(0.0784)	(0.0169)	(0.176)	(0.0789)
l_RD_INT_w	-0.0262***	-0.0265***	0.202***	-0.0261***	-0.0263***	0.203***
	(0.00407)	(0.00845)	(0.00657)	(0.00407)	(0.00844)	(0.00656)
l_AD_IND_INT_w	-0.286*	1.184*	1.412	-0.284*	1.189*	1.418
	(0.154)	(0.616)	(0.904)	(0.155)	(0.615)	(0.901)
l_BIG4	-0.0143*	0.0311	0.00961	-0.0140*	0.0325	0.0109
	(0.00754)	(0.0351)	(0.0186)	(0.00756)	(0.0354)	(0.0179)
l_FIRM_AGE_w	0.00124	0.0208	-0.0727***	0.000454	0.0175	-0.0761***
	(0.00767)	(0.0180)	(0.0211)	(0.00770)	(0.0188)	(0.0208)
Constant	-0.152***	-0.0734	-0.00779	-0.153***	-0.0767	-0.0109
	(0.0185)	(0.100)	(0.139)	(0.0186)	(0.0995)	(0.135)
Observations	13,770	13,769	13,770	13,770	13,769	13,770
Adjusted R-squared	0.584	0.271	0.311	0.584	0.271	0.311
r clust sic2	YES	YES	YES	YES	YES	YES
i.fyear	YES	YES	YES	YES	YES	YES
i.gvkey	YES	YES	YES	YES	YES	YES

*, **, *** Indicate statistical significance at the 0.1, 0.05, and 0.01 levels, respectively, based on a two-tailed test. Robust standard errors in parentheses. Variables are defined in Appendix A.

Table 6: CSR in the post-disaster period

VARIABLES	(1) CSR_SCORE	(2) EXTERNAL	(3) INTERNAL	(4) f_CSR_SCORE	(5) f_EXTERNAL	(6) f_INTERNAL	(7) f2_CSR_SCORE	(8) f2_EXTERNAL	(9) f2_INTERNAL	(10) f3_CSR_SCORE	(11) f3_EXTERNAL	(12) f3_INTERNAL
aftertreat_sic2	0.147	0.115***	0.0536	0.154	0.165***	0.0384	0.153	0.217***	0.0761	0.0803	0.165**	0.0640
	(0.110)	(0.0207)	(0.0531)	(0.132)	(0.0257)	(0.0675)	(0.128)	(0.0754)	(0.101)	(0.127)	(0.0750)	(0.0893)
l_SIZE_w	-0.000430	-0.0121*	0.00221	0.00653	-0.000604	0.000846	-0.00965	0.00363	-0.0108	0.00771	0.00803	0.00707
	(0.0150)	(0.00614)	(0.0107)	(0.0157)	(0.00687)	(0.0115)	(0.0154)	(0.00908)	(0.0118)	(0.0165)	(0.0110)	(0.0102)
l_ROA_w	0.0422	0.0267	0.0251	0.121*	0.0160	0.0926**	0.102*	-0.0150	0.0784**	0.0662	-0.0210	0.0392
	(0.0462)	(0.0280)	(0.0312)	(0.0609)	(0.0250)	(0.0422)	(0.0511)	(0.0208)	(0.0373)	(0.0585)	(0.0266)	(0.0414)
l_MB_w	-0.000695	0.000464	-0.00124*	-0.00145	-0.000114	-0.000738	0.000252	0.000808	8.19e-05	0.000151	0.000537	-0.000227
	(0.00112)	(0.000578)	(0.000692)	(0.00103)	(0.000827)	(0.000757)	(0.00113)	(0.000714)	(0.000922)	(0.00144)	(0.000847)	(0.000968)
l_LEV_w	0.0408	0.0218	-0.0141	0.0531	0.00962	-0.00198	0.0764	0.0191	0.00730	0.00323	-0.0200	0.00466
	(0.0436)	(0.0262)	(0.0311)	(0.0473)	(0.0279)	(0.0300)	(0.0522)	(0.0346)	(0.0304)	(0.0529)	(0.0339)	(0.0464)
l_CH_w	0.0286	0.0453**	0.00189	0.00823	0.0159	0.00521	0.0913	0.0167	0.0585	0.0585	0.0108	0.0543
	(0.0766)	(0.0208)	(0.0603)	(0.0776)	(0.0218)	(0.0587)	(0.0774)	(0.0343)	(0.0528)	(0.102)	(0.0368)	(0.0738)
Constant	-0.184*	0.0221	-0.155**	-0.304***	-0.0743*	-0.204**	-0.253**	-0.0841	-0.156**	-0.337***	-0.0866	-0.265***
	(0.102)	(0.0370)	(0.0725)	(0.103)	(0.0424)	(0.0765)	(0.0955)	(0.0607)	(0.0761)	(0.106)	(0.0732)	(0.0648)
Observations	16,281	16,281	16,281	12,968	12,968	12,968	10,591	10,591	10,591	8,712	8,712	8,712
Adjusted R-squared	0.589	0.482	0.590	0.605	0.491	0.608	0.616	0.503	0.618	0.624	0.502	0.622
r clust sic2	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
i.fyear	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
i.gvkey	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

*, **, *** Indicate statistical significance at the 0.1, 0.05, and 0.01 levels, respectively, based on a two-tailed test. Robust standard errors in parentheses. Variables are defined in Appendix A.

Table 7: CSR in the post-disaster period depending on pre-disaster CSR

VARIABLES	(1) CSR_SCORE	(2) EXTERNAL	(3) INTERNAL	(4) f_CSR_SCORE	(5) f_EXTERNAL	(6) f_INTERNAL	(7) f2_CSR_SCORE	(8) f2_EXTERNAL	(9) f2_INTERNAL
					FULL SAMPLE				
aftertreat_sic2	0.147	0.115***	0.0536	0.154	0.165***	0.0384	0.153	0.217***	0.0761
	(0.110)	(0.0207)	(0.0531)	(0.132)	(0.0257)	(0.0675)	(0.128)	(0.0754)	(0.101)
Observations	16,281	16,281	16,281	12,968	12,968	12,968	10,591	10,591	10,591
Adjusted R-squared	0.589	0.482	0.590	0.605	0.491	0.608	0.616	0.503	0.618
					CSR $<$ P(50)				
aftertreat_sic2	0.296**	0.178**	0.152*	0.358**	0.250***	0.165	0.384***	0.294*	0.241*
	(0.115)	(0.0682)	(0.0775)	(0.147)	(0.0751)	(0.0981)	(0.132)	(0.162)	(0.139)
Observations	15,556	15,556	15,556	12,334	12,334	12,334	10,040	10,040	10,040
Adjusted R-squared	0.599	0.494	0.592	0.617	0.502	0.610	0.627	0.512	0.622
					CSR $\geq$ P(50)				
aftertreat_sic2	0.0211	0.0646***	-0.0308	-0.0364	0.0885***	-0.0765**	-0.0547	0.147***	-0.0700
	(0.0998)	(0.0183)	(0.0311)	(0.109)	(0.0234)	(0.0349)	(0.110)	(0.0113)	(0.0569)
Observations	2,067	2,067	2,067	1,844	1,844	1,844	1,627	1,627	1,627
Adjusted R-squared	0.532	0.489	0.594	0.555	0.500	0.612	0.564	0.501	0.614
CONTROLS	YES	YES	YES	YES	YES	YES	YES	YES	YES
r clust sic2	YES	YES	YES	YES	YES	YES	YES	YES	YES
i.fyear	YES	YES	YES	YES	YES	YES	YES	YES	YES
i.gvkey	YES	YES	YES	YES	YES	YES	YES	YES	YES

*, **, *** Indicate statistical significance at the 0.1, 0.05, and 0.01 levels, respectively, based on a two-tailed test. Robust standard errors in parentheses. Variables are defined in Appendix A.

