Received: 6 May 2017 | Revised: 16 July 2018 | Accepted: 25 July 2018 | Published on: 6 September 2018

DOI: 10.1002/smj.2935

WILEY STRATEGIC MANAGEMENT JOURNAL

RESEARCH ARTICLE

Winning us with trifles: Adverse selection in the use of philanthropy as insurance

Jiao Luo | Aseem Kaul | Haram Seo

Strategic Management & Entrepreneurship Department, Carlson School of Management, University of Minnesota, Minneapolis, Minnesota, 55455

Correspondence
Jiao Luo, Strategic Management & Entrepreneurship Department, Carlson School of Management, University of Minnesota, 321 19th Avenue South, Minneapolis, MN 55455.
Email: luoj@umn.edu

Research Summary: We study the use of corporate philanthropy as a form of reputation insurance, developing a formal model of such insurance to examine how the terms of insurance in equilibrium change under different assumptions about the firm and its stakeholders. We then test the predictions from this model in the U.S. petroleum industry and find that philanthropic donations offer insurance-like benefits, but are also positively associated with subsequent oil spills—firms that give more, spill more—with this association being stronger for spills that are under firms' control and in states with low civic capacity. These results are consistent with an adverse selection/moral hazard equilibrium and suggest that the use of philanthropy as reputation insurance may benefit firms at the cost of society.

Managerial Summary: Firms that donate to social causes develop a reputation for being socially responsible, and are often given the benefit of doubt when negative information about them comes to light. But are philanthropic firms truly more responsible? We argue that firms that donate more may be more likely to do harm—those that expect to do harm later are likely to give more now, and those that know their reputation protects them may become less careful. Evidence from the U.S. petroleum industry is consistent with this argument, with firms that give more having more subsequent oil spills, but only the type of spills that are under the firm's control, and only in states where the firm faces weaker scrutiny.

KEYWORDS

corporate philanthropy, formal model, information asymmetry, insurance, social impact

1 | INTRODUCTION

The study of Corporate Social Responsibility (CSR) is an important topic in the strategy literature, with a growing body of work arguing that firms can do well by doing good, that is, firms that share value with key stakeholders through the provision of positive externalities or the abatement of negative externalities (Dorobantu, Kaul, & Zelner, 2017; Porter & Kramer, 2006) may achieve superior financial performance (Barnett & Salomon, 2006, 2012; Flammer, 2015; Henisz, Dorobantu, & Nartey, 2014; Waddock & Graves, 1997). In particular, several recent studies have highlighted the role of corporate philanthropy[1] as an important form of CSR with the potential to yield financial benefits for firms (Cuypers, Koh, & Wang, 2015; Lev, Petrovits, & Radhakrishnan, 2010; Singh, Teng, & Netessine, 2017; Wang, Choi, & Li, 2008; Wang & Qian, 2011). A prominent argument in this literature is that philanthropy serves as a form of reputation insurance; firms that undertake philanthropy accrue a moral reputation, which helps protect them when adverse events occur or negative information about the firm comes to light, because stakeholders are more apt to give the benefit of the doubt to firms they judge to be responsible or trustworthy (Bansal & Roth, 2000; Godfrey, 2005). This insurance effect has emerged as one of the key mechanisms[2] linking CSR to superior financial performance (Barnett & Salomon, 2006, 2012; Jayachandran, Kalaignanam, & Eilert, 2013; Madsen & Rodgers, 2015), with several empirical studies[3] showing evidence for it (Barnett, Hartmann, & Salomon, 2017; Flammer, 2013; Godfrey, Merrill, & Hansen, 2009; Hong & Liskovich, 2017; Koh, Qian, & Wang, 2014; Minor & Morgan, 2011; Peloza, 2006; Shiu & Yang, 2017).

While this work clearly establishes the potential for philanthropy to provide insurance-like benefits, the central assumption driving this effect—that philanthropic firms are more responsible, and therefore worthy of society's trust—remains largely unexamined. This is a serious omission, given that problems of adverse selection and moral hazard are endemic to insurance markets (Pauly, 1974; Rothschild & Stiglitz, 1976), so that it may be the case that it is the less responsible firms that undertake philanthropy in order to offset the substantial harm they expect to do in the future (Brammer & Pavelin, 2005; Minor & Morgan, 2011). More generally, if "good deeds earn chits" (Godfrey, 2005, p. 777), then we need to know the terms at which these chits are exchanged in order to assess whether investing in these chits is a sound strategy, that is, whether the benefits of doing so exceed the costs (Bromiley & Marcus, 1989). On one hand, if the price firms pay for the coverage they receive from their philanthropic activities is too high, then philanthropy may, in fact, be value destroying for shareholders, with the benefit of lower penalties when accidents happen being offset by the cost of philanthropy when they don't. On the other hand, if the price firms pay for the protection they receive for their philanthropic acts is too low, then corporate philanthropy may benefit firms at the cost of social welfare, with the damage that firms are forgiven when accidents happen being greater than the philanthropy they do to earn that forgiveness. In the absence of a better understanding of the terms at which the reputation insurance associated with corporate philanthropy is offered, it is therefore impossible to determine its consequences for either firms or society.

[1]We recognize that CSR includes both corporate philanthropic giving and more integrated forms of social activities such as pollution abatement or the adoption of sustainable technologies and practices. For the purposes of this article, however, we focus primarily on corporate philanthropy so as to be consistent with Godfrey's original (2005) work.

[2]There are, of course, many other ways in which philanthropy may drive superior financial performance. While we focus largely on the insurance mechanism in this article, we discuss spillovers from these other mechanisms in our discussion on exogenous philanthropy later in the paper.

[3]A few recent studies question the significance and sustainability of this benefit (Kang, Germann, & Grewal, 2016; Shiu & Yang, 2017).

This study attempts to address this lacuna. We develop a formal model of corporate philanthropy as reputation insurance, building off classic economic models of insurance markets under information asymmetry (Pauly, 1974; Rothschild & Stiglitz, 1976). Specifically, we model a "market", where firms buy reputation insurance from society, that is, they undertake philanthropy in exchange for the promise of lower penalties should they have an accident that is harmful to society. We assume that firms differ in their probability of having an accident—specifically, that there are "clean" firms and "dirty" firms, where the latter are more likely to have an accident than the former—and that this probability is known to the firms, but may or may not be known to society. We then derive the equilibrium price and quantity of insurance exchanged in this market for philanthropy as insurance, and examine how this equilibrium changes depending on whether (a) society is informed or uninformed, (b) society is able or unable to collectively organize, and (c) firms undertake philanthropy exclusively for insurance purposes or give for other reasons as well.[4] The model generates a set of alternate equilibria based on these assumptions, with each equilibrium mapping to different predictions about the relationships among philanthropy, accident probability, and accident penalty, and the benefits of philanthropy for both firms and society.

We test the predictions from our model in the context of the U.S. petroleum industry in order to see which of the alternate equilibria holds in practice. We find that philanthropy weakens the negative relation between oil spill amounts and stock market reactions, consistent with prior work on philanthropy as a form of insurance. However, we also find a positive association between a firm's philanthropic donations and the subsequent amount and number of its oil spills: firms that give more, spill more. These observed associations are consistent with an adverse selection/moral hazard equilibrium, where less socially responsible firms undertake more philanthropy so as to offset the negative consequences of doing greater damage. Moreover, supplementary analyses show that the positive association between donations and future spills is stronger for in-state donations than for out-of-state donations, for spills that are under firms' control than for spills related to natural disasters or adverse weather conditions, and in states with low civic capacity than in those with high civic capacity, lending further support to the adverse selection/moral hazard case.

Our study contributes to our understanding of corporate philanthropy, and its insurance-like benefits, in several ways. First, our use of a formal model allows us to consider not only the firm's demand for such insurance coverage (Godfrey, 2005), but also the supply of insurance by society and the terms on which it is made available, making room for strategic action on the part of society as well as the firm, and deriving the equilibrium price and quantity of insurance exchanged. In doing so, our study not only responds to the call for further exploration of the contingencies underlying the insurance-like benefits of philanthropy (Barnett et al., 2017), it also joins a stream of recent work that seeks to offer a more complete and rigorous theoretical account of nonmarket strategic actions through the use of formal models (Baron, 2001; Chatain & Plaksenkova, 2018; Fosfuri, Giarratana, & Roca, 2016; Kaul & Luo, 2018).

Second, our study extends the literature on philanthropy as reputation insurance by going beyond this work's traditional focus on the benefits of such insurance for firms (Godfrey et al., 2009; Koh et al., 2014; Minor & Morgan, 2011), to simultaneously examine its implications for social welfare (Ballesteros, Useem, & Wry, 2017; Barnett, 2016; Kaul & Luo, 2018). Theoretically, our use of a formal model allows us to map underlying assumptions to equilibrium outcomes, thus defining the

[4]As we discuss later, firms may give for ethical reasons (Freeman, 1984; Maitland, 1994), in response to institutional pressures (Aguilera, Rupp, Williams, & Ganapathi, 2007; Marquis, Glynn, & Davis, 2007; Marquis & Tilcsik, 2016), or because philanthropy benefits them in other ways (Bénabou & Tirole, 2010; Dorobantu et al., 2017).

conditions under which the use of philanthropy as a form of insurance is potentially beneficial for firms, for society, or for both. Empirically, we show that philanthropic donations not only moderate the stock market's negative reaction to an accident, but are also positively associated with subsequent accidents. Our study thus highlights the potential for adverse selection and moral hazard in the use of philanthropy as a form of reputation insurance (King & Lenox, 2000), raising serious questions about the claim that such philanthropy allows firms to do well while doing good, and suggesting that, at least in some cases, the use of philanthropy as reputation insurance may be a means for firms to benefit their shareholders at the cost of society (Akerlof & Shiller, 2015).

2 | THEORY

2.1 | Model assumptions and setup

As previously discussed, our theoretical approach in this article is to develop a formal model of the use of philanthropy as a form of reputation insurance. Specifically, we model a "market"[5] for such insurance, where firms compete to establish relational contracts with key stakeholders, whereby they undertake philanthropy today on the understanding that doing so will earn them forgiveness for negative events in the future. Our goal in doing so is to derive the terms of exchange for philanthropy as reputation insurance, that is, the extent of reputation insurance the firm will receive per unit of corporate philanthropy, in an equilibrium where both firms and stakeholders are satisfied. Our use of a formal model allows us to rigorously evaluate the net benefit of the use of philanthropy as insurance for both firms and society, allowing for strategic action on the part of both, and clearly laying out the assumptions underlying our results. Moreover, it allows us to conceptually examine the effect of changing these assumptions, and to develop alternate predictions regarding the equilibrium price and quantity of reputation insurance under these distinct assumptions.

Our model and subsequent empirical analysis deal specifically with corporate philanthropy, consistent with prior work (Godfrey, 2005). While other forms of CSR could also offer similar insurance-like benefits (Koh et al., 2014; Minor & Morgan, 2011), focusing on philanthropy offers several analytical advantages, both theoretically and empirically. For one thing, the cost of philanthropy to the firm is relatively easy to model and observe—being simply the dollar amount of philanthropy in the case of cash donations—compared to the costs of other forms of CSR activities, which involve changes to the firm's operating activities. Relatedly, because corporate philanthropy is largely independent of the firm's business operations (Kaul & Luo, 2018), the extent of a firm's giving is less likely to directly impact the probability of an accident in its operations. Conceptually, this means that we can model the insurance benefits of philanthropy without either having to specify a "production function" linking the benefit received by society to firm costs, or having to account for changes in accident probabilities as a result of CSR. Empirically, it means that we can more cleanly estimate the cost of reputation insurance to the firm, without having to account for operational changes as a potential omitted variable. Nevertheless, we believe the general logic of our arguments would apply equally to other forms of CSR as to philanthropy.

To develop our model, we begin with Godfrey's work on the optimal level of philanthropy (Godfrey, 2005). We build on this work by considering the supply-side of insurance (i.e., the terms

[5]This is not to suggest that there is a real market where insurance contracts between firms and society are actively traded; our use of a market concept is simply to help us analytically arrive at the terms of exchange between firms and society. We discuss how this exchange may depart from competitive market conditions in more detail in the following.

on which insurance is made available to the firm) and by incorporating the possibility that the likelihood of an accident may be endogenous to the firm. To do so, we draw heavily on prior work on insurance markets with imperfect information; specifically, we draw on models of insurance developed by Pauly (1974) and Rothschild and Stiglitz (1976) in that we adopt a similar conceptualization of insurance contracts and use a similar solution concept. We also draw on these economic models in that we assume that firms are risk neutral (see also Bromiley & Marcus, 1989), but individuals are risk averse, and model the utility of individuals in the same way as this prior work. The key difference between these insurance models and ours is that they are concerned with multiple individuals entering into a formal insurance contract with a single provider (firm) to offset potential costs of negative events in the future, while we are concerned with multiple firms entering into a relational contract (as described later) with society for the same purpose. Given this difference, we additionally assume that society's risk preferences are a simple aggregation of those of individuals, that is, society as a whole is risk-averse as well.[6]

Consider a profit-maximizing, risk-neutral firm that faces the hazard of having an accident that generates a negative externality for society. In the event of an accident, the firm is required to compensate society for this damage, that is, it is required to pay F to society.[7] This payment could take the form of direct reparations to those harmed, either voluntary or involuntary (e.g., court-mandated), or of fines or penalties levied by the state on society's behalf, or of additional expenditures the firm must make to rehabilitate its reputation after the accident and maintain its competitive position. Whatever the form of the compensation, the firm faces an additional cost F in case of an accident, where the probability of an accident occurring is $p \in (0, 1)$. We assume that F is exogenously determined—that is, the firm has no control over the damage it may cause—but that it can partially influence the probability of an accident p by undertaking some prevention measures. Specifically, we assume that the probability of the firm experiencing an accident is a function $p(m)$, where m is the amount the firm spends on internal actions to lower its accident risk—including safety training, preventive maintenance, and investments in better equipment (Kölbel, Busch, & Jancso, 2017)—such that $p'(m) < 0$ and $p''(m) > 0$ (Pauly, 1974). We define $p^0 = p(0)$ as the probability of accident when $m = 0$, that is, if the firm makes no attempt at prevention.[8]

We further assume that there are two types of firms, C and D, which are identical in all other respects except that type D (dirty) firms are more likely[9] to have accidents than type C (clean) firms, that is, $p_c(m) < p_d(m)$ for all m. This difference in type may be the result of several factors. It may reflect age and quality of the firm's equipment, the types of technologies it uses, and the experience and expertise of its personnel. It may also reflect differences in the conditions in which it operates, with firms that operate in more challenging contexts having a higher likelihood of accidents. In addition, firms experimenting with newer, more risky technologies and processes may have higher

[6]Strictly speaking, our analysis requires only that society be relatively risk-averse compared to firms; if that were not the case, there would be no basis for society to exchange a large but uncertain payoff in the future for a small but certain payoff today, so philanthropy as reputation insurance would not work. It may be the case that society is less risk-averse than individuals; so long as it is still somewhat risk-averse and firms are risk-neutral, our analysis is unaffected. Our main conclusions would also hold if society were risk-seeking, but firms were even more risk-seeking than society.

[7]This compensation may or may not be equal to the damage caused by the accident; it is simply the amount the firm must pay society.

[8]We assume that the net effect of prevention is to increase cost. If prevention were associated with increased profits through other means (King & Lenox, 2001), then firms would rationally maximize prevention independent of insurance costs, so that, for the purposes of our analysis, the probability of accident would be effectively exogenous. In other words, any profit-increasing prevention would already be reflected in p^0.

[9]For simplicity, we vary only the probability of an accident and not the extent of damage caused, that is, $F_c = F_d = F$. Varying only the probability is sufficient to have the expected damages for dirty firms be higher than for clean firms, which is what we are trying to model.

accident probabilities, at least in the short run. Whatever the underlying cause, there are some firms that have a higher accident probability. The proportion of such firms in the total population is θ, where $0 \leq \theta \leq 1$. Each firm knows its own type with certainty, but society may or may not be able to tell each firm's type.[10] This may be because several of the factors that drive firm type—the firm's equipment, capabilities, and so on—may be hard for outsiders to observe. In addition, firms would also naturally have a superior understanding of, and expertise in, their area of operations, so they may be better able to assess the probability of an accident, given the same information.

In addition to lowering the probability of an accident by undertaking prevention expenditure, the firm can also try to shield itself from the cost of an accident by investing in philanthropy as a form of reputation insurance. More specifically, a firm can enter into a relational contract $\alpha = (S, W)$ with society whereby it makes an ex ante donation of $S \geq 0$ in exchange for paying a total amount of W if an accident occurs (Rothschild & Stiglitz, 1976), where $S \leq W \leq F$. In other words, the firm pays S irrespective of whether an accident occurs or not, but pays an additional penalty of $W - S$ if an accident occurs, instead of paying a penalty of F. We may thus think of $W - S$ as the firm's "co-pay" in case of an accident. Where $W = S$ the firm enjoys full insurance; effectively, it pays S in each period, and nothing additional if an accident occurs. The contract α thus implies an understanding that should an accident occur, society will forgive the firm an amount R of the total damages F, where $R = F + S - W$, with $\frac{S}{R}$ being the price of reputation insurance, that is, the amount of philanthropy the firm must undertake per unit of insurance it receives. Note that, in our case, α is a relational contract (Baker, Gibbons, & Murphy, 2002; Poppo & Zenger, 2002) in that it is not formally defined or enforceable, but relies on the mutual self-interested cooperation of both parties. In fact, while we speak of a "market" for reputation insurance, with firms as buyers and society as the seller (Godfrey, 2005), it may be more accurate to think of firms undertaking philanthropy as conditional cooperators (Frey & Meier, 2004; Ostrom, 2005), proactively benefiting society in the expectation of being rewarded for doing so later (Bosse, Phillips, & Harrison, 2009; Dorobantu et al., 2017).

In order to determine the equilibrium level of philanthropy as reputation insurance, we follow Pauly (1974) and adopt a two-stage approach, with the firm determining the optimal amount of prevention to undertake (and therefore, the probability of an accident) in the first stage, and then the optimal amount of philanthropy in the second, based on the insurance contract made available to it by society. We derive these equilibrium conditions using primarily a verbal or graphical approach in our main article, though a fuller mathematical treatment may be found in Appendix S1.

Consider a firm of type $i = \{c, d\}$ trying to determine the optimal amount of prevention to undertake. In general, the firm will choose an amount of prevention such that the marginal decrease in the expected cost of an accident from an additional unit of prevention is just equal to the cost of that additional unit of prevention. In other words, in the absence of reputation insurance, the firm would choose an amount of prevention m_i^* such that $-\frac{\partial\left(p_i^* F\right)}{\partial m} = 1$, where $p_i^* = p_i\left(m_i^*\right)$ is the probability of an accident at this optimal level of prevention.

Now consider the case where insurance is a possibility. The firm then chooses an amount of prevention $\dot{m}_i$ such that $\frac{\partial \dot{p}_i}{\partial m_i} = -\frac{1}{W_i - S_i}$ (see Appendix S1 for a derivation of this result), where $\dot{p}_i = p_i(\dot{m}_i)$. We further assume that $\dot{m}_i = 0$ if $W_i = S_i$, implying that with full insurance the firm would prefer to

[10]We assume that this is not only true ex ante, but may also be true ex post, at least in the short run, that is, even after society observes an accident it cannot tell which type of firm it's dealing with. The intuition for this result is that even "clean" firms will have some probability of having an accident, so that observing a single accident may not help distinguish firm type. Assessing the true probability from observed events is cognitively challenging (Barnett, 2014) and may require many periods of observation, even if this probability were unchanging, and the challenge may be even greater if the true probability itself were evolving over time.

undertake no prevention, and therefore, $\dot{p}_i = p_i^0$. Note that $F > W_i - S_i \Rightarrow \dot{m}_i < m_i^*$, meaning that the profit-maximizing firm always chooses a lower level of prevention (and therefore, a higher probability of accident) when it can invest in philanthropy as reputation insurance. The intuition is that a firm has less incentive to try and prevent accidents if it can protect itself by undertaking philanthropy. In particular, in the full insurance case where the firm pays no additional penalty in case of an accident (i.e., $W_i = S_i$), it has no incentive to undertake prevention. This decrease in prevention expenditure (and consequent increase in accident probability) reflects the problem of moral hazard (Pauly, 1974) in the use of philanthropy as reputation insurance.

Having determined the optimal amount of prevention for a given insurance contract (note that $\dot{m}_i$ is a function of both W_i and S_i), we now turn to the second stage where the firm and society mutually determine the optimal insurance contract between them. To do so, consider, first, the value of philanthropy as reputation insurance to the firm. For a firm that enters into contract α_i, the expected cost[11] from an accident is $Z(\alpha_i, m_i) = p_i(m_i)W_i + (1 - p_i(m_i))S_i + m_i$. The value of entering into the reputation insurance contract to the firm is then the difference between this expected cost with or without the contract (i.e., with or without philanthropy), and is given by:

$$\pi(\alpha_i) = Z\left(0, m_i^*\right) - Z(\alpha_i, \dot{m}_i) = p_i^* F + m_i^* - \dot{p}_i W_i - (1 - \dot{p}_i) S_i - \dot{m}_i. \quad (1)$$

Clearly, a profit-maximizing firm will seek to maximize $\pi(\alpha_i)$ when choosing a contract, and will be willing to accept any contract where $\pi(\alpha_i) \geq 0$. A risk-neutral, profit-maximizing firm will only pursue philanthropy as a form of reputation insurance if the expected payoff from doing so is nonnegative, that is, if it believes that the benefits it receives in the form of forgiven damages in case an accident occurs are at least equal to the amount it pays out in philanthropy.[12]

Next, consider the benefits of such insurance to society, and the consequent supply of reputation insurance. We assume that society's preferences[13] are defined by a utility of money income U such that $U' > 0$ and $U'' < 0$ (Rothschild & Stiglitz, 1976). If the firm does not undertake philanthropy, then society receives a payment F with probability p_i, and its expected utility is $p_i U(F)$. If the firm does enter into the relational contract, then expected utility is $p_i U(W_i) + (1 - p_i)U(S_i)$. The value to society is the difference between the two expected utilities:

$$V(\alpha_i) = (1 - p_i)U(S_i) - p_i[U(F) - U(W_i)]. \quad (2)$$

The use of philanthropy as reputation insurance only benefits society if $V(\alpha_i) > 0$, that is, if the expected utility it receives from philanthropic payments in the absence of an accident ($(1 - p_i)U(S_i)$) is greater than the loss of utility from the reduction in payments in the event of an accident ($p_i[U(F) - U(W_i)]$). If $V(\alpha_i) = 0$, the use of philanthropy as reputation insurance leaves social welfare unaffected, and if $V(\alpha_i) < 0$, then society is worse off as a result of philanthropy.

The range of contracts acceptable to both firms and society may be represented using a figure that graphs the amount the firm pays society in the absence of an accident (S) on the horizontal axis and the amount it pays in the event of an accident (W) on the vertical axis as shown in Figure 1. Every

[11]Throughout this analysis, we assume away the issue of time value of money. This is equivalent to assuming that F and R_i represent the (discounted) present value of future cash flows.

[12]In the special case where the probability of an accident is entirely exogenous, the benefit to the firm is $\pi(\alpha_i) = p_i^0 F - p_i^0 W_i - \left(1 - p_i^0\right) S_i = p_i^0 R_i - S_i$, and the firm will accept any insurance contract where $p_i^0 R_i \geq S_i$.

[13]We can think of this social utility function as reflecting the aggregation of individual preferences. Since the accident results in a negative externality, that is, the effects of the accident are assumed to be nonexcludable and nonrivalrous, and perceived as harmful by all members of society, society must respond to the hazard collectively as a unit, and there is no possibility of risk-spreading across individuals. Specifically, as previously discussed, our assumption is that society's risk preferences have the same form as that of an individual.

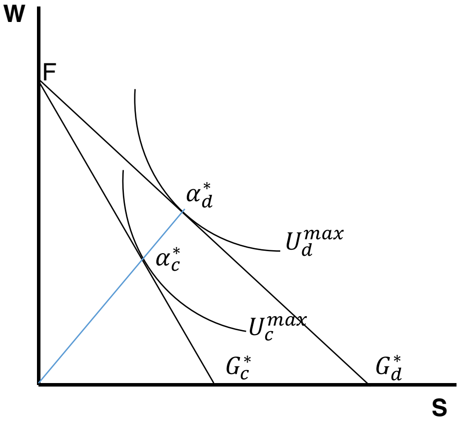

FIGURE 1 Separating equilibrium with active and fully informed society. Society offers contracts α_c^* and α_d^* to clean and dirty firms, respectively. Firms receive full insurance ($S_i^* = W_i^* = p_i^* F$), with dirty firms undertaking more philanthropy, and therefore, paying a higher price for insurance. Firms make no profit from philanthropy as reputation insurance, but social welfare is enhanced

point on the graph thus corresponds to an insurance contract. The point F on the vertical axis represents the initial state, that is, the case where the firm does not undertake philanthropy and pays the full amount of compensation if the accident occurs. The line FG_i^* then represents the set of contracts where $\pi(\alpha_i) = 0$, that is, the set of contracts that leave the firm indifferent. This line may be thought of as the budget line for the firm looking to buy reputation insurance since all contracts above and to the right of the line would result in a net loss to the firm, and will therefore be unacceptable to it. We can show quite easily that the slope of this line is $(1 - p_i(m_i))/p_i(m_i)$, that is, it is equal to the odds ratio of not having an accident to having an accident. Further, it follows from $p_c(m) < p_d(m)$ that $(1 - p_c(m))/p_c(m) > (1 - p_d(m))/p_d(m)$, that is, the slope of the line for clean firms will be steeper than that for dirty firms, at least for equivalent levels of prevention spending by both types of firms. Similarly, society's utility function can be represented as a set of indifference curves that are level sets of $p_i U(W_i) + (1 - p_i)U(S_i)$.

Having described the setup of the model, we now turn to derive the equilibrium conditions under a range of different assumptions about the firms and society. Note that in doing so, we deliberately choose to focus on the extreme cases. In practice, of course, a range of intermediate positions are possible. By focusing on the extreme cases, we seek to model and describe the relevant range of potential outcomes (MacDonald & Ryall, 2004), while also keeping the model parsimonious. Introducing additional parameters—such as the extent of society's information or its relative bargaining power—to allow for intermediate states would only complicate the model without altering the basic logic of its arguments or its core conclusions.

2.2 | Alternate equilibria and empirical predictions

As previously mentioned, we use the model to derive a set of alternate equilibria based on different assumptions about firms and society. In order to conform to journal length requirements, we briefly outline the assumptions and logic behind the alternate equilibria in the main text,[14] but provide a more detailed discussion of each equilibrium in Appendix S2.

[14]While this section sketches out the logic for the different equilibria in broad terms, we strongly recommend that interested readers refer to Appendix S2 for a full development of the arguments underlying each equilibrium.

We start by considering the case where society is fully informed, that is, it can distinguish clean firms from dirty firms, and can observe the extent of prevention the firm undertakes. As Figure 1 shows, this case yields a separating equilibrium with society offering (full insurance) contracts α_c^* and α_d^* to clean and dirty firms, respectively, at the point where each firm's budget line becomes tangent to the corresponding indifference curve. Under these contracts, dirty firms undertake more philanthropy than clean firms, and therefore (given that with a full insurance contract $S_d^* = W_d^*$), pay higher penalties in the event of an accident. In other words, dirty firms pay a higher price for reputation insurance, which is only logical, given that they are more likely to have an accident, and society is aware of this. Since the equilibrium contracts lie on the firms' budget lines, firms make no profit from reputation insurance in this case, but social welfare is enhanced, because society has moved to a higher indifference curve than the one through F.

The financial and welfare consequences of the equilibrium previously described are critically dependent on the society being able to observe firm prevention, and therefore, to ensure that firms do not slack off in their prevention efforts. A slight variant of this equilibrium arises where society can tell dirty firms from clean firms but cannot observe their prevention efforts. In such a case, society would still offer the contracts α_c^* and α_d^*, but since these are full insurance contracts, firms would cut their (externally unobservable) prevention expenditures to zero. As a result, firms would now make some profit from reputation insurance by cutting back on prevention (specifically, they would make profits equal to m_i^*). Such cuts in maintenance would increase the probability of accident, however, reducing social welfare, though probably not enough to make reputation insurance harmful to society (see Appendix S2, especially Figure A1, for a discussion of this point).

If society is uninformed, that is, it cannot tell the difference between clean and dirty firms, then we have the pooling equilibrium shown in Figure 2. In this case, the best society can do is to offer the optimal contract for clean firms α_c^* to all firms, allowing dirty firms to mingle with clean firms. Both types of firms donate equal amounts and pay equal penalties in this case, with dirty firms benefiting substantially from being mixed in with clean firms, though clean firms also make some profit by cutting back on prevention expenses (in light of full insurance). As discussed at greater length in Appendix S2, however, society will only offer this insurance contract if it believes the

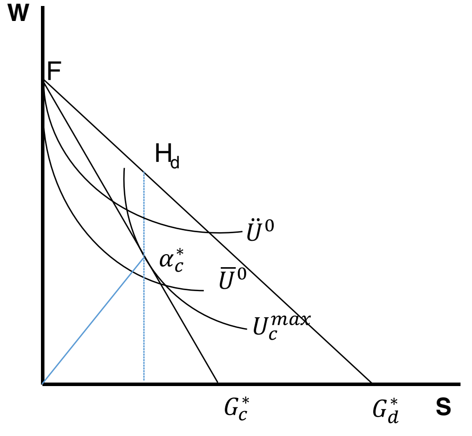

FIGURE 2 Pooling equilibrium with active but uninformed society. Dirty and clean firms undertake equal philanthropy and pay equal penalties in the event of an accident. Both firms make profit from moral hazard, with dirty firms realizing additional gains proportional to $H_d - W_c^*$ from being mixed in with clean firms. Equilibrium exists so long as α_c^* lies above $\overline{U}^0$, but is only welfare-enhancing if it lies above $\ddot{U}^0$

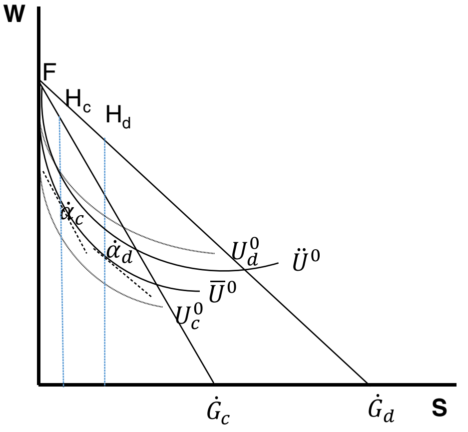

FIGURE 3 Separating equilibrium with passive society. Dirty firms undertake more philanthropy ($\dot{S}_d > \dot{S}_c$) and pay lower penalties in the event of an accident ($\dot{W}_d < \dot{W}_c \Rightarrow (\dot{R}_d > \dot{R}_c)$). Both firm types make profit from philanthropy as reputation insurance, with dirty firms benefiting more than clean firms. Social welfare declines since both $\dot{\alpha}_c$ and $\dot{\alpha}_d$ lie below $\ddot{U}^0$

proportion of dirty firms is relatively small; otherwise, it may not offer reputation insurance at all. Even if reputation insurance is offered, however, the welfare consequences are ambiguous at best, because of the increased accident probabilities associated with moral hazard, as well as the mixing of dirty firms with clean ones.

While the first two scenarios assume that society is actively seeking to maximize welfare, leveraging its superior bargaining power (there are many firms but only one society) to dictate the best contract terms possible, our third scenario considers the case where such collective action fails, that is, where social stakeholders are unable to organize in their own self-interest. In particular, we consider the case where society is passive—willing to accept any insurance contract it sees as not reducing welfare—and firms are free to choose the terms of their own insurance contracts. In this case, clean and dirty firms choose contracts $\dot{\alpha}_c$ and $\dot{\alpha}_d$, respectively, corresponding to the points where the slope of the societal indifference curve through point F just equals the firms' odds ratio of having an accident, as shown in Figure 3. In this equilibrium, dirty firms undertake more philanthropy than clean firms, but also pay lower penalties, that is, this is a case with adverse selection. Profits for both clean and dirty firms are maximized in this case, but social welfare unambiguously declines as moral hazard leads to lower prevention and therefore a higher probability of accidents.

Finally, Figure 4 considers the case where society is active but uninformed (as in the second case), but firms undertake philanthropy for reasons exogenous to reputation insurance.[15] In particular, we assume that clean firms exogenously undertake an amount of philanthropy S^A in excess of dirty firms. This causes the budget line for clean firms to shift outward, as shown in Figure 4, and enables a separating equilibrium where society offers contracts α_c^{**} and α_d^{*} to clean and dirty firms, respectively. In this case, clean firms do more philanthropy but pay lower penalties in the case of an accident (once exogenous philanthropy is accounted for). Both firms make some profit from lower prevention, with clean firms potentially making additional profit for their philanthropy from other sources. Social welfare is enhanced, however, with society moving to a higher indifference curve,

[15]These may include ethical or moral concerns (Freeman, 1984; Maitland, 1994), institutional pressures (Marquis et al., 2007; Marquis & Tilcsik, 2016), managerial preferences (Chin, Hambrick, & Trevino, 2013; Marquis & Lee, 2013), or the prospect of being rewarded for philanthropy by other stakeholders, such as customers (Kaul & Luo, 2018; Lev et al., 2010) or employees (Bode, Singh, & Rogan, 2015; Burbano, 2016; Flammer & Luo, 2017).

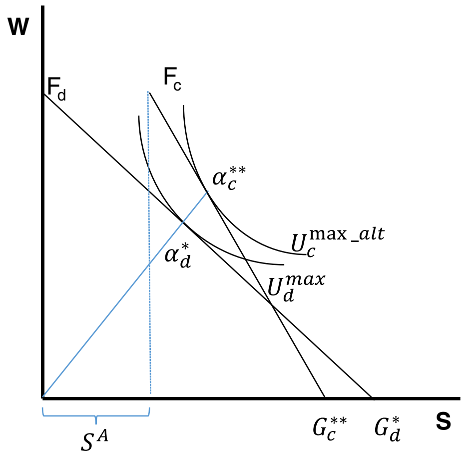

FIGURE 4 Separating equilibrium with exogenous philanthropy. Clean firms undertake more philanthropy and pay lower penalties (net of exogenous giving amount S^A) in the event of an accident. Both firm types profit from philanthropy as reputation insurance and social welfare may also be enhanced

and the exogenous philanthropy of clean firms forcing dirty firms to up their game, even though society cannot tell the two firm types apart.