Table 8: Accrual-based earnings management and readability in the post-disaster period.
*, **, *** Indicate statistical significance at the 0.1, 0.05, and 0.01 levels, respectively, based on a two-tailed test. Robust standard errors in parentheses. Variables are defined in Appendix A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	f_EXTERNAL	f2_ABS_DA	f3_ABS_DA	f4_ABS_DA	f5_ABS_DA	f_EXTERNAL	f_bogindex	f2_bogindex
aftertreat_sic2	0.155***					0.127**		
	(0.0349)					(0.0481)		
f_EXTERNAL_w_hat		0.0594***	0.0654***	0.0968**	0.180*		7.549*	9.385**
		(0.0174)	(0.0215)	(0.0438)	(0.0940)		(4.089)	(4.534)
l_SIZE_w	0.0104	-0.00248	-0.00117	0.00202	-0.00249	0.00252	-0.0225	0.0983
	(0.00842)	(0.00191)	(0.00189)	(0.00194)	(0.00225)	(0.00658)	(0.0805)	(0.0822)
l_MB_w	-0.000423	9.16e-05	0.000478	0.000313	0.000107			
	(0.000953)	(0.000267)	(0.000339)	(0.000346)	(0.000302)			
l_ROA_w	0.0293	0.00136	-0.00929	-0.00366	-0.00228			
	(0.0260)	(0.00871)	(0.0123)	(0.0147)	(0.0157)			
l_LEV_w	0.0422	-0.00532	-0.0166*	-0.00249	-0.00482			
	(0.0307)	(0.00944)	(0.00980)	(0.0108)	(0.00727)			
l_CH_w	-0.0146	0.0147***	0.0139**	0.0121	0.00218			
	(0.0246)	(0.00497)	(0.00536)	(0.00845)	(0.00729)			
l_COMBINED_RAM	-0.0144	-0.00545	0.00347	0.00138	0.00916*			
	(0.0140)	(0.00430)	(0.00481)	(0.00386)	(0.00496)			
l_RD_INT_w	0.00333	0.00567***	0.00719***	0.00653***	-0.00933***			
	(0.00219)	(0.000842)	(0.000735)	(0.000928)	(0.00123)			
l_AD_IND_INT_w	0.522	-0.000410	0.179	0.0990	-0.134			
	(0.432)	(0.0650)	(0.147)	(0.0720)	(0.184)			
l_GOVERNANCE_w	0.0192***	-0.00166*	-0.00241**	-0.00282	-0.00398			
	(0.00623)	(0.000893)	(0.00112)	(0.00173)	(0.00238)			
l_BIG4	0.00470	0.00177	0.00387	-0.0126*	-1.43e-05			

Continued on next page

Continued from previous page

VARIABLES	(1) f_EXTERNAL	(2) f2_ABS_DA	(3) f3_ABS_DA	(4) f4_ABS_DA	(5) f5_ABS_DA	(6) f_EXTERNAL	(7) f_bogindex	(8) f2_bogindex
	(0.0226)	(0.00456)	(0.00595)	(0.00736)	(0.00590)			
l_FIRM_AGE_w	-0.0646***	0.00159	0.000545	0.00194	0.00654			
	(0.0164)	(0.00341)	(0.00492)	(0.00592)	(0.00832)			
l_MTB_w						0.000145	-0.0105	-0.0144*
						(0.00101)	(0.00931)	(0.00818)
l_AGE						-0.0540***	0.496	0.670*
						(0.0145)	(0.302)	(0.342)
l_SI_w						0.00202	-1.799***	-1.081**
						(0.0460)	(0.544)	(0.461)
l_GEOSEG						0.00473	-0.462***	-0.235
						(0.0234)	(0.153)	(0.177)
l_BUSSEG						0.0148	-0.102	-0.252
						(0.0177)	(0.249)	(0.271)
LN_NITEMS						-0.146*	4.505***	4.549***
						(0.0753)	(1.407)	(1.232)
Constant	-0.0387	0.0582***	0.0556**	0.0471***	0.0922***	0.831**	58.10***	57.78***
	(0.0671)	(0.0190)	(0.0259)	(0.0158)	(0.0288)	(0.411)	(7.969)	(6.833)
Observations	11,080	11,040	9,306	7,801	6,460	13,304	15,959	15,149
Adjusted R-squared	0.500	0.276	0.299	0.292	0.296	0.492	0.869	0.874
r clust sic2	YES	YES	YES	YES	YES	YES	YES	YES
i.fyear	YES	YES	YES	YES	YES	YES	YES	YES
i.gvkey	YES	YES	YES	YES	YES	YES		
i.cusip							YES	YES

Table 9: Accrual-based earnings management and readability in the post-disaster period depending on pre-disaster CSR

VARIABLES	(1) f_EXTERNAL	(2) f2_ABS_DA	(3) f3_ABS_DA	(4) f4_ABS_DA	(5) f5_ABS_DA	(6) f_EXTERNAL	(7) f_bogindex	(8) f2_bogindex
				FULL SAMPLE				
aftertreat_sic2	0.155***					0.127**		
	(0.0349)					(0.0481)		
f_EXTERNAL_w_hat		0.0594***	0.0654***	0.0968**	0.180*		7.549*	9.385**
		(0.0174)	(0.0215)	(0.0438)	(0.0940)		(4.089)	(4.534)
Observations	11,080	11,040	9,306	7,801	6,460	13,304	15,959	15,149
Adjusted R-squared	0.500	0.276	0.299	0.292	0.296	0.492	0.869	0.874
				CSR $<$ P(50)				
aftertreat_sic2	0.230**					0.238***		
	(0.0862)					(0.0665)		
f_EXTERNAL_w_hat		0.0390***	0.0544***	0.0748**	0.125***		2.720**	3.964***
		(0.0102)	(0.0187)	(0.0276)	(0.0413)		(1.359)	(0.995)
Observations	10,506	10,440	8,778	7,340	6,067	12,760	15,351	14,560
Adjusted R-squared	0.512	0.269	0.293	0.286	0.290	0.498	0.868	0.873
				CSR $\geq$ P(50)				
aftertreat_sic2	0.0891***					0.0336		
	(0.0134)					(0.0262)		
f_EXTERNAL_w_hat		0.110**	0.0956**	0.158	0.307		39.43*	45.72
		(0.0412)	(0.0357)	(0.0939)	(0.219)		(22.04)	(27.59)
Observations	10,558	10,490	8,826	7,379	6,099	12,867	15,487	14,693
Adjusted R-squared	0.485	0.272	0.291	0.286	0.291	0.486	0.867	0.872
CONTROLS	YES	YES	YES	YES	YES	YES	YES	YES
r clust sic2	YES	YES	YES	YES	YES	YES	YES	YES
i.fyear	YES	YES	YES	YES	YES	YES	YES	YES
i.gvkey	YES	YES	YES	YES	YES	YES		
i.cusip							YES	YES

*, **, *** Indicate statistical significance at the 0.1, 0.05, and 0.01 levels, respectively, based on a two-tailed test. Robust standard errors in parentheses. Variables are defined in Appendix A.