These four alternate equilibria are summarized in Figure 5, which describes the equilibrium conditions corresponding to the cases where (I) stakeholders are active and fully (or partially) informed, (II) stakeholders are active but uninformed, (III) stakeholders are passive, and (IV) philanthropy is

	Equilibrium I	**Equilibrium II**[a]	**Equilibrium III**	**Equilibrium IV**
Assumption about society	Active and fully (partially) informed	Active but uninformed	Passive and uninformed	Active but uninformed
Assumption about firms	Strategic	Strategic	Strategic	Clean firms undertake exogenous philanthropy
Clean firm benefit[b]	Zero or Low $\pi(\alpha_c^*, m_c^*)$ or $\pi(\alpha_c^*, 0)$	Low $\pi(\alpha_c^*, 0)$	High $\pi(\dot{\alpha}_c, \dot{m}_c)$	Low $\pi(\alpha_c^*, 0)$
Dirty firm benefit	Zero or Low $\pi(\alpha_d^*, m_d^*)$ or $\pi(\alpha_d^*, 0)$	Moderate $\pi(\alpha_c^*, \dot{m}_d)$	High $\pi(\dot{\alpha}_d, \dot{m}_d)$	Low $\pi(\alpha_d^*, 0)$
Social welfare[c]	High or Moderate $(1-\theta)V(\alpha_c^*, m_c^*) + \theta V(\alpha_d^*, m_d^*)$ or $(1-\theta)V(\alpha_c^*, 0) + \theta V(\alpha_d^*, 0)$	Ambiguous $(1-\theta)V(\alpha_c^*, 0) + \theta V(\alpha_c^*, \dot{m}_d)$	Negative $(1-\theta)V(\dot{\alpha}_c, \dot{m}_c) + \theta V(\dot{\alpha}_d, \dot{m}_d)$	Moderate $(1-\theta)V(\alpha_c^*, 0) + \theta V(\alpha_d^*, 0)$
Relationship between philanthropy and accident probability	Positive	None	Positive	Negative
Relationship between philanthropy and accident penalty	Positive	None	Negative	Negative

a. Equilibrium may not exist
b. $0 = \pi(\alpha_i^*, m_i^*) < \pi(\alpha_i^*, 0) < \pi(\dot{\alpha}_i, \dot{m}_i)$
c. $V(\alpha_i^*, m_i^*) > V(\alpha_i^*, 0) > V(\dot{\alpha}_i, \dot{m}_i)$ and $V(\dot{\alpha}_i, \dot{m}_i) < 0$

FIGURE 5 Summary of equilibria

partly exogenous, and summarizes the implications of these assumptions for both firm profitability and social welfare. Figure 5 shows that the use of philanthropy as a form of reputation insurance is most likely to benefit society where social stakeholders are active, aggressively demanding better terms from firms, and that such stakeholder activism is most likely to enhance social welfare where either stakeholders are well-informed or firms undertake philanthropy exogenously. Where stakeholders are relatively passive, profit-maximizing firms may take advantage of this to boost their own profits while leaving society worse off. Conversely, the figure also suggests that the potential for additional profit from the use of philanthropy as a form of reputation insurance comes primarily from firms taking advantage of the information asymmetry between themselves and stakeholders—either by cutting back on prevention expenses once they are protected by reputation insurance (the moral hazard problem) or by strategically hiding their true type (the adverse selection problem)—with this increase in profits potentially benefiting irresponsible (dirty) firms more than responsible (clean) ones. Where stakeholders are active and fully informed, firms can still use philanthropy as a form of reputation insurance, but are unlikely to see much financial benefit from doing so. Thus, the mere fact that philanthropy serves as a form of insurance (Godfrey, 2005) does not, by itself, ensure that it is beneficial to either firms or society. Overall, Figure 5 thus raises serious questions about the theoretical validity of the claim that the use of philanthropy as a form of reputation insurance is a means for firms to do well while doing good. In fact, it suggests that this claim is only likely to be true when philanthropy is undertaken exogenously, that is, it is not intended to serve as a form of insurance (Equilibrium IV).

In addition to summarizing the assumptions and welfare implications of each equilibrium, Figure 5 also offers two distinct sets of empirical predictions to help tease apart these various equilibria empirically. One set of predictions focuses on the relationship between the amount of philanthropy a firm undertakes and the penalty it subsequently pays in case of an accident, with Equilibrium I predicting a positive relationship, while Equilibria III and IV predict a negative relationship. A second set of predictions focuses on the relationship between the amount of philanthropy and the likelihood of an accident, with Equilibria I and III predicting a positive relationship, while Equilibrium IV predicts a negative relationship. In both cases, Equilibrium II predicts an insignificant relationship, given that it predicts that all types of firms will donate equal amounts. These empirical predictions thus provide a way to tease apart the alternate equilibria and identify, in any given context, the prevailing equilibrium. While we have developed these predictions using comparative statics to compare across firms, the same predictions would apply within firms over time if the probability of an accident changed exogenously. So, for instance, a firm in Equilibrium III that experienced an exogenous increase in its accident probability would increase its philanthropic giving, and thus reduce the penalty it would pay in case of an accident.

Note that Equilibria I and II are inconsistent with the preponderance of existing empirical evidence which shows that firms that undertake more philanthropy face lighter penalties in the aftermath of an accident (Godfrey et al., 2009; Hong & Liskovich, 2017; Koh et al., 2014; Minor & Morgan, 2011). This prior work does not, however, tell us whether the firms that receive lighter penalties are genuinely more responsible firms that are using philanthropy to signal their true type to active social stakeholders (Equilibrium IV), or dirty firms that are investing in philanthropy so they can cut back on prevention expenses and get away with doing greater damage (Equilibrium III). While the former would predict a negative relationship between philanthropy and accident probabilities, the latter would predict a positive relationship, which is what we turn to consider next.

3 | DATA AND METHODOLOGY

3.1 | U.S. oil industry

We test the empirical predictions in Figure 5 using oil spills in the U.S. petroleum industry from 2004 to 2011. Several factors make this a good context for our purpose. First, oil spills are a frequent occurrence in the United States—our sample period saw an average of 883 spills a year, involving a total of 1.10 million barrels of oil spilled every year—ensuring sufficient variance for our empirical tests in a single industry in a single country. Relatedly, U.S. oil companies are also among the most actively engaged in corporate philanthropy, with the firms in our sample donating a total of $105 million annually during our study period. Second, reporting on spills is mandatory, with federal law requiring companies to immediately report all releases of oil and hazardous substances to the National Response Center (NRC), and failure to do so carrying fines of up to $250,000 for individuals and $500,000 for organizations as well as a maximum prison sentence of 15 years (Environmental Protection Agency, 1999). This means that we can be relatively confident that we observe all negative occurrences, not only those where the firm happens to get caught (Bromiley & Marcus, 1989). Third, oil spills involve substantial social and economic costs to society, with these damages estimated to average between $80 and $985 per gallon, which equals $3,360–$41,370 per barrel (Etkin, 2004, p. 10),[16] making the setting extremely relevant for an analysis of the welfare consequences of corporate philanthropy as reputation insurance. Finally, the negative externalities generated by oil spills are, in most cases, locally bounded—as, previous studies suggest, is much of corporate philanthropy (Galaskiewicz, 1997; Marquis et al., 2007)—allowing us to examine the relationship between philanthropy and oil spills at the local (state) level.

The goal of our empirical analysis is to demonstrate how the empirical predictions in Figure 5 may be applied to a real-world context to tease apart the four alternate equilibria. We are not claiming that our chosen industry is representative of all others, or that the relationships we find in this industry will apply universally; on the contrary, we think it is entirely plausible that the equilibrium conditions may be different in different industries. That is why we have offered no hypotheses about which equilibrium we expect to see supported.

3.2 | Data sources

Our sample of oil spills in the United States is constructed using data from the National Response Center (NRC), which represents the most comprehensive source of information on oil spills available. The data include information about the time, location, and severity (i.e., amount spilled) of the incident as well as the name of the responsible firm. The data also include information on the weather conditions at the time of the spill as well as the proximate cause of the spill as reported by the firm (both of which we will use in supplementary analyses to follow).

Consistent with the sampling strategy in Helfat (1994, 1997), we keep our analysis tractable by applying a minimum size threshold to all firms in the oil and gas sample of the National Establishment Time-Series (NETS) Database, which provides establishment-level data on firm operations. Thus, we include firms that account for over 1% of total U.S. sales in 2003 (i.e., at the start of our study) and then restrict our sample to those that have at least one spill during our study period, resulting in a final sample of 35 oil firms that together account for 56% of the total assets and 65% of the

[16]The social and environmental cost of spills is determined by a variety of factors, such as spill amount, oil type, location-specific characteristics, and response methodology (Etkin, 2004). Later, we estimate the economic significance of our result using the full range of potential values.

revenue of the industry. Where spills occur at subsidiaries of these oil companies, we hand match the names of responsible firms with the LexisNexis Corporate Affiliation Directory to identify parent companies. We use the spill location to map accidents to corresponding states.

We obtain data on corporate philanthropy from the Foundation Directory Online (FDO), a nonprofit that compiles information on U.S. grant makers and grants. Specifically, it contains information on corporate giving of U.S. firms through their company-sponsored foundations and corporate giving programs. The grant data is drawn from the IRS Form 990s (Returns of Organization Exempt from Income Tax), which 501(c)3 nonprofit organizations such as corporate foundations are required to report annually.[17] These include cash donations, but not in-kind contributions and employee volunteerism. FDO records grant-level information on donor name and location, recipient name and location, amount given, and subject area. We match the recipient organization to the state where it is located in order to calculate grant amounts at the firm-state-year level. Given that donation data from FDO are available to us from 2003 to 2010, we limit our empirical analysis to 2004–2011 (accounting for a one-year lag). In addition to these two main sources, we also draw on a variety of other sources, including Compustat, CRSP, and NETS databases as detailed in Appendix S3.

3.3 | Empirical approach

In order to determine which of the four equilibria in Figure 5 prevails in our empirical context, we undertake two sets of analyses, consistent with the two discriminating predictions in Figure 5. First, we try and replicate prior results for the insurance-like benefits of philanthropy (Godfrey et al., 2009; Hong & Liskovich, 2017; Koh et al., 2014; Minor & Morgan, 2011) in order to confirm that philanthropy does protect firms in the face of negative events. Following these studies, we focus on the effect of negative events on the firm's stock market performance and the moderation of this effect by past philanthropy. We choose to focus on stock market performance as our outcome measure to be consistent with prior work (Godfrey et al., 2009; Hong & Liskovich, 2017; Minor & Morgan, 2011), and because it captures the expected value of a range of potential penalties the firm may face, including lost sales, regulatory and legal penalties, and clean-up costs. We also look at alternative measures of accident penalties as described in a later section.

Specifically, we study the cumulative abnormal return (CAR) around the spill, which is estimated using the Carhart four-factor model (Carhart, 1997). We regress the CAR on the size of the oil spill, the firm's prior philanthropy, and the interaction between the two, as well as a number of controls (discussed in more detail in the following), with year- and firm-fixed effects.[18] If the results from prior studies hold in our context, we would expect to see a negative main effect of the size of the oil spill on CAR, moderated by a positive interaction of oil spill size with past philanthropy, which would be consistent with Equilibria III or IV. If, on the other hand, we see a negative interaction term, that would be consistent with Equilibrium I, where more philanthropic firms face higher penalties.

Our second (and main) analysis examines the correlation between philanthropic donations by oil companies and subsequent oil spills. Specifically, we look at the relation between philanthropy and subsequent oil spills not just at the firm level, but at the firm-state level, consistent with our expectation that the effects of both will be largely local. The model for our main regression is:

---

[17]We check the accuracy of the FDO data against the 990s forms, and rely on the 990s in rare cases of inconsistency.

[18]Consistent with the prior literature (Flammer, 2013; McWilliams & Siegel, 1997), we do not include spill events for which market reaction is not available due to market closure in our sample, and treat cases of multiple oil spills by a firm on a single date as one observation, with information used for analysis aggregated across same-date spills. The CAR analysis also excludes four firms from our main sample that are not listed on U.S. stock exchanges.

$$Spill_{ijt+1} = f\left(Grant_{ijt}, Controls_{ijt+1}, Controls_{it}, FE_i, FE_j, FE_t\right), \tag{3}$$

where $Spill_{ijt+1}$ is the logged amount of spills by firm i in state j in year $t + 1$, and $Grant_{ijt}$ is the logged amount of philanthropic giving by firm i to recipient organizations located in state j in year t. If the coefficient of $Grant_{ijt}$ were positive, this would be consistent with the adverse selection/moral hazard equilibrium (Equilibrium III in Figure 5), while if it were negative, that would support the exogenous philanthropy equilibrium (Equilibrium IV). Note that a positive relation between philanthropy and subsequent oil spills could be the result of firms increasing their philanthropy as they observe the probability of future spills increasing (the adverse selection case), or of the probability of future spills increasing as the coverage provided by philanthropy makes firms laxer in prevention (the moral hazard case). Our empirical approach in no way distinguishes between these two possibilities; we are only seeking to establish an association between philanthropy and subsequent spills without testing or implying a causal relationship.

We include a vector of firm-level characteristics $Controls_{it}$ in year t (logged assets, ROA, cash holdings, and capital expenditure) to deal with the concern that time-varying firm characteristics may drive both donations as well as oil spills. In addition, we include firm-state level characteristics $Controls_{ijt+1}$—number of production facilities, number of nonproduction facilities, and logged sales of firm i in state j in year $t + 1$—to control for the level of the firm's activities in the state. Last, we include firm, state, and year-fixed effects, FE_i, FE_j, FE_t, with standard errors clustered at the firm level (our results are similar if standard errors are clustered at the state level). A full description of controls is provided in Appendix S3, and Appendix S4 provides summary statistics and correlations for both the CAR analysis (Panel A) and the oil spill analysis (Panel B).

4 | RESULTS

4.1 | Replicating the insurance benefits of philanthropy

We start by replicating the insurance-like benefits of philanthropy in our setting, that is, by confirming that philanthropy does limit the negative impact of spills on firms' financial performance. Table 1 presents the results of this analysis. With year-fixed effects and a two-day event window (0, +1) for measuring the CAR in Model 1, we find that corporate donations moderate the negative relationship between the size of the spill and stock market reaction, that is, stock market reactions following a spill are more negative the greater the amount of the spill ($\beta = -0.0512$, $p = .0577$), but this effect is weaker the higher a firm's corporate donations in the previous year (for the interaction term, $\beta = 0.0052$, $p = .0098$). The interaction effect is robust if we measure the CAR in a three-day event window (−1, +1) in Model 2, or if we replace firm-level donation with donations by the firm in the state where the spill occurred in Model 3. Further, in Models 4–6, we include firm-fixed effects and find continued support within firms, that is, for a given firm, the more it donates, the less the stock market penalizes the firm for the occurrence of spills. Overall, results in Table 1 are consistent with the prior literature and confirm that philanthropy does serve as a form of insurance in our context, ruling out the possibility that we are in Equilibrium I.

In addition to looking at stock market reactions, we consider two alternate measures of accident penalty (as shown in Appendix S5). First, we look at the probability that a firm faces a (civil or criminal) clean-up lawsuit from the EPA related to its oil spills.[19] Within-firm analyses show a negative

[19]As Appendix S5 reports, we find a total of 18 lawsuits (13 civil, 5 criminal) filed by the EPA against just seven of our sample firms within our sample period, so the statistical power of this test is low.

TABLE 1 Stock market reaction to spills and the insurance effects of philanthropy

	Model 1 CAR(0,+1) Firm-level donation	Model 2 CAR(−1,+1) Firm-level donation	Model 3 CAR(−1,+1) Firm-state level donation	Model 4 CAR(0,+1) Firm-level donation	Model 5 CAR(−1,+1) Firm-level donation	Model 6 CAR(−1,+1) Firm-state level donation
Amount of oil spill	−0.0512	−0.0441	−0.0011	−0.0650	−0.0610	−0.0050
	(.0577)	(.1445)	(.9482)	(.0083)	(.0453)	(.7785)
Corporate donation_l1	−0.0061	0.0010		0.0045	0.0270	
	(.2148)	(.8961)		(.6086)	(.0342)	
Amount of oil spill × Corporate donation_l1	0.0052	0.0056		0.0061	0.0065	
	(.0098)	(.0227)		(.0011)	(.0089)	
Firm-state corporate donation_l1			−0.0088			0.0013
			(.4442)			(.9240)
Amount of oil spill × Firm-state corporate donation_l1			0.0080			0.0078
			(.0477)			(.0536)
Production facilities	0.0026	0.0109	0.0095	−0.0066	−0.0021	0.0006
	(.6547)	(.2618)	(.2981)	(.1843)	(.7652)	(.9489)
Nonproduction facilities	−0.0364	−0.1061	−0.0727	−0.0690	−0.0692	−0.0610
	(.8172)	(.7034)	(.7757)	(.6730)	(.7977)	(.8354)
Sales	0.0059	0.0066	0.0064	0.0068	0.0090	0.0062
	(.4313)	(.6101)	(.5915)	(.3294)	(.4490)	(.5947)
Total assets_l1	0.0588	0.0530	0.0572	0.5122	0.5219	0.6490
	(.0317)	(.1929)	(.1321)	(.0360)	(.0355)	(.0203)
ROA_l1	−1.8475	−3.1826	−3.0160	−1.6868	−3.4122	−3.7470
	(.0238)	(.0056)	(.0109)	(.1041)	(.0458)	(.0602)
Cash holding_l1	1.6327	2.0100	2.2777	0.4599	−0.0278	0.6970
	(.0960)	(.0895)	(.0637)	(.7592)	(.9902)	(.7538)
CAPEX_l1	0.3185	−0.1514	−0.1612	0.3316	0.2160	0.3179
	(.5741)	(.8190)	(.8048)	(.7167)	(.8737)	(.8092)
Constant	−0.6027	−0.4574	−0.5034	−5.9758	−6.0941	−7.1724
	(.1042)	(.3185)	(.2253)	(.0370)	(.0378)	(.0310)
R^2	0.0069	0.0084	0.0078	0.0171	0.0196	0.0170
YearFE	Yes	Yes	Yes	Yes	Yes	Yes
FirmFE				Yes	Yes	Yes
Observations	3,460	3,460	3,460	3,460	3,460	3,460

Notes. Dependent variables are the cumulative abnormal return (CAR) over a two-day window (0, +1) or three-day window (−1, +1) around the date of oil spill. Estimated by Carhart four-factor model. Coefficients of OLS regressions. Standard errors are clustered at the firm level. *P*-values are reported in parentheses.

and significant effect of past philanthropy on the probability of subsequent clean-up lawsuits, controlling for spill amounts, consistent with philanthropy providing an insurance benefit. Second, we look at the relationship between philanthropy and the provisions firms make on their balance sheets for future settlements as a proxy for the penalties firms expect to pay. We find no evidence of an effect of either the amount of spill or of past philanthropy on these provisions, however, perhaps because we are forced to aggregate up to the firm-year level rather than look at the penalty to individual spills as we can with the CAR analysis. In any case, the analysis in Table 1, as well as the direct negative effect of past philanthropy on the probability of clean-up lawsuits, supports the conclusion that philanthropy is negatively related to accident penalties.

4.2 | Main results

Having replicated prior results for insurance benefits of philanthropy, we now turn to our main analysis, which examines the relation between philanthropy and subsequent spills, in order to distinguish between the adverse selection/moral hazard equilibrium (Equilibrium III in Figure 5) and the exogenous philanthropy equilibrium (Equilibrium IV). Our main results based on OLS fixed-effects regressions are reported in Table 2. All models use the logged amount of oil spills as the dependent variable. Models 7, 8a, and 8b are baseline analyses at the firm level. In Model 7, we include only the control variables. In Model 8a, we add our main predictor—the logged donation amount—and run a between-effects model. As can be seen, the coefficient on the corporate donation amount ($\beta = 0.2605$, $p = .0014$) supports a significant and positive association between donation and subsequent spills. This implies that, in the cross-section, firms that give more in one year tend to spill more in the subsequent year. In Model 8b, we then run a within-effects model, and again find a positive and strong coefficient for donation amount ($\beta = 0.0684$, $p = .0077$). This implies that in the year following an increase in a given firm's overall donation amount, the amount of oil it spills increases. Overall, these results are consistent with an adverse selection/moral hazard equilibrium (Equilibrium III) with higher levels of philanthropy being associated with greater subsequent spills both within and across firms.

Models 9 and 10 are our main models, and examine the relationship between donation and subsequent spills at a more granular, firm-state level. Model 9 contains controls with year-, state-, and firm-fixed effects. In Model 10,[20] we include the donation amount variable in the regression, which enters the model with a positive and significant coefficient ($\beta = 0.0370$, $p = .0033$). This further confirms a positive association between donations and subsequent spills—firms that give more, subsequently spill more—in line with an adverse selection/moral hazard equilibrium.

These results are economically significant. Based on our point estimates in M10, a one standard deviation increase in the donation amount is associated with a 0.11 standard deviation increase in the subsequent amount of oil spilled. While that may appear small in percentage terms, it is important to consider the average level of oil spills, and their high economic and social costs relative to average donation amounts. Starting from the sample average, a $100,000 increase in donations is associated with an increase of 39.3 barrels in oil spills, which corresponds to social and environmental costs of around $132,000 to $1.6 million based on the estimated damages of $3,360 to $41,370/barrel previously discussed (Etkin, 2004). Our estimates thus imply that the terms of exchange for philanthropy as reputation insurance may be unfavorable to society on average, with increases in donations in one year being more than offset by the associated increases in subsequent damages in the following year.

We conduct several robustness checks. One concern is that the association we observe may be the result of unobserved firm-level changes driving both donations and spills. If that were the case, we would expect to see a similar spurious association between donations in other states and subsequent spills in the focal state. In Model 11 in Table 2, we add Out-of-state donation amount as a control, and find that Out-of-state donation amount is not significant, while the effect of Within-state donation amount remains stable. A second concern is that firms' previous levels of spills may act as an omitted variable, predicting both the donation amount in the previous year as well as spills in the current period. We deal with this concern by including a lagged dependent variable in the regression in Model 12, and find that our main results are robust to the inclusion of lagged spills. In unreported

[20] Appendix S6 reports the value of firm-fixed effects corresponding to this regression.

TABLE 2 Main results

	DV: Logged firm-level amount of oil spill			DV: Logged firm-state level amount of oil spill			
	Model 7	Model 8a (Between effects)	Model 8b (Within effects)	Model 9	Model 10	Model 11	Model 12
Corporate donation_l1		0.2605	0.0684				
		(.0014)	(.0077)				
Firm-state corporate donation_l1					0.0370	0.0375	0.0254
					(.0033)	(.0035)	(.0052)
Firm-other-state corporate donation_l1						−0.0025	
						(.4709)	
Firm-state amount of oil spill_l1							0.2922
							(.0000)
Production facilities				0.0177	0.0159	0.0159	0.0114
				(.0949)	(.1273)	(.1282)	(.1284)
Nonproduction facilities				0.8164	0.7617	0.7600	0.5358
				(.0021)	(.0046)	(.0046)	(.0079)
Sales				0.0332	0.0317	0.0317	0.0231
				(.0000)	(.0000)	(.0000)	(.0001)
Total assets_l1	−0.1288	1.3222	−0.2986	−0.0385	−0.0566	−0.0507	−0.0500
	(.7699)	(.0085)	(.5173)	(.3042)	(.1594)	(.2305)	(.1355)
ROA_l1	−0.0136	−14.6866	0.3156	−0.0927	−0.0970	−0.1090	−0.1481
	(.9960)	(.3686)	(.9076)	(.7106)	(.7080)	(.6616)	(.5305)
Cash holding_l1	−0.6075	−11.6016	−0.7239	−0.1903	−0.1303	−0.1253	−0.0158
	(.8314)	(.3419)	(.7942)	(.0992)	(.3664)	(.3709)	(.9242)
CAPEX_l1	−2.4374	−10.9202	−2.8140	0.0013	−0.0631	−0.0504	−0.0602
	(.0629)	(.1043)	(.0516)	(.9878)	(.6611)	(.7145)	(.6311)
Constant	5.4021	−4.7258	6.7113	0.4324	0.5324	0.5022	0.4521
	(.2183)	(.4442)	(.1430)	(.2403)	(.1814)	(.2094)	(.1591)
R^2	0.0203	0.5405	0.0352	0.2300	0.2378	0.2379	0.3022
YearFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FirmFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
StateFE				Yes	Yes	Yes	Yes
Observations	278	278	278	13,900	13,900	13,900	13,900

Notes. Dependent variables are measured at the firm level (Models 7, 8a, and 8b) and firm-state level (Models 9–12) in a given year. Firm-level Models 7, 8a, and 8b, and firm-state level Models 9–12. Coefficients of OLS fixed-effects regressions. Standard errors are clustered at the firm level. *P*-values are reported in parentheses.

analyses, we also run AR1 models and confirm that our results are robust to controlling for auto-regression. In addition to the robustness checks in Table 2, we run several other models with alternate dependent variables, specifications, samples, and measures, and continue to find consistent results (in the interests of space, these are reported in Appendix S7).

4.3 | Supplementary analyses

While the preceding results confirm the existence of a strong positive association between donation and subsequent spills that is consistent with an adverse selection/moral hazard equilibrium (Equilibrium III), we further test this interpretation through two sets of supplementary analyses. Our approach in doing so is to see how the strength of the positive association between donations and subsequent spills varies as we vary the theoretical drivers underlying Equilibrium III.

First, we expect the positive relation between donation and subsequent spills to be stronger in the case of spills that are largely under firms' control (i.e., those that might have been prevented with greater safety training/maintenance). If donations result in more of such controllable spills, that would be strongly consistent with the moral hazard argument, which suggests that firms do less to prevent spills once they are assured of the protection of positive reputation; it would also be consistent with adverse selection since it is precisely the failures in their internal operations or equipment that firms would be best able to predict. On the other hand, oil spills that occur due to external causes such as natural disasters, or more generally, bad weather conditions, are presumably far more difficult for firms to either predict or prevent, so they should be less subject to moral hazard or adverse selection. Thus, if the positive association between donations and subsequent spills reflects problems of adverse selection/moral hazard, we would expect this association to be stronger for spills within firms' control than for spills outside it.

To test this, we divide the overall spill amount into the amount of spills due to natural disasters and that due to other internal causes[21] (as reported to the NRC), and run separate models using each as an alternate dependent variable. The results are reported in Models 13a and 13b of Table 3. The results confirm our expectation; compared to the effect of donation on spills caused by natural disasters ($\beta = 0.0074$, $p = .0438$), its effect on internally caused spills is stronger and larger in effect size ($\beta = 0.0333$, $p = .0080$). Similarly, Models 14a and 14b break spill amounts based on whether the spill happened during bad weather or during good weather (wind speed above or below 22 knots, respectively), on the assumption that spills in good weather are easier to predict and prevent than those in bad weather. We continue to see a significant and positive effect of donation on spills during good weather ($\beta = 0.0373$, $p = .0031$), whereas the coefficient of donation on bad weather spills is not significant ($\beta = 0.0001$, $p = .7244$). Finally, we look at the moderating effect of capital expenditure on the relationship between philanthropy and subsequent spills—if we are truly seeing adverse selection/moral hazard at work, we would expect this relationship to be stronger for firms that invest less on capital expenditure, and therefore, have older, less well-maintained equipment that is more susceptible to failure. Model 15 confirms this, with the interaction between philanthropy and capital expenditure taking a negative and significant coefficient ($\beta = -0.1885$, $p = .0009$). The results in Table 3 thus suggest that the positive association between donations and subsequent spills comes mainly from spills that are relatively controllable and closely linked to firms' internal prevention efforts, further supporting our moral hazard/ adverse selection arguments.

Second, we look at the moderating effect of a community's civic capacity. Since the underlying assumption driving Equilibrium III is that society is unable to organize and act in its collective interest, it follows that the positive association between donations and subsequent donations should be strongest in communities where this is true. To empirically test this, we exploit state-level differences in the level of civic capacity, that is, the extent to which the local community is able to organize to demand better terms from firms.

Table 4 shows the results of this analysis, using three alternate proxies for the level of civic capacity in the state. In each case, we conduct split sample analyses, using the median value of these proxies to divide our overall sample. Models 16a and 16b present the results based on a proxy of social capital developed by Putnam (2000). Consistent with our expectation, we find that the

[21]In unreported analyses, we distinguish between spills due to equipment failure, operator error, or transport accidents, and those due to all other causes, and confirm that our findings are robust to this alternate coding. We also examine whether different spill causes impact accident penalties (as in Table 1) differently, but find no significant differences.

TABLE 3 Supplementary analyses (I)

	Model 10 DV: Logged firm-state amount of oil spill	Model 13a DV: Logged firm-state amount of oil spill due to natural disasters	Model 13b DV: Logged firm-state amount of oil spill due to other causes	Model 14a DV: Logged firm-state amount of oil spill taking place during bad weather	Model 14b DV: logged firm-state amount of oil spill taking place during good weather	Model 15 DV: Logged firm-state amount of oil spill
Firm-state corporate donation_l1	0.0370	0.0074	0.0333	0.0001	0.0373	0.0749
	(.0033)	(.0438)	(.0080)	(.7244)	(.0031)	(.0002)
Production facilities	0.0159	0.0086	0.0120	0.0002	0.0159	0.0171
	(.1273)	(.0538)	(.2087)	(.2412)	(.1273)	(.0982)
Nonproduction facilities	0.7617	0.2946	0.7889	−0.0109	0.7688	0.7315
	(.0046)	(.0650)	(.0030)	(.1534)	(.0045)	(.0060)
Sales	0.0317	0.0044	0.0287	0.0002	0.0316	0.0313
	(.0000)	(.0297)	(.0001)	(.6113)	(.0000)	(.0000)
Total assets_l1	−0.0566	−0.0364	−0.0279	−0.0023	−0.0546	−0.0468
	(.1594)	(.0405)	(.3912)	(.5674)	(.1845)	(.2435)
ROA_l1	−0.0970	−0.2939	0.0752	−0.0257	−0.0726	−0.0500
	(.7080)	(.2125)	(.6944)	(.4214)	(.7861)	(.8523)
Cash holding_l1	−0.1303	0.2810	−0.3294	0.0059	−0.1376	−0.0719
	(.3664)	(.0446)	(.1061)	(.7747)	(.3579)	(.6444)
CAPEX_l1	−0.0631	−0.0970	−0.0053	0.0180	−0.0783	0.2850
	(.6611)	(.1225)	(.9702)	(.2345)	(.5654)	(.0579)
Firm-state corporate donation_l1 * CAPEX_l1						−0.1885
						(.0009)
Constant	0.5324	0.3312	0.2360	0.0475	0.4854	0.4127
	(.1814)	(.0502)	(.4816)	(.3699)	(.2242)	(.2927)
R^2	0.2378	0.0762	0.2282	0.0088	0.2390	0.2408
YearFE	Yes	Yes	Yes	Yes	Yes	Yes
FirmFE	Yes	Yes	Yes	Yes	Yes	Yes
StateFE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13,900	13,900	13,900	13,900	13,900	13,900

Notes. Dependent variables are measured at firm-state level in a given year. Model 10 is the same as that in Table 2. Coefficients of OLS fixed-effects regressions. Standard errors are clustered at the firm level. *P*-values are reported in parentheses. Causes as well as weather conditions of the oil spills are coded from the National Response Center (NRC) database recorded at the time of spill reporting. Natural disasters including earthquake (0.4%), explosion (0.05%), flood (0.2%), hurricane (9.1%), natural phenomenon (4.8%), tornado (0.01%), and vessel sinking (0.04%); while other causes include dumping (0.3%), equipment failure (45%), operator error (8%), other (10.3%), over pressuring (0.3%), transport accident (0.4%), and unknown (21%). Bad weather is coded 1 if the wind speed is above 22 knots, and 0 otherwise.

coefficient of donation is greater in states with low civic capacity ($\beta = 0.0420$, $p = .0000$) than in states with high civic capacity ($\beta = 0.0071$, $p = .0769$), with a Chi-square test confirming the significant difference between these two coefficients ($p = .0056$). Models 17a and 17b then use the level of political disagreement (measured as the standard deviation in county-level voting percentages for the Democratic candidate in the 2008 presidential election) as a (negative) proxy for civic capacity, with the rationale that states whose citizens are more politically polarized will find it harder to collectively organize in their own interest. As before, we see that the positive association between donations and subsequent spills is stronger in states where people diverge on their political values ($\beta = 0.0580$,

TABLE 4 Supplementary analyses (II)

	Model 10 All sample	Model 16a Split sample (States with high civic capacity)	Model 16b Split sample (States with low civic capacity)	Model 17a Split sample (States with low political disagreement)	Model 17b Split sample (States with high political disagreement)	Model 18a Split sample (Blue states)	Model 18b Split sample (Red states)
Firm-state corporate donation_l1	0.0370	0.0071	0.0420	0.0129	0.0580	0.0269	0.0525
	(.0033)	(.0769)	(.0000)	(.0013)	(.0000)	(.0000)	(.0000)
Production facilities	0.0159	0.0925	0.0135	0.0458	0.0112	0.0506	0.0113
	(.1273)	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)
Nonproduction facilities	0.7617	0.1188	0.8842	−0.1987	1.0416	0.7628	0.6215
	(.0046)	(.3306)	(.0000)	(.0397)	(.0000)	(.0000)	(.0000)
Sales	0.0317	0.0100	0.0340	0.0268	0.0333	0.0255	0.0310
	(.0000)	(.0001)	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)
Total assets_l1	−0.0566	0.0090	−0.1151	−0.0462	−0.0820	−0.0103	−0.1202
	(.1594)	(.8303)	(.0927)	(.3230)	(.2248)	(.8293)	(.0730)
ROA_l1	−0.0970	0.0114	0.0004	−0.1608	0.1343	−0.2338	0.1129
	(.7080)	(.9616)	(.9992)	(.5430)	(.7246)	(.3874)	(.7657)
Cash holding_l1	−0.1303	−0.3879	−0.1720	−0.2675	−0.1515	0.0178	−0.3394
	(.3664)	(.1795)	(.7159)	(.4073)	(.7450)	(.9571)	(.4634)
CAPEX_l1	−0.0631	0.1644	−0.3001	0.0022	−0.2043	0.0832	−0.2476
	(.6611)	(.3225)	(.2701)	(.9907)	(.4463)	(.6619)	(.3524)
Constant	0.5324	−0.1897	1.0085	0.2904	0.4677	−0.0371	0.9343
	(.1814)	(.5824)	(.0740)	(.4510)	(.4006)	(.9251)	(.0906)
Wald test comparison of coefficients		7.68 (p-value = .0056)		14.68 (p-value = .0001)		3.21 (p-value = .0730)	
R^2	0.2378	0.1073	0.2920	0.1220	0.3121	0.1858	0.2879
YearFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FirmFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
StateFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13,900	6,672	6,672	6,950	6,394	7,784	6,116

Notes. Dependent variable is the logged firm-state amount of oil spill in a given year. Model 10 is the same as that in Table 2. Coefficients of OLS fixed-effects regressions. Standard errors are clustered at the firm level. *P*-values are reported in parentheses. In Models 16a and b, states with high/low civic capacity are defined as those with above/below median levels of social capital (Putnam, 2000), respectively. In Models 17a and b, states with high political disagreements are those states with a higher-than-median standard deviation in political opinion, measured by the county-level percentage voting toward the Democratic party in the 2008 presidential election; in Models 18a and b, Blue (Red) states are defined as states that voted for the Democratic (Republican) party candidate in the 2008 presidential election. The Wald test evaluates the differences of the coefficients of Firm-state corporate donation_l1 across split samples for M16, M17 and M18.

p = .0000), which present greater challenges for collective action, than for states where people have similar political leanings, and thus, can more easily act together ($\beta = 0.0129$, $p = .0013$), with a Chi-square test confirming that the two coefficients are significantly different ($p = .0001$). Finally, we split the sample by Red and Blue states, based on whether the state voted for the Republican or Democratic candidate in the 2008 presidential election, respectively. While party affiliation is not directly a measure of civic capacity, prior work suggests that Red states tend to be more pro-business (Pe'Er & Gottschalg, 2011), so that we would expect stakeholders in such states to be less critical and demanding of firms than those in Blue states. Consistent with this expectation, Models 18a and 18b show that the increase in subsequent spills associated with donations is greater in Red states than in Blue states ($\beta = 0.0525$ and 0.0269, respectively), with the difference between them being statistically significant ($p = .0730$). Overall, across all three split-samples we see a clear pattern of the positive association between donations and subsequent spills being less pronounced in states where stakeholders are more able or likely to collectively organize and demand better terms from corporations, consistent with the assumptions underlying Equilibrium III.

5 | CONCLUSIONS AND DISCUSSION

Our study critically examines the use of philanthropy as a form of reputation insurance both theoretically and empirically. Theoretically, we develop a rigorous formal theory of such insurance, considering not only the demand for such insurance by firms (Godfrey, 2005), but also the supply of such insurance by society, thus deriving the terms on which such insurance is offered, and the implications of those terms for both firm profits and social welfare. This formal analysis allows us to map alternate assumptions about reputation insurance to the corresponding equilibria, highlighting the role of informed stakeholder activism, as well as the importance of philanthropy being undertaken exogenously, that is, without consideration of its insurance benefits. It also shows that the fact that philanthropy serves as a form of reputation insurance does not by itself mean that it is beneficial to either firms or society. Moreover, our formal model produces a set of empirical predictions that allow us to distinguish between these alternate equilibria.