Table 10: Leads and lags model

VARIABLES	(1) EXTERNAL	(2) EXTERNAL	(3) EXTERNAL
aftertreat_sic2_-3	-0.0302***	-0.0304***	-0.0331***
	(0.00678)	(0.00679)	(0.00699)
aftertreat_sic2_-2	-0.0492***	-0.0498***	-0.0534**
	(0.0178)	(0.0175)	(0.0201)
aftertreat_sic2_-1	-0.0710***	-0.0715***	-0.0755***
	(0.0216)	(0.0213)	(0.0236)
aftertreat_sic2_0	0.00905	0.0125	-0.0296**
	(0.0120)	(0.0124)	(0.0119)
aftertreat_sic2_+1	0.0265	0.0191	
	(0.0228)	(0.0217)	
aftertreat_sic2_+2	0.153**		
	(0.0633)		
aftertreat_sic2_3+	0.268***		
	(0.0594)		
aftertreat_sic2_2+		0.176**	
		(0.0747)	
aftertreat_sic2_1+			0.139***
			(0.0287)
Constant	0.0188	0.0173	0.0168
	(0.0349)	(0.0344)	(0.0335)
CONTROLS	YES	YES	YES
Observations	16,281	16,281	16,281
Adjusted R-squared	0.487	0.487	0.486
r clust sic2	YES	YES	YES
i.fyear	YES	YES	YES
i.gvkey	YES	YES	YES

*, **, *** Indicate statistical significance at the 0.1, 0.05, and 0.01 levels, respectively, based on a two-tailed test. $aftertreat\_sic2\_-3$ ($aftertreat\_sic2\_-2$ or $aftertreat\_sic2\_-1$) is a dummy variable that equals to 1 for the treated industries three (two or one) years before the disaster and 0 otherwise. $aftertreat\_sic2\_+1$ ($aftertreat\_sic2\_+2$) is a dummy variable that equals to 1 for the treated industries one (two) year(s) after the disaster and 0 otherwise. $aftertreat\_sic2\_1+$ ($aftertreat\_sic2\_2+$ or $aftertreat\_sic2\_3+$) is a dummy variable that equals to 1 for the treated industries for all years except the first one (the firms two; or the first three) after the disaster and 0 otherwise. The rest of the variables are as described in Appendix A.

Figure 2: Placebo test. Probability density function of the placebo coefficients (random industries)

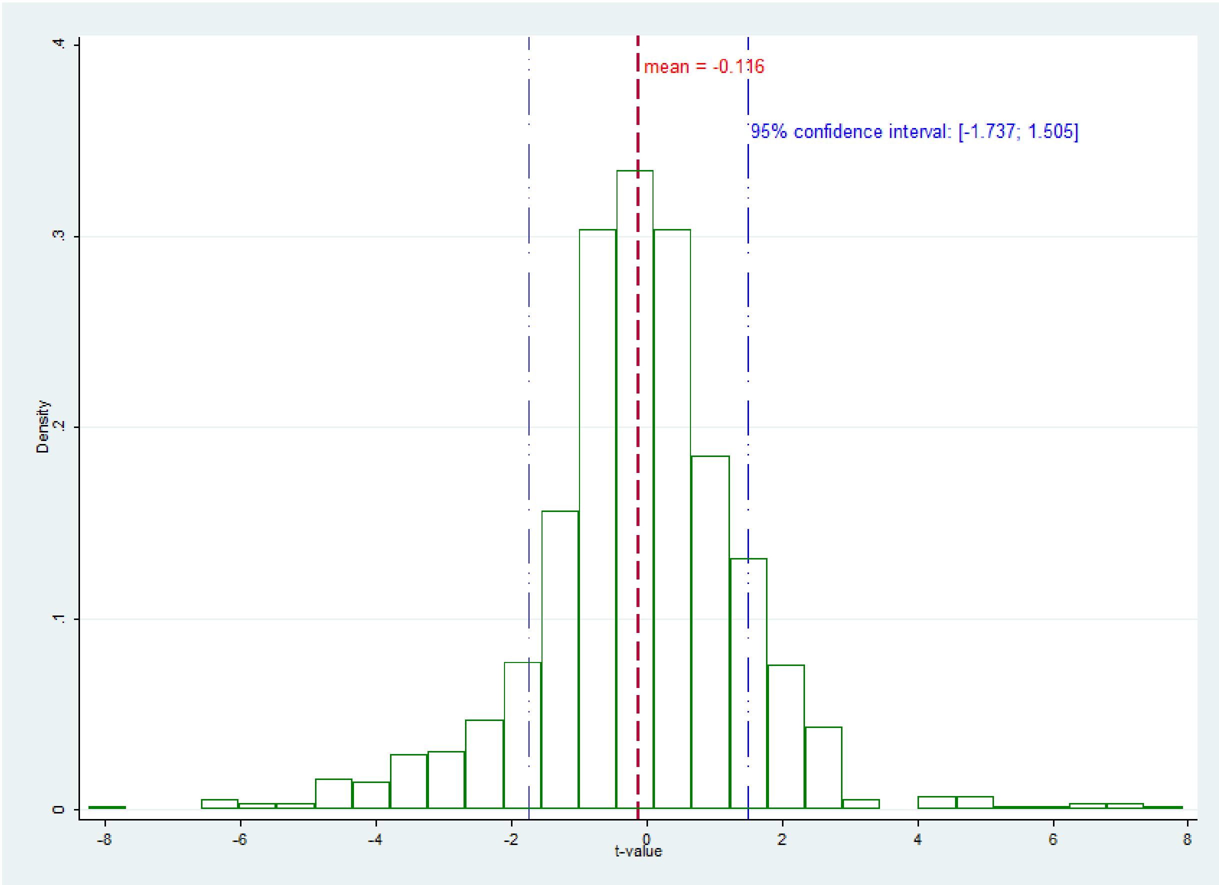


Table 11: SGAE in the post-disaster period depending on pre-disaster CSR

VARIABLES	(1) SGA	(2) SGA_AD	(3) SGA_R	(4) f_SGA	(5) f_SGA_AD	(6) f_SGA_R	(7) f2_SGA	(8) f2_SGA_AD	(9) f2_SGA_R
FULL SAMPLE									
aftertreat_sic2	0.00192	0.00175	-0.00732	0.0101	0.00250	0.00779	-0.0159	0.00622	0.0217**
	(0.0536)	(0.00307)	(0.00923)	(0.0682)	(0.00392)	(0.00577)	(0.0595)	(0.00430)	(0.00905)
Observations	14,565	14,565	14,497	13,923	13,923	13,865	13,183	13,183	13,137
Adjusted R-squared	0.977	0.903	0.903	0.977	0.908	0.913	0.977	0.913	0.927
CSR $<$ P(50)									
aftertreat_sic2	-0.0513	-0.00253	-0.0215	-0.0197	0.000548	0.00721	-0.0213	0.00643	0.0313**
	(0.0572)	(0.00398)	(0.0161)	(0.0640)	(0.00398)	(0.00654)	(0.0472)	(0.00511)	(0.0142)
Observations	14,224	14,224	14,159	13,596	13,596	13,540	12,873	12,873	12,828
Adjusted R-squared	0.978	0.875	0.901	0.978	0.891	0.911	0.978	0.897	0.925
CSR $\geq$ P(50)									
aftertreat_sic2	0.102	0.00707**	-0.0209*	0.0976	0.00597*	-0.00643	0.0442	0.00648**	0.0155***
	(0.0609)	(0.00331)	(0.0117)	(0.0915)	(0.00354)	(0.00870)	(0.100)	(0.00319)	(0.00531)
Observations	14,186	14,186	14,118	13,549	13,549	13,491	12,816	12,816	12,770
Adjusted R-squared	0.977	0.899	0.907	0.977	0.905	0.917	0.978	0.910	0.928
CONTROLS	YES	YES	YES	YES	YES	YES	YES	YES	YES
r clust sic2	YES	YES	YES	YES	YES	YES	YES	YES	YES
i.fyear	YES	YES	YES	YES	YES	YES	YES	YES	YES
i.gvkey	YES	YES	YES	YES	YES	YES	YES	YES	YES