Empirically, our findings from the U.S. oil industry show that while philanthropy does ameliorate the negative consequences of oil spills for firms, consistent with prior work, it is positively associated with the amount and likelihood of subsequent spills—firms that give more, spill more. While not strictly causal, this result is robust to a variety of specifications, controls, and levels of analysis. Moreover, supplementary analyses show that this association is stronger for spills that the firm could plausibly predict or prevent; and is stronger in states with low civic capacity, or those that are more likely to view corporations favorably. Taken together, these results provide strong evidence for the presence of an adverse selection/moral hazard problem in the use of philanthropy as reputation insurance in our context. At the very least, our findings call attention to the possibility that the use of philanthropy as reputation insurance may benefit firms at the cost of social welfare, and that firms that invest in philanthropy for this purpose may "win us with honest trifles, to betray's/in deepest consequence" (Shakespeare, n.d.).

Our article makes several contributions to the existing literature. First, it provides a more complete and coherent theoretical account of philanthropy as a form of reputation insurance. While this idea has been widely cited and discussed in the strategy literature (Godfrey, 2005; Godfrey et al., 2009), the theory for such reputation insurance remains incomplete, with prior work considering neither the terms on which such insurance is made available, nor the possibility of strategic action on the part of society, nor the information problems involved. Our study addresses these gaps, providing a more analytically rigorous theory of philanthropy as reputation insurance, one that both derives equilibrium conditions and examines how they change under alternate assumptions. In doing so, our study not only answers the call for a more nuanced understanding of the use of philanthropy as insurance (Barnett et al., 2017) and the associated cost-benefit trade-offs (Singh et al., 2017), it also joins a small but growing body of work that seeks to bring more formal rigor to nonmarket strategy research (Chatain & Plaksenkova, 2018; Fosfuri et al., 2016; Kaul & Luo, 2018).

Second, by emphasizing the social impact of philanthropy as reputation insurance, our study also contributes to a stream of recent work concerned with critically assessing the welfare consequences of CSR both theoretically and empirically (Ballesteros et al., 2017; Barnett, 2016; De Bettignies & Robinson, 2018; Kaul & Luo, 2018). While prior work on philanthropy as reputation insurance has focused almost exclusively on the benefits of such insurance for firms (Flammer, 2013; Godfrey, 2005; Godfrey et al., 2009; Koh et al., 2014), we shift the focus to the welfare implications of philanthropy as insurance, arguing that—to the extent that the protection from such insurance makes firms more lax in preventing accidents—society may be left worse off. Our formal theory thus suggests that the assumption that lies at the heart of reputation insurance—that more philanthropic firms are more moral or trustworthy (Godfrey, 2005)—may only make sense when philanthropy is not motivated by the quest for insurance. Where firms are aware of the insurance potential of philanthropy and invest in it strategically, that assumption may be naïve, allowing firms to take advantage of stakeholders (Akerlof & Shiller, 2015) to buffer their own profits. Our article not only highlights the possibility of such an adverse selection/moral hazard problem, it provides empirical evidence in support of it. By showing that firms that give more, spill more, our study thus highlights the hazards of relying on corporate reputations (Williamson, 1985), and suggests the need for skepticism when dealing with for-profit philanthropy (Bhanji & Oxley, 2013; Kaul & Luo, 2018; Milgrom & Roberts, 1986). Moreover, while our focus is on philanthropic reputation, our study has implications for work on corporate reputation more generally, raising the possibility that other forms of reputation may also be susceptible to adverse selection/moral hazard.

Third, our study also highlights the importance of informed and active stakeholders in ensuring that corporate philanthropy is truly to society's benefit, thus contributing to a growing literature on private politics (Abito, Besanko, & Diermeier, 2016; Baron, 2001; Baron & Diermeier, 2007) and the importance of strong oversight of firms' CSR efforts (Delmas & Montes-Sancho, 2010; Marquis, Toffel, & Zhou, 2016; Short & Toffel, 2010). Specifically, our arguments suggest—and our supplementary analyses on the role of civic capacity confirm—that social activists (Hiatt & Park, 2013; Ingram, Yue, & Rao, 2010; King & Soule, 2007; McDonnell & King, 2013; Sine & Lee, 2009), media outlets (Luo, Meier, & Oberholzer-Gee, 2012), self-regulatory collectives (King & Lenox, 2000; Luo & Kaul, 2018; Yue, Luo, & Ingram, 2013), and public disclosure requirements (Reid & Toffel, 2009) may all have a critical role to play in ensuring social welfare by subjecting CSR activities to stronger scrutiny.

As with any study, ours has several limitations. As one of the first studies to offer a formal account of philanthropy as reputation insurance, our model is relatively simple, and future work could certainly extend it in several ways, for example, by adding parameters for the relative bargaining power of firms versus society, by building in dynamics over time, or by allowing society to invest in learning about firm type. Future work could also move away from the assumption of uniform preferences on the part of society and engage with problems of social choice (Arrow, 1951) or explore the implications of alternate ways of aggregating individual risk preferences to the societal level. On the empirical side, our findings are preliminary in that they only show an association between donations and subsequent spills. More work is needed to understand the causal relationship between the two, and therefore, to determine whether the association is the result of adverse selection or moral hazard. Moreover, one plausible alternate explanation for this association is that firms that undertake riskier or more experimental projects also contribute more to philanthropy in anticipation of future failures.[22] We think this is unlikely, both because such an explanation would be consistent with Equilibrium I, which we did not find support for, and because it would predict a positive moderation of the relationship between philanthropy and accident probability by capital expenditure (firms that invested more would give more, and have more spills), and our empirical results show a negative

[22]We are grateful to an anonymous reviewer for suggesting this alternate explanation.

moderation. Nevertheless, our current analyses cannot entirely rule out this possibility, and future work could explore it further.

Our empirical findings are also limited in that they are based on evidence from a single industry, one that is often regarded as being highly polluting. It is certainly not our intent to claim that our empirical results generalize to other industries, or that the use of philanthropy as reputation insurance is always associated with adverse selection/moral hazard. On the contrary, we expect that the equilibrium conditions will vary with context, and we would welcome future work that tests our framework and predictions in other settings. Finally, while we expect that our logic and findings should extend beyond corporate philanthropy, we would welcome future work that extends both our formal model and our empirical analysis to directly consider other forms of CSR.

To conclude, our study offers a formal theoretical account of the use of philanthropy as a form of reputation insurance, deriving the terms at which such insurance is offered in equilibrium, and studying how this equilibrium changes with changing assumptions about the exogeneity of philanthropy, the civic capacity of stakeholders, and the information context. Using our formal model, we highlight the possibility of an adverse selection/moral hazard equilibrium where philanthropy is positively associated with the likelihood of an accident. Consistent with this, empirical results from the U.S. oil industry show that firms that give more, subsequently spill more. Our study thus calls into question the value to society of corporate philanthropy as a form of reputation insurance, suggesting that, at least in some cases, it may be a way for firms to benefit their shareholders at the cost of social welfare.

ACKNOWLEDGEMENTS

We thank Editor Sendil Ethiraj and two anonymous reviewers for their generous and insightful feedback throughout the review process. We are also grateful to Alan Benson, Connie Helfat, Vit Henisz, Gwendolyn Lee, Marvin Lieberman, Rich Makadok, Alfie Marcus, Myles Shaver, Wes Sine, Jasjit Singh, Joel Waldfogel, and Sarah Wolfolds as well as to participants at the 2016 Strategic Management Society Conference Junior Faculty Paper Development Workshop, the 2017 Academy of Management Conference, the 2017 Society for Institutional & Organizational Economics Conference, the 2017 Strategic Management Society Conference, and seminar participants at Cornell University, Nanyang Technological University, the National University of Singapore, Purdue University, the University of Florida, and the University of Minnesota for providing comments and feedback. Luo acknowledges research and funding support from Gurneeta Vasudeva Singh and Nicholas Poggioli; some part of the data used in this article were collected in the course of a separate collaboration with Singh and Poggioli that ended in July 2015, and are being used for this study as per mutual agreement with Singh and Poggioli. In fulfilment of this agreement, Luo also acknowledges that once, in the summer of 2015, she, Singh, and Poggioli ran an empirical model similar to one of those reported here on a subset of the data used for this project, while examining a theoretically unrelated question, the findings from which were inconclusive. All errors are our own.

ORCID

Jiao Luo http://orcid.org/0000-0002-8003-3958

Aseem Kaul http://orcid.org/0000-0003-1455-6897

REFERENCES

Abito, J. M., Besanko, D., & Diermeier, D. (2016). Corporate reputational dynamics, private regulation, and activist pressure. In J. M. De Figueiredo, M. Lenox, F. Oberholzer-Gee, & R. G. Vanden Bergh (Eds.), *Strategy beyond markets (Advances in strategic management, Volume 34)* (pp. 235–299). Bingley, U.K.: Emerald Group Publishing Limited.

Aguilera, R. V., Rupp, D. E., Williams, C. A., & Ganapathi, J. (2007). Putting the S back in corporate social responsibility: A multilevel theory of social change in organizations. *Academy of Management Review*, *32*(3), 836–863.

Akerlof, G. A., & Shiller, R. J. (2015). *Phishing for phools: The economics of manipulation and deception*. Princeton, NJ: Princeton University Press.

Arrow, K. J. (1951). *Social choice and individual values*. New York, NY: Wiley.

Baker, G., Gibbons, R., & Murphy, K. J. (2002). Relational contracts and the theory of the firm. *Quarterly Journal of Economics*, *117*(1), 39–84.

Ballesteros, L., Useem, M., & Wry, T. (2017). Masters of disasters? An empirical analysis of how societies benefit from corporate disaster aid. *Academy of Management Journal*, *60*(5), 1682–1708.

Bansal, P., & Roth, K. (2000). Why companies go green: A model of ecological responsiveness. *Academy of Management Journal*, *43*(4), 717–736.

Barnett, M. L. (2014). Why stakeholders ignore firm misconduct: A cognitive view. *Journal of Management*, *40*(3), 676–702.

Barnett, M. L. (2016). The business case for corporate social responsibility: A critique and an indirect path forward. *Business and Society* (in press). https://doi.org/10.1177/0007650316660044

Barnett, M. L., Hartmann, J., & Salomon, R. (2017). Have you been served? Extending the relationship between corporate social responsibility and lawsuits. *Academy of Management Discoveries*, *4*(2), 109–126.

Barnett, M. L., & Salomon, R. M. (2006). Beyond dichotomy: The curvilinear relationship between social responsibility and financial performance. *Strategic Management Journal*, *27*(11), 1101–1122.

Barnett, M. L., & Salomon, R. M. (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.

Baron, D. P. (2001). Private politics, corporate social responsibility, and integrated strategy. *Journal of Economics & Management Strategy*, *10*(1), 7–45.

Baron, D. P., & Diermeier, D. (2007). Strategic activism and nonmarket strategy. *Journal of Economics & Management Strategy*, *16*(3), 599–634.

Bénabou, R., & Tirole, J. (2010). Individual and corporate social responsibility. *Economica*, *77*(305), 1–19.

Bhanji, Z., & Oxley, J. E. (2013). Overcoming the dual liability of foreignness and privateness in international corporate citizenship partnerships. *Journal of International Business Studies*, *44*(4), 290–311.

Bode, C., Singh, J., & Rogan, M. (2015). Corporate social initiatives and employee retention. *Organization Science*, *26*(6), 1702–1720.

Bosse, D. A., Phillips, R. A., & Harrison, J. S. (2009). Stakeholders, reciprocity, and firm performance. *Strategic Management Journal*, *30*(4), 447–456.

Brammer, S., & Pavelin, S. (2005). Corporate reputation and an insurance motivation for corporate social investment. *Journal of Corporate Citizenship*, *1*(20), 39–51.

Bromiley, P., & Marcus, A. (1989). The deterrent to dubious corporate behavior: Profitability, probability and safety recalls. *Strategic Management Journal*, *10*(3), 233–250.

Burbano, V. C. (2016). Social responsibility messages and worker wage requirements: Field experimental evidence from online labor marketplaces. *Organization Science*, *27*(4), 1010–1028.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, *52*(1), 57–82.

Chatain, O., & Plaksenkova, E. (2018). NGOs and the creation of value in supply chains. *Strategic Management Journal* (in press). https://doi.org/10.1002/smj.2938

Chin, M. K., Hambrick, D. C., & Trevino, L. K. (2013). Political ideologies of CEOs: The influence of executives' values on corporate social responsibility. *Administrative Science Quarterly*, *58*(2), 197–232.

Cuypers, I. R., Koh, P. S., & Wang, H. (2015). Sincerity in corporate philanthropy, stakeholder perceptions and firm value. *Organization Science*, *27*(1), 173–188.

De Bettignies, J. E., & Robinson, D. T. (2018). When is corporate social responsibility socially desirable? *Journal of Labor Economics* (in press).

Delmas, M. A., & Montes-Sancho, M. J. (2010). Voluntary agreements to improve environmental quality: Symbolic and substantive cooperation. *Strategic Management Journal*, *31*(6), 575–601.

Dorobantu, S., Kaul, A., & Zelner, B. A. (2017). Non-market strategy research through the lens of new institutional economics: An integrative review and future directions. *Strategic Management Journal*, *38*(1), 114–140.

Environmental Protection Agency. (1999). Oil spill program update. *The U.S. EPA's Oil Program Center Report*, *2*(2), 1–14.

Etkin, D. S. (2004). *Modelling oil spill response and damage costs*. U.S. EPA Archive Document. Retrieved from https://archive.epa.gov/emergencies/docs/oil/fss/fss04/web/pdf/etkin2_04.pdf.

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 corporate social responsibility lead to superior financial performance? A regression discontinuity approach. *Management Science*, *61*(11), 2549–2568.

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

Fosfuri, A., Giarratana, M. S., & Roca, E. (2016). Social business hybrids: Demand externalities, competitive advantage, and growth through diversification. *Organization Science*, *27*(5), 1275–1289.

Freeman, R. E. (1984). *Strategic Management: A Stakeholder Approach*. Boston, MA: Pitman.

Frey, B. S., & Meier, S. (2004). Pro-social behavior in a natural setting. *Journal of Economic Behavior & Organization*, *54*(1), 65–88.

Galaskiewicz, J. (1997). An urban grants economy revisited: Corporate charitable contributions in the Twin Cities, 1979–81, 1987–89. *Administrative Science Quarterly*, *42*(3), 445–471.

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., Merrill, C. B., & Hansen, J. M. (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.

Helfat, C. E. (1994). Evolutionary trajectories in petroleum firm R&D. *Management Science*, *40*(12), 1720–1747.

Helfat, C. E. (1997). Know-how and asset complementarity and dynamic capability accumulation: The case of R&D. *Strategic Management Journal*, *18*(5), 339–360.

Henisz, W. J., Dorobantu, S., & Nartey, L. J. (2014). Spinning gold: The financial returns to stakeholder engagement. *Strategic Management Journal*, *35*(12), 1727–1748.

Hiatt, S. R., & Park, S. (2013). Lords of the harvest: Third-party influence and regulatory approval of genetically modified organisms. *Academy of Management Journal*, *56*(4), 923–944.

Hong, H., & Liskovich, I. (2017). *Crime, punishment and the value of corporate social responsibility* (Working Paper). Retrieved from https://ssrn.com/abstract=2492202.

Ingram, P., Yue, L. Q., & Rao, H. (2010). Trouble in store: Probes, protests, and store openings by Wal-Mart, 1998–2007. *American Journal of Sociology*, *116*(1), 53–92.

Jayachandran, S., Kalaignanam, K., & Eilert, M. (2013). Product and environmental social performance: Varying effect on firm performance. *Strategic Management Journal*, *34*(10), 1255–1264.

Kang, C., Germann, F., & Grewal, R. (2016). Washing away your sins? Corporate social responsibility, corporate social irresponsibility, and firm performance. *Journal of Marketing*, *80*(2), 59–79.

Kaul, A., & Luo, J. (2018). An economic case for CSR: The comparative efficiency of for-profit firms in meeting consumer demand for social goods. *Strategic Management Journal*, *39*, 1650–1677.

King, A. A., & Lenox, M. J. (2000). Industry self-regulation without sanctions: The chemical industry's responsible care program. *Academy of Management Journal*, *43*(4), 698–716.

King, A. A., & Lenox, M. J. (2001). Does it really pay to be green? An empirical study of firm environmental and financial performance. *Journal of Industrial Ecology*, *5*(1), 105–116.

King, B. G., & Soule, S. A. (2007). Social movements as extra-institutional entrepreneurs: The effect of protests on stock price returns. *Administrative Science Quarterly*, *52*(3), 413–442.

Koh, P.-S., Qian, C., & Wang, H. C. (2014). Firm litigation risk and the insurance value of corporate social performance. *Strategic Management Journal*, *35*(10), 1464–1482.

Kölbel, J. F., Busch, T., & Jancso, L. M. (2017). How media coverage of corporate social irresponsibility increases financial risk. *Strategic Management Journal*, *38*(11), 2266–2284.

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

Luo, J., & Kaul, A. (2018). *Private action in public interest: The comparative governance of social issues* (Working Paper). Retrieved from https://ssrn.com/abstract=2769368.

Luo, J., Meier, S., & Oberholzer-Gee, F. (2012). *No news is good news: CSR strategy and newspaper coverage of negative firm events* (Working Paper). Cambridge, MA: Harvard Business School.

MacDonald, G., & Ryall, M. D. (2004). How do value creation and competition determine whether a firm appropriates value? *Management Science*, *50*(10), 1319–1333.

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

Maitland, I. (1994). The morality of the corporation: An empirical or normative disagreement. *Business Ethics Quarterly*, *4*(4), 445–458.

Marquis, C., Glynn, M. A., & Davis, G. F. (2007). Community isomorphism and corporate social action. *Academy of Management Review*, *32*(3), 925–945.

Marquis, C., & Lee, M. (2013). Who is governing whom? Executives, governance, and the structure of generosity in large U.S. firms. *Strategic Management Journal*, *34*(4), 483–497.

Marquis, C., & Tilcsik, A. (2016). Institutional equivalence: How industry and community peers influence corporate philanthropy. *Organization Science*, *27*(5), 1325–1341.