*, **, *** Indicate statistical significance at the 0.1, 0.05, and 0.01 levels, respectively, based on a two-tailed test. Robust standard errors in parentheses. Variables are defined in Appendix A.

Table 12: STRENGTHS and CONCERNS in the post-disaster period depending on pre-disaster CSR

VARIABLES	(1) STRENGHTS	(2) CONCERNS	(3) f_STRENGHTS	(4) f_CONCERNS	(5) f2_STRENGHTS	(6) f2_CONCERNS
			FULL SAMPLE			
aftertreat_sic2	0.124***	-0.0200	0.118***	-0.0335	0.120***	-0.0294
	(0.0404)	(0.0763)	(0.0397)	(0.101)	(0.0300)	(0.110)
Observations	16,281	16,281	12,968	12,968	10,591	10,591
Adjusted R-squared	0.737	0.614	0.743	0.629	0.754	0.636
			CSR $<$ P(50)			
aftertreat_sic2	0.213***	-0.0239	0.210***	-0.0820	0.203***	-0.0829
	(0.0776)	(0.0475)	(0.0696)	(0.0825)	(0.0458)	(0.0748)
Observations	15,616	15,616	12,393	12,393	10,095	10,095
Adjusted R-squared	0.743	0.618	0.748	0.634	0.759	0.640
			CSR $\geq$ P(50)			
aftertreat_sic2	0.0336	-0.0185	0.0109	0.0250	0.0220	0.0317
	(0.0214)	(0.108)	(0.0219)	(0.127)	(0.0207)	(0.154)
Observations	15,572	15,572	12,339	12,339	10,039	10,039
Adjusted R-squared	0.738	0.587	0.744	0.601	0.755	0.606
CONTROLS	YES	YES	YES	YES	YES	YES
r clust sic2	YES	YES	YES	YES	YES	YES
i.fyear	YES	YES	YES	YES	YES	YES
i.gvkey	YES	YES	YES	YES	YES	YES

*, **, *** Indicate statistical significance at the 0.1, 0.05, and 0.01 levels, respectively, based on a two-tailed test. Robust standard errors in parentheses. Variables are defined in Appendix A.

APPENDIX A: Variable Definitions

Variable	Variable Definitions
Dummy Variables	
$aftertreat\_sic2\_0$	Is a dummy equal to 1 for the treated industries in the year of the disaster and 0 otherwise.
$aftertreat\_sic2$	Is a dummy equal to 1 for the treated industries in the years after the disasters (including the year of the disaster) and 0 otherwise.
CSR Variables	
CSR_SCORE	Net score of CSR ratings, measured as total strengths minus total concerns in five social rating categories of KLD ratings data: community, diversity, employee relations, environment, and product.
$ENVIRONMENT$	Net score of CSR ratings, measured as environment strength minus environment concerns of KLD rating data.
$COMMUNITY$	Net score of CSR ratings, measured as community strength minus community concerns of KLD rating data.
$DIVERSITY$	Net score of CSR ratings, measured as diversity strength minus diversity concerns of KLD rating data.
$EMPLOYEE$	Net score of CSR ratings, measured as employee relations strength minus employee relations concerns of KLD rating data.
$PRODUCT$	Net score of CSR ratings, measured as product strength minus product concerns of KLD rating data.
$GOVERNANCE$	Net score of CSR ratings, measured as governance strength minus governance concerns of KLD rating data.
$HUMAN$	Net score of CSR ratings, measured as human strength minus human concerns of KLD rating data.
$STRENGHTS$	Net score of CSR ratings, measured as sum of total strengths in five social rating categories of KLD ratings data: community, diversity, employee relations, environment, and product.
$CONCERNS$	Net score of CSR ratings, measured as sum of total concerns in five social rating categories of KLD ratings data: community, diversity, employee relations, environment, and product.
$INTERNAL$	Net score of CSR ratings, measured as total strengths minus total concerns in two social rating categories of KLD ratings data: diversity and employee relations.

Continued on next page

Continued from previous page

Variable	Variable Definitions
$EXTERNAL$	Net score of CSR ratings, measured as total strengths minus total concerns in three social rating categories of KLD ratings data: community, human rights, and environment.
Earnings Management Variables	
AAC	Signed discretionary accruals, where discretionary accruals are computed through the cross-sectional modified Jones model adjusted for performance (Kim et al., 2012).
ABS_DA	Absolute value of discretionary accruals, where discretionary accruals are computed through the cross-sectional modified Jones model adjusted for performance (Kim et al., 2012).
AB_CFO	The level of abnormal cash flows from operations (Kim et al., 2012).
AB_PROD	The level of abnormal production costs, where production costs are defined as the sum of cost of goods sold and the change in inventories (Kim et al., 2012).
AB_EXP	The level of abnormal discretionary expenses, where discretionary expenses are the sum of R&D expenses, advertising expenses, and SG&A expenses (Kim et al., 2012).
$COMBINED\_RAM$	$COMBINED\_RAM = AB\_CFO - AB\_PROD + AB\_EXP$ (Kim et al., 2012).
Readability Variables	
$bogindex$	The Bog Index (Bonsall et al., 2017).
Control Variables	
$SIZE$	Natural logarithm of the market value of equity (MVE) (Kim et al., 2012).
MB	Market-to-book equity ratio, measured as MVE/BVE, where BVE is the book value of equity (Kim et al., 2012).
ADJ_ROA	Industry-adjusted ROA, where ROA is measured as income before extraordinary items, scaled by lagged total assets (Kim et al., 2012).
LEV	Long-term debt scaled by total assets (Kim et al., 2012).
RD_INT	R&D intensity (R&D expense/net sales) for the year (Kim et al., 2012).
AD_IND_INT	Advertising intensity for the two-digit SIC code industry for the year (Kim et al., 2012).
CH	Cash holding is the ratio of cash and short-term investments to the book value of assets (Flammer, 2015).
ROA	Is the ratio of income before extraordinary items to the book value of assets (Flammer, 2015).
ROE	Is the ratio of income before extraordinary items to the lagged Common/Ordinary Equity-Total.
$SALE\_G$	Sales growth. (sale-l.sale)/l.sale.

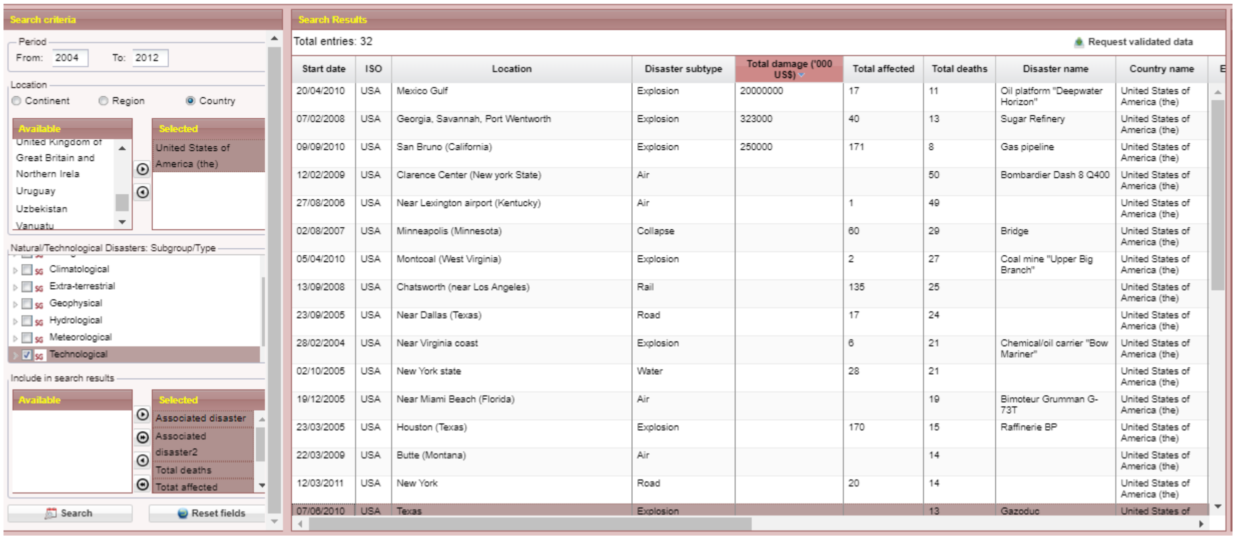
Continued on next page

Continued from previous page

Variable	Variable Definitions
BM	Book-to-market ratio measured at the fiscal year-end (Huang et al., 2014).
MTB	Market-to-book ratio measured at the fiscal year-end (Li, 2008).
AGE	log(1+#years since a firm appears in CRSP monthly file) (Huang et al., 2014).
$BUSSEG$	log(1+# of business segments) (Huang et al., 2014).
$GEOSEG$	log(1+# of geographic segments) (Huang et al., 2014).
SI	The amount of special items scaled by book value of assets (Li, 2008).
$NITEMS$	The logarithm of the number of non-missing items in Compustat as a proxy for financial complexity (Li, 2008).
SGA	Log SG&A expenses (Di Giuli and Kostovetsky, 2014).
SGA_AD	Log(SG&A-advertising) (Di Giuli and Kostovetsky, 2014).
SGA_R	SG&A/revenue (Di Giuli and Kostovetsky, 2014).

APPENDIX B: EM-DAT Output

Figure 3: EM-DAT output



Start date	ISO	Location	Disaster subtype	Total damage ('000 US$)	Total affected	Total deaths	Disaster name	Country name	E
20/04/2010	USA	Mexico Gulf	Explosion	20000000	17	11	Oil platform "Deepwater Horizon"	United States of America (the)	
07/02/2008	USA	Georgia, Savannah, Port Wentworth	Explosion	323000	40	13	Sugar Refinery	United States of America (the)	
09/09/2010	USA	San Bruno (California)	Explosion	250000	171	8	Gas pipeline	United States of America (the)	
12/02/2009	USA	Clarence Center (New york State)	Air			50	Bombardier Dash 8 Q400	United States of America (the)	
27/08/2006	USA	Near Lexington airport (Kentucky)	Air		1	49		United States of America (the)	
02/08/2007	USA	Minneapolis (Minnesota)	Collapse		60	29	Bridge	United States of America (the)	
05/04/2010	USA	Montcoal (West Virginia)	Explosion		2	27	Coal mine "Upper Big Branch"	United States of America (the)	
13/09/2008	USA	Chatsworth (near Los Angeles)	Rail		135	25		United States of America (the)	
23/09/2005	USA	Near Dallas (Texas)	Road		17	24		United States of America (the)	
28/02/2004	USA	Near Virginia coast	Explosion		6	21	Chemical/oil carrier "Bow Mariner"	United States of America (the)	
02/10/2005	USA	New York state	Water		28	21		United States of America (the)	
19/12/2005	USA	Near Miami Beach (Florida)	Air			19	Bimoteur Grumman G-73T	United States of America (the)	
23/03/2005	USA	Houston (Texas)	Explosion		170	15	Raffinerie BP	United States of America (the)	
22/03/2009	USA	Butte (Montana)	Air			14		United States of America (the)	
12/03/2011	USA	New York	Road		20	14		United States of America (the)	
07/06/2010	USA	Texas	Explosion			13	Gazoduc	United States of	

APPENDIX C: MSCI ESG KLD STATS

KLD data starts in 1991 and is based on the assessment of how well firms perform in environmental, social, and governance issues. KLD covers five universes presented in Table 14.

Table 14: KLD Universes

Data Set Universe	Time Series	Number of Companies**	Inclusion in Our Study
Universe A			
MSCI KLD 400 Social Index + MSCI USA Index	1991 - present	650	From 2003
Universe B			
Largest 1000 U.S. companies by market capitalization	2008-2013	1000 (*discounted**)	No
Universe C			
MSCI KLD 400 Social Index + 1000 Largest U.S. Companies	2001-2013	1100 (*discounted**)	No
Universe D			
MSCI USA IMI Index	2003-present	2400	Yes
Universe E			
Non-U.S. Universe	2013 - present	2600	No

* Universes B and C have been discounted as of STATS-2014 Data Set.
** Number of firms is an approximate average for the time series.

Universe A covers firms that are included in MSCI KLD 400 Social Index & MSCI USA index. Universe D covers the 3000 largest US firms measured by market capitalization and are not covered by Universe A. RiskMetrics acquired KLD in 2009 and Morgan Stanley Capital International (MSCI) acquired RiskMetrics in 2010. In 2010 the industry-based key issue rating model was introduced to KLD. Prior 2010 all of the positive ESG performance indicators were searched for all of the companies. Starting from 2010, all companies are assessed for limited set of industry specific positive ESG indicators. Thus, our results can be affected by the change in the methodology in 2010. We acknowledge this caveat as follows.

First, we apply differences-in-difference approach (DiD) which eliminates two main concerns. First, if the conclusion is done only based on the difference between treated and control groups in the post-disaster period. In this case, the final result may capture only the permanent difference between treatment and control. Second, if the conclusion is done only based on the difference between treatment group before and after the disaster. This result can be driven by the trends in the database. Thus, if in 2009 or (and) 2010 some methodological aspects of KLD were changed, this difference is captured by the DiD design.

Second, we are consistent with the prior literature that uses KLD databases before and after 2010 and (or) 2011. Some selected examples are Flammer and Luo (2017) (1991-2013), Petrenko et al. (2016) (1997-2012) and Marano and Kostova (2016) (2007-2011).