Marquis, C., Toffel, M. W., & Zhou, Y. (2016). Scrutiny, norms, and selective disclosure: A global study of greenwashing. *Organization Science*, *27*(2), 483–504.
McDonnell, M. H., & King, B. (2013). Keeping up appearances: Reputational threat and impression management after social movement boycotts. *Administrative Science Quarterly*, *58*(3), 387–419.
McWilliams, A., & Siegel, D. (1997). Event studies in management research: Theoretical and empirical issues. *Academy of Management Journal*, *40*(3), 626–657.
Milgrom, P., & Roberts, J. (1986). Relying on the information of interested parties. *RAND Journal of Economics*, *17*(1), 18–32.
Minor, D., & Morgan, J. (2011). CSR as reputation insurance: Primum non nocere. *California Management Review*, *53*(3), 40–59.
Ostrom, E. (2005). *Understanding institutional diversity*. Princeton, NJ: Princeton University Press.
Pauly, M. V. (1974). Overinsurance and public provision of insurance: The roles of moral hazard and adverse selection. *Quarterly Journal of Economics*, *88*(1), 44.
Pe'Er, A., & Gottschalg, O. (2011). Red and Blue: The relationship between the institutional context and the performance of leveraged buyout investments. *Strategic Management Journal*, *32*(12), 1356–1367.
Peloza, J. (2006). Using corporate social responsibility as insurance for financial performance. *California Management Review*, *48*(2), 52–72.
Poppo, L., & Zenger, T. R. (2002). Do formal contracts and relational governance function as substitutes or complements? *Strategic Management Journal*, *23*(8), 707–725.
Porter, M. E., & Kramer, M. R. (2006). Strategy and society: The link between competitive advantage and corporate social responsibility. *Harvard Business Review*, *84*(12), 1–16.
Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. New York, NY: Simon & Schuster.
Reid, E. M., & Toffel, M. W. (2009). Responding to public and private politics: Corporate disclosure of climate change strategies. *Strategic Management Journal*, *30*(11), 1157–1178.
Rothschild, M., & Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics*, *90*(4), 629–649.
Shakespeare, W. (n.d.) *Macbeth* from Folger Digital Texts. Retrieved from www.folgerdigitaltexts.com
Shiu, Y.-M., & Yang, S.-L. (2017). Does engagement in corporate social responsibility provide strategic insurance-like effects? *Strategic Management Journal*, *38*(2), 455–470.
Short, J. L., & Toffel, M. W. (2010). Making self-regulation more than merely symbolic: The critical role of the legal environment. *Administrative Science Quarterly*, *55*(3), 361–396.
Sine, W. D., & Lee, B. H. (2009). Tilting at windmills? The environmental movement and the emergence of the US wind energy sector. *Administrative Science Quarterly*, *54*(1), 123–155.
Singh, J., Teng, N., & Netessine, S. (2017). Philanthropic campaigns and customer behavior: Field experiments on an online taxi booking platform. *Management Science* (in press). https://doi.org/10.1287/mnsc.2017.2887.
Waddock, S. A., & Graves, S. B. (1997). The corporate social performance-financial performance link. *Strategic Management Journal*, *18*(4), 303–319.
Wang, H., Choi, J., & Li, J. (2008). Too little or too much? Untangling the relationship between corporate philanthropy and firm financial performance. *Organization Science*, *19*(1), 143–159.
Wang, H. C., & Qian, C. (2011). Corporate philanthropy and corporate financial performance: The roles of stakeholder response and political access. *Academy of Management Journal*, *54*(6), 1159–1181.
Williamson, O. E. (1985). *The economic institutions of capitalism*. New York, NY: Free Press.
Yue, L. Q., Luo, J., & Ingram, P. (2013). The failure of private regulation: Elite control and market crises in the Manhattan banking industry. *Administrative Science Quarterly*, *58*(1), 37–68.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Luo J, Kaul A, Seo H. Winning us with trifles: Adverse selection in the use of philanthropy as insurance. *Strat Mgmt J*. 2018;39:2591–2617. https://doi.org/10.1002/smj.2935

Winning us with trifles:

Adverse selection in the use of philanthropy as insurance

Jiao Luo
e-mail: luoj@umn.edu

Aseem Kaul
e-mail: akaul@umn.edu

Haram Seo
e-mail: seoxx209@umn.edu

Carlson School of Management
University of Minnesota
321 19th Avenue South
Minneapolis, MN 55455

ONLINE APPENDIX

S1: Derivation of Equilibrium Conditions

S2: Discussion of Alternative Equilibria

S3: Main Variables

S4: Summary Statistics and Correlation

S5: Alternate Measures of Accident Penalty

S6: Firm Fixed Effects

S7: Robustness checks

Appendix S1. Derivation of Equilibrium Conditions

Stage 1: Firm choice of prevention

In the first stage, the firm chooses the optimal level of prevention expenditure m_i. If reputation insurance were unavailable, the firm would choose an amount that minimized the joint cost of prevention and accidents, *i.e.,* it would solve

$$\min_{m_i} p_i(m_i)F + m_i$$

Taking first order condition, this means that the firm would choose a level of prevention $\dot{m}_i$ such that $\frac{\partial p_i^*}{\partial m} = -\frac{1}{F}$ where $p_i^* = p_i(m_i^*)$. This is the same as the $-\frac{\partial(p_i^* F)}{\partial m_i} = 1$ condition derived in the main text, and follows logically from the idea that the firm will set prevention expense such that the marginal return in reduced accident costs from every unit of prevention ($-\frac{\partial(p_i^* F)}{\partial m_i}$) is equal to the marginal cost of prevention.

Next, consider the case where insurance is available. The optimal level of prevention spending m_i that the firm should undertake for any given insurance contract α_i, may be determined by solving

$$\max_{m_i} \pi(\alpha_i, m_i)$$

where $\pi(\alpha_i, m_i) = p_i^* F + m_i^* - p_i(m_i)W_i - \left(1 - p_i(m_i)\right)S_i - m_i$.

The first order condition is:

$$\frac{\partial \dot{p}_i}{\partial m} = -\frac{1}{W_i - S_i}$$

where $\dot{p}_i = p_i(\dot{m}_i)$ and $\dot{m}_i$ is the optimal prevention spending given α_i. We assume further that $\dot{m}_i = 0$ if $W_i = S_i$, *i.e.*, given full insurance the firm prefers to undertake no prevention, and define $p_i^0 = p_i(0)$ as the exogenous risk of an accident in the absence of prevention.

Note that since $F \geq W_i - S_i$, $\frac{\partial \dot{p}_i}{\partial m} \leq \frac{\partial p_i^*}{\partial m} \Rightarrow \dot{p}_i \geq p_i^* \Rightarrow \dot{m}_i \leq m_i^*$. Thus, access to reputation insurance tends to reduce prevention spending and increase the probability of an accident.

Stage 2: Optimal contract choice

In stage 2, an optimal insurance contract is chosen. We consider three alternate scenarios:

Active and fully informed stakeholders

If stakeholders are active, then the choice of the optimal contract shall be made by society. Thus, society will choose the contract that maximizes its gain from philanthropy as reputation insurance without harming the firm, *i.e.*, it will solve:

$$\max_{S_i} V(\alpha_i, \dot{m}_i)$$

$$s.t. \, \pi(\alpha_i, \dot{m}_i) = 0$$

where $V(\alpha_i, \dot{m}_i) = (1 - \dot{p}_i)U(S_i) + \dot{p}_i U(W_i) - \dot{p}_i U(F)$

and $\pi(\alpha_i, \dot{m}_i) = p_i^* F + m_i^* - \dot{p}_i W_i - (1 - \dot{p}_i)S_i - \dot{m}_i$

The first order condition is:

$$(1 - \dot{p}_i)U'(S_i^*) + \dot{p}_i U'(W_i^*) \frac{\partial W_i^*}{\partial S_i} = \frac{\partial \dot{p}_i}{\partial S_i}[U(F) - U(W_i^*) + U(S_i^*)]$$

where α_i^* is the welfare maximizing contract for society. If society is fully informed, *i.e.*, it can monitor and enforce that the firm undertakes prevention equal to m_i^* then it will logically choose the contract that corresponds to that level of prevention and its associated accident probability, and the firm will have no choice but to maintain its pre-insurance level of accident prevention. Substituting p_i^* for $\dot{p}_i$ above, and recalling that $\frac{\partial p_i^*}{\partial S_i} = 0$ since p_i^* is independent of S_i, the first order condition reduces to

$$\frac{\partial W_i^*}{\partial S_i} = -\frac{(1 - p_i^*)}{p_i^*} \frac{U'(S_i^*)}{U'(W_i^*)}$$

But setting $\pi(\alpha_i, \dot{m}_i) = 0$ and differentiating with respect to S_i gives us $\frac{\partial \dot{W}_i}{\partial S_i} = -\frac{1}{\dot{p}_i}\left[\frac{\partial \dot{p}_i}{\partial S_i}\left(\dot{W}_i - \dot{S}_i\right) + \frac{\partial \dot{m}_i}{\partial S_i} + (1 - \dot{p}_i)\right] = -\frac{(1-\dot{p}_i)}{\dot{p}_i}$, substituting $\frac{\partial \dot{p}_i}{\partial S_i} = \frac{\partial \dot{p}_i}{\partial m_i}\frac{\partial \dot{m}_i}{\partial S_i}$ and $\frac{\partial \dot{p}_i}{\partial m} = -\frac{1}{W_i - S_i}$. It follows that

$$\frac{\partial W_i^*}{\partial S_i} = -\frac{(1-p_i^*)}{p_i^*}$$

Thus $\frac{U\prime(S_i^*)}{U\prime(W_i^*)} = 1 \Rightarrow W_i^* = S_i^*$. In equilibrium, the firm buys full insurance. Moreover, $\pi(\alpha_i^*, m_i^*) = 0$ implies that $S_i^* = p_i^* F$, that is, the firm undertakes an amount of philanthropy exactly equal to its expected cost of an accident, after it has maximized its attempts at prevention. The value to the firm $V(\alpha_i^*, m_i^*) = U(p_i^* F) - p_i^* U(F)$, which is positive, given our assumptions about societal risk aversion. The basic intuition in this case is that society uses its strong bargaining power and full information about the firm to maximize the value from philanthropy as reputation insurance for itself, by offering the firm a contract that leaves it just indifferent between buying or not buying insurance (*i.e.*, undertaking or not undertaking philanthropy). The firm, having no real choice on its level of prevention (given society's ability to observe its true accident probability) does the best it can by purchasing the full insurance it is offered, though it makes no additional profit from doing so.

Active and partially informed stakeholders

Next, consider the case where stakeholders are active but unaware of the firm's realized accident probability, *i.e.*, they are aware of p_i^* and can therefore distinguish between clean and dirty firms, but cannot monitor or know the true probability that the firm has an accident, once it has adjusted its prevention expense in the first stage ($\dot{p}_i$). Given that stakeholders are active, the basic solution here is similar to that in the previous case, *i.e.*, society still seeks to maximize its overall welfare, which now means maximizing $V(\alpha_i, m_i^*)$, and thus chooses to offer contract α_i^*. The key difference is that since society cannot measure the firm's true accident probability, the firm can now adjust its accident prevention levels down to $\dot{m}_i$, thus raising its profits. Specifically, since α_i^* is a full-insurance contract, the firm undertakes no prevention, and its probability of having an accident rises to p_i^0. As a result, the firm makes profit of $\pi(\alpha_i^*, 0) = m_i^*$. Conversely, the true welfare gain for society is $V(\alpha_i^*, 0) = U(p_i^* F) - p_i^0 U(F)$. This is less than it would have gained in the fully informed case, since

$p_i^0 > p_i^* \Rightarrow V(\alpha_i^*, 0) < V(\alpha_i^*, m_i^*)$. In fact, even though society is choosing the contract it believes will maximize its welfare, it is possible that $V(\alpha_i^*, 0) < 0$ if $U(p_i^* F) < p_i^0 U(F)$. Intuitively, this would be the case where there is substantial potential for prevention, *i.e.*, p_i^0 is much greater than p_i^*.

Passive and uninformed stakeholders

Finally, consider the case where stakeholders are passive, *i.e.*, the terms of the insurance contract are chosen not by society but by the firm, subject to the constraint that society does not believe itself to be worse off and is therefore willing to accept the contract. In this case, the problem for the firm is:

$$\max_{S_i} \pi(\alpha_i, \dot{m}_i)$$

$$s.t. V(\alpha_i, m_i^*) = 0$$

where $\pi(\alpha_i, \dot{m}_i) = p_i^* F + m_i^* - \dot{p}_i W_i - (1 - \dot{p}_i) S_i - \dot{m}_i$, as before, and

$V(\alpha_i, m_i^*) = (1 - p_i^*) U(S_i) + p_i^* U(W_i) - p_i^* U(F)$

The first order condition is:

$$\frac{\partial \dot{W}_i}{\partial S_i} = -\frac{1}{\dot{p}_i}\left[(1 - \dot{p}_i) + \frac{\partial \dot{m}_i}{\partial S_i} + \frac{\partial \dot{p}_i}{\partial S_i}(\dot{W}_i - \dot{S}_i)\right]$$

where $\dot{\alpha}_i$ is the profit-maximizing insurance contract for the firm. Substituting $\frac{\partial \dot{p}_i}{\partial S_i} = \frac{\partial \dot{p}_i}{\partial m_i}\frac{\partial \dot{m}_i}{\partial S_i}$ and recalling that $\frac{\partial \dot{p}_i}{\partial m_i} = -\frac{1}{W_i - S_i}$, this simplifies to:

$$\frac{\partial \dot{W}_i}{\partial S_i} = -\frac{(1 - \dot{p}_i)}{\dot{p}_i}$$

From the binding condition on social acceptance we have that

$$\frac{\partial \dot{W}_i}{\partial S_i} = -\frac{(1 - p_i^*)}{p_i^*}\frac{U'(\dot{W}_i)}{U'(\dot{S}_i)}$$

Thus the profit-maximizing insurance contract corresponds to the point where

$$\frac{(1 - \dot{p}_i)}{\dot{p}_i} = \frac{(1 - p_i^*)}{p_i^*}\frac{U'(\dot{W}_i)}{U'(\dot{S}_i)}$$

In other words, the equilibrium contract lies at a point where the budget line corresponding to the accident probability $\dot{p}_i$ chosen by the firm becomes a tangent to the indifference curve corresponding to p_i^*. The intuition for this result is that with passive and uninformed stakeholders the firm is free to choose the combination of prevention and insurance that maximizes its gains. Note that at $\dot{\alpha}_i$ the firm does not buy full insurance since $p_i^* \neq \dot{p}_i \Rightarrow \dot{W}_i \neq \dot{S}_i$. The firm realizes a profit from philanthropy as reputation insurance of $\pi(\dot{\alpha}_i, \dot{m}_i) = p_i^* F + m_i^* - \dot{p}_i \dot{W}_i - (1 - \dot{p}_i)\dot{S}_i - \dot{m}_i$, and the effect on social welfare $V(\dot{\alpha}_i, \dot{m}_i) = (1 - \dot{p}_i)U(\dot{S}_i) + \dot{p}_i U(\dot{W}_i) - \dot{p}_i U(F)$, which is unambiguously negative since $p_i^* < \dot{p}_i \Rightarrow V(\dot{\alpha}_i, m_i^*) > V(\dot{\alpha}_i, \dot{m}_i)$ and the binding constraint for social acceptance means that $V(\dot{\alpha}_i, m_i^*) = 0$. Thus, in this case society is unambiguously worse off, even though it believes it has suffered no loss.

Note that if stakeholders were passive but fully informed, *i.e.*, they were aware of the firm's realized probability of an accident but were unable to organize to demand better terms, then the firm would choose a different contract $\hat{\alpha}_i$ based on the binding constraint $V(\hat{\alpha}_i, 0) = 0$. $\hat{\alpha}_i$ would be a full insurance contract, *i.e.*, $\widehat{W}_i = \hat{S}_i$, and would lie at the point where the budget line corresponding to p_i^0 met the corresponding indifference curve through F. At $\hat{\alpha}_i$ the social welfare effect of philanthropy as reputation insurance would be zero, and firm would make additional profit $\pi(\hat{\alpha}_i, 0) = p_i^* F + m_i^* - p_i^0 \widehat{W}_i - (1 - p_i^0)\hat{S}_i$. We do not consider this case in our main manuscript, however, because we think it unlikely that stakeholders who are unable to organize in their own interest would be so well informed or able to monitor firms so closely.

Appendix S2. Discussion of Alternative Equilibria

As discussed briefly in the text, we consider four alternative equilibria in the market for philanthropy as reputation insurance, corresponding to alternate assumptions about firms and society.

Equilibrium I: Active and fully informed stakeholders

We begin with the assumption that society is an active and fully-informed player, *i.e.*, it actively seeks to maximize its own utility from the use of philanthropy as a form of reputation insurance, and can tell clean firms from dirty firms. Further, we assume that society can observe the extent of prevention the firm undertakes, or at least the realized probability of an accident, *i.e.,* it is aware of $\dot{p}_i$. Such full information may be the result of signaling by clean firms, that may be motivated to offer voluntary disclosure of accident probabilities in order to differentiate themselves from dirty firms, either individually (Spence, 1973; Reid and Toffel, 2009; Bagnoli and Watts, 2017), or collectively, to establish a shared reputation (King and Lenox, 2000; King, Lenox, and Barnett, 2002; Barnett and King, 2008).

Since there are many firms of each type in the market seeking to purchase reputation insurance, while society has a monopoly on the supply of such insurance, it is reasonable to expect that the bargaining power in this case lies with society, *i.e.*, competition among firms will bid up the terms of the contract to the point where the gains from reputation insurance go entirely to society. Thus, society will offer the contract that maximizes $V(\alpha_i)$, subject to the firms' budget constraints. Further, since stakeholders can distinguish between clean and dirty firms, they will offer a different contract to each (Rothschild and Stiglitz, 1976).