APPENDIX D: 2SLS Estimations

We perform IV estimation with $COMBINED\_RAM$ and $READABILITY$ as dependent variables. The first stage for $COMBINED\_RAM$:

$$\begin{aligned} EXTERNAL_{t+1} = &\beta_0 + \beta_1 aftertreat\_sic2 + \beta_2 SIZE_{t-1} + \beta_3 ROA_{t-1} + \beta_4 MB_{t-1} \\ &+ \beta_5 LEV_{t-1} + \beta_6 CH_{t-1} + \beta_7 ABS\_DA_{t-1} + \beta_8 RD\_INT_{t-1} \\ &+ \beta_9 AD\_IND\_INT_{t-1} + \beta_{10} GOVERNANCE_{t-1} + \beta_{11} BIG4_{t-1} \\ &+ \beta_{12} FIRM\_AGE_{t-1} + \epsilon_t, \end{aligned} \tag{5}$$

The second stage for $COMBINED\_RAM$:

$$\begin{aligned} RAM = &\beta_0 + \beta_1 f\_EXTERNAL\_w\_hat_t + \beta_2 SIZE_{t-1} + \beta_3 ROA_{t-1} + \beta_4 MB_{t-1} \\ &+ \beta_5 LEV_{t-1} + \beta_6 CH_{t-1} + \beta_7 ABS\_DA_{t-1} + \beta_8 RD\_INT_{t-1} \\ &+ \beta_9 AD\_IND\_INT_{t-1} + \beta_{10} GOVERNANCE_{t-1} + \beta_{11} BIG4_{t-1} \\ &+ \beta_{12} FIRM\_AGE_{t-1} + \epsilon_t, \end{aligned} \tag{6}$$

where, RAM is a proxy for real activities manipulation ($COMBINED\_RAM$) in $t+1$, $t+2$, and $t+3$. The first stage for $READABILITY$:

$$\begin{aligned} EXTERNAL_{t+1} = &\beta_0 + \beta_1 aftertreat\_sic2 + \beta_2 LEV_{t-1} + \beta_3 CH_{t-1} + \beta_4 EARN_{t-1} \\ &+ \beta_5 RET_{t-1} + \beta_6 SIZE_{t-1} + \beta_7 BM_{t-1} + \beta_8 STD\_RET_{t-1} \\ &+ \beta_9 AGE_{t-1} + \beta_{10} BUSSEG_{t-1} + \beta_{11} GEOSEG_{t-1} \\ &+ \beta_{12} D\_EARN_{t-1} + \beta_{13} AFE_{t-1} + \beta_{14} AF_{t-1} + \beta_{15} lLOSS_{t-1} + \epsilon_t, \end{aligned} \tag{7}$$

The second stage for $READABILITY$:

$$\begin{aligned} READABILITY = &\beta_0 + \beta_1 f\_EXTERNAL\_w\_hat_t + \beta_2 LEV_{t-1} + \beta_3 CH_{t-1} + \beta_4 EARN_{t-1} \\ &+ \beta_5 RET_{t-1} + \beta_6 SIZE_{t-1} + \beta_7 BM_{t-1} + \beta_8 STD\_RET_{t-1} \\ &+ \beta_9 AGE_{t-1} + \beta_{10} BUSSEG_{t-1} + \beta_{11} GEOSEG_{t-1} \\ &+ \beta_{12} D\_EARN_{t-1} + \beta_{13} AFE_{t-1} + \beta_{14} AF_{t-1} + \beta_{15} lLOSS_{t-1} + \epsilon_t, \end{aligned} \tag{8}$$

where, $READABILITY$ is $bogindex$ in $t+1$ and $t+2$.

APPENDIX E: PROPENSITY SCORE MATCHING

Table 15: Summary statistics for treated and matched control group

Variable	Unmatched Matched	Mean Treated	Mean Control	%bias	%reduct $\mid bias \mid$	t-test t	t-test p>$\mid t \mid$
CSR_SCORE	U	-0.49018	-0.23177	-54.4		-5.68	0.000
	M	-0.49018	-0.50147	2.4	95.6	0.19	0.846
SIZE	U	7.5982	6.7341	53		6.69	0.000
	M	7.9053	7.974	-4.2	92.1	-0.36	0.719
MB	U	2.2929	2.8058	-18.9		-1.96	0.05
	M	2.2013	2.3128	-4.1	78.3	-0.43	0.665
LEV	U	0.30454	0.19999	58		6.15	0.000
	M	0.30569	0.31861	-7.2	87.6	-0.66	0.507
ROA	U	0.00522	0.00387	1		0.12	0.907
	M	0.01705	0.01367	2.6	-150.1	0.3	0.767
CH	U	0.0479	0.21814	-108.8		-10.18	0.000
	M	0.04214	0.05115	-5.8	94.7	-1.17	0.242

This table reports summary statistics for treated and match control group based on single nearest-neighbour (without caliper), 1-to-1 matching without replacement. The nearest neighbor is calculated based on six firm-level characteristics: CSR_SCORE, size ($SIZE$), market-to-book (MB), leverage ratio (LEV), return on asset (ROA), and cash holdings (CH). Matching variables are computed as average in the three years preceding the disasters. A bias before and after matching is calculated for each variable and the change in this bias is stated. This "bias" is defined as the difference of the mean values of the treatment group and the (not matched / matched) non treatment group, divided by the square root of the average sample variance in the treatment group and the not matched non treatment group. Variables are defined in Appendix A.