Figure 1 shows the equilibrium in this case, with contracts α_c^* for clean firms, and α_d^* for dirty firms. These contracts correspond to the points of tangency between the budget lines for the two types of firms and their associated indifference curves, based on probabilities p_c^* and p_d^*. Social utility has increased because society has moved from the indifference curves that passed through the point F to a higher indifference curve with utility level U_i^{max}; and it is impossible for society to get higher utility, since all points above and to the right of the indifference curve U_i^{max} would be unacceptable to the firm. Note that at α_c^*, firms are fully insured, *i.e.*, $S_i^* = W_i^*$. Moreover, the amount of firm philanthropy exactly equals the expected cost of an accident, *i.e.* $S_i^* = p_i^* F$ (see Appendix 1 for a derivation of this result), consistent with the prediction for the 'optimal' amount of philanthropy from prior work that focused only on the demand side (Godfrey, 2005)[1]. Firms make no profit at this point, *i.e.*, $\pi(\alpha_i^*, m_i^*) = 0$, and are therefore indifferent between buying and not buying reputation insurance. Conversely social welfare is maximized with society gaining $V(\alpha_i^*, m_i^*) = U(p_i^* F) - p_i^* U(F)$ from each contract[2], *i.e.*, the utility of the expected accident payment rather than the expected utility of the accident payment, which is positive given risk aversion.

Note further that in this case dirty firms undertake more philanthropy ($S_d^* > S_c^*$), which, given full insurance, means they also pay more in the event of an accident ($W_d^* > W_c^*$); in other words, they pay a higher price for reputation insurance ($R_c^* = R_d^* = F \Rightarrow S_d^*/R_d^* > S_c^*/R_c^*$), which is only reasonable given that they are at higher risk of having an accident. This would be the case, for instance, if a firm trying out risky or experimental new technology were to compensate society for its higher risk of accident by undertaking more philanthropy. Social welfare in this case is maximized, with society using its bargaining power and full information to extract the maximum possible utility.

The case above assumes that society not only knows firm type, it also knows how much prevention the firm undertakes and the resulting probability of an accident—which may be an

[1] Specifically, in the special case where accident probability is entirely exogenous, the equilibrium level of philanthropy is given by $p_i^0 F = S_i^*$,which is identical to the $\alpha L = p$ condition arrived at by Godfrey (2005, p.791) where $\alpha, L,$ and p in his notation correspond to p_i^0, F, and S_i^* in ours, respectively. Once we consider the terms on which insurance is supplied, however, it becomes clear that far from being optimal, this is a level of philanthropy at which the firm makes no additional profit.

[2] Given the mix of firms, the average gain to society is $(1-\theta)V(\alpha_c^*, m_c^*) + \theta V(\alpha_d^*, m_d^*)$.

extreme assumption. An equilibrium similar to that in Figure 1 is obtained if we assume that society cannot observe the actual probability of a firm having an accident given its prevention expenditure, but only the expected probability if the firm were investing in full prevention, *i.e.*, it can observe p_i^* but does not know $\dot{p}_i$. Even in this partial information case, society will still choose the optimal contract α_i^*, with the same philanthropic payments by clean and dirty firms as before.

The key difference between this case and the full information case is in the extent of prevention firms actually undertake. While in the full information case the firm had no choice but to maintain prevention expenditure at m_i^*—given that society would know if it tried to cut back and would adjust the terms of insurance accordingly—in the partial information case the firm may choose to lower its prevention spending. In particular, since α_i^* is a full insurance contract (meaning the firm incurs no additional penalty in case of an accident), the firm may choose to undertake no prevention, with the result that its probability of accident will rise to p_i^0. This increase in accident probability reflects moral hazard on the part of the firm (Pauly, 1974), and will have consequences for both firms and society. On one hand, firms now make at least some profit from philanthropy as insurance on account of the lower prevention they are able to undertake; specifically, they realize profit equal to $\pi_i(\alpha_i^*, 0) = m_i^*$. On the other, the welfare consequences for society are somewhat less clear in this case, since $V(\alpha_i^*, 0) = U(p_i^* F) - p_i^0 U(F) < V(\alpha_i^*, m_i^*)$ and may or may not be greater than zero.

Figure A1 below shows this case for a single firm type. It shows that while the equilibrium contract remains at α_i^*, the true budget line for the firm has shifted out from FG_i^* to FG_i^0, so that the firm makes a profit at α_i^*. Similarly, the true welfare baseline for society (the indifference curve passing through F) is now $\dot{U}_i^0$ rather than U_i^0, and while α_i^* lies unambiguously above the latter, it may or may not lie above the former. For the purposes of the current study, however, we focus on the case where α_i^* lies above $\dot{U}_i^0$, *i.e.*, we assume that the potential for firms to alter their accident probabilities (and therefore the potential for moral hazard) is sufficiently modest to not lead to negative welfare outcomes by itself[3], *i.e.*, $V(\alpha_i^*, 0) > 0$.

Equilibrium II: Active but uninformed stakeholders

Next, consider the case where society is active but uninformed, *i.e.*, it seeks to maximize its utility, but cannot tell clean firms from dirty firms[4]. This lack of information may result from a lack of attention or expertise on the part of stakeholders in evaluating firm disclosures (Milgrom and Roberts, 1986; Barnett, 2014), a lack of adequate sanctioning mechanisms (Delmas and Montes-Sancho, 2010; Short and Toffel, 2010) that may cause firms to make selective disclosures (Lyon and Maxwell, 2011; Doshi, Dowell, and Toffel, 2013; Marquis, Toffel, and Zhou, 2016), or the threat of greater scrutiny that may cause even clean firms to limit disclosure (Lyon and Montgomery, 2013; Kim and Lyon, 2014). For all these reasons, competitive market pressures may prove insufficient to drive full disclosure (Bromiley and Marcus, 1989), leaving stakeholders uninformed. In such a case, the combined probability of an accident perceived by society is $\bar{p} = \theta p_d^* + (1-\theta)p_c^*$, which is the probability that a firm of unknown type will have an accident, and the associated indifference curves have the slope $[(1-\bar{p})U'(S_i)]/[\bar{p}U'(W_i)]$. Because society cannot discriminate between clean and dirty firms, its best option may be a pooling equilibrium where it offers the optimal contract for clean firms (α_c^*) to all firms. Figure 2 shows this case graphically. At α_c^* clean firms engage in no prevention (as discussed above) and therefore realize some additional profits; however, dirty firms potentially do even better, because not only are they able to lower prevention from p_d^* to $\dot{p}_d$, they are also able to realize additional profit of $\dot{p}_d(H_d - W_c^*)$. The intuition behind this result is that dirty

[3] This assumption, which we return to several times below, is made simply to provide a more positive account of philanthropy as reputation insurance; if it did not hold (and there is no logical necessity why it should), then the use of philanthropy as reputation insurance would be welfare destroying with anything short of full transparency around firm prevention efforts.

[4] We assume that society is aware that there are two types of firms and knows the relative proportion of each type, it just cannot tell whether a given firm is a dirty firm or a clean firm. If society were active but unaware of θ it would presumably not choose to offer reputation insurance at all, being unable to assess the likely benefit from doing so.

firms can take advantage of society's inability to distinguish them from clean firms to earn additional profit for themselves. Clearly, at this pooling equilibrium all firms undertake equal amounts of philanthropy and receive equal coverage.

Note, however, that this pooling equilibrium is not always feasible. Specifically, the pooling equilibrium is only achieved where α_c^* lies above $\bar{U}^0$, the combined indifference curve passing through the initial endowment F. Where this is not the case, *i.e.*, where $\bar{p}U(F) > p_c^* U(W_c^*) + (1 - p_c^*)U(S_c^*)$, society recognizes that it is better off not offering reputation insurance at all than offering α_c^*. Logically, this is more likely to be the case the higher the proportion of dirty firms in the population (θ) and the higher the probability of accidents for dirty firms relative to that of clean firms (p_d^*/p_c^*). The intuition being that it only makes sense for society to let dirty firms take advantage of being mixed in with clean firms if there are relatively few dirty firms and they are not too harmful; as the proportion of dirty firms, or the potential damage they may cause, increases, society is better off being skeptical of any attempt to procure reputation insurance through philanthropy, and trusting no firm's philanthropic efforts. Further, while the existence of the equilibrium depends on whether α_c^* lies above $\bar{U}^0$, making it acceptable to society, the welfare consequences depend on whether α_c^* lies above $\ddot{U}^0$, which is the indifference curve corresponding to $\ddot{p} = \theta\dot{p}_d + (1 - \theta)p_c^0$. Where this is not the case (and it may well not be), the pooling equilibrium may exist but be welfare destroying (as shown in Figure 2).

Equilibrium III: Passive stakeholders

Our assumption thus far has been that society plays an active role in determining the terms of the reputation insurance contract, *i.e.*, stakeholders are able to collectively organize in their own interest, and demand the best possible terms from firms in exchange for offering them insurance. The assumption that society can organize and act strategically in its own interest in this way is somewhat questionable, however, given the long-standing literature on the challenges of organizing successful collective action (Olson, 1965; Sen 1967; Ostrom, 1990; Yue, Luo, and Ingram, 2013). For society to demand the most favorable relational contract, individual stakeholders must not only have a shared interest in offering such an insurance contract (as we continue to assume), they must also be capable of organizing as a community to both credibly threaten negative consequences for firms should they have an accident, and credibly commit to offering insurance-like benefits in return for corporate philanthropy. Such collective organization on society's part could take the form of formal institutions (North, 1990), such as local or state governments that exercise appropriate oversight over firms; or it could occur through less formal means, such as the extra-institutional actions of social activists (Baron, 2001; King and Soule, 2007; Sine and Lee, 2009; McDonnell and King, 2013), who may act as society's representatives, entering into relational contracts with firms; or even through more generalized norms of civic engagement (Putnam, 2000), which create the potential for spontaneous collective action by concerned stakeholders. While a detailed exploration of the ways in which society can collectively organize—and the costs of such organization (Hansmann, 1996)—is beyond the scope of the current study, the key point is that society's ability to collectively organize and take civic action (Longhofer, Negro, and Roberts, 2018) is a necessary condition for the equilibria we have discussed thus far.

Where society is unable to organize and act in its collective interest, we can think of it as being essentially passive, allowing firms to dictate the terms of their insurance contracts, as long as these contracts do not leave it worse off[5]. Specifically, we assume that society's behavior is not maximizing but satisficing (Simon, 1947; Winter, 1971; 2000), *i.e.*, it does not seek to maximize the utility from the provision of reputation insurance, but simply seeks to maintain some minimum level of utility. In other words, society accepts any bid for a contract that satisfies $V(\alpha_i, m_i^*) \geq 0$; so long

[5] We assume that society is at least able to safeguard its interests to the point where insurance contracts do not obviously reduce its utility. If this were not the case, reputation insurance would only be more harmful to society.

as the contract does not leave society worse off, it is willing to supply reputation insurance at those terms[6]. We can think of this scenario as one where firms anticipate full reciprocity on society's part—*i.e.*, they expect society to forgive just as much loss of utility in the aftermath of an accident as the amount of utility it received from their philanthropic efforts—and make choices about how much philanthropy to undertake accordingly.

The equilibrium associated with this case is shown in Figure 3. In this case both types of firms choose the contract that maximizes their profits, *i.e.*, they choose contracts $\dot{\alpha}_c$ and $\dot{\alpha}_d$, which lie at the points where the slope of the combined indifference curve passing through the initial endowment $\bar{U}^0$ is equal to the odds ratio for clean and dirty firms respectively. These points represent a profit-maximizing equilibrium for the firms because in order to make any additional profit, firms would have to choose a contract that lay below and to the left of the odds ratio line passing through $\dot{\alpha}_i$, but any such point would be unacceptable to society since it would leave society worse off than it was before the contract (see Appendix 1 for a mathematical derivation of this result). Thus $\pi(\alpha_i)$ is maximized at $\dot{\alpha}_i$. The intuition being that with passive stakeholders, the firm is free to make choices on both prevention and insurance contracts that are in its own best interest. Note also that $\dot{\alpha}_i$ is not a full insurance contract—given full flexibility on the choice of prevention and insurance, the firm can more precisely balance the two, and will therefore choose to retain some level of accident penalty in exchange for less philanthropy up-front.

Conversely, the welfare consequences for society from this equilibrium are unambiguously negative. Since contracts $\dot{\alpha}_c$ and $\dot{\alpha}_d$ lie on the combined indifference curve through the origin $\bar{U}^0$ they necessarily lie below the indifference curve $\ddot{U}^0$, which is the true counterfactual taking into account moral hazard on the part of the firms. Thus, even though passive stakeholders may believe

[6] We assume that if society is passive it also cannot observe realized probabilities, and even if it had information about firm type it would not use it. The case where society is aware of realized probabilities but passive is discussed briefly in Appendix 1.

they are no worse off, this is only because they are unable to observe the reduction of prevention by firms as a result of the availability of insurance. In fact, stakeholders would be better off not offering insurance to firms at all[7]. Thus, the equilibrium with passive stakeholders represents a case where firms are doing well by doing harm rather than good[8].

Moreover, since $\dot{p}_d > \dot{p}_c$, point $\dot{\alpha}_c$ lies further along the combined indifference curve than point $\dot{\alpha}_d$, meaning that dirty firms undertake more philanthropy ($\dot{S}_d > \dot{S}_c$), but also pay lower penalties in case of an accident ($\dot{W}_d < \dot{W}_c \Rightarrow \dot{R}_d > \dot{R}_c$), implying that this is an equilibrium with adverse selection—firms that expect to do more damage buy more insurance and get more coverage. It is also worth noting that while clean firms do benefit from reputation insurance on account of society's inability to act collectively in its own interest, this benefit is lower than it would have been in the absence of dirty firms, and, conversely, dirty firms benefit more than they would have in the absence of clean firms. This follows from the fact that the indifference curve U_c^0, which is the social indifference curve passing through the initial endowment for clean firms alone, lies below $\bar{U}^0$ while U_d^0 lies above it.

Equilibrium IV: Exogenous philanthropy

Returning to the case where stakeholders are active but uninformed, it is also important to consider the firm's motives in undertaking philanthropy. Our discussion thus far assumes that the only reason firms undertake philanthropy is for its insurance benefits. But firms may undertake philanthropy for a variety of reasons. First, they may be genuinely altruistic, moved to donate to social causes out of an ethical or moral concern for their stakeholders (Freeman, 1984; Maitland, 1994; Windsor, 2001; Freeman, Wicks, and Parmar, 2004). Second, they may be responding to

---

[7] In the special case where accident probabilities were exogenous and there was no moral hazard problem, society would be no worse off at $\dot{\alpha}_i$, but it would also be no better off.

[8] We assume throughout that firm's philanthropic efforts are substantive and not merely symbolic (Kim and Lyon, 2011; Marquis *et al.*, 2016; Kaul and Luo, 2018). If society were forgiving firms for doing substantive damage in exchange for symbolic efforts, this would clearly result in a further decrease in social welfare.

institutional pressures (*e.g.,* Marquis *et al*., 2007; Marquis and Tilcsik, 2016) or the ideological preferences and structures of top management teams (*e.g.*, Chin, Hambrick, and Trevino, 2013; Marquis and Lee, 2013; Gupta, Briscoe, and Hambrick, 2017). Third, they may benefit from philanthropy in ways other than reputation insurance (Bénabou and Tirole, 2010), *e.g.*, through stronger demand from customers (*e.g.,* Lev, Petrovits, and Radhakrishnan, 2010; Elfenbein, Fisman, and McManus, 2012; Kaul and Luo, 2018), greater employee support and effort (*e.g.,* Bode, Singh, and Rogan, 2015; Burbano, 2016; Flammer and Luo, 2017), or improved relations with key stakeholders (*e.g.,* Wang and Qian, 2011; Henisz, Dorobantu, and Nartey, 2014; Dorobantu, Henisz, and Nartey, 2017). Whatever the reason, in all these cases, the firm undertakes some base level of philanthropy that is exogenous to our model, in the sense that it is not determined by the potential insurance benefits, but would be undertaken even if there were no insurance contract involved.

Specifically, we assume that clean firms undertake an exogenous level of philanthropy S^A, while dirty firms undertake no exogenous philanthropy[9]. The assumption that clean firms donate more for other reasons than dirty firms seems logical, both because the fact of being a clean firm may be associated with a stronger social conscience, and because it may be safer for clean firms to rely on philanthropy as a way to appeal to key stakeholders (customers, employees, etc.) given that they are less likely to have an accident that will undermine their credibility. Figure 4 shows the case with exogenous philanthropy by clean firms, with the budget line for the clean firms being accordingly shifted out to the right by S^A.

This shift in the budget line of the clean firms is important because it enables society to arrive at a separating equilibrium, since there are now a set of contracts that are acceptable to clean

[9] While we only model exogenous philanthropy by clean firms, our results would be identical if both type of firms gave exogenously, so long as clean firms gave more than dirty firms. We can thus think of S^A as the difference between the exogenous philanthropy of clean and dirty firms, *i.e.*, $S^A = S_c^A - S_d^A$.

firms but not to dirty firms[10]. In particular, as Figure 4 shows, society can now offer two contracts α_c^{**} and α_d^*, at the points where the budget lines for clean and dirty firms become tangent to their relevant indifference curves respectively, such that α_c^{**} is only acceptable to clean firms but not to dirty firms. In this equilibrium, clean firms undertake more philanthropy than dirty firms ($S_c^{**} > S_d^*$), and pay lower penalties once the base rate of philanthropy is accounted for ($W_d^* > W_c^{**} - S^A$). Both firm types make some profit from reputation insurance[11] as a result of the reduction in prevention expenses. Social welfare is likely enhanced however, at least so long as we assume, as we did in the partially informed equilibrium above, that the moral hazard problem is not severe enough to lower social welfare by itself. Note that by creating the conditions for the separating equilibrium, exogenous philanthropy by clean firms not only benefits society directly, it also contributes to social utility indirectly by forcing dirty firms to raise their game. Thus, the market for reputation insurance amplifies the benefits from philanthropic actions that clean firm would have undertaken in any case.

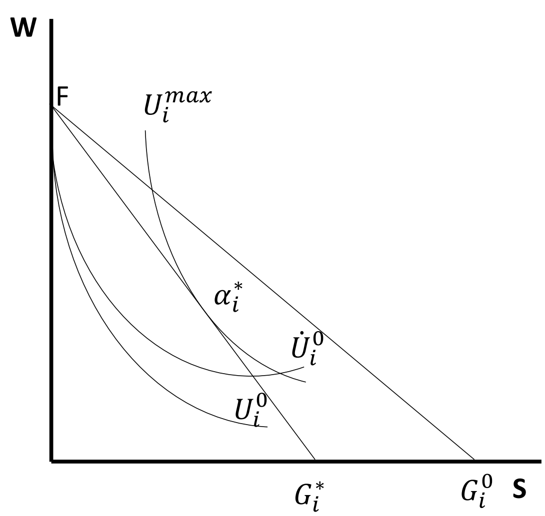

Figure A1: Moral hazard problem

Firm i reduces its prevention expense, increasing accident probability from p_i^* to p_i^0, with corresponding shifts from FG_i^* to FG_i^0 and U_i^0 to $\dot{U}_i^0$. Equilibrium contract is at α_i^* which is unambiguously above U_i^0 but less of an improvement relative to $\dot{U}_i^0$(and could potentially lie below it).