Figure 4: Dot graph of covariate balance

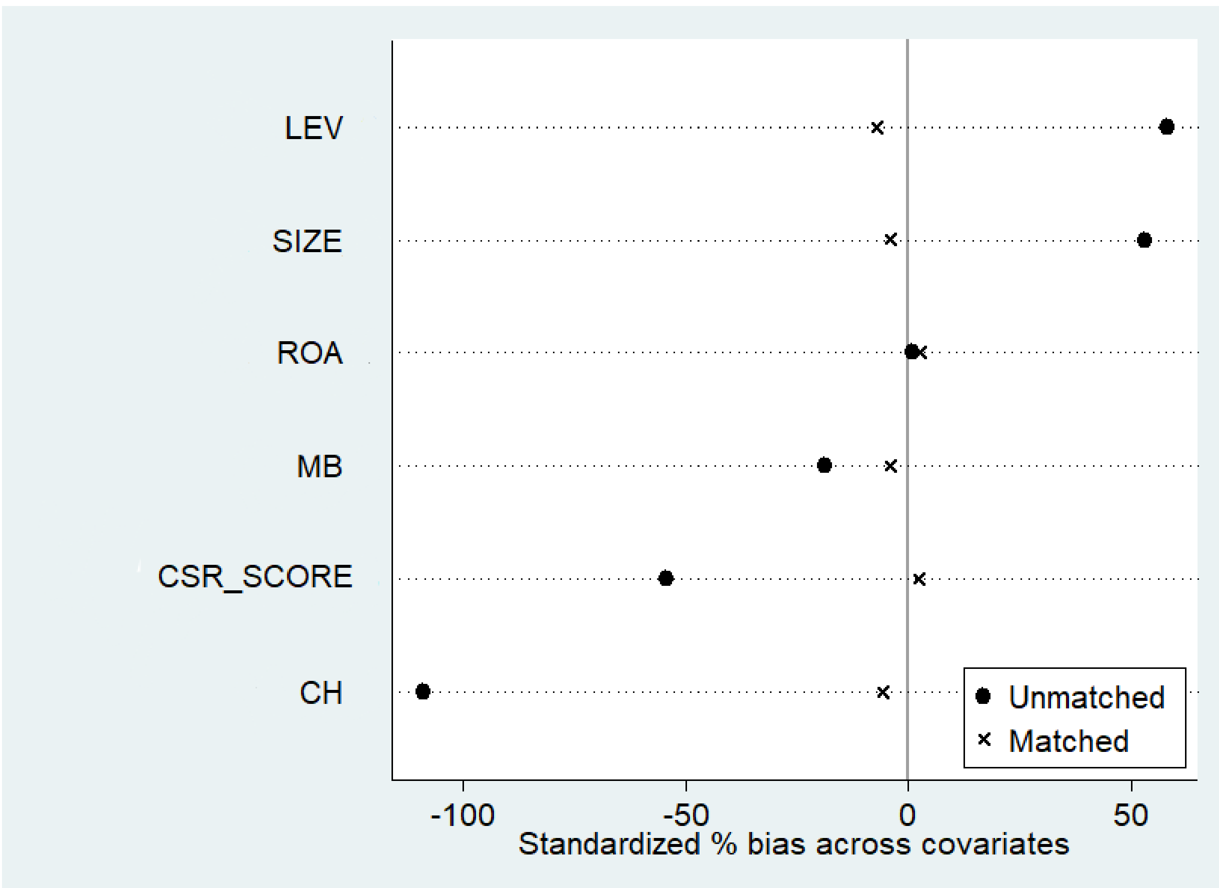


Table 16: PSM Sample. CSR in the post-disaster period

VARIABLES	(1) CSR_SCORE	(2) EXTERNAL	(3) INTERNAL	(4) f_CSR_SCORE	(5) f_EXTERNAL	(6) f_INTERNAL	(7) f2_CSR_SCORE	(8) f2_EXTERNAL	(9) f2_INTERNAL	(10) f3_CSR_SCORE	(11) f3_EXTERNAL	(12) f3_INTERNAL
aftertreat_sic2	0.218	0.245	-0.0856	0.284	0.341*	-0.0446	0.479	0.464	0.207	0.398	0.413	0.207
	(0.657)	(0.258)	(0.285)	(0.597)	(0.173)	(0.320)	(0.480)	(0.273)	(0.418)	(0.526)	(0.294)	(0.487)
l_SIZE_w	-0.221	-0.105	-0.0548	-0.213	-0.0521	-0.0766	-0.269	-0.0255	-0.0890	-0.127	0.0166	0.0284
	(0.159)	(0.0785)	(0.113)	(0.201)	(0.0831)	(0.143)	(0.186)	(0.0872)	(0.142)	(0.170)	(0.0915)	(0.122)
l_ROA_w	0.923	0.512	0.138	2.306***	0.821	0.808**	2.562***	0.352	1.201***	1.083	-0.469	0.745
	(0.603)	(0.519)	(0.298)	(0.805)	(0.582)	(0.317)	(0.775)	(0.373)	(0.373)	(0.930)	(0.280)	(0.566)
l_MB_w	0.00166	0.0106	-0.00970	-0.00169	0.00653	-0.0129	0.0110	0.0104	-0.0109	0.0156	-0.00908	0.000422
	(0.0104)	(0.00985)	(0.0111)	(0.00940)	(0.0109)	(0.00837)	(0.0180)	(0.0157)	(0.00913)	(0.0249)	(0.0207)	(0.0128)
l_LEV_w	-1.368**	-0.522	-0.873**	-0.816	-0.473	-0.522	-0.454	-0.526	-0.369	-0.129	-1.001*	0.293
	(0.610)	(0.470)	(0.365)	(0.648)	(0.505)	(0.329)	(0.731)	(0.533)	(0.387)	(0.923)	(0.568)	(0.486)
l_CH_w	0.180	0.205	0.0833	0.838	0.209	0.715	2.140**	0.305	1.857***	2.574	0.913	2.002**
	(0.965)	(0.483)	(0.679)	(0.890)	(0.633)	(0.683)	(0.920)	(0.813)	(0.670)	(1.535)	(0.701)	(0.843)
Constant	1.333	0.447	0.711	0.845	-0.114	0.661	0.702	-0.419	0.542	-0.521	-0.514	-0.567
	(1.156)	(0.535)	(0.821)	(1.412)	(0.526)	(1.021)	(1.342)	(0.634)	(1.110)	(1.348)	(0.731)	(0.984)
Observations	2,583	2,583	2,583	2,305	2,305	2,305	2,032	2,032	2,032	1,768	1,768	1,768
Adjusted R-squared	0.543	0.499	0.610	0.557	0.503	0.625	0.571	0.505	0.631	0.576	0.498	0.623
r clust sic2	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
i.fyear	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
i.gvkey	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

*, **, *** Indicate statistical significance at the 0.1, 0.05, and 0.01 levels, respectively, based on a two-tailed test. Robust standard errors in parentheses. Variables are defined in Appendix A.

Table 17: PSM Sample. CSR in the post-disaster period depending on pre-disaster CSR

VARIABLES	(1) CSR_SCORE	(2) EXTERNAL	(3) INTERNAL	(4) f_CSR_SCORE	(5) f_EXTERNAL	(6) f_INTERNAL	(7) f2_CSR_SCORE	(8) f2_EXTERNAL	(9) f2_INTERNAL
FULL SAMPLE									
aftertreat_sic2	0.218	0.245	-0.0856	0.284	0.341*	-0.0446	0.479	0.464	0.207
	(0.657)	(0.258)	(0.285)	(0.597)	(0.173)	(0.320)	(0.480)	(0.273)	(0.418)
Observations	2,583	2,583	2,583	2,305	2,305	2,305	2,032	2,032	2,032
Adjusted R-squared	0.543	0.499	0.610	0.557	0.503	0.625	0.571	0.505	0.631
CSR $<$ P(50)									
aftertreat_sic2	1.127	0.608***	0.368	1.355**	0.742***	0.513	1.629***	0.803*	0.923
	(0.704)	(0.199)	(0.393)	(0.649)	(0.192)	(0.448)	(0.546)	(0.464)	(0.599)
Observations	1,917	1,917	1,917	1,714	1,714	1,714	1,514	1,514	1,514
Adjusted R-squared	0.601	0.549	0.623	0.618	0.552	0.640	0.630	0.550	0.653
CSR $\geq$ P(50)									
aftertreat_sic2	-0.669	-0.102	-0.536**	-0.763	-0.0383	-0.602***	-0.649*	0.137	-0.496**
	(0.564)	(0.311)	(0.195)	(0.478)	(0.212)	(0.191)	(0.355)	(0.171)	(0.229)
Observations	1,970	1,970	1,970	1,757	1,757	1,757	1,547	1,547	1,547
Adjusted R-squared	0.555	0.468	0.628	0.567	0.475	0.638	0.579	0.476	0.637
CONTROLS	YES	YES	YES	YES	YES	YES	YES	YES	YES
r clust sic2	YES	YES	YES	YES	YES	YES	YES	YES	YES
i.fyear	YES	YES	YES	YES	YES	YES	YES	YES	YES
i.gvkey	YES	YES	YES	YES	YES	YES	YES	YES	YES

*, **, *** Indicate statistical significance at the 0.1, 0.05, and 0.01 levels, respectively, based on a two-tailed test. Robust standard errors in parentheses. Variables are defined in Appendix A.