[10] Strictly speaking, this is only true if $S^A > (p_d - p_c)F$.

[11] Clean firms may realize additional profit from their exogenous philanthropy through mechanisms other than insurance, as described above, or incur a loss if the exogenous philanthropy is truly altruistic.

Additional References (not included in main text)

Bagnoli M, Watts SG. 2017. Voluntary assurance of voluntary CSR disclosure. *Journal of Economics and Management Strategy* **26**(1): 205-230

Barnett, ML, King AA. 2008. Good fences make good neighbors: A longitudinal analysis of an industry self-regulatory institution. *Academy of Management Journal* **51**(6): 1150-1170.

Dorobantu S, Henisz WJ, Nartey LJ. 2017. Not all sparks light a fire: Stakeholder and shareholder reactions to critical events in contested markets. *Administrative Science Quarterly* **62**(3): 561–597.

Doshi AR, Dowell GWS, Toffel MW. 2013. How firms respond to mandatory information disclosure. *Strategic Management Journal* **34**(10): 1209–1231.

Elfenbein DW, Fisman R, McManus B. 2012. Charity as a substitute for reputation: Evidence from an online marketplace. *Review of Economic Studies* **79**(4): 1441–1468.

Freeman RE, Wicks AC, Parmar B. 2004. Stakeholder theory and "the corporate objective revisited". *Organization Science* 15(3): 364-369.

Gupta A, Briscoe F, Hambrick DC. 2017. Red, blue, and purple firms: Organizational political ideology and corporate social responsibility. *Strategic Management Journal* **38**(5): 1018-1040.

Hansmann H. 1996. *The Ownership of Enterprise.* Belknap Press: Cambridge, MA.

Kim EH, Lyon TP. 2011. Strategic environmental disclosure: Evidence from the DOEompvoluntary greenhouse gas registry. *Journal of Environmental Economics and Management* **61**(3): 311-326

Kim EH, Lyon TP. 2014. Greenwash vs. brownwash: Exaggeration and undue modesty in corporate sustainability disclosure. *Organization Science* **26**(3): 705-723.

King AA, Lenox MJ, Barnett ML. 2002. Strategic responses to the reputation commons problem. In *Organizations, Policy and the Natural Environment: Institutional and Strategic Perspectives*, Hoffman A, Ventresca M (eds.). Stanford, CA: Stanford University Press: 393-406.

Longhofer W, Negro G, Roberts PW 2018. The changing effectiveness of local civic action: The critical nexus of community and organization. *Administrative Science Quarterly* forthcoming.

Lyon TP, Maxwell JW. 2011. Greenwash: Corporate environmental disclosure under threat of audit. *Journal of Economics & Management Strategy* **20**(1): 3-41.

Lyon TP, Montgomery AW. 2013. Tweetjacked: The impact of social media on corporate greenwash. *Journal of Business Ethics* **118**(4): 747–757.

North, DC. 1990. A transaction cost theory of politics. *Journal of theoretical politics* **2**(4), 355-367.

Olson M. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Boston, MA: Harvard University Press.

Ostrom E. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action.* Cambridge University Press.

Sen AK. 1967. Isolation, assurance and the social rate of discount. *Quarterly Journal of Economics* **81**(1): 112–124.

Simon H. 1947. *Administrative behavior*. New York: Free Press.

Spence M. 1973. Job market signaling. *Quarterly Journal of Economics* **87**(3): 355-374.

Windsor D. 2001. 'Corporate social responsibility: A theory of the firm perspective': Some comments. *Academy of Management Review* **26**(4): 502–504.

Winter SG. 1971. Satisficing, selection, and the innovating remnant. *Quarterly Journal of Economics* **85**(2): 237.

Winter SG. 2000. The satisficing principle in capability learning. *Strategic Management Journal* **21**(10/11): 981–996.

Appendix S3: Main Variables

Variables	Measurement	Source
Amount of spill	Logged barrel amount of spills at firm-(state) level	National Response Center (NRC)
Corporate donation	Logged amount of donation (in $Mil) at firm-(state) level	Foundation Directory Online (FDO)
Firm-Otherstate Corporate donation	Logged amount of donation (in $Mil) at firm-state level in the states where spill did not occur	FDO, NRC
Production facilities	Number of production facilities at firm-state level	National Establishment Time Series (NETS)
Non-production facilities	Number of non-production facilities at firm-state level	NETS
Sales	Logged Sales at firm-state level	NETS
Total assets	Logged book value of total asset at firm level	Compustat
ROA*	Operating income before depreciation/ the book value of total assets at firm level	Compustat
Cash holding*	Cash and short-term investments/ the book value of the total assets at firm level	Compustat
Capex	Capex investments/ the value of property, plant and equipment	Compustat
CAR	Cumulative Abnormal Return (CAR) over a two-day (0, +1) or three-day (-1, +1) window around the date of spill, estimated by Carhart four-factor model	Center for the Research on Stock Prices (CRSP)

* Both ratio variables are winsorized at the 1st and 99th percentiles of their empirical distribution.

Appendix S4. Summary Statistics and Correlation

Panel A – CAR Analysis

	Mean	S.D.	Min	Max
(1) CAR (-1, +1)	0.024	2.230	-14.580	18.087
(2) Amount of Oil Spill (Logged)	1.241	2.154	0.000	15.405
(3) Corporate Donation_l1 (Logged)	10.557	6.882	0.000	19.417
(4) Firm-state Corporate Donation (Logged)	1.453	4.071	0.000	15.191
(5) Production Facilities	1.435	4.871	0.000	63.000
(6) Non-Production Facilities	0.028	0.165	0.000	1.000
(7) Sales (Logged)	4.524	7.309	0.000	21.542
(8) Total Assets_l1 (Logged)	11.048	1.248	7.078	12.684
(9) ROA_l1	0.082	0.054	-0.119	0.198
(10) Cash Holding_l1	0.050	0.042	0.001	0.160
(11) CAPEX_l1	0.188	0.073	0.000	0.702

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) CAR (-1, +1)	1.000										
(2) Amount of Oil Spill (Logged)	0.002	1.000									
(3) Corporate Donation_l1 (Logged)	0.008	-0.077	1.000								
(4) Firm-state Corporate Donation (Logged)	0.015	0.029	0.312	1.000							
(5) Production Facilities	0.020	0.032	-0.085	0.119	1.000						
(6) Non-Production Facilities	0.008	0.003	0.151	0.096	-0.002	1.000					
(7) Sales (Logged)	0.027	0.065	-0.026	0.207	0.537	0.276	1.000				
(8) Total Assets_l1 (Logged)	0.026	-0.016	0.270	0.246	0.071	0.156	0.256	1.000			
(9) ROA_l1	-0.010	-0.019	0.246	0.124	0.044	0.113	0.049	0.301	1.000		
(10) Cash Holding_l1	0.022	0.024	0.188	-0.115	-0.065	0.070	-0.033	0.221	0.286	1.000	
(11) CAPEX_l1	0.010	-0.121	-0.044	-0.098	-0.081	-0.112	-0.150	-0.206	-0.067	-0.151	1.000

Notes: Variables followed by _l1 are lagged by one year.

Panel B – Main Models

		Mean	S.D.	Min	Max
(1)	Firm-state Amount of Oil Spill (Logged)	0.206	1.099	0.000	15.405
(2)	Firm-state Corporate Donation_l1 (Logged)	1.038	3.182	0.000	18.133
(3)	Firm-Otherstate Corporate Donation_l1 (Logged)	6.036	6.978	0.000	19.417
(4)	Production Facilities	1.248	5.891	0.000	125.000
(5)	Non-Production Facilities	0.018	0.157	0.000	3.000
(6)	Firm-state Sales (Logged)	2.590	5.813	0.000	21.775
(7)	Total Assets_l1 (Logged)	10.159	1.151	7.078	12.684
(8)	ROA_l1	0.077	0.056	-0.141	0.196
(9)	Cash Holding_l1	0.053	0.050	0.000	0.211
(10)	CAPEX_l1	0.221	0.097	0.045	0.702

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1)	Firm-state Amount of Oil Spill (Logged)	1.000									
(2)	Firm-state Corporate Donation_l1 (Logged)	0.216	1.000								
(3)	Firm-Otherstate Corporate Donation_l1 (Logged)	0.098	0.407	1.000							
(4)	Production Facilities	0.270	0.224	0.058	1.000						
(5)	Non-Production Facilities	0.243	0.168	0.05	0.343	1.000					
(6)	Firm-state Sales (Logged)	0.338	0.207	0.094	0.374	0.307	1.000				
(7)	Total Assets_l1 (Logged)	0.148	0.114	0.095	0.051	0.075	0.203	1.000			
(8)	ROA_l1	0.014	-0.010	-0.032	0.044	0.033	0.037	0.103	1.000		
(9)	Cash Holding_l1	-0.002	0.055	0.147	0.095	0.008	-0.040	0.118	0.150	1.000	
(10)	CAPEX_l1	-0.077	-0.040	-0.060	0.044	-0.020	-0.076	-0.272	0.080	0.007	1.000

Notes: Variables followed by _l1 are lagged by one year.

Appendix S5. Alternate Measures of Accident Penalty

	AM1	AM2	AM3
	DV: Dummy of EPA lawsuit filing against the firm		DV: Amount of settlement provisions
Corporate Donation_l1	-0.1816	-0.2695	1.3018
	(0.0238)	(0.0006)	(0.2572)
Amount of Oil Spill	0.0475		-0.2324
	(0.6914)		(0.8093)
Amount of Oil Spill_l1		0.3493	
		(0.1034)	
Total Assets_l1	-0.8153	0.2670	-7.9509
	(0.6695)	(0.8629)	(0.6497)
ROA_l1	13.0359	19.4659	-4.4057
	(0.1881)	(0.0225)	(0.9514)
Cash Holding_l1	50.8368	44.6243	-15.1423
	(0.0238)	(0.0196)	(0.7139)
CAPEX_l1	22.3814	33.3930	-25.5512
	(0.1809)	(0.1105)	(0.4540)
Constant			74.1660
			(0.6531)
R^2			0.0489
BIC	51.7	48.8	
Observations	56	56	278

Notes: AM1 and 2 use the dummy variable of EPA lawsuit as a dependent variable, which is coded as 1 if cleanup lawsuits were filed by EPA against the firm in a given year and 0 otherwise, and run conditional FE logit regressions. AM3 uses the logged amount of settlement provisions as the dependent variable, and runs OLS FE regressions. All models employ year and firm fixed effects. Robust standard errors are used. P-values are reported in parentheses. We manually matched listings on cleanup enforcement lawsuits from the EPA website (www.epa.gov/enforcement/data-and-results) against our sample companies by firm names. Cleanup lawsuits are rare, filed only for the most significant cases of spills. During our study period of 2004 to 2011, 13 civil lawsuits and 5 criminal prosecutions were filed against 7 companies in our sample, which include Anadarko, BP, Chevron, Conocophilips, Exxon, Shell, and Sunoco. To invoke sufficient variation, AM1 and 2 are thus based on the sub-sample of firms that experienced lawsuits at least once during the study period. Note that all the civil suits were settled, with the median settlement amount of $11 Mil. The median size of fines companies were ordered to pay in criminal prosecutions was $13 Mil.

Appendix S6. Firm Fixed Effects (Omitted firm: Baker Hughes)

	Firm FE	Std.Err.
ConocoPhilips	0.8453	(0.0000)
Chevron	0.7662	(0.0000)
Exxon Mobil	0.6801	(0.0001)
Royal Dutch Shell	0.6592	(0.0001)
BP	0.6273	(0.0001)
Marathon Oil	0.5483	(0.0000)
Sunoco	0.5097	(0.0000)
Anadarko Petroleum	0.3728	(0.0002)
Valero Energy	0.3726	(0.0001)
Diamond Offshore Drilling	0.3105	(0.0042)
Statoil ASA	0.3092	(0.0083)
Repsol SA	0.2926	(0.0100)
Lukoil	0.2834	(0.0138)
Apache	0.2741	(0.0007)
Nexen	0.2547	(0.0023)
Suncor Energy	0.2451	(0.0090)
BG Group	0.2419	(0.0101)
Encana	0.2357	(0.0225)
Noble Energy	0.2236	(0.0058)
Talisman Energy	0.2031	(0.0177)
Murphy Oil	0.1728	(0.0306)
Schlumberger	0.1675	(0.0108)
Santos	0.1626	(0.0365)
Tesoro	0.1585	(0.0641)
XTO Energy	0.1524	(0.0732)
Hess	0.1452	(0.1052)
Newfield Exploration	0.1362	(0.0554)
Devon Energy	0.1343	(0.1637)
Chesapeake Energy	0.1253	(0.1590)
Occidental Petroleum	0.1125	(0.2242)
EOG Resources	0.0982	(0.1899)
Freeport-McMoRan Oil & Gas	0.0769	(0.2772)
Halliburton	0.0673	(0.0632)
Pioneer Natural Resources	0.0498	(0.5105)

Notes: This table reports the coefficients and the standard errors of firm fixed effects from Table 2, M10.

Appendix S7. Robustness checks

As mentioned in the text, we run a number of robustness tests to confirm our main findings, as shown in the table below. First, we use a number of alternative dependent variables, confirming that our results are robust to using the count of spills in a firm-state (AM4) or just a dummy indicator for the presence of a spill in a firm-state as outcome variables (AM5). We also confirm in AM6 that our results continue to hold if we only consider large spills (those involving spills of 10,000 barrels or more), setting all other spill amounts to 0, as well as (in unreported analysis) to limiting our sample to spills that result in negative and significant CARs. Second, we use several alternative specifications. Specifically, to alleviate the concern that time varying state characteristics or state-specific shocks (such as policy incentives for drilling) may be driving both donation activities and spills, we re-run our main analysis with state-year fixed effects (AM7) as well as additional controls for state specific time trends (AM8) and find consistent results. Third, we confirm in AM9 that our results are robust to dropping BP out of the sample as a potential outlier with high influence (given its Gulf of Mexico accident in 2010), as well as to dropping firm-states with no corporate donations in AM11. The findings are also robust in firm-states with and without production operations (AM10a/b), and hold for both firms listed on US stock exchanges whose primary business is petroleum, as well as firms listed on foreign exchanges or in other primary industries who have substantial US oil operations (AM12a/b). Fourth, to account for the possibility that a positive reputation may be acquired slowly over time (Barnett and Salomon, 2012; Barnett, 2016), we include in AM13 longer lags of donation amount, and continue to see a positive association between donations two years ago and spills in the current year, though the size of the effect attenuates somewhat over time. Lastly, our result is robust to using Tobin's Q as a control measure for firm profitability instead of ROA (AM14), as well as to including patents applied for by the firm as a measure of firm investment in new technologies (AM15).

	Panel A: Alternative DVs			Panel B: Alterative Specifications		Panel C: Alternative Samples						Panel D: Other Controls		
	AM4	AM5	AM6	AM7	AM8	AM9	AM10a	AM10b	AM11	AM12a	AM12b	AM13	AM14	AM15
Firm-state Corporate	0.0801	0.0756	0.0329	0.0372	0.0373	0.0305	0.0511	0.0133	0.0250	0.0448	0.0057	0.0221	0.0266	0.0385
Donation_l1	(0.0000)	(0.0000)	(0.0055)	(0.0035)	(0.0032)	(0.0165)	(0.0000)	(0.0000)	(0.0105)	(0.0000)	(0.0198)	(0.0458)	(0.0429)	(0.0077)
Tobin's Q_l1													-0.0000	
													(0.6648)	
Number of Patents														-0.0001
														(0.7214)
Firm-state Corporate												0.0176		
Donation_l2												(0.0237)		
Firm-state Corporate												0.0156		
Donation_l3												(0.1194)		
Constant	-6.8775		0.5428	0.5364	0.5020	0.9301	0.5970	0.1276	-0.8839	0.8405	-0.2591	0.2929	0.9541	0.8807
	(0.0015)		(0.1929)	(0.2664)	(0.2368)	(0.0667)	(0.6972)	(0.5326)	(0.4666)	(0.0653)	(0.4377)	(0.5431)	(0.0322)	(0.1705)
R^2			0.1864	0.2548	0.2408	0.2359	0.2906	0.0427	0.3416	0.2595	0.1699	0.2427	0.2351	0.2285
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
YearFE			Yes		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FirmFE		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
StateFE			Yes		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
StateYearFE				Yes										
StateTR					Yes									
Year-, State- Dummies		Yes												
Firm-, State-, Year- Dummies	Yes													
Observations	13900	13900	13900	13900	13900	13500	4000	9900	3464	10700	3200	10400	12700	10100

Notes: Coefficients of unconditional negative binomial regression with firm-, year-, and state- dummies (AM4), conditional logistic regression with firm fixed effects (AM5), and OLS fixed effects regressions (AM6-AM15). Dependent variables are measured at firm-state level in a given year and are the same as in the main analyses unless specified otherwise. P-values are reported in parentheses. **Panel A:** AM4 uses the count of spill occurrence as the DV; AM5 uses the dummy variable of spill occurrence as the DV; AM6 uses the total amount of spills bigger than 10,000 barrels as the DV. **Panel B:** AM7 includes firm-fixed effects and state-year fixed effects. AM8 includes firm-fixed effects and state-specific time trend. **Panel C:** AM9 drops BP from the sample. AM10a uses the subsample of firm-state with production facilities; while AM10b uses the subsample of firm-states without production facilities. AM11 drops firm-states that receive no donations. AM12a uses the subsample of firms that are mainly in the petroleum business and listed on a US stock exchange, while AM12b uses other firms. **Panel D:** AM13 controls for lags of donations in the last two periods. AM14 controls for Tobin's Q instead of ROA. AM15 controls for the count of patents applied for by the firm.