Table 18: PSM Sample. Accrual-based earnings management and real activities manipulation in the post-disaster period. For firms with low CSR before the disaster

*, **, *** Indicate statistical significance at the 0.1, 0.05, and 0.01 levels, respectively, based on a two-tailed test. Robust standard errors in parentheses. RAM - COMBINED_RAM. Variables are defined in Appendix A.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	f_EXTERNAL	f2_ABS_DA	f3_ABS_DA	f4_ABS_DA	f5_ABS_DA	f_EXTERNAL	f_RAM	f2_RAM	f3_RAM
aftertreat_sic2	0.197***					0.198***			
	(0.0586)					(0.0597)			
f_EXTERNAL_w_hat		0.0249	0.0319	0.0704***	0.133***		-0.195*	-0.249**	-0.198**
		(0.0150)	(0.0190)	(0.0253)	(0.0439)		(0.112)	(0.117)	(0.0906)
l_SIZE_w	0.00428	0.00315	-0.00148	-0.00562	-0.000199	0.00183	0.000430	0.00188	-0.00839
	(0.0330)	(0.00373)	(0.00399)	(0.00645)	(0.00521)	(0.0325)	(0.0157)	(0.0131)	(0.0121)
l_MB_w	-0.000260	0.000275	-0.000494	5.74e-05	-0.000744*	-0.000370	0.00283	0.00176	0.00221
	(0.00333)	(0.000565)	(0.000465)	(0.000692)	(0.000379)	(0.00346)	(0.00166)	(0.00161)	(0.00157)
l_ROA_w	0.0586	0.00890	0.0288	0.0148	-0.0144	0.110	0.114*	0.0228	-0.0437
	(0.162)	(0.0198)	(0.0188)	(0.0356)	(0.0341)	(0.176)	(0.0576)	(0.0917)	(0.0845)
l_LEV_w	-0.0474	0.0212*	-0.0240	0.00181	-0.00535	-0.0493	0.137	0.0587	0.0199
	(0.146)	(0.0119)	(0.0205)	(0.0348)	(0.0306)	(0.149)	(0.141)	(0.0662)	(0.0727)
l_CH_w	0.0431	0.105***	0.0967***	0.0260	-0.0714*	0.0428	0.0792	0.423***	0.381***
	(0.154)	(0.0339)	(0.0331)	(0.0402)	(0.0381)	(0.156)	(0.121)	(0.112)	(0.120)
l_COMBINED_RAM	0.00162	-0.00431	-0.0188**	-0.0275***	0.00806				
	(0.0819)	(0.00935)	(0.00732)	(0.00970)	(0.0118)				
l_RD_INT_w	-0.0326	-0.156***	-0.0558**	0.0241	-0.0474	-0.0443	-0.144***	-0.221	-0.174
	(0.0339)	(0.0430)	(0.0235)	(0.0206)	(0.0484)	(0.0367)	(0.0325)	(0.231)	(0.172)
l_AD_IND_INT_w	1.690	0.155	0.301	0.234	0.199	1.748	0.156	0.751	0.920

Continued on next page

Continued from previous page

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	f_EXTERNAL	f2_ABS_DA	f3_ABS_DA	f4_ABS_DA	f5_ABS_DA	f_EXTERNAL	f_RAM	f2_RAM	f3_RAM
	(1.358)	(0.275)	(0.339)	(0.334)	(0.533)	(1.366)	(1.842)	(1.035)	(1.254)
l_GOVERNANCE_w	0.0514***	-0.00111	-0.000511	-0.00151	-0.00784	0.0517***	0.0190**	0.0176**	0.0116
	(0.0158)	(0.00157)	(0.00181)	(0.00351)	(0.00518)	(0.0155)	(0.00831)	(0.00794)	(0.00792)
l_BIG4	0.150*	-0.00705	0.00372	-0.0153	-0.0271	0.148*	-0.00661	0.0877***	0.0387
	(0.0860)	(0.0104)	(0.00795)	(0.00991)	(0.0169)	(0.0866)	(0.0608)	(0.0295)	(0.0361)
l_FIRM_AGE_w	-0.126***	-0.0122	0.00188	0.00113	0.0115*	-0.125***	-0.0316	-0.0128	-0.0343*
	(0.0327)	(0.00860)	(0.0112)	(0.0123)	(0.00586)	(0.0330)	(0.0276)	(0.0303)	(0.0193)
l_ABS_DA						0.212	0.396***	0.236	-0.0337
						(0.163)	(0.129)	(0.169)	(0.147)
Constant	-0.00481	0.0456	0.0513	0.0998	0.0848*	0.00428	-0.0522	-0.191	0.0236
	(0.228)	(0.0433)	(0.0576)	(0.0631)	(0.0457)	(0.220)	(0.175)	(0.164)	(0.0921)
Observations	1,617	1,718	1,521	1,334	1,146	1,617	1,774	1,719	1,522
Adjusted R-squared	0.521	0.262	0.295	0.292	0.282	0.521	0.747	0.750	0.764
r clust sic2	YES	YES	YES	YES	YES	YES	YES	YES	YES
i.fyear	YES	YES	YES	YES	YES	YES	YES	YES	YES
i.gvkey	YES	YES	YES	YES	YES	YES	YES	YES	YES

Table 19: PSM Sample. Leads and lags model

VARIABLES	(1) EXTERNAL	(2) EXTERNAL	(3) EXTERNAL
aftertreat_sic2_-3	0.0160	0.0165	0.0114
	(0.0164)	(0.0164)	(0.0182)
aftertreat_sic2_-2	-0.0174	-0.0176	-0.0237
	(0.0250)	(0.0251)	(0.0277)
aftertreat_sic_-1	-0.0274	-0.0277	-0.0340
	(0.0266)	(0.0266)	(0.0294)
aftertreat_sic2_0	0.0170	0.0220	-0.0215
	(0.0372)	(0.0378)	(0.0346)
aftertreat_sic2_+1	-0.0595	-0.0667	
	(0.0409)	(0.0410)	
aftertreat_sic2_1+			0.0617
			(0.0470)
aftertreat_sic2_+2	0.109*		
	(0.0567)		
aftertreat_sic2_2+		0.139*	
		(0.0710)	
aftertreat_sic2_3+	0.257***		
	(0.0644)		
CONTROLS	YES	YES	YES
Observations	2,583	2,583	2,583
Adjusted R-squared	0.488	0.487	0.484
r clust sic2	YES	YES	YES
i.fyear	YES	YES	YES
i.gvkey	YES	YES	YES

*, **, *** Indicate statistical significance at the 0.1, 0.05, and 0.01 levels, respectively, based on a two-tailed test. $aftertreat\_sic2\_-3$ ($aftertreat\_sic2\_-2$ or $aftertreat\_sic2\_-1$) is a dummy equal to 1 for the treated industries three (two or one) years before the disaster and 0 otherwise. $aftertreat\_sic2\_+1$ ($aftertreat\_sic2\_+2$) is a dummy equal to 1 for the treated industries one (two) year(s) after the disaster and 0 otherwise. $aftertreat\_sic2\_1+$ ($aftertreat\_sic2\_2+$ or $aftertreat\_sic2\_3+$) is a dummy equal to 1 for the treated industries for all years after the disaster except the year of the disaster (the year of the disaster and one year after; the year of the disaster and two years after) and 0 otherwise. The rest of the variables are as described in Appendix A.